

INSTITUTO DE ARTE E COMUNICAÇÃO SOCIAL
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
DOUTORADO EM CIÊNCIA DA INFORMAÇÃO

RAQUEL LUISE PRET COELHO

DO QUÊ SE TRATA? DOCUMENTO, INDEXAÇÃO E PRODUÇÃO DE SENTIDO
NOS ARQUIVOS DE UNIVERSIDADE



Niterói
2019

RAQUEL LUISE PRET COELHO

DO QUÊ SE TRATA? DOCUMENTO, INDEXAÇÃO E PRODUÇÃO DE SENTIDO NOS
ARQUIVOS DE UNIVERSIDADE

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação do Instituto de Arte e Comunicação Social da Universidade Federal Fluminense, como requisito para obtenção do título de Doutor em Ciência da Informação.

Área de concentração: Fluxos e Mediações Sócio-Técnicas

Orientadora: Profa. Dra. Rosa Inês de Novais Cordeiro

Niterói

2019

P922q Pret, Raquel Luise
Do quê se trata? : Documento, indexação e produção de
sentido nos arquivos de universidade / Raquel Luise Pret ;
Rosa Inês de Novais Cordeiro, orientador. Niterói, 2019.
277 f. : il.

Tese (doutorado)-Universidade Federal Fluminense, Niterói,
2019.

DOI: <http://dx.doi.org/10.22409/PPGCI.2019.d.09219272741>

1. Indexação. 2. Arquivo (Universidade). 3. Protocolo. 4.
Produção de sentido. 5. Produção intelectual. I. Cordeiro,
Rosa Inês de Novais, orientador. II. Universidade Federal
Fluminense. Instituto de Arte e Comunicação Social. III.
Título.

CDD -

RAQUEL LUISE PRET COELHO

DO QUÊ SE TRATA? DOCUMENTO, INDEXAÇÃO E PRODUÇÃO DE SENTIDO NOS
ARQUIVOS DE UNIVERSIDADE

Tese apresentada ao Programa de Pós-Graduação
em Ciência da Informação do Instituto de Arte e
Comunicação Social da Universidade Federal
Fluminense, como requisito para obtenção do título
de Doutor em Ciência da Informação.

Niterói, 28 de março de 2019.

Prof. Dra. Rosa Inês de Novais Cordeiro (orientadora)
Universidade Federal Fluminense - UFF

Profa. Dra. Mariângela Spotti Fujita
Universidade Estadual de São Paulo - UNESP

Profa. Dra. Vânia Lisbôa da Silveira Guedes
Universidade Federal do Rio de Janeiro - UFRJ

Profa. Dr. Renato Tarciso de Sousa
Universidade de Brasília - UnB

Profa. Dra. Joice Cleide Cardoso Ennes de Souza
Universidade Federal Fluminense - UFF

Profa. Dra. Anna Carla Mariz
Universidade Federal do Estado do Rio de Janeiro – UNIRIO (suplente)

Prof. Dra. Natália Bolfarini Tognoli
Universidade Federal Fluminense – UFF (suplente)

Para Maria Noemia, Waninha e Raphael

AGRADECIMENTOS

Agradeço imensamente à Rosa Inês, minha orientadora e amiga, que acreditou nesse projeto e tem me acompanhado ao longo desse caminho, corrigindo rotas, aconselhando e oferecendo todo tipo de apoio necessário nesse difícil exercício de produção de uma tese. Agradeço a todos os meus amigos e colegas do campo que estiveram dispostos em estabelecer um diálogo contínuo sobre a pesquisa, apresentando muitas vezes suas concordâncias e questionamentos que instigaram essa investigação. Especialmente a Clarissa Schmidt, primeira a acreditar nessa realização, antes mesmo de mim.

Meu agradecimento também à Universidade Federal Fluminense, instituição que além de contribuir para a minha formação com a graduação em Arquivologia e agora o meu doutoramento, é o meu lugar de trabalho onde tanto as reflexões em sala de aula, quanto os projetos de pesquisa e de extensão, juntos aos seus arquivos levaram ao desenvolvimento desse trabalho.

Amigos que também me ajudaram a entender melhor o processo de doutoramento, compartilhando experiências e mostrando atalhos neste caminho fazem parte desse agradecimento como Natália Tognoli e Renato Mattos..

Agradeço a minha mãe e amiga, Maria Noemia, por todo apoio, dedicação, paciência e grande investimento ao longo de todos esses anos.

Agradeço a Zenir (*in memoriam*), minha avó e madrinha, segunda mãe que participou ativamente de minha formação e educação.

Ao meu pai Jair pelo entusiasmo, e ao meu irmão Raphael que sempre foi um camarada para todas as horas.

Minha gratidão à Waninha, amor de uma vida inteira, companheira e paciente que sempre me encorajou aos novos desafios e perdoou minhas ausências ao longo dessa trajetória.

Não sou nada.

Nunca serei nada.

Não posso querer ser nada.

À parte isso, tenho em mim todos os sonhos do mundo.

(Álvaro de Campos, Tabacaria)

RESUMO

Pesquisa sobre o processo de indexação no âmbito dos arquivos das universidades federais brasileiras na fase de produção documental. O foco da análise encontra-se nas relações da indexação com os usos dos documentos em unidades de trabalho das universidades, considerando a importância dos protocolos como sistemas de registro dos documentos. Na fase de produção documental, os usos informativos estão relacionados aos próprios processos de trabalho da entidade produtora. Eles poderão ajudar na tomada de decisão, no controle fiscal, no aprimoramento de funções, no histórico sobre ações realizadas, na identificação do trâmite dos processos de trabalho, entre outros usos. As universidades como entidades produtoras possuem especificidades em relação às suas dinâmicas organizacionais, estruturas hierárquicas, práticas decisórias que diferem de outros tipos de organização. Os processos de trabalho são mais autônomos, respondendo muitas vezes apenas a seu departamento, instituto ou faculdade. Estudo da natureza e dimensão histórica do conceito de indexação nas correntes da Ciência da Informação. Este trabalho utiliza como metodologia além de uma ampla revisão bibliográfica acerca da indexação, dos arquivos, dos documentos de arquivo; a análise da legislação sobre sistemas de arquivos, acesso à informação e implementação do sistema protocolo integrado; investigação de como as universidades brasileiras estão padronizando o registro de seus documentos nos seus manuais de protocolo; e a observação dos sistemas de recuperação da informação da Universidade Federal Fluminense como campo empírico da tese. Os resultados preliminares apontam que tanto os documentos indexados, quanto a operação da indexação são práticas sociais realizadas no uso da língua. Ademais, de uma forma *stricto sensu* a indexação parte do reconhecimento ou identificação de assuntos dos documentos e nas demandas por informação, ou de uma maneira mais *lato sensu*, a indexação poderá apontar um “prognóstico” de identificação de domínios de conhecimento/análise de domínio instrumental que se estabelecem em uma determinada área do conhecimento, de uma comunidade discursiva. Os protocolos funcionam para registrar, controlar o fluxo documental, informar sobre o trâmite dos documentos nos arquivos, mas também devem ser explorados como uma instância de poder da universidade que subsidia a tomada de decisão da sua gestão visando o planejamento de ações na instituição.

Palavras-chave: Indexação; Arquivo (Universidades); Protocolo; Produção documental (Universidades); Produção de Sentido; Documentos de arquivo (Universidades).

ABSTRACT

This research is about the indexing process in university archive systems when the records are in production phase. The intend is analyzing the indexing relations with the record uses by work office units. It is also considerate the indexing implications in information retrieval. In record production phase, the information uses are associated with the producer activities. This uses may help the producers to make strategic decisions, to control finances, to improve functions, to inform about historical actions and to identify business transactions. The universities as producers have specific organizational dynamics with particular hierarchical structures and decision-making practices. The business trasactions are more independent than other institutions. The managers most part of times respond only one department, institute or school. This reality is also presented by this study. The approach is about the nature and the history of indexing concept. It also presents different tradictions of indexing theory in Science Information field. A literature review of indexing process, archive systems, records, production of meaning are made, besides that a large research about registry systems, federal universities, archives and information acess is produced. It is also investigated the new electronic registry system, called Sistema Protocolo Integrado, a government iniciative, and the accession by Universidade Federal Fluminense, the empirical field of this research. The preliminary results provide that the indexing records are social practices characterized by language use. Moreover, in *strictu sensu*, the indexing is a technique able to recognize and identify concepts inside the records. Although, *in lato sensu*, the indexing may dignose a knowledge domain and offer a instrumental domain analysis whose knowledge domain of discursive community can be recognized. The registry systems make the control of the workflow and inform about business transactions in university archive systems, but they also became a potential power instance for the managers. They can provide all kind of crucial information to the administrators make strategic decisions and plan important action for the university development.

Keywords: Indexing, Archive (Universities), Registry; Record Production (Universities); Production of Meaning; Record (Universities)

LISTA DE SIGLAS E ABREVIATURAS

AACR - Anglo American Cataloguing Rules
ABNT - Associação Brasileira de Normas Técnicas
BNB - British National Bibliography
CA - Chemical Abstracts
CDD - Classificação Decimal de Dewey
CDU - Classificação Decimal Universal
CEART - Centro de Artes
CI - Ciência da Informação
CONARQ - Conselho Nacional de Arquivos
CRG - Classification Research Group
DIBRATE - Dicionário Brasileiro de Terminologia Arquivística
GED – Gestão Eletrônica de Documentos (software)
GPCA - Gerência Plena de Comunicação Administrativa
HUAP - Hospital Antônio Pedro
KO - Knowledge Organization
KWIC - Keyword in Context
KWOC - Keyword out Context
ICA - International Council of Archives
ISAD(G) - General International Standard Archival Description
ISO - International Organization of Standardization
JD - Journal of Documentation
LISA - Library and Information Science Abstracts
LTM - Large-term memory
MAD - Manual of Archival Description
MJ - Ministério da Justiça
MP - Ministério do Planejamento, Desenvolvimento e Gestão
NDC - Núcleo de Documentação
NOBRADE - Norma Brasileira de Descrição Arquivística
PEN - Processo Eletrônico Nacional
PPGCI - Programa de Pós-graduação em Ciência da Informação
PRECIS - Preserved Context Index System
PROAC - Pró-reitoria de Assuntos Acadêmicos
PROAD - Pró-reitoria de Administração

PROAES - Pró-reitoria de Assistência Estudantil
PROGEPE - Pró-reitoria de Gestão de Pessoas
PROGRAD - Pró-reitoria de Graduação
PROPLAN - Pró-reitoria de Planejamento
PROPPI - Pró-reitoria de Pesquisa, Pós-graduação e Inovação
PSAS - Protocolo Setorial da Área de Saúde
PSG - Protocolo Setorial do Gragoatá
PSPV - Protocolo Setorial da Praia Vermelha
PSV - Protocolo Setorial do Valonguinho
PUC - Pontifícia Universidade Católica
RAD - Rules of Archival Description
REUNI – Programa de Reestruturação e Expansão das Universidades Federais
SCA - Serviços de Comunicações Administrativas
SCP - Sistema de Controle de Processos
SDC - Superintendência de Documentação
SEI - Sistema Eletrônico de Informação
SETIC – Secretaria de Tecnologia da Informação e Comunicação
SIGAD - Sistema Informatizado de Gestão Arquivística de Documentos
SOC – Sistema de Organização do Conhecimento
SPI - Sistema Protocolo Integrado
SRI - Sistema de Recuperação da Informação
STI - Superintendência de Tecnologia da Informação
STM - Short-term Memory
UBC - University of British Columbia
UERJ - Universidade do Estado do Rio de Janeiro
UFF - Universidade Federal Fluminense
UFRJ - Universidade Federal do Rio de Janeiro
UnB - Universidade de Brasília
UNESP - Universidade Estadual de São Paulo
UNICAMP - Universidade de Campinas
UNIRIO - Universidade Federal do Estado do Rio de Janeiro
UP - University of Pittsburgh
USP - Universidade de São Paulo

SUMÁRIO

1 INTRODUÇÃO.....	13
2 O PROCESSO DE INDEXAÇÃO E SUA APROXIMAÇÃO DA TEORIA DA CLASSIFICAÇÃO.....	27
2.1 A GENEALOGIA DA INDEXAÇÃO.....	27
2.2 A INFLUÊNCIA DO <i>CLASSIFICATION RESEARCH GROUP</i> (CRG) NA TEORIA DA INDEXAÇÃO.....	37
2.3 OS SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO COMPUTADORIZADOS E A TEORIA MATEMÁTICA DA COMUNICAÇÃO NO PROCESSO DA INDEXAÇÃO.....	51
2.4 O COGNITIVISMO NA INDEXAÇÃO E A PERSPECTIVA DO USUÁRIO.....	59
3 DO QUÊ SE TRATA? AS FILOSOFIAS DA LINGUAGEM NA ATRIBUIÇÃO DOS ASSUNTOS.....	73
3.1 O PARADIGMA REALISTA: VER E DIZER O REAL.....	75
3.2 MENTALISMO: A LINGUAGEM COMO INSTRUMENTO DO PENSAMENTO.....	78
3.3 A TRADIÇÃO PLATÔNICA-ARISTOTÉLICA E SUA INFLUÊNCIA NOS ESTUDOS DA LINGUAGEM.....	79
3.4 AS CONTRIBUIÇÕES DOS ESTUDOS DA LINGUAGEM PARA A INDEXAÇÃO.....	85
3.5 AS APROPRIAÇÕES CONTEMPORÂNEAS DO REALISMO E MENTALISMO: UMA ANÁLISE A PARTIR DO <i>JOURNAL OF DOCUMENTATION</i> E O <i>KNOWLEDGE ORGANIZATION</i>	94
3.6 A CRÍTICA DO PRAGMATISMO À FILOSOFIA PLATÔNICA-ARISTOTÉLICA DA LINGUAGEM.....	102
4 O DOCUMENTO, OS ARQUIVOS E SEUS USOS.....	109
4.1 A ESPECIFICIDADE DO DOCUMENTO DE ARQUIVO.....	113
4.2 A GESTÃO DE DOCUMENTOS E SUAS ABORDAGENS NO CAMPO ARQUIVÍSTICO.....	117
4.3 OS USOS DOS DOCUMENTOS DE ARQUIVO.....	126
4.4 A ENTIDADE PRODUTORA E OS PROCESSOS DE TRABALHO.....	128
4.5 OS DOCUMENTOS DE ARQUIVO NATO-DIGITAIS E SUA GESTÃO.....	135
4.6 AS FUNÇÕES ARQUIVÍSTICAS: DIFERENÇAS APROXIMAÇÕES.....	141
5 PROCEDIMENTOS METODOLÓGICOS.....	157
5.1 PESQUISA BIBLIOGRÁFICA.....	157
5.2 LEGISLAÇÃO.....	160
5.3 DICIONÁRIOS, NORMAS E MANUAIS.....	161
5.4 O CAMPO EMPÍRICO.....	163
6 AS UNIVERSIDADES FEDERAIS BRASILEIRAS E O SISTEMA PROTOCOLO INTEGRADO.....	166
6.1 AS UNIVERSIDADES E O ORDENAMENTO JURÍDICO.....	167

6.2 O PROGRAMA DE REESTRUTURAÇÃO E EXPANSÃO DAS UNIVERSIDADES FEDERAIS (REUNI).....	172
6.3 O PROTOCOLO E SEU LUGAR ESTRATÉGICO.....	177
7 RESULTADOS E DISCUSSÃO.....	184
7.1 O SISTEMA PROTOCOLO INTEGRADO E O SEI.....	184
7.2 AS UNIVERSIDADES FEDERAIS E O SEI.....	190
7.3 A UFF E O MAPEAMENTO DOS PROCESSOS DO TRABALHO.....	196
7.4 A INDEXAÇÃO NOS ARQUIVOS DE UNIVERSIDADE: POSSIBILIDADES DE PRODUÇÃO DE SENTIDO A PARTIR DOS USOS DOS DOCUMENTOS.....	208
8 CONCLUSÃO.....	234
REFERÊNCIAS.....	240
ANEXOS.....	262

1 INTRODUÇÃO

Pensemos ainda, em particular, da formação dos conceitos. Toda palavra torna-se um conceito justamente quando não deve servir, como recordação, para a vivência primitiva, completamente individualizada e única à qual deve o seu surgimento, mas ao mesmo tempo tem de convir a um sem-números de casos, mais ou menos semelhantes, isto é, tomados rigorosamente, nunca iguais, portanto, a casos claramente desiguais. Todo conceito nasce pela igualação do não-igual (NIETZSCHE, 1999, p. 56).

O olhar lançado por essa pesquisa ao campo da indexação não possui a pretensão de compreender sua cientificidade, delimitando conceitos, conteúdos, métodos ou conhecimentos produzidos acerca da temática na Ciência da Informação. A proposta está mais associada à insurreição sobre tudo isso. Não se trata de buscar nos saberes acerca da indexação em sua unidade, unanimidade, centralidade ou univocidade, mas o contrário, afastar-se de tentativas nessa direção.

No entanto, tal insurreição não se configura como uma proposta relativista ou niilista de desconstruir uma área do conhecimento ou inaugurar tendências e modismos. Trata-se de caminhar por desconstruções e análises voltadas à compreensão das discursividades em torno da indexação inserida nos campos da Organização do Conhecimento, da Ciência da Informação e da Arquivologia. Assim como propôs Foucault (1998) em seu método genealógico, nossa intenção é compreender as práticas discursivas, as maneiras em que os enunciados são articulados e como os sujeitos escolhem, validam, nivelam, combatem, interditam conceitos, teorias acerca da indexação, dos arquivos, das universidades e de seus documentos. Trata-se, então, de entender como os jogos e as regras das discursividades se dão acerca do nosso objeto analisado.

Dessa forma, nossa análise parte do entendimento de que a indexação é um processo inserido em um conjunto de práticas discursivas que regulam métodos, teorias, domínios, critérios, condutas e objetos que constituem e delimitam a atuação de seus sujeitos sejam pesquisadores, teóricos, indexadores ou usuários de sistemas de informação. Nosso esforço está em procurar compreender como certos enunciados, discursos, conceitos aparecem, são coordenados, articulados, subordinados e transformados em teorias e abordagens científicas, legitimadas como tal pela Ciência da Informação.

Nessa lógica de análise das práticas discursivas, compreendemos a informação à luz do conceito de Frohmann (2004), trata-se daquilo que é informativo para alguém, ou seja, depende do envolvimento dos sujeitos, da conjuntura, dos discursos e de suas validações. Portanto, não pode ser considerada como uma entidade autônoma ou fenômeno isolado em si, atemporal, impessoal e universal. O conhecimento tampouco se trata de uma singularidade a ser desvelada, formado pela sedimentação de axiomas incontestes determinados pela completa compreensão da essência das coisas, cujas mutações são regularmente mapeadas, descobertas e informadas por teorias e métodos racionalmente rigorosos na verificação de suas verdades.

Seguindo as pistas de Arendt (2007), o conhecimento pode ser considerado parte da condição humana, parte do que a autora chamou de labor, a necessidade de inventar do homem, de modificar a sua realidade e está vinculada à necessidade de viver em pluralidade, de transcender à existência individual. Por conseguinte, não é possível produzir conhecimento sem que haja convívio com uma comunidade discursiva e toda ação humana se dá na legislação da linguagem, pois como afirma Nietzsche (2010), a necessidade de viver em rebanho envolve a linguagem em todo fluxo da vida humana. O conhecimento faz parte das relações culturais, sociais, morais, econômicas, religiosas, históricas, jurídicas do homem, logo, do cotidiano, da rotina humana. Não está apartado, ocupando um *telos* em nome de uma racionalidade transcendental.

Outrossim, a indexação é compreendida como objeto inserido na área da organização do conhecimento, campo dedicado a lidar com o desejo do homem em criar ordenamentos sobre as coisas para melhor compreender a si mesmo.

Tal como afirma Benjamin (1987, p. 228) "toda ordem é precisamente uma situação oscilante à beira do precipício", entendemos que a organização do conhecimento pende entre a ordem e o caos. A organização do conhecimento procura delimitar um universo, unificar conceitos e discursos sobre determinado objeto, visa o mínimo de completude muitas vezes relacionada a projetos futuros: novas necessidades, novas estruturas, novos caminhos. Toda organização do conhecimento carrega consigo a pretensão de reproduzir minimamente um todo, ainda que seja necessário o abandono de muitas partes. No entanto, cabe ressaltar que tais ordens impostas pelos Sistemas de Organização do Conhecimento (SOCs) são reflexos dos contextos sócio-históricos que estão inseridos. Eles fazem parte também das relações subjetivas, das afetividades da comunidade discursiva que pertencem. Consequentemente, estes sistemas não estão apartados da dinâmica da vida, os sujeitos misturam-se a esses complexos sistemas.

A necessidade de identificar, mapear, catalogar, categorizar o conhecimento, tem apartado esses sistemas das relações humanas pela tentativa de estabilizar, unificar, centralizar, cristalizar e disponibilizar saberes. Uma quimera que não consegue mascarar as transformações e descontinuidades da própria dinâmica da produção do conhecimento. Sendo assim, a possibilidade de trabalhar com SOCs, no qual o processo de indexação está inserido, converte-se em compreender a interação entre sujeitos, documentos, informações, usos que fazem parte dele. Nesse sentido, podemos incluir as instituições arquivísticas e seus agentes.

Destarte, a partir dessas premissas, construímos a análise do nosso objeto: a indexação na fase de produção dos documentos nos arquivos de universidades. Compreendemos os arquivos como instituições historicamente marcadas pelas relações jurídico-administrativas que seus documentos possuem.

Os documentos de arquivo servem para comunicar, informar, mas também viabilizar ações administrativas, muitas vezes reguladas e legitimadas pelo ordenamento jurídico em que a instituição produtora faz parte. Diversas funções e atividades são realizadas pela mediação desses dispositivos. Então, não são apenas artefatos produzidos por suas instituições de origem para atestar uma ação, mas eles são produzidos porque há imposições, coerções, consensos, múltiplas relações que exigem a sua produção. São dispositivos produzidos, regulados, distribuídos e escalonados por convenções socialmente instituídas. Eles têm a capacidade de estabilizar a informação, preservar e estender no tempo a memória e a evidência daquilo que foi registrado. Logo, atestam ações e transações de seu contexto sócio-histórico.

A Arquivologia para atender às demandas informacionais de seus usuários desenvolveu ao longo dos séculos a descrição como função arquivística capaz de possibilitar o acesso aos documentos de arquivo. A descrição trata de uma complexidade maior, de um todo. Yeo (2013) afirma que a descrição é uma representação do documento formulada pelo arquivista e essa formulação envolve processos, métodos de pesquisa complexos e particulares que convergem para a compreensão de um arquivo. Assim, percebe-se como atividades arquivísticas inseridas no processo de descrição a elaboração de mecanismos de acesso; a explicitação dos relacionamentos entre os documentos de arquivo e as funções e atividades que lhes deram origem; e, também, o controle do acervo.

A preocupação com o contexto do documento de arquivo, característica da descrição enquanto operação de representação da informação, atende a um dos dois níveis presentes nos arquivos, relatados por Rousseau e Couture (1998), a informação do conjunto documental que revela as ações, atividades e funções das atividades produtoras. De acordo com os autores canadenses, há outro nível de informação presente nos arquivos, que é a informação contida no documento em si, isoladamente, relacionada mais ao conteúdo do que ao contexto (ROUSSEAU, COUTURE, 1998, p. 251).

A literatura arquivística ao abordar a fase de produção documental tem dedicado-se às abordagens das funções arquivísticas de classificação e avaliação, constatação que apresentamos ao longo da tese, assim como os motivos que indicam as necessidades de tais estudos como por exemplo a teoria da gestão de documentos voltada para a economia e eficácia no gasto de recursos. Já a representação dos documentos na fase de produção com fins de sua recuperação têm recebido menor atenção em comparativo com as funções mencionadas, embora a indexação de documentos seja uma prática recorrente nos arquivos.

Na fase de produção documental, apesar dos usos dos documentos serem múltiplos e, muitas vezes, simultâneos, inclusive ampliados com a lei nº 12. 527 de 18 de novembro de 2011 que regula o acesso à informação, estes estão mais relacionados às próprias atividades da entidade

produtora. Tais usos estão relacionados aos processos de trabalho e podem facilitar o seu desenvolvimento, ajudam nas tomadas de decisão, permitem o controle fiscal, aprimoram funções, comunicam decisões tomadas, identificam em que momento estão os processos de trabalho, dentre inúmeros usos realizados pelas unidades de trabalho.

O valor primário dos documentos na fase de produção continua sendo uma questão importante aos arquivos, isto é, estão relacionados às utilizações imediatas e administrativas que lhes deram os seus criadores (ROUSSEAU E COUTURE, 1998, p. 117), seu valor está vinculado às razões para as quais os documentos foram criados (SCHELLENBERG, 2006).

O protocolo, unidade de registro surgida na Prússia entre os séculos XVII e XVIII, procura registrar não somente a entrada e/ou produção de um documento, mas o trâmite dos processos de trabalho a fim de facilitar a própria organização dos documentos (FOSCARINI, 2009, p. 24). O registro dos trâmites pelo protocolo também facilita a percepção do contexto de produção e de uso dos documentos dentro da entidade produtora.

O protocolo ao longo do tempo foi se transformando em importante unidade informacional pelo controle da produção documental que realiza, sendo adotado em diversas administrações em países como Alemanha, Canadá, Austrália, Inglaterra e Brasil.

É importante destacar que há duas dimensões relacionadas ao protocolo: uma relacionada à unidade de trabalho, a instância física que faz parte da estrutura da instituição produtora e como as demais unidades de trabalho possui suas atividades específicas. A outra dimensão diz respeito a suas funções em um sistema como o registro, a autuação de processos, a classificação dos documentos, controle da tramitação, a comunicação das ações.

Há uma ampla literatura na Arquivologia cujas abordagens preocupam-se com a organização dos arquivos para evitar o acúmulo de massa documental e viabilizar o desenvolvimento pleno das atividades e funções nas unidades de trabalho. A gestão de documentos tem sido matéria da área desde o início do século XX, demarcada pelo grande fluxo informacional do pós-guerra, e se firmou a partir de teorias e dos modelos desenvolvidos para o tratamento dos documentos de arquivo, sobretudo em sua fase de produção. No entanto, os estudos sobre gestão de documentos destacam-se na análise dos processos de avaliação e classificação dos documentos, visando a eficiência e eficácia na administração da entidade produtora. Essas apropriações, como veremos no desenvolvimento da tese, procuram responder as demandas de economia de recursos, diminuindo os custos com armazenamentos desnecessários e evitando a perda de documentos importantes à administração (INDOLFO, 2007, p. 38).

No século XXI, com a nova realidade tecnológica, o foco dos estudos de gestão de documentos tem sido em manter a confiabilidade do documento de arquivo, considerando as suas possibilidades de migração para diversos suportes e plataformas. Todavia, a aplicação de técnicas de

representação temática dos documentos oriundas da Biblioteconomia e da Ciência da Informação, tal como a indexação realizada em larga escala nos arquivos, sobretudo na fase de produção documental para viabilizar a recuperação da informação nas unidades de trabalho, e, desse modo, fazendo parte da gestão de documentos, ainda não atingiram a mesma proporção na literatura arquivística.

Os processos de trabalho realizados para o desenvolvimento das atividades nas universidades têm exigido cada vez mais sistemas de recuperação da informação cujos procedimentos estejam baseados nos usos dos documentos pelas rotinas administrativas. Estas demandas estão impulsionando programas de gestão de documentos a desenvolverem métodos de representação que possibilitem a recuperação destes. Os processos de trabalho precisam dos documentos e, muitas vezes, do seu conteúdo informacional para se desenvolverem.

Com o advento das novas tecnologias, em particular com a implementação do Sistema Protocolo Integrado (SPI), instituído pela portaria interministerial MJ/MP nº 2.320, de 30 de dezembro de 2014, os órgãos e entidades da administração pública brasileira, no âmbito do poder executivo federal, as atividades têm sido desenvolvidas e registradas em processos eletrônicos. Os protocolos nas universidades, ao registrarem esses processos, estão realizando também a sua autuação na fase de produção, e, por conseguinte, a indexação desses documentos, pois os sistemas de registro estão incorporando esta forma de representação temática do conteúdo dos documentos. No entanto, tal prática não tem sido problematizada pelos estudos da representação da informação na mesma medida.

Abordar a indexação a partir do seu *corpus* teórico, considerando as estratégias de indexação, em suas duas dimensões - a análise conceitual dos documentos e a representação (tradução) de seus assuntos para linguagens documentárias - é fundamental, uma vez que esse processo faz parte dos sistemas de arquivos das universidades, usado como técnica capaz de viabilizar a recuperação do conteúdo de seus documentos.

Destarte, a partir dessa realidade apresentada, a questão de partida desta pesquisa é: como a se configura a indexação nos arquivos das universidades, sobretudo em sua fase de produção documental? Esta indagação desdobra-se em outras como: em que medida a área da Ciência da Informação tem produzido pesquisa e, por conseguinte, literatura acerca dos usos dos documentos em arquivos de universidades? Que tipo de literatura é essa? Quais as teorias da indexação mais debatidas na Ciência da Informação? Elas atendem às demandas informacionais dos usos dos documentos nos arquivos de universidade? Como as técnicas de indexação tendem a ser aplicadas na análise conceitual dos documentos e sua representação visando às necessidades de recuperação do documento nos arquivos de universidade?

Desse modo, destacamos que o objetivo principal desta tese é analisar e compreender o processo de indexação na fase de produção documental nos arquivos das universidades federais brasileiras, considerando os usos dos documentos feitos pela comunidade discursiva. Este se desdobra em objetivos específicos que procuram contemplá-lo como um todo:

- Identificar na literatura as teorias/filosofias que norteiam às práticas de indexação no estabelecimento do assunto dos documentos;
- Apresentar o impacto das filosofias da linguagem no processo de indexação, tendo em vista a dimensão da produção de sentido na análise e seleção de assuntos;
- Elicitar as especificidades das universidades como entidades produtoras e seus processos de trabalho;
- Apontar os contrastes entre a indexação, a classificação e a descrição arquivísticas no tocante à representação dos documentos de arquivo;
- Destacar as relações entre o protocolo, os processos de trabalho e a representação dos documentos;
- Perceber os múltiplos usos dos documentos na sua fase de produção, as demandas informacionais cotidianas e as formas de busca nos sistemas de recuperação da informação nos arquivos de universidade.

Partimos de alguns pressupostos para a construção desta tese, o principal consiste em entender a indexação como um processo usado pelos arquivos das universidades públicas federais brasileiras para a representação das informações contidas nos documentos, visando a recuperação dos mesmos.

A indexação dos documentos de arquivo se dá na fase de produção quando acontece a autuação dos processos, principal tipo documental que registra as atividades e tarefas das unidades de trabalho. No entanto, observa-se que muitas vezes o processo de indexação é baseado no empirismo, na experiência de cada autuador de processo nos protocolos. As universidades não seguem políticas de indexação o que ocasiona perdas informacionais importantes e por vezes inviabiliza as próprias atividades da entidade produtora.

Há pouca literatura acerca da indexação em arquivos na Ciência da Informação e na Arquivologia. O processo geralmente é abordado de maneira mais ampla, considerando o conceito *lato* de documento, ou seja, o registro de informação independente do suporte. Entretanto, há especificidades relacionadas aos documentos de arquivo que precisam ser consideradas a exemplo dos seus usos administrativos na fase de produção. Outrossim, as universidades como entidades produtoras possuem formas administrativas singulares que impactam na própria produção e uso dos documentos na fase corrente. Todavia, tanto na área da Ciência da Informação, quanto na

Arquivologia, a indexação em sistemas de arquivo na fase de produção documental, prática recorrente, tem recebido pouca atenção por parte dos pesquisadores e acadêmicos.

Outro pressuposto é que as unidades de trabalho na fase de produção documental possuem especificidades no uso da informação e dos documentos que refletem na própria busca e precisam ser consideradas na indexação. Este processo deve estar pautado nos usos dos documentos por sua comunidade discursiva. Esses usos da informação e do documento estão vinculados aos processos de trabalho da universidade enquanto entidade produtora. Por isso, é necessário incorporar a compreensão desses processos de trabalho na elaboração das políticas de indexação da instituição.

Outra importante hipótese é que a forma de compreender a linguagem, em particular a produção de sentido, influencia no processo de indexação na atribuição dos assuntos dos documentos e, por conseguinte, na sua eficiência nos sistemas de recuperação da informação.

Há hipóteses secundárias a serem analisadas como: a indexação não possui o compromisso de manter o contexto de produção do documento, embora possa realizá-lo. Este é dever da classificação arquivística, função de organização dos documentos de arquivo que tem como objetivo central garantir a organicidade e o vínculo arquivístico entre eles. O objetivo primeiro da indexação é representar os assuntos dos documentos nos diversos índices sejam estes digitais ou não, para fins de recuperação e acesso dos documentos e suas informações.

Os usos dos documentos na fase de produção documental estão associados, em sua maioria, aos processos de trabalho. Portanto, as formas de busca e os interesses das unidades de trabalho no documento precisam ser considerados no processo de indexação.

Apresentamos então um resumo com a estrutura que norteou o desenvolvimento desta tese:

Questão Geradora:	Qual o papel da indexação na fase de produção documental, considerando os arquivos das universidades?
Objetivo Geral:	Analisar o processo de indexação na fase de produção documental nos arquivos das universidades federais brasileiras
Objetivos Secundários	<ul style="list-style-type: none"> • Identificar as teorias/filosofias que norteiam às práticas de indexação no estabelecimento do assunto dos documentos; • Demonstrar o impacto das filosofias da linguagem no processo de indexação, considerando a dimensão da produção de sentido na análise e seleção de assuntos; • Evidenciar as especificidades das universidades como entidades produtoras e seus processos de trabalho; • Mapear as diferenças entre a indexação, a classificação e a descrição arquivísticas no tocante à representação dos documentos de arquivo; • Destacar as relações entre o protocolo, os proces-

	<p>sos de trabalho e a representação dos documentos;</p> <ul style="list-style-type: none"> • Perceber o múltiplos usos dos documentos na sua fase de produção, as demandas informacionais cotidianas e as formas de busca nos sistemas de recuperação da informação nos arquivos de universidade.
Principal hipótese norteadora:	<p>A indexação é um processo de análise conceitual e representação dos assuntos dos documentos capaz de atender às demandas informacionais surgidas na fase de produção documental. No entanto, sua eficiência está associada ao mapeamento dos usos dos documentos por sua comunidade discursiva.</p>
Hipóteses Secundárias:	<ul style="list-style-type: none"> • O desenvolvimento da técnica da indexação está associada às próprias transformações culturais, econômicas, científicas, tecnológicas ocorridas na sociedade ocidental, na perspectiva maussiana de fato social total; • As teorias da indexação possuem interlocução com a teoria da classificação, sobretudo a corrente inglesa influenciada pelo trabalho do <i>Classification Research Group</i>; • A filosofia da linguagem de inspiração platônica-aristotélica possui grande repercussão nas abordagens acerca da indexação, principalmente naquelas que realizam interlocuções com a linguística na problematização da análise de assunto, em especial com os estudos semânticos; • As universidades como entidades produtoras possuem especificidades em seus processos de trabalho que precisam ser consideradas no mapeamento dos usos dos documentos na fase de produção, pois estão vinculados e definem as demandas informacionais; • A indexação e a descrição arquivística são processos distintos e autônomos que podem estar associados um ao outro ou não. A indexação não está necessariamente subordinada à descrição. • A indexação, por possibilitar a representação da informação de grandes volumes documentais e poder adotar um método de análise conceitual voltado aos usos dos documentos pela comunidade discursiva, apresenta-se como uma importante técnica capaz de viabilizar a recuperação da informação. • No Brasil, com a adoção do Sistema Protocolo Integrado e o Processo eletrônico Nacional pelo Governo Federal, faz-se necessário repensar os sistemas de registro no controle da produção documental. • Urge pensar políticas de indexação que facilitem a recuperação dos documentos pelos arquivos.

O primeiro passo no desenvolvimento desta tese foi identificar as categorias operatórias que serviriam como elementos norteadores da pesquisa. Para delimitá-las procuramos autores que representam escolas de pensamento a fim de trabalhar com a literatura já analisada, avaliada e validada pelos pares, procurando minimizar confusões e equívocos conceituais. Destacamos alguns autores e categorias que fazem parte da fundamentação teórica desta tese.

Nossa escolha metodológica foi apresentar tal fundamentação ao longo das seções conjuntamente com as análises de nosso objeto e dos resultados da investigação, em vez de destacá-la em uma seção exclusiva.

A primeira categoria elencada foi a indexação, para compreendê-la trabalhamos com a perspectiva histórica de Wellisch (2000), com os princípios elencados por Lancaster (2003) e com seus exemplos de aplicabilidade do mesmo, com as análises e propostas de Borko (1967; 1977; 1978), Bernier (1978), Fugman (1985) e Beghtol (1986; 2002), além das propostas inovadoras de Frohmann (2004; 2001; 1990), Hjørland (2002; 1999; 1998; 1992); e Albrechtsen (1999), Barité (2007), Mai (2005) e Tennis (2016; 2012; 2008). Já para compreender a realidade brasileira, recorreremos a Fujita (2013; 2012; 2004; 2003; 1989), Lara (2011; 2007; 2003), Guimarães (2008; 2003) e Cesarino (1978).

Na abordagem sobre os arquivos e documentos de arquivo, buscamos compreender as abordagens de Schellenberg (2006), Duchein (1986), Cook (1989), Posner (1959), Rousseau e Couture (1998), Cruz Mundet (1994), Heredia Herrera (1991), Cortés Alonso e Mendo Carmona (2004; 1995). No Brasil, os trabalhos que nos influenciaram foram os de Bellotto (2006; 2002); Camargo (2003; 2000) e Sousa (2003; 2005). Esses autores também contribuíram para a nossa compreensão das funções arquivísticas e suas diferenças em relação às metodologias encontradas na área da Ciência da Informação.

A gestão de documentos e a teoria do ciclo de vida também foram importantes para o desenvolvimento desta tese, neste caso, autores como Llansó I Sanjuan (2006; 1993), Cruz Mundet (1994), Rousseau e Couture (1998), Rhoads (1989), Alberch I Fugueras (2015; 2004; 2001) serviram como referências. Jardim (2015; 1995; 1987), Lopes (1996), Indolfo (2007) e Rodrigues (2008; 2012) embasaram a compreensão da abordagem brasileira.

Para compreender o protocolo, suas funções na fase de produção documental, inclusive como unidades de informação, recorreremos à Foscarini (2009), Sabourin (2001), Miller (2003), McKemmish (2001), Ketelaar (2005), Menne-Haritz (2005) e Duarte (2009). As universidades, suas funções, atividades e processos de trabalho, além da percepção de sua produção documental e as características de seus sistemas de recuperação da informação foram investigados a partir da leitura de Cohen e March (1986), Yakel e Bost (1994), Bantin e Bernbom (1996), Marsden (1997) e Wilson (1976).

As filosofias da linguagem que ocupam também um lugar de destaque nesta pesquisa foram compreendidas a partir da leitura de autores como Wittgenstein (1999, 1968), Fish (1992), Swales (1990) Austin (1990), Nietzsche (1999), Searle (1979), Bazerman (2005), Chomsky (1968; 1972; 1988;), Chierchia (2003), Frege (1976), Russel (1986), Lakoff e Johnson (2002).

Esse quadro teórico serviu para orientar o desenvolvimento da pesquisa. O olhar lançado às fontes analisadas partiu da compreensão de tais categorias a partir dos autores supracitados, procurando refletir sobre o objeto desta tese, ou seja, a indexação nos arquivos das universidades na fase de produção documental, a partir das dimensões humanas elencadas ao longo da tese - o fazer,

as relações entre os sujeitos e os documentos em sua comunidade, a produção de sentido nessas relações. Logo, procuramos desenvolver uma metodologia de pesquisa que refletisse este olhar.

A indexação mostrou-se como um dos primeiros conceitos que precisamos definir para podermos realizar a nossa abordagem acerca do objeto. Compreendemos a indexação como um processo de análise conceitual e representação da informação de documentos, baseado nos usos que os atores sociais de determinada comunidade discursiva fazem destes. Essa definição está baseada em uma ampla revisão de literatura apresentada ao longo das seções, no entanto, alguns autores como Frohmann (1990), Hjørland (1992; 2002) e Day (2005) foram importantes para a nossa compreensão do processo da indexação.

A aproximação desses autores dos estudos de linguagem, em especial, da corrente pragmatista, influenciou a nossa apropriação de conceitos correlacionados à indexação como documento, informação e comunidade discursiva. Influenciado pelos estudos semânticos, Frohmann (1990) reflete sobre os documentos e a informação em suas relações sincrônicas e não sucessivas. Assim como a significação não é algo representativo de um signo relativo ao mundo, a informação também não o é.

A informação constitui-se pela e na linguagem, se processa nas relações entre os homens, as coisas e os signos. A crítica de Frohmann (1990) à abordagem da indexação centrada no documento, à concepção que é possível extrair os assuntos dos documentos, é um importante referencial, pois nos levou a pesquisar como a área da indexação construiu seus aportes teóricos e a identificar três paradigmas importantes neste campo: o realismo, de inspiração platônica, o mentalismo aristotélico e o pragmatismo.

A abordagem de Hjørland (1992) aproximando o conceito de assunto às abordagens das filosofias da linguagem instigou-nos a relacionar a etapa da análise conceitual do processo de indexação, na qual a apropriação do assunto é fundamental, à compreensão do que seria a produção de sentido em um documento. A partir da leitura de Wittgenstein (1999) acerca da produção de sentido que esta se dá no uso da linguagem, procuramos compreender como se dá a atribuição dos assuntos dos documentos pelos sistemas de informação. Dialogamos dessa forma com o conceito de jogos de linguagem (WITTGENSTEIN, 1999), para enfatizar que a produção de sentido se dá pelos múltiplos acordos simbólicos que a comunidade discursiva em questão possui tacitamente ou não. Neste processo está inserida a indexação, suas fases de análise conceitual e representação temática que precisam pautar-se no entendimento dessa construção dos sentidos que se manifesta em lugares e momentos múltiplos.

Comunidade discursiva é uma apropriação feita a partir do conceito de Swales (1990), aproximando tal conceito da comunidade interpretativa de Fish (1992). A comunidade interpretativa é caracterizada pelo conjunto de atores sociais que compartilham visões de mundo, concepções,

signos, regras, práticas, atividades e estabelecem teias de relações entre eles. Isso pode acontecer de forma expressa ou tácita, consciente ou não, de forma coercitiva ou deliberada. Swales (1990) diferencia a comunidade discursiva da interpretativa justamente pela necessidade de escolha de seus membros. As relações, regras, acordos, formas de agir são conhecidas, delimitadas por contextos específicos como o ambiente corporativo ou acadêmico, e constantemente revalidadas por seus membros. Swales (1990) aponta para a necessidade de persuasão e a vontade de se relacionar em grupo, aderindo aos mesmos fatos sociais totais para a manutenção da comunidade discursiva.

É importante associarmos à indexação aos conceitos de comunidade interpretativa e comunidade discursiva, pois, para Fish (1992) e Swales (1990), o texto não possui significado fora de um conjunto de pressuposições construídas no seio da comunidade que o produziu e onde ele circula. Isso inclui a intencionalidade do produtor do documento, mas não está limitado a esse aspecto. O contexto cultural de seleção, escalonamento, repetição, interdição, validação de práticas discursivas e de suas formas de criação, circulação e acolhimento desses documentos vão moldando os sentidos e suas apropriações. Fish (1992) afirma que a forma de nos apropriarmos de um documento está vinculada ao nosso pertencimento a uma comunidade interpretativa.

Outro importante referencial é a concepção materialista de Buckland (1991) acerca da informação, segundo este autor a informação apenas pode se dar a partir de sua configuração material. É a sua materialidade, o registro em algum suporte que torna possível a Ciência da Informação, uma vez que as atividades relacionadas à informação estão vinculadas ao seu suporte - recuperação, classificação, indexação, retórica, análise estatística, entre outros. Nessa perspectiva, a indexação tem por objeto o documento em seu sentido lato.

De acordo com Foscarini (2009, p. 26), os protocolos são unidades essenciais em um sistema de arquivo confiável, seu uso deve ser obrigatório a todos os órgãos públicos. O registro de um documento no protocolo significa certificar o exato momento em que ele foi produzido ou recebido, identificando-o de modo único, inserindo-o na classificação e realizando a indexação de forma precisa, colocando a unidade dentro de seu contexto processual e arquivístico.

A perspectiva de Bourdieu (2003) também foi importante para compreendermos o protocolo como instância que faz parte do processo de socialização dos sujeitos, ou seja, um espaço que media múltiplas relações sociais e permitem a aparição do *habitus*, que segundo o sociólogo francês trata-se de “um sistema de disposições duráveis e transponíveis que, integrando todas as experiências passadas, funciona a cada momento como uma matriz de percepções, de apreciações e de ações – e torna possível a realização de tarefas infinitamente diferenciadas, graças às transferências analógicas de esquemas” (BOURDIEU, 2003, p. 65). Então, o protocolo revela relações dinâmicas entre instituições e agentes sociais distintamente posicionados em função de sua visibilidade e dos recursos disponíveis aos mesmos. O protocolo dentro dos arquivos evidenciam as relações de

interdependência entre as diferentes instâncias e os agentes sociais. Pode, inclusive, corroborar relações de continuidade ou de ruptura.

A noção de conjunto documental orgânico é o que especifica o arquivo. É nesta perspectiva que se estrutura o conceito de arquivo consolidado pela Arquivologia como sendo "conjunto de documentos de qualquer natureza, reunidos orgânica e automaticamente por um corpo administrativo, uma pessoa física ou jurídica, em razão de suas funções ou de suas atividades. E que são conservados a título de referências" (ROUSSEAU, COUTURE, 1998, p. 52). Por este contexto, justifica-se a afirmativa de Heloisa Bellotto (2006) de que uma das principais características do documento de arquivo é ser presunção de prova. Já Heredia Herrera (1991, p. 30) afirma que o documento de arquivo está marcado por duas coordenadas: o produtor que o gera e a função/atividade que o determina.

Esta abordagem é necessária, pois para pensarmos a indexação dos documentos de arquivo, é fundamental analisarmos os usos que os sujeitos fazem destes. Dessa maneira, identificamos preliminarmente alguns usos historicamente instituídos nos arquivos: o de evidência, o de registro para lembrar, o informativo, o de mediador de relações sociais no sentido de promover processos comunicacionais e viabilizar atividades. Usos historicamente constituídos, oriundos da configuração assumida pelos arquivos ao longo do tempo e das práticas dos sujeitos que produzem e manuseiam tais documentos. Não buscamos a universalidade dos usos dos documentos de arquivo, pois eles podem ser múltiplos, e qualquer tentativa neste sentido já nasceria fracassada, no entanto, a diversidade de autores que apresentam tais usos e em diferentes "universos arquivísticos", instigam-nos a reflexão sobre eles na abordagem da indexação nos arquivos.

Acerca do conceito de sistemas de informação (SI), Saracevic (1974) esclarece que se trata da reunião integrada de diversos elementos diferentes, que podem ser classificados e descritos de inúmeras maneiras para disponibilizar informações. Formado pela combinação de elementos individuais que interagem entre si para produzir um conjunto único de propriedades. Estas incluem (a) redundância de informação (b) associação entre elementos de informação (classificação) (c) ordenação dos elementos de informação e (d) existência de relações "quasi-métricas" entre elementos de informação. Sem estas propriedades não há sistema de informação (ou este não pode funcionar eficientemente e por isso não existirá por muito tempo).

Já os Sistemas de Recuperação da Informação (SRI) tratam da representação, ordenamento, organização e localização das informações registradas em documentos. De acordo com Souza (2006), o armazenamento, a gestão física e/ou lógica dos documentos e de suas representações; a representação das informações contidas nos documentos e expressas pelos processos de indexação e descrição dos documentos; e a recuperação das informações ali contidas e dos próprios documentos armazenados no sistema de informação são funções dos SRIs. Nessa perspectiva, os sistemas de

arquivo devem incorporar tais funções, sobretudo no que se refere à representação da informação e dos documentos visando a sua recuperação por sua comunidade discursiva.

Compartilhamos da compreensão de Cendón (2005) de que as linguagens documentárias revelam diferentes sistemas semânticos. Uma mesma unidade de informação, a exemplo do cabeçalho de assunto, não cobre a mesma noção de um vocabulário para outro, e a gramática das linguagens documentárias é diferente de um sistema para o outro, pois dependem do uso que a comunidade discursiva faz da linguagem em seu sentido *lato*.

O conceito de gestão de documentos é importante para compreendermos a delimitação do campo na fase de produção documental, a partir das abordagens dos autores Llansó e Sanjuan, Mendo Carmona e Jardim, compreendemos que se trata de uma importante teoria arquivística que procura racionalizar a criação, manutenção, utilização e descarte dos documentos, visando a eficiência e eficácia administrativa. Essas abordagens que também incorporam a teoria do ciclo de vida dos documentos, procuram dividir em três fases a administração dos documentos de arquivo corrente (produção), intermediária e permanente. Essa divisão, como destacam os autores citados, serve para facilitar a manipulação e a compreensão da massa documental, além de atender as especificidades de uso dos documentos em cada estágio. De todo modo, não se deve perder de vista a integralidade dos documentos.

Partimos da compreensão de entidade produtora a partir de sua caracterização de sua missão, funções e atividades. Essa perspectiva foi encontrada em Schellenberg (2006) que sugere a compreensão das funções, atividades e atos para entendermos como se configura uma entidade produtora de documentos e sua missão. Paul Saborin também traz a perspectiva funcional para a caracterização da entidade produtora e sua finalidade. Já Foscarini (2009), ressalta a importância dos processos de trabalho para se compreender a entidade produtora e seus documentos.

Abordando as universidades como entidades produtoras, Elizabeth Yakel e Laura Bost (1994) mostram que a maioria das abordagens da Arquivologia sobre arquivos universitários está voltada para a investigação do comportamento e das necessidades dos usuários nos arquivos históricos, no entanto, os administradores, em muitos casos, são a primeira população usuária dos arquivos. As autoras afirmam que os usuários administrativos possuem demandas informacionais diferentes dos usuários dos arquivos permanentes, estas estão relacionadas na execução das suas atividades - as respostas precisam ser imediatas devido às restrições administrativas de tempo no desenvolvimento dos processos de trabalho.

Além da ampla revisão de literatura que permitiu a criação de nosso quadro teórico acerca das categorias operatórias aqui mencionadas e apresentada ao longo das seções para a análise de nosso objeto, realizamos também uma ampla pesquisa bibliográfica na base de dados *Library and Information Science Abstracts* (LISA) e na Base de Dados Referencial de Artigos de Periódicos em

Ciência da Informação (BRAPCI) para compreendermos como o tema indexação tem sido abordado, em especial no campo dos arquivos. Como nosso objetivo era acompanhar os estudos contemporâneos e as tendências, delimitamos nossa pesquisa entre os anos de 2006 e 2017. Como o grau de revocação foi alto e não seria possível analisar qualitativamente, optamos por realizar um estudo nos periódicos que mais publicaram sobre indexação. Dessa forma, analisamos *The Indexer*, *Journal of Documentation* e *Knowledge Organization*, da área da Ciência da Informação e *Archivaria* e *The American Archivist*, do campo da Arquivologia.

Essa pesquisa também analisou os dicionários terminológicos do campo da Arquivologia, as normas ISO de padronização, além dos manuais acerca da descrição dos documentos de arquivo com o objetivo de melhor compreender as diferenças e similitudes acerca da descrição arquivística e da indexação.

Os documentos de arquivo são altamente regulados pelo ordenamento jurídico, nesta perspectiva, realizamos a análise da legislação referente aos arquivos públicos, aos protocolos e ao acesso à informação. Destacam-se a lei 12.529 de 30 de novembro de 2011, sobre acesso à informação, a lei 8.159 de 8 de janeiro de 1991, considerada a lei dos arquivos e a portaria interministerial MJ/MP 2.320, de 30 de dezembro de 2014 que institui o Sistema Protocolo Integrado. Analisamos também as normativas referentes às instituições federais de ensino superior (IFES) e os manuais de protocolo sobre registro e autuação dos processos.

Delimitamos o nosso campo exploratório às universidades públicas federais, usando como campo empírico a Universidade Federal Fluminense no momento de implementação do Sistema Protocolo Integrado (SPI), por meio da incorporação dos processos de trabalho no Sistema Eletrônico de Informações (SEI).

Os aspectos metodológicos são melhor elicitados na seção cinco deste trabalho onde cada etapa é descrita, assim como as fontes que compõe esta análise do objeto proposto. Ressalta-se que a revisão de literatura é desenvolvida ao longo das seções desta tese onde é relacionada com os resultados da investigação e com o desenvolvimento da abordagem de nosso objeto.

A próxima seção dedica-se a abordar a genealogia da indexação, assim como a sua aproximação com a teoria da classificação, em especial a partir dos estudos e propostas do *Classification Research Group*.

2 O PROCESSO DE INDEXAÇÃO E SUA APROXIMAÇÃO DA TEORIA DA CLASSIFICAÇÃO

A multiplicidade das complexas teorias acerca do processo de indexação leva a inúmeros caminhos que por vezes ocasionam omissões ou confusões pela imensidão dos territórios que podem ser explorados sobre tais práticas discursivas, enunciados e diferentes escalas de análise possíveis. Assim, sem a pretensão de apresentar uma totalidade acerca das teorias da indexação, pois a tentativa de uma exegese apenas se converteria em um desejo megalomaniaco não possível de ser concretizado, nossa proposta é realizar um sobrevoo em complexas teorias considerando suas repercussões nas próprias práticas da técnica da indexação.

Esse exercício cartográfico não impede certa simplificação ou mesmo arbitrariedades nas escolhas das teorias a serem abordadas, ao contrário, está bem próximo desses enganos. No entanto, de forma semelhante ao processo de indexação, nossa tentativa é possibilitar índices, acesso, contato com as teorias apresentadas nesta seção, fornecer indícios a partir de um estudo genealógico que não as esgota, apenas procura lançar alguns olhares como convite aos leitores a construir suas próprias compreensões sobre o tema abordado.

2.1 A GENEALOGIA DA INDEXAÇÃO

Nesse percurso genealógico procuramos entender a indexação como uma atividade realizada por sujeitos ao longo da história, entramos em contato com uma vasta literatura que aponta o seu surgimento como uma técnica ligada a objetivos práticos e imediatos. O desenvolvimento desta técnica está associado às necessidades de recuperar partes específicas do conteúdo de um ou mais documentos.

Weinberg (2010), ao apresentar a história da indexação, aponta que a operação de analisar um documento e sintetizá-lo é uma prática encontrada ainda na Antiguidade, tais vestígios podem ser percebidos em hieróglifos que datam de 3.000 a. C. Gil Leiva (2008) afirma que ações de análise e sumarização dos conteúdos de documentos já estavam presentes nos afazeres dos escribas da Baixa Mesopotâmia, em textos hebraicos do mesmo período e também nos códices latinos dos primeiros anos da era cristã. Weinberg (2010) associa inclusive algumas formas de escrita como as por ideogramas à indexação, pois sintetizam ideias, expressões, descrições, etc. O conceito da ordem conhecida (*known order*) para a entrada de índices já era uma prática realizada nos mosteiros para facilitar o acesso aos documentos manuscritos no período da Idade Média (WITTY apud WEINBERG, 2010, p. 2282).

Rouse afirma que a indexação de assuntos foi inventada na França, com as compilações de obras realizadas no século XIII, sendo atribuídas características como uma espécie de redes de referências cruzadas (ROUSE apud WEINBERG, 2010, p. 2283). As bíblias, que passaram a ser

traduzidas do latim para outras línguas no período da Reforma Protestante, também apresentavam índices com indicações dos números de capítulos e versículos (localizadores). Essa técnica, segundo autores como Witty (1973), facilitou a compreensão do próprio livro.

É importante ressaltar que embora as técnicas de análise e sumarização estejam presentes desde tempos remotos, não é possível delimitá-las como indexação, uma vez que seria anacrônico a seu próprio processo de formação. Apenas podemos perceber que a prática de analisar os documentos e condensar a informação é milenar na tradição ocidental.

No entanto, o processo de criar índices para recuperar partes de um documento ou documentos que abordassem o mesmo assunto ganhou maior robustez no século XV com a invenção da imprensa. Nesse período se utilizava numerais romanos como formas de entrada para localizar assuntos no *Incunabulum*¹(WELLISCH, 1994, p. 7).

A história da construção de índices está intimamente relacionada à história do livro e da imprensa, uma vez que o aumento da produção documental e a relação da sociedade com os documentos, sobretudo os textuais, fizeram surgir a demanda de recuperar "partes" destes sem a necessidade de ter que lê-los em sua totalidade.

É preciso compreender essa demanda dentro de um contexto maior como o próprio desenvolvimento do capitalismo, uma vez que a forma de lidar com o tempo e a necessidade de viabilizar atividades mercantis de forma ágil e hábil levaram ao aparecimento de diversas técnicas, inclusive da indexação.

Weinberg (2010) mostra que os primeiros índices foram realizados tanto em documentos em sua unidade, como nos primeiros livros impressos, quanto em conjunto de documentos, como nos primeiros jornais. Todavia, o objetivo era o mesmo, recuperar assuntos "contidos" em tais documentos. Essa configuração já aponta para a temporalidade característica da cultura capitalista de otimização do tempo, poupando recursos e visando maior produtividade.

Thompson (2011), ao analisar o impacto da cultura capitalista no cotidiano familiar, nas tarefas domésticas e nos processos familiares, alerta sobre a transformação ocorrida do tempo da natureza para o tempo do relógio, demonstrando como essa mudança na concepção da temporalidade influenciou os modos dos indivíduos se relacionarem entre si e com o mundo ao seu redor.

Os primeiros séculos do capitalismo na Europa foram marcados pela transformação do tempo da igreja no tempo do mercador, como assinalou Le Goff (1995). Segundo o historiador francês, entre os séculos XIV e XVI, duas concepções de tempo digladiavam-se: de um lado, o tempo do cristianismo consagrado na Idade Média, e, do outro, o tempo pragmatista manipulado

¹ Formato de livro que surgiu nos primeiros anos da imprensa na Europa, seu nome está associado justamente à característica de ser impresso e não manuscrito.

pelos mercadores. A percepção do devir temporal por parte dos clérigos medievais consistia em revesti-lo de obras à salvação pessoal alcançada pela conduta moral do indivíduo, tornando o tempo uma sucessão teleológica - de Gênesis a Apocalipse -, tendo a eternidade como objetivo a ser alcançado. Já o mercador via o tempo como um artefato profano, isto é, parte de sua rotina e elemento essencial na organização de sua vida. O mercador o conquista, impõe-lhe um preço, trata-o como um objeto de cálculo, de projeção de riscos e de realização de lucros: um tempo mensurável, mecanizável, e sobre o qual “agem a inteligência, a habilidade, a experiência e a manha do mercador” (LE GOFF, 1995, p. 54). No capitalismo, a concepção de tempo que passa a predominar, é o tempo do mercador.

A preocupação do tempo na realização de uma tarefa, seja de qualquer natureza, foi intensificando-se e levou à necessidade de sincronização do trabalho. De acordo com Thompson (2011), esta demanda surgiu com as primeiras manufaturas na Inglaterra e foi aumentando durante a fase do capitalismo industrial. Era necessário padronizar e criar rotinas para possibilitar a produção em larga escala, inclusive sincronizando as atividades dos homens com as das máquinas.

Essa forma de conceber o tempo impactou diversos campos da sociedade, uma vez que as relações de produção, as técnicas, estavam inseridas nas relações sociais. A produção e organização do conhecimento também estavam inseridas em tais dinâmicas sociais, sendo oportuno entender a indexação como um processo dentro de contextos mais amplos. A lógica da economia e otimização do tempo pode ser percebida na construção de índices surgidos para poupar o tempo do leitor interessado em assuntos específicos e não no documento como um todo.

As transformações na forma de compreender o mundo pelo impacto das obras de Maquiavel, Copérnico, Newton, Descartes, Galileu, Hobbes, Bacon, Leibniz entre outros também inauguraram um tempo em que há uma necessidade da investigação os fenômenos das ciências naturais, suas propriedades, assim como a humanidade em si e suas relações com tais fenômenos. O Renascimento Cultural que deslocou as atenções de Deus para o homem, criou a necessidade de estudos sobre a sua “natureza”.

A chamada Primeira Revolução Científica ficou marcada pela releitura das filosofias platônica e aristotélica em diversos campos do saber. Os documentos milenares, diversas vezes traduzidos pelos mouros, pelos monges católicos, passaram a ser consultados com uma frequência considerável. Ainda que seu acesso fosse para poucos privilegiados como os próprios membros do clero ou nobres com grande influência junto a seus reis, a clausura e o sigilo em que os documentos estavam submetidos passaram a ser colocados em xeque pelas forças políticas que atuavam no período como a lógica capitalista, o poder dos monarcas e o próprio desenvolvimento dos saberes (DUARTE, 1983).

O período da Revolução Científica, entre os séculos XVI e XVIII, também ficou marcado pelo surgimento das primeiras sociedades científicas, tratavam-se de associações com o objetivo de promover a pesquisa científica, criando consensos entre os pares sobre teoria e métodos a serem empregados em suas áreas. Uma teoria científica, para se afirmar, necessitava de ser comunicada a um conjunto de sábios que a podiam criticar livremente.

A *Accademia dei Lincei*, em que Galileu foi membro e nela divulgou as suas descobertas astronômicas; a *Accademia del Cimento*, fundada em Florença, em 1657, que conseguiu reunir a maior coleção de equipamentos científicos da época ao longo dos seus dez anos de existência, destacando-se as mais variadas experiências nas áreas da pneumática, som, magnetismo e movimento; a *Royal Society*, criada em Londres, em 1645, que possuía um grupo de personalidades cientistas que mantinham correspondência com os principais cientistas europeu; a *Académie des Sciences*, criada em Paris em 1666, para um reforço da importância da investigação científica francesa; e em Berlim, Leibniz criou em 1700 a *Akademie der Wissenschaften*, patrocinada pelo rei de Brandeburgo-Prússia, Frederico I são alguns exemplos dessas sociedades do conhecimento que faziam circular suas “descobertas” entre si e estão na genealogia da disciplinarização dos saberes investigada por Foucault (2005).

A divulgação das ideias científicas debatidas por essas sociedades era realizada pela publicação de livros, normalmente escritos na língua franca da ciência – o latim –, pela correspondência trocada entre seus membros e pelos periódicos científicos surgidos no século XVII. Destacam-se o francês *Journal de Savants* e o britânico *Philosophical Transactions*, que tiveram suas primeiras edições no ano de 1665 (PEIXOTO, 1997). Essa dinâmica de circulação dos novos conhecimentos levou à necessidade da seleção de conteúdos, inclusive para diminuir os custos da tradução e publicação em diversos idiomas.

Nessa nova mentalidade também emergia a necessidade de padronização e sistematização das técnicas de trabalho. Nesse sentido, a sumarização de conteúdos e construção de índices também passaram a ser pensadas a partir de regulações, padronizações e controle para sua melhor eficácia, processo que se aprofunda nos séculos XVIII e XIX com a expansão do capitalismo e a disciplinarização dos saberes.

O século XVIII foi o século da disciplinarização dos saberes, ou seja, a organização interna de cada saber como disciplina tendo, em seu campo próprio, a um só tempo critérios de seleção que permitem descartar o falso saber, o não-saber, formas de normalização e de homogeneização dos conteúdos, formas de hierarquização e, enfim, uma organização interna de centralização desses saberes em termo de um tipo de axiomatização de fato. Logo, organização de cada saber como disciplina e, de outro lado, escalonamento desses saberes assim disciplinados do interior, sua intercomunicação, sua distribuição, sua hierarquização recíproca numa espécie de campo global ou de disciplina global a que chamam precisamente de a “ciência” (FOUCAULT, 2005, p. 208).

Foucault (2005), ao abordar a disciplinarização dos saberes no século XIX, esboça as etapas em que estes passam a ser escalonados, distribuídos e nivelados de acordo com as relações de poder dos sujeitos históricos, sobretudo das instituições, como o próprio Estado. De acordo com o filósofo, o Estado intervém direta ou indiretamente, mediante a quatro procedimentos por meio de suas instituições educacionais, culturais, ou mesmo de correção: primeiro, a eliminação, ou seja, a desqualificação do que é considerado inútil e irreduzível, para economizar esforços dispendiosos, isto é, de saberes considerados menores; segundo, a normalização dos saberes considerados pertinentes, no sentido de ajustá-los uns aos outros, criando uma comunicação entre os mesmos e os homogeneizando, aproximando técnicas; terceiro, classificando tais saberes, estabelecendo hierarquias entre eles - saberes formais, gerais, específicos, materiais; e todos esses procedimentos permitem o estabelecimento do quarto que é a centralização piramidal, esta permite o controle dos saberes, realizando seleções, exclusões, interdições, padronizações, nivelamentos, fixações e transmissões.

As enciclopédias são citadas por Foucault (2005) como exemplos desse movimento de organização dos saberes tecnológicos a partir de práticas institucionais de controle. Podemos também mencionar os índices e catálogos, na medida em que diversas operações de seleção, exclusão, escalonamento, distribuição e controle são realizadas a partir de critérios que correspondem igualmente a um empreendimento de normalização de saberes técnicos.

Antes do século XVIII não existia "a ciência", mas ciências, no plural, saberes, técnicas e questionamentos que se comunicavam entre si e que constantemente se questionavam por meio da filosofia. Funcionavam como um sistema de comunicação entre os diversos saberes, a exemplo das academias científicas, uns em relação aos outros. No entanto, a partir do século XVIII surgiu o conhecimento-verdade, um tipo de saber disciplinado no qual suas práticas e conteúdos passaram a ser selecionados, regulados e hierarquizados. Essa forma de disciplinar os saberes, centralizando-os, hierarquizando-os, distribuindo-os em um campo global, fez surgir a grande Ciência.

A ciência, como domínio geral, como policiamento disciplinar dos saberes, tomou o lugar da filosofia, e passou a formular problemas específicos ao policiamento disciplinar dos saberes: problemas de classificação, problemas de hierarquização, problemas de vizinhança, etc. (FOUCAULT, 2005, p. 218)

A demanda por parte da Ciência de desenvolver novos conhecimentos baseados em "fontes confiáveis" levou ao surgimento de práticas de representação temática para desenvolver as consideradas "grandes áreas do conhecimento" neste período como a medicina, a astronomia, a matemática, a química, a biologia e a física. A virada do século XVIII para o XIX ficou marcada

pelo cientificismo e por filosofias que propunham abordagens a partir de critérios, regulações e padronizações dos fenômenos naturais e sociais tal como o positivismo (FOUCAULT, 2005).

Para o positivismo, os fatos fossem eles naturais, fossem eles sociais obedeciam a lógicas inatas, leis universais que poderiam ser identificadas e compreendidas por seus pesquisadores a partir de observações, análises, experimentos, construções de modelos que chegassem à comprovação de suas regularidades, seus funcionamentos, seus regimentos operativos. Caberia ao pesquisador o papel de intérprete de tais leis, de descobridor das leis que regiam o fato abordado (COMTE, 1988, p. 23). Não haveria espaço para a crítica, o julgamento, as adjetivações, pois o que importava à Ciência Positiva era a compreensão dos fenômenos tal como eles eram, em sua natureza, sem a interferência externa. A verdade dos fatos se revelaria a partir do estudo minucioso de sua formação e regularidade (COMTE, 1988, p. 23).

Segundo Coblans (1957, p. 42), neste período ocorre a separação da Biblioteconomia ortodoxa da Documentação, uma vez que o corpus epistemológico e a práxis permitem o surgimento de críticas e de discursos acerca da profissão. O saber biblioteconômico passa a ser considerado algo erudito, generalista e voltado para um aspecto “civilizatório” e educacional, aos moldes da mentalidade da era vitoriana. Já a documentação era considerada como sinônimo do desenvolvimento de técnicas para o aprimoramento da construção de bibliografias por assunto aplicada principalmente no controle de periódicos científicos e técnicos.

Para Bradford (1961), Langridge (1977) e Rayward (1997), os processos de catalogação, classificação e indexação não eram tão delimitados, com fronteiras tão visíveis como ficaram evidentes a partir da segunda metade do século XX. Esses autores, ao abordarem o empreendimento de Otlet e La Fontaine em procurar sistematizar a área da Documentação, demonstram que o projeto inicial, datado de 1892, tinha como meta a organização bibliográfica da produção científica no âmbito internacional. Então, foi a partir da ideia de realizar o controle bibliográfico universal que foi criada a Classificação Decimal Universal (CDU). Ela foi elaborada para atender às necessidades de tratamento da informação do Repositório Bibliográfico Universal, idealizado por Otlet e La Fontaine e depois mantido pela Federação Internacional de Documentação (FID).

De acordo com Silva e Fujita (2004, p. 136), o conceito de indexação surgiu a partir da elaboração de índices num período em que não estava tão demarcada a separação entre representação descritiva e representação temática. Bell (1997), ao analisar a história da indexação no Reino Unido, mostra que no século XIX a catalogação por assunto já demonstrava ser insuficiente para atender as necessidades de organização da produção científica. As bibliografias nacionais em países como França, Inglaterra, Alemanha (Prússia), Holanda e Estados Unidos e as universais passaram a ser produzidas com maior frequência.

As circunscrições entre os campos da bibliografia, da catalogação e da indexação ficaram mais evidentes na segunda metade do século XX, com as transformações na produção do conhecimento e no campo informacional após a Segunda Guerra Mundial. Ainda assim, Lancaster (2003, p. 12) menciona que o termo indexação fora tratado como sinônimo de descrição bibliográfica pelo *Oxford English Dictionary*, em 1968.

Segundo Mey (1987, p.5), os códigos de classificação norte-americanos durante muitos anos abordavam o ângulo conteúdo como parte da representação descritiva. No entanto, como apenas pontos de acesso para assuntos. Tais pontos faziam parte de uma descrição do todo que incluía entradas como o nome do autor, o título da obra, editores, tradutores, local de produção e ano, ou seja, os elementos principais de uma descrição bibliográfica. "A classificação e a indexação, embora componentes da catalogação, alcançaram desenvolvimento próprio, com grandes avanços relacionados à teoria da informação" (MEY, 1987, p. 4-5).

Mesmo que, em sentido amplo, o termo catalogação envolva todos os procedimentos para a representação de um item em catálogos, bibliografias ou outros produtos provenientes da análise documental, com o tempo, o termo catalogação foi comumente atribuído apenas à descrição bibliográfica e aos pontos de acesso de título e de responsabilidade. Portanto, a representação da informação pode ser subdividida em representação descritiva e representação temática. A primeira representa as características específicas do documento, denominada descrição bibliográfica, que permite a individualização do documento. A segunda detém-se na representação dos assuntos dos documentos a fim de aproximá-los, tornando mais fácil a recuperação de materiais relevantes que dizem respeito a temas semelhantes (MAIMONE; SILVEIRA; TÁLAMO, 2011, p. 28)

Os estudos de Cutter nos Estados Unidos, sistematizados em sua obra *Rules for dictionary catalog*, de 1876, caracterizam-se também por seu esforço em definir princípios para elaboração de regras para serem usadas na representação dos documentos visando o uso dos catálogos de obras. Autores que teorizaram sobre a indexação como Vickery, Coates, Foskett e Lancaster reconheceram o pioneirismo do bibliotecário norte-americano em procurar estabelecer padrões para a catalogação de assuntos, vislumbrando a possibilidade de acesso aos documentos também por seus conteúdos.

Embora tenha criado apenas 29 regras para determinar os assuntos das obras até a quarta e última edição publicada em vida das 369 que elaborou no total para a catalogação, como destaca Martinho (2010), Cutter e sua proposta de sistematização dos cabeçalhos de assunto exerceram grande influência e estimularam o desenvolvimento de diversos estudos no campo da indexação a exemplo das teorias elaboradas pelos autores citados e elicitadas nesta seção.

Os princípios sindético, de especificidade e de uso, previstos em sua obra, continuam servindo de inspiração para a abordagem dos documentos visando a sua indexação. Ainda que a preocupação de Cutter tenha sido no que se considera hoje como representação descritiva (MEY,

2003), procurando padronizar a catalogação dos documentos pensando a suas múltiplas formas de acesso e, por conseguinte, de entradas.

Silva e Fujita (2004) destacam a importância da sua obra em identificar pela primeira vez as limitações da indexação alfabética. Esta questão foi retomada por autores ingleses como Vickery, Coates, Foskett que passaram a articular a teoria da classificação com a indexação na tentativa de mitigar os problemas da indexação alfabética.

Como veremos, essa aproximação dos estudos da classificação com os da indexação, sobretudo com o *Classification Research Group* (CRG), possibilitaram o surgimento de diversas abordagens acerca do tratamento temático da informação e do documento.

A ordem alfabética guia o critério sistêmico. Cutter aproximou a construção de índices ao processo de classificação, sugeriu, de fato, que os termos para um índice de assunto, e suas interconexões, fossem tomados das tabelas de classificação (VICKERY, 1955, p. 3). Os defensores do índice de assunto alfabético foram levados a introduzir algum tipo de mapa classificatório para resolver seus problemas. Há outro aspecto do índice alfabético a ser considerado - o que Cutter chamou de 'nome composto', que se tornam mais frequentes à medida que se passa da indexação de livros para a indexação de artigos. A regra de Cutter deixa o indexador (e o usuário) julgar, em cada caso, qual é a palavra "mais significativa". Ele acrescenta: "Deve-se confessar que esta regra é um tanto vaga" (CUTTER apud VICKERY, 1955, p. 3).

Cutter, como afirma Cesarino (1978, p. 274), foi o primeiro a sistematizar os princípios que deveriam ser considerados na elaboração de índices, dentre eles podemos citar o princípio de uso - que considera os usos dos documentos; o princípio sindético - construção de uma rede de referências cruzadas compreendida como estruturas sindéticas a fim de resolver o problema recorrente nos cabeçalhos de assunto que são as entradas em ordem alfabética; e o princípio específico - o termo a ser escolhido deve ser o mais específico para a determinação do assunto.

Vickery (1950) ao abordar a indexação alfabética, afirma que seu êxito está no uso de uma boa classificação, pois revela a hierarquia dos assuntos, as classes, as subclasses e suas interrelações. Segundo o autor, a primeira proposta da indexação é indicar a existência e a localização de cada item de informação no campo em que é designado. Essa missão é estabelecer contato entre o usuário/pesquisador (*seacher*) e a informação que buscou, exatamente, exaustivamente e diligentemente.

Vickery aponta que uma indexação alfabética só pode esperar satisfazer as exigências de um índice exaustivo se: "(a) seu desenvolvimento de síntese considerar uma extensão muito grande, e (b) se a síntese for baseada em uma extensa ordem relacional oculta. A análise lógica é absolutamente necessária para a consistência e uniformidade das entradas dos assuntos" (VICKERY, 1950, p. 147). Isto, segundo o autor, justificaria o uso generalizado nos Estados Unidos

de listas padronizadas e publicadas de cabeçalhos de assunto, que presumivelmente já foram analisadas, sendo as inter-relações dos termos (bem ou mal) exibidas por referências cruzadas.

Destarte, percebemos a aproximação da teoria da classificação dos estudos de indexação, desde o início do século XX e se aprofundado a partir do pós-guerra. É importante destacarmos o pioneirismo das propostas de Cutter para indexação, a qual nomeava de catalogação de assunto, ainda no final do século XIX que estavam em sintonia com o momento de transformações das tecnologias, das técnicas de trabalho, das formas de produção do conhecimento e das maneiras das pessoas se relacionarem entre si e com o mundo ao seu redor que refletiram em diversas áreas, inclusive na Biblioteconomia e Documentação que passam a se diferenciar cada vez mais.

As inovações científicas e tecnológicas, além do desenvolvimento da chamada administração científica e o aumento da burocracia estatal, levaram ao aumento exponencial da documentação. Como demonstra Buckland (1996), o século XX ficou marcado pelo aprimoramento dos processos documentários. Ortega (2014) ao abordar a importância dos estudos e suas aplicações da Escola de Chicago, destaca que estes procuravam deslocar o foco nos processos para a função, procurando consolidar modelos, tecnologias, técnicas e habilidades de gerenciamento que fornecessem um efetivo e eficiente serviço de biblioteca no tocante ao acesso e recuperação dos documentos.

O capitalismo contemporâneo do século XX incorpora os conhecimentos científicos aos processos produtivos. As corporações industriais passaram a ter uma considerável dependência tecnológica para executar suas tarefas e viabilizar sua produção. A incorporação de sistemas de comunicação exigia o processamento de um número considerável de informações. Como demonstrou Bachelard (1986), o desenvolvimento dos transportes e da comunicação além de ter possibilitado o próprio desenvolvimento da Ciência, demandou também novas formas de conhecimento e tecnologias para a expansão das duas áreas e do próprio capitalismo. O telégrafo e as ferrovias são marcas da expansão do capitalismo norte-americano que para seu funcionamento dependia de meios rápidos e eficientes do escoamento da produção e do tráfego de informações que garantissem tal logística (TAIULE, 1982).

A produção em larga escala da indústria de bens duráveis que levou à quebra da Bolsa de Nova Iorque em 1929, exigiu um complexo sistema produtivo que contava com a criação de núcleos de informação, padronização das formas de comunicação, desenvolvimento tecnológico e aprimoramento das técnicas de trabalho visando a economia dos recursos, a otimização do tempo, a eficiência da mão de obra para que as corporações garantissem o lucro em seus negócios. Assim, a segunda fase do capitalismo industrial, marcada pelo desenvolvimento dos meios de comunicação (telégrafo, telefone, televisão, cinema), da energia elétrica, dos meios de transporte (trens, aviões, automóveis) e da farmacologia (vacinas, antibióticos e novas substâncias químicas) exigiu do

campo da Documentação o aprimoramento das técnicas de organização, tratamento e recuperação da informação que acompanhassem o acelerado desenvolvimento dessas áreas.

De acordo com Saracevic (1996), a II Guerra Mundial levou ao aprofundamento dessas demandas no que se refere ao desenvolvimento de áreas estratégicas como a indústria armamentícia, a incorporação dos primeiros computadores no processamento de dados, e o desenvolvimento de programas estratégicos financiados por governos como o norte-americano e o britânico para controlar a explosão informacional e documental do período, sobretudo na ciência e na tecnologia.

A lógica estratégica original que fundamentou tais programas e esforços era a seguinte: uma vez que a ciência e a tecnologia são críticas para a sociedade (por exemplo, para a economia, saúde, comércio, defesa) é também crítico prover os meios para o fornecimento de informações relevantes para indivíduos, grupos e organizações envolvidas com a ciência e a tecnologia, já que a informação é um dos mais importantes insumos para se atingir e sustentar o desenvolvimento em tais áreas (SARACEVIC, 1996, p. 43).

A nova temporalidade marcada pela velocidade do sistema produtivo, pelo intenso fluxo de informações, pela diminuição das distâncias, pela compressão do tempo que começou a se delinear no início do século XX, consolida-se com a II Guerra Mundial. A necessidade da logística para o avanço das tropas e seu abastecimento, da codificação das mensagens que permitissem as informações chegarem aos aliados e não aos inimigos, o desenvolvimento de armas de destruição em massa, o conhecimento estratégico da região a ser invadida e ocupada, todas essas ações demandavam o desenvolvimento de conhecimentos específicos e em uma velocidade nunca antes demandada.

Vannemar Bush, em 1945, identificou o problema da “explosão informacional”, apontando que o crescimento exponencial da informação e dos documentos, aliada a demanda da ciência e tecnologia para sua recuperação eficaz e eficiente (SARACEVIC, 1996). Essa figuração permanece no contexto da Guerra Fria com as disputas política, econômica, ideológica, tecnológica entre os Estados Unidos e a União Soviética. Estes governos investiram massivamente para o desenvolvimento de áreas estratégicas como as engenharias naval, elétrica, química, de comunicações, de transportes, na medicina, na tecnologia.

A produção do conhecimento nessas áreas e a disputa pela primazia dos "grandes feitos" e das "grandes descobertas" demandou do campo emergente, a Ciência da Informação, o desenvolvimento de técnicas e sistemas de representação e recuperação da informação, que incorporassem as tecnologias do período - como a inclusão das máquinas no processamento da informação e de seus registros. Essa tendência espalhou-se pela Europa, a exemplo da Inglaterra que também passou a explorar formas de tratamento, organização e representação da informação e dos documentos.

Na próxima subseção apresentaremos algumas das principais abordagens acerca da indexação e a formação de seu corpus teórico.

2.2 A INFLUÊNCIA DO *CLASSIFICATION RESEARCH GROUP* (CRG) NA TEORIA DA INDEXAÇÃO

O *Classification Research Group* (CRG), formado na Inglaterra em 1952 por A. J. Wells e B. C. Vickery, surgiu a partir das discussões acerca da classificação bibliográfica levantadas na *Royal Society Scientific Information Conference* de 1948. Sua proposta inicial era reunir bibliotecários interessados em discutir os sistemas de classificação de suas instituições, sobretudo a aplicabilidade de princípios classificatórios.

O grupo sentia a necessidade de compartilhar suas experiências na organização sistemática de documentos, divulgando abordagens teóricas e aplicações práticas bem sucedidas. Os debates registrados em seus boletins foram importantes referenciais para os estudos do tratamento temático da informação, em especial para a indexação.

Segundo Foskett, os membros do CRG "[...] estavam ansiosos para explorar a base da indexação de assuntos a partir das discussões e do compartilhar das experiências" (FOSKETT, 1962, p. 127, tradução nossa). Nomes como Vickery, Farrandane, Faithorne, Coates, Foskett, Langridge, Austin e Kyle procuraram analisar e debater a teoria da classificação, os esquemas classificatórios e suas operações no âmbito das práticas realizadas nas instituições que trabalhavam.

Dessa forma, os princípios da classificação facetada, o uso de atributos, os termos coordenados, a classificação cruzada, a ordem combinada de facetas foram alguns temas abordados com o objetivo de organizar a documentação, visando a representação e recuperação da informação. O grupo ficou ativo até 1968 com diversos artigos de seus membros publicados neste período, além do registro de suas discussões em seus boletins anuais.

Guimarães (2003) associou esse movimento à corrente do tratamento temático da informação que o autor denominou *indexing*. De acordo com o autor, a proposta do CRG era problematizar e elaborar metodologias de tratamento temático da informação. O primeiro esforço estava centrado em questões de representação temática e a avaliação da performance dos instrumentos tais como as linguagens documentárias. Os tesauros receberam uma atenção especial por possibilitarem a combinação dos cabeçalhos de assunto - que permitiam o acesso do assunto dos documentos por meio dos índices alfabéticos - com os catálogos sistemáticos - que apresentavam a estrutura hierárquica dos conceitos a partir do sistema classificatório adotado (GUIMARÃES, 2003, p. 107).

Em seu quarto boletim, o CRG, ao debater os problemas relativos à recuperação da informação, aproximou a classificação da indexação. Era consenso no grupo que a recuperação da

informação deveria ser consequência do tratamento temático dado aos documentos no momento da entrada nos sistemas. Nesse sentido, a base de um sistema de recuperação de informação seria uma rede de termos interligados.

As interconexões de cada termo seriam potencialmente ilimitadas, e cada sistema de recuperação deveria selecionar certas interconexões para exibição. Esta seleção baseava-se em postulados quanto aos tipos de termo e de relação que deviam ser exibidos - postulados relativos ao nível semântico dos termos, categorias de termos a serem utilizados, relações genéricas, coordenadas e conjuntivas a serem exibidas, os tipos de operação de busca a serem conduzidos. A primeira meta estabelecida para representar os documentos visando sua recuperação foi explorar a variedade de postulados que poderiam ser consolidados e avaliar a sua adequação em vários sistemas de recuperação da informação (Classification Research Group Bulletin, 1958, p. 137).

Destarte, nesse primeiro momento, o CRG dedicou-se em analisar os problemas semânticos na entrada dos documentos em um sistema, associando o processo de indexação à classificação. O CRG, ao discutir a técnica de análise de assuntos, deliberou que "a maneira correta de indexar um assunto específico é analisá-lo em termos classificatórios - conceitos que têm uma interpretação contemporânea e reconhecidamente sem ambiguidade - e mostrar as relações entre esses termos" (Classification Research Group Bulletin, 1958, p. 137). Para o grupo, a forma mais útil de esquema de classificação para a indexação seria aquela capaz de agrupar os termos em categorias bem definidas, que poderiam ser usadas de maneira independente para formar compostos.

Chegou-se à conclusão que haveria pelo menos duas maneiras de expressar relações entre termos classificatórios: (i) inserindo operadores entre os termos que atribuiriam significados diretos a essas relações, e (ii), agrupando os termos em categorias, citando estes em uma ordem particular que permitisse inferir a relação entre eles a partir do contexto. Estas duas técnicas de análise não são mutuamente exclusivas, e poderia ser desejável utilizar ambas.

Faithorne, uma das lideranças do grupo, ampliou a discussão acerca das técnicas que aproximavam a indexação da classificação, contrastando a identificação de um documento por descrição e por relevâncias. A descrição foi relativamente estável, no entanto, de acordo com o pesquisador, as relevâncias variaram de acordo com tempo e lugar (Classification Research Group Bulletin, 1958, p. 141).

A aproximação do processo de classificação da indexação feita pelo CRG baseou-se nas discussões sobre a aplicação do esquema de facetas, elaborado na *Colon Classification*, em 1933, por Ranganathan. Desde de suas primeiras reuniões, os membros do CRG relatavam a suas experiências e propunham métodos que pudessem ser mais efetivos no tratamento temático dos documentos, visando a sua entrada nos sistemas.

Kyle, por exemplo, procurou trabalhar com a questão geral das divisões de forma. Ela levantou práticas possíveis para a entrada dos documentos considerando a forma: a) realizar as subdivisões dominantes, de maneira a recolher todos os materiais sob formas especiais, quer no início (materiais de abordagem), quer no final (materiais suplementares) de todo o índice, (b) torná-los secundários ao assunto, mas assim mesmo anteriorizando ou posteriorizando sua inclusão nos SRIs (Classification Research Group Bulletin, 1958, p. 142).

A preocupação do CRG nas décadas de 1950 e 1960 estava centrada em elaborar princípios para o estabelecimento das hierarquias nos sistemas classificatórios que fosse facilmente compreendidas por seus usuários.

No sexto boletim, o grupo alertava para os problemas na elaboração de classificações especializadas, particularmente questões de ordem das facetas, que eram resolvidas por referência à natureza do assunto central ou ao ponto de vista predominante de uma “clientela” relativamente homogênea. No entanto, nenhum destes fatores estava disponível para uma classificação geral e era evidente que seria necessário um novo princípio para ordená-la.

O principal problema diagnosticado foi que o estabelecimento de hierarquias para os sistemas classificatórios baseado em um grupo de usuários desconsiderava, trazia prejuízos para todos os outros. Kyle, então, buscou um princípio independente, que, embora não produzindo ordem de máxima utilidade para qualquer grupo de usuários, seria facilmente compreensível para todos aqueles que acessassem o sistema (Classification Research Group Bulletin, 1961, p. 157). A classificação geral exigiria do indexador que reorientasse os assuntos de determinado campo específico para uma visão mais ampla.

As facetas eram exaustivamente discutidas considerando sempre sua aplicação em sistemas de recuperação de informação. Ao procurar qualificar as categorias “coisa” e “atividade”, o CRG propôs o uso de níveis integrativos. O grupo percebeu a existência de "formas sucessivas de ordem numa escala de complexidade e organização" (Classification Research Group Bulletin, 1961, p. 158) e de uma sucessão de níveis em que as agregações se tornavam inteiras, agindo como unidades envolvidas em processos característicos de níveis particulares.

Foi no âmbito do CRG que surgiu a discussão que resultou na teoria dos níveis integrativos, consolidada depois por J. K. Feibleman e Joseph Needham, oferecendo uma repartição preliminar viável da lista de coisas físicas e materiais, a partir da ordem das formas sucessivas, embora ainda não fique claro no acompanhamento das discussões do grupo como as abstrações, ou as formas de energia seriam montadas na série de níveis.

De acordo com Foskett (1962), esta teoria foi debatida ao longo da década de 1950 nas ciências naturais e sociais, sob vários nomes, e provavelmente provém originalmente de teorias baseadas no darwinismo e na evolução da biologia. Reduzido para termos básicos, significa que o

mundo das coisas se desenvolveria a partir de estruturas simples em direção às complexas, acumulando propriedades novas e divergentes. Em certos pontos específicos ocorreriam então mudanças que transformariam a entidade de um membro de um grupo ou classe em um membro de um novo grupo. A nova entidade teria propriedades próprias, características do novo nível de organização dentro dele, e se comportaria de uma maneira similarmente nova e característica. Não poderia, sem perder a sua identidade, ser reduzida às suas partes constitutivas que existiriam em um nível mais baixo.

Não é, por assim dizer, esclarece Foskett (1962, p. 136) que um átomo eventualmente cresça em uma molécula pela simples evolução; é mais que uma agregação de átomos, começa-se a desenvolver modos de comportamento que a transformam de uma agregação de átomos em uma molécula, e o comportamento de uma molécula não é explicável sem resíduo em termos de átomos sozinhos, pois suas transformações levaram a surgir uma nova entidade.

A teoria de níveis integrativos propunha soluções para a questão dos termos compostos baseados em termos isolados extraídos de classificações existentes em várias classes principais. Este era um dilema frequente na classificação de uma bibliografia geral ou na construção de índices. Segundo os relatos do sexto boletim do CRG, tal teoria surgiu pela necessidade de encontrar uma solução acerca de tais termos compostos que não era oferecida por nenhum tipo de faceta dentro de um sistema de classe que usasse a *Colon Classification*.

Dessa forma, o fator semântico foi um tópico que surgiu pela primeira vez nas discussões em grupo sobre classificações especiais e apareceu novamente em conexão com a proposta de classificação geral. A teoria dos níveis integrativos realiza essa aproximação com a semântica ao analisar o significado de um termo complexo em um número de termos mais generalizados que denotam o próprio conceito. O valor da análise semântica para fins de recuperação é que ele pode ser feito para exibir as relações existentes entre os conceitos que possuem propriedades em comum.

Segundo os apontamentos do CRG, uma cadeia de classes genéricas que são reveladas a partir de um conceito específico representa uma série de propriedades do conceito em si e, conseqüentemente, representa uma análise semântica (Classification Research Group Bulletin, 1961, p. 160).

O CRG considerava as técnicas de faceta uma ajuda real na análise sistemática de documentos de áreas técnicas. O indexador logo formaria o hábito de classificar rapidamente o assunto em categorias, selecionando os números de classe de unidade e combinando os símbolos de acordo com a ordem preferida da comunidade de usuários de seu SRI. Este procedimento, juntamente com a construção semiautomática do índice de assunto, aceleraria o trabalho da catalogação uma vez que o conteúdo técnico do documento seria mais facilmente compreendido. (Classification Research Group Bulletin, 1962, p. 86)

Foskett (1962) relata que os membros do CRG em suas discussões chegaram à conclusão de que a indexação por assunto, fosse alfabética, facetada ou mecânica, dependeria de sua eficiência na coordenação de conceitos ou ideias expressos nos documentos, de modo que um conjunto de conceitos declarados como o requerimento de um pesquisador/usuário (*searcher*) possa ser combinado com o mesmo ou similar conjunto já existente na biblioteca ou sistema de informação. Logo, "o bibliotecário [indexador] tem que fazer uma tradução ou transformação do conjunto de termos em que a solicitação é declarada em um conjunto principal de esteriótipos de significados em que os documentos estão organizados" (FOSKETT, 1962, p. 128)

O próprio Ranganathan teorizou acerca do uso das facetas no processo de indexação. Para ele, a escolha do nome do assunto de um documento e a representação do nome no título da entrada de um assunto específico pode ser obtido por análise de faceta com base em postulados e princípios da *Colon Classification*.

De acordo com Ranganathan (1964), a análise das facetas é um trabalho feito no plano das ideias e no plano verbal. Ele critica ainda a concepção que trata a análise das facetas como um método exclusivamente concebido para a classificação. Depois de demonstrar que os cabeçalhos de assunto constituem uma linguagem artificial, aponta que o uso da análise de faceta para o cabeçalho de assunto não equivale a usar o número de notação, mas aplicar a própria teoria facetada. Neste trabalho, ele estabelece três usos possíveis para as facetas: a) determinação do número da classe; b) Serviço de Referência; e c) determinação do título do assunto.

O êxito da *Colon Classification* no processo de classificação dos documentos permitiu a sua adaptação em outras operações como o serviço de referência - a exigência de um leitor, particularmente de um leitor especialista. No momento em que é formulada, de forma pontual, sua busca, ela pode ser atendida pela análise de faceta - isto é, averiguando, através da interação entre o bibliotecário e o usuário, as facetas de exigência do requerente, o foco exato no assunto básico e cada uma de suas facetas isoladas, respectivamente. A determinação do objeto de um documento e seu título de assunto também tem sido uma operação em que a aplicação das facetas tem se revelado como método bem sucedido. Ela pode ser feita de uma forma sistemática e padronizada a partir da análise das facetas e atribuição dos títulos a partir de suas combinações (RANGANATHAN, 1964, p. 111).

No entanto, Ranganathan alerta que, mesmo seguindo a tais critérios para as cinco facetas selecionadas, há 120 arranjos possíveis para representar um assunto específico. Diante dessa perspectiva, ele apresenta princípios que podem nortear, estabelecer diretrizes para a escolha dos assuntos de um documento. O princípio da predominância, que estabelece ser a predominância do termo que decide a posição de cada entrada no índice; o princípio do sentido procurado, estabelece que o primeiro termo a constar no vocabulário de assunto deve ser o que dá razão, sentido, à

existência do documento; o princípio da estatística de pesquisas ou opiniões, esse princípio aponta para a conjectura e a experiência dos sistemas de informação, a definição da posição dos termos pode ser baseada em um levantamento objetivo estatístico dos termos utilizados para a busca do documento ou nas opiniões dos usuários.

Segundo Ranganathan (1964), só havia duas escolas de pensamento dominantes - uma escolhendo os termos do cabeçalho de assunto de forma perspectiva (*forward rendering*) e a outra escolhendo os termos de forma retrospectiva (*reverse rendering*). De acordo com Ranganathan, a disputa entre estas duas escolas de pensamento deveria ser decidida por um estudo estatístico ou por uma grande quantidade de modelagens. Caso contrário, prosseguir a discussão sobre esta questão seria apenas alimentar a polêmica sem a objetividade de resolver a questão. Esta polêmica poderia não somente ficar restrita às duas escolas dominantes, mas variar entre cinco e 120 escolas de pensamento por conta das possibilidades de combinação. Desse modo, tais escolhas deveriam estar baseadas na experiência e nos dados estatísticos dos sistemas (RANGANATHAN, 1964, p. 116).

Uma vez estabelecida a escolha e a representação do assunto específico de um documento para um cabeçalho com base no conjunto de postulados da análise das facetas e dos princípios para a sequência facetada, o número mínimo de referências cruzadas, a exemplo do “ver também”, nos cabeçalhos de assunto seria igualmente e mecanicamente derivado do cabeçalho de assunto com o auxílio do procedimento em cadeia. Nessa teoria, as decisões acerca da escolha das facetas que irão representar o assunto deverão ser feitas primeiramente no plano de ideias. Em seguida, as regras do procedimento em cadeia podem ser enquadradas para implementar no plano verbal, implementar a política adotada no plano das ideias (RANGANATHAN, 1964, p. 118).

As entradas de cabeçalho de assunto, em conjunto, formariam uma linguagem artificial. As palavras usadas seriam retiradas da linguagem natural, mas a sintaxe de posição seria diferente da sintaxe da linguagem natural a partir da qual as palavras são tomadas. É sabido que a sintaxe varia de uma linguagem natural para outra. Para Ranganathan (1964), os princípios para a análise das facetas e de sua ordenação ajudariam em grande medida a atribuição de assuntos específicos do documento por fornecer uma espécie de sintaxe universal, transformando-os em cabeçalho de assunto, e, inclusive, estabelecendo seus números de classe.

As tarefas de catalogação e de classificação são igualmente beneficiárias desses postulados e princípios. A utilização de um único e mesmo procedimento na catalogação e na classificação não garante que a presunção do cabeçalho de assunto seja derivada do número de classe ou do número de classe sendo derivado do título de assunto, mas facilita uma padronização na representação do documento. (RANGANATHAN, 1964, p. 119, tradução nossa)

Ranganathan associa o surgimento do que ele chama de Escola de Assunto Específico ao trabalho de Cutter, de 1891, aproximando novamente a catalogação da indexação. Ranganathan pondera que no final do século XIX, Cutter estava preocupado apenas com a organização dos livros e não com a organização dos documentos multifacetados e fragmentados, característicos da contemporaneidade da década de 1960. Mesmo assim, a atribuição do assunto de um livro era nebulosa e, por conseguinte, a sua representação.

Para Ranganathan, nenhum procedimento definitivo era possível pela falta de maturidade teórica do campo. No entanto, alegava que na década de 1960, a posição era melhor por conta do desenvolvimento teórico e pela experiência com os sistemas de informação. Novamente, enfatiza a necessidade do levantamento estatístico da abordagem do leitor ao primeiro assunto, bem como a sequência dos outros assuntos em uma declaração de assunto múltipla. Argumenta que apesar de não ser uma pesquisa fácil, era o único trabalho a ser feito, uma vez que os princípios já eram conhecidos (RANGANATHAN, 1964, p. 119).

A atuação do *Classification Research Group* e de seus membros no desenvolvimento de teorias acerca da organização e representação do conhecimento impactou não somente a teoria da classificação, mas também o processo de indexação que, na vertente inglesa de tratamento temático da informação, possui um grande vínculo. De acordo com o CRG, a organização da documentação visando a sua recuperação passa por importantes etapas como: primeiro, a definição do assunto do documento e a terminologia a ser adotada para classificar todo universo documental, depois o levantamento das facetas a serem adotadas.

Ranganathan ao elaborar os princípios da *Colon Classification* estabeleceu cinco facetas, chamadas de *PMEST* - personalidade, matéria, energia, espaço, tempo, respectivamente. De acordo com Spiteri (1998, p. 19), O CRG ao longo de seus debates e troca de experiências percebeu que era necessário ampliar a quantidade de facetas para dez e com a seguinte ordem: tipo de produto final, partes, materiais, propriedade, processo, operação, agente, espaço, tempo e forma de apresentação. É importante destacar que o CRG, ao contrário dos postulados de Ranganathan, chegou a conclusão de que a escolha das categorias fundamentais e a ordem de citação deveriam derivar da natureza dos assuntos encontrados no acervo a ser representado.

A natureza do assunto seria revelada pelo próprio contexto em que o assunto estava inserido. Assim sendo, não deveria haver qualquer imposição de listas de categorias fundamentais apriorísticas, ainda que fossem exaustivas. Essa imposição invalidaria o benefício das possibilidades de múltiplas combinações na representação dos assuntos que o uso das facetas permite.

Para Foskett (1962, p. 137), o CRG, a partir dessa perspectiva, procurou contemplar uma espécie de *UR-classification*, uma espécie de sistema classificatório generalista que poderia ser

adaptado por diversos sistemas classificatórios por meio dos níveis integrativos, ou seja, um conjunto primordial de termos em facetas, sub-arranjados de acordo com a teoria dos níveis integrativos, mas não indo mais fundo na síntese de complexos termos do que a própria teoria. A partir de tal esquema ideal, que provavelmente ninguém poderia realmente usar, poderia traçar-se listas de termos, análises de termos complexos que precisariam ser usados como conjuntos em certos campos, relações básicas e atividades.

Vickery (1955), um dos autores de destaque no *CRG*, preocupou-se com as mudanças da linguagem e o surgimento de novos conhecimentos, alertando que um esquema classificatório jamais estaria inteiramente pronto. Dessa forma, procurou elaborar possibilidades de futuras expansões nos esquemas classificatórios, considerando sub-facetas a partir da aplicação de novos princípios de divisão como a teoria dos níveis integrativos já abordada nesse trabalho.

Ao analisar a trajetória da prática de indexação entre 1945 e 1955, em seu artigo *Developments in subject indexing* (1955), Vickery afirma que houve uma enxurrada de artigos, livros, palestras, simpósios e discussões sobre indexação, classificação, recuperação mecânica, análise de assuntos, recuperação da informação nesses dez anos. De acordo com o autor, essas publicações demonstram que uma geração anterior usou palavras e cabeçalhos de assunto, já os estudos da década de 1950 passou a usar termos isolados, facetas, analetos, descritores e unitermos.

Para Vickery, a discussão de diferentes técnicas de indexação estava sendo quase sempre conduzida de forma polêmica, cada novo sistema que surgia era acompanhado por críticas mordazes de seus antecessores e rivais. As características que distinguem um sistema particular de outros eram enfatizadas, e os princípios comuns a todos os sistemas obscurecidos. Isto prejudicava o desenvolvimento das técnicas de indexação.

A indexação alfabética, a classificação e a recuperação mecânica estavam sendo abordadas como processos completamente diferentes e descritos em termos tão diferentes que os profissionais de um campo eram às vezes ininteligíveis para aqueles em outros. Vickery procurou aproximar tais tendências, revelando que todas essas formas de classificação, seleção mecânica e indexação de assunto estavam enfrentando problemas semelhantes e estavam começando a resolvê-los de maneiras semelhantes.

O primeiro problema citado por Vickery foi a eliminação de títulos sinônimos ou sobrepostos. O fato dos cabeçalhos estarem interligados leva o indexador a fazer um número considerável de entradas “ver também” (VICKERY, 1955, p. 1). Conforme a lista cresce, a necessidade de alguma ação para assegurar consistência e uniformidade torna-se óbvia. Influenciado pelos estudos de Pettee (1947), Vickery passou a defender uma análise lógica do

assunto para a indexação dos documentos. Essa análise seria feita a partir de uma árvore conceitual classificatória que mostraria as interconexões de um campo particular do conhecimento².

Destarte, Vickery considerava oportuno na construção de índices alfabéticos de assuntos - simplesmente para evitar cabeçalhos mal definidos e sobrepostos - construir uma série de mapas classificatórios, cobrindo diferentes partes do conteúdo dos documentos. Um dos problemas identificados por Pettee e enfatizado por Vickery foi que, geralmente, os indexadores começavam o cabeçalho de assuntos com "o carrinho alfabético diante do cavalo, mas logo se descobre que, embora o carro deva carregar a carga, é o cavalo que deve puxá-lo" (PETTEE apud VICKERY, 1955, p.2). Este deve ser o princípio básico para o estabelecimento de um cabeçalho de assunto: sua ordenação deve estar baseada no objeto principal do documento, este seria o "cavalo" para Pettee que achava necessário construir uma série de mapas classificatórios para os vários campos abrangidos pelo índice.

Vickery retomou Cutter mostrando que já no século XIX havia uma pirâmide de referências, em *Rules for a dictionary catalog* para representar os assuntos subordinados àqueles estabelecidos como gerais. Uma primeira questão a ser operacionalizada era: como o indexador poderia saber, à medida que entrasse um novo termo no índice, quais eram os termos genéricos e os coordenados? A partir de quais termos ele deveria fazer referência ao recém-chegado?

Para Vickery (1955, p.2), existia apenas uma maneira certa de indexar o novo termo: relacionando-o com os seus genéricos e coordenados: elaborando uma tabela sinóptica, como Cutter a chamava, mostrando os termos e suas interconexões em um mapa classificatório. Desse modo, o novo termo poderia ser ajustado e as referências cruzadas apropriadas feitas.

Ao aludir Cutter, Vickery lembra que esse imenso esforço não é recomendado pelo bibliotecário norte-americano para um único catálogo. No entanto, "o que parece ser demais para um catálogo pode ser lucrativo para todos os catálogos" (CUTTER apud VICKERY, 1955, p.3), na forma de um sistema universal de classificação, complementa Vickery.

Cutter sugeriu, de fato, que os termos para um índice de assunto, e suas interconexões, fossem tomados das tabelas de classificação. O próprio Cutter considerou certos exemplos de títulos compostos, e ofereceu algumas sugestões quanto à ordem de citação. Ele aconselhou: "descreva um nome de assunto composto pela sua primeira palavra, invertendo a frase somente quando alguma outra palavra é decididamente mais significativa. Nessa perspectiva um trabalho sobre genoma humano dificilmente seria procurado pelo termo 'humano'. Por outro lado, o termo mais significativo em revolução bolchevique deveria vir à frente, ou seja, 'bolchevique, revolução' ”³ (CUTTER apud VICKERY, 1955, p.3). A regra de Cutter deixaria o indexador (e o usuário) julgar,

²Ver anexo A

³ Termos substituídos para dar o mesmo sentido à ideia de Cutter na língua portuguesa, originalmente Cutter utilizou os termos *digestive organs* e *ancient history*.

em cada caso, qual é a palavra "mais significativa". Ele acrescenta: "Deve-se confessar que esta regra é um tanto vaga" (CUTTER apud VICKERY, 1955, p. 3).

Vickery aprofundou as relações entre a teoria da classificação e a indexação, sobretudo na década de 1960, como demonstram Cordeiro e La Barre (2011, p. 182). Com o intuito de possibilitar as possibilidades de organização e representação do conhecimento, aprimorando a recuperação da informação e do documento, conseqüentemente, Vickery passou a teorizar sobre a normalização da linguagem dos documentos (conteúdos registrados), considerando também a normalização das perguntas (*queries*) dos usuários dos Sistemas de Recuperação da Informação.

A proposta de Vickery procurava acertar a combinação entre as representações dos documentos, realizadas na entrada dos sistemas pela técnica da indexação, e as questões formuladas pelo usuário na saída dos SRIs. A linguagem da indexação teria a função de representar o conteúdo do documento e, nessa perspectiva, os esquemas classificatórios forneceriam parâmetros para a definição dos conceitos e suas interrelações, facilitando a escolha dos termos na fase de análise conceitual (CORDEIRO; LA BARRE, 2011, p. 182).

A indexação alfabética e a seleção mecânica são técnicas de indexação largamente difundidas na década de 1950 nos SRIs que começavam, por assim dizer, de baixo para cima (*bottom-up*), a partir dos termos individuais. Sua abordagem seria chamada de indutiva. Ambas as técnicas haviam sido conduzidas, no decurso do seu desenvolvimento, a introduzir métodos de agrupamento de termos individuais.

O efeito destas operações figurava, em princípio, na construção de uma série de mapas classificatórios, nos quais os termos eram agrupados primeiramente em várias categorias, e dentro de cada categoria são arranjados nas árvores de família conceituais. A partir desses mapas, os termos de indexação e suas interconexões são derivadas. Segundo Vickery (1955, p. 7), esse marco na transformação da indexação que passou a ser baseada na adoção de mapas conceituais já eram nítido nos sistemas europeus e norte-americanos onde essa forma de aplicar a técnica já era realizada na indexação alfabética e na seleção mecânica dos termos.

Farradane, outro importante nome no CRG, propôs que as relações entre os diferentes termos do cabeçalho pudessem ser expressas por meio de um pequeno número de operadores-padrão, definidos uma única vez. Essas relações operacionais poderiam ser definidas de modo que, na maioria dos casos, só existisse uma possível ordem de citação que correspondesse ao que Farradane chamou de ordem totalmente dedutiva (FARRADANE, 1952, p.74).

As conclusões acerca do uso das facetas tanto para classificar os documentos, quanto para construir notações, debates recorrentes ao CRG durante os anos de 1950, reverberaram também nos debates acerca da indexação dos documentos. Farradane, por exemplo, passou a teorizar acerca do sistema conceito-coordenação na construção de índices. O autor defendia que o assunto de um

documento fosse analisado em seus termos constituintes. Sobre os estudos de Farradane, Foskett resumiu da seguinte forma:

Originalmente, sua técnica se aplicava apenas à elaboração de índices alfabéticos, mas é um sistema conceito-coordenação, sobretudo. O assunto de um documento é analisado em seus termos constitutivos ou isolados, que são então reunidos por meio de uma série de relações de símbolos ou operadores. Estes operadores são aplicados em uma sequência fixa, e a série de termos isolados e seus operadores é chamada de 'analetos' (FOSKETT, 1962, p. 132).

Ao aproximar a classificação da indexação, Farradane justifica tal feito argumentando que a classificação representa uma teoria da estrutura do conhecimento, isto é, das relações entre as diferentes áreas do conhecimento. O problema básico é determinar quais são essas relações e como elas ligam os diferentes conceitos do conhecimento em uma estrutura coerente (FARRADANE, 1952, p.74).

A subdivisão arbitrária ou dedutiva de um conjunto de conhecimentos, segundo o autor, não poderia dar uma representação verdadeira dessas relações, pois estes não consistiriam apenas em agrupamentos de uma classe e seus membros, ou na divisão de um todo em suas partes. Farradane (1952) argumenta que uma classificação deve ser construída de forma indutiva, ou de baixo para cima (*bottom-up*), reunindo fragmentos conhecidos de relações. Partindo de elementos de conhecimento singularmente definíveis, denominados "isolados", estabelecendo as relações entre eles, chamadas de "operadores", sendo representados linearmente por símbolos denominados analetos (*analets*) (FARRADANE, 1952, p.74). Estes analetos poderiam ser rearranjados de acordo com uma técnica simples para proporcionar várias permutações desejáveis e permissíveis que, dispostas em ordem alfabética dos primeiros isolados, produziriam um índice de assunto.

Selecionando as combinações que estão em ordem totalmente dedutiva e selecionando-as para agrupar os isolados, será possível, contudo, construir uma classificação do material disponível. Com uma amostra suficientemente representativa de todos os conhecimentos, será obtida uma classificação universal. Argumentou-se que este é o único método possível pelo qual uma verdadeira classificação, ou representação totalmente lógica da estrutura do conhecimento, pode ser obtida. Em quase todas as classificações tem sido indubitavelmente o objetivo de seus construtores apresentarem um sistema lógico, agrupar os termos relacionados nas classes corretas, mostrar a subdivisão de um todo em suas partes e trazer nosso conhecimento disperso em um único sistema ordenado. A classificação é, de fato, uma atividade humana fundamental. É o método pelo qual, em nossos primeiros contatos com o mundo, conseguimos um senso de familiaridade e segurança em nosso ambiente (FARRADANE, 1952, p. 74, tradução nossa).

Segundo Farradane (1952), quando podemos alocar um novo objeto ou uma nova experiência em uma classe de conceitos previamente conhecidos, ou quando podemos encontrar um fator comum em um conjunto de fenômenos elencados previamente, removemos um elemento de estranheza, uma possível fonte de medo do incompreensível, do nosso ambiente. "É uma maneira de

alcançar a harmonia em nosso processo contínuo de adaptação a um mundo de outra forma hostil e incalculável" (FARRADANE, 1952, p. 74).

Nesse sentido, os operadores, ou seja, sistema de relações, representam relações lógicas. Os significados dos operadores, que representam as combinações básicas de percepções mentais (das quais, no estado atual da ciência, não podemos fazer mais análises), não são fáceis de expressar em palavras simples. Os nomes que foram dados aos operadores são as palavras da linguagem cotidiana que parecem descrever as situações relacionais mais comuns que, a partir de um exame de uma grande variedade de assuntos típicos, parecem se adequar às exigências lógicas. Esses nomes não devem, no entanto, ser tomados demasiado literalmente ou muito singularmente (FARRADANE, 1952, p. 83)

Para Farradane, nossa linguagem comum é um meio muito pobre para a expressão dos novos conceitos apresentados pela ciência. Por esta razão os cientistas necessitam em grande parte criar novas palavras para novos conceitos. No campo da classificação, no entanto, poderia ter sido demasiado confuso elaborar muitas palavras novas para descrever os operadores, e os nomes usados devem ser considerados como sendo apenas típicos dos significados possíveis de cada operador, tendo em conta a lógica do significado de cada operador,

No âmbito do CRG, os debates acerca do tratamento da informação e dos documentos visando a melhor eficiência dos SRIs passaram a incorporar as discussões acerca das possibilidades de interconexões entre diferentes centros de documentação e informação. O quarto boletim (1958) já apresentava a preocupação de seus membros com a construção de sistemas interconectados. Para os membros do CRG, a base de um sistema de recuperação de informação é uma rede de termos interligados.

As interconexões de cada termo são potencialmente ilimitadas, e cada sistema de recuperação deve selecionar certas interconexões para acesso aos documentos. A padronização e seleção dessas interconexões deveriam basear-se em princípios quanto aos tipos de termo e às suas relações - princípios relativos ao nível semântico dos termos, das categorias a serem utilizadas, das relações genéricas, coordenadas e conjuntivas a serem representadas e dos tipos de operação de busca a serem conduzidos.

O CRG preocupava-se em explorar a variedade de princípios que poderiam ser estabelecidos para a construção de sistemas classificatórios e a ordenação na construção de vocabulários controlados, além de avaliar a sua adequação em várias situações ocorridas nos sistemas de recuperação da informação.

O fator semântico, questão que recebeu grande atenção por parte do CRG, tornou-se um problema a ser resolvido uma vez que o grupo procurava formas de classificação universais, ou seja, que fossem aplicações possíveis a qualquer SRI. O grupo chegou à conclusão de que seria a

análise do significado de um termo complexo em um número de termos mais generalizados que denotariam a própria questão abordada.

O valor da análise semântica para fins de recuperação se trataria então da capacidade de exibir as relações existentes entre os conceitos e suas propriedades em comum na representação dos documentos. Uma cadeia de classes genéricas escolhida a partir de um conceito específico representa uma série de propriedades do conceito e, dessa maneira, uma análise semântica. Os termos da cadeia, no entanto, não produzem uma análise semântica completa. Eles não incluem todas as propriedades pertencentes ao conceito, mas apenas as propriedades que são necessárias para distingui-lo de outros conceitos. Na lógica formal, uma propriedade usada na estrutura de uma classificação é conhecida como diferença, enquanto uma propriedade não tão usada é chamada de acidente.

A proposta foi a adoção da teoria dos níveis integrativos, de J. K. Feibleman e Joseph Needham (FEIBLEMANN, 1954). Esta teoria observou a existência de “formas sucessivas de ordem numa escala de complexidade e organização” (Classification Research Group Bulletin, 1961, p. 160) e de uma sucessão de níveis em que as agregações se tornam inteiras agindo como unidades e envolvidas em processos característicos de níveis particulares. Isto ofereceria uma repartição preliminar viável da lista de coisas físicas e materiais, embora ainda não estivesse claro como as abstrações, ou as formas de energia, seriam montadas na série de níveis.

Nesta perspectiva, o conceito de nível integrador parece oferecer soluções razoáveis para a questão dos assuntos compostos baseados em termos isolados extraídos de classificações existentes em várias classes principais. Este era um dilema frequente, por exemplo, na classificação de uma bibliografia geral ou de uma biblioteca, e nenhuma solução era oferecida pelo tipo de faceta dentro de um sistema de classe usado na *Colon Classification*.

Embora a inserção de classes canônicas deva ser considerada como um dos problemas menos bem resolvidos da estrutura integrativa de níveis, o CRG acreditava que esta teoria oferecia naquele momento soluções mais satisfatórias do que as classificações gerais existentes - "consegue representar um documento no que se refere a todos os seus aspectos de uma 'coisa' particular (Classification Research Group Bulletin, 1961, p. 159).

No sétimo boletim (1962), os relatos sobre a utilização da indexação por cadeia de facetas tal como foi sugerido por Ranganathan mostrou-se bem sucedida para os centros de documentação que contavam com publicações altamente especializadas como nas áreas da saúde e da segurança, analisadas neste encontro pelo grupo.

A prática era preparar as entradas de cadeia necessárias e referências cruzadas para novos códigos ao mesmo tempo em que o sistema de classificação estava sendo elaborado (Classification

Research Group Bulletin, 1961, p. 67). O acervo foi abordado através da criação de analetos, ou representações conexas relacionadas linearmente, para um grande número de assuntos.

Os analetos poderiam ser omitidos uma vez que a ordem das facetas tomariam seu lugar. Em alguns casos especiais, os assuntos poderiam ter uma ordem lógica diferente, e algumas indicações da quebra da ordem poderiam ser necessárias (por exemplo, inserção de um termo como “cólón” ou “acidente vascular cerebral”); tais assuntos não são frequentemente relevantes para o campo de interesse (Classification Research Group Bulletin, 1961, p. 78). Uma vez definidas as facetas, todos os termos que surgissem no campo poderiam facilmente ser colocados em suas classes corretas e dados de notação apropriados.

O CRG preocupou-se com a dispersão das facetas numa indexação em cadeia, pois a ordem de combinação das categorias geralmente dava prioridade aos produtos e não a seus materiais. Para superar esta dificuldade com a ordem preferida, na tentativa de reduzir a dispersão, as regras para a construção das cadeias na indexação deveriam ser estabelecidas em certas partes das tabelas que limitavam estritamente o uso das subdivisões previstas. Deveria haver uma preocupação em limitar o estabelecimento das entradas de assunto por cadeia, uma vez que um amplo índice em cadeia poderia levar a dispersão dos assuntos e, por conseguinte, a impossibilidade da recuperação do documento.

O grupo percebeu que diversos programas haviam abandonado a indexação em cadeia, pois as dificuldades encontradas quando as buscas eram feitas por determinadas combinações de termos causavam uma fadiga excessiva pela quantidade de documentos recuperados. Embora esta crítica da indexação em cadeia tenha sido aceita como comprovada em diversos relatos de seus membros como Cranfield e Cleverdon, isto não fez com que esta fosse inviabilizada completamente (Classification Research Group Bulletin, 1964, p. 147).

Dois pontos emergiram desta discussão: o primeiro foi que a linguagem descritiva mais detalhada trazia maiores chances de erro tanto para o indexador, quanto para o usuário; e o segundo foi a conclusão de que o conceito relevância, da forma em que estava sendo aplicado nos sistemas de indexação, era bastante amplo na atribuição dos assuntos dos documentos.

Um cabeçalho de assunto poderia corresponder a um termo de pesquisa, mas a referência recuperada ainda poderia ser não pertinente se, por exemplo, o documento estivesse em um nível inadequado para o leitor ou fosse, por algum motivo, uma apresentação incompreensível para ele. Assim, recomendava-se que a indexação deveria primar pela especificidade e dessa forma ajudar o pesquisador a tornar seus requisitos de busca mais claros para si mesmo.

Langridge na mesma reunião do grupo levantou a questão do uso dos termos “genérico” e “hierárquico”. Verificou que estes termos criavam dificuldades e sugeriu a utilização do termo “relação parte-todo” com o argumento de que parte-todo e espécie-gênero eram ambas relações de

inclusão e podiam, por conseguinte, ser descritas pelo mesmo termo (Classification Research Group Bulletin, 1964, p. 148). No entanto, as relações não poderiam ser resumidas a apenas este aspecto. O Grupo procurou realizar uma análise para descobrir até que ponto o uso desses termos poderia alcançar certa consistência entre eles para: (1) para atribuir títulos de assuntos acordados a partir da literatura da área e (2) para analisar as relações entre os conceitos. Os relatos mostraram que desacordos surgiam na fase de análise conceitual, mas eram resolvidos no processo de tradução para a linguagem do sistema, uma vez que os conceitos haviam sido acordados (Classification Research Group Bulletin, 1964, p. 149).

2.3 OS SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO COMPUTADORIZADOS E A TEORIA MATEMÁTICA DA COMUNICAÇÃO NO PROCESSO DA INDEXAÇÃO

A incorporação dos computadores na Segunda Guerra Mundial nos meios de comunicação impactou também as formas de organização e tratamento documental. O surgimento dos primeiros centros de documentação e dos primeiros sistemas de informação sofreram a influência da filosofia predominante na área, isto é, pensar a padronização de códigos, de dados, de informações para o funcionamento adequado, eficaz e eficiente dos sistemas, no sentido, em que as máquinas pudessem compreender e executar os comandos dados e que o fluxo de informações ocorresse sem interrupções.

Essa concepção é marcada pela influência da chamada teoria matemática da comunicação, de Shannon e Weaver, em 1948, na qual a comunicação se dá por meio de sistemas fechados em que há um emissor, a mensagem, um receptor e um destinatário. O emissor produz o sinal a ser transmitido através do canal. O receptor decodifica a mensagem e a reconstrói a partir do sinal. Já o destinatário trata-se daquele a quem a mensagem é direcionada.

A comunicação é vista como algo que pode ser matematicamente entendida e calculada. É uma teoria que se concentra na medição da informação, ou seja, trata do estudo quantitativo da informação em mensagens e do fluxo de informação entre emissores e receptores. Essas mensagens, sequências de valores transmitidos e que podem ser calculados probabilisticamente, trafegam por canais previstos e que devem ser padronizados a fim de diminuir a perda de valores, dados, transmitidos. Os ruídos, excesso de dados nos canais de comunicação, ou os silêncios, falta de dados suficiente para a compreensão de uma mensagem, poderiam causar falhas, levando a erros na decodificação das mensagens (FREIXO, 2006)⁴.

No campo do tratamento temático da informação, as mensagens deveriam garantir o valor semântico da informação (BORKO, 1977), este seria extraído dos documentos por critérios confiáveis de determinação dos assuntos. Dessa maneira, as representações dos documentos

⁴ O modelo está representado no anexo B deste trabalho.

poderiam ser padronizadas e incluídas nos sistemas de recuperação da informação, atendendo a demanda dos usuários e garantindo a recuperação dos mesmos sem erros. A eficácia dos SRIs seria medida pelos *feedbacks* de seus usuários, ou seja, a influência da cibernética que compreende a retroação (respostas dos destinatários) como elemento regulador da circularidade da informação.

A *Terminological Continuum*, de Frederick Jonker (1957), é uma teoria da indexação em consonância com esses pressupostos. A partir de operações matemáticas, Jonker procurou controlar as significações das palavras, tentando criar um vocabulário em que cada palavra tivesse apenas um significado, evitando assim os equívocos nas mensagens. As 26 letras do alfabeto poderiam gerar múltiplas palavras e uma infinidade de combinações, a *terminological continuum* procurava estabelecer leis para controlar essas variações na indexação, expressando a relação entre o tamanho do vocabulário e a especificidade com que um conceito poderia ser descrito.

Nesta perspectiva, a indexação deveria reconhecer os termos mais significativos de um documento para auxiliar a recuperação da informação. Jonker enfatizou ainda que as linguagens técnicas deveriam esforçar-se para diminuir o problema dos diversos significados que uma palavra pode ter, criando um vocabulário em que cada palavra tivesse apenas um significado. Tal modelo já expressava a dificuldade de prever e descrever o alcance da recuperação da informação para o usuário.

O processamento da indexação se daria por meio da representação de todo o documento a partir de termos únicos (*uniterms*) pós-coordenados. Um axioma da teoria de Jonker é o *connective continuum* que mostra as relações hierárquicas e terminológicas em um sistema de recuperação da informação. O *connective continuum* estabelece uma relação funcional entre a especificidade de um conceito a ser descrito e o número de palavras na indexação dos termos.

A teoria da indexação de Jonker permite compreender e mapear a disposição das relações entre o tamanho do vocabulário, a organização hierárquica da terminologia e a especificidade com que os conceitos podem ser descritos.⁵

A incorporação dos computadores pelos SRIs levou ao desenvolvimento das teorias de indexação automática. Alguns trabalhos começaram a aparecer na década de 1960 como o de Cleverdon (1961) que procurou pensar a mecanização do processo de indexação. O autor procurava transformar o processo intelectual da indexação, ou seja, "a operação que consiste em uma pessoa ler o documento e avaliar o conteúdo da matéria, depois decidir quais são os assuntos relevantes para a finalidade do índice e, finalmente, decidir quais seriam incorporados nos sistemas de informação" (CLEVERDON, 1961, p. 107, tradução nossa), em uma operação sistemática e rotineira. Isso seria possível a partir do mapeamento da frequência do uso de cada palavra e do

⁵ O modelo *terminological continuum* está representado no anexo c.

controle terminológico que estipulariam os termos a serem indexados. Dessa maneira, o computador poderia realizar toda a operação da indexação.

Segundo Cleverdon (1961), neste esquema, poderia ser previamente selecionada uma lista de palavras-chave e a instrução dada ao computador seria que as dez palavras-chave que ocorressem com mais frequência fossem usadas como termos de indexação automaticamente. O autor argumenta que para um humano fazer isso seria uma tarefa hercúlea, no entanto, esta operação poderia ser facilmente realizada por uma máquina, como um computador de alta velocidade (CLEVERDON, 1961, p. 107).

Um dos exemplos mais conhecidos deste modelo é a indexação realizada pela *Chemical Abstracts* neste período, no qual os assuntos passaram a ser atribuídos pelo sistema *Key-Words-in-Context*, comumente conhecido como *KWIC*. Neste caso, o computador realizava o processo descrito anteriormente de encontrar todas as palavras-chave nos títulos, ele então classificava todo acervo documental em um arranjo alfabético a partir da rotação das palavras constantes de título de determinado documento.

O grande mérito deste método particular, tal como era utilizado nas entradas dos documentos da Química, é que permite que informações sobre novos artigos sejam disponibilizadas muito mais rápido do que se houvesse os inevitáveis atrasos da elaboração de resumos e de índices. Cleverdon (1961) reconhecia a limitação do processo de indexação automática como as sinonímias, a generalização de documentos que abordavam assuntos de grande especificidade e a dificuldade de estabelecer correlações entre os documentos. No entanto, a rapidez na realização da indexação de grandes volumes documentais que tal processo permitia, segundo ele, precisava ser considerada. Cabe mencionar, que tal recurso foi bastante utilizado nos índices alfabéticos de assuntos das bibliografias brasileiras.

Outra proposta para indexação dos documentos influenciada pela teoria matemática da comunicação é a de Blagden (1966) que procura eliminar os ruídos nos sistemas de indexação que trabalham com a atribuição de termos coordenados. Em seu estudo, mostra o nível alto de ruído em sistemas que usavam os *uniterms* como descritores, ou seja, na indexação o ruído se expressa na porcentagem de documentos irrelevantes sobre a porcentagem do número total de documentos recuperados.

Os testes apresentados por Blagden(1966) consideravam as perguntas (*queries*) feitas pelos usuários no processo de busca nos SRIs e a capacidade de revocação (*recall*) dos documentos. O ruído tratava-se dos documentos irrelevantes para o usuário relativos a sua busca e o silêncio dos documentos não recuperados pelo sistema que faziam parte da base de dados e seriam relevantes para o usuário. O autor propôs que a indexação por unitermo (*uniterm*) fosse reformulada, procurando eliminar os sinônimos a partir de vínculos criados por relações hierárquicas e

coordenadas, desenvolvendo um conjunto complexo de onze indicadores de função, estabelecendo um vocabulário com relações superordenadas, coordenadas e subordinadas entre os todos os termos usados na indexação. Segundo o autor, esse sistema foi utilizado com êxito pelo Engineers Joint Council na década de 1960.

Os anos 1960 foram marcados pela avaliação da eficácia dos sistemas de recuperação da informação e da indexação automática dos documentos. Essas análises, influenciadas em grande medida pela teoria matemática da comunicação, procuravam aferir o grau de revocação e precisão na recuperação dos documentos.

Lynch (1966) realizou um importante estudo avaliativo na *Chemical Abstracts (CA)*, um dos sistemas altamente desenvolvido do período e referencial para vários SRIs. De acordo com o autor, a linguagem de indexação de assuntos do CA altamente estruturada havia se convertido em uma poderosa forma de expressão, pois possibilitava a representação de assuntos complexos e das relações entre eles.

O vocabulário da linguagem de indexação representava um admirável compromisso entre o controle necessário para limitar o número de documentos recuperados e da variedade de assuntos que permitiam o acesso do usuário a uma quantidade razoável de documentos pertinentes a sua busca. "A indexação realizada pelo CA é de uma qualidade e consistência comprovada em muitos anos de uso" (LYNCH, 1966, p. 168, tradução nossa).

O primeiro objetivo deste estudo foi descobrir a lógica implícita na relação entre a entrada de assuntos individuais, tal como aparece no índice publicado pela *Chemical Abstracts*, e a forma normal, ou título, da qual pode ser considerada derivada. De acordo com Lynch (1966, p. 169), um alto grau de organização na indexação é alcançado pelo uso de vários níveis de sintetização, não apenas de modificações completas sob um título de assunto comum, mas também de partes de modificações que têm palavras iniciais comuns ou sub-frases.

O interesse deste estudo centrou-se nestas entradas com modificações que incluíssem informação química não estrutural, ou seja, não serviam apenas para completar a descrição de um composto químico, mas para descrever as relações entres os compostos. O grau de especificidade requerida pelos usuários da CA fez com que os métodos de indexação das publicações na área da Química fossem aprimorados, a lógica adotada foi de apresentar as relações como uma árvore binária, em cada nível sucessivo do qual um componente adicional era incluído ao agregado já formado.

Assim, era realizada a equivalência entre a ordem dos componentes da entrada na forma normal e as posições relativas da função. No entanto, Lynch (1966) afirmou que em alguns casos não era possível resolver certas ambiguidades, sobretudo em relação às entradas não regulares, de um ou mais dos componentes químicos que não possuísem uma indicação inequívoca da posição

que deveria assumir no sistema e, conseqüentemente, se deveria realizar uma análise mais profunda para se obter uma associação correta. Em muitos destes casos, a referência a outras entradas relativas ao mesmo documento poderia ajudar na resolução.

Essa forma de indexar poderia ser automatizada a partir dos padrões de entrada, de funções e de relações estabelecidos pelos indexadores humanos, aliviando-os do trabalho de registrar as entradas individuais em uma variedade de formas sob diferentes títulos de assunto. Destarte, pelo menos parte do trabalho de indexação poderia também ser assumida pelas máquinas.

Este modelo simples abriu caminho para a indexação automática de assuntos em centros de documentação, com procedimentos consolidados, não de forma intuitiva como a maioria dos SRIs estava operando a indexação dos assuntos de seus documentos, mas por padrões de entrada estabelecidos de forma apriorística. Então, o computador escolheria dentre as alternativas possíveis cada entrada e, por conseguinte, as posições secundárias e subseqüentes de acordo com os indicadores (LYNCH, 1966, p. 184).

No entanto, os títulos dos artigos são muito variáveis na estrutura e no vocabulário para permitir que este método seja aplicado todo automaticamente. É sugerido, contudo, que as línguas que compreendam os termos de vocabulários de indexação coordenada, com o uso apropriado de preposições e conectivos, e outras palavras auxiliares que não apareçam como cabeçalhos de assunto ("efeito", "formação", etc., são exemplos desses no contexto da Química), utilizem tais conectivos, facilitando assim o processamento automático, produzindo índices de assuntos de alta qualidade.

Estes índices de assunto produzidos automaticamente poderiam ser comparáveis, segundo Lynch (1966), em qualidade com os índices produzidos totalmente por indexadores humanos, mas com considerável economia de tempo e trabalho. Para ele, esta forma de indexação automática, certamente superaria a indexação por *KWIC* (*keyword in context*) e *KWOC* (*keyword out of context*) em facilidade de uso, uma vez que o grau de organização possível dos documentos com esses métodos era severamente limitado.

Borko (1977), ao argumentar sobre a necessidade do desenvolvimento de uma teoria da indexação, apresentou os trabalhos de Landry (1971) e Salton (1975) que aproximam tal processo da teoria matemática da comunicação, utilizando como soluções o desenvolvimento de uma linguagem documentária baseada na lógica notacional para diminuir os ruídos e silêncios nos SRIs.

Landry (1971) estabelece três classes conceituais - conjuntos de documentos, conjuntos de atributos e conjunto de relações - para expressar conexões entre os documentos e os atributos. Para Landry (1971), o papel da indexação, que é ao mesmo tempo processo e sistema, está totalmente voltado para especificar os dados dos elementos e das relações do fluxo documental por meio da preservação da ordem de suas transformações.

A linguagem documentária seria elaborada a partir de uma notação lógica com uma rigorosa representação matemática que padronizasse as entradas dos documentos, permitindo que qualquer entidade fosse representada por palavras, por um rol de palavras, por títulos, etc. O documento seria bem indexado por um conjunto de dados elementares representados por essa linguagem documentária lógica.

Segundo Landry (1971), a confiabilidade da recuperação dos documentos depende da exatidão do processo de indexação e esta é garantida pelo desenvolvimento de uma linguagem documentária lógica que represente os dados elementares dos documentos e as relações entre si, controlando as imperfeições da linguagem natural. Dessa forma a comunicação efetiva seria possível nos SRIs.

Gerard Salton (1975) baseia-se na utilização do computador para realizar a conexão entre a quantidade de documentos armazenados e os sistemas de recuperação, com ênfase na indexação do vocabulário. Para desenvolver seu trabalho, Salton sentiu a necessidade de desenvolver aspectos teóricos importantes como as funções dos termos de indexação, o estabelecimento de procedimentos capazes de aferir a qualidade da indexação dos termos (boa ou ruim), a definição de técnicas que possibilitem a modificação da indexação do vocabulário com a finalidade de melhorar a eficácia da recuperação do documento.

Salton (1975) parte dos pressupostos de que a indexação consiste em atribuir os termos de um documento. Estes termos são atributos escolhidos para representar o conteúdo informacional do documento. Os atributos também podem ter valores, de modo que em um arquivo de registros de pessoal, os atributos podem ser o nome do funcionário, a classificação de cargos, salário, etc., e o valor do atributo seria o nome da pessoa, os nomes dados pela classificação das funções, etc. Salton (1975) utiliza a *Colon Classification*, de Ranaganathan, para estabelecer como os principais atributos as facetas PMEST - personalidade, matéria, energia, espaço e tempo.

Os valores de atributo seriam os termos específicos usados para descrevê-lo. Na teoria da indexação de Salton (1975), os documentos são reunidos em grupos (*clusters*) e podem ser facilmente separados por suas classes (atributos). Nesta distribuição, a distância entre dois documentos é inversamente proporcional à similaridade de seus vetores de indexação. Quando os documentos são separados no espaço dos objetos é possível obter alta precisão se não for alta a revocação, porque cada documento é recuperável sem também recuperar todos os seus possíveis vizinhos não relevantes. Salton enfatiza que as propriedades da indexação de um vocabulário são a exaustividade e a especificidade.

A especificidade denota o nível de detalhamento em que um conceito é representado em seus vetores. A exaustividade designa a completude com que os tópicos relevantes de um documento são representados na indexação de um vocabulário e no vetor índice-documento.

Quando se recupera documentos de coleções pela indexação específica, as pesquisas são de alta precisão nas quais um número relativamente pequeno de documentos são recuperados, mas altamente relevantes. Já a indexação exaustiva leva à revocação de um grande número de documentos, no entanto, os documentos relevantes e não relevantes são recuperados na mesma proporção. De acordo com a teoria de Salton (1975), a especificidade e a exaustividade são sempre inversamente proporcionais no que se refere à recuperação de documentos.

Essas teorias procuraram controlar a linguagem natural a fim de automatizar o processo de indexação, com base em critérios eficazes que permitissem acelerar a inclusão dos documentos nos SRIs, no entanto, garantindo a eficiência e eficácia em relação às buscas realizadas e os documentos recuperados, evitando a perda informacional.

Os sistemas de recuperação da informação cada vez mais automatizados fizeram surgir teorias da indexação preocupadas em desenvolver linguagens que facilitassem a interface entre os processos realizados pelo trabalho humano e pelos computadores.

Coates (1970) procurou desenvolver um mecanismo para tornar a indexação da informação, originalmente introduzida numa linguagem de indexação particular, imediatamente utilizável por outra instituição que empregasse uma linguagem de indexação diferente.

Inspirado na proposta dos documentalistas franceses do Grupo de Estudos sobre a Informação Científica do CRNS, de Marselha, procurou desenvolver uma linguagem de indexação mediadora, uma linguagem de indexação intermediária, nas palavras do próprio autor, uma linguagem de comutação (COATES, 1970, p. 103).

O debate sobre o desenvolvimento de um léxico intermediário foi realizado no CRG em 1968 e o grupo procurou pensar na construção de uma ferramenta ou software no qual os termos de indexação que fossem aplicados a uma dada entrada de documento em um sistema de indexação pudessem ser convertidos para os seus equivalentes conceituais em qualquer outro sistema de indexação por meio de operações rotineiras e com o mínimo de esforço intelectual. A conversão empregaria o método de concordância indireta entre sistemas, isto é, existiria uma concordância de equivalências entre cada um dos sistemas de indexação e o léxico intermediário (Classification Research Group Bulletin, 1968, p. 292)

Um esboço de tal linguagem foi preparado para aplicar na literatura da Ciência da Informação. A versão inglesa foi chamada de Léxico Intermediário (COATES, 1970, p. 103). Considerou-se para a escolha da primeira área a ser aplicada o léxico intermediário, a utilidade imediata do uso de uma linguagem de comutação. Se uma linguagem de comutação funcionasse com êxito para a Ciência da Informação, o mesmo mecanismo deveria funcionar pelo menos igualmente bem para outros assuntos de outros campos menos problemáticos em relação à terminologia. Coates pensava em desenvolver uma linguagem de indexação intermediária que

facilitasse a interoperabilidade entre diversos SRIs de inúmeras instituições, de diversas áreas e, inclusive, de diversos países.

A tarefa essencial de uma linguagem de comutação é facilitar a conversão do assunto informativo indexado, originalmente dado em uma linguagem de indexação particular, em entradas que empregam os formulários de indicação de assunto de uma segunda linguagem de indexação igualmente particular (COATES, 1970, p. 103, tradução nossa).

No nível mais simples, seria fácil conceber a construção de um par de tabelas de equivalência que ligariam cada termo de cada linguagem com seu equivalente conceitual em outra, à maneira de um dicionário bilíngue. Com tais tabelas de equivalências o processo de conversão precisaria ser apenas uma operação rotineira. Para Coates (1970), as tabelas de equivalência poderiam desempenhar um papel central na transmissão de informações bibliográficas, quebrando as barreiras de incompatibilidade entre diferentes linguagens de indexação. No entanto, a realização prática de uma linguagem de comutação precisaria procurar isolar as propriedades de linguagens de índice de remetente e de destinatário que iriam interagir (talvez negativamente) com o sistema de conversão (COATES, 1970, p. 104).

O projeto do Léxico Intermediário, que se originou com o grupo francês de Marselha, adquiriu rapidamente um caráter internacional, incluindo participantes dos Estados Unidos, Grã-Bretanha, Bélgica, Dinamarca, Alemanha Ocidental e Polônia com encontros de representantes dos respectivos países ocorridos entre 1965 e 1968. Um dos principais esforços foi incluir as teorias da classificação, desenvolvidas pelo *Classification Research Group* para a Biblioteconomia, na *Library and Information Science Abstracts* (LISA).

Este foi um grande intento, uma vez que o projeto LISA de base de dados, implementado em 1969, procurava reunir a maioria dos periódicos da Biblioteconomia, Documentação, Ciência da Informação, Informática, Telecomunicações e áreas afins de diversos países, com línguas diferentes. Era necessário desenvolver uma linguagem de comutação que permitisse a interação do sistema inglês com os sistemas dos demais países. Estas diferenças dificultavam a aplicação operacional e praticamente nenhuma alternativa ao esquema CRG estava disponível para LISA, dada a exigência de uma linguagem tanto para o campo amplo quanto para a pesquisa de temas específicos (COATES, 1970, p. 104).

O projeto do Léxico Intermediário, de Coates (1970), inclui uma eloquente e persuasiva exoneração de responsabilidade sobre a redução semântica. A redução semântica seria provavelmente essencial até ao nível dos cabeçalhos de assunto. Poderia ser necessário estabelecer uma regra que considerasse um conceito composto a fim de incluir várias noções listadas pelo grupo de cabeçalhos representados. Menos tratáveis de que os problemas decorrentes da diversidade de formas e estruturas de indexação seriam aqueles que resultariam das particularidades sobre a

quantidade de informações que cada sistema de indexação daria acerca dos documentos de origem.

Políticas contrastantes quanto à especificidade, à exaustividade ou ao uso ou não de termos relacionais seriam, naturalmente, refletidas nas linguagens de indexação utilizadas pelos centros de informação que poderiam desejar trocar informações por meio de uma linguagem de comutação.

Se atribuíssemos à linguagem de indexação emissora a propriedade de detalhamento, de acordo com o grau de ligação que permitiria a expressão do detalhe, então certamente, que na situação de transferência, a mensagem que seria processada, por sua vez, pela linguagem de comutação, poderia ser menos detalhada, o que se refletiria na linguagem de indexação receptora. Esta última, certamente, seria a menos detalhada das três linguagens. As informações que seriam relativamente detalhadas na linguagem de indexação emissora seriam representadas se a linguagem de comutação fosse generalista em comparação com a linguagem de indexação emissora, podendo transmitir a forma representada para a linguagem de recepção (COATES, 1970, p.109).

Coates (1970) alertava que a mesma tarefa de indexação dos mesmos documentos estava sendo realizada simultaneamente por várias instituições, esforço desnecessário e duplicado que poderia ser eliminado. A barreira principal a tal economia sempre foi as exigências muitas vezes de originalidade e especificidade na indexação pelos centros de documentação especializados. O conceito de linguagem de comutação tentava conciliar economia com diversidade de representações.

2.4 O COGNITIVISMO NA INDEXAÇÃO E A PERSPECTIVA DO USUÁRIO

A partir da década de 1970, com o uso crescente dos sistemas de recuperação da informação automáticos, a avaliação deslocou-se do desempenho tecnológico destes para o atendimento das necessidades de seus usuários. A visão cognitivista de grande influência nesse período em diversas áreas a partir das contribuições de Piaget (1985) e Vygotsky (1991) na Psicologia Social, também impactaram nos estudos sobre indexação. A informação e o conhecimento passaram a ser compreendidos como resultados de processamentos cognitivos que combinariam a apreensão da realidade e os padrões da mente humana.

Essa abordagem cognitivista na indexação fez crescer o interesse por estudos de usuário, de suas necessidades e do comportamento de busca. Mauperon apresentou um dos primeiros estudos sobre o comportamento dos usuários do Centro de Informação e Documentação da *Euratom*, cujo acervo estava relacionado à área da Energia Nuclear. Para Mauperon (1967), os usuários eram uma matéria-prima privilegiada no tocante à avaliação dos SRIs.

Por usuários eu não considero necessariamente somente aqueles que usam a informação, mas também aqueles que podem se beneficiar de seu uso. Os *potenciais usuários* são ativos em numerosos domínios - investigação e desenvolvimento, engenharia, economia, direito,

seguros, educação e formação, gestão, documentação - para citar apenas alguns (MAUPERON, 1967, p. 232, tradução nossa).

Esta diversidade de atividades sugere uma correspondente diversidade de perspectivas e necessidades de informação. Isso não significa realizar um estudo psicológico de cada uma dessas categorias de usuários. Isso seria demasiado oneroso em tempo e dinheiro.

Mauperon propunha o mapeamento e a análise dos estudos de usuários que diversas organizações especializadas estavam realizando. Muitas centenas de estudos de usuários sobre informações científicas e técnicas foram relatados por Auerbach que menciona 676 (AUERBACH apud MAUPERON, 1966, p. 232). No entanto, Mauperon criticava tal metodologia, pois confundia as necessidades dos usuários com os padrões de uso. Para o autor, esta visão levava a análises ruins, recolhimento de dados equivocados e amostragens irreais, pois os hábitos dos usuários, sua falta de treinamento ou habilidade, os efeitos de recursos documentais insuficientes, ou experiências infelizes tornavam as estatísticas de "uso" pouco confiáveis. "O que importa mais do que o usuário é, em muitos casos, o que ele deveria fazer" (MAUPERON, 1966, p. 233).

A proposta era avaliar a ordem de grandeza do número de consultas que poderiam os SRIs esperar receber todos os dias. O projeto realizado na *Euratom* tinha como objetivo acompanhar os hábitos de seis pesquisadores de nacionalidades diferentes, procurando perceber a ampla gama de padrões. Destarte, o número de consultas, devido a diversidade de países da comunidade representados, seria proporcional ao número total de potenciais usuários, sendo o fator de aproximadamente um décimo, ou seja, uma consulta para cada dez cientistas por ano.

A teoria de Mauperon (1966) afirma que para alcançar uma boa cobertura a melhor solução seria adquirir todas as publicações dentro do escopo (havia mais de dez mil periódicos no mundo que lidavam parcial ou inteiramente com energia nuclear já naquela época e mais de doze mil relatórios todo ano). Entretanto, os custos seriam altíssimos em relação a compra e espaço físico.

Desse modo, o estudo de usuário tornou-se importante, uma vez que o Centro de Informação e Documentação da *Euratom* decidiu usar as fontes secundárias, ou seja, índices e resumos de outras instituições para a partir da demanda efetuar a aquisição da assinatura dos principais periódicos. Contrariamente a uma suposição bastante difundida, a utilização de um resumo em vez de um texto completo num sistema de documentação orientada para a recuperação não causa uma perda significativa de informação, é importante mencionar que essa compreensão de Mauperon não é consensual na literatura sobre o processo de indexação. Por outro lado, o intervalo de tempo que é inerente à publicação de periódicos secundários é por vezes considerável e impede que os documentalistas que os utilizam exclusivamente forneçam aos seus clientes literatura recente (MAUPERON, 1966).

Mauperon (1966) enfatiza a importância da indexação dos documentos científicos e técnicos. Em sua visão, os resumos desses documentos, escritos em linguagem natural, deveriam ser transformados em informações expressas em termos de uma linguagem de indexação prontamente compreensível por um computador.

Basicamente, dois caminhos estavam abertos para nós. Poderíamos decidir analisar um número suficientemente grande de documentos - digamos vários milhares - atribuindo-lhes os termos encontrados ao ler os documentos. Por contagem de frequências e por ordem alfabética listagem dos termos assim recolhidos, seria então possível seleccionar os mais adequados. Isto, naturalmente, teria sido muito longo, e, uma vez que o tesouro tivesse sido elaborado, teria exigido a análise dos mesmos documentos novamente. A outra maneira era examinar e estudar listas de palavras-chave já seleccionadas por outras organizações que trabalhassem nos mesmos domínios ou em domínios afins. Este processo foi, obviamente, muito mais rápido e apresentou a vantagem adicional de aproveitar um considerável repertório de experiências (MAUPERON, 1966, p. 236, tradução nossa).

Foi nesse contexto que surgiram as operações de recuperação da informação envolvendo estratégias de pesquisa que usavam operadores booleanos (e, ou, não) nos SRIs automatizados. O arranjo conveniente de palavras-chave na formulação da consulta reduzia o tempo no computador ao mínimo necessário: processar vinte consultas em um acervo de 500.000 documentos levava então uma média de dezoito minutos. O número de referências esperadas para ser recuperado numa dada pesquisa podia ser previamente estimado utilizando uma fórmula empírica, compreendendo a frequência de atribuição dos descritores na formulação de consulta.

Mauperon (1966) procurou adaptar as necessidades dos usuários do Centro de Informação e Documentação da *Euratom* a partir de um sistema assistido por computador que pudesse responder rapidamente à maioria de perguntas científicas e técnicas sobre energia nuclear, listando as referências das publicações que abordavam o assunto. Descobriram que os usuários queriam mais do que apenas referências bibliográficas, por isso passaram a fornecer resumos também.

Nas décadas de 1970 e 1980, as teorias da indexação, além da preocupação com os fatores semânticos, de criar linguagens de indexação com princípios e postulados que garantissem a delimitação dos assuntos dos documentos, não baseada no empirismo, mas a partir de critérios compreensíveis aos indexadores e usuários dos SRIs, também estavam voltadas a desenvolver aplicações mais eficazes de indexação que permitissem a recuperação do documento de acordo com a demanda do usuário. Não bastava mais analisar o documento para a extração de seu conteúdo e representação mais adequada, utilizando os princípios da teoria da classificação. Intensificaram-se então os estudos dos perfis dos usuários dos SRIs para melhor qualificar os graus de relevância e pertinência dos documentos em um sistema.

Segundo Foskett (1970), a relevância, de qualquer parte do conhecimento, pode ser concebida como aquilo que se enquadra no padrão geral de uma área maior, que é uma parte reconhecida e estabelecida pelo consenso entre os especialistas nessa área. Já a pertinência significa

aquilo que se enquadra em um padrão particular que um indivíduo tenta construir em sua própria mente. A relevância é o que determina o conteúdo de qualquer serviço de indexação ou resumo publicado, por exemplo, os indexadores não estão relacionando sua seleção de material a um indivíduo, mas a um consenso. A pertinência, por outro lado, é o que determina a seleção oferecida por um bibliotecário ou profissional da informação a um investigador específico, que é matéria mais subjetiva, menos consensiosa. "Pode-se razoavelmente dizer, penso eu, que a contribuição que um indivíduo faz ao avanço do conhecimento pode ser julgada pela medida em que ele consegue transformar o que é pertinente para ele em o que é relevante para o consenso" (FOSKETT, 1970, p. 91, tradução nossa).

Foskett argumenta que as chamadas Ciências Exatas e Naturais, pela necessidade do desenvolvimento de conhecimentos específicos, conseguiram um alto grau de consenso em suas terminologias e estas passaram a ser incorporadas na indexação dos documentos nos mais variados centros de documentação. No entanto, essa era uma realidade distante nas Ciências Sociais. Para ele, uma terminologia convencionada é um vocabulário controlado. Não se deve usar um método de vocabulário descontrolado, ainda que a alegação seja da utilização da linguagem natural dos autores, pois cada um possui uma maneira de ver a sua própria obra e não haverá qualquer padronização dos índices.

A visão de que a análise de faceta fornece uma base sólida para todas as formas de controle de informação já era de grande aceitação nos EUA, em grandes programas do *American Institute of Physics*, por exemplo, e no *Biological Sciences Communication Project da George Washington University*, no início dos anos de 1970. Foskett enfatizava a necessidade da construção de vocabulários controlados baseados no que mais tarde Beghtol (1986) chamou de garantia literária. Segundo ele, a indexação *KWIC* era pouco confiável. A falta de referências, mesmo das mais elementares, a grande incidência de sinônimos, e a falta de termos relacionados, levava ao fracasso, ou melhor, a ausência de possibilidades de fornecer qualquer tipo de orientação para ajudar o usuário a estruturar seu pensamento.

Parece-me ser um recuo de tudo o que a boa indexação se propõe a ser. Quando essa técnica [KWIC] é usada para fornecer um índice de assunto a um arranjo classificado, é claro, que esta crítica não se mantém na mesma medida, embora ainda se possa facilmente ignorar itens pertinentes e relevantes por não passar numa mesma palavra o pensamento como um autor usou em seu título (FOSKETT, 1970, p. 94, tradução nossa).

Foskett relata que o sistema de indexação baseado no *KWIC* responsabiliza o usuário sobre o sucesso ou fracasso em sua busca por assuntos, uma vez que seus criadores entendem que os conceitos e as categorias devem estar já bem elaborados na mente do usuário e, portanto, que não faz parte do trabalho de um indexador ajudar a pensar os conceitos. No entanto, para Foskett faz

parte da função do indexador facilitar as buscas dos usuários, simplificando seus processos de recuperação dos documentos por meio de cabeçalhos de assunto baseados na classificação facetada (1970, p. 95).

Portanto, a tecnologia implementada nos SRIs possibilitou a indexação de múltiplas entradas de assunto dos documentos com uma coordenação sistemática de conceitos. As linguagens documentárias, cada vez mais refinadas e padronizadas em relação aos conceitos, permitiam essa interação entre homens e máquinas, diminuindo os ruídos e silêncios. O autor cita a adoção da Teoria Geral dos Sistemas e dos Níveis Integrativos na construção dessas linguagens formais e que foram amplamente discutidas no âmbito do CRG, o que havia facilitado a integração entre diversos SRIs.

Preocupado com interação dos sistemas com os usuários e os indexadores, Foskett abordou a forma em que os indivíduos compreendiam a realidade, o universo, os objetos ao seu redor impactava na maneira de classificar os conceitos, as categorias, os assuntos.

JS Bruner, JP Guilford e Jean Piaget demonstraram que é investigando o ambiente imediato e construindo um esquema conceitual que a criança vem a compreender o seu mundo e a desenvolver a capacidade de abstrair, assim é capaz de lidar com circunstâncias que ela nunca conheceu antes. É óbvio, no entanto, que esses objetos familiares, ou entidades, eles próprios derivam de outras entidades já existentes. Não podemos classificá-los com precisão sem levar essas outras entidades em consideração. A cadeira vem da madeira, a madeira da árvore, a árvore da semente, e assim por diante (FOSKETT, 1970, p. 97, tradução nossa)

Para Foskett (1970), a teoria dos níveis integrativos facilita a compreensão das relações entre as entidades conceituais, pois permite ordená-las em uma sequência linear com classes principais num esquema de classificação. As entidades em cada nível podem se combinar para formar novas entidades, novos conceitos a partir do rearranjo de unidades únicas, embora constituído por uma série de unidades separadas reconhecidas anteriormente. As unidades constituintes e as relações que venham a existir entre elas podem ser conectadas, estas vão determinar a natureza da nova unidade de nível superior. O todo é maior do que a soma de suas partes.

Ao sugerir a adoção da teoria dos níveis integrativos para a representação das publicações da área das Ciências Sociais, Foskett a associa à própria teoria funcionalista durkheimiana, no sentido em que os indivíduos fazem parte de um todo e que cada célula e seu bom funcionamento leva a constituição da sociedade e as suas combinações garantem o seu desenvolvimento. A própria forma de representação por meio dos níveis integrativos, facilita o desenvolvimento das pesquisas no campo da Ciências Sociais uma vez que estão em consonância com as regras do método sociológico, de Durkheim.

A teoria do nível integrativo forneceria um conjunto de hipóteses para explicar que as crenças na possibilidade das unidades exibirem de um modo particular o todo e que a organização poderia ser totalmente explicada pela análise das partes isoladamente não eram possíveis de se concretizarem. As integrações, as relações, os vínculos são importantes para compreender o homem e a sociedade (FOSKETT, 1970, p. 98).

Foskett enfatiza a importância da contribuição intelectual do indexador no processo de construção de linguagens artificiais que serão utilizadas no tratamento dos documentos e das informações pelos SRIs. Para o autor o aumento exponencial dos documentos e das informações levaria à necessidade de sistemas cada vez mais eficientes e confiáveis que não sobrecarregassem os usuários no momento de suas buscas, mas que fornecessem acessos prontos e rápidos. Dessa maneira, o processo de indexação precisaria incorporar as máquinas para tratar as massas de informação, na leitura e processamento dos dados, considerando o trabalho intelectual do indexador que transformaria essas massas em possibilidades de acesso de acordo com as necessidades dos pesquisadores que utilizam esses SRIs.

A teoria cognitiva foi adaptada para a classificação e indexação a partir da apropriação da abordagem de Van Dijk e Kintsch acerca do processo de leitura, desenvolvida entre os anos de 1978 e 1983⁶. A compreensão de que o processo cognitivo se dá em uma estrutura flexível e sensível que é a mente humana onde acontecem as interações a partir de etapas como o *bottom-up*, ou seja, o momento em que as percepções filtram e preenchem inferencialmente, e, depois, adaptam a seu repertório as impressões e variações do ambiente. Por esta perspectiva, entender um texto implica em fazer conexões entre as ideias expressas em um documento e um conhecimento relevante anteriormente adquirido.

Na cognição acontece uma espécie de seleção mental de estímulos que são externos à consciência, utilizando como critério natural a procura de coerência no que o indivíduo observa. Além disto, na teoria cognitiva de Van Dijk e Kintsch, afirma-se que há dois fatores principais que influenciam o processo de compreensão: o que está relacionado com os objetivos e motivações do indivíduo e as características específicas da situação em que a compreensão está ocorrendo.

No modelo mental de Van Dijk e Kintsch, o processo se dá por meio de representações e das integrações destas. A construção do sentido “correto” ou contextualmente apropriado resulta do processo de integração que, rapidamente, desativa as construções contextualmente não apropriadas. Logo, há três níveis de representação cognitiva relacionados à compreensão de textos escritos: (1) o código de superfície, o qual conserva o léxico exato e a sintaxe das orações; (2) o texto base, que

⁶O modelo mental proposto por Van Dijk e Kintsch foi apresentado no livro dos pesquisadores **Strategies of discourse comprehension**. San Diego: Academic Press, 1983. Destaca-se o trabalho de Brookes (1977; 1980) sobre a apropriação das teorias cognitivas na Ciência da Informação e o de Belkin (1982; 1984) sobre os modelos cognitivos para o processamento da informação.

contém as proposições do texto explícito de uma maneira que mantenha o seu significado, mas não o léxico ou a sintaxe exata do texto (inclui algumas poucas inferências necessárias para estabelecer a coerência do texto) e (3) o modelo situacional, ou seja, a representação mental daquilo que está explicitamente mencionado ou inferencialmente sugerido no texto (GRAESSER; ZWAAN, 1995). No entanto, o modelo final de representação é unitário.

Nesse processo da leitura, de acordo com o modelo de Van Dijk e Kintsch, há três divisões convencionais da memória: o registro sensorial, a memória de curto prazo (de trabalho) e a memória de longo prazo. A função do registro sensorial é transmitir informações visuais para a memória de trabalho. Já a memória de curto prazo (*short-term memory/STM*) realiza o processamento e armazenamento limitado dessas informações. Em contrapartida, a memória de longo prazo (*long-term memory/LTM*) compreende "a totalidade do conhecimento e crenças de um indivíduo sobre o mundo" (VAN DIJK; KINTSCH, 1983, p. 310, tradução nossa).

Beghtol, em seu célebre artigo, *Bibliographic classification theory and text linguistics*, de 1986, procura pensar as implicações desse modelo de representação cognitiva para a teoria da classificação bibliográfica. Sua proposta de compreender os sistemas de classificação como um texto que precisa ser compreendido em seu contexto, uma vez que a sua leitura possibilitará o acesso aos demais documentos, teve grande repercussão no campo da indexação uma vez que, a partir do modelo do processo cognitivo de leitura de Van Dijk (1976), a autora propõe a análise dos assuntos pertinentes ao documento.

Beghtol procurou compreender o modelo descritivo do processo cognitivo da classificação de documentos. Sua proposta foi aproximar a linguística textual da teoria e prática da classificação, investigando as relações entre elas. Beghtol (1986) discute a questão da tematicidade (*aboutness*) e do significado (*meaning*) a partir de dois aspectos: a compreensão extensional, ou seja, quando o assunto é inerente ao documento, o seu entendimento se dá a partir dele; e a compreensão intensional, quando a razão ou propósito para o seu entendimento é externo ao documento, está mais associada às causas em que ele foi incorporado a um SRI ou solicitado por um usuário.

Para Van Dijk (1976), há diferentes graus de relevância que o leitor atribui às informações fornecidas por um documento e que estão relacionados aos graus estruturais e semânticos de um determinado texto que são identificados no processo de leitura. Essa noção é recuperada por Beghtol para abordar o processo cognitivo durante a representação dos documentos. Para a autora, um mesmo documento pode ter diferentes significados inclusive para o mesmo leitor em momentos e contextos diferentes. No entanto, o documento em si é imutável, possui um assunto fundamental, a chamada tematicidade (*aboutness*).

O tratamento dessa tematicidade do documento tem sido objeto dos tradicionais estudos sobre classificação e a compreensão de que todo documento carrega em si sua questão fundamental

é o seu pressuposto basilar. Os sistemas de classificação foram criados a partir dessa perspectiva, procurando contemplar esquemas de notação hospitalares para que qualquer conhecimento contido em um documento pudesse ser representado (FOSKETT, 1973). No entanto, estes estudos não consideram que a questão inerente ao documento muda de acordo com a leitura de seu classificador, de seu usuário, com a incorporação de outro determinado significado ou com outro uso sendo atribuído.

Para Beghtol (1986), os estudos de classificação precisavam contemplar o processo cognitivo em sua abordagem aos documentos. Pela influência dos estudos de classificação nas abordagens sobre indexação desde os trabalhos do CRG, essa contribuição também teve grande ressonância na aplicação desta técnica, uma vez que a reflexão central trata-se da análise da tematicidade dos documentos (*aboutness analysis*).

A incorporação dos estudos cognitivos para o tratamento temático dos documentos fez com que a prática da indexação passasse a incorporar em sua fase de análise conceitual a preocupação com a forma em que as pessoas leem os documentos e quais caminhos percorrem para identificar os seus principais assuntos. O modelo de Van Dijk (1976), adotado por autores como Beghtol, retoma a abordagem da compreensão de que um texto pode se dar de duas maneiras: pelo método de cima para baixo (*top-down*), baseado em esquemas dedutivos, e pelo método de baixo para cima (*bottom-up*), baseado em esquemas indutivos. Neste segundo método, a compreensão de um documento pelo seu leitor resulta de sua capacidade de reduzir a informação em um texto a proporções gerenciáveis e possíveis de serem registradas na memória.

Nosso comportamento linguístico mostra que podemos dizer que um discurso, ou parte dele, era sobre algo, ou seja, somos capazes de produzir outros discursos, ou partes de discursos, expressando o "assunto principal" em resumos, títulos, conclusões ou declarações sob qualquer forma (VAN DIJK apud BEGHTOL, 1986, p. 90).

Sendo assim, identificar o sentido, a questão semântica, de um discurso trata-se de um processo cognitivo de uma redução automática, processo de síntese que permite ao leitor construir durante a leitura uma noção do texto em tópicos e armazená-los na memória em estruturas dispostas hierarquicamente para posterior recuperação.

Essas ações cognitivas de compreensão de um texto para gerar uma afirmação sematicamente precisa do assunto principal de um discurso são, segundo Van Dijk (1976), regidas por macro-regras que permitem que as informações menos importantes sejam retiradas da memória durante a leitura e a macroanálise do texto. Van Dijk identifica cinco macro-regras: a deleção fraca, a deleção forte, a regra zero, a regra de generalização e a regra da construção.

Farrow (1991) apropria-se dessas macro-regras para abordar o processo de indexação. Propõe o aprendizado-orientado, no sentido em que a leitura do indexador é uma tarefa-orientada, é

uma leitura do documento com um propósito específico e imediato: de produzir a sua indexação. Os processos de compreensão dos textos por classificadores, produtores de resumo, indexadores, entre outros profissionais dessa área, é similar, têm como objetivo a representação dos documentos lidos. No entanto, há diferenças nas propostas de análise de conteúdo.

O classificador procura representar o documento como um todo e traduzi-lo para a linguagem do sistema de classificação. O produtor de resumo, preocupa-se em representar os principais argumentos do documento. Já o indexador lê o texto dividindo-o em informações relevantes e periféricas, e emprega o mesmo processo *top-down* (conceitual) e *bottom-up* (perceptual) da leitura normal, fluente.

Em termos cognitivos, para Farrow (1991), o processo de leitura é fortemente localizado junto à área de processamento humano da informação. A informação do mundo é absorvida pelos sentidos em um registro sensorial. Essas informações são captadas de forma não processada durante um período muito curto. A informação pode ser selecionada por um processamento adicional porque a mente reconhece esta como um padrão já previamente armazenado ou porque não é um padrão reconhecido e a mente deseja processar o novo. Farrow (1991), utilizando o modelo Van Dijk e Kintsch, afirma que o processamento da informação se dá por meios cognitivos na interação entre memória de curto prazo (STM) e memória de longo prazo (LTM)⁷.

O processamento de texto pela mente seria uma parte do processamento da informação. Um texto, de acordo com o modelo baseado nessa teoria cognitiva, consiste em uma sequência coerente de proposições, sendo uma proposição a menor unidade de discurso que é capaz de ser avaliada. O ponto de partida do modelo é a habilidade da linguagem dos usuários de ler uma passagem do texto e resumir seu tema e assunto, do que se trata, em uma sentença intuitivamente e convencionalmente aceitável. O status específico das sentenças numa passagem do texto é a busca do seu significado e das relações com os outros significados das próximas sentenças na sequência.

As macroestruturas do texto propiciam ao indexador bases para decidir quais partes do texto devem ser indexadas. Farrow (1991) defende que a indexação utilize como parâmetros de seu processo as macroestruturas como regras. Então, na delação fraca - a informação seria simplesmente deixada de fora, apenas as informações acidentais podem ser eliminadas, isto é, informações que podem ser omitidas sem mudar o significado, nem influenciar na interpretação das sentenças subsequentes do discurso; já na delação forte - haveria a exclusão de informações localmente relevantes, as informações excluídas seriam as com especificidades normais ou associações previsíveis; na regra de delação zero - não haveria qualquer tipo de redução; na regra de generalização - vários objetos ou propriedades de uma mesma classe superordenada seriam

⁷ Os termos estão abreviados do inglês da forma que são internacionalmente reconhecidos *short-term memory* (STM) e *long-term memory* (LTM).

referenciados globalmente, com o nome de uma classe superordenada; e na regra de construção - haveria a combinação ou integração de informações que denotaria propriedades, causas, componentes, consequências essenciais, etc.

A “leitura dinâmica” (*skimming text*) é outro modelo de compreensão adaptado da teoria cognitiva por Farrow (1991) para a indexação, cuja leitura ocorre em uma taxa média de 300 palavras por minuto, e os indexadores voltados ao livro (*back-to-book*) poderiam ser treinados ou não a utilizá-la, dependendo dos prazos estabelecidos.

A produção da indexação é considerada como produção textual, em termos cognitivos, o conjunto de entradas indexadas é o mais primitivo e simples de todos os tipos de produção textual. Há duas razões principais para isso: em todas as situações normais de indexação, a produção de entradas advém imediatamente da compreensão do texto, de modo que a perda e interferência do sentido são minimizadas; além disso, em comparação com discurso normal, os termos indexados são em grande parte desprovidos de sintaxe, permitindo que o indexador concentre-se na semântica para a produção de entradas (FARROW, 1991, p. 157).

Farrow (1991) enfatiza que, independentemente do modelo adotado, o processo de produção da indexação precisa levar em consideração as demandas dos usuários e também prever essas possíveis demandas.

Fugman (1985) ao pensar nas possibilidades do tratamento temático dos documentos para as publicações na área da Química, compartilhou dessa abordagem cognitiva ao propor seus cinco axiomas para uma teoria da indexação. Os cinco axiomas sugeridos são: (1) Definibilidade (*definability*) - o conjunto de respostas relevantes de um tópico somente pode ser determinado pela extensão com que o usuário-requerente (*querier*) o define em termos conceituais e em suas relações; (2) Ordem (*order*): qualquer conjunto de respostas relevantes para um tópico é um processo de criação de uma ordem; (3) Grau suficiente de ordenamento (*sufficient degree of order*): as demandas de busca e as frequências destas criam o grau de ordenamento e o aumento do tamanho da coleção; (4) Previsibilidade representacional (*representational predictability*): a exatidão de qualquer busca direcionada a textos relevantes (especialmente em relação à revocação) depende da previsão de modelos de expressão para conceitos e relações conceituais na pesquisa; (5) Fidelidade representacional (*representational fidelity*): a exatidão de qualquer busca direcionada a textos relevantes (especialmente em relação à precisão) depende da fidelidade de modelos de expressão para conceitos e relações conceituais na pesquisa.

O trabalho de Fugman (1985) caracteriza-se pela análise do comportamento dos usuários nas buscas. O autor chega à conclusão que a expansão dos SRIs por conta do aumento do acervo documental leva à diminuição da capacidade de recuperação pelo usuário, pois a proporção de buscas realizadas em relação à totalidade dos documentos é alterada. O tempo utilizado na pesquisa

geralmente é o mesmo, não importando o tamanho do acervo. Então, a capacidade de realizar a varredura de partes menores vai se perdendo, este comportamento do usuário exige um alto grau de ordenamento dos documentos⁸.

Fugman (1985) apresenta duas formas diferentes de graus de ordenamento: um que procura evitar a perda de informação (linguagem de indexação necessária) e outro que procura evitar o ruído na recuperação da informação. O autor a partir de seus cinco axiomas propõe a indexação imperativa (mandatória) (*mandatory indexing*) que usa os termos mais apropriados e as expressões relacionadas a estes termos, por isso, a importância da sintaxe que fornecerá a hierarquia e as relações de coordenação e subordinação dos termos. A escolha pelo termo de indexação é sempre pelo termo mais específico do vocabulário, além de indexar também o termo equivalente mais próximo.

Tanto Beghtol (1986), quanto Farrow (1991) e Fugman (1985) se basearam na teoria cognitiva para pensar suas abordagens da análise de assunto nos documentos. Esses autores partem do pressuposto que o processamento do discurso precisa considerar a interação entre os sistemas, classificadores, indexadores, resumidores e usuários.

A análise de um texto perpassa por macro-estruturas que permitem o indivíduo reconhecer os assuntos relevantes de um documento lido e esse reconhecimento, de acordo com a teoria cognitiva comportamental, passa necessariamente por seus interesses, objetivos, mas também pela influência de seu ambiente. A análise dos micro e macro-níveis de um documento precisa considerar a posição do leitor no momento em que ele está realizando a redução, a condensação das informações durante o processo cognitivo.

De acordo com Beghtol (1986, p. 93), as descrições dos documentos precisam ser formuladas considerando como as pessoas, incluindo aqueles que realizam o tratamento documental, formulam mentalmente os assuntos dos textos escritos. É preciso considerar nos processos de classificação, indexação, produção de resumos questões como: as relações entre os documentos classificados entre si; as relações entre os sistemas de classificação e os documentos classificados; o processo cognitivo de classificar os documentos para a compreensão de texto e seu armazenamento em memória semântica.

Resgatando as abordagens influenciadas pela teoria cognitiva para o estabelecimento do assunto, o leitor processa um texto de baixo para cima (*bottom up*) e, simultaneamente, insere o conhecimento pré-existente do quadro de cima para baixo (*top-down*) na leitura do texto. Acredita-se que esses processos complementares resultem em uma compreensão do que é o texto. Outros elementos podem, no entanto, entrar na compreensão do leitor de um texto. A compreensão pode,

⁸ O diagrama de Fugman está representado no anexo D do presente trabalho.

por exemplo, "ser influenciada por outros fatores cognitivos, como interesse, tarefa, propósito, conhecimento, normas ou opiniões e atitudes" (BEGHTOL, 1986, p. 98, tradução nossa).

As discussões de interpretação geralmente se centravam na inserção de vários tipos de conhecimento no quadro diretamente da mente do leitor. Entretanto, Beghtol (1986), Farrow (1991) e Fugman (1985) estavam preocupados com o processo cognitivo na representação dos documentos nos SRIs, sobretudo, para a compreensão de seus usuários.

Uma teoria do processo cognitivo de classificação dos documentos deve explicar como o classificador sequencialmente: 1- transforma a estrutura de superfície do documento em sua estrutura lógica proposicional profunda; 2 - transforma a estrutura superficial do sistema de classificação em sua estrutura lógica proposicional profunda; 3 - junta essas duas estruturas proposicionais profundas usando o atual sistema de quadros de classificação e instruções como um sistema virtual artificialmente construído para expressar os sistemas reais de linguagem natural dos documentos primários; e 4 - transforma a estrutura proposicional única resultante do 'sistema de classificação aplicado ao documento' de volta à estrutura de superfície dos quadros de classificação, que gera a notação apropriada e conclui que o documento foi adequadamente colocado na classe com outros documentos aos quais é quase intelectualmente e intertextualmente similar. Esta declaração geral do processo cognitivo de classificação de documentos abrange o que se aprende a fazer quando se aprende a classificar (BEGHTOL, 1986, p. 102, tradução nossa)

Nesta visão cognitiva, há dois tipos de leituras de relevância que a mente realiza durante a compreensão de um documento. No primeiro, acontece a atribuição da relevância textual normal, que ocorre no nível de micro-análise (sentença) ou macro-análise (texto). Em ambos os níveis, a atribuição da relevância normal distingue-se da atribuição da relevância contrastiva ou diferencial.

No nível macro, a relevância normal é a pertinência independente de um discurso e a relevância diferencial é a base para o que se chama de significado do texto para um usuário. A atribuição da relevância diferencial ocorre quando aquele que está realizando a análise seleciona diferencialmente itens para tratamento especial entre itens similares, isto é, itens no mesmo nível hierárquico.

Tal tratamento especial é pensado para resultar de maior interesse, um propósito imediato ou uma tarefa especial, entre outros. No segundo, acontece a relevância contextual, em que o contexto cognitivo (quem também é social e comunicativo) define quais elementos do texto são considerados importantes pelo leitor. No caso da classificação, indexação e produção de resumos dos documentos, esses elementos de relevância contextual teriam a contribuição da garantia cultural e classificatória dos sistema de recuperação da informação para serem identificados.

Nessa perspectiva, o *aboutness* de um documento é o principal tema que o classificador determina durante a análise conceitual e o significado (*meaning*) do assunto de um documento é a causa, o motivo pelo qual um usuário pode querer recuperá-lo. Neste caso, a relevância conceitual do *aboutness* e do seu significado (*meaning*) no sistema de classificação influencia a análise

conceitual do documento e pode gerar um assunto⁹, diferencialmente relacionado à estrutura disponível, às classes e aos dispositivos notacionais de classificação.

Para Beghtol (1986), se as avaliações de relevância diferencial variassem, devido, por exemplo, à complexidade conceitual de diferentes documentos ou por causa dos diferentes níveis de quadros do conhecimento que diferentes classificadores possuem, a classificação poderia se tornar inconsistente.

A análise consistente do *aboutness* foi considerada crucial para a recuperação eficiente de documentos relacionados intertextualmente a partir de sistemas de informação. Por conseguinte, para esses autores influenciados pela teoria cognitiva, é justificável buscar possíveis causas da análise inconsistente do assunto na esperança de que o conhecimento das causas possa levar a métodos de redução da instabilidade ou de anomalias na representação temática dos documentos.

Outras correntes filosóficas aproximaram os estudos da indexação das teorias da classificação além da inglesa e da americana, como é o caso da chamada análise documentária, de origem francesa. Na década de 1960, autores como Gardin, Chaumier e Coyaud sentiram a necessidade de um maior embasamento teórico-epistemológico para respaldar as práticas do tratamento temático dos documentos. Destarte, esses teóricos foram buscar na linguística subsídios por entenderem que a “matéria-prima” da representação dos documentos e da informação é a linguagem.

Percebe-se que nas abordagens como *Eléments d'un modèle pour la description des lexiques documentaires* (1966) e *Analyse documentaire et théorie linguistique* (1974) de Gardin, *Analyse et langages documentaires* (1982), de Chaumier, e *Introduction à l'étude des langages documentaires* (1966), de Coyaud, possuem grande influência da semântica formal na apropriação da linguística para o tratamento dos documentos e construção de linguagens controladas.

A influência da semântica formal nos estudos de indexação será explorada na próxima seção. No entanto, acerca da corrente francesa não iremos aprofundar nossa análise por entendermos que há diversos trabalhos publicados no Brasil neste sentido, destacando-se a atuação do grupo Temma nas décadas de 1980 e 1990, vinculado à Escola de Comunicação e Artes da Universidade de São Paulo e ao Departamento de Informação e Cultura da Universidade Estadual de São Paulo, no campus de Marília, fundado por pesquisadores que se dedicaram aos estudos de indexação.

As reflexões do Grupo acabaram por afirmar a condição da Análise Documentária como disciplina metodológica que sugere procedimentos para a análise de textos com o objetivo de selecionar conteúdos informativos que poderão ser representados, recuperados e disseminados (TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N.Y. LIMA, V.M.A.). Embora a proposta original de Gardin dirigia-se explicitamente à análise de textos técnico-

⁹ As diferenças entre *aboutness*, assunto e significado (*meaning*) serão abordadas no início da seção 3 deste trabalho.

científicos, na apropriação brasileira esses procedimentos foram generalizados para outros tipos de textos, como por exemplo para a análise de imagem e de dados estatísticos, com adaptações (LARA, 2011, p. 99).

Autores da chamada escola espanhola também dedicaram-se à abordagem do tratamento temático do documento e da informação a partir da perspectiva da análise documentária, consagrada por Coyaud e Gardin, sendo definida como:

uma operação semântica, mesmo que ela não obedeça a uma regra precisa, e que cada organismo, cada analista [...] se limite a ver na ocorrência uma certa regularidade interna fundada mais na experiência ou hábito do que sobre algum procedimento explícito [...] (GARDIN, 1970, p. 631 apud LARA, 2011, p. 97).

Percebe-se que, de forma semelhante à corrente francesa, autores como López Yepes (1984), Isquierdo Arroyo (1990; 1983) e Sagredo Fernández (1983) trabalham com o reconhecimento de unidades lexicais como sendo núcleos de sentido dos documentos e que, portanto, a partir delas a representação de seu conteúdo deve ser feita. Os elementos textuais e não-textuais da linguagem precisam ser considerados para extrair as noções fundamentais para estabelecer as categorias e realizar a sumarização dos documentos. Os documentos possuem estruturas internas que devem ser mapeadas a fim de representá-los, além de uma lógica formal que deve ser identificada na análise de seu conteúdo para posterior tradução em linguagens controladas.

É importante perceber que ambas as correntes, a francesa e a espanhola, acerca do tratamento temático dos documentos e da informação possuem uma intersecção com a linguística e a lógica formal em suas abordagens. A influência dessas filosofias da linguagem nos estudos de indexação será analisada na próxima seção. No entanto, é importante compreender que tanto estas, quanto a corrente inglesa, foco da revisão de literatura apresentada nesta seção, consideram a dimensão simbólica da representação dos documentos e da informação e que esta é o fator permissionário da comunicação entre os indivíduos nos usos dos documentos.

Cabe destacar, tal como aponta Cordeiro (2000), que a representação documentária e a recuperação da informação devem ser analisadas como um processo comunicacional interativo e inserido em um contexto situacional de múltiplas relações sociais, culturais, políticas como ressalta Ingwersen (1996) e inserido nos acordos e tensionamentos da comunidade discursiva a qual pertencem como destacam Hjørland e Albrechtsen (1999).

Tais estudos sobre indexação procuraram nas filosofias da linguagem aporte teórico e científico que pudessem embasar as aplicações de tal processo. Na próxima seção analisaremos a influência de correntes filosóficas da linguagem, sobretudo aquelas que refletiram sobre a produção de sentido, procurando perceber como os teóricos da indexação se apropriaram das mesmas para pensar em princípios e critérios para a análise e representação dos assuntos dos documentos.

3 DO QUÊ SE TRATA? AS FILOSOFIAS DA LINGUAGEM NA ATRIBUIÇÃO DOS ASSUNTOS

O título desta seção e da tese apresenta uma certa duplicidade de sentido proposital: uma indagação sobre qual o objeto da indexação, uma questão que procuramos responder ao longo de todo trabalho, procurando evidenciar o compromisso desta técnica com a análise conceitual e representação dos documentos para possibilitar a recuperação dos mesmos. A outra possibilidade de interpretação da pergunta aponta para uma questão central no processo de indexação: estabelecer o assunto do documento a ser representado.

De acordo com Fujita (2013), a indexação consiste na análise do assunto do documento e representação dos conceitos identificados para a linguagem de indexação do sistema de recuperação da informação. No campo da indexação a análise dos conceitos sempre se converteu em um processo complexo com área de sombras e incertezas em relação ao estabelecimento dos assuntos. Como afirma Begthol (2002), a segunda etapa da indexação, considerada por muitos como a de tradução (Unisist, Lancaster, Chaumier, Vickery, Van Slype) ou de representação (Fujita, Gil Leiva) possui diversos estudos e possibilidades de aplicação para a construção de linguagens documentárias. No entanto, a primeira etapa de análise conceitual do documento para o estabelecimento dos assuntos continua sendo uma operação delicada, pois a própria maneira de abordar o documento enfrenta entraves como a compreensão do que se configura como assunto e como este pode ser determinado.

A própria definição de assunto ganhou destaque nos anos sessenta e setenta quando a tentativa de delimitar essa categoria fez surgir uma dualidade entre o próprio termo e o *aboutness*. Segundo Joudrey (2005), o *aboutness* surgiu na indexação para definir aquilo que seria selecionado em um documento como sendo de grande relevância, em outras palavras, os assuntos considerados de grande relevância para a compreensão do documento. Definição essa que se afastava do termo assunto que passou a ganhar um *status* filosófico e epistemológico na abordagem conceitual do conteúdo do documento. O *aboutness* tornou-se o termo de preferência utilizado pelos indexadores na operação de análise conceitual, relacionado à prática de atribuir os assuntos dos documentos na rotina desses processos de indexação nos sistemas de recuperação da informação.

Como demonstra Begthol (1986, p. 84-85), diversos autores procuraram definir o *aboutness* a exemplo de Faithorne, entretanto não há consenso no campo da indexação acerca dos termos *aboutness* e assunto. Diversos autores procuraram definir circunscrever essas diferenças como Faithorne que dividiu o *aboutness* a partir de duas questões: a extensional, quando é o assunto inerente ao documento, e a intensional, quando é a razão ou propósito para o qual foi adquirido por uma biblioteca ou solicitado por um usuário.

Para Boyce, o *aboutness* seria a topicalidade de um documento, já o seu significado estaria associado a sua informatividade. Van Dijk distinguiu entre o *aboutness* e o *meaning* em um sentido ligeiramente diferente com os termos "atribuição de relevância normal" e "atribuição de relevância diferencial" para o caso especial em que um leitor está lendo um texto em particular. Van Dijk usa "relevância" para significar as relações percebidas pelos leitores entre os vários elementos estruturais e semânticos de um determinado texto enquanto está sendo lido. Uma vez que este uso difere do modo como a "relevância" é convencionalmente usado em discussões de sistemas de recuperação de informações, parece melhor evitar essa possível fonte de confusão.

Cabe ressaltar que no Brasil, como afirma Martinho (2010), o termo *aboutness* foi traduzido de diversas formas como tematicidade, atinência, sobrecidade. No entanto, a dificuldade de tradução tem levado ao uso do próprio termo *aboutness* no Brasil para se referir ao assunto selecionado como sendo o principal tema abordado pelo documento, determinado mediante a análise conceitual, primeira etapa da indexação.

O próprio binômio assunto e aboutness revela o quanto que o processo de indexação está associada à produção de sentido. Determinar o assunto de um documento, seja em seu sentido *lato*, filosófico, associado ao significado (*meaning*), seja em seu sentido *stricto*, de tematicidade do documento, envolve a questão do sentido. Determinar o assunto de um documento é trabalhar com atribuição de sentido. Desse modo, não é possível falar de análise de assunto sem falar de produção de sentido.

A necessidade de estabelecer o assunto e/ou o *aboutness*, assim como apontar os principais conceitos ou categorias de um documento está diretamente relacionada à questão da linguagem: qual é o sentido do documento? Instintivamente, o indexador responde a esta questão a partir de sua concepção filosófica, mesmo sem conscientemente perceber quais os princípios que governam a sua decisão. Identificamos três filosofias da linguagem na abordagem dos estudos de indexação: o que identifica parcelas da realidade; o que partilha eventos mentais entre falantes e ouvintes; o que experencia a partir do fluxo das práticas de uma comunidade linguística.

A questão do sentido aqui se coloca a partir da apresentação de três perspectivas filosóficas que se entende de grande influência nos estudos de indexação. As formas em que a linguagem é concebida e o sentido é apropriado pelo indexador vão impactar na análise conceitual, na seleção dos assuntos e suas representações, ou seja, no processo de indexação dos documentos propriamente dito.

As teorias da indexação apresentadas na seção anterior mostram o impacto da tradição filosófica de inspiração platônica-aristotélica na produção de sentido. O essencialismo, o universalismo e o representacionismo são características indelévels nestas abordagens que denotam a forma de compreender a linguagem e, por conseguinte, o sentido. A intenção é elicitare essas

influências, além de abordar a crítica do pragmatismo a tal tradição e a sua perspectiva de produção de sentido a partir da dinâmica das comunidades discursivas, das relações entre sujeitos, objetos e o mundo, na perspectiva dos usos da linguagem.

A Semântica sempre foi uma importante questão para indexação. Vickery (1952) e Borko (1967) já haviam identificado a necessidade de se recorrer aos estudos linguísticos sobre produção de sentido para procurar definir princípios e critérios que melhor direcionassem a prática da atribuição dos assuntos dos documentos. No entanto, é importante perceber que há diferentes abordagens acerca da produção de sentido, com diferentes filosofias acerca da linguagem que impactam na própria forma de analisar o documento e traduzir o seu conteúdo.

Percebemos que o representacionismo de tradição platônica-aristotélica tem influenciado há bastante tempo os estudos de indexação e que aqui desdobramos em dois paradigmas para melhor compreender a sua influência neste campo: o realismo e o mentalismo.

3.1 O PARADIGMA REALISTA: VER E DIZER O REAL

O realismo pode ser considerado um paradigma surgido da díade platônica entre linguagem e mundo real. Nessa concepção, a linguagem possui uma vocação, é um instrumento do homem usado para expressar o real. O sentido é aquilo que descreve a realidade e as suas coisas. Em **Crátilo**, Platão apresenta sua compreensão sobre como o mundo é formado. Para ele, há um mundo ideal onde a essência de todas as coisas reside. Esta essência transcende à experiência humana.

É possível identificar conceitos, leis, lógicas universais porque as coisas do mundo são criadas a partir da mesma substância e o que muda é a maneira em que a essência é configurada no mundo sensível onde vivemos. Consequentemente, a linguagem seria uma reprodução, uma descrição do mundo real, uma correspondência direta entre as palavras e as coisas. Para entender o realismo, é imprescindível compreender os postulados platônicos sobre o mundo.

Em sua **Teoria das Ideias**, Platão defende que tudo o que existe na realidade possui uma essência apriorística, uma substância que origina tais coisas. Essa essência faz parte do mundo das ideias na qual não tocamos, apenas temos impressões a partir das formas materializadas no mundo sensorial e imperfeito em que vivemos. Apenas acessamos as aparências que estão ao alcance de nossos sentidos, neste mundo corpóreo e mutável.

Qualquer objeto ou ser pode mudar, se decompor ou ser destruído pela ação do tempo e do espaço. No entanto, de acordo com a filosofia platônica, a essência destes permanece inalterada, sendo atemporal e universal, sendo sempre única. Independentemente de sua forma, uma rosa será sempre uma rosa por conta de sua essência.

Para Platão, existiria uma outra dimensão que serviria de abrigo para todas estas entidades autônomas, as substâncias do mundo. Dessa maneira, a eficácia da linguagem estaria em cumprir a

sua vocação, ou seja, dizer como as coisas são, chegar o mais próximo de sua essência. Os enunciados, dessa forma, teriam uma autonomia simbólica.

Na compreensão realista da semântica, a linguagem descreve, informa, representa o real. Tanto na **República**, quanto em **Crátilo**, Platão argumenta que a verdade prevalece sobre o consenso porque sua essência pode ser conhecida. A designação da linguagem é declarar as coisas do mundo e não importa quais as palavras serão escolhidas para isso. “O nome, por conseguinte, é instrumento para informar a respeito das coisas e para separá-las, tal como a lançadeira separa os fios da urdidura” (**Crátilo**, 388c).

No diálogo entre Sócrates e Hermógenes, Platão estabelece a função essencial da linguagem: informar sobre as coisas. A linguagem desempenha bem a sua função se representar de forma fiel o real. A fidelidade estaria na proximidade dessas representações da essência do mundo e de suas partes.

Os nomes, os conceitos, as categorias, as classes devem transcender às opiniões, aos juízos, e representar o real em sua dimensão fixa e atemporal. Nomes, classes, categorias, conceitos, são representações universais, autônomas, transcendentais, entidades com características próprias, pois são extratos da essência das coisas identificadas no mundo. Nesta perspectiva, o sentido é algo inerente à palavra, independe dos indivíduos, é anterior e externo.

No paradigma realista há a necessidade de estabelecer nas expressões, nos discursos, nos textos, nos conceitos uma objetividade essencialista que está no nexo em que estes elementos estabelecem ou não com o real. Esta filosofia está presente nos métodos de indexação que estabelecem a tematicidade unicamente pela análise dos documentos. O sentido se encontra no texto e cabe ao indexador identificá-lo. O conteúdo dos documentos descreve a parcela da sua realidade de maneira autônoma e objetiva.

Soergel (1985) ao abordar a indexação centrada no documento afirma que é uma forma de indexação que foca no documento enquanto entidade e na sua descrição fiel. A compreensão é que o indexador extrai o assunto a partir da análise conceitual do documento.

Uma boa indexação, por conseguinte, é aquela que consiga sintetizar e representar de forma mais sincera possível, garantindo que a representação do conteúdo seja válida por um longo período de tempo, pois está descrevendo aquilo que o documento contém. “O indexador deve ficar atento ao texto e às reivindicações do autor do documento” (LANCASTER, 2003, p. 37).

Muitas abordagens no campo da indexação ao estabelecerem as formas do tratamento do conteúdo dos documentos e a sua tradução para linguagens dos sistemas de recuperação partem da compreensão realista do sentido, na qual a função maior da linguagem é descrever ou representar o real. Consequentemente, podemos identificar o assunto, as ideias e os conceitos principais do documento, pois o sentido está no texto.

Essa visão instrumentalista da linguagem na qual sua principal função é informar sobre as coisas possui grande ressonância nos estudos da indexação, sobretudo nas décadas de 1950 e 1960, como vimos na revisão literária realizada na segunda seção desse trabalho, que consideram estar nos documentos, e somente neles, os assuntos, as informações e os conhecimentos expressos, independentemente se serão identificados ou não. Da mesma maneira que o nome identifica a coisa nomeada, os documentos identificam as coisas registradas por eles. O propósito dos documentos seria registrar informações sobre as coisas, pois seriam produtos da linguagem.

Na indexação centrada no documento, como esclarece Mai (2005), os assuntos podem ser atribuídos independentemente de qualquer contexto ou uso específico do documento. Esta análise baseia-se nos atributos do documento. Os atributos permitiriam a identificação dos principais assuntos e conceitos contidos no documento indexado.

Uma abordagem centrada no documento para a indexação depende da análise dos indexadores de uma série de atributos do documento como base para o estabelecimento dos documentos em questão. Os atributos exatos que as diretrizes de indexação ou os livros didáticos recomendam que o indexador examine variam, mas os exemplos arquetípicos são: o título, o resumo, a tabela de conteúdos, os títulos dos capítulos, os subtítulos dos capítulos, o prefácio, a introdução, o próprio texto, as referências bibliográficas, Entradas de índice, ilustrações, diagramas e tabelas e suas legendas. As recomendações exatas variam de acordo com o tipo de documento que está sendo indexado (monografias versus artigos periódicos, por exemplo) (MAI, 2005, p. 602, tradução nossa).

Desse modo, o processo de indexação funcionaria como um transporte que extrairia do documento os principais conceitos e traduziria para representações documentárias, sejam os resumos produzidos em linguagem livre, sejam os vocabulários controlados ou linguagens de indexação com palavras carregadas de conceitos onde cada termo escolhido é um componente significativo. A filiação com o paradigma realista de inspiração platônica na indexação faz conceber os documentos como objetos de conteúdo estável e objetivo. São informações registradas e constatáveis em sua relação com o mundo.

A indexação torna-se um processo de analisar o conteúdo informacional dos registros do conhecimento e expressá-lo em uma linguagem própria que na terminologia da área é chamada de linguagem documentária ou de indexação. As operações de indexação seriam: identificar os assuntos de um documento, traduzi-lo (representá-lo) para uma linguagem controlada e incluir ou não as relações entre os termos. Este processo é baseado na compreensão de que a linguagem é um sistema de representações tal como no realismo.

De forma semelhante ao realismo platônico, a concepção aristotélica da linguagem acredita que as coisas que formam o mundo possuem essências, que são entidades externas ao sujeito (MARTINS, 2009, p. 461). Existe um laço que une as palavras às coisas e de forma estável. A divergência está na descrença de Aristóteles de que haja uma divisão entre o mundo ideal e o

mundo sensível. Existe uma recusa da filosofia aristotélica à teoria das formas. As formas ideais não possuem qualquer tipo de existência em um outro mundo. Existe apenas a alma, a razão humana, que abstrai as essências universais do mundo por meio das experiências particulares, mediadas pelos sentidos.

3.2 O MENTALISMO: A LINGUAGEM COMO INSTRUMENTO DO PENSAMENTO

A díade platônica é transformada em tríade aristotélica, na qual a mediação entre o real e a linguagem é feita pelo pensamento. A substância das coisas, suas essências, não fazem parte de um mundo ideal tal como concebeu Platão, mas o universal é possível porque a substância que compõe os seres humanos é a mesma. A mente é universal, pois somos semelhantes, o espírito humano partilha da mesma essência. A faculdade racional ocupa um lugar central no desenvolvimento do pensamento e a língua uma posição secundária, subordinada à mente. É importante destacar que o mentalismo aristotélico partilha do pressuposto platônico de que a linguagem é um instrumento para nomear os objetos, no entanto, a diferença é que para Aristóteles a linguagem não diz o real, mas diz o que pensamos do real. Neste sentido, a preocupação do filósofo é compreender a lógica em que compreendemos o mundo.

A Lógica foi instituída pelo filósofo como saber capaz de entender a essência do mundo. Em **Metafísica**, Aristóteles afirma que há leis universais constitutivas de todas as coisas e a maior felicidade do homem seria a descoberta de como elas funcionam. Esse seria o objetivo da Lógica, identificar e compreender as leis universais que constituem os seres humanos e o mundo a sua volta. A linguagem enquanto artefato criado pelo homem para se comunicar com os outros e com o mundo possui uma lógica que é subordinada ao pensamento.

A linguagem do pensamento, na qual podemos abstrair os conceitos universais, é a mesma para todos e, por esta razão, a padronização, a tradução e o controle terminológico são possíveis. As variações se dão apenas em termos de estrutura de cada língua e maneirismos. Não importa se em português denominamos “flor” e inglês “*flower*”, a referência no mundo aponta para a mesma coisa.

Parte-se do pressuposto de que o pensamento é autônomo e universal às línguas, funciona como uma linguagem interna, uma espécie de mentalês, que contém todos os recursos necessários para qualquer conhecimento que os seres humanos possam compreender, pensar ou expressar. Então, o conteúdo dos documentos seria um sistema nuclearmente literal e declarativo que representaria o pensamento daquele que o produziu. A representação seria um sistema de significados essenciais, universais e fixos, compartilhados tacitamente pelos indivíduos.

Este mentalismo de origem aristotélica compreende que o pensamento funciona partir de regras lógicas, de padrões preestabelecidos combinados entre si a partir de cada situação

apresentada pela realidade. A linguagem seria a manifestação mais ampla e acessível da nossa vida mental e por isso constituiria uma porta de acesso insubstituível ao estudo da mente humana. É, conseqüentemente, pela influência do inatismo que estes estudos mentalistas abordam o real. Haveria na mente um mecanismo complexo capaz de se organizar em módulos para analisar, sistematizar, programar e solucionar problemas, questões, conflitos etc.

Por este prisma, a linguagem universal da mente seria o conjunto desses padrões, dessas representações universais, comum a todos os humanos, ou seja, haveria um mentalês que antecederia a própria forma dos indivíduos se comunicarem. A variação se encontraria nas combinações que podem ser realizadas, logo, a mente seria uma grande máquina processadora dessas combinações infinitas. De acordo com Pinker, as coisas nos forneceriam informações que correlacionaríamos com nossos conhecimentos que, por sua vez, as correlacionariam com entidades mentais, "uma operação matemática e lógica" (1997, p. 77). Os processos mentais funcionariam como sistemas de *inputs e outputs* cabendo à mente manipular as representações formais dos símbolos. As línguas seriam apenas expressões codificadas desses pensamentos que possuem autonomia e, por isso, a tradução seria uma operação não somente possível como simples.

3.3 A TRADIÇÃO PLATÔNICA-ARISTOTÉLICA E A SUA INFLUÊNCIA NOS ESTUDOS DA LINGUAGEM

García Gutiérrez (1998) aponta que para o estabelecimento de metodologias de indexação grande parte dos estudos na área da Ciência da Informação, preocupada com a compreensão do processo de tratamento temático da informação, tem procurado referências nos estudos linguísticos, sobretudo no campo da Semântica. O realismo e o mentalismo como paradigmas da linguagem influenciaram diversos estudos semânticos, destacamos aqui dois deles que compreendemos ter grande influência no campo da indexação: a semântica formal de inspiração fregeana e a gramática gerativista chomskeana.

Frege, filósofo matemático, herdeiro da tradição aristotélica, procurou resolver problemas da Lógica criando um sistema de representação formal. Seu trabalho objetiva o controle do significado nas equações matemáticas. A partir desta perspectiva, ele escreve **Sentido e Referência** (1976), um livro sobre a necessidade de controlar os sentidos na linguagem, removendo suas imperfeições. A lógica representacional da semântica de Frege prioriza as relações formais, ou seja, os aspectos objetivos da linguagem.

De acordo com Frege, os enunciados informativos, isto é, aqueles que não são tautologias, são sobre os mesmos objetos, mas apresentam ideias diferentes. Os objetos são designados pelas referências. Conseqüentemente, as referências são aspectos objetivos do significado. Pela influência aristotélica em sua obra, Frege preocupa-se com os referenciais atribuídos a partir da designação

das coisas no mundo. A existência dos objetos referenciais independe do seu observador. As referências de expressões nominais, na semântica formal de inspiração fregeana, apontam para valores de verdade correspondentes. Já os sentidos são modos de apresentação das referências, meios para encontrá-las. "É, pois, a busca da verdade, onde quer que seja, o que nos dirige do sentido para a referência" (FREGE, 1976, p. 69). São os estudos lógicos acerca dos aspectos objetivos do significado, isto é, das referências e de seus modos de apresentação (os sentidos) o que importa à semântica formal.

Em Frege (1976), as relações entre sentenças e seus elementos constitutivos poderiam ser melhor compreendidas por uma linguagem lógica. A notação valida sentenças a partir de critérios que estabelecem o que é verdadeiro ou falso, eliminando as dúvidas e os mal entendidos da linguagem natural. O sentido é objetivo porque "a humanidade possui um tesouro comum de pensamentos" (FREGE, 1976, p.65). Nesta perspectiva, é importante refrear a "euforia da linguagem" para que os sistemas de representação possam ser eficientemente comunicativos. A padronização sempre opera a partir de relações binomiais entre palavras, conceitos, termos, sentenças e expressões como: certo e errado, verdadeiro e falso, válido e inválido.

No sistema lógico, utilizado pela semântica formal, as palavras representam sentidos que são modos de apresentação dos objetos, são meio de encontrar a referência. Os sentidos são dados objetivos que devem ser comunicados. Frege busca métodos da lógica matemática para controlar os sentidos expressos pela linguagem. O sentido determina a referência, ou seja, o objeto nomeado e o método da lógica serviria para identificá-lo e eliminar os elementos secundários da linguagem.

Frege procurou desenvolver um sistema lógico unificado que permitisse reconhecer os diferentes modos com que se apresenta um objeto. Oliveira (2009) esclarece que a teoria fregeana apresenta duas contribuições aos estudos semânticos: a distinção entre sentido e referência e o conceito quantificador (OLIVEIRA, 2009, p. 20). A partir da Lógica, Frege concebe seu formalismo apropriando-se de quantificadores aristotélicos - universal \forall e existencial \exists - como mecanismos para solucionar as imperfeições da linguagem natural como as ambiguidades.

A semântica formal de inspiração fregeana influenciou diversos estudos que aproximaram a lógica da linguística como a filosofia analítica de Russell que procurou desenvolver uma linguagem lógica artificial. Esta teria como principal função corrigir e eliminar as imperfeições da linguagem natural, procurando a exatidão e a precisão dos conceitos. Tal como Frege, Russell (1986) via a linguagem natural como defeituosa, pois os termos seriam vagos e imprecisos, não representando adequadamente a "verdade" que se constata na realidade. Uma linguagem baseada na lógica formal poderia por meio de notação alcançar representações perfeitas.

Numa linguagem logicamente perfeita haverá uma palavra, e não mais, para cada objecto simples, e tudo aquilo que não é simples será expresso por uma combinação de palavras, -

por uma combinação, é claro, derivada das palavras das coisas simples que a constituem, ou seja, uma palavra para cada. Simples componente. Uma linguagem dessa natureza será completamente analítica e mostrará, num só relance, a estrutura lógica dos factos afirmados ou negados (RUSSELL, 1986, p. 197-198).

Russell procurou por meio da lógica construir uma linguagem ideal que servisse de instrumento privilegiado de representação ontológica do real. Sua preocupação central sempre estava voltada para o problema do significado como as ambiguidades e a falta de precisão. Tanto Frege, quanto Russell, sentiram-se levados a refletir sobre a linguagem natural para mostrar a sua insuficiência intrínseca em traduzir o pensamento. Partindo das premissas mentalistas, procuraram construir uma linguagem notacional para evitar outras possibilidades de compreensão dos termos, conceitos, palavras. O objetivo passou a ser o de criar condições para elaborar um cálculo exato expresso por meio de símbolos que seriam substitutos exatos do pensamento.

Frege foi o primeiro a aproximar a noção de conceito da noção de função matemática, pois ambos são estruturas indeterminadas com percursos de valor. Procurar aproximar os conceitos das funções para o filósofo seria chegar em conceitos exatos onde não haveria espaço para ambiguidades ou imprecisões.

Na perspectiva da semântica formal, que inclusive inspirou a primeira fase dos estudos de Wittgenstein sobre a linguagem, resultando no *Tractatus logico-philosophicus*, de 1921, a linguagem determina um conjunto de elementos – nomes, preposições, conceitos – que, combinados entre si e de uma maneira particular, têm uma significação própria, possuem vida autônoma. Tais elementos saem de si próprios para evocar outros objetos e outras relações que vão compor diversas situações que compõem o mundo. Há nos elementos linguísticos algumas propriedades que são inerentes e comuns a todos eles, entre elas está o fato de representarem algo. A semântica formal trabalha com a percepção mentalista de que há uma diferença entre o que a linguagem e o que é representado por ela.

O primeiro Wittgenstein compartilha desta compreensão e constrói a teoria da proposição, uma vez que a linguagem se caracteriza não por um conjunto de elementos materiais, mas pelo conjunto de funções que servem para representar. Se há uma diferença entre a representação e o representado, a proposição não é o fato em si, mas uma imagem do fato, representa uma maneira de enxergar o fato.

O caráter lógico da linguagem estaria em precisar melhor essas proposições, criar critérios de distinção que melhor delimitassem essas proposições, elementos que incluam relações simbólicas como espaço, tempo, circunstância. Na lógica semântica, é possível calcular valores de verdade das proposições, pois ela impõe uma condição de semelhança entre a representação e o representado a ser verificado no mundo. “Chamo de signo proposicional o signo pelo qual exprimimos o pensamento. E a proposição é o signo proposicional em sua relação projetiva com o

mundo” (WITTGENSTEIN, 1968, p. 62). Destarte, os fatos seriam combinações de objetos, expressos por proposições que sem seu apoio no mundo ficam desprovidas de sentido.

A base fixa da linguagem para a semântica formal é a própria realidade, os objetos possuem formas fixas e estáveis, a *substância*. A proposição seria a articulação lógica que pode representar os fatos. Essa representação pode ser realizada por meio de uma notação, ou seja, uma escrita conceitual ideográfica que colocassem em evidência as articulações lógicas possíveis.

Wittgenstein, em *Tractatus logico-philosophicus*, apresenta a linguagem como um conjunto de formas proposicionais possíveis que procura representar significativamente os fatos. Na compreensão de que “os limites da minha linguagem denotam os limites do meu mundo” (WITTGENSTEIN, 1968, p. 111), a linguagem lógica ideal possibilitaria a descrição de diversos mundos particulares. A forma lógica, fixa e exata, da linguagem permitiria a expressão do pensamento.

Em contrapartida, ao analisarmos os estudos cognitivos da linguagem, percebemos que a gramática gerativista de Chomsky (1968) apropria-se da compreensão aristotélica da existência de uma tríade universal entre o real, a mente e a própria linguagem.

Diferentemente da semântica formal de inspiração fregeana, as palavras não apenas descreveriam o real, mas representariam também as “afecções da alma”, algo no interior do próprio homem. Então, estes estudos cognitivos apresentam o conceito de representação distinto daquele fregeano: em vez de ser uma percepção individual que deveria ser desconsiderada pelos processos linguísticos em sua dimensão lógica nuclear, a representação seria um sistema de significados essenciais, universais e fixos, compartilhados tacitamente pelos indivíduos. As representações “são as mesmas para todos” (MARTINS, 2009, p. 467).

Neste cognitivismo filiado ao paradigma mentalista, o mundo possui estruturas autônomas, racionais e únicas que formam entidades, categorias de entidades definidas por propriedades essenciais. A mente reproduz essa ordem do mundo por relações objetivas entre entidades e categorias. O canal de comunicação entre o pensamento e o mundo é a linguagem.

Podemos associar a gramática gerativista ao cognitivismo uma vez que, segundo Chomsky (1968), o aparato combinatório e interpretativo da gramática consiste em um conjunto de princípios e regras, inconscientes, mas que podem ser analisados precisamente porque, de algum modo, eles são análogos a operações matemáticas.

De acordo com a teoria da Gramática Gerativa, há uma gramática comum a todas as línguas, a mente. A Gramática Gerativa fornece um esquema geral: regras e princípios universais governam o comportamento da linguagem e ajuda a gerá-la. “A diferença das línguas está no léxico” (CHOMSKY, 1968, p.46). Esta Gramática Universal, segundo Chomsky, é um componente

biológico humano. Se "a língua é a manifestação mais ampla da nossa vida mental" (CHIERCHIA, 2003, p.24), é necessário compreender as estruturas mentais para padronizá-las.

Se a maneira na qual todas as pessoas pensam é semelhante, se possuímos aparelhos cognitivos que funcionam da mesma forma, logo o que precisamos fazer é a correspondência necessária entre palavras e pensamentos, de acordo com as especificidades de cada língua.

Chomsky (1968) afirma que as estruturas linguísticas variam pouco de um país para outro. A maior diversidade está nas possibilidades de combinação entre as estruturas. Nesta concepção, nós possuímos um algoritmo inconsciente capaz de combinar estruturas finitas da linguagem, criando seqüências infinitas de palavras, sentenças, expressões e textos. A maneira com que amalgamamos os significados das palavras depende exclusivamente da maneira com que as combinamos sintaticamente.

Na perspectiva chomskyana, os seres humanos nascem equipados com a faculdade da linguagem, ou seja, estruturas de conhecimento e presunções acerca dela que se desenvolvem ao longo da vida. É da competência do indivíduo a capacidade de se articular gramaticalmente. É uma capacidade cognitiva particular do homem desenvolver formas de se expressar e se comunicar enquanto espécie.

A linguagem configura-se como uma propriedade dos seres cognoscentes, uma capacidade mental de pleno direito. Segundo o linguista, o indivíduo possui a capacidade de relacionar os padrões internos com as condições externas, extra-linguísticas, ou seja, saber articular a faculdade da linguagem com o mundo para construir ou interpretar mensagens, a esta aptidão chamou de competência gramatical. Essa propriedade é natural aos indivíduos a ponto de permanecer inacessível à consciência, as combinações são feitas automaticamente. É um conhecimento tácito que se desenvolve instintivamente de acordo com a demanda exigida pelas relações com os outros e com a realidade.

No exercício da linguagem, há um aspecto criativo quando emitimos e interpretamos as mensagens que se dá continuamente e com naturalidade, diversas frases, orações e textos podem ser produzidos de forma inédita. Existem, segundo Chomsky (1968), possibilidades infinitas de novas formulações gramaticais.

[...] o homem tem uma capacidade específica, um tipo de organização intelectual que não pode atribuir-se a órgãos exteriores nem relacionar-se com a inteligência geral e que se manifesta no que podemos denominar o 'aspecto criador' do uso da linguagem corrente, e cuja propriedade consiste em ser ilimitado em quanto alcance em não precisar de estímulo (CHOMSKY, 1968, p. 28).

Percebemos, desse modo, uma filiação do gerativismo chomskiano com inatismo descartiano, ao conceber que a linguagem existe para a livre expressão do pensamento, hábil o

suficiente para possibilitar a resposta adequada a qualquer nova situação que apareça. A adesão de Chomsky ao cartesianismo se manifesta em sua compreensão de que o falante possui conhecimentos inatos que se manifestam na prática ordinária da comunicação e que independem de qualquer estímulo externo.

No gerativismo chomskiano, o desenvolvimento da linguagem é um amadurecimento do sistema cognitivo, ou seja, um empreendimento pessoal. Neste aspecto retoma a tradição platônica-aristotélica de perceber o conhecimento como uma virtude a ser perseguida no mais íntimo do ser. “Os homens virtuosos não estão em condições de fazer os demais como eles, pois são tal como são por obra do conhecimento” (PLATÃO, **Menón**, 99a.). Chomsky substituiu a virtude por linguagem ao entender que o desenvolvimento desta se dá por um desenvolvimento pessoal e sem controle consciente, é uma capacitação progressiva.

Existe um pressuposto inatista na gramática gerativista chomskiana de que a forma geral do conhecimento está fixada aprioristicamente, uma predisposição da mente cuja função da experiência torna-se simplesmente possibilitar oportunidades para que esta estrutura esquemática seja realizada de modo preciso. Há uma correspondência com o mentalismo ao perceber a formulação do pensamento como algo essencialista, inata ao sujeito cognoscente, e ao entender a linguagem como expressão do pensamento materializada no mundo.

As pré-disposições mentais precisariam do estímulo da experiência para se concretizar. A experiência não modularia a mente dos sujeitos, mas se tornaria uma condição para a manifestação de sua reserva de potenciais pensamentos. Chomsky retoma Descartes para negar o experencialismo, ou seja, a mente não responde automaticamente aos condicionamentos do meio natural, pois é essencialmente livre.

A atualização chomskiana do inatismo descartiano implica na aceitação de que a mente possui uma configuração biológica determinante no uso da linguagem. “Certos aspectos do nosso conhecimento e compreensão são inatos, parte de nossa herança biológica, geneticamente determinada, como braços e pernas em vez de asas” (CHOMSKY, 1988, p. 14, tradução nossa).

Na gramática gerativista, a linguagem é um sistema de conhecimento interiorizado na mente do indivíduo, um sistema cognitivo altamente especializado e determinado biologicamente. A linguagem por ser uma faculdade mental, armazena, organiza e manipula dados, códigos e itens lexicais de maneira a construir significados compreensíveis para os demais. Uma compreensão mentalista de que a linguagem é o instinto da razão que caracteriza o homem como espécie.

Se a linguagem é uma faculdade mental, biologicamente determinada, inata aos seres cognoscentes, é possível uma gramática universal, pois o cérebro funciona de forma semelhante em todos os indivíduos. Para Chomsky, qualquer propriedade que se pode dizer é aplicável a todas as

línguas, pois há expectativas universais anteriores a qualquer experiência linguística. Há princípios universais por sermos da mesma espécie.

Esse sistema de princípios são padrões matriciais delimitadores que transformam todas as línguas humanas em compreensíveis e concebíveis, inclusive as que são geradas artificialmente, pois o que se modificam são as combinações. “As línguas do mundo parecem ser radicalmente distintas uma das outras em muitos aspectos, mas sabemos que devem ser baseadas no mesmo molde, que suas propriedades essenciais devem ser determinadas pelos princípios fixos da gramática universal” (CHOMSKY, 1988, p. 59).

Ao propor os aspectos de uma teoria da sintaxe, Chomsky os coloca como universais, pois são propriedades intrínsecas do sistema humano de aquisição da linguagem. Estas permitem um esquema que se aplica aos dados e que os determina em uma gramática. Esses traços substanciais que se apresentam de forma muito geral vão se apresentar em suas variáveis de modo restrito nas experiências de cada língua.

Logo, os princípios da gramática universal são constituintes da faculdade da linguagem, possuem parâmetros que se manifestarão em diferentes combinações na experiência de cada língua. “A gramática universal é um conjunto de princípios que caracteriza a classe das gramáticas possíveis. É uma teoria das gramáticas. Um conjunto de hipóteses empíricas relativas à faculdade da linguagem biologicamente determinadas” (CHOMSKY, 1999, p. 241).

3.4 AS CONTRIBUIÇÕES DOS ESTUDOS DE LINGUAGEM PARA A INDEXAÇÃO

Alguns teóricos da indexação como Vickery (1955) ao problematizar o processo apontaram para os problemas da linguagem natural como as ambiguidades e os vazios conceituais de alguns termos, da mesma maneira que a semântica formal os identificou. Preocupado com a construção de linguagens de indexação que controlassem estas imperfeições, Vickery propôs uma análise lógica que combinasse a classificação filosófica do conhecimento pelo método em árvore.

Então, as hierarquias entre classes, conceitos, assuntos seriam explícitas e as relações entre eles melhor evidenciadas. É uma perspectiva, tal como a semântica formal, de procurar criar uma linguagem artificial que pudesse controlar os sentidos, baseada na mesma compreensão de que os conceitos são universais e atemporais, pois representam a realidade do mundo, facilmente verificável nele próprio.

O controle semântico por meio da construção de linguagens artificiais é uma preocupação presente em diversas abordagens da indexação como vimos na segunda seção desta tese como os níveis integrativos debatidos pelo CRG na década de 1950 e abordados por Kyle (1964), procurando estabelecer uma ordenação que invertia a forma da teoria da classificação de *top-down*

para *bottom-up*. Não se trata de fragmentar uma classe genérica diversas vezes até chegar às suas propriedades.

A perspectiva dos níveis integrativos vai se aproximar da semântica formal ao compreender que os conceitos são formados pelo acúmulo de propriedades e características, suas articulações irão compor o assunto. Portanto, os conceitos seriam entidades formadas pela integração de suas partes.

A notação dessas linguagens deveria seguir a ordem do mais simples para o mais complexo. Tal como a semântica formal, a teoria dos níveis integrativos procura trabalhar com representações a partir de elementos simples que se articulam para a composição de um sentido, de um assunto, de um conceito, de uma categoria. Os níveis integrativos permitiriam diversas combinações por descreverem os elementos constituintes dos conceitos, estando, conseqüentemente, a representação mais próxima do que Wittgenstein, em sua primeira fase chamou de substância.

Farradane (1952), também preocupado com a questão semântica na abordagem do processo de indexação, criou os analetos (*analets*) procurando estabelecer relações lógicas entre os conceitos previamente identificados. A proposta era estabelecer a lógica do significado utilizando os métodos de notação inspirados na semântica formal. Novas representações conceituais seriam possíveis porque o sentido se forma pela combinação de propriedades formalmente descritas.

A teoria *Terminological Continuum*, de Frederick Jonker (1957) postula que a operação da indexação deve reconhecer os termos mais significantes de um documento para auxiliar a recuperação da informação. A partir de operações matemáticas, baseadas nos estudos da lógica formal, Jonker procura controlar as significações das palavras, tentando criar um vocabulário em que cada palavra tenha apenas um significado. A *terminological continuum* procura estabelecer leis para controlar essas variações na indexação, expressando a relação entre o tamanho do vocabulário e a especificidade com que um conceito pode ser descrito.

A Análise Documentária e a Lógica têm pouco a oferecer uma a outra, entretanto, é preciso mencionar que da Lógica a Análise Documentária pode apoiar-se na Lógica Formal porque, segundo Pinto Molina (1994, p.128), a Lógica Formal 'se limita a um sistema de símbolos, relacionando o exercício do raciocínio lógico a um cálculo algébrico' (SILVA; FUJITA, 2004, p. 153).

A linguagem representacional deste modelo matemático de indexação assemelha-se com a semântica formal proposta por Frege (1976), marcada pela influência do paradigma realista da linguagem. Tanto na teoria terminológica da indexação de Jonker, quanto na semântica formal de Frege, a linguagem é compreendida como representação da realidade. Dessa forma, o sentido seria o conteúdo expresso, no caso da indexação, expresso no documento que nos remete às suas referências. As referências estariam no âmbito do real, nas relações das descrições dos objetos com o mundo.

Jonker (1957) propõe a criação de uma linguagem de indexação, baseada na lógica, que permita o controle terminológico contínuo da linguagem natural. Este controle dos sentidos das palavras seria possível na indexação, pois como no mentalismo e na semântica formal, a *terminological continuum* parte do pressuposto que é possível criar uma terminologia universal, porque os domínios da mente humana são os mesmos, ou seja, possuímos repertórios de significados semelhantes, o mundo é o mesmo para todos e a mente dos indivíduos funciona de maneira semelhante.

Por conseguinte, é de grande importância conter as "euforias da linguagem" (FREGE, 1976, p. 68) para que os sistemas de representação sejam meios confiáveis de comunicação. Do contrário, a racionalidade do pensamento não conseguiria ser expressa pela linguagem e, conseqüentemente, não poderia ser registrada nos documentos, pois se perderia nas especificidades de cada língua.

Já o Cognitivismo na Ciência da Informação é uma abordagem associada às teorias da mente computacional e da inteligência artificial em analogias diretas entre os processamentos humanos da informação e da computação (INGWERSEN, 1992, p 19).

A grande ênfase desta visão encontra-se na percepção de que o conhecimento é ativamente construído pela matéria cognitiva, isto é, por combinações de padrões de pensamento da mente humana que são semelhantes a todos os indivíduos, pois biologicamente possuímos um processador igual - o cérebro. As padronizações da informação são possíveis porque seguem a mesma lógica da construção dos pensamentos. Uma compreensão de linguagem muito próxima à gramática gerativista de Chomsky.

Frohmann (1990) ao analisar a influência do paradigma mentalista na teoria da recuperação da informação, aponta o quanto as operações de análise e tradução na indexação são influenciadas pela concepção da linguagem enquanto representação do conteúdo mental. Segundo o autor, a indexação compreendida como duas operações distintas - a primeira envolve tanto a representação implícita, quanto explícita de um documento por uma declaração de assunto, ou seja, a declaração em linguagem livre de conceitos resultantes do conteúdo do documento; a segunda envolve a tradução ou representação dos conceitos em frases lexicais para um vocabulário controlado ou não, resguardadas as devidas relações semânticas e sintáticas com a linguagem documentária.

A visão cognitivista na indexação, criticada por autores como Frohmann (1990), Hjørland (2002) e Day (2005), que a chamam também de mentalismo, possui, segundo os mesmos, influências dos estudos da Psicologia e da Linguística. A apropriação em comum dos estudos cognitivos nessas três áreas é analisar seus objetos - a informação, o pensamento, a linguagem - a partir do pressuposto de que eles são construídos por meio de processamentos cognitivos que combinam a apreensão da realidade e os padrões da mente humana.

Percebemos que algumas teorias acerca da representação da informação assemelham-se com esta abordagem cognitivista. Autores como Fogl (1979), entre outros, percebem as palavras como correspondências de conceitos, representando o real e determinando um campo do conhecimento. A informação é compreendida como interpretação da forma material do conhecimento e possível de ser representada por uma linguagem porque o repertório cognitivo é semelhante a todos os seres humanos. Nessa abordagem mentalista que encontramos na representação da informação, seus esforços estão concentrados em descrever conteúdos informacionais, mapear leis e propriedades de objetos informativos, analisar a assimilação dos fenômenos informativos da realidade objetiva na mente humana.

Essa abordagem cognitivista sobre a informação, realizada por diversos autores como Shera e Egan (1961), Soergel (1999), Vickery (1973), entre outros, trata os enunciados informativos como aspectos objetivos do significado, produtos do conhecimento enquanto entidade mental. A representação seria um sistema de significados essenciais, universais e fixos, compartilhados tacitamente pelos indivíduos, as representações seriam as mesmas para todos.

Dessa maneira, o conhecimento se daria a partir de regras lógicas, de padrões preestabelecidos combinados entre si a partir de cada situação apresentada pela realidade. Nesta concepção instrumentalista, as palavras possuem funções, elas designam as coisas do mundo. Logo, caberia ao profissional da informação compreender, organizar e descrever tais combinações.

Tais abordagens partem do pressuposto que há objetividade no conhecimento e se propõem a entender a relação entre informação e usuário a partir das características e possibilidades da informação. Consideradas estruturalistas por Day tomam a linguagem como uma ferramenta capaz de criar representações precisas, modelos, “inventários sistemáticos de gramática para os elementos e suas possibilidades de combinações no registro da informação” (DAY, 2005, p. 576).

Algumas definições sobre indexação, mapeadas por Gil Leiva (2008), possuem grande proximidade com esse paradigma da linguagem como é o caso daquela estabelecida pela Unisist, “descrever e caracterizar um documento com a ajuda de representações dos conceitos contidos no dito documento para permitir uma busca eficaz das informações contidas num fundo documental” (1975, p. 2). Outra é a estabelecida por Borko e Bernier, “analisar o conteúdo informacional dos registros do conhecimento e expressar o conteúdo informacional na linguagem do sistema de indexação” (1978, p. 8). Além da definição estabelecida pela norma 5963/1985, criada pela *International Organization for Standardization (ISO)*¹⁰, “ação de descrever ou identificar um documento no que diz respeito ao seu conteúdo”.

¹⁰ Cf ISO (2018) esta versao foi confirma pela ISO em 2015. Disponível em: <<https://www.iso.org/standard/12158.html>>. Acesso em: 25 de janeiro de 2018.

Em todas essas definições há um pressuposto implícito de que os assuntos são extraídos do documento por ser o sentido uma entidade presente na informação, cabendo ao analista a tarefa de identificar tais atributos e descrevê-los/traduzi-los para uma linguagem de indexação.

Muitas técnicas de indexação partem dessa concepção cognitivista para construir suas políticas e linguagens artificiais. A exemplo de Esteban Navarro (1999) que compreende a indexação como um processo destinado a identificar e descrever ou caracterizar o conteúdo informativo de um documento, mediante a seleção das matérias sobre as quais versa (indexação sintética) ou dos conceitos presentes (indexação analítica), para sua expressão em língua natural e sua reunião em índice. Seu objetivo é permitir a posterior recuperação dos documentos pertencentes a um acervo documental ou a um conjunto de referências documentais, como resposta a uma demanda acerca do tipo de informação que este contém.

Percebemos que essas maneiras de aplicação da indexação largamente difundidas tratam os sentidos como dados objetivos possíveis de serem retirados do texto de forma sistemática, naturalizando a produção de sentido feita na análise conceitual e representação do documento.

Lara (2007), entre outros autores, aponta que princípios de categorização no campo da Ciência da Informação historicamente estiveram associados à semântica cognitiva de inspiração mentalista a exemplo de teóricos como Ranganathan, Dahlberg e nos trabalhos do *Classification Research Group*. A autora aponta que as noções de categoria são baseadas em princípios lógicos e universais. Essa influência do cognitivismo acaba por reproduzir a concepção do homem como uma máquina processadora que se utiliza da classificação, da categorização do mundo para melhor entendê-lo, decifrá-lo e descrevê-lo.

Richard Moss na década de 1960 ao analisar as categorias e relações na construção de esquemas de classificação mostrou a influência das categorias aristotélicas na obra de Ranganathan e a apropriação feita da abordagem de Russel pelo campo da Ciência da Informação no desenvolvimento da classificação e da indexação para a representação dos documentos.

Algumas categorias de Aristóteles e Ranganathan são idênticas. Então, Lugar e Tempo em Aristóteles são as mesmas que Espaço e Tempo em Ranganathan – uma era que pensa em termos de Lugar quando ligado à superfície da terra é substituída por pensar em termos de Espaço na era do Sputnik. A posição de Aristóteles é supérflua – uma duplicação do Espaço. Ação agora possui uma expressão mais moderna que é a Energia. Substância, no sentido aristotélico, passa a ser usada como Matéria. Então, como apontou Bertrand Russel, em uma história não inteiramente imparcial, essas categorias de Aristóteles também podem ser consideradas para outras como Vida e Personalidade (MOSS, 1964, p. 296, tradução nossa)

A teoria da classificação de inspiração platônica-aristotélica, que compartilha de seu essencialismo em sua abordagem clássica, define tal operação como “um processo mental pelo qual coisas, seres ou pensamentos são reunidos segundo as semelhanças ou diferenças que apresentam”

(BARBOSA, 1969, p.13). Portanto, uma atividade humana inata. Essa definição aparece em diversos autores no campo da Ciência da Informação como Vickery (1973; 1980); Foskett (1970; 1973); Langridge (1973); Piedade (1977).

De acordo com Piedade (1977), Ranganathan se apropria dos postulados aristotélicos sobre as dez categorias que compõem o ser para estabelecer a teoria facetada da classificação. Em sua obra, *Prolegomena to Library Classification* (1957), Ranganathan utiliza os mesmos conceitos de Aristóteles para definir característica – qualquer atributo ou conjunto de atributos cujo vínculo ou falta de vínculo de entidades podem ser determinados e ao menos dois deles são diferentes; classe – um grupo de elementos reunidos por suas características ou atributos; e ordem – sequência de classes derivada da organização baseada em uma única característica e classificadas de acordo com os seus arranjos.

Essas definições serviram para estabelecer a base da construção dos seus cânones gerais para a classificação dos documentos, dentre os quais destacamos o da diferenciação, da relevância, da permanência, da consistência, da exaustividade, e da exclusividade. Que estabelecem respectivamente que: cada característica usada deve diferenciar, isto é, deve dar origem a pelo menos duas classes; cada característica deve ser relevante para o propósito da classificação; cada característica deve continuar a ser assertiva e inalterada, desde que não haja mudança no propósito da classificação; a sequência de aplicação das características escolhidas deve ser consistentemente obedecida; as classes em qualquer ordem de classes devem ser totalmente exaustivas de seu universo comum imediato; as classes em uma ordem de classes devem ser mutuamente exclusivas.

La Barre (2010) ao analisar a teoria facetada de Ranganathan aponta que é fundamental o estabelecimento de categorias compostas por conceitos, termos e assuntos que determinarão as características, atributos ou funções que servirão de base para a análise de um conjunto de facetas. Essa base da teoria facetada é originada da lógica aristotélica.

Para Aristóteles, existem dez categorias que elencam as propriedades de qualquer objeto no mundo, são o que ele chamou de predicados, mas que podemos associar a ideia de característica do conceito em Dahlberg (1978), ou das facetas em Ranganathan (1987). De acordo com o filósofo grego, qualquer coisa que existe no mundo pode ser classificada por seus predicados (categorias) identificadas pela lógica como naturais ao ser. Aristóteles então elenca substância, quantidade, qualidade, relação, lugar, tempo, situação, posse, ação e paixão. Ranganathan compartilha desse essencialismo aristotélico para estabelecer as facetas: personalidade, matéria, energia, espaço e tempo.

Segundo Ranganathan, analisar um assunto por facetas significa que cada aspecto desse assunto pode ser visto como as manifestações de certas características ou facetas que

obedecem a postulados pré-determinados. O sistema torna-se, assim, multidimensional e ilimitado (BARBOSA, 1969, p. 166).

A base do paradigma mentalista de compreender o pensamento como a formulação de combinação desses padrões, ora chamado por Aristóteles de predicados, por Chomsky de sintagmas, por Ranganathan de facetas e por Dahlberg de características, está presente em todas essas teorias. Uma vez identificados os padrões, as propriedades dos conceitos, todas as suas combinações podem ser mapeadas e classificadas, ainda que tendam ao infinito.

Encontramos a influência da filosofia platônica-aristotélica de linguagem também na definição de conceito estabelecida por Dahlberg. De acordo com a autora, os conceitos "identificam as características, os atributos das coisas designadas" (DAHLBERG, 1978, p. 102). Segundo esta teoria, o processo de análise dos conceitos somente é possível "coletando-se os enunciados verdadeiros que sobre determinado objeto se podem formular" (DAHLBERG, 1978, p. 102).

Influenciada pelos pressupostos platônicos-aristotélicos, Dahlberg indica que nos conceitos há elementos que constituem enunciados verdadeiros sobre as coisas, passíveis de serem identificados e classificados de acordo com suas características naturais. Nesta concepção, existe uma autonomia simbólica dos enunciados que compõem os textos. Os elementos, as características e os atributos vão compor os conceitos, a tessitura bem sucedida dessas combinações feitas pelos indivíduos é que vai indicar se a representação do real será bem ou mal sucedida. A eficácia dos conceitos está relacionada diretamente à fidelidade de sua descrição do real pelo homem.

No paradigma mentalista, há a necessidade de estabelecer nas expressões, nos discursos, nos textos, nos conceitos uma objetividade essencialista que está no nexo em que estes elementos estabelecem ou não com o real. Dahlberg inicia seu artigo apresentando a linguagem como instrumento do homem para comunicar o pensamento. "Desde que o homem foi capaz de pensar e de falar, empregou palavras (conjunto de símbolos) para designar os objetos de sua circunstância e para traduzir os pensamentos formulados sobre os mesmos (1978, p. 101).

Numa perspectiva muito próxima ao gerativismo chomskeano, Dahlberg compreende que os conceitos são formulados no processo de cognição no qual o homem identifica propriedades, características das coisas, dos objetos no mundo e as processa, realizando possibilidades infinitas de combinações entre elas. Para Dahlberg existem dois tipos de características: aquelas que são constitutivas, essenciais às coisas em si e as acidentais, aquelas que podem ser adicionadas, que são possíveis de inserção. A combinação entre elas formam os conceitos. Neste mesmo artigo, ela apresenta as relações lógicas possíveis entre as características que podem compor conceitos. A teoria do conceito de Dahlberg apresenta uma percepção da linguagem próxima à gramática gerativista chomskeana na qual a linguagem possui meios finitos para possibilidades infinitas de expressão.

Chomsky vai dedicar-se ao estudo dos sintagmas em diferentes línguas para a formulação de conceitos, frases, orações; Dahlberg vai teorizar sobre as características do conceito, suas funções e relações. Embora os sintagmas tenham sido criado por Saussure em 1969 para melhor compreender como a linguagem se estrutura na sociedade, Chomsky se apropriou do termo utilizando como partes constituintes da língua que podem gerar inúmeras combinações de palavras, frases, orações.

Desse modo, para Chomsky, os sintagmas são termos relativos à boa formação da mensagem. Já Dahlberg irá se preocupar com a ordem das características para a formulação dos conceitos. Ambos partilham da compreensão de que a linguagem expressa um sistema de representações mentais que possuem padrões, propriedades, características, sintagmas finitos, mas que podem ser combinados de formas incalculáveis e que desta maneira conseguem decifrar o mundo e descrevê-lo.

Jerry Fodor ao procurar explicar a gramática do mentalismo, teoria de que é partidário como Chomsky e Dahlberg, diz que é uma forma de explicar como os objetos mentais - os pensamentos e conceitos - podem providenciar domínios para os processos mentais como o raciocínio. Destarte, os processos mentais são espécies de computações, ou seja, operações formais de decodificação do mundo para sua compreensão e retorno por meio de representações estruturadas sintaticamente (FODOR, 2007).

Nessa dinâmica a representação é um tipo de relação mundo-mente constituída pela interação entre os pensamentos. Conseqüentemente, a função do sentido é explicar como essas representações são determinadas pela sintaxe, ou seja, essas infinitas combinações entre os padrões finitos da mente. “Os processos mentais são operações sintáticas definidas sobre representações mentais” (MARTINS, 2009, p. 465), ou seja, são manipulações formais de símbolos, como se o pensamento fosse algorítmico, uma máquina que processa a linguagem.

É importante ressaltar que o paradigma da mente computacional não está eliminando a subjetividade ou a sensibilidade estética da lógica do pensamento, mas centraliza tal lógica em uma estrutura computacional. Neste paradigma, pensar é essencialmente computar. A mente, dessa forma, é modular, e o nosso cérebro é dotado de uma organização extremamente sofisticada não passível de explicação pela exterioridade da sociedade. As línguas são em sua essência, ou seja, sistemas de representação da vida mental. O pensamento é entendido como autônomo em relação às línguas faladas no mundo, uma vez que o próprio pensamento funciona como uma linguagem interna da mente.

O realismo e o mentalismo de filiação platônica-aristotélica por partilharem a concepção de que existe uma essência apriorística nas coisas que compõem o mundo, que o homem compartilha de representações mentais por possuir um aparelho biológico semelhante e que este reproduz a ordem do mundo, assumem que conceitos, categorias e a produção de sentido podem ser universais

e atemporais, pois a linguagem é um sistema nuclearmente literal e declarativo do pensamento e, por conseguinte, é pouco afetado por fatores sócio-culturais.

Embora as teorias da classificação e do conceito não sejam objetos desta tese, elas são importantes para compreender o próprio estado da arte da indexação, uma vez que foram apropriadas para o tratamento temático dos documentos. Estas teorias estão presentes nas técnicas de indexação que estabelecem o *aboutness* unicamente pela análise dos documentos.

O paradigma mentalista na indexação mostrou grande êxito, sobretudo com a introdução dos computadores nos sistemas de recuperação da informação. A indexação automática passou a ser desenvolvida a partir do estabelecimento de padrões de entrada e as possibilidades de combinação entre tais padrões de entrada. Os princípios que são programados para o estabelecimento dos assuntos nos documentos têm variado desde 1950, no entanto, a lógica é a mesma: o design com os módulos conceituais é arquitetado e as possíveis combinações entre eles são computadas pelo sistema.

Um exemplo bem sucedido de indexação automática é o *PRECIS (Preserved Context Index System)* que procurou estabelecer padrões operativos a partir de estruturas sintáticas e semânticas da língua para a construção de índices e de linguagens documentárias. Desenvolvido por Derek Austin, em 1971, para o processamento automático das entradas e cabeçalhos do índice de assunto da British National Bibliography (BNB), este sistema, analisado por Fujita (1989), procura estabelecer os padrões de entrada a partir da sintaxe da língua à qual a base de dados pertence. Esses padrões são combinados a partir de funções pré-estabelecidas que irão reproduzir as funções sintáticas nos operadores. O computador realizará a correspondência dos termos automaticamente a partir da pergunta feita pelo usuário do sistema.

Os operadores de função atribuídos aos termos extraídos do documento, formarão uma *cadeia* organizada em seqüência lógica sob os preceitos da gramática do PRECIS, assegurando a uniformidade de organização para diferentes indexadores. Seus valores numéricos crescentes indicam a ordem dos termos nas entradas, além de determinar a tipografia, a pontuação e o formato das entradas (FUJITA, 1989, p. 123)

Fujita (1989) ao analisar o *PRECIS* afirma que o sucesso do sistema na língua inglesa abriu precedentes para sua aplicação em outras como a francesa, alemã, a dinamarquesa e a polonesa. Percebemos que na pretensão de universalidade desse sistema há a compreensão de que o conteúdo do documento é produto do conhecimento e do pensamento do seu autor, é possível representá-los em qualquer linguagem porque esta é concebida como instrumento de representação das ideias.

3.5 AS APROPRIAÇÕES CONTEMPORÂNEAS DO REALISMO E MENTALISMO: UMA ANÁLISE A PARTIR DO *JOURNAL OF DOCUMENTATION* E O *KNOWLEDGE ORGANIZATION*

Procuramos compreender como a influência da tradição platônica-aristotélica e, por conseguinte, e os estudos da linguagem apresentados nesta seção estão influenciado as abordagens da indexação na contemporaneidade. Relembrando a metodologia aludida na introdução deste trabalho, realizamos um levantamento bibliográfico na base de dados *LISA*, considerando as produções entre os anos de 2006 e 2016. Devido ao alto número de artigos recuperados e a necessidade de trabalhar com dados qualitativos, optamos por direcionar a nossa análise aos periódicos *Journal of Documentation* e *Knowledge Organization*.

A metodologia de análise é apresentada na quinta seção da tese. No entanto, os resultados relacionados às abordagens da indexação à luz dos paradigmas realista e mentalista são aqui explicitados.

No levantamento realizado no *Journal of Documentation* acerca da influência da tradição platônica-aristotélica nos estudos da indexação entre (2006-2016), identificamos 36 artigos sobre indexação automática como o trabalho de Tudhope *et al* (2006), *Query expansion via conceptual distance in thesaurus indexed collections*, o estudo de Zhang e Zeng (2014), *A new similarity measure for subject hierarchical structures; Classifications and concepts: towards an elementary theory of knowledge interaction*, de Smiraglia e Heuvel (2013); *Untangling search task complexity and difficulty in the context of interactive information retrieval studies*, de Wildemuth, Freund, Toms (2014) são alguns exemplos da influência do paradigma realista na abordagem acerca da indexação e da recuperação da informação.

Esses artigos possuem em comum nas suas abordagens uma concepção de linguagem semelhante à semântica formal. Tais estudos possuem um princípio em comum que é procurar estabelecer melhores maneiras - seja na entrada dos documentos, seja na sua saída dos sistemas - para que o computador realize a melhor conexão entre os documentos armazenados pelo sistema e a pergunta do pesquisador.

Nestes, a ênfase no controle terminológico é uma condição primordial. A maioria dos artigos que versam sobre indexação automática nos últimos dez anos compartilham o pressuposto de que é necessário a definição de técnicas que possibilitem a modificação do vocabulário com a finalidade de melhorar a eficácia da recuperação do documento. Estes termos são atributos escolhidos para representar o conteúdo informacional do documento.

Outrossim, a construção de modelizações, leis, padronizações, para o controle de termos, conceitos e palavras pela indexação automática filia-se à semântica formal ao procurar traduzir a linguagem para uma representação lógica, tentando evitar o que Frege chamava de imperfeições da

linguagem. Problemas também identificados na indexação como a sinonímia, a polissemia e os sentidos fora do contexto eram questões que a semântica formal procurava eliminar a partir do controle da linguagem pela notação lógica e o cálculo de predicados.

Já os estudos mais recentes sobre indexação automática que propõem a utilização de métodos como a padronização por sintagmas como o artigo de Ibekwe-Sanjuan (2006), *Constructing and maintaining knowledge organization tools*, de Rafferty (2011), *Knowledge Representation in the Social Semantic Web* são marcados pela influência da gramática gerativista chomskyana.

A Gramática Gerativa tem sido amplamente utilizada como suporte teórico dos sistemas de indexação automática. Algumas técnicas como a indexação por extração automática, a indexação por atribuição semântica, a identificação de palavras in full context e a indexação automática sintácticabuscam estabelecer padrões de conceitos, de relações hierárquicas entre as palavras, termos, sintagmas. As leis, as fórmulas, os princípios de combinação são baseados na filosofia mentalista na qual a linguagem representa os conceitos universais da mente humana.

Esta abordagem da indexação a partir da gramática gerativista e, por conseguinte, do paradigma mentalista, facilitou a interface entre o indexador e o computador, possibilitando o aparecimento de diversas bases de dados com graus elevados de precisão e revocação na recuperação dos documentos.

Já no levantamento realizado no *Knowledge Organization*, considerando o mesmo período, percebemos que as publicações se fundamentam no conceito de organização e representação do conhecimento, diferentemente do *Journal of Documentation* que privilegia o tratamento temático da informação e o comportamento de uso. Conseqüentemente, operações como a indexação estão inseridas nas discussões, teorias e problematizações da própria área da Organização do Conhecimento.

Percebemos que a tendência majoritária dos artigos publicados nos últimos dez anos no *Knowledge Organization* apresenta esta maneira semelhante de abordar a indexação e a organização do conhecimento.

O trabalho, *Semantic relations in knowledge organization systems*, de Brascher (2014,) faz alusão ao trabalho de Dahlberg sobre a teoria do conceito na construção da categoria *organização do conhecimento*, enfatiza a necessidade do mapeamento das estruturas e funções semânticas e sintáticas dos KOS a fim de facilitar a identificação e, por conseguinte, a representação dos conceitos para facilitar a padronização da linguagem artificial dos KOS e a recuperação da informação.

Tal artigo, aponta para caminhos semelhantes à gramática gerativista de Chomsky (1968) ao compreender os conceitos como unidades universais e representativas do pensamento do indivíduo.

Ao associar os modelos cognitivos às unidades conceituais, a autora deixa transparecer a influência da teoria gerativista na qual a linguagem descreve/representa sistemas de conceitos elaborados e organizados pelo pensamento, produto da cognição característica comum a todos os indivíduos.

Outros trabalhos como *Modes of growth of subjects*, de Satija; Madalli e Dutta, de 2014, que aborda o crescimento dos assuntos e busca métodos universais para o controle do significado, influenciados também pela teoria de Ranganathan, aproximam-se igualmente do paradigma realista acerca da linguagem. Esta pesquisa afirma ser possível criar classificações e atualizá-las uma vez que os conceitos são entidades universais. Possuem existência autônoma, uma vez que qualificam e designam as coisas do mundo e, portanto, possuem substâncias próprias.

Encontramos também a influência da teoria gerativista em artigos publicados pelo **KO** na abordagem da indexação tais como o *Using a semantic analysis tool to generate subject access points*, de Zeng; Grace e Zumer (2014); *Method for selecting specialized terms from a general language*, de Anguiano Peña e Naumis Peña (2015); *A survey of the coverage and methodologies of schemes and vocabularies use to describe information resources*, de Hider (2015); além do artigo *Terminology web services*, de Binding e Tudhope (2010).

A maioria das publicações analisadas vai tratar de questões voltadas à organização e representação do conhecimento a partir da compreensão de que existem conceitos universais que expressam o pensamento tal como vimos em Dahlberg (1978).

Sendo assim, os artigos vão refletir sobre formas de classificar, enquadrar, estabelecer hierarquias e nivelamentos do conhecimento. Há uma predominância dos estudos sobre ontologias - encontramos 56 artigos abordando esta temática dentre eles *Formal Ontology and the Foundation of Knowledge Organization*, de Herre (2013), *Organizing Acts and Objects*, de Budd (2014) e *Boundary Objects*, de Marchese e Smiraglia (2013).

Ademais, as linguagens documentárias e taxonomias também se destacam como temas de muitos artigos no **KO**. São 30 artigos abordando a questão da taxonomia na organização do conhecimento como *Towards a Taxonomy of KOS*, de Souza, Tudhope e Almeida (2012), *Concepts and Terms in the Faceted Classification*, de Broughton (2010), *Naming and Reframing*, de Gross (2015). Todos os artigos citados possuem em comum a filosofia de que o conhecimento é uma entidade autônoma com estruturas passíveis de serem classificadas, codificadas e organizadas de maneira a promover o seu acesso. Esses métodos possuem uma grande proximidade do paradigma mentalista.

Os SOCs, assim como a mente humana, possuem padrões limitados, inclusive de estruturas linguísticas tal como os sintagmas, no entanto podem ser utilizados de infinitas maneiras, de acordo com as combinações feitas de forma criativa. Os algoritmos tão utilizados na Matemática e na Ciência da Computação são exemplos dessas possibilidades, aliás a arquitetura dos sistemas de

recuperação da informação passou a incorporar algoritmos cujos padrões são baseados em sintagmas.

A sequência finita de regras, instruções bem definidas e não ambíguas, cada uma das quais devendo ser executadas mecânica ou eletronicamente em um intervalo de tempo finito e com uma quantidade de esforço finita que podem gerar infinitas combinações. Tal como a mente é capaz de realizar computações sobre os pensamentos, os computadores são capazes de realizar indexações que permitam a recuperação de documentos úteis ao usuário, eliminando o ruído e o silêncio, problemas da linguagem natural “imperfeita”.

O paradigma mentalista nos estudos de indexação também está presente não apenas nas abordagens próximas à mente computacional, mas também no experencialismo. Há sensíveis diferenças entre ambas as abordagens, no entanto, é possível classificá-las como filiadas à tradição aristotélica por compreenderem a linguagem como um instrumento que descreve, informa, representa o pensamento do indivíduo acerca do mundo. No entanto, diferentemente da semântica formal e do gerativismo, o experencialismo compreende a linguagem como resultado da interação do indivíduo com o seu meio.

A linguagem é resultado de diversos fatores relacionados entre si: biológico, histórico, cultural, sociológico, psicológico, emocional, entre outros. É um canal de interação entre os indivíduos cujo significado depende do contexto em que está inserido. O mundo não é composto de estruturas autônomas e racionais, com propriedades e características fixas no tempo e no espaço, mas um lugar em permanente transformação por seus agentes (físicos, químicos, biológicos, etc).

Ela é uma faculdade biológica, no entanto, altamente afetada pelas injunções sociais, históricas e culturais. O conhecimento linguístico não pode se separar da cognição geral e a linguagem é um sinal de como funciona nossa forma de produzir conhecimento. As proposições do experencialismo são que a realidade não é autônoma de quem a vivencia; que o pensamento não é abstrato, mas que está vinculado ao corpo que sente os efeitos do mundo, produto de uma mente formada pelo consciente e inconsciente, por isso metafórico e imaginativo.

Os estudos mentalistas que adotam um paradigma experencialista compreendem que o sentido de um texto ou de qualquer expressão verbal é construído a partir de estruturas e esquemas cognitivos tomados como gestalts (LAKOFF; JOHNSON, 2003, p. 70). Destarte, uma tentativa de tratamento objetivista dos processos cognitivos seria improdutiva, já que um entendimento de seu mecanismo teria de considerar a variedade e a multiplicidade de fenômenos experimentais cuja abrangência seria cada vez mais restritiva, até ser reduzida ao individual.

A proposta de caráter experencialista não deve ser vista como elogio ao subjetivismo, pois além deste ser também considerado, nessa linha de pensamento, pernicioso (por tratar o que vai externo à mente como uma espécie de projeção desta), não despreza completamente os modelos

objetivistas de estudo. Lakoff e Johnson (2003, p.223), por exemplo, veem neles uma contribuição valiosa para o entendimento das capacidades intelectuais humanas.

A base da experiência permite a criação de esquemas imagéticos estruturados que servem para organizar a realidade (continente/conteúdo, parte/todo, ligação, percurso, centro/periferia, em cima/embaixo, frente/trás, etc.). A metáfora também é um elemento essencial na categorização da realidade e nos processos cognitivos: mediante o símil, projetamos conceitos concretos em outros mais abstratos (usamos, por exemplo, o espaço para falar do tempo, ou uma planta para falar de conhecimento).

Na Ciência da Informação, o experencialismo também é chamado de construtivismo social. De acordo com Talja, Tuominen e Savolainen (2005), o construtivismo social é uma posição metateórica que parte da compreensão de que a mente constrói a realidade em sua relação com o mundo, esse processo mental é significativamente (in)formado por influências recebidas de convenções sociais, históricas e interação com outros sentidos.

O trabalho de Ingwersen (1982; 1992) que procurou desenvolver um modelo de recuperação da informação incorporando o ambiente sócio-organizacional e o comportamento do usuário na busca pela informação é um exemplo de aplicação baseada na abordagem experencialista.

O construtivismo social é a base filosófica dos estudos voltados aos comportamentos de busca e uso da informação nos sistemas de recuperação da informação para orientar a indexação. Além disso, o construtivismo tem sido especialmente utilizado como teoria que fundamenta os estudos de desenvolvimento de sistemas de gerenciamento de informações pessoais e personalizados como as bibliotecas digitais (BERGMAN et al, 2003; KUHLTHAU; TAMA, 2001 apud TALJA; TUOMINEN; SAVOLAINEN, p. 85,2005).

No experencialismo ou construtivismo social, o conhecimento é incorporado pela ação. Ele funciona como um determinante interno para outras ações, que por sua vez modifica o conhecimento interno do indivíduo. Desta forma, o indivíduo como ator constrói o conhecimento interno de fatos, valores e procedimentos através da interação constante entre o seu conhecimento internalizado e sua participação no mundo externo. O conhecimento é tanto explícito, na medida em que pode ser comunicado através de linguagem e implícito ou tácito por poder ser encaixado dentro atividades particulares (JACOB; SHAW, 1998).

Nas abordagens da indexação pelo viés experencialista, os processos de informação devem ser inseridos em contextos sociais, organizacionais e profissionais, a partir dessa compreensão é que o processo vai se deslocar do documento para o usuário. Eles deslocam a atenção em estruturas de conhecimento individuais para "os produtores de conhecimento, o compartilhamento de conhecimentos e as comunidades" no que tange a absorção de conhecimento (JACOB; SHAW, 1998, p. 142). Mai (2005, p. 600) ao analisar as abordagens acerca da indexação, apresenta a

orientada no usuário com características semelhantes ao experientialismo. De acordo com o autor, a ideia básica dessa forma de indexar é que o indexador precisa compreender as necessidades e as terminologias das informações dos usuários no momento de atribuir o assunto e os termos do índice para o documento.

Outra questão levantada por Farrow (1991) que corrobora a percepção experientialista na indexação é a leitura do indexador do documento. O assunto não existe no documento, pronto para ser descoberto e retirado pelo indexador. O indexador necessariamente inclui conhecimento ou dúvida sobre o uso potencial dos documentos e as necessidades dos usuários em sua análise do documento. A decisão final do indexador interagindo com o seu contexto irá impactar substancialmente na escolha dos assuntos e dos termos a serem indexados e, por conseguinte, na recuperação da informação.

A segunda seção dessa tese apresentou também as contribuições de Beghtol (1986) para a teoria cognitiva da indexação que, influenciada pelo trabalho de Lee (1976), procurou consolidar a garantia de uso (*user warrant*), sobretudo na perspectiva da garantia cultural, como um dos princípios para a realização da indexação. Para a autora, era necessário não somente considerar a garantia literária para a produção de vocabulários controlados e linguagens documentárias, instrumentos da indexação. Não bastava considerar a literatura de determinada área do conhecimento e sua terminologia para extrair e validar os termos que seriam incluídos em um sistema de recuperação da informação.

Na definição de Beghtol, a garantia diz respeito mais do que apenas a seleção de termos, mas sim ao mapeamento de um sistema de classificação para o contexto e usos a que será colocado. Beghtol (2002), trata do problema de sistemas de classificação culturalmente orientados e mostra as barreiras ao acesso à informação por causa do que é entendido. O que pode ser fácil acesso por pessoas de uma cultura pode não ser assim para outras. É necessário integrar os conhecimentos culturais, geográficos e linguísticos, e a recuperação multicultural é necessária. Beghtol aproxima-se do experientialismo dos estudos da linguagem ao prever uma necessidade de examinar as relações entre os conceitos lexicais e sua implementação em várias culturas.

Percebe-se a influência do mentalismo aristotélico nesta corrente experientialista de Beghtol, quando a autora mapeia oito conjuntos de conhecimentos universais. No entanto, diferentemente da gramática gerativista e da semântica formal, nesta perspectiva o tempo, o lugar, a cultura influenciam a forma dos indivíduos lidarem com as informações. Aqui, tempo e lugar são usados para examinar as complexidades e especificidades da cultura no processo de indexação dos documentos.

Segundo Beghtol (1986), para tal compreensão[dos conceitos de um documento], é necessário: conhecimento do assunto primário e estrutura do conhecimento; o leitor

indexador deve conhecer o assunto e possuir conhecimento extratextual tomando por base as instruções de uso do sistema linguístico (SILVA; FUJITA, 2004, p. 147).

De acordo com Barité (2007, p. 4, tradução nossa), “A garantia de uso (*user warrant*) se sustenta na ideia de que o processo de seleção dos termos deve ser associado às formulações de busca que os usuários realizam ante um sistema de informação, e à maneira em que elas se manifestam.” Já a garantia cultural trata-se da compreensão de os indivíduos em culturas diferentes possuem necessidades de informação diferentes e interpretações distintas do conhecimento estabelecido.

Beghtol exemplifica com o calendário que é um artefato cultural não universal, e que ajuda a criar os diferentes aspectos sintáticos (estruturais) e semânticos (de significado) do pensamento e da mentalidade de uma sociedade. Assinala que ‘o uso de um calendário particular é uma eleição significativa de uma cultura e a reforma de um calendário é vista como uma intromissão dentro de tradições culturais estabelecidas porque um novo calendário muda a cultura.’ Agrega que apesar da especificidade cultural de calendários diferentes, a globalização da informação eletrônica tem revelado que a habilidade para traduzir um calendário a outro é importante em um número de áreas (por exemplo, banca, negócios, informática, investigação histórica) (BARITÉ, 2007, p. 5, tradução nossa).

Na abordagem experencialista da indexação, a leitura documentária ganha grande ênfase no processo de seleção e validação dos termos para a representação dos documentos. O leitor do documento, seja o indexador, seja usuário, torna-se protagonista no processo de indexação.

Pinto Molina (1994), ao abordar o processo cognitivo da indexação estabelece como estruturas básicas necessárias: o texto, o contexto, os dados do conhecimento do analista, objetivos documentários e um método de avaliação (PINTO MOLINA apud SILVA; FUJITA, 2004, p. 153). Giasson (1993, p. 21) afirma que o processo de leitura é uma interação entre o texto, leitor e contexto. Fujita atribui aos estudos cognitivos o enfoque na leitura e na participação do leitor no processo de compreensão da informação. “A atitude do leitor frente ao texto, anteriormente vista como recepção passiva de mensagens, passou a considerar o processamento mental de informação da compreensão e evoluiu para uma perspectiva de interação entre o leitor e o texto” (FUJITA, 2004, p. 3). Para a autora, a análise do assunto no processo de indexação deve considerar as variáveis apresentadas por Giasson e não somente as possíveis combinações entre o leitor e o indexador.

Neste sentido, o autor como emissor e o texto como mensagem transmitem uma suposição de objetividade que leva o leitor interagir com o texto. Para Fujita (2004), é essa relação triádica que permite a transmissão da informação como algo significativamente coerente. É importante considerar o conhecimento de mundo, linguístico e das estruturas sintáticas que o leitor possui para a análise do assunto. Fujita alerta que o próprio texto é produzido em um contexto específico que influencia também na leitura. Segundo a autora, todos esses aspectos precisam ser considerados na

leitura profissional realizada pelo indexador. A leitura profissional se difere da leitura comum, pois possui um objetivo principal, metas a serem alcançadas, métodos específicos e limitação de tempo.

O leitor profissional tem objetivos definidos para a leitura, conhecimentos de procedimentos de análise, de estratégias de análise e, sobretudo, de políticas de indexação e de demanda do usuário do sistema de informação. É preciso, então, que o indexador agregue ao seu conhecimento prévio, conhecimentos específicos para a atividade de leitura documentária em análise de assunto e isso precisa ser melhor estudado (FUJITA, 2004, p. 9)

Fujita ainda aponta que um dos principais problemas diagnosticados pelos indexadores está relacionado à tradução dos conceitos para uma linguagem utilizada pelo sistema de recuperação da informação. A autora atribui tal dificuldade a falta de conhecimento pleno da estrutura e da funcionalidade da linguagem por parte dos indexadores. Esta percepção filia-se aos estudos cognitivos de viés experiencialista, uma vez que compreende a linguagem como uma faculdade mental, com estruturas definidas, que se moldam a partir da interação dos indivíduos com a sua cultura, meio ambiente, contexto sócio-histórico, enfim, com a sua experiência de ser e estar no mundo.

Nos levantamento realizado na *Journal of Documentation* dos artigos publicados entre 2006 e 2016, percebemos que esta abordagem experiencialista encontra-se bastante presente nas publicações acerca da indexação. Dos 147 artigos produzidos entre 2006 e 2016 sobre o comportamento de busca da informação (*behavior information seeking*), 87 procuram estudar o impacto do contexto de uso da informação na forma em que o usuário (*querier*) busca a informação e 42 apresentam propostas de incorporar tais estudos na recuperação da informação, no sentido de elaborar padrões de entrada do conteúdo dos documentos na operação da indexação que considerem tais mapeamentos dos comportamentos de uso da informação.

A padronização desloca-se da linguagem para o comportamento do usuário. Tratam-se de estudos acerca de como o cognitivo do indivíduo funciona a partir dos estímulos do mundo e como as influências físicas, sociais, subjetivas afetam a maneira em que os indivíduos buscam a informação. Portanto, os modelos mentais consistem em esquemas, roteiros e estruturas do conhecimento. Esses modelos podem se transformar e se tornar em estruturas mais sofisticadas de acordo com os novos dados sensoriais e as novas situações em que o indivíduo se depara.

Trabalhos como *Language in the information-seeking context: a study of US scholars using non-English sources*, de Sabbar e Xie (2016), *Seek, share, or withhold: information trading in MMORPGs*, de Harviainen e Hamari (2015), *The Information Seeking and Communication Model: A study of its practical application in healthcare*, de Robson e Robinson (2015) são recentes estudos de caso que procuram mapear comportamentos semelhantes em contextos e conjunturas que são iguais.

O construtivismo social, encontrado nos artigos que analisam o comportamento de busca da informação, aborda os processos de informação descrevendo as necessidades de informação, buscando os critérios de relevância que afetam os indivíduos ou de como os indivíduos são dirigidos por seus estados emocionais e cognitivos atuais, além das situações e tarefas de trabalho para estabelecer as formas de busca da informação. Os estudos sobre as necessidades, a busca e o uso da informação, a interação na pesquisa e recuperação da informação, o comportamento de busca na internet e as concepções acerca da literatura de informação são preocupações dessa filosofia, orientada para o usuário e que se assemelha com o paradigma mentalista.

Percebemos que, mesmo se afastando da influência da filosofia aristotélica de compreender a linguagem como um sistema universal de representações do pensamento, há ainda a crença em categorias universais no construtivismo social. As experiências são diferentes, mas o aparelho cognitivo é o mesmo, a mente é um aparelho biológico comum a todos os indivíduos. Não seria a mente presa a um corpo, tal como concebe o mentalismo computacional, mas o corpo na mente, *body into the mind*. Portanto, há a necessidade de se mapear os padrões para que a partir deles as linguagens artificiais dos SRIs sejam estabelecidas.

Ao realizarmos a mesma análise no *Knowledge Organization*, notamos que alguns artigos chegam a abordar a indexação a partir do paradigma mentalista baseado no construtivismo social, sobretudo os estudos influenciados pela teoria de Beghtol, a exemplo da edição de 2010 em sua homenagem que contou com seis artigos revisionistas de sua obra.

Os artigos de Kwasnik (2010), Howarth (2010), Martínez-Ávila e Sansegundo (2013) direcionam sua abordagem acerca da indexação realizando a interface entre conceitos, termos e relações na construção de linguagens controladas com estudos do comportamento de busca da informação tal como a predominância dos artigos analisados na última década no *JD*.

As garantias para estabelecer os termos e os controles dos significados nestes estudos estão baseados na construção do conhecimento a partir da experiência dos indivíduos em seus contextos de busca e uso da informação na qual a linguagem é possível de ser padronizada, pois há comportamentos passíveis de serem mapeados por ser semelhante as formas de cognição dos indivíduos.

3.6 A CRÍTICA DO PRAGMATISMO À FILOSOFIA PLATÔNICA-ARISTOTÉLICA DA LINGUAGEM

A proposta de Day (2005) em abordar as concepções sobre os processos de representação da informação, aproximando-as dos estudos semânticos pela tendência pragmatista, parece ser oportuna, pois aponta para reflexões sobre a própria *práxis* da indexação. As análises pragmatistas

focam-se na materialidade da informação, do documento e das relações sociais entre os sujeitos de determinada comunidade discursiva no contexto dos SRIs.

Day ao criticar a abordagem cognitivista na indexação aponta três problemas importantes: em primeiro lugar, a tentativa de usar contexto como uma explicação causal e não como fator que interfere na dinâmica das representações; segundo, a tentação de ver as questões de enraizamento como objetos em contextos estruturais rígidos; e terceiro, a tendência de sobre-generalizar sujeitos e objetos, ou trivializá-los em nome de uma estrutura de narrativa comum (DAY, 2005, p. 579).

A principal premissa pragmatista é a negativa em assegurar uma correspondência exata entre o significado desejado pelo o emissor de um texto e o compreendido por seu leitor. A crítica do “segundo” Wittgenstein, em **Investigações Filosóficas** (1999), feita ao representacionismo aristotélico que influenciou a visão cognitivista da linguagem nos é útil para tratarmos a informação, uma vez que a base da informação é a linguagem.

De acordo com o filósofo, o problema de associarmos cada palavra a um conceito nos leva a crença de que os significados são entidades autônomas e correspondências diretas do real. Essa ânsia por generalidades cria o mito de que a essência das coisas está nos conceitos, sublimando a linguagem do pulsar da vida, apagando os contextos de uso das palavras. De forma semelhante, os estudos representacionistas da informação baseiam suas operações em sistemas de regras ou mecanismos bem definidos, que pré-determinam os usos concretos da informação, sublimando-os.

Para Day (2005), o sucesso da imposição de vocabulários transcendentais como mediação de dispositivos para a estabilização de sentidos sempre foi difícil de se atingir, porque os usuários tendem a implantar os seus próprios termos para relacionar significantes e significados (palavras e conceitos).

Mais recentemente, com a tecnologia digital, os vocabulários pós-coordenados têm se tornado mais comuns, embora os problemas de sobrecarga e baixa precisão persistam (p. 580). Esta percepção funciona de alerta para a abordagem da indexação dos documentos de arquivo em sua fase de produção, pois chama atenção de que representações construídas sem considerar os múltiplos usos da informação tendem a fracassar.

Aproximamos a análise dos estudos pós-estruturalistas de Day (2005) à concepção sobre informação e documento de Frohmann (2004). Influenciado pelos estudos semânticos, Frohmann reflete sobre os documentos e a informação em suas relações sincrônicas e não sucessivas. Logo, como a significação não é algo representativo de um signo relativo ao mundo, a informação também não o é. A informação constitui-se pela e na linguagem, se processa nas relações entre os homens, as coisas e os signos.

Os estudos pragmáticos semânticos nos dão importantes contribuições para trabalhar o processo de indexação. O processo de indexação poderá ser potencializado se além dos conceitos-

chave extraídos do documento em vez de procurar extrair do texto o assunto, procurando localizar o núcleo de qualquer documento, como se houvesse um caráter substantivo da entidade informação, basearmos a indexação a partir das práticas cotidianas, daquilo que as pessoas fazem com a informação e na análise do domínio.

A presunção da homogeneidade dos signos para elaborar índices, vocabulários controlados, linguagens documentárias é uma crítica realizada por esses teóricos da corrente pós-estruturalista na representação dos documentos. Em vez disso, apontam que nossas atenções deveriam voltar-se para os significantes e os destinatários, mais do que pensar em significados universais, sublimados do cotidiano das unidades de informação.

Sendo assim, procuramos analisar os arquivos de universidade, considerando a entidade produtora como uma comunidade discursiva que possui seus próprios usos dos documentos e da linguagem. A informação e seu sentido não possuem uma essência apriorística, eles se dão em seu uso. "A informação é aquilo que pode ser considerado informativo para alguém" (FROHMANN, 2001, p. 14). O que pode ser informativo para uma comunidade discursiva, em um contexto sócio-histórico, em uma instituição, não necessariamente servirá a outras comunidades em contextos distintos.

Para Frohmann (1990), a derivação de um texto e as representações do seu conteúdo ou de seu *aboutness*, é, naturalmente, altamente relevante para o primeiro passo, o passo mais fundamental do processo de indexação. No entanto, a explicação do poder de internalização tácita de regras mentais depende da suposição de que elas funcionam como sistemas internos de orientação que determinam nossas ações. Segundo o autor, o mentalismo na indexação parte do pressuposto que a derivação das frases indexadas dos textos é explicada pela invocação de regras mentais que automaticamente oferecem representações em forma de linguagem.

Os sistemas de indexação transformam a informação armazenada e os sistemas de recuperação da informação em coisas análogas ao sistema de recuperação da informação da mente humana. A mente de um indexador humano através de processos normais de percepção, combina estes contrapondo com aqueles estocados na mente, determina quais conceitos são representados e quais serão importantes, depois escolhe os símbolos para representar esses conceitos na indexação.

A crítica ao paradigma mentalista na indexação direciona-se à restrição de abordagens e possibilidades de outros métodos, limitando o processo à representação de setenças declarativas que compreendem uma fração muito pequena de documentos do mundo.

A imagem do mentalismo acerca do conteúdo da mente influencia a favor do conhecimento, pois os conteúdos mentais são considerados como as formas para se retirar as proposições. Quando se acredita que os conteúdos mentais são representações, eles naturalmente se transformam em fatos representados. Os conteúdos mentais são tidos como declarações que pouco a pouco são capazes de

mostrar a verdade ou falsidade dos casos. Uma fixação de formas propositivas pode ser quebrada pela atenção à grande variedade de formas linguísticas nos documentos atuais, mas quebrar este feitiço requer abandonar o mentalismo como um todo.

Frohmann realiza a crítica ao mentalismo na indexação apropriando-se do pensamento de Wittgenstein acerca do representacionismo da linguagem. A crítica de Wittgenstein (1999) à teoria representacionista de tradição platônica-aristotélica, procura elicitá-la como ela enraizou-se na sociedade ocidental e como os modos de ser, pensar, agir do sujeito no mundo estão imersos nessa filosofia.

Wittgenstein (1999) procura mostrar que a proto-imagem da linguagem feita pelo mentalismo direciona para uma única concepção possível que é esta: um instrumento de comunicação capaz de apresentar as ideias sobre as coisas. O logocentrismo presente neste paradigma faz acreditar que o significado está no conceito contido nas palavras. Nesta concepção instrumentalista, as palavras possuem uma função básica: elas designam as coisas do mundo. O filósofo critica a ânsia por generalidade do mentalismo. O aprendizado das palavras não é a captação apriorística da essência das palavras. O engano está em afastar a linguagem do fluxo das relações humanas, dos seus usos. Isto seria um reducionismo semântico, pois ao realizarmos tal sublimação, transformamos uma forma de utilizar a linguagem - para designar objetos - na única possível ou na mais importante.

A proposta de Wittgenstein (1999) é uma subversão da concepção da linguagem: em vez de voltarmos para o apagamento forçado dos contextos de uso das palavras em prol de conceitos universalizantes, mas precários, pensemos na linguagem a partir das experiências e da práxis cotidianas.

Os significados vão forjando-se na dinâmica da vida. As atividades humanas são inúmeras e a linguagem atravessa todas elas, são relações mutuamente constitutivas. Wittgenstein não nega a racionalidade da linguagem, mas também não a considera de forma demiúrgica. A qualidade demiúrgica é da linguagem.

As regras, as padronizações, a gramática não são entidades que decodificam e elucidam as formas que devemos empregar as palavras, são invenções e convenções dos homens estabelecidas nos usos da língua. Tais convenções não podem ser consideradas regularidades superlativas e abstratas, mas fazem parte dos jogos de linguagem.

Cada sentença é sempre enunciada em uma situação nova, ainda que possa ser semelhante a alguma das infinitas experiências humanas já conhecidas e mapeadas. Dizer algo é sempre jogar com possibilidades de acerto, mas também de erro. Aprender tais convenções não significa realizar conexões entre os conceitos e o real, mas, como diz Martins "saber jogar" (2009, p. 36).

Frohmann (1990) transpõe as críticas de Wittgenstein para o campo da indexação e da recuperação da informação. De acordo com o autor, o sucesso ou fracasso destes sistemas de grande elegância, poder e utilidade, tanto na Ciência da Informação, quanto em outras disciplinas, não estão em padrões determinados biologicamente, mas na ressonância que possuem no trabalho com regras expressas ou tácitas no uso da linguagem da qual a informação é fruto. Frohmann (1990) propõe que em vez de se procurar descobrir padrões de pensamento e a combinação entre eles, sobretudo no desenvolvimento da indexação, pensar em como as regras são construídas. Perceber que tanto os documentos indexados, quanto a operação da indexação são práticas sociais realizadas no uso da língua.

A teoria da indexação é confrontada com o desafio, não de descobrir regras produzidas segundo o inconsciente, mas de compreender a construção coerente, com os propósitos declarados, das regras formuladas e explícitas, rigorosas que podem ser utilizadas para produzir frases de indexação dos textos.

O problema da inconsistência na indexação, por exemplo, não é resolvido primeiramente pela descoberta e depois pelo ordenamento do conhecimento tácito das regras seguidas inconscientemente pelos indexadores, essa crença é substituída pela proposta de perceber que as regras são convenções vagas e instáveis. Os indexadores não mais fornecem as diretrizes do que expressa o conteúdo do texto em uma declaração concisa, os indexadores forçosamente interpretam, de forma variada, com regras suficientemente precisas para servir como justificativas, como padrões de correção, e como instrumentos de treinamento do indexador.

A mudança da “descoberta” para a “construção” das regras implica também em uma mudança metodológica. A percepção da construção da regra como prática social, diferentemente da descoberta da regra como padrão da mente, é simplesmente contingencial, em vez de necessariamente relacionada com o comportamento de indexação. Uma vez que o pressuposto das operações intelectuais como resultado da adesão inconsciente de regras mentais é abandonado, não há nenhuma razão para supor que a observação dos indexadores é fundamental na ausência de regras estritas, que é uma tarefa especialmente relevante para formular e melhorar as regras.

Regras de indexação são regras de representação de um texto para fins de sua recuperação. No entanto, a recuperação do texto designa um conjunto de práticas sociais particulares. Consequentemente, a construção dos princípios de indexação ou os seus tipos de facilidades, particulares e práticas de recuperação dependem, acima de tudo, de um entendimento preliminar de que as práticas sociais constituem a recuperação do texto num histórico, social, atual e real mundo.

De acordo com Frohmann (1990), a filosofia de Wittgenstein é útil à indexação por lembrar que seguir uma regra é uma *práxis*, não um hábito privado. Para construirmos e seguirmos tais regras é fundamental entender o jogo que estamos jogando.

A categoria comunidade discursiva, de Swales (1990), apropriada por Hjørland e Albrechtsen (1995) ao propor a metodologia de análise de domínio, parece-nos apropriada para abordar a questão da indexação. Em oposição às abordagens puramente mentalitas, os autores destacam o papel da esfera social para além do indivíduo existir, pensar e agir em sociedade. O social afeta o sujeito, pois a carga simbólica do mundo nomeado do qual ele faz parte afeta suas formas de pensar, agir e se relacionar.

A sociedade não é uma parte que compõe a vida do indivíduo tal como no experencialismo. O indivíduo está imerso em meio a ações, relações, forças, símbolos e trânsitos que afetam sua percepção e interação com o mundo. Nessa perspectiva, ele aprende a “ler” o mundo do qual faz parte, ou seja, aprende a interagir com a sua comunidade discursiva. Para além dos acordos da instrumentalidade de sua língua, o indivíduo aprende como sua comunidade significa o mundo, os acordos simbólicos existentes, as agências envolvidas nessas interpretações das coisas, dos seres e da vida.

Não trabalharemos aqui com a análise de domínio proposta por Hjørland e Albrechtsen (1995), isso se deve pelo desconforto em delimitarmos uma área do conhecimento como comunidade discursiva. Em nosso entendimento, ainda que toda delimitação seja um exercício de entender um retrato feito em determinado período para facilitar a compreensão daquela dada realidade, a sugestão dos autores não nos parece aplicável à realidade dos arquivos, pois tratar um campo do conhecimento como uma comunidade linguística é uma demasiada “igualação do não igual” em nosso estudo.

Os autores mostraram ser possível tal enquadramento em seus estudos ao repensarem a epistemologia da Ciência da Informação. No entanto, tal empreendimento bem sucedido pouco se aplica à abordagem do processo de indexação nos sistemas de arquivo das universidades na fase de produção documental. Já foi apresentado no presente trabalho exemplos de diversas abordagens sobre tal técnica que possuem inclusive genealogias filosóficas distintas. Portanto, procurar consenso em comunidades discursivas a partir do domínio, ou seja, do campo do conhecimento formado, parece-nos não ser possível na realidade em que analisamos.

Todavia, o conceito de comunidade discursiva mostra-se importante ao percebemos que a realidade dos sistemas de arquivo das universidades em sua fase de produção apresenta tais características apontadas pelos autores para configurar tal categoria.

As formas de administração em que estão subordinados os sistemas de arquivo, as instituições que fazem parte, os agentes que desempenham funções e atividades por meio dos documentos, os tipos dos documentos produzidos, os processos de trabalho e as relações mediadas pelos documentos possuem configurações semelhantes às comunidades discursivas. Esses elementos circulam por caminhos previstos devido aos acordos historicamente construídos e às

forças que atuam nelas, reproduzindo o campo de disputa simbólico e social a que pertencem. Essas comunidades discursivas interagem por meio de fatos sociais totais, compartilhando diversos acordos simbólicos, para além da importância lexical, mas que operam vários trânsitos, inclusive disputando, hierarquizando, distribuindo, escalonando, interditando práticas discursivas que não somente comunicam, informam e representam conteúdos de seus documentos, mas realizam, impedem, fracionam, repetem e negociam direitos, deveres, ações e relações.

Neste sentido, a indexação não deveria ser uma técnica abstrata, universal, aplicada de forma atemporal a partir de representações de conceitos autônomos, de entidades cuja essência está em si e não em relação ao mundo e aos sujeitos históricos. A indexação como processo desenvolvido em comunidades discursivas, processo de análise conceitual e representação da informação ou dos documentos em linguagens documentárias, foca-se na compreensão daquilo que sua comunidade entende como informação e quais os usos que realiza dos documentos.

As comunidades discursivas podem ter uma ampla rede de conexões e partilhar signos, significados, acordos, práticas discursivas. Essa possibilidade pode ampliar a efetividade da indexação seja no espaço e tempo, prevendo inclusive a inclusão da tecnologia no processamento e difusão de dados. No entanto, a manutenção ou não dessa estabilidade dos sistemas de representação deve-se muito a manutenção dos acordos, relações e trânsitos entre os sujeitos históricos da comunidade discursiva e das suas redes construídas. Nesta abordagem, as formas de cognição moldam-se a essas relações e não o contrário.

O consenso é construído nas comunidades discursivas e não na revelação da essência das coisas no mundo. “Não achando boi anônimo, posso eu dar nome aos bois?” (ROSA, 2001, p. 24). A pergunta de Guimarães Rosa chama-nos à realidade em que o mundo em que nascemos e aprendemos a interagir já está todo nomeado, interpretado e significado. Seguimos nele aprendendo convenções, acordos, relacionamentos e formas de ser e estar no mundo em comunidade. Na indexação, a percepção dessas produções de sentido e a aproximação destas e de seus usos resultam na efetividade do processo nos sistemas de recuperação da informação.

Diante da perspectiva pragmatista da linguagem, torna-se fundamental conhecer o contexto social, histórico e cultural da comunidade discursiva que produz seus signos, significados e significantes em sua relação com o mundo. Portanto, a seguir procuraremos compreender o universo dos arquivos, as universidades como comunidades discursivas e o estatuto do documento nesse contexto.

4 O DOCUMENTO, OS ARQUIVOS E SEUS USOS

A importância do documento para a Ciência da Informação é um debate que permeia o campo desde sua constituição como tal. O final da década de 1940, com a influência das teorias da comunicação e da documentação, ficou marcado pelo debate de qual seria o objeto da Ciência da Informação. Houve o aparecimento de duas correntes majoritárias: os partidários da teoria matemática da comunicação que defendiam a informação, sua composição por dados e seleção de mensagens, isto é, um objeto imaterial dessa nova área; e os adeptos da tese que definia a informação a partir de sua materialidade e nessa perspectiva o documento era tão importante quanto a própria informação por registrá-la.

De acordo com Buckland (1991), a informação apenas pode se dar a partir de sua configuração material. É a sua materialidade, o registro em algum suporte que torna possível a Ciência da Informação, uma vez que as atividades relacionadas à informação estão vinculadas ao seu suporte - recuperação, classificação, indexação, retórica, análise estatística, entre outros.

Ao procurar qualificar a categoria documento, Buckland, amplia a noção de materialidade da informação: "num determinado aspecto a informação é usada como evidência do aprendizado – como base para a compreensão. Conhecimento e opiniões são afetados pelo que se vê, lê, ouve e experimenta" (BUCKLAND, 1991, p. 354). Para o autor, a informação como coisa, tal como a evidência, não faz nada ativamente, propositalmente, são os "seres humanos que fazem coisas com a informação e para a informação" (BUCKLAND, 1991, p. 354).

Cabe aqui lembrar a crítica de Bourdieu (1987) ao fetichismo da evidência na construção de teorias científicas, tão caracterizada pelos estudos da Escola de Chicago no início do século XX, sobretudo na Sociologia, que recusa trabalhos empíricos, negando a própria definição de evidência, reconhecendo apenas parcelas ínfimas de dados, somente aqueles consagrados pelas tradições metodológicas de pesquisa. De certa forma, compreendemos que Buckland aproxima-se deste retorno à definição reclamada por Bourdieu ao ampliar os usos do termo evidência.

O conceito de documento também é ampliado por Buckland, não sendo apenas representações de informações. O autor defende que os documentos são artefatos produzidos pelo homem que evidenciam suas experiências, inclusive expressam a sua vontade de informar-se.

Já Frohmann (2004) analisa a potencialidade de agenciamento do documento/informação, no sentido de não somente evidenciar, informar, comunicar, mas também operar trânsitos, inventar sujeitos, produzir ações, ou seja, viabilizar relações o tempo inteiro. Frohmann, ao criticar a abordagem da informação como uma "substância nobre intencional presente no mundo" (FROHMANN, 2004, p. 392), propõe que a filosofia da informação passe a investigar as práticas informativas e documentárias.

Buckland (1991) e Frohmann (2004) aproximam-se em suas teorias ao destacarem a importância dos fenômenos circunstanciais para a informação ser considerada como válida. "Portanto, qualquer objeto particular, documento, dado ou evento pode ser considerado como informativo dependendo das circunstâncias, do mesmo modo que a 'relevância' de um documento ou um fato que é circunstancial dependendo da questão e da sagacidade do pesquisador" (BUCKLAND, 1991, p. 358). Buckland afirma ainda que a informação como coisa pode ser percebida a partir de consensos. "Onde há um consenso de julgamento, o consenso é algumas vezes tão forte que o status dos objetos, especialmente documentos, transformam a informação em inquestionável, exemplo, lista telefônica, painel de horário de voo e livros de exercício" (BUCKLAND, 1991, p. 359).

Frohmann (2004) enfatiza que esses consensos, tal como os jogos de linguagem teorizados por Wittgenstein (1999), são construções sócio-históricas a partir de práticas discursivas e informativas. Os autores destacam a importância da análise dos contextos sócio-históricos da produção e uso dos documentos para melhor compreender o seu alcance informativo.

As práticas documentárias possuem certas propriedades particularmente notáveis em tais descrições. A primeira é a materialidade: já que os documentos existem em alguma forma material, sua materialidade configura práticas com eles. As fotografias excluem de suas narrativas aqueles que não enxergam, pois é a faculdade exigida para a sua compreensão, por exemplo. Uma segunda propriedade das práticas documentárias é o quanto profundamente incorporadas elas estão nas instituições. Muito da autoridade da informatividade dos documentos depende dos locais institucionais de sua produção, um argumento apresentado por Foucault em sua discussão sobre a importância das instituições para a formação de modalidades enunciativas (FOUCAULT, 1996, p. 51-52).

Os documentos de arquivo são altamente regulados, desde sua produção pelo ordenamento jurídico e pelos atos administrativos, até a sua guarda em arquivos com processos rígidos de custódia. Uma terceira propriedade é a disciplina social: práticas documentárias, como muitas outras, exigem treinamento, ensino, correção e outras medidas disciplinares; o argumento é reforçado pelo papel do treinamento em diversos jogos de linguagem de Wittgenstein e enfatizado pelo elo de Foucault entre o aparato disciplinar e o campo da documentação (FOUCAULT, 2007, p. 189-194). Uma quarta propriedade é a historicidade: práticas surgem, desenvolvem-se, entram em declínio e desaparecem – tudo isso sob circunstâncias históricas específicas. Essas quatro propriedades restritivas das práticas documentárias são somente noções analíticas; descrições completas de tais práticas retrataram interações entre algumas ou todas elas (FROHMANN, 2004, p. 392). Essas ideias de materialidade, instituições, disciplina social e história fornecem uma reflexão útil para a análise das práticas documentárias.

De acordo com essas propriedades elencadas por Frohmann (2004) das práticas documentárias, podemos perceber como os documentos são mediadores de relações sociais, pois em todo momento eles estão viabilizando ações, regulando discursos, sistematizando pessoas como demonstra Bazerman em seu livro, **Gêneros textuais, tipificação e interação** (2005). Segundo o autor, os documentos organizam pessoas e atividades. Os documentos e suas interações com as atividades criam padrões interativos, atitudes, relações. Percebe-se, dessa maneira, que os documentos não apenas comunicam, descrevem ou registram informações, por meio deles diversas ações são viabilizadas.

Por exemplo, um aluno somente poderá colar grau se houver em seu histórico a integralização da carga horária e dos créditos referentes às disciplinas de seu curso. Para que isso ocorra, ele deverá ter sido aprovado em todas elas, aprovação que somente pode ser comprovada pelo lançamento das notas por seus professores nos diários de classe, mas antes mesmo do nome do discente constar nos diários de classe, faz-se necessário a sua inscrição em disciplinas que somente é viabilizada pelo registro da matrícula. Entretanto, tal registro e, por conseguinte, o ingresso do aluno na universidade, só será possível mediante a apresentação de diversos documentos como, no caso do Brasil, o diploma de conclusão do ensino médio, a inscrição no Sistema de Seleção Unificada (SiSU) e a identidade do discente. Nessa perspectiva, notamos que os documentos não apenas estão inseridos nas atividades, ações e relações sociais, mas também as produzem, operam diversos trânsitos, fabricam sujeitos, determinam caminhos, o que Marcel Mauss chamou de fato social total.

Existe aí um enorme conjunto de fatos. E fatos que são muito complexos. Neles, tudo se mistura, tudo o que constitui a vida propriamente social das sociedades que precederam as nossas – até às da proto-história. Nesses fenômenos sociais ‘totais’, como nos propomos chamá-los, exprimem-se, de uma só vez, as mais diversas instituições: religiosas, jurídicas e morais – estas sendo políticas e familiares ao mesmo tempo –; econômicas – estas supondo formas particulares da produção e do consumo, ou melhor, do fornecimento e da distribuição –; sem contar os fenômenos estéticos em que resultam estes fatos e os fenômenos morfológicos que essas instituições manifestam. (MAUSS, 2003, p. 187).

Os documentos nesse sentido possuem agência uma vez que desencadeiam relações sociais. Eles nunca param de circular em nossa sociedade, circulam entre indivíduos, indivíduos e coisas, indivíduos e direitos/deveres, misturam-se nos fazeres da sociedade. Os documentos não apenas informam ou registram informação ou ações, mas participam ininterruptamente da vida social dos sujeitos e das coisas.

Diversos atos, direitos, deveres, conquistas e interdições acontecem por sua mediação. Existe um caráter performático do documento para além do seu uso informativo, declarativo. Seguindo a lógica de Bazerman (2005), os indivíduos criam novas realidades de significações,

relações e conhecimento fazendo os usos dos documentos. Os documentos organizam pessoas e atividades.

Austin, ao abordar os atos de fala, realiza uma distinção entre os constatativos e performativos. Em seu livro **Quando dizer é fazer** (1990), o linguista inglês afirma que dizer é comunicar, declarar, informar, descrever, mas também é uma forma de realizar, interagir com o outro e com o mundo atuando na linguagem. Os enunciados performativos produzem ações e nada declaram ou informam. Quando frases do tipo "Aceito esta mulher como minha legítima esposa", ou "Batizo esse navio com o nome de Rainha Elizabeth", "Aposto cem cruzados que vai chover amanhã", exemplos utilizados pelo próprio Austin, são ditas, os sujeitos não descrevem atos, mas os estão praticando, ou seja, dizer faz parte da ação casar, batizar e apostar.

Geralmente o proferimento de certas palavras é uma das ocorrências, senão a principal ocorrência, na realização de um ato, cuja realização também é o alvo do proferimento, mas está longe de ser, ainda que excepcionalmente o seja, a única coisa necessária para a realização do ato (AUSTIN, 1990, p. 26).

Nota-se que muitos atos exigem o proferimento de certas sentenças e a existência das mesmas somente se dá pela necessidade imposta pelo ato. No entanto, Austin alerta sobre a necessidade de determinadas circunstâncias para que os atos de fala enquanto performativos aconteçam. Eles não bastam em si, precisam satisfazer critérios para que sejam bem-sucedidos e não se tornem sem efeito ou desencadeiem ações indesejadas. Austin então apresenta as condições para performativos felizes: deve existir procedimento convencionalmente aceito, que apresente efeito familiar e que inclua o proferimento de determinadas palavras, ditas por pessoas certas, em circunstâncias apropriadas; o procedimento deve ser completo e tem de ser executado de maneira correta por todos os seus participantes (AUSTIN, 1990, p. 31).

Nessa perspectiva, os documentos e as atividades mediadas pelos mesmos apresentam-se como sistemas articulados e bem estruturados, circulam por caminhos previstos, com consequências familiares e de fácil compreensão, legitimando, permitindo e interditando sejam discursos, sejam pessoas (BAZERMAN, 2005). Muitas ações dependem da existência de documentos anteriores a elas ou existem, simplesmente, pois o documento permitiu, autorizou, propiciou o seu surgimento.

Os documentos viabilizam realidades ou fatos por estarem inseridos num modo de vida social organizado também por meio deles. Portanto, cada documento encontra-se encaixado em atividades sociais estruturadas e depende da existência de documentos anteriores que influenciam nas atividades e na organização social.

Para Searle (1979), o indivíduo nasce em um mundo dotado de linguagem e aprende a ser um ente social incorporando acordos, convenções, formas de expressão, normas, estruturas, caminhos para compreender e ser compreendido e assim viabilizar a sua própria existência na

sociedade. Então, a mediação realizada pelos documentos precisa obedecer certos padrões de comunicação que direcionam o sentido e a performance dos atos subsequentes.

Rousseau e Couture (1998) argumentam que o Estado regula e controla os cidadãos por meio de documentos. Posner (1959) afirma que os documentos são os meios em que o Estado se relaciona com a sociedade. Schellenberg (2006) demonstra como a Administração Pública necessita dos documentos para viabilizar suas atividades e funções.

4.1 A ESPECIFICIDADE DO DOCUMENTO DE ARQUIVO

Os documentos de arquivo demonstram a perspectiva de regulação e mediação das ações de forma explícita por sua vinculação ao ordenamento jurídico e à administração do Estado. Suas relações históricas traçadas com essas esferas os configuram como reguladores de relações sociais. Os documentos de arquivo registram ações que seus produtores desejam e/ou precisam registrar de acordo com o sistema administrativo-legal em que fazem parte.

Diversas funções e atividades são realizadas pela mediação destes dispositivos. Consequentemente, os documentos de arquivo não são apenas artefatos produzidos por suas instituições de origem para atestar uma ação, mas eles são produzidos porque há imposições, coerções, consensos, múltiplas relações que exigem a sua produção. Pensar em documentos de arquivo é pensar no sistema jurídico-legal da sociedade em que tais arquivos estão inseridos. Os documentos de arquivo são dispositivos produzidos, regulados, distribuídos e escalonados por convenções socialmente instituídas.

De acordo com Camargo e Machado (2000), os arquivos surgiram da necessidade prática de conservar registros das ações e fatos seja pela demanda de comprovar algo, seja para conservar uma memória, seja para informar sobre determinado fato.

Esse é o sentido original e primário da acumulação de documentos em toda e qualquer entidade, pública ou privada. Produzidos no decorrer de ações administrativas e conservados a título de prova, os documentos são capazes de refletir, como um espelho fiel, o próprio curso dos atos que lhes deram origem. Atos e registros são elementos inseparáveis (CAMARGO; MACHADO, 2000, p. 12).

Por comprovar atos e fatos, serem produzidos no decorrer de atividades administrativas, os documentos de arquivo estão vinculados a sua contextualização, eles não são sem seus contextos de produção. Para a Arquivologia, um documento de arquivo somente pode cumprir seu estatuto probatório por sua natureza contextual. Segundo Camargo (2003, p. 12), a contextualização é a própria genealogia do documento que evidencia os vínculos entre as ações, funções, estruturas e os próprios documentos. Esta necessidade de contextualização, própria do documento de arquivo, que

o distingue de todos os demais documentos, reside no “vínculo natural e necessário que mantém, a título de prova, com o produtor/acumulador” (CAMARGO, 2003, p. 11).

Para Bellotto (2002), os documentos de arquivo são testemunhos inequívocos da vida das instituições, pois registram informações sobre o estabelecimento, a competência, as atribuições, as funções, as operações e os processos de trabalho de sua entidade produtora. Estes decorrem de relações administrativas, políticas e sociais mantidas pela entidade produtora, tanto no âmbito interno como no externo, sejam com outras entidades de seu mesmo nível, ou com as que lhe são, hierarquicamente, superiores ou inferiores (BELLOTTO, 2002, p. 197).

Mendo Carmona (1995) também percebe o documento de arquivo inscrito em práticas sociais que expressam consensos estabelecidos previamente na sociedade. Por eles se estabeleceu que os usos dos documentos de arquivo podem ser de prova, de testemunho histórico, de registro de ações administrativas, de evidências do passado. Esses consensos são regulados por meio do sistema jurídico. De certa forma, pensar nos contextos de produção e usos dos documentos de arquivo é trabalhar na direção de se pensar as experiências e as práxis cotidianas.

Um dos principais usos dos documentos de arquivo, analisados tanto pela Diplomática Contemporânea, quanto pela Arquivologia, é o de evidência. Para a Diplomática, a força probatória do documento está associada aos elementos que o configuram. Segundo Carucci (1987, p. 34), o fato de tratar-se de um testemunho escrito; a natureza jurídica do ato que o gera e a forma de sua redação são elementos que legitimam o documento de arquivo como prova na sociedade. Para Rodrigues e Garcia (2012, p. 35), esta capacidade de provar o fato que lhe dá origem é resultado da especial relação que o documento tem com o órgão que o produz e do vínculo que se revela no conteúdo pela atividade registrada, que constitui o núcleo de sua identidade.

Embora a Diplomática Clássica procure identificar os elementos formais que configuram o documento em sua gênese e unidade como prova na sociedade, a Arquivística procura absorver seus métodos para compreender o valor probatório no conjunto documental orgânico, característica que especifica o documento de arquivo. É nesta perspectiva que se estrutura o conceito de arquivo, isto é, "conjunto de documentos de qualquer natureza, reunidos orgânica e automaticamente por um corpo administrativo, uma pessoa física ou jurídica, em razão de suas funções ou de suas atividades." (ROUSSEAU, COUTURE, 1998, p. 52).

As relações dos documentos entre si no desenvolvimento da atividade que os produziu é uma condição *sine qua non* para a existência do documento de arquivo. Tal condição foi qualificada por Duranti (1997a) como vínculo arquivístico. Trata-se da relação que vincula cada documento a outro no desempenho de uma ação até que esta seja realizada de forma plena.

O vínculo arquivístico, de acordo com Rodrigues e Garcia (2012, p. 35) "é uma parte essencial e principal componente identificador do documento, que torna documentos idênticos em

distintos, depois que o adquirem. Esta relação que se manifesta no conteúdo do documento, que o transforma em documento de arquivo.” Por tais razões os documentos de arquivo só podem ser analisados em seus contextos de produção a partir de processos documentais que ordenam as fases de realização da ação e, por conseguinte, a própria composição do documento de arquivo.

Esta configuração de documento de arquivo apresentada é fundamental para analisarmos os usos que os sujeitos que os produzem e os manipulam fazem dele. Identificamos alguns usos como: o de evidência, o mediador de ações e o informativo.

O campo da Arquivologia preocupa-se em pensar os usos dos arquivos e seus documentos ao longo do tempo. Desde a Antiguidade Clássica havia locais destinados a guarda de documentos considerados importantes para o governo das cidades-estado em Atenas no século V a. C. A forma e o local em que os documentos eram guardados conferiam-lhes autenticidade. Eles eram depositados no *archéion*, uma das dependências do *Métroon*, templo de Mnemósine, deusa da memória, mãe da Ágora de Atenas. Pela sacralidade do lugar, os documentos do *archéion* eram considerados de grande valor e possuidores de verdades incontestes (ROUSSEAU & COUTURE, p.33, 1998).

Os usos relacionados à administração pública sempre estiveram presentes nos arquivos enquanto instituições desde sua gênese. Os seus usos como prova de ação e meio para o desenvolvimento de atividades administrativas, debatido anteriormente, sempre estiveram presentes na constituição dos arquivos.

A partir do século XIV com a configuração do Estado Moderno, os arquivos consolidaram-se como lugares de salvaguarda de documentos considerados importantes para o desempenho das ações do Estado (POSNER, 1959, p. 11), uso que perdurou até o século XVIII. Com a criação dos *Archives Nationales*, em 1790, no contexto da Revolução Francesa, a missão dessas instituições, segundo a própria ata da Assembléia Nacional que o funda, seria salvaguardar os documentos relativos às “glórias e conquistas” da Nova França (SCHELLENBERG, 2006, p. 26). No século XIX, com o cientificismo e a disciplinarização dos saberes como relata Foucault, em seu livro **Em defesa da sociedade** (2005), a História procurou estabelecer-se como ciência utilizando métodos do positivismo comteano e do historicismo rankeano.

O positivismo comteano, usado como arcabouço teórico nas ciências humanas, defendia que o conhecimento se explica por si mesmo, necessitando apenas seu estudioso recuperá-lo e colocá-lo à mostra. Não foram poucos os que seguiram a corrente positivista: Auguste Comte, na Filosofia; Émile Durkheim, na Sociologia; Fustel de Coulanges, na História, entre outros, contribuíram para fazer do Positivismo e da cientificação do saber um posicionamento poderoso no século XIX.

A teoria positivista ganhou força num momento em que se acreditava ser a verdade um fim a ser atingido e que isso somente seria possível por meio de métodos científicos como a investigação, a observação e comprovação das leis do mundo – princípios naturais que regiam as coisas e que

estavam no universo prontos a serem descobertos assim como foram as leis newtoneanas (LÖWY, 2003).

Na França, em 1841, elaborou-se um dos fundamentos da Arquivologia: o princípio de respeito aos fundos, que consiste em manter agrupados os documentos de arquivo de uma mesma instituição, sem misturá-los a outros. Os arquivos provenientes de uma administração, de uma instituição ou de uma pessoa física ou jurídica são unidades singulares (DUCHEIN, 1986, p. 14). Dessa maneira, o campo arquivístico passou a consolidar suas especificidades enquanto saber. Seria de competência e responsabilidade dos arquivos classificar, ordenar e guardar documentos, sobretudo os do Estado, da administração pública, documentos de interesse dos cidadãos a partir de critérios criados por sua ciência, a Arquivologia.

Cabe, no entanto, destacar as diferenças conceituais entre o Princípio da Proveniência e de Respeito aos Fundos à luz dos argumentos apresentados por Schmidt (2012). A autora destaca o caráter prático do princípio de *respect des fonds*, criado pelo francês Nataly de Wailly em 1841, que estabelecia o respeito aos documentos de uma instituição, ou seja, tratava-se exclusivamente de não permitir que documentos fossem agrupados com outros de entidades produtoras diferentes. Já o princípio da proveniência está associado às formas administrativas encontradas na Prússia no século XVIII, “sua origem teórica se remete à obra de Philipp Ernst Spiess de título *Von Archiven* (SCHMIDT, 2012, p. 109).

O princípio prussiano da proveniência está relacionado à forma da burocracia prussiana organizar seus documentos de administração. Nesse sentido, o *registratur prinzip* compunha o princípio da proveniência. A forma de administração prussiana valorizava o registro de seus documentos demarcando sua origem, sua entrada e saída das repartições.

Nesse sentido, autores como Martín Ponzuello Campillos (1996), Cruz Mundet (1994), Londolini (1993), Bellotto (2006) e Schmidt (2012) vão tratar o princípio da ordem original, que estabelece a necessidade dos documentos serem mantidos na ordem que sua administração de origem teria dado, como desdobramento do princípio da proveniência. Esses princípios de organização dos documentos de arquivo foram consolidados com em 1898, a publicação do **Manual de arranjo e descrição de arquivos**, por Muller, Feith e Fruin, marcou a disciplinarização da Arquivologia, firmando o conceito de arquivo, a organização dos documentos, a descrição, os inventários, o uso de normas etc.

Os dois princípios, formalmente proclamados como normas jurídicas, converteram-se no núcleo histórico da ciência Arquivologia. O princípio de proveniência (*principio di pronienza*, em italiano; *provenienzprinzip*, em alemão; *herkomstbeginsel*, em holandês; *respect des fonds*, em francês; *principle of provenance*, em inglês) e o princípio da ordem original (*metodo storico*, *strukturprinzip* ou *registraturprinzip*, *respect de Vordre primitif*,

principle of original order) evoluíram a partir de uma obrigação legal até uma opção científica e cultural (SOUSA, 2003, p. 249).

A consolidação dos princípios da proveniência e da ordem original, além dos estritos laços com a Diplomática Clássica para definir a autenticidade e fidedignidade de seus documentos transformaram os arquivos em celeiros da história como afirma Bellotto (2002, p. 9). A objetividade, a minuciosidade, o detalhe e a dedicação impessoal, então, foram as grandes lições da escola positivista para o estudo de áreas que influenciaram o campo da História no século XIX. Essa metodologia de pesquisa para investigar o passado encontrou nos arquivos fontes confiáveis. Por mais que os arquivos nacionais guardassem os documentos referentes à Administração Pública, até o final da Segunda Guerra Mundial, a racionalidade de suas classificações não estava voltada para o atendimento prioritário deste campo, e sim para a pesquisa dos historiadores. Os documentos de arquivo passaram a ser considerados testemunhos históricos confiáveis. Esses usos persistem nos arquivos permanentes. Percebemos desse modo que os usos dos documentos são diferentes nos arquivos permanentes daqueles dos arquivos correntes.

Na fase de produção documental, a maior parte dos usos dos documentos está relacionada aos processos de trabalho. Os documentos costumam ser recuperados para viabilizar as atividades das unidades de trabalho que compõem a entidade produtora.

4.2 A GESTÃO DE DOCUMENTOS E SUAS ABORDAGENS NO CAMPO ARQUIVÍSTICO

A Arquivologia ao longo de sua história desenvolveu a teoria do ciclo de vida dos documentos de arquivo com o objetivo de facilitar a sua organização e tratamento. Composto por três períodos que se baseiam na frequência de uso dos documentos pela entidade produtora, o ciclo de vida divide os arquivos em: corrente, intermediário e permanente.

Há documentos em que sua guarda justifica-se por mediar possíveis direitos, deveres ou validações que ainda não prescreveram. No entanto, ao término deste período não possuem valor de guarda permanente, pois seus usos cessarão. Não são considerados testemunhos históricos, não terão potencial uso como evidência pelo ordenamento jurídico e o produtor não os considera como parte da memória institucional de suas rotinas administrativas. É importante destacar que tal abordagem está em consonância com os princípios da proveniência e da ordem original.

A divisão visando a delimitação de volumes documentais possíveis para a organização e tratamento não significa a ruptura com o contexto de produção desses. No entanto, o aumento da burocratização, a ressonância da teoria da administração científica e o desenvolvimento do capitalismo industrial, sobretudo no século XX, questões contextualizadas na segunda seção desta tese, levaram aos arquivistas voltarem suas atenções para o documento de arquivo em sua fase de produção.

A Segunda Guerra Mundial configura-se como um marco no campo da Arquivologia no tocante ao tratamento documental, uma vez que o aumento da burocracia foi exponencial e a demanda pela rápida recuperação dos documentos para a tomada de decisões estratégicas era um imperativo do período. Os Estados passaram a utilizar intensamente os documentos na mediação das suas relações com os cidadãos. Schellenberg denominou a esses documentos por tais características de serem volumosos e pela necessidade de sua organização para viabilizar as atividades imediatas da entidade produtora, neste caso o governo dos Estados Unidos, alvo de seu estudo, de documentos modernos (2006, p. 179). O valor administrativo do documento voltou a ser uma questão central na abordagem dos arquivos e não somente o de testemunho histórico tal como no século XIX.

Schellenberg, ao retomar a proposição de Jenkinson (1937, p. 4) do documento de arquivo como evidência de certas ações administrativas ou executivas realizadas no passado que preservam referências de responsabilidade das pessoas envolvidas com tais atos com o valor administrativo dos documentos, diferencia os documentos chamados por ele de modernos em duas categorias: os que possuem valor primário e aqueles que possuem valor secundário.

De acordo com o arquivista americano, os documentos surgem para viabilizar as funções pelas quais uma entidade foi criada, sejam por meio de suas atividades administrativas, legais, fiscais ou executivas (SCHELLENBERG, 2006, p. 180). Dessa maneira, os primeiros usos estão vinculados aos seus produtores. Já o valor secundário é aquele que persiste ao longo do tempo e não está associado à razão pelo qual foi criado. Schellenberg aponta que o valor de prova e o valor informativo de tais documentos justificam a sua preservação.

Todavia, Schellenberg afirma que na conjuntura de seu tempo, ou seja, logo após a Segunda Guerra Mundial, na década de 1950, era necessária uma nova definição de arquivo. No século XIX, o manual dos arquivistas holandeses, as contribuições de Jenkinson e a própria constituição dos arquivos públicos como “celeiros da história” consagraram métodos de tratamento documental como o arranjo e a descrição destinados aos documentos capazes de fornecer indícios sobre a origem, o desenvolvimento e o funcionamento das instituições humanas, tendo considerável valor de pesquisa como seus usos no decorrer da história demonstraram.

No entanto, as necessidades da nova realidade levavam aos arquivistas a pensarem novos métodos e abordagens para atender a tais demandas. “Uma definição que tenha surgido da observação de material da Idade Média não poderá atender às necessidades dos arquivistas que trabalham principalmente com documentos modernos. A recíproca desse fato também é verdadeira” (SCHELLENBERG, 2006, p. 40).

O autor reivindica uma maior atenção ao tratamento da massa documental na fase de produção, uma vez que este era o maior problema do arquivista daquele tempo. Então, o conceito de

ciclo de vida dos documentos surge de forma pragmática para possibilitar o tratamento de documentos que juntos não seria viável.

De acordo com Llansó I Sanjuan (1993), a teoria do ciclo de vida dos documentos surgiu no contexto norte-americano com as iniciativas como as Comissões Hoover (1947), a criação dos *records centers* e a instituição da *Federal Records Act* (1950). Essas transformações acabaram culminando na gestão de documentos, importante teoria no campo da Arquivologia.

O governo norte-americano, no contexto da Guerra Fria, procurou otimizar a burocracia estatal a partir de medidas políticas que visavam a melhoria dos usos dos documentos, influenciado pelos princípios da administração científica, as palavras de ordem eram economia, eficiência e eficácia. Instituídas em 1947 e lideradas por Schellenberg, as Comissões Hoover buscavam desenvolver formas que melhorassem a administração pública através da melhoria dos métodos de trabalho e do trâmite da documentação nos órgãos e departamentos do poder executivo federal dos Estados Unidos.

Um de seus produtos foram os próprios *records centers*, chamados atualmente de arquivos intermediários. Foi a solução encontrada para o principal problema do período que era o acúmulo de massa documental nas unidades de trabalho. Apesar de não serem utilizados com frequência, os documentos não podiam ser descartados por ainda gerarem efeitos legais, administrativos ou fiscais. No entanto, não podiam ser destinados aos arquivos históricos, pois ainda estavam vinculados a seu valor primário. Em 1948, foram criados os primeiros *records centers*, arquivos destinados aos documentos com baixa frequência de utilização nas unidades de trabalho, mas com valores administrativos. O ganho de espaço físico e a economia de gastos com pessoal, material e equipamento fizeram com que os documentos passassem a ser organizados em arquivos correntes, intermediários e permanentes.

Dessa forma, concomitantemente aos arquivos intermediários, elaborou-se a teoria do ciclo de vida dos documentos que os divide em três períodos: corrente, ativo, vivo, setorial, primeira idade - fase em que os documentos estão em vigência, servindo para mediar ações e relações necessárias ao desenvolvimento de atividades e funções da entidade produtora; intermediário, semi-ativo, purgatório, central, segunda idade - fase em que os documentos não são utilizados com tanta frequência, mas ainda não prescreveram; permanente, de custódia, inativo, morto, histórico, terceira idade - fase em que grande parte da massa documental já foi eliminada, sendo recolhidos a estes arquivos apenas os documentos que possuam efeitos jurídico-legais por tempo indeterminado ou que a entidade produtora julgue terem estas possibilidades de usos referenciais, ou seja, servem como testemunhos históricos (LLANSÓ I SANJUAN, 2006).

A teoria do ciclo de vida surgiu para facilitar a organização dos documentos, no entanto, como ressalta Cortés Alonso (2011), essa divisão serve para facilitar a manipulação e a

compreensão da massa documental, não deve se perder de vista a integralidade dos documentos. Outra importante teoria arquivística associada às três idades é a gestão de documentos. O *Dictionary of Archival terminology* (2012, p. 65) conceitua gestão de documentos (*records management*) como: "a área de gestão administrativa geral responsável em alcançar economia e eficiência na criação, manutenção, utilização e descarte dos documentos".

De acordo com Jardim (1987), as iniciativas dos governos dos Estados Unidos e do Canadá para o aprimoramento da administração pública culminaram na teoria do ciclo de vida e na gestão de documentos. Segundo o autor, Phillip Brooks foi quem estabeleceu o termo ciclo de vida dos documentos ao defender que os arquivistas deveriam voltar a suas atenções para o desenvolvimento de técnicas e métodos que subsidiassem políticas públicas para a administração e guarda dos documentos nos arquivos, em vez de se atentarem exclusivamente com os usos científicos dos arquivos históricos. Indolfo (2007, p. 31) afirma que a teoria se consolidou com a criação de programas de gestão de documentos e na implementação de arquivos intermediários, além de uma legislação própria que padronizava os procedimentos nos arquivos da administração federal.

Se a arquivologia tem suas raízes em alguns países europeus do século XIX, a gestão de documentos está associada aos Estados Unidos, especialmente após a II Guerra Mundial. Tal como a arquivologia a partir do século XIX, a gestão de documentos emerge, em meados do século XX, com um forte referencial estatal. No entanto, a gestão de documentos pode ser também compreendida fora das dinâmicas e reestruturações do Estado norte-americano, no cenário da organização capitalista no pós-Guerra e da influência da administração como campo científico (JARDIM, 2015, p. 20).

Para Mendo Carmona (1995), a gestão de documentos institucionalizou-se nos Estados Unidos a partir da aprovação da Lei de Arquivos, em 1950, que obrigava a todo órgão do Governo Federal a ter um programa permanente e ativo de gestão dos documentos. A autora indica ainda que a teoria de ciclo de vida documental e a gestão de documentos estão intrinsecamente ligadas desde seus surgimentos e consolidaram-se nos sistemas de arquivo a partir das operações de avaliação e de transferência dos documentos desde as unidades de trabalho até o arquivo permanente, passando previamente pelo arquivo intermediário.

Schmidt (2012) afirma que a promulgação da *Federal Records Act*, em 1950 foi um marco importante no estabelecimento de uma doutrina americana de gestão de documentos que cindiu a fase de produção da fase de custódia permanente. Os documentos na fase de produção passaram a ser concebidos como *records* que necessitavam de práticas arquivísticas diferentes daquelas dispensadas aos *documents*, caracterizados por seu valor secundário.

De acordo com Llansó I Sanjuan, a chave para se compreender a definição do conceito de gestão de documentos reside no conceito de ciclo de vida dos documentos, o objetivo de sua

formulação é garantir a presença do arquivista e os métodos que aplica, para que os documentos recebam o tratamento adequado em cada uma das idades (LLANSÓ I SANJUAN, 2006, p. 42).

Os estudos acerca da gestão de documentos dedicaram seus esforços para investigar as formas mais adequadas de organização e avaliação dos documentos, sobretudo nos arquivos correntes e intermediários, visando a eficiência e a eficácia na execução das atividades da entidade produtora.

Os documentos são eficientemente administrados quando, uma vez necessários, podem ser localizados com rapidez e sem transtorno ou confusão; quando conservados a um custo mínimo de espaço e manutenção enquanto indispensáveis às atividades correntes; e quando nenhum documento é preservado por tempo maior do que o necessário a tais atividades, a menos que tenham valor contínuo para fins de pesquisa e outros fins (SCHELLENBERG, 2006, p. 63).

Cruz Mundet define a gestão de documentos como o conjunto de tarefas e procedimentos orientados a conseguir uma maior eficácia e economia na utilização dos documentos por parte das administrações. Este processo se estende ao ciclo de vida completo dos documentos, desde sua produção até a sua eliminação final ou seu envio ao arquivo para sua conservação permanente (CRUZ MUNDET, 1994, p. 143). O objeto está dirigido a assegurar a preservação de uma documentação adequada, evitar o não essencial, simplificar os sistemas de criação e uso da papelada, melhorar a forma em como se organizam e recuperam os documentos, proporcionar o cuidado adequado e o armazenamento a baixo custo dos documentos nos arquivos intermediários e assegurar a eliminação adequada dos documentos.

Llansó I Sanjuan defende que um bom programa de gestão de documentos deve incluir: a) plano de classificação, b) programa de conservação ou tabela de temporalidade, c) gestão dos arquivos intermediários e dos documentos semi-ativos, d) transferência e seleção de documentos inativos destinados a ser conservados, e) eliminação de documentos inativos sem valor histórico (LLANSÓ I SANJUAN, 2006, p. 46).

Jardim (2015) apresenta diferentes concepções acerca do termo gestão de documentos como as oriundas das tradições americana, inglesa, canadense, espanhola, francesa e brasileira. O conceito validado pelo dicionário de terminologia da Sociedade Americana de Arquivistas descreve como “o controle sistemático e administrativo dos documentos, pensando em seu ciclo de vida para garantir a eficiência e a economia em sua criação, uso, manuseio, controle, manutenção e disposição” (SAA apud JARDIM, 2015, p. 22).

A compreensão inglesa assemelha-se com a americana ao focar-se no ciclo de vida para a garantia da eficiência e eficácia na administração da entidade produtora. De acordo com o *The management of public sector records*, editado por Michael Roper e Laura Millar (1999), gestão de

documentos é a “área da gestão administrativa geral que se ocupa da obtenção de economia e eficiência na criação, manutenção, utilização e eliminação dos documentos de uma organização ao longo de todo o seu ciclo de vida e na disponibilização das informações que contêm para apoiar os negócios dessa organização” (IRMT apud JARDIM, 2015, p. 22).

O projeto liderado pelos canadenses, o Interpares, conceitua em sua proposta a gestão de documentos como “o domínio da gestão responsável pelo eficiente e sistemático controle da criação, recebimento, manutenção, uso e disposição dos documentos, incluindo processos para capturar e manter evidências de informações sobre processos de trabalho (*business activities*) e ações (*transactions*) na forma de documentos” (The Interpares Project 3 apud JARDIM, 2015, p. 23). Aqui se nota uma preocupação com os procedimentos de tratamento arquivístico que garantam o valor de prova de ação do documento.

A norma técnica espanhola de gestão documental também expressa tal preocupação com a preservação das características do documento de arquivo em sua definição a ver “todas as funções, atividades e processos que se aplicam na organização dos documentos ao longo de sua vida para garantir a sua produção, autenticidade, integridade, conservação, confiabilidade e sua disponibilidade para maior uso e menor trabalho” (UNE 1589 apud JARDIM, 2015, p. 26, tradução nossa).

O conceito francês inclui a dimensão política e informacional ao estabelecer a gestão de documentos como “conjunto de políticas, procedimentos e atividades que permitem uma organização para estabelecer controle sobre a informação que geram os documentos, desde sua criação até o descarte ou recolhimento aos arquivos permanentes” (PIAF 2.0 apud JARDIM, 2015, p. 24).

No Brasil, o **Dicionário Brasileiro de Terminologia Arquivística**, define como “o conjunto de procedimentos e operações técnicas referentes a sua produção, tramitação, uso, avaliação e arquivamento em fase corrente e intermediária, visando a sua eliminação ou recolhimento para guarda permanente” (CONARQ, 2005, p. 100). Já o **Dicionário de Terminologia Arquivística**, editado por Ana Maria Camargo e Heloísa Bellotto, conceitua como “conjunto de medidas e rotinas visando a racionalização e eficiência na criação, tramitação, classificação, uso primário e avaliação de arquivos” (AAB/SP, 2012, p. 52). Nesta segunda concepção notamos a influência da tradição norte-americana, sobretudo dos pressupostos de Schellenberg, ao atribuir à gestão de documentos o tratamento na fase corrente e a preocupação com o processo de avaliação.

Estas definições sobre gestão de documentos são importantes, pois mostram a influência do modelo norte-americano em diversos países no tocante aos métodos arquivísticos, sobretudo na divisão entre a fase de produção e a fase permanente. Apesar da grande ressonância no campo da Arquivologia da proposta canadense da arquivística integrada no final da década de 1990, a prática

da gestão de documentos é desenvolvida sobretudo nos arquivos correntes e intermediários pela própria genealogia do conceito.

Na arquivística integrada, influenciada pelas propostas de Hugh Taylor (1984 e 1993), Ducharme e Rousseau (1980) e formalizada por Rousseau e Couture (1998), a preocupação era manter a organicidade dos documentos de arquivo, uma das principais características que os especificam. Na perspectiva canadense, a gestão de documentos deve englobar as três fases da *vida* do documento (corrente, intermediária e permanente) com o objetivo de garantir a sua unidade e permitir a recuperação do seu contexto de produção.

Taylor (1984), ao revisar os modelos de avaliação e descrição mais utilizados na Europa e na América do Norte, criticou a separação entre os *records* e os *documents*. Segundo o autor, era necessário pensar novas estratégias de gestão que integrassem ambos, pois o potencial informativo global estava sendo perdido. O maior desafio do século XX para os arquivistas seria preservar os princípios arquivísticos numa era de complexidade dos processos informacionais (1993, p. 228). Os arquivos precisavam ser vistos como sistemas integrados de informação.

Para Rousseau e Couture (1998), a gestão integrada garante o respeito aos princípios da proveniência e da ordem original tão caros à Arquivologia. Ademais, os canadenses questionam os usos dos documentos na fase de produção, criticando a divisão entre valor primário e secundário. De acordo com os canadenses, as unidades de trabalho para a execução de suas atividades muitas vezes demandam por informações e documentos que não possuem somente valor primário, mas também secundário. A arquivística integrada procura:

garantir a unidade e a continuidade das intervenções do arquivista nos documentos de um organismo e permitir assim uma perspectiva do princípio das três idades e das noções de valor primário e secundário; permitir a articulação e a estruturação das atividades arquivísticas numa política de organização de arquivos; integrar o valor primário e o valor secundário numa definição alargada de arquivo (ROUSSEAU; COUTURE, 1998, p. 70).

A perspectiva canadense de gestão global da informação orgânica abre uma discussão importante que é a preocupação com o uso da informação integrada, ampliando a visão sobre os usos dos documentos de arquivo. O nosso enfoque de tais usos na fase de produção converge com esta abordagem, pois diversas informações circulam e são necessárias na execução dos processos de trabalho.

A abordagem australiana também é uma importante contribuição para o desenvolvimento de possibilidades de gestão dos documentos de arquivo, como afirmam Costa Filho e Sousa (2017). Nesta, os arquivos são concebidos como organismos em constante mutação e formação. Peter Scott ao romper com o modelo de *record group*, propôs na década de 1960 que os arquivos fossem vistos como línguas que se moldam às circunstâncias, produtos e agentes daqueles que os operam, fazendo

parte das dinâmicas sociais, econômicas, ideológicas, suas interações são tanto sincrônicas, quanto diacrônicas (CUNNINGHAM; MILLAR; REED, 2012).

Cook (1997) argumenta que a concepção de Scott foi fundamental para o desenvolvimento da ideia do *continuum*, pois seu modelo de descrição de séries documentais evidenciou os múltiplos relacionamentos entre documentos entre si e entre diversos produtores, contestando o modelo hierárquico e diacrônico consagrado no campo dos arquivos.

Upward (1996) retoma a concepção dos *records continuum* na década de 1990 para enfrentar o desafio da gestão documental em um período que os sistemas estão conectados à internet e os documentos natos digitais já eram uma realidade na administração australiana. Baseado então na teoria de estruturação de Giddens, propõe quatro dimensões para a compreensão do *records continuum*:

1ª dimensão: criação – caracteriza-se pela multiplicidade de atores e ações, relacionada às atividades, funções e missão da entidade.

2ª dimensão: captura – trata-se do componente intrínseco do documento que define a(s) evidência(s) específica(s) de acordo com o propósito para o qual foi criado o documento. Engloba sujeitos e unidades de trabalho da instituição, responsáveis pela captura dos documentos associados a suas atividades com fins de comprovação destas.

3ª dimensão: organização – estruturação de todo conjunto documental de uma entidade, conforme suas funções e processos de trabalho. A documentação é vista como suporte à entidade para o desempenho de suas funções.

4ª dimensão: pluralização – a entidade insere-se na dimensão social e cultural de que faz parte. Torna-se parte da memória coletiva. Diz respeito às relações institucionais com a sociedade. Compromisso com o acesso à informação.¹¹

Para Upward (1996), a gestão de documentos precisa ser tratada de maneira global e contínua. É preciso obter-se um entendimento prático das entidades enquanto culturas ou organismos vivos que produzem e usam informações e documentos. Aqueles que criam, organizam e controlam os documentos de arquivo são orientados por suas próprias necessidades e visões de mundo.

A gestão de documentos baseada em estruturas hierárquicas por meio de relações de superior-subordinado entre documentos, unidades de trabalho, agentes acaba por camuflar significativas relações formais e informais numa entidade. Estas, quando reunidas, revelam a missão, as estruturas, as redes e as atividades da entidade.

¹¹ O diagrama do *records continuum*, de Upward (1996), encontra-se no anexo E.

McKemmish (2001) afirma que a concepção dos *records continuum* é baseada no estabelecimento de um regime integrado de gestão de processos para toda a existência dos documentos. Portanto o tratamento do documento de arquivo incluir:

- A análise da criação dos documentos no fluxo dos processos de trabalho, o que significa assegurar apenas a produção dos documentos necessários;
- O design, o estabelecimento e a operação dos sistemas de arquivo na perspectiva do *continuum*
- A gestão dos documentos usados nos processos de trabalho (tradicionalmente separados como do domínio da administração pela influência da teoria do ciclo de vida) e nos arquivos (tratados como arquivos de valor mediato).

Nesta abordagem, os gestores de documentos estão operando em ambientes multidimensionais em termos de criação, estruturas, procedimentos, fluxo de trabalho, relações entre os sujeitos. Não se deve separar tempo e espaço da gestão de documentos.

O *continuum* não rejeita as visões cíclicas da realidade. De acordo com McKemmish (2001), os documentos de arquivo estão constantes estágios de ser e vir a ser. Eles assumem novas formas e estruturas durante o uso, o tratamento e o arquivamento. Agregam inclusive novos processos que os formam, seja pelo rearranjo ou novos padrões administrativos que os alteram em si, o controle de seu trâmite e a atenção que eles recebem.

Os documentos podem inclusive ter múltiplas vidas no mesmo tempo e espaço como os contextos que estão a sua volta de seu uso e controle podem ser múltiplos, dando novas possibilidades de ação e envolvendo sua remodelagem e a renovação de seus ciclos de criação e disposição (UPWARD, 1996)

Bearman (1995) propõe que o tratamento documental considere:

- interseções de regiões e uma amplitude espacial de contextos imediatos de interação;
- rotinização de procedimentos de gestão documental que ofereça tal amplitude;
- distanciação de formas rígidas de concepção de tempo e espaço;
- formas de totalidade social

Portanto, os arquivos são dinâmicos e abrigam múltiplas realidades. Os sujeitos envolvidos nos processos de criação dos documentos são diversos, assim como suas ações. Os processos de criação, organização, uso e controle dos documentos englobam: estrutura, conteúdo, ordenação, representação, recuperação, armazenamento e segurança (MCKEMMISH, 1997).

Concordamos com Costa Filho e Sousa (2017) de que a visão da gestão de documentos fornecida pelos australianos, na perspectiva do *continuum*, fornece possibilidades para a resolução

dos limites do modelo do ciclo documental, rompendo com a estrutura hierárquica, com as amarras de tempo e espaço, além de considerar os múltiplos usos dos documentos de arquivo.

4.3 OS USOS DOS DOCUMENTOS DE ARQUIVO

Se a literatura arquivística, entre o século XIX e a segunda metade do século XX, dedicou-se a consolidação dos princípios arquivísticos e as funções que os garantissem como a classificação e a descrição, respostas às demandas relacionadas aos usos dos documentos nos arquivos históricos, após a Segunda Guerra Mundial, passou a produzir estudos acerca da gestão de documentos, procurando estabelecer padronizações de procedimentos e modelos a serem aplicados, dando ênfase na avaliação como função arquivística estratégica para uma boa administração.

No entanto, as novas tecnologias, a mudança dos suportes dos documentos, as grandes transformações nas interações entre sujeitos, sistemas e os processos de trabalho têm levado ao campo a reavaliar o estatuto do documento de arquivo, seus princípios e as estratégias de representação dos documentos visando a sua recuperação, além de outros desafios frente a nova realidade.

O foco de nossa pesquisa encontra-se justamente em compreender os usos dos documentos nos arquivos e as demandas informacionais para a sua melhor representação e, por conseguinte, facilitar o desenvolvimento dos processos de trabalho, das atividades e das funções. Nesta perspectiva, embora reconheçamos as contribuições da arquivística integrada, é necessário destacar que na fase de produção documental, existem especificidades em relação aos usos dos documentos e às demandas por informação que são diferentes daquelas da fase permanente.

Os próprios Rousseau e Couture (1998, p. 114) definem os arquivos correntes como a fase de atividade na qual os documentos são imprescindíveis ao desenvolvimento das atividades da entidade produtora. Os autores recomendam que eles estejam o mais próximo possível de seu utilizador, ou seja, do próprio produtor, e devem ser fácil e rapidamente acessíveis. A razão para isso é que estes documentos estão viabilizando os processos de trabalho o tempo inteiro em seu cotidiano.

Os arquivistas canadenses ressaltam que cada entidade produtora possui uma realidade e um universo documental e que estas fases não podem ser tratadas de forma cartesiana, sem considerar a fluidez dessas fronteiras, uma vez que tal forma de organização é uma construção artificial e a realidade das organizações e as demandas inesperadas podem fazer com que os documentos transfigurem-se de fases.

Tanto Rousseau e Couture (1998), quanto Schellenberg (2006) afirmam que a diferença entre as fases de produção e guarda permanente está no uso dos documentos atrelados ao valor primário e secundário. Para estes autores, o tratamento documental deve estar ligado diretamente

aos usos que se fazem deles. É importante perceber as especificidades dos usos primeiros dos documentos na sua fase de produção. O objetivo dos documentos nas unidades de trabalho é servir às finalidades para as quais foram criados.

O aspecto mais importante da administração de arquivos correntes relaciona-se com o uso dos documentos no curso das operações governamentais. Pouco se faz dentro de uma repartição que não se torne objeto de um documento. Tanto os administradores de cúpula, que tratam dos principais programas, quanto os funcionários de categorias inferiores, que executam operações de rotina, precisam de documentos no seu trabalho (SCHELLENBERG, 2006, p. 68).

A compreensão de Schellenberg em que os documentos na fase corrente estão envolvidos nos processos de trabalho, nas atividades, nas funções da entidade produtora, sendo produzidos e produzindo tais ações e relações, é importante para compreender as necessidades de seus utilizadores.

Os usuários nos arquivos correntes, em sua maioria, são os próprios sujeitos que atuam nas unidades de trabalho. A eficiência na administração dos documentos na fase de produção revela-se na rapidez em se localizar o documento sem qualquer tipo de confusão, além de serem descartados aqueles que já cumpriram seus objetivos e não possuem valor secundário, não gerando lixo informacional desnecessário.

Nas unidades de trabalho, os documentos operam diversos trânsitos: transmitem orientações; relatam a execução das atividades, registram as decisões e deliberações, autorizam ações, legitimam e validam atos. Schellenberg (2006, p. 69) alerta ainda para a dificuldade de administrar os documentos considerados mais importantes pela entidade produtora nas unidades de trabalho, uma vez que estão em constante circulação pelas demandas dos processos de trabalho.

Rousseau e Couture (1998, p. 118) lembram ainda que os documentos na fase de produção são criados também para concretizar e veicular uma ou várias informações importantes para o desenvolvimento das atividades da entidade produtora. Estes são sempre demandados e, por isso, precisam com frequência ser recuperados rapidamente.

Os documentos gerados vão mediar, auxiliar, validar as ações que precisam ser realizadas para o desenvolvimento das atividades e, por conseguinte, o exercício das funções e cumprimento da missão de sua entidade produtora, inclusive gerando outros documentos no decorrer dessas transações.

Destarte, a atuação da entidade produtora depende da mediação de seus documentos. Então, a recuperação de um documento na fase de produção pode atender a diversos usos: para dar continuidade ao desenvolvimento de determinada atividade, comprovar a realização de algum fato, informar sobre rotinas administrativas realizadas, revisar procedimentos, organizar condutas,

consultar decisões tomadas, entre outros. Tais usos estão pautados não somente no documento em si, mas no conteúdo que registram. É importante qualificar a entidade produtora para compreendermos as suas especificidades e as diferenças nos usos dos documentos.

4.4 A ENTIDADE PRODUTORA E OS PROCESSOS DE TRABALHO

A entidade produtora caracteriza-se por sua missão, funções e atividades. A missão diz respeito a própria razão da existência da entidade, a sua principal finalidade que justifica a sua necessidade na sociedade (SANTOS, 2013, p. 137). As funções, chamadas pela doutrina do Direito Administrativo Brasileiro de competências, correspondem às responsabilidades que em seu conjunto são capazes de cumprirem a missão da entidade (SCHELLENBERG, 2006, p. 84).

Sendo assim, entendemos como função o conjunto de atividades destinadas a um propósito, uma abstração e, como tal, precisa de uma estrutura feita por regras convencionadas e recursos para se materializar. Concordamos com Paul Saborin (2001, p. 144) que define função como qualquer finalidade de alto nível, responsabilidade, tarefa ou atividade que, dentro da hierarquia da instituição produtora, definida na agenda de sua instituição com regulações legais, políticas ou mandatárias.

Consequentemente, as funções de uma instituição são o próprio motivo de sua existência, uma vez que podem ser compreendidas como um conjunto ou série de atividades, ou como define Foscarini (2009), processos de trabalho, que, quando realizado de acordo com uma sequência prescrita, irão resultar em uma produção institucional ou individual de resultados esperados, isto é, a execução de serviços delegados.

Já as atividades correspondem ao conjunto de procedimentos ou ações desenvolvidas pelas unidades de trabalho da entidade produtora para cumprir, desempenhar suas funções. Schellenberg (2006) insere neste universo as transactions, termo traduzido no Brasil, em seu livro como atos, outros autores como Santos (2013), Araújo (2013) e Rodrigues (2008) como tarefas, e outras abordagens de Jardim (2015), Negreiros (2007), Lopez *et al.* (2011) utilizam como sinônimo processos de trabalho. Já Cruz Mundet os chama de expedientes, ou seja, servem para encaminhar uma resolução de determinado assunto ou matéria.

Portanto, tratam-se do “conjunto ordenado de documentos e atuações que servem de antecedentes e fundamentos à resolução administrativa, assim como as diligências encaminhadas a executá-las” (SIERRA apud CRUZ MUNDET, 1994, p. 152). No entanto, tais termos são utilizados para designar as ações que são tomadas para normatizar, direcionar, executar ou dar prosseguimento a determinadas atividades. Neste sentido, são semelhantes à definição de *transactions* dada por Schellenberg (2006, p. 85). A mesma literatura que trabalha tal conceituação esclarece que nem toda tarefa produz documentos de arquivo.

Para melhor compreendermos as diferenças entre ação, função e atividade, utilizamos a definição clássica de Schellenberg:

Os documentos de arquivo são produto de uma ação. Uma ação pode ser tratada em termos de funções, atividades e atos (transactions). O termo função é usado com referência a todas as responsabilidades atribuídas a um órgão a fim de atingir os amplos objetivos para os quais foi criado. Geralmente, essas funções são definidas nas leis ou regulamentos que criam o órgão. [...] Cada função pode ser subdividida em diversas atividades, consideradas como uma série de ações, levadas a efeito de uma função específica (SCHELLENBERG, 2006, p. 84)

A complexidade da organização de uma entidade produtora está vinculada à sua missão. As unidades de trabalho, os níveis de hierarquia, a autonomia dos setores, as autoridades existentes são estabelecidas pelo volume de trabalho e pelas características sociais, econômicas e culturais de suas atividades-fim, tendo em vista que estas que especificam a entidade produtora.

De acordo com Santos (2013), as unidades de trabalho, chamadas pelo autor de unidades organizacionais, são determinadas pelas competências da entidade produtora. “O nível da unidade depende de sua importância em relação aos objetivos da organização ou do órgão de cuja estrutura interna participa” (SANTOS, 2013, p. 138). Logo, cada competência pode originar uma unidade de trabalho.

As competências ou funções geralmente estão estabelecidas nos atos normativos da entidade produtora (regimentos, estatutos, regulamentos, portarias, decretos, contratos sociais, entre outros) e a partir delas que a estrutura organizacional deve ser instituída, originando outros importantes documentos constitutivos como o organograma, o plano diretor, o manual de governança, o planejamento estratégico, etc.

Nos processos de trabalho, os documentos vão sendo vinculados uns aos outros formando provas, ditames, acordos, autorizações, execuções de ações, comunicados, notificações e outros trâmites que irão viabilizar determinada atividade. Durante os chamou de vínculo arquivístico (*archival bond*): “o vínculo arquivístico diz respeito à rede de relacionamentos que cada documento possui com outros documentos produzidos como parte do mesmo processo de trabalho ou atividade e localizados dentro de um mesmo grupo” (DURANTI, 1997a, p. 215-216, tradução nossa).

Portanto, os vínculos arquivísticos são componentes estruturais dos processos de trabalho, os seus relacionamentos fazem parte do desenvolvimento das atividades que se dão em cadeia, na tramitação de tais processos em relações que Cruz Mundet (1994, p. 152) chamou de “causa-efeito” pela forma em que se desenvolvem.

Nesta fase de produção documental, no desenvolvimento dos processos de trabalho que viabilizam a execução das atividades da entidade, o conteúdo dos documentos são de grande importância para a realização destes, pois neles estão informações importantes que fundamentam

determinadas decisões, as próprias decisões, as medidas tomadas anteriormente, as diretrizes para a execução das próximas ações, entre outros conteúdos específicos cujo acesso determina a boa ou má realização de um processo de trabalho.

Cruz Mundet (1994), ao abordar a fase de produção documental, afirma que uma boa gestão de documentos nesta primeira fase se revela como algo fundamental ao desenvolvimento pleno das atividades da entidade produtora. É uma ferramenta essencial para reduzir custos e evitar a perda informacional.

Diversos problemas são elencados como exemplos pelo autor que a falta de gestão de documentos pode ocasionar na fase de produção como “(...) o atraso na resolução de assuntos, o que ocasiona prejuízos à própria organização e a terceiros, o desperdício do tempo de trabalho, salários, motivação, mal gastos em buscas de informação difíceis, senão impossíveis, se traduz nos mesmos termos” (CRUZ MUNDET, 1994, p.148, tradução nossa). O autor ainda destaca o caráter performático do uso da informação na fase de produção documental.

A informação acessada quando demandada permite a resolução de problemas, o desenvolvimento de ações, a tomada de decisões estratégicas. No entanto, a falta dela, sua incompletude ou defeito acarreta consequências negativas que podem inclusive impedir o desempenho das funções de uma entidade produtora.

Desse modo, a gestão de documentos mais que objetivar a economia de custos, a otimização de espaço e a eficiência dos serviços, deve realizar um mapeamento das necessidades informacionais de cada unidade de trabalho, os usos que realizam dos documentos em suas atividades, além de identificar o volume e os tipos documentais produzidos. Esse trabalho deve embasar o desenvolvimento do sistema de gestão arquivística de documentos cuja arquitetura contemple um sistema de recuperação da informação integrado. “Os arquivos das unidades de trabalho são subsistemas integrados num sistema global, o do arquivo, pelo qual devem gestionar-se com a vida posta no futuro de essa documentação (que não seja destruída arbitrariamente, que permaneça organizada, etc)” (CRUZ MUNDET, 1994, p. 1994).

Percebe-se que a informação registrada no documento na fase de produção está intrinsecamente ligada à realização do trabalho e a sua qualidade, acessibilidade e rapidez de seu fluxo influem no desempenho deste. Existe um trânsito entre as informações e as ações nas unidades de trabalho, uma via de mão dupla na qual a ação gera informação e a informação influem no desenvolvimento da ação. Por esta razão a documentação deve ser organizada de modo a refletir os expedientes de trabalho, como estabelece Cruz Mundet (1994, p. 151) e, acrescentamos: a indexação deve refletir os usos dos documentos pelas unidades de trabalho.

Outrossim, como esclarece o autor, há especificidades nos processos de trabalho na Administração Pública que precisam ser considerados como, por exemplo, serem regulados por leis

cujos atos de seus servidores são vinculados. O tipo documental mais comum na Administração Pública é o processo que pode ser iniciado de ofício, por necessidade da própria entidade produtora, ou por requerimento da parte interessada.

No trâmite do processo, há diversos procedimentos cujo registro é necessário como qualquer decisão, despacho, recurso, deliberação, alegação, inclusive com a possibilidade de inclusão de documentos que comprovem ou justifiquem tais medidas. Na execução dos atos, se o sistema em que os processos são registrados indexar os documentos que são juntados ao processo, sua consulta para o desenvolvimento das próprias atividades da entidade será facilitada. “A localização de um documento se resolve com a busca do mesmo através da relação do conteúdo, de modo que não é necessário revolver todos os papéis e sim ir direto ao documento” (CRUZ MUNDET, 1994, p. 156, tradução nossa).

As universidades como entidades produtoras possuem especificidades que a distinguem da maioria dos órgãos administrativos no que tange as suas dinâmicas organizacionais, estruturas hierárquicas e práticas decisórias. Não há os trâmites comuns de uma organização administrativa com estrutura hierárquica rígida que demarca claramente as subordinações e coordenações das unidades de trabalho. Os processos de trabalho são mais autônomos, respondendo muitas vezes apenas a seu departamento, instituto ou faculdade. Os trâmites da documentação nem sempre seguem os caminhos usuais pela própria cultura de liberdade de ação e independência na tomada de decisão que a universidade possui.

A autonomia das unidades de trabalho acaba por não facilitar a padronização de rotinas administrativas e leva a dificuldade de estabelecer ordenamentos que regulem tais práticas. Cohen e March (1986) nomeiam as universidades por essas características como anarquias organizadas.

Wilson (1976) ao abordar a realidade dos arquivos de universidades canadenses relatou algumas especificidades que podemos identificar também no Brasil. Segundo o autor, pela autonomia das universidades, pela pulverização dos arquivos em centros urbanos e pela escassez de recursos direcionados a essas unidades, há uma grande dificuldade de desenvolver sistemas de gestão arquivística padronizados que atendam as demandas de suas comunidades. A própria geografia das universidades - com unidades de trabalho como faculdades, institutos e departamentos, coordenações, cursos espalhados em diversos campi - pode dificultar a aplicação e desenvolvimento de um sistema integrado. Ademais, nem todas as unidades de trabalho possuem forma de registro de sua produção documental.

Embora Wilson esteja descrevendo uma realidade da década de setenta e que tais distâncias e problemas de comunicação possam ser mitigadas pelas tecnologias, persiste esse modelo institucional, ou seja, há liberdade para atuação das unidades de trabalho que muitas vezes

funcionam administrativamente de maneira praticamente independente, com relações de subordinação nem sempre precisas.

As unidades de trabalho como coordenações, departamentos, institutos e faculdades muitas vezes respondem a conselhos e colegiados, órgãos de deliberação, e não a setores administrativos como acontecem em outras entidades públicas e privadas, com estruturas altamente hierárquicas a exemplo de ministérios, empresas públicas, sociedades de economia mista ou mesmo multinacionais.

A própria cultura universitária, como enfatiza Wilson (1976) de liberdade e autonomia como instituições sociais promotoras de conhecimento e reflexões críticas sobre a realidade nas mais variadas áreas, transforma em uma questão delicada a implementação de sistemas arquivísticos integrados, uma vez que enseja, como qualquer determinação de implementação e execução de um sistema, o ferimento de um dos princípios basilares de qualquer universidade: a autonomia em todas as suas esperas - didática-científica, administrativa, disciplinar, econômica e financeira.

Por isso, os sistemas de gerenciamento arquivístico, da mesma forma que os sistemas de recuperação da informação que podem estar inseridos no primeiro e cujo processo de indexação deve fazer parte, precisam ser desenvolvidos nas universidades a partir de políticas arquivísticas e políticas de informação baseadas na construção de consensos, ou seja, com formulação de ações, envolvimento de atores sociais, definição de agenda e avaliação das políticas implementadas.

Ao lidar com os arquivos universitários canadenses, as generalizações são difíceis e potencialmente enganosas à medida que as condições mudam rapidamente. Cada universidade tem seu próprio caráter e auto-imagem, produzido por uma mistura de tradição, faculdade, interesses curriculares e objetivos, além do espírito dos ex-alunos. Essa diversidade se reflete na extensão dos recursos, no tempo de mandato em que o arquivista está alocado e no equilíbrio definido entre os documentos de universidade e as coleções privadas (WILSON, 1976, p. 17, tradução nossa).

Embora cada universidade possua a sua natureza documental e suas particularidades, no Brasil algumas preocupações são compartilhadas como a posição secundária dos arquivos dentro da estrutura administrativa, a dificuldade de aplicabilidade de sistemas gerenciamento arquivístico dos documentos, além da pouca inclusão dos arquivos, dos cursos de Arquivologia e dos professores da Ciência da Informação nas políticas de informação e de arquivos.

A informalidade e a falta de rigidez gerencial são reflexos do estado da administração universitária, sobretudo nas instituições públicas, pois, geralmente, os administradores são acadêmicos que acumulam ambas funções. Ademais, aqueles que trabalham nas unidades, em sua maioria, são servidores públicos que respondem a um regimento jurídico único, o seu agir na Administração não está subordinado às organizações, mas à sociedade. "Os servidores não

obedecem às organizações da Administração Pública, mas participam delas" (CARVALHO FILHO, 2005, p. 42).

Há ainda universidades que ao padronizarem os sistemas de informação subordinam os arquivos a departamentos de documentação e biblioteconomia. É importante que os sistemas de informação de uma entidade produtora sejam integrados, no entanto, respeitando as especificidades de cada documentação.

Wilson (1976) relata que no Canadá, por este tipo de organização, os investimentos, prioridades e objetivos das bibliotecas universitárias muitas vezes se sobrepunham aos programas de gerenciamento arquivístico de documentos. Os relatórios, os orçamentos, as necessidades no tocante a pessoal qualificado e a espaço ficavam escamoteados por estarem inseridos em uma estrutura na qual os arquivos não eram reconhecidos. Acabavam por se enquadrar em um lugar que não eram considerados como pertencentes, tampouco se identificavam.

Essa realidade descrita por Wilson (1976) assemelha-se com a das universidades federais brasileiras. Ademais, os procedimentos e metodologias como a classificação e a avaliação são essencialmente incompatíveis entre arquivos e bibliotecas de universidades pelas próprias características de sua documentação.

Até que tais questões sejam julgadas e avaliadas com base em prioridades arquivísticas ou mesmo administrativas em vez de prioridades de setores, coordenações ou mesmo superintendências de documentação e/ou bibliotecas, os programas de gerenciamento de documentos nas universidades permanecerão acidentais, informais e ineficientes. Wilson afirma ainda que as universidades podem estar entre as poucas instituições onde a ineficiência administrativa é valorizada como sinônimo de liberdade e auto-gestão (1976, p. 23).

Os administradores universitários são duplamente cautelosos. Nas universidades, há uma parcela justa daqueles que se apegam aos velhos hábitos e às culturas burocráticas pouco eficientes e de grande informalidade, além disso, os administradores também são prudentes, para que suas propostas de mudança não sejam interpretadas como uma conspiração com o objetivo de aumentar seu poder às custas da universidade.

O sistema de planejamento descentralizado comum à maioria das universidades, o esforço para consultar o mais amplamente possível a comunidade universitária na tomada de decisões, a relação desconfortável entre administradores, professores e servidores, além de uma suspeição geral quase inerente dos sistemas de gestão combinam-se para erguer-se um grande obstáculo para a inauguração de um programa de gestão de documentos nos arquivos universitários (WILSON, 1976, p. 23).

As universidades federais, sobretudo após a promulgação da lei nº 12.527, de acesso à informação, e da portaria interministerial MJ/MP nº 2.320 que institui o Sistema Protocolo

Integrado – que abordaremos nas próximas seções - estão procurando melhorar seus sistemas de gestão de documentos, tentando fazer o melhor uso destes e reduzir o gasto com recursos. Sistemas de informação e serviços administrativos centralizados estão sendo formulados e programas de gestão de documentos mais sistemáticos seguem em implementação.

O trabalho nos arquivos das universidades pela necessidade de adequação a tais normais passou a ter certa visibilidade. No entanto, a preocupação da comunidade universitária com os arquivos continua sendo mais aparente na fase permanente, com a documentação custodiada que é alvo de interesse por pesquisas acadêmicas. O trabalho dos arquivos na fase de produção documental ainda é pouco conhecido, sendo alvo de pouco interesse ou preocupação pelas instituições.

Dessa forma, a comunidade universitária pouco se conscientiza do papel dos arquivos como administradores de sua documentação, preservando aqueles que possuem valor legal, fiscal e de informação, para além do valor histórico. Como resultado, a comunidade também pouco se interessa com a aderência rigorosa dos arquivos aos princípios da prática arquivística – como a proveniência e o respeito à ordem original .

Elizabeth Yakel e Laura Bost, em seu estudo sobre os usos dos arquivos pelos serviços administrativos das universidades, *Understanding Administrative Use and Users in University Archives* (1994), mostram que a maioria das abordagens da Arquivologia sobre arquivos universitários está voltada para a investigação do comportamento e das necessidades dos usuários nos arquivos históricos, no entanto, os administradores, em muitos casos, são a primeira população usuária dos arquivos.

Yakel e Bost (1994) afirmam que os arquivistas têm sido lentos em estudar os usos administrativos dos arquivos de universidades. As autoras qualificam como uso administrativo qualquer demanda direcionada aos arquivos para dar assistência aos administradores universitários, ao pessoal administrativo, aos docentes, ou qualquer atividade ou função relacionada a sua missão.

Percebemos que as pesquisas realizadas nos arquivos pelos administradores, geralmente, são para resolver questões específicas, com começo e fim imediatos. Os documentos de arquivo em unidades de trabalho apresentam características peculiares: podem ser soluções potenciais para problemas administrativos, assim esperam por encaminhamentos ou podem estar em circulação, ao mesmo tempo, os arquivistas precisam disponibilizar esses documentos nos lugares certos para atender as demandas de uso, inclusive informacional, geradas pelo administrador.

Elizabeth Yakel e Laura Bost argumentam que muitos problemas administrativos nas universidades não são resolvidos pela falta de comunicação entre os arquivos e as unidades de trabalho. Segundo as autoras isto se dá por causa da falta de gestão de documentos e da má administração da base de dados dos arquivos que não analisam os usos de sua documentação.

O sistema de recuperação da informação em um arquivo precisa basear seus procedimentos e mecanismos a partir da análise da maneira em que os pesquisadores-administradores definem os seus processos de trabalho, atividades e funções em termos de assunto. Isto se dá porque os administradores têm usos particulares da informação, uma vez que, muitas vezes, eles são os próprios produtores das informações que demandam.

Yakel e Bost mostram em seu estudo que os usuários administrativos possuem demandas informacionais diferentes dos usuários dos arquivos permanentes, estas estão relacionadas na execução das suas atividades - as respostas precisam ser imediatas devido às restrições administrativas de tempo no desenvolvimento dos processos de trabalho (YAKEL; BOST, 1994, p. 602). Por esses usos informativos dos documentos de arquivo pelos administradores é que a gestão de documentos deve priorizar à recuperação da informação, da mesma forma que faz com a avaliação dos documentos e sua organização.

Os processos de trabalho precisam dos documentos e muitas vezes do seu conteúdo informacional para viabilizar suas ações. Os usos administrativos e os usos informativos dos documentos são simultâneos e intrínsecos, não devem tratados de formas isoladas nos arquivos. A teoria arquivística dedicou-se durante um longo período aos estudos sobre organização dos arquivos e à adoção dos princípios da proveniência e da ordem original, assegurando a organicidade dos documentos de arquivo e possibilitando a recuperação do contexto de produção. No entanto, urge pensar técnicas que melhorem a recuperação dos documentos e das informações contidas neles. Estas técnicas devem estar pautadas nas análises dos usos que a comunidade discursiva faz de seus documentos.

4.5 OS DOCUMENTOS DE ARQUIVO NATO DIGITAIS E SUA GESTÃO

O imperativo tecnológico, o advento dos sistemas informatizados e as novas formas de se relacionar por meio deles, com outras características e outras demandas, levaram ao crescimento de estudos relacionados à compreensão destes novos ambientes de trabalho. A interação entre os indivíduos passa também pela mediação de máquinas que operam diversos sistemas interconectados em rede, conseqüentemente se tornou necessário refletir sobre tais configurações.

As formas dos documentos, os meios em que circulam, as maneiras em que viabilizam ações e transações nesta nova realidade passou a ser uma preocupação também da Ciência da Informação e da Arquivologia. A configuração dos documentos em novos suportes, as formas de validação jurídica e diplomática desses documentos, sua tramitação em rede, a custódia a partir de repositórios digitais, passaram a ser problematizações frequentes nos estudos sobre arquivos.

No âmbito das universidades, desde a década de noventa, algumas iniciativas destacaram-se no sentido de oferecer alternativas que articulassem a gestão de documentos com o tratamento

temático da informação contida nos documentos a fim de possibilitar a sua recuperação. O projeto desenvolvido pela Universidade de Indiana, em 1995, procurou trabalhar seus documentos como evidências que viabilizam os processos de trabalho.

Esse projeto, sob liderança de David Bearman e Richard Cox, intitulado *The Indiana University Electronic Records Project*, foi desenvolvido em quatro etapas: a primeira sendo a análise funcional das unidades de trabalho, identificando os processos de trabalho e as categorias das informações básicas necessárias para estabelecer evidências nos próprios processos de trabalho registrados. A segunda tratando-se da identificação e descrição dos sistemas existentes para a manutenção dos documentos (ou sistemas de informação que executam funções de manutenção de documentos), para os processos de trabalho previamente identificados. A terceira etapa consistindo em avaliar os sistemas existentes em termos dos "requisitos funcionais para evidências na manutenção de documentos" e os "metadados com especificações derivadas dos requisitos funcionais desenvolvidos na Universidade de Pittsburgh" (projeto em que a Universidade de Indiana associou-se) (BANTIN; BERNBOM, 1996, p. 252). A quarta e última etapa tratando-se do próprio desenvolvimento de táticas de implementação.

Os objetivos básicos eram rever as modificações dos sistemas de gerenciamento de documentos em termos de táticas - políticas, padrões, projetos, etc. - mais adequados para garantir a sua implementação e pleno funcionamento.

De acordo com Bantin e Bernbom (1996), mais do que dados e informações, os arquivos coletam evidências. Bearman define a evidência como “resultado da conjunção de dados (ou seja, ‘a gravação das palavras, números, imagens e sons realmente criados pelo criador’), estrutura (ou seja, ‘as relações entre esses dados empregadas pelo criador de registro para transmitir significado’) e o contexto (ou seja, ‘a relação entre o documento e a atividade de que surgiu)” (BEARMAN apud BANTIN; BERNBOM, 1996, p. 249). Se algum desses atributos estiver faltando, o resultado são dados, ou seja, não são documentos de arquivo.

Na fase de produção, os documentos registram e viabilizam os processos de trabalho. Um documento não é apenas uma coleção de dados, mas a consequência de uma transação. Esta redefinição, segundo Bantin e Bernbom (1996), teve um grande impacto na elaboração dos projetos de gestão de documentos no final dos anos 1990.

No relato dos autores acerca da proposta para os arquivos da Universidade de Indiana, o foco muda da análise dos tipos de documentos a serem mantidos e da análise de conteúdo para a forma e o conceito em que o documento foi criado e seu uso. Isso não significa que os tipos documentais e os conteúdos dos documentos não sejam considerados na gestão de documentos, mas a ênfase se encontra nos processos de trabalho, sobretudo na fluidez do ambiente eletrônico.

Destarte, o importante é analisar os arquivos como sistemas integrados que capturam, mantêm e recuperam documentos. Os sistemas precisam ser consistentes e confiáveis para o desenvolvimento dos processos de trabalho.

Por esta perspectiva, o projeto de gerenciamento de documentos nos arquivos da Universidade de Indiana foi todo orientado para a análise funcional como meio para entender a criação de documentos; definindo documento como evidência de transação, evidências definidas em termos de conteúdo, contexto e informações estruturais; e, finalmente, visando um objetivo maior: a análise do sistema de arquivo considerando a sua manutenção (BANTIN; BERNBOM, 1996, p. 251).

Pelo escopo de nossa tese se encontrar no processo de indexação na fase de produção documental visando a recuperação do documento para viabilizar os processos de trabalho, a primeira fase do projeto de gestão de documentos da Universidade de Indiana nos é útil justamente pelo seu foco na análise das funções, na identificação dos processos de trabalho e nas evidências desses processos de trabalho – em sua tríplice dimensão: dados, contexto e estrutura. Ao realizar a análise funcional, é necessário abordar três questões básicas: como se caracteriza uma função na entidade produtora? Como obter as informações necessárias para fazer essa análise? Quais são os produtos deste processo?

Como já relatado, as funções podem ser identificadas por meio dos documentos constitutivos da entidade produtora. No caso das universidades, esses documentos podem ser leis que criam tais instituições, decretos que definem suas estruturas e estatutos, seus regimentos gerais, as atas de seus conselhos, portarias do Ministério de Educação, entre outros exemplos. Ademais, as rotinas administrativas das unidades de trabalho podem ser mapeadas tanto pelos próprios tipos documentais, quanto por entrevistas realizadas junto aos servidores. Essa análise pode compor três níveis: funções, atividades e processos de trabalho.

Na indexação dos documentos de arquivo em sua fase de produção, a análise dos processos de trabalho é fundamental, pois são eles que pautam a demanda pela recuperação dos documentos. O processo de identificação dos processos de trabalho é a ação considerada em uma análise funcional proposta por Bearman e Cox (BANTIN; BERNBOM, 1996, p. 257). Por isso, o conceito de processo de trabalho deve estar fundamentado nas ações realizadas no desenvolvimento de uma atividade, considerando toda a tramitação nas unidades de trabalho. Não se tratam de seleções arbitrárias, mas de procedimentos delimitados por cada unidade de trabalho, na execução de suas tarefas.

De acordo com Bantin e Bernbom (1996), definir a potencialidade de registro de um documento, significa identificar o processo de trabalho e preservar sua evidência (dados, contexto e estrutura). Identificar as categorias de informação não é o mesmo que identificar os dados, pois

nesta compreensão, as informações estão vinculadas aos processos de trabalho, não são apenas dados registrados pelos sistemas computadorizados.

As informações estão atreladas aos valores atribuídos aos documentos na fase de produção e esses valores são caracterizados pelos usos de tais documentos nas unidades de trabalho. Portanto, para desenvolver um consistente processo de indexação na fase de produção documental é necessário compreender qual tipo de informação será demandada nos processos de trabalho, nas atividades e funções como evidências.

Desse modo, a questão passa a ser: que tipo de informações precisamos para desenvolver os processos de trabalho? A resposta para tal questão passa por políticas de indexação, modelos de aplicação de indexação e construção de linguagens de indexação baseadas não em formas abstratas e universais, mas na análise da comunidade discursiva da universidade, considerando suas funções, atividades e processos de trabalho.

Outros dois projetos de grande repercussão no campo da Arquivologia acerca da gestão arquivística de documentos de universidade foram os desenvolvidos pela Universidade de Pittsburgh e a Universidade de British Columbia (UBC), no Estados Unidos e Canadá respectivamente.

No final da década de noventa, os projetos surgiram da preocupação das universidades com o impacto da tecnologia e dos sistemas eletrônicos, sobretudo na criação, manutenção e uso de seus documentos. Os documentos digitais representavam um desafio fundamental para os métodos arquivísticos, para as teorias na Arquivologia e para as instituições produtoras de documentos de arquivo, assim como para as pessoas que trabalhavam na área.

Os projetos partilhavam da preocupação com as diretrizes, critérios e métodos que os arquivos iriam adotar frente a nova realidade documental que ameaçava princípios basilares como o da proveniência e o da ordem original. “Os arquivistas que avaliam, adquirem e descrevem documentos digitais hoje têm muito pouco subsídios para ajudá-los em seu trabalho. Em grande parte, isso decorre devido ao ritmo de mudança tecnológica; tudo é novo e os sistemas que avaliamos hoje podem ter pouca semelhança com os de cinco anos atrás” (MARSDEN, 1997, p. 159).

O que originou os projetos na Universidade de Pittsburgh e na Universidade de British Columbia foi o mesmo contexto: a necessidade de criar requisitos e modelos teóricos e práticos bem documentados para enfrentar a interseção da tecnologia com a gestão de documentos e informações. O Projeto Pittsburgh, liderado por David Bearman e Ken Sochats, intitulado **Requerimentos Funcionais para Evidência Eletrônica** (*Functional Requirements for Electronic Evidence*)¹²,

¹² As diretrizes do projeto estão disponíveis no site: <http://www.archimuse.com/papers/nhprc/BACartic.html>. Acessado em 04 de janeiro de 2018.

apresentou como seu principal objetivo "definir explicitamente quais requisitos devem ser atendidos pelos sistemas de gerenciamento arquivístico de documentos para que [os arquivistas] possam intervir na política organizacional, no projeto de sistemas e na implementação do programa para garantir a criação de documentos, preservar sua integridade e fornecer acesso" (MARSDEN, 1997, p. 159, tradução nossa).

Já o projeto da UBC, liderado por Luciana Duranti, Terry Eastwood e Heather MacNeil, intitulado **Requisitos para Gerenciamento Arquivístico** (*Manage Archival Framework*)¹³, possuía como meta "definir os métodos para garantir a confiabilidade e autenticidade dos registros eletrônicos com base em conceitos e princípios diplomáticos e arquivísticos" (DURANTI, EASTWOOD, MACNEIL, 1997, tradução nossa). Ambos possuíam como centro de suas preocupações a manutenção da integridade e dos elementos essenciais da natureza dos documentos de arquivo ao longo de sua existência, fossem pela perspectiva da teoria do "ciclo de vida" (UBC), ou pela do "continuum" (Pittsburgh).

Os focos do trabalho que se revelavam diferentes. Enquanto a Universidade de Pittsburgh preocupava-se com "a precisão, a compreensão, o significado e a coerência" (MARSDEN, 1997, p. 160, tradução nossa) dos documentos nos sistemas de gerenciamento arquivísticos eletrônicos, a Universidade de British Columbia estava voltada para as questões de "autenticidade, confiabilidade, integridade" (MARSDEN, 1997, p. 160, tradução nossa) dos documentos de arquivo nos sistemas.

O projeto da Universidade de Pittsburgh estava direcionado em manter os documentos de arquivo como evidência dos processos de trabalho. Ao se concentrar no documento como evidência de uma transação, Pittsburgh encorajava os arquivistas a olhar não apenas para o conteúdo dos documentos, mas também para o contexto de sua criação e sua estrutura.

Com uma forte base em análise dos processos de trabalho e design de sistemas, os membros do projeto viam documentos de arquivo como resultado de processos de trabalho previsíveis, como qualificar um cliente para receber benefícios, comprar ou vender ou fazer um pedido de patente entre outros exemplos dados pelo relatório da pesquisa. Este não é um conceito ou redefinição inteiramente nova de documento de arquivo, pois os pesquisadores afirmam que o trabalho trata-se de uma reformulação vigorosa do conceito de valor probatório já há muito trabalhado por Jenkinson (1937) e Schellenberg (2006).

A definição do documento como evidência de uma transação é uma das mais características distintivas do Projeto Pittsburgh e já ganhou adeptos em várias jurisdições, particularmente na Austrália, onde a sua aplicação foi promulgada na **Política de Gerenciamento de Mensagens Eletrônicas como Documentos** (*Policy on Managing Electronic Messages as Records*) e nas **Diretrizes para gerenciamento de mensagens eletrônicas como documentos** (*Guidelines on*

¹³ Disponível em: <<http://www.interpares.org/UBCProject/index.htm>>. Acesso em: 04 de janeiro de 2018.

Managing Electronic Messages as Records). Nesta perspectiva, o valor informacional do documento não está desvinculado do seu valor como evidência, prova de ação nos processos de trabalho.

Os sistemas informatizados de gestão arquivística, denominação que receberam no Brasil, precisam capturar as evidências dos processos de trabalho nos ambientes eletrônicos. No projeto da Universidade de Pittsburgh, o documento como evidência do processo de trabalho é o que norteia as atividades de identificação, classificação, avaliação e descrição nos sistemas informatizados de gestão arquivística de documentos. É uma percepção de gestão arquivística centrada nos processos de trabalho.

A padronização necessária aos sistemas informatizados de gestão arquivística de documentos uma vez voltada para os processos de trabalho torna-se menos complexa uma vez que os documentos criados como resultado de processos de trabalho geralmente já são estruturados e, por conseguinte, facilmente modelados. As transações têm variações finitas e seu "documento" é conduzido por exercícios, demandas e rotinas de trabalho reconhecidos. Nessa linha, a criação dos documentos deve se dar no contexto de produção em suas unidades de trabalho. O projeto então estabeleceu como diretriz o mapeamento dos processos de trabalho da entidade produtora e suas relações com os documentos durante o trâmite.

Os documentos deveriam estar disponíveis para o acesso nestes sistemas, sendo confiáveis e prontos a informar ou provar sobre determinado processo de trabalho ao longo do tempo. Os sistemas informatizados de gerenciamento arquivístico deveriam garantir a produção e manutenção dos documentos visando este objetivo.

O projeto da Universidade de British Columbia discordava do foco para o estabelecimento de requisitos no desenvolvimento de sistemas informatizados de gerenciamento arquivístico. Em vez de estarem voltados para o documento de arquivo como evidência do processo de trabalho, para os idealizadores do projeto UBC, o conceito de documento deveria contemplar outros potenciais usos a longo prazo como os associados a valores legais, fiscais, sociais ou mesmo de prova de ação.

Esses usos poderiam ser mascarados, pois na perspectiva de Pittsburg, há um deslocamento das atividades de criação e manutenção dos documentos das unidades de arquivo para as unidades de trabalho, o que poderia afetar o próprio contexto de produção uma vez que as unidades de trabalho estariam voltadas para o desenvolvimento e execução de suas atividades e não com a gestão de seus documentos.

A UBC enfatizou o uso do protocolo como sistema de registro dos documentos digitais em sistemas eletrônicos para padronizar os documentos desde sua entrada nos sistemas informatizados de gestão arquivística. O documento somente seria criado a partir de seu registro no protocolo,

unidade vinculada ao sistema informatizado de gestão arquivística adotado pela entidade produtora (MARSDEN, 1997, p. 166).

Dessa forma, a UBC procurava conciliar os princípios da Diplomática e os princípios arquivísticos – base epistemológica no projeto – com os modelos que previam a inclusão dos processos de trabalho e os trâmites documentados numa dimensão integrada dos sistemas informatizados de gerenciamento arquivístico de documentos.

A intenção era criar uma familiaridade de procedimento que considerasse as especificidades das entidades produtoras, uma vez que serão elas a utilizar os sistemas. “Documentos são preservados, inventários são feitos, e esquemas de classificação são criados para possibilitar o acesso e o uso dos documentos, para isso os sistemas de gerenciamento eletrônico de arquivos são projetados” (MARSDEN, 1997, p. 166, tradução nossa).

4.6 AS FUNÇÕES ARQUIVÍSTICAS: DIFERENÇAS E APROXIMAÇÕES

A) CLASSIFICAÇÃO

O contexto do documento, tema tão caro no campo da Arquivologia, justifica-se pela relação orgânica entre o documento de arquivo e a função que o originou, um não existe sem o outro. Na maioria das vezes, a função não é exercida em um único ato, há inúmeros processos para o desenvolvimento desta, incluindo inúmeras ações e atividades. Os documentos, além das atividades e das funções, fazem parte da vida da instituição produtora, estão no cotidiano de seus produtores. Por estes usos dos documentos de arquivo, por esta relação mutuamente constitutiva entre o documento e a função é que a classificação funcional se justifica.

Admite-se que toda classificação é arbitrária. De acordo com Nietzsche (1999), toda classificação é a igualação do não igual, pressupõe o apagamento das diferenças e o direcionamento forçoso do olhar para as características que as coisas que estão sendo classificadas possuem em comum, amalgamando-as por esses vínculos.

A teoria da classificação na Ciência da Informação está associada ao paradigma platônico-aristotélico. Na seção anterior abordamos os paradigmas que influenciaram a concepção da linguagem nessa área, estes também se fazem presentes nas diversas práticas como a própria classificação.

O Código de Classificação Decimal de Dewey, o Código de Classificação Universal, elaborado por Otlet e La Fontaine, e a *Colon Classification*, de Ranganathan possuem o mesmo princípio filosófico: procurar estabelecer categorias universais. Por conseguinte, o conhecimento é uma estrutura autônoma, com entidades definidas por propriedades essenciais e relações objetivas entre entidades e categorias.

As propriedades das coisas reproduzem a ordem intrínseca do mundo e podem ser mapeadas e representadas. As subdivisões criadas por essas classificações se apresentam como subdivisões do mundo, naturalmente impostas e reconhecidas pela percepção humana. É possível criar representações universais do conhecimento porque os processos mentais são operações semelhantes em todos os indivíduos, pois possuímos o mesmo aparelho cognitivo. O mundo é uma estrutura completamente autônoma e compreensível a partir da lógica formal que identifica e descreve suas propriedades.

Essa lógica essencialista já abordada anteriormente possibilitou que inúmeros estudos da teoria da classificação em Ciência da Informação trabalhassem com a perspectiva de captar a essência das coisas no mundo, classificá-las para melhor entendê-las. Dahlberg, ao analisar a Organização do Conhecimento enquanto campo científico estabelece que sua organização é um domínio da Biblioteconomia e da Ciência da Informação comprometido em organizar, ordenar e dispor a informação (1993, p. 211). Este pensamento está em consonância com a sua teoria do conceito que estabelece as formas em que os indivíduos organizam suas formas de pensar.

Segundo Dahlberg, os conceitos "identificam as características, os atributos das coisas designadas" (DAHLBERG, 1978, p. 102). O processo de análise dos conceitos somente é possível "coletando-se os enunciados verdadeiros que sobre determinado objeto se podem formular" (DAHLBERG, 1978, p. 102).

Destarte, como nos pressupostos platônicos, Dahlberg indica que nos conceitos há elementos que constituem enunciados verdadeiros sobre as coisas, passíveis de serem identificados e classificados de acordo com suas características naturais. Nesta concepção, existe uma autonomia simbólica dos enunciados que compõem os textos. Os elementos, as características e os atributos vão compor os conceitos, a tessitura bem sucedida dessas combinações é que vão indicar se a representação do real será bem ou mal sucedida. A eficácia dos conceitos está relacionada diretamente à fidelidade de sua descrição do real. Logo, encontramos a influência do pensamento realista de tradição platônica na definição de conceito estabelecida por Dahlberg.

No paradigma realista há a necessidade de estabelecer nas expressões, nos discursos, nos textos, nos conceitos uma objetividade essencialista que está nonexo em que estes elementos estabelecem ou não com o real. Esta filosofia estará presente nos métodos de indexação que estabelecem a tematicidade unicamente pela análise dos documentos. O sentido se encontra no texto e cabe ao indexador identificá-lo. O conteúdo dos documentos descreve a parcela da sua realidade de maneira autônoma e objetiva.

No entanto, a teoria da classificação aplicada aos documentos se desenvolve num primeiro momento no âmbito das bibliotecas públicas cuja missão era dar acesso ao acervo, inclusive considerando a dimensão educacional dessa instituição, sobretudo na era vitoriana, mas ganha

envergadura após a Segunda Guerra Mundial com o desenvolvimento dos centros de documentação e a exigência da circulação de informações e documentos para desenvolver áreas específicas como já abordamos anteriormente.

Vimos que a adoção dessa filosofia da linguagem para a prática da classificação e a exigência da recuperação de documentos pela temática abordada para o desenvolvimento de determinada técnica, produto, processo ou tecnologia, levou ao campo a privilegiar a organização dos documentos, via classificação para a organização do conhecimento, por meio de coleções relativas ao mesmo assunto em detrimento do contexto.

Susan Stewart (1984), ao abordar o conceito de coleção, afirma que ao contrário do *souvenir*, esta se caracteriza pela destruição do contexto do objeto. Podemos transpor essa concepção para a classificação dos documentos com o objetivo de organizar as áreas do conhecimento por meio da formação de coleções a partir do enquadramento dos assuntos, a exemplo do campo da Ciência da Informação que ao longo de sua história tem procurado desenvolver métodos de classificação para atender as demandas da organização do conhecimento. É do estatuto da coleção o esquecimento, o apagamento dos vínculos entre os documentos para a construção de uma nova ordem.

A coleção realiza esse deslocamento para unificar em uma única categoria, em um único assunto os documentos. Este tipo de classificação rompe com o contexto de produção do documento, rompe com o passado para projetar um futuro estabelecido a partir da ordem dada pelo assunto. A classificação por assunto sempre apresenta um universo documental completo, pois os unifica em uma classe, deslocando-os de seus contextos. Procura sempre estabelecer um mínimo de documentos que formem um conjunto compreensível. A ordem imposta pela coleção torna viável o próprio esquema classificatório por categorias.

Nos arquivos, a classificação é uma função que possui uma lógica completamente distinta da ideia de coleção e de apagamento do contexto, aliás, o processo de classificação é de grande importância na organização dos arquivos justamente por seu compromisso em manter o contexto dos documentos desde sua produção e o vínculo entre eles. A condição da existência do documento de arquivo é a manutenção de sua própria organicidade.

De acordo com Cortés Alonso (2008), os documentos de arquivo, resultado da atividade humana e testemunho e informação da vida dos homens, são unidades que apresentam peculiaridades muito concretas, distintas de outras espécies de documentos (bibliográficos, arqueológicos, artísticos) e tais características fazem que seu tratamento deva atendê-las. Por esta razão, a classificação precisa manter as marcas resultantes da atividade pública ou privada que o gerou, que o circunscreve totalmente enquanto a sua configuração e que, atento a essa atividade, o faz semelhante aos que como resultado da mesma se vão produzindo no espaço e no tempo.

Tal condição, dependente dos atos jurídicos, administrativos, comerciais, familiares que os homens vão desenvolvendo através dos anos, dos séculos, faz com que os documentos apresentem outra característica muito particular, a de constituir séries de unidades relacionadas entre si por algum elemento, que as convertem em conjuntos orgânicos facilmente identificáveis e, por tanto, formando unidades complexas de milhares de documentos simples.

Não existe documento de arquivo se o seu contexto não for preservado. A classificação facilita a compreensão das dinâmicas em que os documentos, as ações, os agentes e suas relações participaram. Ao contrário das classificações na Ciência da Informação, não é possível uma ruptura com o passado do documento para o estabelecimento de uma ordem imposta pelo assunto. A função da classificação em arquivos é justamente garantir a rede de mediações que o documento participou ou ainda participa.

Cortés Alonso afirma que "se reunimos os documentos produzidos por um só indivíduo ou instituição ao longo de sua vida ativa, o resultado documental de uma existência longa ou curta, simples ou com grandes variações de atuação, teremos o que chamamos de arquivo" (2008, p. 412). Dessa maneira, o documento de arquivo para se configurar como tal precisa manter a sua organicidade.

Segundo Sousa (2005), a classificação é a função matricial nos arquivos, pois é ela que dá sentido e preserva o caráter orgânico do conjunto documental. Esta perspectiva coaduna com a compreensão de que é preciso preservar o contexto do documento na classificação, pois ela acaba por revelar sua própria condição de mediador de relações sociais.

Para Reed (1997), os documentos são agentes de ação, ativos participantes nos processos de trabalho e que só podem ser descritos se pensarmos nas séries de interações que eles participam. A classificação é a função que garantirá a identificação desses processos de trabalho com os documentos produzidos no seu desenvolvimento.

Autores da Arquivologia, que compartilham desta visão acerca do processo de classificação, advertem que por esse compromisso com o contexto é que em seu processo os princípios da proveniência e da ordem original devem ser respeitados.

Cruz Mundet afirma que "classificar consiste em agrupar hierarquicamente os documentos de um fundo mediante grupos ou classes, desde os mais amplos aos mais específicos, de acordo com os princípios de proveniência e ordem original" (1994, p 239). Lopes (1996) estabelece que os princípios de proveniência e ordem original incidem diretamente sobre a função da classificação, eles fundamentam tal processo e servem para definir estratégias nestes esquemas. Duchein (1986) afirma que a separação dos documentos provenientes de um organismo e a integridade interna de um fundo são os princípios basilares que devem guiar a classificação. Essa manutenção da estrutura interna, chamada de princípio da proveniência, e problematizada por Duchein em sua aplicação na

classificação, corresponde "a manutenção dos traços gerais com a organização da administração da qual o arquivo provém" (MULLER, FEITH, FRUIM apud DUCHEIN, 1986, p. 23).

Considerando as dificuldades de identificar as estruturas internas de certas entidades administrativas sejam pelas mudanças sofridas ao longo de sua existência, seja pela complexidade da própria instituição ou pela própria falta de classificação sistemática baseada em tais princípios, Duchein aponta que a classificação dos documentos a partir das funções que a entidade produtora desenvolvia é a forma mais eficaz de manter a ordem original dos documentos, pois segundo ele nada mais estável que a função por estar associada a própria razão da existência da entidade produtora.

Lodolini (1993) ao abordar a classificação dos documentos de arquivo por séries afirma que esta é uma prática existente na Itália desde o século XVII e consistia em agrupar os documentos por matérias de competência de cada instituição descritas em um quadro chamado *titilario*. Heredia Herrera (1991) diz que o respeito ao princípio da proveniência acontece em dois níveis: o primeiro trata-se da identificação da estrutura ou do funcionamento da entidade produtora; o segundo corresponde a identificação das séries documentais, que de acordo com a mesma, são os testemunhos das atividades realizadas naquela estrutura. Então, a classificação dos documentos de arquivo é a organização hierárquica de séries documentais vinculadas às funções desenvolvidas pela entidade produtora ao longo de sua existência.

De acordo com Foscarini (2009), a prática da classificação de documentos vem da necessidade de tornar explícito o vínculo arquivístico entre os documentos. Este vínculo existe entre todos aqueles registros que participam da mesma atividade, desde o momento da criação destes. Através do ato de classificação, a rede de relações inerentes à natureza de qualquer documento não só é trazida à luz, mas também é estabelecida e perpetuada.

Deste modo, o significado de cada documento em relação a todos os outros, bem como a estrutura do conjunto de documento (isto é, o fundo arquivístico) pode ser compreendida e transmitida ao longo do tempo (FOSCARINI, 2009, p. 2). A autora mencionada defende ainda um cuidadoso estudo sobre as atividades e as funções da entidade produtora dos documentos para a construção do sistema de classificação no arquivo. Entendemos que esta defesa se trata de uma reflexão acerca dos usos que a entidade produtora tem de seus documentos, uma vez que eles estão o tempo inteiro viabilizando suas ações, atividades e funções.

O que distingue um arquivo (no sentido de uma pluralidade de documentos) como um todo orgânico, ou *universitas rerum*, ou seja, uma entidade estruturada, são as circunstâncias contingentes da sua criação, a partir da formação de um mero conjunto ou soma de itens únicos, unidos artificialmente para a realização de quaisquer fins. Este link incremental é original e necessário (conhecido como "vínculo arquivístico"), ele existe entre todos os documentos que pertencem juntos ao mesmo processo, porque eles são originados durante,

e em virtude da mesma atividade ou processo de negócios (FOSCARINI, 2009, p.2, tradução nossa)

A classificação nos arquivos serve, sobretudo, para a instituição produtora compreender seu universo documental e as ações, atividades e funções amalgamadas por eles. O compromisso da classificação dos documentos de arquivo não é com o conteúdo destes, em sua unidade, é com os vínculos arquivísticos, ou seja, com seu registro e fixação em um plano de classificação. Classificar na fase de produção documental, nesta lógica, é um exercício de compreensão da entidade de si mesma, de seu cotidiano, de suas práticas que justificam o seu ser e o seu estar no mundo.

Portanto, a classificação funcional para os documentos de arquivo justifica-se por estar baseada nos seus usos pelo seu produtor. A compreensão das funções, das atividades, das ações e dos documentos, ou seja, da *práxis* da entidade produtora que motiva a operação da classificação funcional.

A classificação evidencia a rede de relações mediadas por qualquer documento e a perpetua. Ao classificar, a instituição produtora não aproxima documentos semelhantes e separa os diferentes, como a interpretação representacionista da classificação tenta sugerir, a classificação constitui-se como uma tentativa do produtor compreender melhor a si mesmo e suas práticas dentro da sociedade e do ordenamento jurídico da qual faz parte.

Foscarini (2009) e Sabourin (2001) defendem que a classificação mais eficaz nos arquivos é a funcional, pois além de facilitar o respeito aos dois princípios que são os pilares da Arquivologia, o da proveniência e o da ordem original, ela permite a compreensão do universo documental, da estrutura da entidade e suas ações ao longo do tempo.

Schellenberg ao abordar o processo de classificação nos arquivos correntes, afirma que tal função arquivística é uma condição para uma eficiente administração de documentos modernos e que ela deve estar baseada nos usos que seus produtores fazem dos documentos.

Todos os outros aspectos de um programa que vise ao controle de documentos dependem da classificação. Se os documentos são adequadamente classificados, atenderão bem às necessidades das operações correntes. E, para tanto, devem ser arranjados em função do uso que têm em determinadas unidades administrativas de um órgão. Em todos os casos, desde os relativos a importantes assuntos de política até os de operações de rotina, deverão ser agrupados em relação ao seu uso (SCHELLENBERG, 2006, p. 83).

Schellenberg ao identificar três tipos de classificação mais utilizados nos arquivos públicos americanos - o funcional, o estrutural e por assuntos - defende que a melhor forma é a funcional pela estabilidade das funções devido a própria existência da entidade produtora justificar-se pelo exercício dessas.

Para o arquivista americano, as entidades podem modificar sua organização, no entanto, suas funções continuarão as mesmas. Já as classificações estruturais, também conhecidas como organizacionais, podem ser utilizadas para facilitar a compreensão da entidade de seu universo documental. "A estrutura que se imprime a um órgão é determinada, em geral, pelos objetivos ou funções a que se destina. Assim, a organização, frequentemente corresponde à função" (SCHELLENBERG, 2006, p. 86). No entanto, o mesmo adverte que tal forma de classificar deve ser utilizada apenas quando as entidades produtoras possuem estruturas estáveis com funções e processos de trabalho bem definidos.

A classificação estrutural, chamada por Cruz Mundet (1994) de classificação orgânica, agrupa os documentos em séries de acordo com as diferentes divisões administrativas ou estrutura orgânica da entidade, reproduzindo seus serviços, seções, unidades. Somente a partir deste nível é possível estabelecer uma classificação estrutural, a qual se realizará hierarquicamente desde as unidades administrativas básicas até as divisões mais amplas (1994, p. 241).

A classificação mais criticada pelos autores e que deve ser evitada pelos arquivos é a por assuntos, justamente por forçar um apagamento do contexto, dos vínculos entre os documentos e de tudo o que caracteriza a especificidade do documento de arquivo. Segundo Schellenberg, os tipos de documentos que possuem apenas usos referenciais e que tenham perdido a sua organicidade podem ser classificados dessa forma (2006, p. 93).

Para Sousa (2014), o documento de arquivo é um artefato humano, cujo conteúdo é delimitado pelo sujeito acumulador. Compreender a forma em que as instituições se estruturam, suas missões, funções e atividades, é entender como os documentos são acumulados, segundo o autor. Dessa maneira, as classificações arquivísticas capazes de evidenciar essas relações facilitam a própria compreensão do documento de arquivo em sua gênese. A identificação arquivística apresenta-se como um importante método nesse sentido, permitindo mapear as funções, as atividades e os tipos documentais da entidade produtora.

Destarte, podemos perceber que a classificação nos arquivos, compreendida como o esforço de organização da entidade produtora de seu universo documental, procura atender em seu processo às necessidades de uso do documento na dinâmica das ações, atividades e funções da entidade produtora. Ao classificar, enquadrar, criar hierarquias e escalonamentos dos seus documentos, das suas funções e das suas atividades, a entidade produtora procura melhor compreender a si mesma e viabilizar suas ações em seu cotidiano, de forma eficiente e eficaz, como aconselha Schellenberg.

O objetivo da classificação é organizar, criar uma compreensão por parte da entidade produtora do todo de seus documentos. Transpondo a alegoria de Walter Benjamin (1987, p. 228) para compreender a classificação, os esquemas classificatórios nos arquivos funcionam como

diques, tentativas de barragem para “segurar a enxurrada” de documentos produzidos a fim de que ações se desenvolvam.

As características elencadas nesta seção revelam-se fundamentais para compreender sua diferença em relação a indexação. Se na Ciência da Informação, a indexação possui uma trajetória muito próxima à teoria da classificação como vimos na segunda seção deste trabalho, a classificação arquivística apresenta especificidades que precisam ser consideradas e que, de certa forma, possui sinais diacríticos em relação à indexação.

A classificação assegura o contexto de produção do documento de arquivo. A indexação pode até viabilizar a recuperação do contexto de produção do documento, mas não é o seu objetivo primeiro. Esta especificidade da indexação nos arquivos em fase da produção documental veremos na última seção dessa tese.

B) DESCRIÇÃO

A Arquivologia para atender às demandas informacionais dos usuários de arquivos no que se refere à recuperação dos documentos desenvolveu ao longo dos séculos a descrição como operação arquivística capaz de realizar tal representação e possibilitar o acesso aos documentos de arquivo.

De acordo com Yeo (2013), a descrição arquivística é um processo intelectual de representação dos documentos de arquivo que tem como finalidade o controle e o acesso aos mesmos. Seu objetivo principal é tornar aparentes os relacionamentos entre o contexto de produção e os documentos em si, além da própria compreensão do arquivo.

A descrição trata de uma complexidade maior, de um todo. Oliveira (2010) afirma que a descrição é uma representação formulada pelo arquivista e a sua formulação envolve processos e métodos de pesquisa complexos e particulares que convergem para a compreensão de um arquivo. Um elemento fundamental para essa compreensão é a reconstrução do contexto arquivístico (OLIVEIRA, 2010, p. 46). A elaboração de mecanismos de acesso; a explicitação dos relacionamentos entre os documentos de arquivo e as funções e atividades que lhes deram origem; também, o controle do acervo são elencadas como atividades arquivísticas inseridas no processo de descrição.

A preocupação com o contexto do documento de arquivo, especificidade da descrição enquanto operação de representação da informação, atende a um dos dois níveis presentes nos arquivos, relatados por Rousseau e Couture: a informação do conjunto documental que revela as ações, atividades e funções das atividades produtoras. De acordo com os autores canadenses, há outro nível de informação presente nos arquivos, que é a informação contida no documento em si, isoladamente, relacionada mais ao conteúdo do que ao contexto (ROUSSEAU, COUTURE, 1998, p. 251), a esta o processo de descrição pouco atende.

Segundo Yeo (2013), a descrição é tanto um processo quanto um produto. Um processo, pois trata-se da função arquivística de representação do documento de arquivo e um produto por gerar instrumentos de pesquisa que inclusive auxiliam na gestão dos conjuntos documentais, a exemplo dos inventários que elencam os documentos detectando possíveis perdas ou extravios. O autor enfatiza que a descrição, acima de tudo, reúne informações sobre o contexto e por essa característica torna-se um elemento essencial no controle dos documentos de arquivo.

Jenkinson, ao abordar a importância do contexto no desenvolvimento da descrição, afirma que o arranjo ao representar a cadeia de custódia, as formas de criação dos documentos e seus usos realiza uma "defesa moral" dos arquivos por proteger a autenticidade dos registros. (JENKINSON, 1937, p. 120).

Na Arquivologia há duas abordagens majoritárias da função arquivística descrição: uma que privilegia o contexto pela compreensão dos documentos de arquivo como presunção de prova, evidência das ações que a entidade produtora realizou; e outra que enfatiza o acesso e a localização dos documentos visando o seu uso informacional.

Segundo Yeo (2016), os profissionais que privilegiam o usuário enxergam a descrição arquivística como análoga à catalogação feita pelas bibliotecas. Para os profissionais que privilegiam os documentos, o material de arquivo não faz parte do universo bibliográfico, os documentos e os arquivos não são meras instâncias de conteúdo informativo descritíveis por meio de técnicas semibibliográficas, mas representações de atividades que ocorrem no mundo real, "átomos de atividades em que eles cumpriram seus papéis indispensáveis"(YEO, 2016 p. 136), intimamente ligados às transações que os engendram.

No entanto, os partidários de ambas as correntes são unânimes na compreensão de que a descrição tem o compromisso de representar o contexto e os vínculos do documento de arquivo para atender o princípio da proveniência. A temática e os assuntos presentes nos conteúdos dos documentos não são objetos da descrição

O MAD, as RAD, a ISAD(G) e a NOBRADE orientam a realização de descrições retrospectivas, ou seja, processos que se dão a partir dos arquivos permanentes e chegam aos processos de custódia e produção.

A descrição se ocupa do hercúleo trabalho de descrever universos complexos, com perspectivas singulares ou múltiplas, procurando representar contextos muitas vezes não delimitáveis e procurando mapear mudanças das instituições que estão em permanente transformação, sobretudo em relação aos suportes que configuram seus documentos e sistemas por onde estes circulam na realidade da obsolescência da tecnologia e na fluidez do mundo contemporâneo (YEO, 2016, P. 147).

De acordo com Yeo (2016), pela descrição ser um processo e um produto de seu tempo e lugar, a forma mais eficaz de desenvolvê-la é considerando seus usos e as demandas dos usuários, procurando perceber as formas que estes interagem com as descrições e com os documentos de arquivo que pesquisam. “Deveríamos reconhecer que diferentes comunidades de interesse lançam diferentes preocupações e percepções sobre os mesmos objetos ou sobre grupos imbricados de objetos, e o desafio será desenvolver sistemas que efetivamente acomodem vários quadros de referência” (YEO, 2016, p. 156).

Autores do campo da Arquivística como Cruz Mundet (1994) e Alberch I Fugueras (2001) compreendem a indexação como uma operação que faz parte da complexidade da função da descrição. De acordo com o primeiro autor, a indexação facilita a compreensão dos instrumentos de descrição que muitas vezes se apresentam demasiadamente abstratos e pouco úteis para os usuários de arquivo, no que se refere a matérias de seus interesses. Justamente pelo compromisso com os princípios da proveniência e da ordem original, frequentemente a descrição revela-se pouco eficaz na recuperação da informação, de acordo com Cruz Mundet.

Demasiado ocupados com a proveniência, com a classificação que desde logo constituem outros pontos de acesso à informação, os arquivos chegaram a afastar-se perigosamente de sua verdadeira razão de ser: os usuários. Por sorte esta atitude está se retificando nos últimos anos e prova disso é o papel crescente que a indexação tem entre os instrumentos de descrição (CRUZ MUNDET, 1994, p. 291, tradução nossa).

Cruz Mundet qualifica a indexação como um processo de representação do conteúdo do documento, baseado na construção de linguagens controladas que estabelecerão termos normalizados a fim de sintetizar as informações, facilitando a recuperação do documento pelo assunto de interesse do usuário (1994, p. 292).

Ele estabelece três etapas para o desenvolvimento da indexação a serem adaptadas no tratamento documental nos arquivos: 1 - Análise do documento e determinação de seu conteúdo, prestando atenção aos seguintes elementos: emissor (ator), receptor, canal (procedimento), mensagem (assunto) e contexto, fazendo alusão aqui à teoria do ato comunicativo de Jakobson (1974); 2 - Identificação e seleção dos conceitos/assuntos principais, extraídos do documento mediante análise intelectual e transformados posteriormente em termos de indexação; 3 - Seleção dos termos de indexação que expressam os conceitos dos documentos, mediante termos contidos neles ou através de outros que sejam representativos, baseados em uma linguagem controlada predefinida (CRUZ MUNDET, 1994, p. 292).

Para Cruz Mundet, o indexador precisa extrair do documento os principais termos e conceitos empregados a fim de simplificar o processo, diminuir a subjetividade do analista e, por conseguinte, possibilitar a recuperação da informação de forma mais precisa.

O autor mostra a sua filiação com os estudos de indexação que possuem inspiração filosófica platônica-aristotélica, partilhando da crença de que o sentido do documento se encontra nele mesmo, é autônomo e atemporal, cabendo ao indexador apenas identificá-lo e representá-lo a partir de categorias universais estabelecidas em uma linguagem padronizada.

Apesar de apontar para as dificuldades da recuperação da informação pela descrição, uma vez que não está voltada para os usos dos documentos de arquivo, apresenta como solução uma forma de indexação que tampouco considera ou investiga tais usos, conforme analisamos na seção anterior. No entanto, destaca-se o esforço de Cruz Mundet em pensar a recuperação da informação nos arquivos.

Alberch I Fugueras (2001), ao preocupar-se com a gestão integrada dos documentos, sobretudo baseada nas normas ISO 15489, de Gestão de Documentos, e a ISO 30300, de Sistemas de Gestão para os documentos, apresenta também a indexação como uma tarefa subordinada à descrição.

A descrição, portanto, é uma função normatizada na área dos arquivos e os critérios estabelecidos pelas normas, não somente as ISOS, mas também a ISAD-G e a ISAAR-CPF, permitem compreender a configuração dos documentos e seus processos de trabalho. Auxilia, dessa forma, a compreensão do contexto em que foi criado o documento de arquivo (trâmites, pessoas envolvidas, unidades de trabalho implicadas, datas de criação e modificação dos documentos, entre outros procedimentos).

Embora seja um tema periférico em seu trabalho que se preocupa com o sistema de gestão de documentos, sobretudo em relação aos documentos de arquivo e processos de trabalho natos digitais e a importância da padronização dos procedimentos, Alberch I Fugueras apresenta a descrição como uma função arquivística que deve estar presente ainda na fase de produção documental. Do mesmo modo que a classificação e a avaliação, a descrição garante o respeito ao princípio da proveniência, questão de grande importância nos arquivos e de grande desafio no ambiente eletrônico, sobretudo porque não há mais translados físicos nestes ambientes, apenas mudanças de competências.

Todavia, ao analisarmos a norma ISO 15489, tanto a edição de 2001 quanto a de 2006, que regula a gestão de documentos, ao estabelecer os processos e controles da gestão de documentos, estas não elencam a descrição dentre eles. Em seus dez itens: determinação dos documentos que deveriam ser incorporados ao sistema; determinação dos prazos de conservação; incorporação dos documentos; registro; classificação; armazenamento e manipulação; acesso; rastreamento; disposições; documentação dos processos de gestão – não se encontra a função da descrição arquivística. Entretanto, a indexação é tratada como um subitem da classificação, definida como

"processo que estabelece pontos de acesso para facilitar a recuperação de documentos ou informação (2006, p. 8, tradução nossa).

A norma ISO 30300, que define os princípios pertinentes a orientação da implementação de Sistemas de Gestão de Documentos, também não aborda o processo de descrição na fase de produção documental. Aliás, apresenta a classificação e a indexação como processos a serem realizados na fase de registro dos documentos conforme mostra a figura extraída da própria norma ao estabelecer os conceitos relativos aos processos de gestão documental¹⁴.

Já a *International Standard for Archival Description* (ISAD-G) silencia acerca da representação temática da informação dos documentos de arquivo. Não há nenhuma operação prevista neste sentido nos elementos das sete áreas estabelecidas - identificação, contextualização, conteúdo e estrutura, condições de acesso e uso, fontes relacionadas, notas, e controle da descrição.

Diferentemente, a **norma brasileira de descrição arquivística** (Nobrade), adaptação da isad-g, criou mais uma área além daquelas previstas pela norma internacional, chamada de pontos de acesso e indexação, com o objetivo de "registrar os procedimentos para recuperação do conteúdo de determinados elementos de descrição, por meio da geração e elaboração de índices baseados em entradas autorizadas e no controle do vocabulário adotado" (2006, p. 59). A Nobrade orienta que os termos a serem definidos pela indexação precisam de estudo, definição e avaliação periódica da entidade custodiadora e devem ser baseados na política empreendida para recuperação sistêmica das informações e na política de indexação de assuntos.

A indexação dos assuntos deve contemplar nomes de entidades, eventos, áreas geográficas, períodos e assuntos tópicos. Os assuntos tópicos devem ser indexados de acordo com as necessidades, possibilidades e objetivos da entidade custodiadora, levando-se em consideração a metodologia e os resultados distintos obtidos pela indexação pré-coordenada, mais comum em sistemas manuais de recuperação de informação, e pela indexação pós-coordenada, mais comum em sistemas informatizados de recuperação de informação (CONARQ, 2006, p. 59)

Dentre os procedimentos estabelecidos para a realização da indexação pela Nobrade, recomenda-se que os elementos obrigatórios produtor e título devam ser usados na geração de índices. Para a normalização das entradas, a Nobrade indica o uso da norma internacional de registro de autoridade arquivística (isaar - cpf), considerando também o código de classificação anglo-americano (aacr2) e as normas da abnt como fontes de referência.

A Nobrade prevê ainda a construção de um vocabulário controlado para a indexação dos assuntos, procurando usar os termos de fontes institucionais validadoras como a *library of congress* e a biblioteca nacional ou tesouros especializados (brasil, 2006, p. 60). Apesar da inovação feita pela Nobrade com a inclusão da área de indexação, este processo não faz parte dos elementos

¹⁴ O diagrama com as etapas da gestão de documentos previstas pela ISO 30300 está no anexo F.

obrigatórios estabelecidos pela mesma - código de referência, título, data(s), nível de descrição, dimensão e suporte, nome(s) do(s) produtor(es), condições de acesso.

Retomando a reflexão em torno do *continuum*, a descrição arquivística é um processo que abarca agentes e ações, independentemente da idade de seus documentos ou das condições de sua guarda. Portanto, deve estar em todas as fases, unindo a produção à custódia (CUNNINGHAM, 2005).

De acordo com Bearman (1992), a descrição retrospectiva é falha em diversos quesitos: exige muitos recursos, causa enormes gargalos de processamento, não registra informações contextuais importantes, deixa de recorrer a banco de dados existentes aos quais boa parte das informações necessárias poderiam ser obtidas.

Assim, a proposta de descrição na perspectiva dos *records continuum* deve ser realizada desde a fase da produção documental, interagindo com os regimes de controle da gestão de documentos, não havendo a separação que o modelo de ciclo vital consolidou. As informações descritivas podem ser agregadas progressivamente, durante a vida do documento e, de forma ideal, reutilizadas e redefinidas de acordo com os contextos, as circunstâncias e as necessidades da instituição (CUNNINGHAM, 2005).

Segundo Bearman (1992), a transição para o ambiente eletrônico, tanto na criação, quanto na organização e arquivamento dos documentos de arquivo, tem transferido boa parte do enfoque descritivo para o documento elementar e não para a série. Mesmo no nível mais básico de sua decomposição em partes componentes, os documentos podem revelar redes complexas de relações. A indexação pode auxiliar essa ação.

A descrição a partir dos *records continuum* é definida pela interação dos processos que captura e liga os dados de autoridade com os documentos criados no contexto social e dos processos de trabalho no momento de sua criação e durante toda a sua vida (MCKEMMISH, 2005). A representação do contexto e das entidades são associadas às complexas relações entre eles.

Além disso, a concepção da descrição no *continuum* procura ampliar o alcance dessa função não apenas voltando as atenções para os contextos formais e administrativos dos documentos, mas também considera as visões de mundo envolvidas no processo – dos produtores, descritores, usuários cujos sujeitos de dada comunidade podem ocupar todos esses papéis. Essa percepção sócio-cultural mostra a descrição como um processo parcial e circunscrito pela conjuntura em que se está realizando tal representação.

Pluralizar as equiparações entre os documentos criados, os sujeitos, as atividades com o que já foi e o que ainda será produzido, permite que a descrição se insira em formas de totalidades sociais. Além dos contextos organizacionais, é possível considerar os atores sociais, as transações, os trânsitos, os fatos sociais totais.

Entendemos que a descrição não se configura como a única forma de representação dos documentos de arquivo, embora sua importância seja incontestável como foi demonstrado nesta seção. As especificidades de cada comunidade, cada entidade, cada universo documental, cada tempo e lugar precisam ser consideradas na escolha das formas de representação. Reconhecemos a ampliação das possibilidades de elicitación das estruturas dos documentos e dos seus sistemas relacionais com a descrição na perspectiva do *continuum*.

No entanto, ressaltamos que o escopo de nossa análise se encontra em pensar a indexação, como representação dos assuntos dos documentos, pode facilitar os usos de tais documentos nos processos de trabalho pela entidade na fase de produção. Entendemos que a descrição e a indexação quando associadas na representação dos documentos de arquivo podem tornar sua recuperação mais eficiente e eficaz. Todavia, tratam-se de processos de representação autônomos e não precisam, necessariamente, estar subordinados entre si.

Consideramos necessário pensar formas de recuperação da informação que não estejam limitadas ao processo de classificação, tampouco à descrição pelos motivos já mencionados. Por estas razões, a indexação pode até ser incorporada no processo de descrição arquivística, no entanto, não está vinculada a esta função. A indexação e a descrição são processos que independem um do outro, embora possam estar associados, inclusive na perspectiva do *records continuum*.

C) INDEXAÇÃO – UM CASO À PARTE

Elencamos aqui a indexação para facilitar a compreensão de suas diferenças em relação às funções arquivísticas aqui abordadas. É importante salientar que a indexação não se trata de uma função arquivística, tanto pela historicidade dessa operação ao longo de sua prática nas unidades de informação, quanto pela formação de seu *corpus* teórico como vimos anteriormente. Por essas questões, o trabalho se propõe a investigar o processo de indexação nos arquivos, mas consideramos equivocado pensar em uma indexação arquivística, ou seja, uma função com matrizes no campo da Arquivologia tal como a classificação e a descrição.

A indexação é um processo que surgiu partir da elaboração de índices num período em que não estava tão demarcada a separação entre representação descritiva e representação temática (SILVA; FUJITA, 2004). Um contexto demarcado pelo grande volume documental e a necessidade de elaborar instrumentos que recuperassem rapidamente os mesmos para o desenvolvimento de áreas estratégicas. Genealogia abordada na segunda seção desse trabalho.

Portanto, a indexação pode ser considerada um processo que surgiu da necessidade de recuperar conteúdos específicos dos documentos para viabilizar diversas atividades de seus requerentes. Sua trajetória não está associada à Arquivologia como área do conhecimento, embora

sua prática tenha sido realizada nos arquivos desde o aparecimento dos instrumentos de pesquisa nos arquivos permanentes com a elaboração de índices.

Apontamos a indexação como processo capaz de representar os documentos de arquivo a partir da análise conceitual e representação dos seus assuntos, com fins de viabilizar a recuperação futura de tais documentos.

Desse modo, para se realizar o processo de indexação, que segundo Beghtol (1986), trata-se de estabelecer os assuntos dos documentos e sistematizá-los em linguagem natural ou controlada, é preciso também considerar o contexto de produção e uso da informação. Em sua fase de produção, os documentos estão em função primordialmente da administração, de seus próprios produtores, para servir ao trâmite dos assuntos e das ações do cotidiano. A indexação deve servir para este fim, isto é, localizar os documentos e as informações que se solicitam com o fim de utilizar seus dados ou certificar os atos que contém (CORTÉS ALONSO, 2011).

As poucas abordagens acerca da indexação na Arquivologia possuem certa semelhança com a corrente tratada por Guimarães (2003) como catalogação de assunto (*subject cataloguing*). A representação temática é vista como uma consequência da representação descritiva. Se no campo da Ciência da Informação, essa abordagem é compreendida como a gênese dos estudos acerca da indexação, na Arquivologia ela ainda é predominante.

A indexação se apresenta como alternativa de recuperar o conteúdo dos documentos a partir da análise centrada nesses, construindo linguagens documentárias, linguagens padronizadas, linguagens controladas – sinônimos dos instrumentos criados com o objetivo de sintetizar e representar o conteúdo dos documentos.

A partir da análise realizada nos periódicos do campo da Arquivologia, especificamente o *Archivaria* e o *The American Archivist*, percebemos que os estudos sobre a gestão de documentos assemelham-se no que diz respeito à preocupação com a integração dos princípios arquivísticos às novas tecnologias, sobretudo no que se refere aos usos de softwares de gestão eletrônica de documentos, a construção de sistemas informatizados de arquivo e a preservação de documentos, todavia, em ambas as abordagens acerca da representação temática dos documentos é escassa.

Uma das principais preocupações contemporâneas têm sido a questão da segurança, sigilo e preservação das características jurídicas dos documentos de arquivo armazenados em nuvens (*cloud computing*). Transpondo o conceito das ciências políticas de Giddens (1997), consideramos que se trata dos efeitos da chamada “modernidade tardia” na qual os imperativos tecnológicos e as novas configurações políticas, sociais, culturais, relativas inclusive à comunicação, têm levado aos sujeitos a procurar respostas na *Ciência* para preencher a ausência e o apagamento das tradições.

Para Giddens (1997), a individuação dos sujeitos, característica de nossa contemporaneidade, e as rupturas e fragmentações das instituições que eram pilares da sociedade

moderna têm levado a reflexão da nova realidade e do futuro. Entretanto, as incertezas surgidas pela volatilidade das mudanças no mundo contemporâneo em diversos campos, muitas vezes fazem aliar reflexões sobre as novas configurações com tentativas de preservar ou manter antigas tradições revisitadas.

Os artigos que podemos analisar em *Archivaria* e *The American Archivist* possuem essa característica: a inquietação com as transformações do campo a partir das novas tecnologias e a aflição em preservar princípios que consolidaram a Arquivologia e suas práticas. Revelam-se como tentativas de manter a confiança e eliminar riscos da desconstrução de um campo do conhecimento, procurando uma segurança ontológica a partir da reinvenção de tradições. As novas configurações e os novos usos dos documentos de arquivo mediados pelas tecnologias e localizados em ambientes digitais, geralmente, são apropriados a partir de princípios e conceitos tradicionais. O exercício teórico é mais descritivo, cuja novidade encontra-se na inclusão de novos termos, sobretudo, relacionados às novas tecnologias, pela linguagem científica.

Destarte, por considerarmos que a indexação é um processo que faz parte da dinâmica de trabalho dos sistemas de recuperação da informação dos arquivos de universidade, sobretudo em sua fase de produção, no entanto que tem sido pouco problematizado pelos campos da Ciência da Informação e da Arquivologia.

5 PROCEDIMENTOS METODOLÓGICOS

A produção desta tese procurou escolher os instrumentos metodológicos mais adequados para a problematização do processo de indexação nos sistemas de arquivos das universidades na fase de produção documental.

Nesta seção descreveremos a estrutura da pesquisa a partir das seguintes etapas: levantamento bibliográfico e revisão literária; legislação e regulamentações e campo empírico. Embora esta pesquisa seja de natureza teórica, a sua temática principal mister uma análise das suas aplicações, por isso escolhemos os protocolos da Universidade Federal Fluminense como campo empírico, com vistas de aprofundar a discussão sobre indexação, recuperação e organização da informação no âmbito da Ciência da Informação e na sua interlocução com a Arquivologia.

Nosso intuito é analisar também a *práxis* da indexação, a partir de uma discussão teórica, considerando as especificidades dos documentos de arquivo, as universidades, sobretudo as federais, como entidades produtoras, os usos dos documentos no contexto de produção e as demandas informacionais por estas comunidades discursivas. É a partir desse detalhamento que iremos analisar o processo de indexação.

5.1 PESQUISA BIBLIOGRÁFICA

Realizamos preliminarmente uma ampla pesquisa bibliográfico em duas bases de dados – na *Library and Information Science Abstracts (LISA)* e na **Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI)**. Procuramos identificar as produções que abordassem a indexação, a indexação em arquivos, os usos da informação e dos documentos, a gestão de documentos e a fase de produção documental nos arquivos.

Essas categorias operatórias foram fundamentais para o desenvolvimento da tese. Neste levantamento preliminar na LISA, ao combinarmos a palavra “indexação” (*indexing*) com “arquivo” (*archive*), encontramos 706 publicações. A busca se deu tanto nos campos “título”, “resumo” e “palavras-chave”. Cerca de 70% das publicações relatam estudos de caso onde os autores analisam os sistemas de indexação automática nos arquivos, sobretudo a partir da década de 1980.

Encontramos relatos abordando os arquivos de Israel, Austrália, Canadá, Inglaterra, Itália, África do Sul e muitos sobre os sistemas utilizados em diversos arquivos dos Estados Unidos. Procuramos refinar a pesquisa, realizando uma busca que, além de combinar as palavras “indexação” e “arquivo”, introduzimos o termo “university”, aplicando o mesmo método de busca. Conseguimos recuperar apenas 24 publicações a partir de seus resumos acerca de indexação em arquivos de universidade. Destes, 14 tratavam-se de catálogos de coleções, o que não possui nenhuma conexão com o nosso trabalho, 8 abordavam bases de dados e a indexação de arquivos

digitais em bibliotecas, os outros 2 artigos tratavam-se de propostas para a indexação de fotografias nos arquivos de universidade, produções que também fogem ao escopo de nossa pesquisa.

Diante desse quadro, procuramos refinar qualitativamente o nosso levantamento bibliográfico, mapeando as publicações acerca da indexação nos periódicos *The Indexer*, *Journal of Documentation* e *Knowledge Organization*, por serem referências consolidadas no campo da Ciência da Informação. Além disso, realizamos o mesmo exercício no campo da Arquivologia com o *The American Archivist* e *Archivaria*.

The Indexer, autointitulado o periódico internacional de indexação, em seu editorial afirma ser um espaço para a publicação de artigos que abordem aspectos da indexação em seus mais variados matizes, relatos de aplicação do processo e análises acerca da organização e acessibilidade do conhecimento e ideias registradas em documentos. Editado desde 1958, o periódico possui 35 volumes e 140 números até o final de 2017.

O *Journal of Documentation* é o periódico mais antigo na área da Ciência da Informação, publicado desde 1945, consolidou-se por sua frequência em divulgar teorias, filosofias, conceitos, modelos sobre documentação, organização do conhecimento, indexação, classificação e outros estudos no mesmo campo. Este periódico, publicado periodicamente, possui 73 volumes e 327 números.

Tradicionalmente, o JD publicou importantes estudos sobre indexação como os de Vickery e Foskett nos anos de 1950 e 1960. Os boletins do CRG também foram reproduzidos ao longo dos anos, sendo o último de número 12 publicado em 1985, igualmente analisados neste trabalho. Essas teorias sobre indexação possuem em comum o fato de pensar formas da padronização da linguagem na entrada dos documentos em um sistema de recuperação da informação.

Neste período, os computadores estavam sendo inseridos nos sistemas de recuperação da informação (SRI). Portanto, os estudos da indexação visavam buscar maneiras de padronização da linguagem em uma linguagem artificial, compreensível para os SRIs, que facilitassem a recuperação dos documentos de tais bases de dados.

Essa busca aproximou os estudos acerca da indexação dos estudos de linguagem. Por esses indícios, fomos também investigar a partir da literatura no campo dos estudos de linguagem, sobretudo na semântica, as principais filosofias que influenciaram as teorias acerca da indexação. Este desdobramento encontra-se na terceira seção desta qualificação. Dessa forma, neste percurso sentimos a necessidade de compreender a semântica formal de inspiração fregeana; o gerativismo chomskiano e o cognitivismo como o de Lakoff e Johnson (2002).

Já o *Knowledge Organization* é a publicação oficial da *International Society for Knowledge Organization* e assim como o *Journal of Documentation* publica pesquisas sobre teorias,

fundamentos filosóficos do conhecimento, operações práticas associadas à indexação e à classificação, além de problemas de terminologia.

Fundado em 1976, possui 4 volumes e mais de 155 números. O que encontramos em comum nos artigos examinados que está presente desde a fundação do periódico, é a abordagem acerca da organização e representação do conhecimento. Diferentemente do JD que privilegia o tratamento temático da informação e o comportamento de uso do documento, as preocupações das publicações no KO giram entorno de elicitare formas do estabelecimento de categorias do conhecimento, mapas conceituais, ontologias, esquemas classificatórios. Logo, processos como a indexação apresentam-se de forma secundária, no sentido de estarem inseridos nas discussões, teorias e problematizações da própria Organização do Conhecimento enquanto área. Exploramos as diferenças entre as publicações também na seção 3 deste trabalho.

Com o objetivo de compreender como o campo arquivístico tem abordado os temas como gestão de documentos, protocolos, arquivos correntes, os usos dos documentos de arquivo e a indexação, realizamos um levantamento preliminar em dois periódicos científicos consolidados na área: *Archivaria* e *The American Archivist*.

O periódico *Archivaria*, produzido pela *Association of Canadian Archivists* (ACA), publicação destinada a comunicar o conhecimento produzido na área dos arquivos, iniciou seus trabalhos em 1975, possuindo 83 revistas publicadas até o presente momento. *The American Archivist*, publicação editada pela *Society of American Archivists*, desde 1938, tornou-se ao longo de sua existência um dos periódicos científicos mais respeitados na área Arquivologia. Até o final de 2017, a publicação possuía 276 números publicados.

O levantamento preliminar que realizamos nos dois periódicos consolidados na Arquivística teve como objetivo perceber como o campo tem abordado as relações existentes entre a gestão de documentos, os protocolos, os arquivos correntes e universitários, além de perceber como a recuperação da informação e a indexação são tratadas pelos autores da área. Constatamos que a gestão de documentos pela quantidade de artigos produzidos, desde a década de 1940, ocupa um lugar de destaque na Arquivística com inúmeros artigos produzidos com múltiplas abordagens. Em um universo de 347 publicações, unindo os dois periódicos científicos, há 171 artigos abordando o tema. Outro número importante são os 128 artigos sobre avaliação. Nossa desconfiância de que a avaliação é um tema central na discussão sobre gestão de documentos

O pouco volume de artigos encontrado acerca da indexação em arquivos frente ao número de publicações nos periódicos da área de Arquivologia, levou-nos a realizar uma comparação com a descrição, pois esta é a função arquivística que realiza a representação do documento de arquivo. No *Archivaria*, encontramos 500 artigos, utilizando o termo “*description*”. E no *The American Archivist* foram encontrados 2.209 de artigos.

Todos esses números são importantes para compreendermos os campos da Ciência da Informação e da Arquivologia e a sua constituição a partir da reflexão do lugar da indexação nos arquivos de universidade. Esse resultado nos permite inferir que os arquivos, sobretudo na fase corrente, enquanto unidades de informação, receberam pouca atenção da área da CI. Em contrapartida, a indexação pouco tem sido pensada como processo de representação do conteúdo dos documentos de arquivo capaz de possibilitar sua recuperação eficiente e eficaz por parte da Arquivologia. Essas abordagens aparecerão na seção 4 que propõe a reflexão sobre a especificidade do documento de arquivo e a diferença entre a indexação e as funções arquivísticas.

5.2 LEGISLAÇÃO

Em relação às unidades de informação, delimitamos a nossa análise do processo de indexação nos arquivos de universidade na fase de produção documental, em especial os protocolos das instituições federais de ensino superior no Brasil. Essa escolha se deu por tais instituições possuem as mesmas funções, estarem subordinadas ao mesmo ordenamento jurídico e político e possuem tipos documentais muito semelhantes.

Então, analisamos a legislação em que as instituições federais de ensino superior e suas unidades de informação estão subordinadas. Faz parte de nosso estudo a lei 8.159, de 08 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados; a lei 12.527, de 18 de novembro de 2011, que regula o acesso às informações; a lei n.º 9.784, de 29 de janeiro de 1999, que regula o processo administrativo no âmbito da Administração Pública Federal.

Analisamos também os decretos como o n.º 7.724, de 16 de maio de 2012, que regulamenta a lei n.º 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição; o de n.º 4.915, de 12 de dezembro de 2003, que dispõe sobre o Sistema de Gestão de Documentos de Arquivo - SIGA, da Administração Pública Federal; o decreto de 15 de setembro de 2011, que institui o Plano de Ação Nacional sobre o Governo Aberto; além do de n.º 4.073, que regulamenta a lei n.º 8.159, de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados; além do decreto 8.539 de 8 de outubro de 2015, que regula o processo administrativo eletrônico nacional.

Já as portarias do Governo Federal que fazem parte da seleção de nossas fontes são: a Portaria Normativa n.º. 05, de 19 de dezembro de 2002, do Ministério da Fazenda, que dispõe sobre os procedimentos gerais para utilização dos serviços de protocolo, no âmbito da Administração Pública Federal, para os órgãos e entidades integrantes do Sistema de Serviços Gerais – SISG; a portaria interministerial MJ/MP 2.320, de 30 de dezembro de 2014, que institui o Sistema Protocolo Integrado.

As deliberações do Conselho Nacional de Arquivos foram observadas a partir das resoluções: nº 38, de 9 de julho de 2013, que dispõe sobre a adoção das **Diretrizes do Produtor - A Elaboração e a Manutenção de Materiais Digitais: Diretrizes Para Indivíduos e Diretrizes do Preservador - A Preservação de Documentos Arquivísticos digitais: Diretrizes para Organizações**; nº 25, de 27 de abril de 2007, que dispõe sobre a adoção do **Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos - e-ARQ Brasil** pelos órgãos e entidades integrantes do Sistema Nacional de Arquivos.

Outras fontes também foram importantes para compreender os documentos produzidos pela entidade e a sua própria estrutura organizacional como o Estatuto da Universidade Federal Fluminense e seu Regulamento Geral, o Regulamento dos cursos de graduação da UFF, além das portarias da Universidade Federal Fluminense nº. 27.089, de 29 de abril de 1999 que designa Comissão para proceder a revisão do Manual de Atos e Comunicações Oficiais da UFF; além da portaria nº 44.482, de 18 de abril de 2011, que versa sobre a reestruturação administrativa da Superintendência de Documentação.

5.3 DICIONÁRIOS, NORMAS E MANUAIS

Destacamos os dicionários terminológicos da área como: o *Glossary of Archival and Records Terminology*, da Sociedade Americana de Arquivistas; o *Multilingual Archival Terminology*, editado pelo Conselho Internacional de Arquivos (ICA); o *Interpares Terminology Database*; o *Diccionario de Terminología Archivística*, dos arquivos estatais do Governo de Espanha, além do *Banco Terminológico*, do Arquivo Geral da Nação Colômbia. No Brasil, utilizamos o **Dicionário Brasileiro de Terminologia Arquivística**, do Arquivo Nacional, e o **Dicionário de Terminologia Arquivística**, da Associação de Arquivistas Brasileiros. Recorremos a estes para entendermos como a Arquivologia tem consolidado termos como arquivo; documento; protocolo; gestão de documentos; indexação e descrição.

As normas internacionais de descrição arquivística (mad, rad, isad-g), também foram importantes para identificarmos as diferenças entre o processo de descrição e de indexação, tal como a Nobrade, esta se difere das demais por ter uma área dedicada à indexação e pontos de acesso.

As diretrizes estabelecidas por algumas entidades referentes às rotinas dos protocolos inseridas na gestão de documentos e nos sistemas de informação foram mapeadas por meio dos manuais de algumas delas. Realizamos um levantamento primeiramente em relação às universidades federais, depois das universidades que possuem programas consolidados e por fim, de órgãos da administração pública. Analisamos os Manuais de Gestão de Documentos, da Universidade de Brasília (2015), da Universidade Federal de São Carlos (2015) e da Universidade

Federal de Minas Gerais (2013), o Manual de Padronização dos Processos Administrativos, da Universidade Federal de Santa Maria (2014), o Manual de procedimentos e normas de protocolo e arquivo, da Universidade Federal do Espírito Santo (2012), e o Manual de protocolo, tramitação e arquivamento de processos, da Universidade Federal do Pará (2000).

A Universidade Federal Fluminense não possui manuais em relação à gestão de documentos, tampouco de procedimentos de protocolo. No entanto, em 1984 a UFF participou, em sua terceira fase, do programa do Ministério da Educação (MEC) para o desenvolvimento e padronização de métodos de organização de procedimentos nas universidades, financiado pelo Banco Internacional de Desenvolvimento (BID).

Os manuais e relatórios produzidos pela Universidade no escopo deste programa que serviram de diretrizes para a implementação do sistema de arquivos da UFF, fazem parte de nossas fontes. Analisamos também a apostila criada pelo Núcleo de Documentação para treinamento das atividades de protocolo aos seus funcionários, de 2010, além do **Manual de atos e comunicações oficiais** (2003). O **Manual do Sistema de Controle de Processos** que instrui a quatro perfis de usuário - gestor de processos, autuador de processos, tramitador e pesquisador - sobre as formas de operar com o próprio sistema SCP e o portal NX, desenvolvido pela Superintendência de Tecnologia da Informação para a gerência de aplicações no ambiente eletrônico, que ainda estão sendo utilizados pelo Sistema Eletrônico de Informações não ter sido implementado completamente, também foram observados nesta pesquisa.

Outros manuais como o do Sistema de Integrado de Administração Acadêmica, o do Sistema de Organogramas da UFF, além de suas relações com as atividades realizadas pelos protocolos da Universidade também serão alvos de nossa pesquisa.

Outrossim, a Universidade Federal do Rio de Janeiro e a Universidade Federal do Estado do Rio de Janeiro também não possuem tais manuais, por esta mesma razão analisamos também seus Manuais de Atos, ambos publicados em 2016. O campus de Macaé da UFRJ lançou um **Manual de procedimentos de protocolo**, que embora não seja adotado por toda universidade também se encontra no escopo de nossa análise.

Além desses, analisamos também a publicação **Gestão documental: procedimentos de protocolo**, produzido pelo Sistema de Arquivos da Universidade de São Paulo (2010); o **Manual de Gestão de processos e expedientes**, da Universidade de Campinas (2012); o **Manual de protocolo e normas processuais, da Universidade Estadual de São Paulo** (2015).

No âmbito da Administração Pública, como já foi dito anteriormente, o Modelo de Requisitos para Sistemas Informatizados de Arquivo, do Conselho Nacional de Arquivos (2011), além da apostila **Gestão de documentos**, elaborada pelo Arquivo Nacional (2011) para o curso de capacitação para os integrantes do Sistema de Gestão de Documentos de Arquivo fazem parte de

nossa análise. O **Manual de Gestão de Documentos e a Cartilha de Gestão de Protocolo**, elaborados pelo Arquivo Público do Estado do Rio de Janeiro (2012), que estabelecem as principais diretrizes para os órgãos do estado, fazem parte do escopo da pesquisa. O **Manual de normas e procedimentos de protocolo para a administração pública do estado de São Paulo**, produzido pelo Arquivo Público do Estado de São Paulo (2013), foi analisado também. Essas fontes serviram não somente para estabelecer comparações, mas para realizarmos um levantamento sobre quais as rotinas de protocolo são consideradas como prioritárias por estas instituições.

5.4 O CAMPO EMPÍRICO

Essa pesquisa estabeleceu como campo empírico os protocolos da própria Universidade Federal Fluminense, procurando compreender as rotinas administrativas e o processo de indexação junto ao sistema de recuperação da informação da instituição. Além disso, desde 2016 essa pesquisa tem acompanhado a implementação do Processo Administrativo Eletrônico Nacional (PEN) e o Sistema Protocolo Integrado (SPI), via a adoção do Sistema Eletrônico de Informações (SEI), nesta universidade.

No âmbito da Administração Pública Federal, a Portaria Interministerial 2.320, de 30 de dezembro de 2014, emitida pelos Ministério do Planejamento e Ministério da Justiça, institui o Sistema Protocolo Integrado para órgãos e entidades desta.

De acordo com o Ministério do Planejamento, através de sua Secretaria de Logística e Tecnologia da Informação¹⁵, o Sistema Protocolo Integrado tem por objetivo disponibilizar um canal do Governo Federal via internet para acompanhamento, pelo cidadão, do andamento de documentos, avulsos ou processos, do seu interesse.

Os Ministérios que instituíram tal Sistema têm incentivado o uso do Sistema Eletrônico de Informações (SEI), software de governo, elaborado pelo Tribunal Federal do Trabalho da 4ª região com o objetivo de executar as diretrizes da portaria para a implementação do Sistema Protocolo Integrado.

O SEI é considerado um dos produtos do projeto Processo Eletrônico Nacional (PEN), iniciativa da Administração Pública, com o intuito de construir uma plataforma pública de processos e documentos administrativos eletrônicos.

A implementação do Sistema Informatizado de Protocolos e do Processo Eletrônico Nacional configuram-se em uma ótima oportunidade de padronizar a indexação dos documentos em seu registro, na fase de sua produção.

¹⁵ Informações disponíveis no próprio site do Ministério do Planejamento, Desenvolvimento e Gestão. Disponível em: <<http://www.planejamento.gov.br/servicos/servicos-do-mp/protocolo-integrado>>. Acesso em: 28 mai 2017.

Os protocolos funcionam para registrar, controlar o fluxo documental, informar sobre o trâmite dos documentos em um sistema de arquivo.

A Universidade Federal Fluminense ao longo dos seus 69 anos tornou-se a maior universidade federal do país em número de alunos na graduação, de acordo com o **Censo da Educação Superior 2017**, documento elaborado pelo Ministério da Educação. Desde sua adesão ao Plano de Reestruturação e Expansão das Universidades Federais (Reuni), em 2008, intensificou seu processo de interiorização, criando novos cursos e ampliando a oferta de vagas em seus campi em Angra dos Reis (RJ), Campos dos Goytacazes (RJ), Macaé (RJ), Nova Friburgo (RJ), Rio das Ostras (RJ), Santo Antônio de Paduá (RJ), Volta Redonda (RJ), Oriximiná (PA), Petrópolis (RJ), além de realizar o mesmo em seu campus sede em Niterói (RJ). O tamanho e a importância dessa instituição federal de ensino foram critérios para sua escolha como campo empírico.

A UFF não possui um sistema de arquivos integrado. Os protocolos não atuam em suas unidades de trabalho. O sistema de registros criados pela UFF baseia-se na geografia de seus campi. A estrutura dos protocolos da Universidade Federal Fluminense é constituída pela Gerência Plena de Comunicações Administrativas (GPCA) que funciona como uma espécie de protocolo central que registra as atividades produzidas pelos setores subordinados diretamente à Reitoria como a Secretaria Geral, as Assessorias, a Ouvidoria, o Gabinete, a Diretoria de Relações Internacionais.

As atividades administrativas das Pró-reitorias (PROGRAD, PROEX, PROPPI, PROAD, PROAES, PROGEPE, PROPLAN) e de seus setores são registradas também pela GPCA. Subordinados a esta gerência estão os protocolos setoriais, divididos por áreas geográficas, de acordo com os campi. A UFF possui quatro protocolos setoriais: o Protocolo Setorial do Gragoatá (PSG), o Protocolo Setorial da Área de Saúde (PSAS), o Protocolo Setorial da Praia Vermelha (PSPV) e o Protocolo Setorial do Valonguinho (PSV). Estas unidades de informação eram subordinadas ao Núcleo de Documentação (NDC), da mesma forma que as demais como as bibliotecas, a Coordenação de Bibliotecas, o arquivo intermediário, o arquivo permanente e a coordenação de arquivos, desde 2004, quando houve a reestruturação dos setores da Universidade Federal Fluminense, com a extinção dos Centros Universitários.

Em março de 2011, após a reestruturação interna, o núcleo transformou-se em Superintendência de Documentação (SDC), os protocolos e demais Serviços de Comunicações Administrativas (SCA) passaram a estar subordinados à Pró-reitoria de Administração (PROAD)¹⁶. Esta decisão rompeu, ao menos administrativamente, com a organicidade dos documentos de arquivo. Esta realidade dos arquivos que afetam os sistemas de informação da Universidade é

¹⁶ Informações retiradas do site da Superintendência de Documentação da Universidade Federal Fluminense que relata o histórico institucional. <http://www.ndc.uff.br/content/institucional>. Acesso em: 22 de setembro de 2014.

importante para compreender os limites e gargalos das instituições federais de ensino como entidades produtoras de documentos de arquivo.

Outrossim, a UFF constituiu uma comissão para a implementação do Sistema Eletrônico de Informação (SEI) em 23 de fevereiro de 2017, através da portaria 58. 196. Uma ação tardia tendo em vista que o decreto 8.539, já neste trabalho citado, estipulou em outubro de 2015 o prazo de dois anos para a adequação dos órgãos e entidades da Administração Pública Federal ao Processo Eletrônico Nacional (PEN). Acompanhamos os trabalhos desenvolvidos por tal comissão desde então com objetivo de analisar as questões propostas pela tese, considerando esta grande transformação na forma de gestão dos documentos pelos órgãos e entidades federais.

Essa pesquisa tem investigado a implementação do Sistema Protocolo Integrado nas universidades federais brasileiras, focando-se no processo de indexação dos documentos, a partir deste campo empírico, os protocolos da Universidade Federal Fluminense.

6 AS UNIVERSIDADES FEDERAIS BRASILEIRAS E O SISTEMA PROTOCOLO INTEGRADO

Para entendermos os arquivos das universidades federais é importante compreendermos as estruturas em que elas estão inseridas, as relações que elas fazem parte, a comunidade discursiva a qual pertencem. Como já fora dito, seus documentos pertencem a essas dinâmicas, são produzidos e produzem ações, relações, operam trânsitos, interditam sujeitos, viabilizam projetos, validam e proíbem atos, ou seja, estão inseridos nos fatos sociais totais que fazem parte deste universo particular das instituições federais de ensino superior.

Essa seção dedica-se a compreender tal universo, entender as dinâmicas das universidades como instituições federais de ensino superior. Seguindo as pistas de Fish (1992) Swales (1990) e Bazerman (2005), no campo da linguística, e Hjørland (1995) na Ciência da informação, nossa proposta é mapear os atores sociais, os fatos sociais totais mediados pelos documentos de arquivo, os caminhos e circuitos, os sistemas de atividades humanas que acontecem na linguagem.

Acompanhamos a circulação dos documentos de arquivo nas universidades federais, tendo como espaço empírico a Universidade Federal Fluminense, identificando possíveis padrões interativos entre ações, sujeitos e documentos nessa comunidade discursiva. Tratamos estas como uma grande comunidade por entendermos que os atores sociais pertencem a um grupo que compartilham atividades em comum, participam de diversas redes de relacionamentos e compartilham concepções, signos, dinâmicas, acordos e tensionamentos. São sujeitos que partilham formas de vida semelhantes e estão vinculados a mesma esfera social.

Guedes (2010) ao aproximar o conceito de comunidade discursiva, de Swales (1990), do campo da Ciência da Informação, o apresenta como um grupo de pessoas que compartilham certas práticas e usos da linguagem. Essas práticas e usos podem ser de duas ordens: convenções estilísticas e interações sociais reguladas. Podemos inserir Bazerman (2005) nesta discussão, por seu conceito de gêneros textuais. De acordo com o autor, as comunidades discursivas criam padrões para realizarem suas atividades com caminhos relativamente previstos que permitem a realização dos atos perlocucionários.

O estabelecimento de formatos padronizados da produção e circulação desses tipos textuais permite a realização de ações, viabiliza relações, opera diversos trânsitos, pois permite o direcionamento do sentido. São acordos, convenções que não se restringem a seara do estilo comunicativo, mas compartilhamentos necessários para a realização dos fatos sociais totais, pois são gêneros definidos em que a comunidade sabe o que esperar de cada um deles e como respondê-los. Cabe ressaltar, que o Bazerman (2005) trabalha com o conceito *lato* de gêneros textuais e sua tipificação.

Livros, artigos, cartas, emails, convites de casamento, processos jurídicos, processos administrativos, cartões-postais, legendas de fotografias, filmes, jornais, atas, diplomas, atestados médicos são exemplos de gêneros textuais. Eles existem em suas comunidades, pois são convenções firmadas e que respondem a certas demandas, são padronizados por certos ritos e desencadeiam ações e relações já esperadas.

Os documentos de arquivo, apesar de toda sua especificidade abordada em seções anteriores, fazem parte dessa dinâmica, pois são altamente regulados, inclusive pela legislação vigente e pelas rotinas administrativas, com fluxos regularizados e circuitos bem delineados. A qualidade mais cara a um documento de arquivo é a evidência. A regulação de sua produção e destinação, padronização de formulários, consolidação de seus caminhos, a ritualização administrativa do qual fazem parte fortalece o seu caráter de evidência.

Na dinâmica dos arquivos, a efetividade do cumprimento dos fatos sociais em que sua comunidade discursiva faz parte, neste caso das funções e atividades da entidade produtora, dá-se pela mediação dos documentos.

6.1 AS UNIVERSIDADES BRASILEIRAS E O ORDENAMENTO JURÍDICO

A tradição brasileira acolhe o direito positivo como filosofia que rege a normatividade da sociedade. Bobbio (1999) aponta seis características que definem o direito positivo: a particularidade – o estabelecimento das normas tem um alcance definido, é aplicado para dada sociedade, em determinado tempo e espaço; a mutabilidade – o direito é dinâmico e responde às ações e relações dos agentes em sua sociedade; o direito emana do povo, portanto é estabelecido pelo povo em seus acordos, costumes e tradições; o direito positivo é uma declaração de vontade, conhecido por sua promulgação; o valor das ações está naquilo que é útil à sociedade e, por último, o objeto do direito são os comportamentos, estes só assumem certa qualificação apenas porque e quando disciplinados pelas ordens promulgadas.

É importante considerar a tradição de nosso ordenamento jurídico, pois nos permite perceber como a mediação das relações sociais acontecem pelos documentos, a configuração da própria sociedade nas relações com suas instituições reguladas pela legislação.

As ações, sobretudo na Administração Pública, por nossa tradição jusnaturalista, carecem do formalismo do ordenamento jurídico para serem desenvolvidas, precisam da validade da forma. Portanto, nossa sociedade é regulada por uma estrutura, um conjunto de normas jurídicas vigentes, conhecida no jusnaturalismo como ordenamento jurídico (BOBBIO, 1999, p. 133). De acordo com esta filosofia, os indivíduos dessa sociedade devem obediência absoluta da lei enquanto tal.

As universidades federais pertencem a esta configuração social. As funções e as atividades desenvolvidas por seus agentes são reguladas por normativas que respondem e são emanadas do

ordenamento jurídico. Os documentos, disciplinados em sua forma de produção, tramitação e destinação, são meios que permitem a legalidade das condutas dos sujeitos inseridos nas dinâmicas das universidades federais.

Estas disposições socialmente constituídas que, enquanto estruturas estruturadas e estruturantes, segundo Bourdieu (2003), compõem um sistema, o *habitus*, princípio gerador e unificador do conjunto de práticas e ideologias características do grupo de agentes. As funções e atividades são realizadas dentro de um fluxo regulado pelo ordenamento jurídico porque essa comunidade discursiva compartilha a região de sentido e significado.

Os diferentes níveis de significação como os níveis da língua, articulam-se nem sistema hierarquizado onde o englobante é por sua vez englobado, o significado por sua vez significante, sistema que a análise percorre nas suas operações ascendentes e descendentes (BOURDIEU, 2003, p. 341)

Percebe-se que o *habitus* não se trata apenas de um código comum, nem do mesmo repertório de respostas a problemas comuns ou conjunto de pensamentos particulares e particularizados. De acordo com Bourdieu (2003, p. 349), o *habitus* é, sobretudo, um conjunto de esquemas fundamentais, previamente assimilados, a partir dos quais outros esquemas particulares se amalgamam e constroem possibilidades infinitas de aplicação de resoluções em situações particulares.

Portanto, para compreender a realidade dos arquivos de universidade, em especial das instituições federais, em sua fase de produção documental, é necessário compreender quem pertence dessa comunidade discursiva; como se formam os significantes e os significados.

De acordo com Delmas (2010), a administração de qualquer órgão, instituição, entidade, se realiza por uma sucessão ou transferência de atribuições e é acompanhada pela transmissão dos registros e dos documentos, ou seja, dos arquivos.

As universidades federais fazem parte da Administração Pública Federal indireta e estão subordinadas ao Ministério da Educação. Então, são reguladas por leis, decretos em nível federal, portarias de ministérios que estabelecem diretrizes para o exercício de suas competências como o Ministério do Planejamento, Desenvolvimento e Gestão, a Casa Civil e o próprio Ministério da Educação. Precisam enquadrar-se na Política Nacional de Educação, na Lei de Diretrizes e Bases (1996), nas políticas de gestão da Administração Pública federal. Entretanto, além do ordenamento jurídico na escala macro, essas instituições também se autogerem com normativas próprias por meio de regulamentos, regimentos, instruções etc (CLÉVE; LORENZETTO, 2016).

As universidades federais são entidades que realizam a vontade da União que, por sua vez, exerce seu controle finalístico com o objetivo de avaliar se essas instituições estão atingindo o fim público para o qual foram criadas. No espectro jurídico, estão subordinadas ao poder uno federal,

mas de forma descentralizada, a fim de desempenhar atribuições específicas deste, isto é, oferecer o ensino superior público. Inseridas nesse contexto legal, as funções e atividades nas universidades federais só podem ser realizadas por meio de atos da administração, regulados pelo direito administrativo (CLÉVE; LORENZETTO, 2016).

Essas instituições estão inseridas em políticas públicas para o ensino superior pelo governo brasileiro que por sua vez procura estar em consonância com as diretrizes globais para o estabelecimento de universidades de classe mundial. A lógica deste conceito encontra-se nesta fase contemporânea do capitalismo neoliberal. A lógica da educação terciária caracterizada pelo estímulo ao empresariamento do ensino superior; a ampliação da privatização interna e o estabelecimento de contratos de gestão que alteram o financiamento e as diretrizes político-pedagógicas das universidades públicas, especialmente as federais (LIMA, 2011, p. 86).

A reconfiguração da educação superior pública aponta para uma visão economicista na qual se abandona o seu princípio de direito social em prol de um modelo de gestão empresarial, percebendo a instituição como organização que oferece um *commodity*. As ações e relações passaram a ter princípios de custo-benefício, eficácia e eficiência, e serem pautadas pelos padrões de qualidade ditados pelo mercado financeiro.

Dessa forma, as políticas públicas voltadas para a educação superior no Brasil passaram a seguir as diretrizes do Banco Mundial e do Banco Internacional de Desenvolvimento. Desde o final da década de 1990, as políticas nacionais para a educação superior têm se alinhado às propostas expressas em documentos como **La enseñanza superior** (1994), o **Relatório sobre o desenvolvimento mundial** (1997) e a **Estratégia para o setor educacional** (1999), publicados pelo Banco Mundial.

As diretrizes sugeridas convergem para a percepção da educação superior como um campo de exploração lucrativa para o capital em crise. Nessa dinâmica do capitalismo tardio, as funções e atividades realizadas nas universidades públicas brasileiras precisam explorar ao máximo os recursos, sobretudo o de mão-de-obra, acelerar a formação profissional com condensação de conteúdos e encurtamento dos períodos para o desenvolvimento de saberes e habilidades, diminuir custos e pulverizar cursos e disciplinas com o objetivo de proliferar certificados de validação de técnicas e conhecimentos que respondam às demandas do mercado de trabalho.

A reforma da educação superior proposta pelo Banco Mundial por meio de seus documentos e seguida pelas políticas educacionais brasileiras, propõe a diversificação das instituições de ensino superior, das modalidades de cursos e das fontes de financiamento.

A noção de público e privado vai sendo diluída nas universidades federais, suas funções e atividades passam a ser geridas, muitas vezes, por convênios de prestação de serviços entre as universidades e empresas de capital privado. O arcabouço político e jurídico tanto na macro-esfera,

em nível de Administração Pública Federal, quanto na micro, na auto-regulação das instituições federais de nível superior passou a ser implementado ao final dos anos de 1990 (LIMA, 2011). Elas criaram meios que possibilitam a implementação das diretrizes privatizantes da educação nas universidades federais.

Essa concepção política, econômica, ideológica instaura também a política de “qualificação” do ensino superior e as formas de avaliação da universidade passam a ser orientadas a partir de critérios de eficiência administrativa e atendimento das exigências dos setores privados.

Nessa dinâmica o Estado é o agente do processo de desenvolvimento social e econômico, mas atua como parceiro catalisador e impulsionador deste processo por meio de suas políticas públicas. O papel principal do Estado converte-se em elaborar e regular de forma eficaz seus serviços com o intuito de melhor aproveitar das “forças do mercado” em benefício dos “bens públicos” e a educação superior está inserida nesse contexto.

As universidades federais passaram então a fazer parte do concerto do estímulo à maior competência da administração pública, por meio de contratações de serviços por mérito e avaliações por produtividade.

O reordamento do papel do Estado, eixo norteador do Banco Mundial para a periferia do capitalismo, consolidou, ao longo da década de 1990 e no início do novo século, um intenso processo de privatização dos serviços públicos, incluindo a educação e, especialmente, a educação superior (PILATTI, 2018, p. 88).

Portanto, percebe-se o esgotamento das características, antes inerentes à universidade pública, caracterizada por Chauí (2003) como instituição social fundada na prática da ação cujo reconhecimento público se dava pela legitimidade de sua atuação na produção do conhecimento e no exercício de suas atribuições como ensino, pesquisa e extensão. A redefinição das atribuições das universidades faz um esforço para aproximá-las do mercado. As diretrizes passam focar na meta de adaptar essas instituições a práticas e a uma cultura afinada como o setor privado (SLAUGHTER; LESLIE, 1997).

A visão utilitarista da universidade passa a permear as políticas públicas para a educação superior e estas passam a atender à temporalidade própria do capital apontando para conhecimentos a serem desenvolvidos que nem sempre atendem ao interesse público, mas às demandas do setor privado. Ciência e tecnologia passaram a ser concebidas como mercadorias de grande valor, fornecidas pela universidade ao mercado, fazendo dessas instituições parte importante da economia. O conhecimento e a informação começaram a fazer parte da dinâmica do capital, sendo considerados bens intangíveis, por conseguinte, passaram a fazer parte da acumulação e reprodução para a produção de riquezas.

A adaptação às suas novas funções pelas universidades para fazer parte dessa lógica compeliu estas instituições a praticarem a racionalização da administração, balizada pelos conceitos de eficiência e eficácia, obedecendo ao primado da gestão, do planejamento, da organização, da coordenação, da previsão, da avaliação por produtividade, do controle e êxito. Essas instituições passaram a ser avaliadas e medidas em sua qualidade pelo mundo dos negócios e o seu financiamento pelas políticas públicas depende do seu desempenho nessas avaliações.

Como afirma Trindade (1999), a crise da universidade pública brasileira é a própria crise da sociedade do conhecimento. O conhecimento deixa de ser uma ação social forjada na dialética das relações humanas para ser um *commoditie* a ser comercializado, quantificado, qualificado a partir de sua resposta às demandas mercadológicas.

O conhecimento torna-se um instrumento a garantir ganhos de produtividade. Nesta nova ordem, a lógica administrativa empresarial impõe a comercialização do saber. As universidades passam a ser espaços que garantem estruturas de sustentação para o desenvolvimento da inovação, com escritórios de transferências de tecnologias, contratos com grandes empresas e organizações curriculares mais aderentes às demandas empresariais (FREITAG, 1986, p. 351)

De acordo com Chanlat (2002), a crise econômica das últimas duas décadas do século XX levou ao esgotamento da concepção Estado-providência. As mudanças afetaram os serviços públicos, influenciando de maneira profunda as políticas, os programas, as estruturas e a cultura das organizações estatais. O movimento reformista que afetou também as universidades federais apresentou como eixo central de sua proposta o abandono do modelo burocrático da administração pública pelo modelo da administração gerencial.

A burocracia estatal recebeu a alcunha de pouco ágil, dispendiosa, que não responde adequadamente às demandas dos cidadãos. Já a administração gerencial é vista como a modernização necessária, pautada na descentralização e flexibilização administrativa, livre das amarras rígidas da burocracia clássica (BRESSER-PEREIRA, 2006).

O sentimento de aversão à burocracia do serviço público passou a ser compartilhado por boa parte da sociedade frente à degradação das instituições públicas, e o discurso de otimização dos serviços prestados pelo Estado por meio do modelo de gestão do setor privado passou a ser amplificado. As universidades federais precisaram se adequar ao modelo de gestão privada para conseguirem recursos e financiamento do Estado e, por conseguinte, se manterem. Assim, os imperativos passaram a ser: inovação, eficácia, produtividade, performance, competência, empreendedorismo, qualidade total, otimização da relação cliente-produto, desempenho, excelência e reengenharia (PAULA, 2005, p. 57).

As universidades públicas brasileiras passaram então a ser avaliadas a partir de critérios padronizados e internacionalmente aceitos nessa lógica do capitalismo tardio. Rankings como o

Academic Ranking of World Universities (ARWU), o *Times Higher Education Ranking* (THE) e o *Quacquarelli Symonds University Ranking* (QS) passaram a ser referências para a avaliação das atividades-fim das universidades brasileiras.

A qualidade da pesquisa, a qualidade da instituição, os resultados da pesquisa e o desempenho per capita passaram a ser critérios de avaliação. Dentre as categorias há a reputação acadêmica dos autores, alunos e equipe acadêmica; orçamento destinado à pesquisa; quantidade de doutores; reputação, investimento e publicação de pesquisas; influência da pesquisa aferida pelos dados bibliométricos; perspectiva de internacionalização da instituição; investimento da indústria a partir do financiamento de projetos de inovação (PILLATTI; CECHIN, 2018).

6.2 O PROGRAMA DE REESTRUTURAÇÃO E EXPANSÃO DAS UNIVERSIDADES FEDERAIS (REUNI)

A lógica do capitalismo tardio de implementação do modelo de gestão privado, transformando as universidades em organizações empresariais, com metas a serem cumpridas e desempenhos a serem avaliados a partir dos critérios citados, norteou o processo de expansão das universidades federais brasileiras. Há o deslocamento da compreensão da educação superior como direito social para a educação terciária onde a sociedade do conhecimento é regida pela lógica de produção da economia mundial.

O programa de Reestruturação e Expansão das Universidades Federais, o Reuni, instaurado pela lei 6.096 de 24 de abril de 2007, foi um processo que procurou transformar as universidades em instituições de ensino terciário, numa perspectiva de ampliação das instituições visando formar catálogos de projetos e conteúdos a serem distribuídos à iniciativa privada.

Nessa perspectiva, as metas do Reuni de aumento do número de estudantes nas universidades federais, de aumento do número de alunos por professor em cada sala de aula da graduação; de diversidade das modalidades dos cursos de graduação, através da flexibilização dos currículos, da educação à distância, da criação de cursos de curta duração e dos ciclos (básico e profissional); da elevação da taxa de conclusão dos cursos de graduação para 90%; do estímulo à mobilidade estudantil entre as instituições de ensino, estão em consonância com as diretrizes do Banco Mundial e do Banco Internacional de Desenvolvimento para a educação superior que visam dinâmicas organizacionais eficientes mercadologicamente, com efetividade, competitividade e sustentabilidade (RIBEIRO; LEDA; SILVA, 2016).

O Reuni inaugura, de certo modo, um período de redefinição das universidades, alinhando cada vez mais à educação superior ao mercado. A expansão das universidades federais apresentava como desafio adaptar as práticas dessas instituições à cultura tipicamente associada ao setor

privado, algo facilitado pela inserção de ações privatizantes nessas instituições desde a década de 1980 (RIBEIRO; LEDA; SILVA, 2016).

O imperativo da lógica da administração gerencial acompanhou o processo de reformulação das universidades federais. A dotação orçamentária estava vinculada aos resultados apresentados, às metas estabelecidas. Ações em consonância com a política do Governo para a educação superior que seguia os ditames capitalistas de produção, utilitarismo, eficiência e eficácia. A temporalidade do imediatismo também é algo presente, soluções e respostas precisam ser rápidas, o tempo para reflexão crítica e desenvolvimento filosófico para a produção do conhecimento é cada vez mais exíguo. Aposta-se na execução de modelos já prontos e adaptados para inúmeras áreas, muitas vezes incompatíveis com as especificidades de cada campo e com a própria realidade da comunidade a qual a universidade pertence.

Apesar da grande injeção de recursos nas universidades federais, o Reuni aprofundou ações na educação superior que passaram a apresentá-la como um serviço que pode estar sob o controle do Estado, tal como a política que estava sendo implementada, mas que não são exclusivos desse agente. As universidades federais vão caminhando para se transformarem em organizações públicas não estatais. Portanto, entidades independentes e controladas, pelo Estado e pela sociedade civil, sendo parcialmente ou totalmente financiadas pelo Estado. A política traçada no Plano Diretor de desresponsabilização do Estado é crescente, inclusive com o Reuni (SERAFIM, 2011).

As universidades federais precisaram apresentar planos de desenvolvimento institucional em acordo com as metas estabelecidas pelo Reuni como a expansão da oferta dos cursos de graduação e pós-graduação; ampliação do acesso na educação superior, sobretudo na graduação; melhor aproveitamento da estrutura física e dos recursos humanos existentes; mobilidade intra e interinstitucional; diminuição da evasão; aperfeiçoamento qualitativo e quantitativo das pesquisas e projetos.

A Universidade Federal Fluminense apresentou, em 2010, o seu **Caderno Técnico – Programa de Expansão e Reestruturação da UFF (2006 a 2012)**. De acordo com a instituição, sua expansão iniciou no ano anterior à instauração do Reuni pelo decreto número 6.096, de 24 de abril de 2007. Neste documento, a UFF procurou responder ao Ministério da Educação e aos requisitos de controle de execução e avaliação das atividades previstas no Plano de Desenvolvimento Institucional (PDI) da universidade, aprovado pelo próprio MEC. O plano apresentado pela UFF seguiu as seis diretrizes estabelecidas e incluiu diagnósticos, metas, estratégias, etapas e indicadores nas seguintes dimensões (UNIVERSIDADE FEDERAL FLUMINENSE, 2010, p. 10):

Ampliação da oferta de educação superior pública

1. Aumento de vagas de ingresso, especialmente no período noturno;

2. Redução das taxas de evasão;
3. Ocupação de vagas ociosas.

Reestruturação acadêmico-curricular

1. Revisão da estrutura acadêmica buscando a constante elevação da qualidade;
2. Reorganização dos cursos de graduação;
3. Diversificação das modalidades de graduação, preferencialmente com superação à profissionalização precoce e especializada;
4. Implantação de regimes curriculares e sistemas de títulos que possibilitem a construção de itinerários formativos;
5. Previsão de modelos de transição, quando for o caso.

Renovação pedagógica da educação superior

1. Articulação da educação superior com a educação básica, profissional e tecnológica;
2. Atualização de metodologias (e tecnologias) de ensino-aprendizagem;
3. Previsão de programas de capacitação pedagógica, especialmente quando for o caso de implementação de um novo modelo.

Mobilidade intra e interinstitucional

1. Promoção da ampla mobilidade estudantil mediante o aproveitamento de créditos e a circulação de estudantes entre cursos e programas, e entre instituições de educação superior.

Compromisso social da instituição

1. Políticas de inclusão;
2. Programas de assistência estudantil;
3. Políticas de extensão universitária.

Suporte da pós-graduação ao desenvolvimento e aperfeiçoamento qualitativo dos cursos de graduação

1. Articulação da graduação com a pós-graduação: expansão quali-quantitativa da pós-graduação orientada para a renovação pedagógica da educação superior

Essa grande transformação impactou profundamente a Universidade Federal Fluminense em suas funções e atividades. Criada pela lei 3.848, de 18 de dezembro de 1960, como Universidade Federal do Estado do Rio de Janeiro (UFERJ), e reestruturada como Universidade Federal Fluminense (UFF), pelo decreto 62.214, de 15 de março de 1962, passou por uma ampliação considerável do número de alunos matriculados, considerando inclusive a sua grande expansão para o interior do estado do Rio de Janeiro.

Ainda nos primeiros anos do PDI, a Universidade já contava com 44.946 alunos matriculados, sendo 28.393 alunos de graduação presencial; 4.396 alunos de graduação à distância;

4.651 alunos de pós-graduação *strito sensu* e 7.506 alunos de cursos *lato sensu*. Para desenvolver as suas funções de pesquisa, ensino e extensão, a instituição contava até agosto de 2009 com 2.563 professores e 3.952 funcionários (UFF, 2010).

No entanto, a expansão da rede federal de educação superior vem sofrendo constantes cortes orçamentários. Em 2017, foi divulgado um estudo que mostra as perdas reais em 63 instituições federais de ensino superior, 90% das universidades federais tiveram o menor repasse de verbas desde 2013, encolhendo 28,5%. No entanto, como as demais universidades federais que têm aumentado o número de alunos matriculados todos os anos na casa dos 10%, a Universidade Federal Fluminense ampliou em mais de 20.000 o número de vagas entre 2009 e 2016.

Segundo o Censo UFF 2016¹⁷, havia 68.248 alunos matriculados. Apesar do número docente ter aumentado proporcionalmente ao número de alunos matriculados, subindo para 3.599, a média é de 20 alunos por professor, incluindo a pós-graduação, além do desenvolvimento de ações, projetos e programas de pesquisa e extensão. O quadro de funcionários aumentou menos de 10%, chegando a 4.306 no final de 2016. Entretanto, a quantidade de terceirizados aumentou consideravelmente, passando para 2.150 (UFF, 2017).

A Universidade Federal Fluminense é uma autarquia da Administração Pública e, portanto desempenha suas funções e atividades, por meio de atos vinculados ao ordenamento jurídico como preconiza o direito administrativo brasileiro.

A lei 9.784, de janeiro de 1999, regula o processo administrativo no âmbito da Administração Pública Federal. Seguindo o princípio da legalidade, a Administração Pública se sujeita obrigatoriamente à lei e ao Direito. Segundo Bobbio (1999), tal ideia é inerente aos postulados de um Estado de Direito e é responsável por determinar a validade das ações desempenhadas pelos governantes.

Neste âmbito, as universidades federais desempenham suas funções por meio de atos administrativos, espécies de atos jurídicos que manifestam, de forma lícita, vontade humana que tem por objetivo imediato adquirir, resguardar, transferir, modificar e extinguir direitos. Esses atos são praticados por agentes credenciados da administração, que tem por efeito imediato a aquisição, o resguardo, a modificação, a transformação ou a extinção de direitos, em matéria administrativa (CRETELLA JÚNIOR, 1998).

Segundo Celso Antônio Bandeira de Mello (2004, p. 356) o ato administrativo é:

[...] declaração do Estado (ou de quem lhe faça as vezes – como, por exemplo, um concessionário de serviço público), no exercício de prerrogativas públicas, manifestada mediante providências jurídicas complementares da lei a título de lhe dar cumprimento, e sujeitas a controle de legitimidade por órgão jurisdicional.

¹⁷Os dados do Censo UFF estão representados em gráficos no anexo G.

Por conseguinte, o ato administrativo caracteriza-se pelas seguintes características apontadas por Mello (2004): 1) trata-se de declaração jurídica que produz efeitos, como a certificação, declaração, criação, extinção, transferência, ou modificação de direitos ou obrigações; 2) provém do Estado, ou de quem esteja investido de prerrogativas públicas; 3) a autoridade que o pratica encontra-se regido pelo Direito Público; 4) é providência jurídica complementar à lei (infralegal ou sublegal) a título de lhe dar cumprimento; 5) submete-se a exame de legitimidade pelo Poder Judiciário (controle judicial), podendo ser invalidados em casos de ilegalidade.

O processo administrativo apresenta-se então como instrumento mais usual para a realização dos atos administrativos, em especial, em nosso campo de análise, as universidades federais. Este recurso é utilizado para manter a ordenação das rotinas dessas instituições e preservar os direitos dos administrados através de protocolos que devem ser seguidos a termo para preservar os princípios do ordenamento jurídico da qual fazem parte.

No campo dos arquivos, o tipo documental processo administrativo tem por essência as características de um documento de arquivo uma vez que registra uma sucessão de atos ordenados os quais compõem uma cadeia, sendo cada um destes atos autônomos concomitantes para culminar determinado fim, ou seja, a execução de determinada atividade seja ela meio ou fim. Cabe aqui ressaltar que nem toda tarefa ou processo de trabalho resultará em um processo administrativo. Entretanto, como ressalta Merkel (2008, p. 22):

Todas as funções estatais e, em particular, todos os atos administrativos são metas que não podem ser alcançadas senão por determinados caminhos. Assim, a lei é a meta que nos leva a via legislativa e os atos judiciais e administrativos são metas que nos conduzem o procedimento judicial e administrativo.

As universidades por meio de seus regimentos e regulamentos normatizam suas funções e atividades, subordinadas hierarquicamente ao Ministério da Educação e observando as diretrizes do Plano Nacional de Educação do Ensino Superior. Outrossim, compõem o sistema federal de ensino e são constantemente avaliadas tanto em si, no desempenho de suas funções, como em seus cursos.

As atividades das universidades federais brasileiras caracterizam-se por viabilizarem o cumprimento de suas finalidades elencadas na legislação brasileira, em sua maioria, via processos administrativos, considerando suas três funções precípuas – pesquisa, ensino e extensão, previstas no artigo 207 da Constituição Federal (1988), além daquelas estabelecidas no artigo 43, da Lei de Diretrizes e Bases da Educação, de 20 de dezembro de 1996, a saber:

I estimular a criação cultural e o desenvolvimento do espírito científico e do pensamento reflexivo; II - formar diplomados nas diferentes áreas de conhecimento, aptos para a inserção em setores profissionais e para a participação no desenvolvimento da sociedade brasileira, e colaborar na sua formação contínua; III - incentivar o trabalho de pesquisa e investigação científica, visando o desenvolvimento da ciência e da tecnologia e da criação e

difusão da cultura, e, desse modo, desenvolver o entendimento do homem e do meio em que vive; IV - promover a divulgação de conhecimentos culturais, científicos e técnicos que constituem patrimônio da humanidade e comunicar o saber através do ensino, de publicações ou de outras formas de comunicação; V - suscitar o desejo permanente de aperfeiçoamento cultural e profissional e possibilitar a correspondente concretização, integrando os conhecimentos que vão sendo adquiridos numa estrutura intelectual sistematizadora do conhecimento de cada geração; VI - estimular o conhecimento dos problemas do mundo presente, em particular os nacionais e regionais, prestar serviços especializados à comunidade e estabelecer com esta uma relação de reciprocidade; VII - promover a extensão, aberta à participação da população, visando à difusão das conquistas e benefícios resultantes da criação cultural e da pesquisa científica e tecnológica geradas na instituição.

Bem assim, as atividades-meio também compõem as ações das universidades federais. Comuns à Administração Pública em geral, são aquelas atividades não pertinentes ao objetivo principal da instituição produtora, tratam-se de serviços necessários, mas que não têm relação direta com a missão da instituição, também costumam tramitar por processos administrativos.

Essas atividades desenvolvidas por meio dos processos de trabalho nas universidades federais, na maioria dos casos, tramitam por processos administrativos, regulados pela legislação brasileira, que registram as ações realizadas no desenvolvimento de cada atividade. Com o objetivo de registrar o fluxo desses processos e os atos realizados em cada etapa, as instituições adotam protocolos como sistemas de registro desses processos administrativos.

6.3 O PROTOCOLO E SEU LUGAR ESTRATÉGICO

O protocolo é uma forma de registro surgido nas administrações governamentais no reino da Prússia. O protocolo surgiu como local para salvaguardar os documentos dentro de uma dinâmica competente aos trabalhos administrativos em si (MILLER, 2003, p. 43). Conhecido como *registratur*, a natureza especial do protocolo é a de arquivo em ação, ou seja, cria o documento, transcreve seu conteúdo e ação, grava, registra e encaminha para as repartições competentes.

Considerado sistema de registro em diversos países como Alemanha, Inglaterra, Estados Unidos, Canadá, Espanha, Portugal entre outros, o protocolo tem a sua origem etmológica na palavra *regere* que em latim significa carregar, trazer de volta. No Latim tardio, a palavra *regista* aparece como particípio passado de *regere* e ganha o significado de lista, catálogo, registro de entradas. Na Idade Média, a palavra foi se transformando e assumindo outras formas como *registrare*, *registtrum*, assim como o termo *registratur*, posteriormente adotado pelos prussianos e que tem um vínculo forte com o princípio da proveniência no âmbito dos arquivos, consagrado pelo **Manual de arranjo e descrição**, de 1898, publicada pela Associação dos Arquivistas Holandeses.

Cabe ressaltar, como demonstra Miller (2003), que as primeiras formas de protocolo surgidas nos monastérios europeus durante a Idade Média eram livros com série de páginas numeradas e vinculadas onde eram registradas as entradas dos documentos. O chamado livro copião é

notoriamente uma das primeiras formas de protocolo surgido em instituições eclesiásticas (MILLER, 2003, p. 44). Dado o valor legal dos documentos recebidos, os ordenadores eclesiásticos passaram a instituir o protocolo como prática necessária na garantia do armazenamento seguro dos documentos originais.

No século XIV já era popular o registro de documentos financeiros, incluindo os livros contábeis, de documentos proferidos pelas cortes e tribunais. O protocolo foi firmando-se como lugar de registro apto a servir como uma espécie de dispositivo de memória por ser uma espécie de diário a serviço de lembrar os documentos que passaram a fazer parte da sua instituição. Esse tipo de registro podia ser usado para checar a autenticidade e a fidedignidade dos documentos, pois eram requeridas as chancelas e validações necessárias aos documentos para que pudessem ser aceitos pelos responsáveis do *registratur* (MILLER, 2003, p. 46).

Autoras como Foscarini (2009) e Miller (2003) assinalam importantes diferenças entre os sistemas de registro, na qual consideramos como sinônimos de protocolo no Brasil, como lugares físicos onde se dá a entrada dos documentos e estes passam a ser organizados de acordo com os processos de trabalho sobre os quais versam ou foram criados e a função dos sistemas de registro que ao longo do tempo se ampliou não sendo apenas uma listagem de entradas de documentos, mas autuam, classificam, indexam, acompanham o trâmite do documento e suas respectivas atividades nas unidades de trabalho.

O protocolo como unidade primeira de arquivamento, responsável por garantir a organicidade dos documentos por meio dos registros realizados, faz parte do modelo clássico prussiano, consagrado na segunda metade do século XVI. Nas administrações dos reinos prussianos, o *registratur* funcionava como uma repartição responsável por organizar os documentos junto às demais repartições e as matérias de negócio, ou seja, junto às ações que geravam os documentos e que eram decorrentes deles também para a execução de dada atividade da entidade.

Durante o curso dos séculos XVIII e XIX, a administração prussiana consolidou o uso do *registratur*, primeiro devido a divisão de áreas de responsabilidade por competência profissional; segundo, por conta das decisões tomadas de forma colegiada; terceiro, pelo uso de códigos de procedimentos administrativos para guiar as formas em que os administradores deveriam tratar as matérias de negócio. O *registratur* possibilitava a integração das atividades e funções pelos registros das ações praticadas (MILLER, 2003, p. 48).

Desde 1652, as formas de lidar com as matérias de negócio foram reguladas por um conjunto de instruções formais que, segundo Miller (2003), eram mais sistematizadas e compreensivas do que as anteriores e surgiram do refinamento das práticas realizadas nessas repartições. Então, essas diretrizes passaram a ser utilizadas no cotidiano do trabalho administrativo realizado. Essa metodologia foi importante para resolver a desordem que se apresentava como obstáculo ao próprio

funcionamento dos reinos prussianos. Essas instruções trataram de maneira particularizada as seções de trabalho, inclusive, as secretarias, chancelarias, protocolos e arquivos (MILLER, 2003, p. 49).

O protocolo foi usado para acompanhar as tarefas em sequência do trabalho administrativo. Primeiro, era capaz de achar os documentos e os arquivos. Segundo, supostamente mantinha registrado a trilha, o trâmite dos documentos, inclusive aqueles que tramitavam fora dos sistemas de registro, pois havia a indicação das suas saídas. Finalmente, possibilitava garantir que os arquivos eram armazenados apropriadamente, seguindo a organicidade e mantidos em segurança.

Destaca-se que o protocolo contemporâneo não precisa necessariamente estar associado a um lugar físico, a uma repartição, há diversas plataformas eletrônicas que mantêm suas funções de registro, autuação, classificação, controle da tramitação dos documentos. O protocolo direciona cada documento para seu processo de trabalho, indica o compasso inteiro em que as ações são desencadeadas em uma atividade pela entidade e fornece uma perspectiva global de como os processos de trabalho vão se integrando no contexto de realização de uma atividade ou função. O protocolo pode configurar-se como uma central de controle de toda complexidade que são os processos de trabalho.

A mentalidade consolidada no capitalismo tardio do século XXI, influenciada pela Administração Científica, pelo neoliberalismo, e na seara dos arquivos, pela teoria da gestão de documentos, impõe velocidade e eficiência no desempenho das atividades e funções pela Administração Pública. O Estado e suas instituições precisam executar suas tarefas de forma ágil e produtiva a fim de atender às demandas da sociedade e do mercado, incorporando a nova tecnologia que facilite a execução de tais deveres e serviços.

Nessa nova dinâmica, os processos de trabalho são cada vez mais globais, sendo realizados muitas vezes, de forma sincrônica, por diferentes agentes e em diferentes lugares. Há múltiplos documentos, múltiplos processos de trabalho, múltiplos objetos em circulação. É necessário organização e controle para que as atividades e funções sejam desempenhadas de maneira, minimamente, satisfatória para todos os envolvidos nos processos de trabalho.

Miller (2003) argumenta que a forma tradicional de protocolo não consegue absorver as transformações em que os Estados e suas administrações passaram desde a segunda metade do século XX. A profusão de documentos após a Segunda Guerra Mundial, a circulação de massas de arquivo, a complexidade do tamanho dos Estados e de suas funções, além do impacto da tecnologia e da internet nas formas de comunicação levaram a uma imensa dificuldade em manter as formas de registro e controle pelos protocolos aos moldes do *registratur*.

A introdução da tecnologia da informação nos sistemas de registro nas administrações federais brasileiras tornou-se mais evidente na década de 1980, apesar de haver uma expectativa de ser

implementada de maneira sistêmica como nos países norte-americanos e europeus, algumas áreas de atuação do Estado foram privilegiadas como a economia, a indústria e a inovação tecnológica, assim como as instituições que atuavam nesses campos.

No Brasil, sobretudo a partir da década de 1990, foi reconhecido o impacto da informatização dos sistemas na execução dos processos de trabalho e, de acordo com cada dotação orçamentária, as entidades no âmbito da União foram desenvolvendo seus sistemas de registro, sua gestão de documentos e o gerenciamento de seus arquivos contando com o apoio da tecnologia da informação.

Os anos 2000 marcam o hibridismo dos documentos envolvidos nos processos de trabalho. Continuamos a trabalhar com documentos físicos, mas cada vez mais com documentos natos digitais. Ademais, os trâmites desses documentos passam por circuitos e sistemas informatizados nos quais os locais físicos tornam-se cada vez mais secundários.

O protocolo por suas funções de registro, controle da tramitação dos documentos e comunicação de ações configura-se como uma importante instância de poder, sobretudo, na administração pública.

Castells (2009), ao discutir o poder do Estado nas sociedades em rede, aponta justamente para o impacto das novas formas de comunicação nas novas formas do Estado exercer o seu poder. Segundo o autor, o poder é uma capacidade relacional que permite a um ator social influir de forma assimétrica nas decisões de outros atores sociais de modo que sejam favorecidas suas vontades, interesses e valores (CASTELLS, 2009, p. 33). Ferramentas como o protocolo permitem que o Estado exerça seu poder pela construção de significados, sem que seja necessário o uso da violência ou coação, consolidando práticas discursivas e maneiras de agir que guiam os demais atores sociais para seus objetivos finalíssimos.

O protocolo faz parte da forma de administração do Estado que se materializa em toda sua estrutura, ou seja, deve ter capacidade relacional que permita a dominação do governo aos demais sujeitos por meio de suas instituições. Estas se caracterizam por sua capacidade de manter relações de poder, baseadas na dominação que conseguem exercer sobre seus indivíduos. Nessa dinâmica, o protocolo faz parte do poder de Estado que não é um atributo, mas uma relação em que os atores sociais precisam aceitar para que suas ações se realizem a partir do domínio e da subordinação. “A relação de poder se estabelece com seus diversos componentes, a complexidade do mecanismo de dominação multi-nível funciona novamente, fazendo da violência mais um fator de um amplo conjunto de determinação” (CASTELLS, 2009, p. 34).

Como fora dito anteriormente, o protocolo também se torna um fator em que facilita a construção de caminhos, formas, validações na qual o Estado determina o exercício de seu poder sobre seus cidadãos.

O protocolo funciona como apoio complementar e recíproco dos principais mecanismos de manutenção do poder, sobretudo por meio da discursividade dos documentos, que legitima a institucionalização das formas de agir do Estado. Essa legitimação acontece em grande medida pelo consentimento da comunidade discursiva obtido mediante a construção dos significados compartilhados. Isso, segundo Castells (2009), envolve todo o processo de ação comunicativa que, entende-se, o protocolo faz parte dele.

Nessa sociedade em rede do século XXI, abordada por Castells (2009), as estruturas são ao mesmo tempo locais e globais. Nas universidades federais, por exemplo, os protocolos estão a serviço tanto de sua comunidade local, ou seja, seus alunos, professores, funcionários e dirigentes, quanto da esfera global, no caso brasileiro, a todos os cidadãos em virtude da lei 12.527 de 18 de novembro de 2017, conhecida como lei de acesso à informação. Neste caso, é de entendimento do ordenamento jurídico brasileiro que toda informação produzida pelo Estado é de caráter público e, portanto, o sigilo deve ater-se em casos excepcionais como em razão de sua imprescindibilidade para a segurança da sociedade e do Estado ou para resguardar a privacidade em caso de informação pessoal.

Portanto, as universidades federais, por pertencerem à esfera do governo federal, precisam adequar-se à realidade de atender às demandas locais de seus funcionários, dirigentes, professores e alunos, envolvidos diretamente em suas funções, atividades e processos de trabalho; mas também às necessidades em escala global, de possíveis pesquisadores ou interessados que necessitem de informações e serviços dessas entidades.

De acordo com Castells (2009), o poder do Estado é melhor exercido quando a construção de seu significado é amplificada por processos e procedimentos que permitam a sua dominação. As rotinas administrativas, como já demonstraram Rousseau e Couture (1998), são formas do Estado perpetuar a sua dominação sobre os membros de sua sociedade. A legitimação desse domínio depende em grande medida do consentimento obtido nos processos de ação comunicativa. Existe uma rede para comunicar as informações do Estado e suas ações, de validá-las, e de consolidar seus significados (CASTELLS, 2009, p. 36).

Essa dinâmica da comunicação mediada pelas tecnologias da informação permitem a combinação de variados tipos de comunicação em diferentes espaços. A autonomia sem precedentes daqueles que produzem e buscam informação multiplica e diversifica os pontos de entrada no processo de comunicação. No entanto, esse potencial de autonomia está modelado, controlado e cerceado pela concentração e interlocução daqueles que produzem os sistemas, os meios e os operadores de rede em todo mundo (CASTELLS, 2009, p. 189).

Essa diversidade (ainda que seja no consumo dos serviços) e o desenvolvimento das tecnologias de auto-comunicação de massas permitem uma maior iniciativa dos sujeitos envolvidos

nos processos informacionais. Embora seja necessário o controle dos fluxos informacionais a fim de organizar o conhecimento e os próprios processos de trabalho, mediados pelos documentos e pelas informações, é reivindicado cada vez mais transparência, participação e pluralidade nos sistemas de informação. Apesar da crescente concentração de poder, capital e produção no sistema global de comunicação, os sujeitos envolvidos nas práticas de comunicação reivindicam processos cada vez mais democráticos e que procurem atender a ampla maioria dos interessados.

Duarte (2009), ao apresentar o conceito de comunicação pública, afirma que esta é oriunda da ideia de comunicação governamental, no entanto com aspirações democráticas características da sociedade brasileira da década de 1980. Anteriormente, as referências de comunicação governamental possuíam viés autoritário como a organização da administração pública do DASP e as políticas de controle da informação durante o Estado Novo, entre 1937 e 1945. Outrossim, no período da Ditadura Civil-Militar, entre 1964 e 1985, o foco do Sistema de Comunicação Social no âmbito do governo focou-se na propaganda e censura.

A Constituição de 1988, a transformação do papel do Estado, o Código de Defesa do Consumidor, a terceirização e a desregulamentação, a atuação de grupos de interesse e movimentos sociais e o desenvolvimento tecnológico estabeleceram um sistema de participação e pressão que forçou a criação de mecanismos para dar atendimento às exigências de informação e tratamento justo por parte do cidadão em sua relação com o Estado e instituições, do consumidor com as empresas e entre todos os agentes sociais. Levou, por exemplo, ao surgimento do conceito de comportamento empresarial socialmente responsável no setor privado (mesmo que muitas vezes subordinado a estratégias comerciais), ao empoderamento do terceiro setor e a uma maior demanda por transparência no setor público. Todas as conquistas devem ser relativizadas, mas permitem identificar importante evolução (DUARTE, 2009, p. 42)

A comunicação pública, portanto, passa a ser uma demanda social do direito à informação, devendo ocorrer no espaço formado pelos fluxos de informação e de interação entre agentes públicos e atores sociais (governo, Estado e sociedade civil – inclusive partidos, empresas, terceiro setor e cada cidadão individualmente) em temas de interesse público. Ela trata de compartilhamento, negociações, conflitos e acordos na busca do atendimento de interesses referentes a temas de relevância coletiva.

Desse modo, o campo da comunicação pública inclui tudo que diga respeito ao aparato estatal, às ações governamentais, partidos políticos, terceiro setor e, em certas circunstâncias, às ações privadas. A existência de recursos públicos ou interesse público caracteriza a necessidade de atendimento às exigências da comunicação pública.

O poder é exercido fundamentalmente construindo significados na mente humana, de acordo com Castells (2009), e isso se dá pelos processos de comunicação que têm lugar nas redes e nos multimeios globais e locais de comunicação, incluindo a auto-comunicação. A era digital fornece a possibilidade do Estado atuar em diversos ambientes de forma sincrônica. Inserido nessa realidade,

o protocolo como sistema de registro, não atua mais de forma centralizada, concentrando suas funções em uma repartição. O imperativo dos processos comunicacionais contemporâneos faz com que o protocolo atue de forma descentralizada, interconectada por rede com as unidades de trabalho ao mesmo tempo. O ambiente eletrônico multifacetado se sobrepôs ao físico, no entanto, as funções de registro, autuação, acompanhamento dos fluxos de trabalho, classificação continuam sendo imprescindíveis aos arquivos.

A Universidade Federal Fluminense (UFF) é um exemplo de comunidade discursiva que tem enfrentado os novos desafios de gerir seus documentos no mundo digital e que vem sofrendo com as transformações que as universidades federais têm passado. Na última seção, apresentaremos nossa análise sobre este campo empírico. Outrossim, a nossa perspectiva acerca da indexação em arquivos de universidade é melhor aduzida, inclusive oferecendo algumas proposta para esta área.

7 RESULTADOS E DISCUSSÃO

Nesta seção apresentamos os resultados obtidos na análise do processo de indexação nos arquivos de universidade, considerando a implementação do Sistema Protocolo Integrado pelas universidades federais brasileiras. Em nosso recorte, optamos em investigar como tem sido a adesão dessas instituições ao SPI, sobretudo a partir do uso do software de governo, conhecido como SEI.

As três primeiras subseções estão voltadas aos relatos analíticos dos resultados, em especial da compreensão do nosso campo empírico, a Universidade Federal Fluminense e seus processos de trabalho.

Na última subseção, procuramos interpretar tais dados, direcionando o nosso olhar para as formas de tratamento temático da informação que estão sendo adotadas para a organização dos documentos de arquivo pelas universidades. Nesta parte, realizamos alguns apontamentos e indicações que possam contribuir para a *práxis* do processo de indexação, em especial na fase de produção documental nos arquivos de universidade.

Ressaltamos que não há qualquer pretensão de estabelecer modelos ou requisitos, e sim, elencar possibilidades de análise e representação dos documentos e de suas informações, fundamentadas pela hipótese principal que norteia esse trabalho: a indexação dos documentos de arquivo possui maior eficácia quando pautada nos usos em que os membros da comunidade discursiva fazem dos seus documentos. Na fase de produção, esses usos estão associados aos processos de trabalho, por conseguinte, a indexação pode partir do mapeamento destes.

7.1 O SISTEMA PROTOCOLO INTEGRADO E O SEI

Em 30 de dezembro de 2014 foi instituído pela portaria interministerial (Ministério da Justiça/Ministério do Planejamento, Orçamento e Gestão), nº 2.320, o Sistema Protocolo Integrado (SPI). Em consonância com a sociedade em rede, a iniciativa do governo federal brasileiro foi passar para uma plataforma eletrônica integrada suas funções e atividades, sobretudo aquelas que tramitam via documentos, sejam eles avulsos ou em processos.

De acordo com a portaria interministerial MJ/MP, nº 1.677, de 7 de outubro de 2015, que regula as atividades de protocolo no âmbito da Administração Pública Federal, os procedimentos das atividades de protocolo são de observância obrigatória no âmbito dos órgãos e entidades da Administração Pública Federal, o que se aplica também às universidades federais. Para tais fins, “(...) consideram-se atividades de protocolo o recebimento, a classificação, o registro, a distribuição, o controle da tramitação, a expedição e a autuação de documentos avulsos para formação de processos, e os respectivos procedimentos decorrentes” (BRASIL, 2015).

Logo, os documentos, avulsos ou em processos, que sejam produzidos no âmbito do órgão ou entidade e que demandem análise, informação, despacho, parecer ou decisão administrativa e que necessitem de tramitação, precisam ser registrados no Sistema Protocolo Integrado.

O SPI procura padronizar os procedimentos gerais para o desenvolvimento das atividades de protocolo no âmbito dos órgãos e entidades da Administração Pública Federal. A normativa ainda estabelece que, do mesmo modo, todos os documentos que sejam produzidos no âmbito do órgão ou entidade e que demandem análise, informação, despacho, parecer ou decisão administrativa e que necessitem de tramitação, recebam um Número Único de Protocolo (NUP) e tramitem no SPI.

Documentos digitais, avulsos ou processos, devem ser recebidos por meio de transmissão entre sistemas integrados, que pode ser realizada de duas maneiras, a critério do órgão ou entidade: 1) centralizada em uma unidade de protocolo/protocolizadora, que procederá o registro e encaminhará o(s) documento(s) para o destinatário; ou 2) diretamente para o destinatário.

De acordo com a portaria interministerial 1. 677, de 7 de outubro de 2015, são ações obrigatórias a cada unidade protocolizadora:

- a) Verificar se o documento, avulso ou processo, destina-se ao órgão ou entidade, independentemente do destinatário ser ou não ocupante do cargo indicado;
- b) Separar os documentos avulsos de caráter oficial daqueles de caráter particular;
- c) Separar os documentos, avulsos ou processos, de caráter oficial em ostensivos ou sigilosos;
- d) Identificar o documento, avulso ou processo, caracterizado como "urgente", que precisa ser priorizado para fins de classificação, registro e distribuição;
- e) Verificar se o documento, avulso ou processo, está acompanhado do(s) respectivo(s) anexo(s) e verificar a integridade dos arquivos recebidos;
- f) Verificar se o documento avulso será objeto de autuação ou não, conforme regulamentação interna do órgão ou entidade;
- g) Classificar o documento, avulso ou processo, relativo às atividades-meio de acordo com as classes aprovadas pelo CONARQ, e o relativo às atividades-fim de acordo com as classes elaboradas por cada órgão ou entidade e aprovadas pelo Arquivo Nacional; e
- h) Registrar o documento, avulso ou processo, em sistema informatizado ou excepcionalmente em formulário, no qual constarão os seguintes dados identificadores: a espécie/tipo do documento; o número e a data de produção do documento; a data de recebimento do documento; o identificador de que o documento é avulso ou processo; o Número Único de Protocolo – NUP; o número de anexo(s); o número de volume(s); o código de classificação e seu respectivo descritor ou o assunto a que se refere o documento; o remetente/interessado/representante legal; e o destinatário.

Na transmissão, deve-se observar o emprego de recursos para apoiar a confiabilidade, autenticidade e acessibilidade dos documentos transmitidos.

Já na distribuição dos documentos, avulsos ou processos, aos seus destinatários é obrigatório seus registros em sistema informatizado ou excepcionalmente em formulário. Estes registros deverão contemplar as seguintes informações: identificação do documento por meio do NUP; remetente/interessado/representante legal; destinatário; especificações do documento, avulso ou processo, encaminhado: espécie/tipo, número e data de produção; identificador de que o documento é avulso ou processo; número de anexos e número de volumes; código de classificação e o respectivo descritor ou o assunto a que se refere o documento; data do encaminhamento; identificação do responsável pelo encaminhamento; data do recebimento; identificação do responsável pelo recebimento; e providências a serem implementadas, quando couber. Esses procedimentos devem ser adotados tanto para documentos digitais, quanto não digitais.

No controle da tramitação dos documentos, deve ser observado se o registro no instrumento de encaminhamento foi realizado de forma correta com as informações previstas no formulário e supracitadas. É vedada a tramitação de documentos se o respectivo registro.

Na fase de expedição, os documentos digitais, avulsos ou processos, devem ser transmitidos (expedidos) por meio de sistemas integrados. Excepcionalmente, documentos avulsos digitais também podem ser expedidos por meio de correio eletrônico ou em mídias removíveis. Na expedição deverão ser seguidos os seguintes procedimentos:

- a) Classificar o documento, avulso ou processo, relativo às atividades-meio de acordo com as classes aprovadas pelo CONARQ, e o relativo às atividades-fim de acordo com as classes elaboradas por cada órgão ou entidade e aprovadas pelo Arquivo Nacional;
- b) Registrar o documento, avulso ou processo;
- c) Verificar se o documento está acompanhado do respectivo anexo;
- d) Registrar a expedição do documento;
- e) Transmitir o documento adotando medidas quanto à segurança, de modo que o mesmo não corra riscos de perda ou corrupção de dados.

Os mesmos procedimentos devem ser observados para documentos não digitais, no entanto, a transmissão, que se trata da forma de expedição por meio eletrônico nessa realidade, não é abordada. O cuidado encontra-se no envio para evitar perda ou extravio.

A autuação dos documentos para formação de processos é obrigatória quando o assunto, tema ou solicitação requerer análises, informações, despachos, pareceres ou decisões administrativas dos órgãos e entidades da Administração Pública Federal.

O processo pode ser iniciado pela unidade protocolizadora, quando existir regulamentação interna do órgão ou entidade, por solicitação da unidade administrativa, por meio de despacho da autoridade competente, ou por requerimento do interessado.

A autuação de processos não digitais somente poderá ser efetuada pela unidade protocolizadora. Nos processos digitais a autuação poderá ser realizada por usuário autorizado, diretamente no sistema informatizado, desde que o procedimento seja monitorado pela unidade protocolizadora.

Nota-se pelas atribuições dadas ao protocolo que, embora ele praticamente deixe de existir fisicamente, na realidade dos documentos nato digitais e da tramitação desses por sistemas eletrônicos de informação, a exemplo do próprio SEI, sua posição torna-se central na viabilização das atividades e funções dos órgãos e entidades da Administração Pública Federal. Não somente ele mantém a sua função de primeira unidade de arquivamento dos documentos, mas passa, com o SPI, a registrar, autuar, distribuir, acompanhar o trâmite e possibilitar a recuperação dos processos.

De acordo com o Ministério do Planejamento, Desenvolvimento e Gestão, o Sistema Protocolo Integrado (SPI) foi instituído para atender à demanda de transparência ativa por parte do Governo Federal, prevista na lei 12.527 de acesso à informação. Nesse sentido, o objetivo primeiro é permitir o acompanhamento pelo cidadão do andamento dos documentos via internet, possibilitando visualizar cada etapa de seu trâmite, seja intra ou inter órgãos e entidades da Administração Pública Federal. A tramitação dos documentos deve acontecer de forma integrada, sob a lógica da eficiência e eficácia, discurso consolidado no campo da administração e da gestão de documentos desde a década de 1980 no Brasil como foi demonstrado ao longo desta tese.

Os procedimentos do SPI normatizados pelas portarias nº 2.320 e nº 1.677 estão em consonância com a iniciativa do governo federal de estabelecer o Processo Eletrônico Nacional (PEN), segundo a Secretaria de Tecnologia da Informação e Comunicação do Ministério do Planejamento, Desenvolvimento e Gestão, trata-se de construir uma infraestrutura integrada entre órgãos e entidades da Administração Pública de processos e documentos administrativos eletrônicos.

O PEN é regulamentado pelo decreto nº 8.539, de 8 de outubro de 2015, que estabelece aos órgãos e às entidades da Administração Pública Federal direta, autárquica e fundacional o uso de sistemas informatizados para a gestão e o trâmite de processos administrativos eletrônicos. O artigo 5º deste decreto afirma que nos processos administrativos eletrônicos, os atos processuais deverão ser realizados em meio eletrônico, exceto nas situações em que este procedimento for inviável ou em caso de indisponibilidade do meio eletrônico cujo prolongamento cause dano relevante à celeridade do processo.

O dispositivo trata também das formas de autenticidade e fidedignidade dos processos administrativos eletrônicos, determinando os sinais de validação e corroboração, assim como suas formas. Então, a assinatura, a autoria, o certificado digital, o recibo eletrônico, as formas de captura são alguns dos critérios utilizados que conferem ao documento digital a chancela de original e com efeitos legais.

Embora haja grande preocupação com as novas formas de comunicação e os efeitos das novas tecnologias da informação, sobretudo com o impacto dos documentos nato digitais na mediação das funções e atividades do governo federal em relação à sociedade, o conceito de documento de arquivo não está restrito a seu suporte e o ordenamento jurídico regula tanto os atos da Administração Pública, quanto os processos oriundos destes meios.

O processo administrativo no âmbito da Administração Pública Federal é regulado pela lei 9.784 de 29 de janeiro de 1999. De acordo com o ordenamento jurídico brasileiro, o processo administrativo é a forma mais comum de atuação do Estado. Os atos estatais derivam de processos decisórios que precisam ser formalizados e estes são materializados por intermédio dos processos.

A sequência de formalidades não é aleatória: há uma ordem a ser observada, um itinerário a seguir. Ademais, os vários passos são necessários: não se pode dar o segundo passo sem que o primeiro tenha sido cumprido. Assim, o processo é o encadeamento necessário e ordenado de eventos (SUNDFELD, 2007, p. 92).

Portanto, o processo administrativo é dinâmico e encadeia atos e fatos que possibilitam a execução de atividades e funções. De acordo com Medauar (2008), o processo deve realizar-se rigorosamente de acordo com esse conteúdo sob pena de tornar-se viciado. Os agentes públicos precisam agir apenas segundo a regra da competência estabelecida pelo ordenamento jurídico. Nessa ótica, há um dever ao processo. O agente público recebe poderes que somente podem ser exercidos para o atendimento de certos deveres.

As competências administrativas são deveres-poderes, de acordo com Mello (2004), pois são atribuídas ao Estado, a seus órgãos e entidades e, por conseguinte, aos agentes neles investidos para que possam atender a certas finalidades e atender os interesses e demandas da coletividade. Cabe ressaltar que nem todo procedimento administrativo resulta em processo, pois há tarefas que não precisam ser realizadas com as formalidades exigidas neste rito. Como afirma Rocha (1997), o processo é um instrumento juridicamente criado para realizar uma determinada finalidade. Já o procedimento é a maneira de fazer ou agir que pode fazer parte do processo ou não. No caso da primeira hipótese, trata-se da forma de conduzir o processo a fim de otimizar a execução da atividade. No entanto, o processo administrativo exige a identificação do objeto e a forma de exercer as atividades administrativas.

Ainda segundo Rocha (1997), o processo administrativo é o que dá visibilidade à responsabilidade administrativa e permite que as atividades da Administração Pública sejam controladas em sua juridicidade e em sua legitimidade. Destarte, devido às características do processo administrativo, a maioria das atividades e funções da Administração Pública, sobretudo no âmbito da União, são realizadas por intermédio dele. E, portanto, existe certa delimitação dos processos de trabalho registrados pelos protocolos dos órgãos e entidades da Administração Pública por serem, em sua maioria, mediados pelos processos administrativos.

Para atender a nova realidade do Sistema Protocolo Integrado, o governo federal disponibilizou o Sistema Eletrônico de Informações (SEI), uma plataforma de gestão de processos e documentos digitais que engloba um conjunto de módulos e funcionalidades a fim de promover a eficiência administrativa.

O SEI é um software desenvolvido pelo IV Tribunal Regional Federal (TRF4) e adquirido pelo governo com o objetivo de padronizar o SPI. Fornecido gratuitamente para os órgãos e entidades da Administração Pública Federal, o SEI tem se tornado o principal sistema utilizado na implementação do SPI.

A padronização da tramitação dos processos administrativos e documentos digitais poderá proporcionar a integração da administração no âmbito do governo federal. A iniciativa é respalda pelo discurso de ganhos com agilidade, produtividade, transparência, redução de custos e satisfação do público usuário.

Apesar de não ser obrigatória a adoção do SEI, a instituição do SPI nos órgãos e entidades da Administração Pública Federal transformou sua adesão praticamente compulsória, pois é a ferramenta disponibilizada pelo Ministério do Planejamento, Orçamento e Gestão, que também instituiu a portaria do SPI. A adaptação dos demais sistemas de gestão ao SPI é de inteira responsabilidade das instituições que usam esses outros GEDs e SIGADs. No entanto, o SPI é modelado para ser integralmente hospiteiro ao SEI.

Então, as instituições que possuem seus próprios geds e sigads ficaram com a difícil escolha de manter seus sistemas, nos quais investiram recursos financeiros, humanos, tempo e tecnologia, e arcarem todo o ônus da adaptação ao SPI, sendo que a arquitetura do sistema, os dados, os códigos a serem usados são todos estabelecidos pela Secretaria de Tecnologia da Informação e Comunicação (STIC) do Ministério do Planejamento, Desenvolvimento e Gestão, ou abandonarem tais sistemas e aderirem ao SEI, software constantemente atualizado pela STIC a fim de possibilitar a tramitação dos processos eletrônicos nacionais e dos demais documentos registrados no SPI relativos às atividades e funções desses órgãos e entidades.

A grande maioria das universidades brasileiras adota o protocolo como sistema de registro a fim de comprovar e viabilizar suas atividades. O protocolo faz parte da gestão de documentos destas

instituições, capaz de racionalizar e controlar o fluxo documental. O registro, o controle do trâmite, a organização e a avaliação dos documentos vigentes são atribuições do protocolo que também permeiam a gestão de documentos. O protocolo facilita a padronização dos processos de trabalho quando faz parte de um programa mais amplo de gestão documental.

Na esfera dos arquivos, o protocolo como sistema de registro favorece a manutenção de características dos documentos de arquivo como a fidelidade, autenticidade e organicidade, além de fornecer um efetivo apoio à administração, sobretudo em instituições marcadas pela autonomia de seus institutos, faculdades, departamentos e coordenações.

7.2 AS UNIVERSIDADES FEDERAIS E O SEI

No Brasil, há 68 universidades federais, algumas procuraram implementar programas de gestão de documentos, sobretudo a partir da década de 1990 com os primeiros sistemas integrados de informação, com a formação de arquivos correntes e permanentes, além da formação das comissões de avaliação de documentos. Os primeiros manuais surgidos destacavam, sobretudo, os processos de classificação e avaliação no programa de gestão de documentos. Havia também certa preocupação com a padronização na produção dos documentos e o registro de seus trâmites. No entanto, essas questões acentuaram-se nos anos 2000 com o surgimento dos documentos nato digitais.

A Universidade Federal de Minas Gerais (UFMG), a Universidade Federal da Bahia (UFBA), a Universidade de Brasília (UnB), a Universidade Federal de São Carlos (UFSCar), a Universidade do Ceará (UFC) procuraram consolidar seus procedimentos de gestão de documentos em manuais. A UFMG procurou estabelecer padrões para a nomeação dos seus tipos documentais a luz dos códigos de classificação elaborados pelo CONARQ tanto para as atividades-meio, quanto para as atividades-fim. Outrossim, lista os formatos para a produção de documentos, além de estabelecer orientações gerais para o trâmite destes. Por último, trata da avaliação e das formas de arquivamento dos documentos (UFMG, 2013).

Já a UFBA concentra-se nas atividades, consideradas por seu corpo técnico como básicas, a serem seguidas na gestão arquivística de documentos. De acordo com sua publicação, **Normas para a Gestão de Documentos Arquivísticos**, tais atividades são: a de identificação, a classificação e a avaliação dos documentos produzidos pela UFBA. O instrumento destaca ainda a importância dos arquivos setoriais na tramitação e avaliação dos documentos, no entanto, não de forma aprofundada e não estabelece diretrizes para suas atividades (FREIXO; TOUTAIN, 2014).

A Universidade de Brasília trata de todo ciclo de vida documental em seu **Manual de Gestão de Documentos (2015)**, desde a produção e tramitação dos documentos, passando pelo registro e rotinas de expedição, pela classificação, uso e conservação e destinação, considerando

tanto a transferência, quanto o recolhimento. O **Guia Prático do SEI na UnB (2016)**, é um manual para o uso da plataforma que possui uma interface com os procedimentos de gestão de documentos padronizados consolidados em seu manual. Neste caso, trata-se da adaptação da gestão dos processos de trabalho e de documentos à plataforma do SEI. Em diversas etapas como “iniciar processo”, o guia recomenda a utilização dos expedientes previstos no manual de gestão de documentos da UnB. Essas medidas são importantes, pois coloca a instituição em uma posição estratégica de agente produtor e gestor de seus processos de trabalho e de seus documentos. Questão pertinente não somente para viabilizar suas atividades e funções na realidade que a cerca, mas também para garantir a sua própria autonomia como instituição social de educação.

A UFSCar em seu manual estabelece diretrizes para a padronização da produção, tramitação e destinação dos documentos da universidade. No entanto, foca-se na definição de protocolo e de suas atividades, em particular na autuação e formação de processos administrativos (UFSCAR, 2015).

Outras universidades produziram manuais específicos para padronizar seus protocolos e suas atividades como o caso da Universidade Federal de Santa Maria (UFSM), da Universidade Federal do Espírito Santo (UFES) e Universidade Federal do Pará (UFPA).

O manual da UFSM dirige-se ao estabelecimento de padrões na formação de processos. São listados todos os processos divididos em quatro grandes áreas: administração geral e as funções que especificam as universidades – extensão, ensino e pesquisa (ROSA, 2014).

A UFES procura estabelecer conceituações sobre documento, processo, unidade protocolizadora, estabelece as atividades do protocolo e suas diferenças daquelas desempenhadas pelo arquivo, assim como as etapas de operacionalização dos processos. Dentre as atividades de protocolo estipuladas estão: recebimento, análise, triagem, registro e distribuição. Acerca das etapas de operacionalização do processo estão: a numeração única, a autuação, a numeração de folhas, o despacho, a tramitação, a retirada de peças, o desentranhamento, o desmembramento, a juntada – seja por anexação ou apensação (UFES, 2014). Cabe destacar que apenas a autuação é de competência exclusiva do protocolo.

O **Manual de Protocolo, Tramitação e Arquivamento de Processos (2000)**, da UFPA trata-se de diretrizes de adaptação das atividades de protocolo como autuação, recebimento e distribuição dos processos ao sistema informatizado PTA que registra as tramitações dos processos da universidade.

As universidades federais do Rio de Janeiro não produziram manuais acerca de procedimentos das instituições de gestão de seus documentos, tampouco diretrizes para seus protocolos. De fato, a Universidade Federal do Rio de Janeiro (UFRJ), a Universidade Federal do Estado do Rio de Janeiro (UNIRIO), a Universidade Federal Rural do Rio de Janeiro (UFRRJ) e a

Universidade Federal Fluminense (UFF) não possuem programas de gestão de documentos, apesar de terem sistemas de registros e arquivos.

A produção, classificação, avaliação, destinação dos documentos não estão normatizadas por estas instituições. No entanto, as universidades federais do Rio de Janeiro acataram a portaria do MEC nº 1.261, de 23 de dezembro de 2013, que determina a obrigatoriedade do uso do Código de classificação e da Tabela de temporalidade e destinação dos documentos de arquivo relativos às atividades-fim das instituições federais de ensino superior, aprovados pela Portaria nº 92, de 23 de setembro de 2011, do Arquivo Nacional.

A resolução nº14 do Conselho Nacional de Arquivos (CONARQ), de 24 de outubro de 2001, que dispõe sobre o código de classificação e a tabela básica de temporalidade de documentos de arquivo para a Administração Pública referentes às atividades-meio, já era usada por essas instituições.

A UNIRIO, a UFRJ e a UFF procuraram padronizar seus documentos a partir da elaboração de manuais de atos e comunicações oficiais. O objetivo era estabelecer diretrizes gerais para a redação oficial de documentos com atenção aos princípios administrativos e às especificidades dos registros. Os três manuais procuram orientar e padronizar a forma e o conteúdo dos documentos a serem produzidos, visando a sistematização das rotinas de trabalho, contribuindo assim para o aprimoramento das atividades administrativas e possibilitando maior eficiência no fluxo de informação e eficácia na produção e gestão da informação produzida no âmbito da Universidade.

No entanto, as unidades de trabalho dessas instituições não seguiram as diretrizes desses manuais para a criação de seus documentos. Outrossim, a padronização da redação de tais documentos não garante por si só a gestão arquivística desses. Conforme foi abordado quarta seção desta tese, a gestão de documentos compreende as fases de padronização da produção documental; controle do fluxo; gestão de arquivos, informações e comunicações; controle da reprodução de documentos; identificação e descrição de séries documentais; estabelecimento de programas de avaliação e destinação (LLANSÓ I SANJUAN, 2006).

As universidades federais do Rio de Janeiro não possuem programas de gestão de documentos que compreendam todas as fases do ciclo vital, inclusive para garantir a economia e eficácia e, sobretudo, evitar a perda de registros vitais para tais instituições. Embora as tentativas de padronização na criação dos documentos; a adesão ao uso dos códigos de classificação e das tabelas de temporalidade, tanto para as atividades-meio, quanto das atividades-fim, elaboradas pelo CONARQ, sejam iniciativas importantes para a gestão de documentos dessas instituições, não são medidas garantidoras da gestão arquivística adequada e, por conseguinte, a recuperação da informação para tomada de decisão ou do documento como evidência de ação.

Com a obrigatoriedade da implementação do Sistema Protocolo Integrado pelos órgãos e entidades da Administração Pública Federal, as universidades precisaram repensar seus processos de trabalho e a gestão de seus documentos. As instituições que não possuíam programas de gestão de documentos e padronização de suas atividades a exemplo da Universidade Federal Fluminense, nosso campo empírico, estão apresentando grandes dificuldades na implementação do Sistema Eletrônico de Informação (SEI) que trata do trâmite dos documentos na fase corrente e dos fluxos de trabalho. O grande gargalo na fase de produção documental da UFF, no que concerne a gestão arquivística, leva a grandes perdas informacionais para a instituição.

Como afirmou o Grupo de Trabalho SEI Avaliação do Arquivo Nacional, o software é um produto para possibilitar a tramitação do processo eletrônico nacional, desenvolvido pelo TRF4 para tratar seus próprios processos judiciais eletrônicos, mas que foi escolhido pela Administração Pública Federal no contexto do Sistema Protocolo Integrado.

O SEI como ferramenta permite a produção, edição, assinatura e trâmite de documentos dentro do próprio sistema. Proporciona a virtualização de processos e documentos, permitindo a atuação simultânea de várias unidades, ainda que distantes fisicamente, em um mesmo processo, reduzindo o tempo de realização das atividades. No entanto, a padronização dessas etapas e o controle do fluxo documental é algo que precisa ser customizado por cada instituição que adotá-lo.

Há iniciativas colaborativas de adequação do SEI ao modelo de requisitos para sistemas informatizados de gestão arquivística de documentos, **e-ARQ Brasil**, sobretudo com a maior adesão e uso do software de governo. No entanto, importantes procedimentos que devem ser adotados para uma consistente gestão arquivística possuem baixa adesão no SEI¹⁸.

Pela análise realizada, os requisitos atendidos são aqueles encontrados em sistemas de gestão eletrônica de documentos ou mesmo de ferramentas de controle de fluxo documental. A maior preocupação da equipe técnica do Arquivo Nacional mostra-se justamente com as questões relacionadas ao armazenamento e segurança dos documentos. Aspectos fundamentais para garantir a autenticidade dos documentos produzidos, tramitados e mantidos no sistema.

O SEI contribui com a identificação, padronização e controle dos fluxos de trabalho da Administração Pública Federal, permitindo inclusive a interconectividade entre instituições e entidades que usam o mesmo sistema. Todavia, cada fluxo de trabalho registrado no SEI recebe o Número Único de Protocolo (NUP), reunindo todos os documentos referentes a tal fluxo em um processo único. Nem todas as tarefas referentes à execução de uma atividade precisam ser registradas, muito menos formar processo.

Portanto, o SEI é uma ferramenta que melhor atende às instituições e órgãos que possuem programas de gestão de documentos em funcionamento, evitando a produção de documentos

¹⁸ A tabela comparativa elaborada pelo GT SEI Avaliação do Arquivo Nacional encontra-se no anexo H.

desnecessários, facilitando o controle e padronização do fluxo de trabalho, garantindo a segurança dos documentos desde sua captura, passando pelo registro até a sua destinação.

A Universidade Federal Fluminense aderiu ao PEN e começou a inserir seus processos no SEI em 12 de setembro de 2017 com a inclusão do reembolso de bilhetes de passagem. Para implementar o SPI, via SEI, a UFF instituiu, pela portaria 58.196 de 23 de fevereiro de 2017, uma comissão mista na qual participam membros de diversas unidades de trabalho. Arquivistas lotados nos protocolos e no arquivo permanente da instituição também fazem parte de sua composição. Dentre as unidades representadas estão: a Coordenação de Arquivos; a Reitoria; as Pró-reitorias de Administração; Planejamento; de Graduação; de Gestão de Pessoas; de Extensão; de Assistência Estudantil; de Pesquisa, Pós-Graduação e Inovação. Ademais, há representantes da Superintendência de Relações Internacionais; de Tecnologia; de Documentação; de Arquitetura e Engenharia; de Comunicação Social. O Hospital Universitário Antônio Pedro; o Centro de Artes; a Editora e o Conselho de Curadores também possuem membros na comissão.

Na Universidade Federal Fluminense, os protocolos não atuam em suas unidades de trabalho. O sistema de registros criados pela UFF baseia-se na geografia de seus campi. A estrutura dos protocolos da Universidade Federal Fluminense é constituída pela Gerência Plena de Comunicações Administrativas (GPCA) que funciona como uma espécie de protocolo central que registra as atividades produzidas pelos setores subordinados diretamente à Reitoria como a Secretaria Geral, as Assessorias, a Ouvidoria, o Gabinete, a Diretoria de Relações Internacionais. As atividades administrativas das Pró-reitorias (PROGRAD, PROEX, PROPPI, PROAD, PROAES, PROGEPE, PROPLAN) e de seus setores são registradas também pela GPCA.

Subordinados a esta gerência estão os protocolos setoriais, divididos por áreas geográficas, de acordo com os campi. A UFF possui quatro protocolos setoriais: o Protocolo Setorial do Gragoatá (PSG), o Protocolo Setorial da Área de Saúde (PSAS), o Protocolo Setorial da Praia Vermelha (PSPV) e o Protocolo Setorial do Valonguinho (PSV).

O PSG além de atender os cursos de Arquivologia, Biblioteconomia e Documentação, Ciências Econômicas, Ciências Sociais, Cinema e Audiovisual, Comunicação Social, Direito, Educação Física, Estudos de Mídia, Filosofia, História, Letras, Pedagogia, Produção Cultural, Psicologia, Antropologia, Artes e Serviço Social; tramita os documentos referentes à Prefeitura da Universidade, o Restaurante Universitário, a Superintendência de Engenharia e a Superintendência de Documentação.

Já o PSPV registra a documentação dos cursos de Arquitetura e Urbanismo, Ciência da Computação, Engenharia Agrícola, Engenharia Civil, Engenharia de Produção, Engenharia de Recursos Hídricos e do Meio Ambiente, Engenharia de Telecomunicações, Engenharia de Petróleo, Engenharia Elétrica, Engenharia Mecânica, Engenharia Química, Física, Geofísica, Ciência

Ambiental, Desenho Industrial, Sistema da Informação, Tecnologias em Sistemas de Computação e Geografia.

O PSAS é responsável pela autuação e trâmite dos documentos dos cursos de Enfermagem, Farmácia, Medicina, Medicina Veterinária e do Instituto da Saúde da Comunidade.

O PSV atende os cursos de Administração, Biomedicina, Ciências Biológicas, Ciências Contábeis, Estatística, Matemática, Nutrição, Odontologia, Química, Química Industrial, Turismo, Ciências Atuariais, Relações Internacionais, além do Instituto de Estudos Estratégicos e o próprio Instituto Biomédico.

Estas unidades de informação eram subordinadas ao Núcleo de Documentação (NDC), da mesma forma que as demais como as bibliotecas, a Coordenação de Bibliotecas, o arquivo intermediário, o arquivo permanente e a coordenação de arquivos, desde 2004, quando houve a reestruturação dos setores da Universidade Federal Fluminense, com a extinção dos Centros Universitários.

Em março de 2011, após a reestruturação interna, o núcleo transformou-se em Superintendência de Documentação, os protocolos e demais Serviços de Comunicações Administrativas (SCA) passaram a estar subordinados à Pró-reitoria de Administração¹⁹. Esta decisão rompeu, ao menos administrativamente, com a organicidade dos documentos de arquivo. Há uma quebra da cadeia custodial, pois a Coordenação de Arquivos (CAR), responsável pelos arquivos intermediário e permanente, não possui qualquer ingerência nos processos de produção e avaliação dos documentos da UFF.

Ademais, os processos administrativos que tramitam nos protocolos setoriais são destinados a um arquivo próprio após a sua vigência. O arquivo de processos da UFF assemelha-se com um depósito onde tais processos esperam sua avaliação e destinação para serem descartados ou incorporados aos arquivos intermediário ou permanente. No entanto, essas funções arquivísticas não são sistemáticas, o que acaba por gerar uma grande massa documental acumulada, além da difícil recuperação da informação dos processos já tramitados.

Igualmente, há diversas atividades que produzem importantes tipos documentais que não estão sendo registradas pelos protocolos como os programas de disciplinas, as folhas de frequência dos alunos inscritos nas disciplinas, as atas de reuniões dos departamentos e colegiados, entre outros. Diversas atividades que não constituem processos administrativos realizadas pelos departamentos, colegiados e institutos não estão sendo registradas pelos protocolos e muitas decisões não estão sendo incorporadas aos sistemas de informação. A falta de gestão integrada dos documentos da Universidade Federal Fluminense tem levado a importantes perdas informacionais e

¹⁹ Informações retiradas do site da Superintendência de Documentação da Universidade Federal Fluminense que relata o histórico institucional. Disponível em <<http://www.ndc.uff.br/content/institucional>>. Acesso em: 22 de setembro de 2017.

ao extravio de documentos que não permitem inclusive um melhor desenvolvimento dos processos de trabalho.

7.3 A UFF E O MAPEAMENTO DOS PROCESSOS DE TRABALHO

Com a adesão do SEI pela UFF, a Comissão estabelecida sentiu a necessidade de mapear seus processos de trabalho para adaptá-los ao software SEI. No entanto, não houve uma padronização dos métodos para a identificação dos processos de trabalho, em um primeiro momento. Cada unidade de trabalho que possuía representantes na Comissão passou a fazer o esforço de descrever seus processos de trabalho.

A metodologia que se mostrou mais adequada para a Comissão foi a realizada pela Pró-reitoria de Graduação. O então pró-reitor, José Rodrigues de Faria Filho, professor do curso de Engenharia de Produção, mobilizou uma equipe de estagiárias para realizar tal identificação que optaram pela engenharia de processos para mapear os processos de trabalho da PROGRAD.

De acordo com Grover e Kettinger (2000), a engenharia de processos é uma espécie de arquitetura que facilita o entendimento, a análise e a melhoria das ações de trabalho dentro e entre as organizações com objetivos como: uniformizar o entendimento da forma de trabalho; gerar uma cultura de integração das atividades; analisar e melhorar o fluxo de informações; explicitar o conhecimento sobre os processos; realizar análise de gestão organizacional; simular cenários e apontar indicadores que facilitem a tomada de decisões.

A engenharia de processos, teoria bastante difundida na Engenharia de Produção, sobretudo para compreender os sistemas de produção (CAULLIRAUX; CAMEIRA, 2000), procura estabelecer orientações metodológicas para as funções de uma organização a partir de uma ótica de atividades sequenciadas lógico-temporalmente (RUMMLER; BRACHE, 1992). Portanto, a organização administrativa se dá por processos de trabalho que agrupados irão desempenhar as funções de determinada organização. Para Salerno (1999), um processo de trabalho é:

Uma cooperação de atividades distintas para a realização de um objetivo global, orientado para o cliente final que lhes é comum. Um processo é repetido de maneira recorrente dentro da empresa. A um processo correspondem: Um desempenho (performance), que formaliza o seu objetivo global (um nível de qualidade, um prazo de entrega etc.); Uma organização que materializa e estrutura transversalmente a interdependência das atividades do processo, durante sua duração; Uma co-responsabilidade dos atores nesta organização, com relação ao desempenho global; Uma responsabilidade local de cada grupo de atores ao nível de sua própria atividade (SALERNO, 1999, p. 105).

A arquitetura para o mapeamento dos processos de trabalho realizada pela PROGRAD se assemelha com o modelo *Business Process Model and Notation* (BPMN). Neste esquema, os processos de trabalho e suas etapas são identificados em notações que utilizam piscinas (pools) e raias (swimlanes). Cada *pool* pode conter apenas um processo de trabalho, já as *swimlanes* são

notações referentes às tarefas para a execução daquele processo de trabalho. Cada raia identificação em sua notação os responsáveis pela execução de cada tarefa no processo de trabalho (WHITE, 2004)²⁰.

Neste modelo de mapeamento de processos, a *pool* trata do processo de trabalho dispensa de disciplina. Cada *swimlane* corresponde ao responsável pelos chamados elementos do processo. Neste caso, o processo de trabalho tem início com o requerimento do aluno. Então, há dois procedimentos identificados como de sua responsabilidade – preencher o requerimento de dispensa de disciplinas; entregar os documentos comprobatórios digitalizados.

Na segunda *swimlane* encontra-se a Secretaria da Coordenação do Curso que por sua vez precisa receber o processo de dispensa de disciplina e encaminhá-lo para a Coordenação do Curso ou para o Departamento de Ensino. Na terceira *swimlane* há a Coordenação de Curso que analisa a documentação; assina a decisão e informa sobre a decisão à Secretaria. Caso haja o deferimento do pedido, a Coordenação necessita incluir as disciplinas dispensadas no histórico do aluno. Se o pedido for indeferido, a Secretaria da Coordenação comunicará a decisão ao aluno para que este possa recorrer dentro do prazo previsto. Na quarta *swimlane* há o Departamento de Ensino que designará o docente responsável em elaborar o parecer técnico; anexará tal parecer e encaminhará o processo à Coordenação de Curso. Na última *swimlane* está o Colegiado de Curso que deverá analisar o recurso impetrado pelo aluno, se houver, e emitir parecer acerca da decisão de acatá-lo ou não.

Estas são as tarefas e os responsáveis identificados pela metodologia BPMN de análise de processos. Há a previsão da elaboração do requerimento de dispensa de disciplinas; da decisão fundamentada da Coordenação; do recurso do aluno; do parecer técnico do professor responsável; do parecer com decisão acerca do recurso impetrado pelo aluno. Esses são os documentos previstos na composição do processo administrativo dispensa de disciplina.

Com o desenvolvimento das atividades pela Comissão de Implementação do SEI da UFF, essa metodologia passou a ser adotada pelas unidades de trabalho que tinham processos de trabalho a serem inseridos no SEI, o que foi facilitado pela adoção do software *Bizage Modeler*, específico para modelar mapas de processos pela metodologia BPMN. Nem todas as tarefas mapeadas nos processos de trabalho são registradas no SEI, justamente por não haver necessidade de evidências e para evitar o acúmulo de informações desnecessárias ao desenvolvimento dos fluxos de trabalho. No entanto, tem sido um exercício importante para as unidades de trabalho da Universidade Federal Fluminense pensar suas rotinas administrativas, inclusive para melhorar seus processos de trabalho.

De acordo com o Relatório de Implementação e Acompanhamento do Sistema Eletrônico de Informações na Universidade Federal Fluminense (2017), a ideia de implantar o SEI no âmbito da

²⁰ Um exemplo de aplicação do modelo BPMN pela UFF encontra-se no anexo I.

UFF nasceu a partir de discussões no Fórum Nacional de Pró-Reitores de Planejamento e de Administração das Instituições Federais de Ensino Superior – FORPLAD, realizado em 10 de novembro de 2016, em Porto Seguro - BA. Na ocasião, foi apresentado o número de instituições que haviam implantado o SEI e/ou estavam com sua implantação em curso, destacando-se a necessidade de informar através de ofício ao Ministério do Planejamento, qual seria a solução de processo eletrônico adotado pelas Universidades; se já existia uma solução implantada ou não, e em caso afirmativo, se todos os processos já eram eletrônicos ou se estava em expansão, ocorrendo gradativamente.

A UFF desde 2012 estava trabalhando no desenvolvimento do Sistema Integrado de Gestão Administrativa (SIGADOC). Tratava-se de uma adaptação realizada pela STI ao sistema desenvolvido pela Seção Judiciária do Rio de Janeiro (SJRJ) que fora cedido à UFF como código livre e que começou a ser implementado em 29 de fevereiro de 2016 por meio da portaria nº 55.750. Ao avaliar se continuaria com o SIGADOC e adaptaria tal sistema às necessidades do Sistema Protocolo Integrado, a STI elaborou um parecer técnico a pedido do Comitê de Gestão da UFF no qual apontou as seguintes dificuldades na continuidade de operacionalizar tal sistema (UFF, 2017, p. 8):

- Suporte insuficiente da Seção Judiciária do Rio de Janeiro - TRF 2ª Região;
- Necessidade de desenvolvimento interno para atendimento das demandas do PEN;
- Não ter sido adotado amplamente por outras instituições, o que dificulta a troca de experiências e a resolução de problemas;
- Uso de tecnologia de banco de dados proprietária;
- Impossibilidade de suporte pelo Ministério do Planejamento.

Já em relação à adesão ao SEI, o mesmo parecer elaborado pela STI aponta como aspectos positivos:

- Suporte do Ministério do Planejamento, Desenvolvimento e Gestão;
- Direito de uso gratuito do software;
- Cooperação entre os diversos órgãos para resolução de problemas;
- Treinamento estruturado e gratuito para os Administradores, Usuários e Corpo Técnico através da Escola Nacional de Administração Pública (ENAP);
- Desenvolvimento de Novos Módulos Previstos para atendimento de demanda Legislativa e da Comunidade do SEI;
- Uso de tecnologias amplamente difundidas, o que facilita a sua instalação/manutenção por parte da Universidade e pelo fato de ser parte integrante do Processo Eletrônico Nacional;

- Solução ofertada e utilizada pelo MPDG, com inclusão de NUP (Número Único de Processo);
- Cumprimento do decreto nº 8.539 de 8 de outubro de 2015 para implantação da gestão eletrônica de documentos na Universidade.

O Comitê de Governança Corporativa da UFF, o Comitê de Gestão da Informação da UFF, a Pró-Reitoria de Planejamento, a Pró-Reitoria de Gestão de Pessoas, a Pró-Reitoria de Administração/GPCA, a Superintendência de Documentação, e a Superintendência de Tecnologia da Informação, levando em consideração: o cumprimento da exigência e do prazo legal apresentado e previsto pelo Decreto nº 8.359 de 8/10/2015; o imperativo de uma administração moderna e eficiente, que incorpore tecnologias de gestão consagrado nas avaliações das universidades; além do princípio neoliberal de maximizar o uso dos recursos públicos e melhor desempenho da gestão, e atender aos padrões de conformidade dos órgãos federais de controle; deliberaram pela adoção do SEI e do PEN em 3 de janeiro de 2017. O acordo de cooperação técnica entre o Ministério do Planejamento, Desenvolvimento e Gestão e a Universidade Federal Fluminense foi firmado em 17 de abril de 2017²¹.

A Universidade Federal Fluminense não elaborou seu próprio manual para a implementação dos processos no SEI ou seu uso por sua comunidade. Em seu site são disponibilizados dois links para dois manuais voltados aos usuários – o primeiro elaborado no âmbito do curso promovido a partir da cooperação técnica entre o Centro de Apoio ao Desenvolvimento Tecnológico (CDT) da Fundação Universidade de Brasília (FUB); o Laboratório Latitude e a Escola Nacional de Administração Pública (ENAP). Este manual aborda os procedimentos operacionais padrões. Já o segundo manual do usuário foi elaborado pelo Departamento de Governo Digital da Secretaria de Tecnologia da Informação e Comunicações do Ministério do Planejamento, Desenvolvimento e Gestão, disponibilizado no site **software público brasileiro**.

Com exceção da breve introdução sobre o que é o SEI, suas facilidades e as boas práticas sobre a utilização do sistema; o segundo manual aborda os mesmos procedimentos do primeiro: acessar o SEI; receber processo; inserir notação; atribuir; iniciar; excluir; relacionar; anexar; sobrestar; duplicar; enviar; concluir processo. Outras funcionalidades como criar, consultar, alterar e excluir modelo de documento; inserir ponto de controle; gerar versão de documento; dar ciência; assinar; encaminhar documento para assinatura em outra unidade; incluir processos em bloco; além de criar e disponibilizar base de conhecimento também se encontram nesses manuais.

A UFF criou e tem atualizado uma base do conhecimento a partir de uma funcionalidade do SEI que permite descrever as etapas de um processo de trabalho e anexar modelos de documento

²¹ As etapas pré-estabelecidas pelo Ministério do Planejamento, Desenvolvimento e Gestão e assumidas pela UFF encontram-se no anexo J.

relacionados aquele “tipo de processo”. Assim, a universidade disponibilizou para cada processo inserido no SEI o mapeamento e sua descrição, realizados pelo método BPMN da engenharia de processos, tendo sido implementados até o presente momento os processos de: reembolso de bilhetes de passagem (024.92); adicional de insalubridade (024.134); adicional de periculosidade (024.133); doação de bens móveis (033.13); autorização para dirigir veículo oficial (042.192); segunda via de diploma de graduação (125.421); segunda via de histórico escolar de graduação (125.41); primeiro cadastro de auxílio-transporte (024.92); cancelamento de auxílio-transporte (024.92); recadastramento de auxílio-transporte (024.92); alteração de auxílio-transporte (024.92); suprimento de fundos (052.22); pagamento de fatura (52.22); pagamento de PASEP (024.151); serviços gráficos (032); afastamento de docente no país (029.21); progressão por capacitação (023.03); afastamento de técnico no país (029.21); programa de qualificação (024.92); exoneração/nomeação de cargo (023.11); dispensa/designação de chefia (023.12 e 023.14); dispensa/designação de substituto eventual de chefia (cargo comissionado) (023.12 e 023.14); dispensa/designação de substituto eventual de chefia (função gratificada) (023.14); pagamento de substituição de chefia (024.121); dispensa de aluno irregular no Enade (110); afastamento de docente no exterior com ônus (029.222); afastamento de docente no exterior sem ônus (029.221); afastamento de técnico no exterior com ônus (029.222); afastamento de técnico no exterior sem ônus (não foi atribuído código de classificação); remanejamento de função (023.02). Estão previstos 257 processos a serem implementados pela UFF no SEI.

A principal vantagem propagada pela UFF a sua comunidade acerca da adesão ao SEI foi o fim do processo em papel, diminuindo o impacto ambiental, desburocratizando as rotinas administrativas e proporcionando economia e eficácia nos fluxos de trabalho²². No entanto, a Comissão ao iniciar o processo de implementação do SPI e PEN via SEI percebeu que o bom funcionamento do sistema requer certos procedimentos de gestão administrativa e gestão de documentos que a Universidade ainda não havia realizado. Ademais, algumas questões políticas e jurídicas não tão bem definidas para a comunidade acabam por dificultar a implementação de certos processos no SEI.

A Universidade Federal Fluminense está a um passo de entrar em uma nova era na sua rotina de trabalho com a implantação do Sistema Eletrônico de Informações (SEI). Os principais objetivos da plataforma digital são agilizar a tramitação e o desempenho dos processos na UFF, aumentando a celeridade, produtividade, transparência e a satisfação do usuário. O uso do sistema representará não só a ampliação da sustentabilidade ambiental, como também a redução de custos no setor público, uma vez que tem como meta eliminar o papel como suporte físico para documentos institucionais nos diferentes setores da universidade (SEI GARANTIRÁ..., 2017).

²² Folder informativo no anexo K.

No entanto, na fase de mapeamento de processos, a Comissão detectou diversos gargalos administrativos que prejudicavam a implementação dos processos eletrônicos e o fluxo de trabalho no SEI. Uma das dificuldades apontadas foi em relação à lotação dos funcionários e dos administradores em cargos de chefia. Há diversos servidores trabalhando em unidades que não estão lotados em sua origem, do mesmo modo, há também chefias cujas funções comissionadas pertencem a outras unidades.

O SEI, da forma em que foi desenvolvido, obriga que a lotação de origem e a lotação de exercício sejam as mesmas no SIORG, de modo que os processos relativos a determinadas unidades de trabalho só possam tramitar nelas. Os processos pertencem às unidades de trabalho responsáveis pela execução de suas ações. Os gestores do processo precisam ter seus SIORGs nas unidades de trabalho correspondentes para receberem tal atribuição. Apenas os servidores lotados nessas unidades de trabalho podem ter processos atribuídos para desenvolver tais ações. Portanto, os servidores, que seriam responsáveis pela execução de determinadas tarefas nos processos de trabalho, não podem realizá-las como anteriormente, pois as lotações de origem e função estão diferentes. A Superintendência de Tecnologia da Informação (STI) e a Coordenação de Pessoal Técnico Administrativo começaram a realizar os ajustes de lotação e de designação de chefia em 2018. Entretanto, há questões políticas e de escassez de recursos humanos que atrapalham tais ajustes.

A questão abriu uma discussão acerca da força de trabalho, ou seja, se há servidores com lotações de origem e exercício diferentes, estes em algum momento foram cedidos. Então, criou-se a problemática se eles devem retornar aos seus setores ou haver alguma contrapartida de força de trabalho.

Outro dilema enfrentado pelas universidades federais na implementação do SEI tem sido em relação ao conflito de competências entre as instituições e as organizações sociais (Oss). No contexto dos processos de mundialização e financeirização do capital e de aplicação das políticas de conteúdo neoliberal, as formas de controle, gestão e organização do trabalho, juntamente com a redefinição do papel do Estado e da regulação sobre o mercado de trabalho, constituíram um novo regime de acumulação, em que a flexibilização/precarização do trabalho passou a ser uma estratégia central.

Na atualidade, a precarização deixou de ser algo residual ou periférico para se institucionalizar, ela se transforma em elemento central, em condição de flexibilidade, portanto, fundamental do ponto de vista da produtividade e da competitividade das instituições (DRUCK, 2013). No serviço público brasileiro, a terceirização, uma realidade já consolidada nas atividades-meio, passará também a estar presente nas atividades-fim a partir da reforma trabalhista instituída pela lei 13.467 de 13 de julho de 2017 e pela decisão do Supremo Tribunal Federal do dia 30 de

agosto de 2018 que pacificou o entendimento de permissão para a terceirização das atividades-fim pela lei.

Nas universidades federais, a exemplo dos hospitais públicos, a forma mais usual de terceirização tem sido por meio das Organizações Sociais (Oss) que são contratadas para desenvolver determinadas atividades dessas instituições, dentre as principais podemos citar segurança, limpeza, manutenção e gestão. Importante destacar que 40 dos 50 hospitais universitários passaram a ser administrados pela Empresa Brasileira de Serviços Hospitalares (EBSERH) (40 DOS 50 HOSPITAIS..., 2018).

A criação da EBSERH está assentada no discurso ideológico neoliberal de que a crise na saúde advém da má gestão dos serviços. Por isso, faz-se necessário reestruturar o Estado com vistas a se tornar eficiente e de qualidade, introduzindo mecanismos de mercado na gestão pública. A empresa foi instituída no contexto de políticas de ajuste estrutural de flexibilização das relações de trabalho e privatização dos serviços sociais públicos como saúde e educação.

Ainda que as OSs sejam uma realidade nas universidades federais, é importante destacar que há diferenças entre elas e a EBSERH no tocante ao ordenamento jurídico em relação à Administração Pública. Embora tanto as Oss, quanto a EBSERH sejam reflexos da privatização dos serviços prestados pelas universidades às suas comunidades, a Empresa Brasileira é entendida como ente estatal, faz parte da administração pública indireta e sua gestão continua sendo pública. Por conseguinte, seu regime jurídico é híbrido. Já as OSs, o regime jurídico e a gestão são privados, mas prestam serviços de responsabilidade do Estado.

A lei 12.550 de 15 de dezembro de 2011 que autoriza a criação da Empresa Brasileira de Serviços Hospitalares traz uma configuração complexa, pois possui um caráter ambivalente, tecnicamente pertence tanto à esfera pública, quanto à privada, criando hiatos na regulação de seus serviços (SODRÉ *et al.*, 2013). Instituída sob a alegação de trazer maior autonomia para a gestão dos hospitais públicos pela desburocratização no uso dos recursos, flexibilização das relações de trabalho e modernização da gestão, trazendo os princípios das organizações privadas, a EBSERH é gerida por fundamentos do direito privado, mas atua muitas vezes como parte da estrutura das universidades como se fosse uma unidade de trabalho.

A Universidade Federal Fluminense aderiu à forma de gestão dos hospitais federais via contratação da Empresa Brasileira de Recursos Hospitalares em 2016. Metade dos servidores que trabalham no Hospital Universitário Antônio Pedro (HUAP) foram cedidos a nova empresa que passou a ser responsável por tais recursos humanos. No entanto, pela sua forma híbrida de atuação, há diversas ações no Supremo Tribunal Federal questionando a sua constitucionalidade, justamente pela falta de lei complementar que defina suas áreas de atuação. Essas relações jurídicas indefinidas em relação à atuação da gestão da EBSERH e dos setores do HUAP que ainda respondem

diretamente à Universidade refletem na própria implementação de seus processos no SEI. Os gestores da EBSEH estão sem acesso ao Sistema por justamente não mais pertencerem ao quadro de servidores da UFF. É um entrave não somente técnico, mas jurídico e político. O próprio hospital universitário atravessa por relações conflitantes pela adesão da gestão da EBSEH não ser uma unanimidade entre a comunidade da UFF.

Em âmbito nacional, diversas entidades se manifestaram contra o modelo de gestão hospitalar via EBSEH como a Frente Nacional contra a Privatização da Saúde, a Federação de Sindicatos de Trabalhadores Técnicos-administrativos em Instituições de Ensino Superior Públicas do Brasil (FASUBRA) e o Sindicato Nacional dos Docentes das Instituições de Ensino Superior (ANDES-SN). De acordo com essas entidades, a gestão da EBSEH fere à autonomia universitária, precariza as relações de trabalho dos servidores e monetariza a saúde, o que deve ser um direito e não um produto.

O Sindicato de Trabalhadores em Educação da Universidade Federal Fluminense (SINTUFF) e a Associação dos Docentes da Universidade Federal Fluminense (ADUFF) também se posicionaram contra a EBSEH. Dentre as questões apontadas por estas entidades estão o desrespeito da empresa ao contrato com a Universidade, sobretudo em relação à reposição de força de trabalho, reabertura de serviços e ampliação de leitos. O hospital ao deixar o modelo de administração por autarquia para ingressar no modelo de empresa pública de direito privado, esvazia duas importantes funções da universidade: a prestação de serviço de saúde público a sua comunidade – pelo não cumprimento do contrato - e o ensino de medicina – a Empresa prioriza gestores e não docentes-médicos.

Esses entraves políticos refletem diretamente no próprio desenvolvimento das atividades no hospital universitário e no fluxo de trabalho. Ademais, por haver duas administrações coexistindo, funcionários da EBSEH e servidores da UFF por vezes realizam seus processos de trabalho de forma diferente e suas carreiras também são geridas de forma diferente. Há os contratos regidos pela CLT dos empregados públicos da EBSEH e os servidores que respondem ao Regime Jurídico Único, na qual a PROGEPE é responsável. Podemos citar como exemplo os processos de adicional de insalubridade; adicional de periculosidade; auxílio transporte; auxílio qualificação entre outros que tramitam de forma diferente.

Ademais, o mapeamento de processos tem retomado questões políticas não pacificadas na comunidade como a extinção de cargos de chefia nas universidades com recolhimento das vagas pelo Ministério da Educação (MEC). Em contrapartida, os cargos de gestão com maior remuneração passaram para a EBSEH, sendo pagos com os recursos da própria Universidade. A questão é vista pelas entidades contra tal forma de gestão como um financiamento a precarização das relações de trabalho na Universidade pela própria instituição.

Ademais, a cultura do papel ainda é presente na UFF. A falta do programa de gestão de documentos e a autonomia das unidades de trabalho faz com que alguns protocolos funcionem quase de maneira independente como é o caso do protocolo do HUAP. A extinção de cargos e a precarização da força de trabalho fazem com que alguns postos de trabalho fiquem sem administradores nomeados e, por conseguinte, com responsabilidades atribuídas de maneira informal como é o caso da chefia provisória do protocolo do HUAP. O hospital possui duas unidades de protocolo – um referente ao hospital em si que registram os prontuários dos pacientes e que não há chefia nomeada; e outra referente às questões dos servidores e da escola de medicina. A falta de padronização de procedimento faz com que o mesmo processo seja aberto no SEI e no formato físico, tramitando de duas formas diferentes com desdobramentos diferentes e pouco efetivos.

O Sistema Protocolo Integrado (SPI) tem sido implementado na UFF de forma descentralizada. Os protocolos setoriais (físicos) servirão para tramitar os processos cujos requerentes não tenham acesso ao SEI e com ilhas de digitalização – para transformar documentos que ainda estejam em formato físico em digitais e que necessitam serem apensados nos processos eletrônicos. Na configuração do SEI, as funções do protocolo são atribuídas ao indivíduo que criou o processo, podendo ser inclusive o próprio requerente. No caso do servidor da universidade, ele é o próprio gestor de seu processo, responsável por registrá-lo, autuá-lo, classificá-lo e acompanhar todo o seu fluxo documental.

A autonomia da abertura de processos a maioria dos membros da comunidade da UFF – todos que possuem o login ao sistema iduff e senha podem abrir processo no SEI – e a facilidade de acesso remoto aos trâmites do processo precisam ser acompanhados da padronização dos procedimentos dos processos de trabalho e do controle das entradas na autuação, além do monitoramento dos fluxos. Caso contrário, há o risco de duplicidade de processos, perda da organicidade entre atividades e documentos, além da quebra do vínculo arquivístico entre os documentos. O lixo informacional que pode ser gerado pode atrapalhar inclusive a própria execução das atividades e funções da Universidade.

A possibilidade de gestão dos processos de trabalho por múltiplos membros da comunidade exige que as rotinas administrativas e os acordos de como serão desenvolvidos os trâmites dos processos sejam de ampla adesão. É necessária uma nova cultura administrativa e informacional baseada na descentralização, na mediação dos sistemas informatizados e na interoperabilidade.

Destarte, as políticas institucionais devem contemplar também políticas informacionais, que por sua vez precisam também abordar políticas de indexação. A publicação de normativas que consolidem as rotinas e acordos e o esforço por sua implementação devem fazer parte das

iniciativas de tais políticas. A customização do SEI deve ser baseada em tais políticas no tocante ao controle dos fluxos dos processos e à padronização das entradas na autuação dos mesmos.

O mapeamento dos fluxos não garante a correta classificação dos documentos e das atividades. O código de classificação das atividades-meio para a administração pública federal, elaborado pelo Arquivo Nacional e aprovado pelo CONARQ, por possibilitar o agrupamento em assuntos, permite que processos de diferentes atividades e unidades de trabalho sejam classificados com a mesma notação a exemplo dos processos de reembolso de bilhetes de passagens e os processos de adicional de qualificação que recebem o mesmo código 024.92, relativo a “AUXÍLIOS”. A indexação não resolveria os problemas de gestão administrativa e de documentos da Universidade Federal Fluminense. No entanto, pode facilitar a recuperação dos documentos e a implementação do SEI na UFF, com a especificação desses processos de trabalho.

O SEI permite ao perfil administrador efetuar cadastros de usuários e unidades, configurações de funcionalidades, hierarquias no fluxo documental e permissões de gestão de processos e acesso, além de parametrizações. O Sistema de Permissões (SIP), associado ao SEI, é responsável pela habilitação dos usuários, unidades de trabalho, hierarquias e permissões²³.

A Portaria Interministerial MP/MJ nº 2321, de 30 de dezembro de 2014, obriga que cada processo ou documento avulso registrado no SEI possua o Número Único de Protocolo (NUP), este deve ser composto por 21 dígitos nos quais 7 primeiros identificam a unidade administrativa a partir da SIORG, que atua como unidade protocolizadora.

O SEI não é personalizado, portanto, todos os usuários de um mesmo perfil têm as mesmas prerrogativas, inclusive assinar documentos. No SEI, o qualificador do servidor que assina um documento não está vinculado ao usuário, mas à unidade à qual o usuário está cadastrado. Além dos problemas relacionados ao SIORG apontados anteriormente, há a possibilidade de alteração e assinatura dos processos de trabalho por usuários que pertençam à mesma unidade de trabalho, mas que não são partes interessadas no processo. Os qualificadores, como cargo, função, descrição das atividades de competência de cada usuário, precisam ser atribuídos para que o controle do fluxo documental seja feito e não haja perda de organicidade entre os documentos e as atividades.

A Universidade Federal Fluminense por não ter definidas suas rotinas administrativas em normativas esbarra em entraves que vão para além da customização do sistema. Os processos são vistos e passíveis de serem modificados, por todos os usuários cadastrados na unidade de trabalho ainda que não tenham qualquer relação com os mesmos. Isso ocorre por não terem sido os qualificadores atribuídos às ações específicas, que por sua vez não estão registradas e padronizadas no ordenamento da Universidade. O departamento de Ciência da Informação (GCI) pode servir

²³ As telas do SEI referentes aos perfis, cadastros, criação de documentos, autuação de processos e suas classificações são apresentadas nos anexos de L e T.

como exemplo, uma vez que todos os docentes e funcionários, assim como o chefe de departamento visualizam todos os processos abertos. Não há necessidade de ser requerente, parte interessada, responsável por alguma tarefa ou necessitar deliberar sobre alguma matéria²⁴.

Outra questão enfrentada pelas universidades federais têm sido a criação de hierarquias da entidade produtora no SEI. O sistema não permite a criação de múltiplas hierarquias, isto é, apenas uma hierarquia ativa (raiz) é permitida para toda a estrutura organizacional. Entretanto, nos seus organogramas, há órgãos que não são considerados como unidades de trabalho, são deliberativos, consultivos, fiscalizadores que funcionam de forma autônoma. Algumas ações dos processos de trabalho são executadas de sincronicamente por tais órgãos que não possuem hierarquia entre si. Embora soluções paliativas tenham sido apresentadas pelos desenvolvedores do SEI como criar ramificações separadas na interface de cadastro da hierarquia (TRF4).

O SEI é um software de gestão de processos de trabalho que necessita ser customizado a partir de requerimentos previstos em um SIGAD, preferencialmente elaborado pela instituição produtora a partir de suas políticas de informação. Dentre as funções previstas no SEI estão: a assinatura dos documentos, inclusive em bloco, essa permissão de assinatura é atribuída a partir da função que o servidor exerce em cada unidade de trabalho.

Os processos ou documentos avulsos são classificados no campo “assunto”. As universidades federais, de acordo com a Portaria 1.261, de 23 de dezembro de 2013, são obrigadas a classificar seus documentos das atividades-fim pelo **Código de classificação das atividades-fim das instituições federais de ensino superior**, aprovado pela Portaria 92, de 23 de setembro de 2011, do Arquivo Nacional. Ademais, a Universidade Federal Fluminense adotou a Resolução 14, de 24 de outubro de 2001, do CONARQ, relativa ao código de classificação de documentos de arquivo da Administração Pública: atividades-meio.

No SEI, como cada usuário é gestor do seu processo, aquele que o produz o classifica. No entanto, o sistema permite a exclusão e inclusão de classes, funcionando tal registro como um protocolo aberto o que pode prejudicar a organicidade e dificultar a recuperação dos documentos.

Ademais, é possível criar modelos de documentos com seções que estruturam os formulários. A UFF a partir do mapeamento de seus processos criou alguns modelos como o requerimento para abertura de processos administrativos. A universidade tem procurado adotar a nomeação de cada modelo pelo tipo documental que ele representa, ou seja, a espécie associada à função a qual pertence.

O SEI ainda prevê a restrição de acesso aos documentos de acordo com as hipóteses legais, inclusive com a base legal já incluída como os artigos da lei 12.527 e da Constituição. No entanto, é possível inserir novas normativas que versem sobre a restrição dos documentos.

²⁴ Exemplo ilustrado no anexo N.

A trilha de auditoria também é uma das funcionalidades previstas pelo SEI, chamada de “histórico” na qual todas as ações realizadas naquele processo são registradas. Essa é uma função importante, pois permite monitorar todas as operações realizadas no fluxo documental. No entanto, requer uma melhor descrição das ações realizadas no tocante as funções e operações arquivísticas como registro, classificação, indexação, arquivamento, armazenamento, recuperação da informação, captura, acesso e uso, destinação e preservação.

O perfil administrador no SEI permite também o controle dos tipos de documentos. É possível criar e controlar a relação de documentos que aparecerão para os demais usuários no formato de combos com barra de rolagem. Na UFF, os processos e documentos avulsos foram incorporados levando em consideração o mapeamento dos processos e a classificação fornecida pelo CONARQ das atividades-meio e o código para as IFES do Arquivo Nacional. Essa medida facilita a classificação pelo usuário do tipo de documento, pois já vem vinculada na sua escolha, embora o campo da classificação seja editável para qualquer perfil.

Todos os documentos do processo são organizados verticalmente, pela ordem de inclusão e o modo de visualização é chamado de árvore. Os processos que possuem muitos documentos são organizados em pastas por conta do tamanho²⁵.

Apesar da lei nº 12.682, de 9 de julho de 2012, que dispõe sobre o armazenamento de documentos eletrônicos, prever a indexação para facilitar a recuperação dos documentos e das informações contida neles, o SEI não obriga, quando da criação dos documentos no sistema, a realização da operação. A falta de indexação, tanto para definir descritores no campo “especificação”, quanto para nomear os documentos juntados no processo, dificulta a recuperação da informação. Ademais, a falta de padronização na nomeação dos documentos, impede inclusive a visualização em árvore das ações realizadas no processo.

Art. 4º As empresas privadas ou os órgãos da Administração Pública direta ou indireta que utilizarem procedimentos de armazenamento de documentos em meio eletrônico, óptico ou equivalente deverão adotar sistema de indexação que possibilite a sua precisa localização, permitindo a posterior conferência da regularidade das etapas do processo adotado (BRASIL, 2012).

O SEI é baseado no Solar BPM (*business process management*), uma plataforma de gestão de processos que automatiza rotinas de trabalho. Voltado para a administração pública, esse produto, desenvolvido pela Softplan, prevê a indexação automática dos documentos tanto natos digitais, quanto dos digitalizados, que tenham passado pelo Reconhecimento Ótico de Caracteres (OCR). A busca é realizada por palavra e na modalidade “pesquisa livre”, todos os documentos que contiverem a palavra selecionada serão recuperados. A modalidade “pesquisa estruturada” fornece

²⁵ A árvore de processos do SEI encontra-se no anexo O.

diversos campos que possibilitam um grau de especificidade maior. No entanto, os descritores utilizados são os mesmos de autuação do processo que necessitam de padronizações no tocante a suas nomeações pelas entidades produtoras.

Na próxima subseção desta tese refletimos sobre as possibilidades de indexação dos documentos de arquivo nas universidades federais brasileiras, considerando a implementação de sistemas eletrônicos de informação. É importante ressaltar que estamos direcionando nossos olhares para a fase de produção documental, sobretudo no momento de registro dos documentos e autuação dos processos, especialmente no protocolo. O nosso campo empírico no qual tais reflexões foram suscitadas é a Universidade Federal Fluminense no processo de implementação do Sistema Protocolo Integrado, via Sistema Eletrônico de Informações.

7.4 A INDEXAÇÃO NOS ARQUIVOS DE UNIVERSIDADE – POSSIBILIDADES DE PRODUÇÃO DE SENTIDO A PARTIR DOS USOS DOS DOCUMENTOS

Ao longo deste trabalho procuramos abordar os documentos como mediadores de relações sociais desde seu surgimento como forma de registro das atividades humanas. Ainda que reconheçamos as limitações desta generalização e reconheçamos também os múltiplos usos sociais dos documentos, tal afirmativa baseia-se sobretudo nos documentos de arquivo, objeto de nossa análise no tocante à indexação.

A interação entre homem-máquina, realidade relatada nessa tese desde a Segunda Guerra Mundial, torna-se um imperativo nas sociedades contemporâneas, consolidado no século XXI. Torna-se necessário analisar o impacto das tecnologias de informação a qualquer estudo sobre sistemas de informação e os processos comunicacionais. Em nosso escopo, ao abordar a fase de produção documental nos arquivos de universidade e o processo de indexação, refletir sobre os documentos nato-digitais, a quantidade de informação produzida em ambiente de rede, o lixo informacional e as formas em que as informações são organizadas com o objetivo de serem recuperadas tornaram-se ações importantes no processo de investigação.

Como afirma Bruno Delmas (1988), o gigantesco volume de documentos e de informação produzidos é um fenômeno contemporâneo. Questão diagnosticada ao final da década de 1980 e que tem aumentado em progressão geométrica. A facilidade de criarem-se documentos e circularem informações em ambiente eletrônico faz surgir uma imensa quantidade de lixo informacional. As pessoas produzem os documentos digitais para viabilizarem suas ações. Não há uma preocupação primeira com a preservação e recuperação destes.

Ademais, o uso dos algoritmos pelos sistemas eletrônicos de recuperação da informação tem sido cada vez mais frequente. A forma em que os padrões são estipulados e as variáveis combinadas começou a ser uma preocupação na última década. O *AI Now Research Institute* publicou em abril

de 2018 um relatório detalhado sobre a avaliação do impacto dos algoritmos nos serviços públicos dos Estados Unidos. O documento aponta inúmeros casos de violações legais, discriminação, desrespeito a direitos e problemas como a reafirmação de estereótipos nas chamadas “decisões automáticas do sistema”.

Kate Crawford, uma das pesquisadoras que assinam o relatório, afirma que dados pessoais, prontuário médico, histórico de compras pela internet têm sido combinados em programas utilizados para conceder ou não empréstimos, permitir ou não a contratação de planos de saúde, estipular preços de apólice de seguro, escolher quem participará da etapa de entrevistas para seleção de empregos e traçar perfis inclusive de futuros suspeitos de crimes que ainda não aconteceram. De acordo com Crawford (2018), as formas em que essas variáveis têm sido combinadas não são claras e abertamente discutidas com a sociedade na qual esses programas estão operando e interferindo, inclusive nas vidas e nas relações das pessoas. “Muitos sistemas baseiam a análise da programação dos algoritmos apenas na avaliação das entradas (inputs), saídas (outputs) e na simples informação de quais algoritmos estão sendo mais usados, sem considerar a falta de acesso aos códigos-fonte camuflados em tais programas” (AI NOW INSTITUTE, 2018, p. 19)

As formas em que criamos os padrões e as infinitas combinações possíveis das variáveis expressas pelos algoritmos estão relacionadas com a cultura e as construções sociais de cada comunidade linguística. De acordo com Crawford (2018), os estereótipos, os preconceitos, as hierarquias sociais estão sendo replicados pelos algoritmos. Portanto, é importante que os governos e a Administração Pública utilizem sistemas abertos, que permitam a constante avaliação externa por suas sociedades, para que suas comunidades participem ativamente de suas remodelagens, identificando inclusive os impactos de tais escolhas nos serviços prestados.

As instituições devem identificar potenciais impactos no público e, em seguida, engajar proativamente as comunidades afetadas para garantir que o sistema atenda aos objetivos de sua comunidade. A avaliação deve articular o porquê, à luz dessas metas, o sistema terá um impacto líquido positivo nessas comunidades (AI NOW INSTITUTE, 2018, p. 18).

Então, a avaliação dos sistemas precisa ser constante. Os requerimentos que irão pautar a programação dos algoritmos dos Sistemas de Recuperação da Informação precisam ser debatidos pela comunidade em que estão inseridos. Um dos principais desafios do processo de avaliação do impacto dos algoritmos é justamente pelo grau de previsibilidade, suposição e antecipação que as instituições responsáveis pelos SRIs precisam incluir em sua análise, sobretudo no tocante aos fatores culturais e sociais cujas dinâmicas variam dentro de suas comunidades. As padronizações dos SRIs são baseadas em prognósticos e probabilidades de êxito. A certeza da eficácia é possível apenas como utopia a ser perseguida.

Para Blair (2003), um sistema de informação é, basicamente, um processo organizado que tenta fornecer informações corretas ou pertinentes em resposta a um pedido formal de algum tipo. O usuário (*inquirer*) geralmente começa o seu processo de busca com algum tipo de pedido que pode ser formulado em linguagem natural. Se pensarmos no pedido como uma pergunta, podemos distinguir um número de tipos básicos: o que, quem, quando, onde, como e por quê. Esse pedido deve então ser traduzido em um formato apropriado para que o sistema de informação possa usá-lo, ou seja, uma consulta formal (*formal query*), uma formulação a partir de uma linguagem de indexação (BLAIR, 2003, p. 285).

Para que a consulta selecione as informações ou dados certos, deve haver alguns dados ou informações para que tal consulta possa ser estruturada e assim facilitar a recuperação da informação ou documento desejados. A normalização é um exemplo de procedimento relativamente difundido e bem compreendido de estruturação de dados para recuperação da informação.

Compartilhamos da visão de Blair sobre os problemas causados por algumas concepções naturalizadas na arquitetura dos sistemas de recuperação da informação. Uma das suposições mais cristalizadas no campo da Ciência da Informação é que um modelo de dados bem estruturado é eficaz em qualquer sistema de recuperação da informação, independente dos tipos de documento e informação gerenciados pelo sistema.

Esse tipo de concepção costuma atribuir a culpa ao usuário e a sua dificuldade de operar o sistema quando os resultados de pesquisa são insatisfatórios. Essa tentativa de aplicar um modelo de dados universal, ou seja, aplicável a qualquer sistema, é bem semelhante à tentativa da semântica formal em estabelecer um modelo lógico e perfeito de linguagem que abordamos na terceira seção deste trabalho, os mesmos problemas no tocante ao sentido da linguagem são encontrados nesses modelos de dados.

Outra concepção que está presente nos sistemas automáticos de recuperação da informação é de que a representação do conteúdo da informação é extraída em sua integralidade, sendo exata e inequívoca. É naturalizada a operação da tradução como um transporte no qual a consulta é transformada para a linguagem utilizada pelo sistema e seu resultado igualmente traduzido em linguagem inteligível ao usuário, sem que haja intervenções ou perda do conteúdo informacional. Novamente, deparamo-nos com a teoria clássica da tradução dos estudos de linguagem que foi sintetizada pelos princípios de um de seus teóricos pioneiros, Alexandre Fraser Tytler:

- 1) a tradução deve reproduzir em sua totalidade a ideia do texto original;
- 2) o estilo de tradução deve respeitar o original;
- 3) a tradução deve ter toda a fluência e naturalidade do texto original (TYTLER apud BASSNETT-MCGUIRE, 1991, p. 63)

Essa concepção de tradução como transporte de significados trata o documento e seu conteúdo como objetos estáveis e, portanto, transportáveis. Seus contornos e sentidos já estão dados e claros, dessa forma, passíveis de serem representados completa e objetivamente. A função da tradução é vista como meramente mecânica. O aprimoramento encontra-se nas formas de garantir que a “carga informacional” do conteúdo traduzido chegue ao seu destino de forma intacta.

Existe uma crença de que o acesso intelectual ao conteúdo do documento é facilitado em grande parte pela melhoria do acesso físico, ou seja, a adoção de tecnologia, que melhore a velocidade do processamento pelas máquinas, garanta maior rapidez e eficiência na recuperação da informação. É inegável que as inovações tecnológicas facilitam a recuperação, no entanto, o determinismo tecnológico não permite que soluções como a avaliação dos sistemas de recuperação da informação a partir dos usos sociais que a sua comunidade tem dos documentos possam ser incorporadas no aprimoramento desse acesso.

Outro pressuposto comum na Ciência da Informação relativo aos SRIs e apontado por Blair (2006) é que o tamanho do sistema não interfere na recuperação da informação. Um grande sistema pode ter a mesma efetividade na recuperação da informação, pois o modelo baseado em padrões universais e suas infinitas combinações é compatível a qualquer base de dados. No entanto, a indeterminação se multiplica em sistemas maiores o que acaba por dificultar a realização de pesquisas específicas ou exaustivas, tornando os resultados não tão bem-sucedidos.

A indexação como processo de análise e representação da informação, segundo Blair (2006), possui dois objetivos fundamentais: a descrição e a discriminação. A descrição trata do chamado “conteúdo intelectual” de um *corpus* particular da informação. Já a discriminação trata da distinção entre a informação mais próxima à busca das informações que possuem um conteúdo intelectual similar.

O julgamento mais comum sobre a qualidade da indexação refere-se a sua capacidade de descrever o documento e suas informações da maneira em que os usuários irão procurar por aquela informação descrita. Entretanto, a indexação também em sua descrição deve discriminar as informações que os usuários consideram utilizáveis daquelas que eles acreditam não sê-las (BLAIR, 2006, p. 310). Os documentos recuperados costumam ser julgados por serem úteis ou não, e alguns documentos podem ser mais úteis do que outros.

O sucesso na recuperação de um documento é medido, geralmente, pelos graus de efetividade e somente sob circunstâncias especiais o documento pode ser considerado correto ou não. “O padrão para avaliar a eficácia da recuperação dos documentos deve ser a utilidade desses e não se eles são corretos ou não” (BLAIR, 2006, p. 314, tradução nossa).

A utilidade é um critério de desempenho muito mais complexo e problemático para medir a relevância de uma informação ou documento do que a exatidão e, normalmente, não há padrões de

utilidade objetivos e independentes do usuário. Cabe destacar que a utilidade não é uma propriedade da informação em si. Segundo Blair (2006), a utilidade de um documento depende do seu histórico de uso, das habilidades daqueles que utilizaram e utilizam suas informações e dos usos a que as informações do documento são disponibilizadas.

A proposta de Blair (2006) para a indexação dos documentos da qual compartilhamos é pensar não apenas na eficácia de sua recuperação, mas no potencial de uso dos documentos e de suas informações. A eficácia está relacionada a exatidão de dados na recuperação do documento. A exatidão é uma propriedade independente, trata-se de cálculos matemáticos com fórmulas padronizadas pautados na filosofia de que o sentido dos documentos está em seu conteúdo.

A partir da última década, grande parte dos SRIs tem desenvolvido formas de recuperação baseadas em algoritmos que combinam cálculos matemáticos voltados à extração do conteúdo dos documentos, na lógica da concepção orientada pelo documento (MAI, 2005), com padrões de mineração de dados que estipulam o comportamento de uso, sob influência do experencialismo. Assim, são traçados perfis de uso, baseados em determinismos biológicos e culturais, que por vezes acabam por reproduzir esteriótipos que nem sempre são úteis ao usuário. Já o potencial de uso, além de considerar o histórico de uso de tal documento, procura avaliar a habilidade de interação daqueles que usam o sistema com as linguagens adotadas pelo mesmo e as intenções nas buscas realizadas.

Trabalhar a indexação a partir da concepção de potencial de uso dos documentos é assumir a concepção pragmática de que o sentido dos documentos e das informações está em seu uso (WITTGENSTEIN, 1999). Então, assim como alertou Fish (1992), não há significados pré-determinados que combinados entre si formam os textos e, conseqüentemente, os documentos. O sentido se dá na interação de interlocutores que agem e se relacionam em um mesmo sistema interpretativo, uma comunidade discursiva (SWALES, 1990).

Circunstâncias, crenças, valores, suposições estão imersas nessa realidade da produção e uso dos documentos. Assim como alerta Wittgenstein (1999), não se trata de mapear o entendimento dos indivíduos das palavras e conceitos em determinado léxico e as regras de combinação dos mesmos. São maneiras de vida, formas de pensar e compartilhar filosofias, ideias, técnicas etc., Portanto, um mundo de objetos, metas, aspirações, intenções, procedimentos, validações compartilhados, ainda que haja embates e antagonismos, mas os usos são compreendidos porque fazem parte da dinâmica da comunidade interpretativa em questão.

A gramática não se trata de regras universais extraídas do conteúdo dos textos, mas são normas da língua que se tratam de acordos estabelecidos na comunidade, constitutivos da própria. O que delimita o sentido são as práticas, as regras institucionalizadas, estabelecidas, organizadas,

validadas e seguidas por seus membros, tacitamente ou não, cujos propósitos e metas são previamente assumidos e conhecidos (FISH, 1992, p. 194).

Os procedimentos de representação da informação tais como a indexação podem reduzir o número de decisões que um usuário deve tomar para realizar a sua busca, mas não as elimina. A redução do número de decisões lógicas pode ser efetuada melhorando a representação dos documentos, melhorando as inferências que um sistema precisa fazer para corresponder solicitações formais a documentos úteis, ou, mediando tais escolhas a partir do mapeamento de usos semelhantes ou correspondentes entre novos usuários do sistema e usuários especializados, já acostumados em operá-lo.

Um “prontuário do aluno” pode ser chamado de “dossiê do aluno”, ou “pasta do aluno”, ou “histórico de atividades do aluno”, no entanto, ainda assim haverá um número relativamente limitado de possibilidade de representar tal documento pelas convenções utilizadas em nossa comunidade, ou seja, o meio universitário brasileiro. Por outro lado, há documentos com assuntos de interesse para a mesma comunidade cujas representações são, praticamente, ilimitadas, como “atividades de pesquisa”. Ao pesquisarmos este termo no AtoM da UnB²⁶, encontramos outros 53 termos, dentre eles: “Ensino e pesquisa”; “Pesquisa Científica”, “Pesquisa”. “Projeto de pesquisa”, “Programa de pesquisa”, “Laboratório de ensino e pesquisa”, “Programa de atividades”, entre outros²⁷.

A consequência desse fenômeno torna cada vez mais difícil para o usuário adivinhar quais as representações correspondem aos documentos que ele quer ou considera útil a sua pesquisa. Um experimento realizado por Swanson (1977) mostrou que quando os indexadores são encorajados a atribuir a quantidade de termos que for possível para os assuntos representados de documentos não muito extensos, eles foram capazes de encontrar descrições de assuntos pertinentes que, praticamente, não ultrapassaram o limite já previsto pelo sistema.

Outro trabalho importante nessa área, de Langendoen e Postal (1984), procurou desconstruir a naturalização da premissa do modelo da máquina de Turing e da Gramática Gerativista de que é possível a representação, criação e reprodução de todas as frases bem formadas em uma linguagem a partir da combinação de padrões estabelecidos tal como os algoritmos. O estudo mostra que a magnitude e complexidade das coleções de frases que compõem uma linguagem natural não podem ser representadas por qualquer código, notação, algoritmo, número finito ou transfinito.

²⁶Sistema utilizado pela Universidade de Brasília para descrever seus documentos de arquivo, cujo principal objetivo, de acordo com o site da própria universidade é : promover a difusão dos documentos da Universidade e fornecer elementos de descrição em vários níveis que permitam aos usuários conhecer as diversas modalidades de acesso ao nosso acervo. Ver: <<https://atom.unb.br/index.php/>>. Acesso em: 17 dez. 2018.

²⁷ A tela com os resultados da pesquisa realizada no Atom-UnB encontra-se no anexo V.

Blair (2006) relata que a consequência do uso desses modelos como os da máquina de Turing e da Gramática Gerativista nos SRIs tem sido a frustração permanente no trabalho com as “redes semânticas”. “A menos que a rede seja extremamente pequena, o número de palavras e frases que podem ser ligadas a um conjunto semântico de ‘significados’ pode tornar-se rapidamente incontrolável” (BLAIR, 2006, p. 317). Como a recuperação determinada de dados depende de uma correspondência bastante direta entre as buscas e as representações, se a busca não conseguir recuperar o que se deseja, é bem provável que os dados desejados não existam no sistema de informação.

Por outro lado, na recuperação de documentos, se a busca feita não localizar os documentos desejados, isso não significa que não haja documentos úteis ao usuário no sistema. Na maioria das vezes, trata-se de uma incompatibilidade entre a forma de consulta realizada pelo usuário e a representação dos documentos úteis. Situação na qual muitas vezes culpa-se o usuário por não ter formulado “corretamente” a pergunta ao sistema.

Uma das características da linguagem natural na perspectiva do gerativismo é a sua “produtividade”. A linguagem é produtiva no sentido de que um vocabulário relativamente pequeno e algumas regras de sintaxe e semântica podem ser combinadas para produzir um número incontável de sentenças ou descrições válidas. A produtividade da linguagem provou ser tanto uma ajuda quanto um obstáculo para os sistemas de recuperação de informação que trabalham a partir da lógica do mentalismo, como tem sido a maioria dos sistemas computadorizados que operam com redes de componentes semânticos baseadas na gramática gerativista. Nesses sistemas, a produtividade é uma ajuda porque é relativamente simples encontrar uma maneira razoável de representar o conteúdo intelectual de qualquer documento; mas essa produtividade também pode ser um obstáculo, porque a linguagem é tão criativa que o número de representações “razoáveis” para o conteúdo intelectual de um documento é virtualmente ilimitado em sistemas que não delimitem os contextos e usos dos documentos e das informações.

Se acreditarmos que cada item de informação tem, potencialmente, uma representação altamente determinada que permitiria ser recuperada por aqueles que o querem em qualquer situação, então muitas das falhas dos sistemas de informação serão atribuídas à indexação incorreta – não representando o conteúdo da informação corretamente ou formulação de consulta de pesquisa incorreta — não representando o que o usuário quer corretamente. Mas, é a tese aqui, que em muitas situações de recuperação da informação, o sistema não tem, real ou potencialmente, representações suficientes para distingui-las daquelas informações que o usuário não quer (BLAIR, 2006, p. 289, tradução nossa).

Então, a indexação será melhor sucedida ou não dependendo de como o documento ou informação requerida é usada. As formas em que os documentos são usados e as escolhas das regras ou convenções a serem seguidas indicarão o nível de determinação dos SRIs. A arquitetura dos sistemas, principalmente no tocante aos processos de representação dos documentos como a

classificação e a indexação, deve ser baseada nos usos dos documentos e nas convenções seguidas por sua comunidade discursiva. Sob essa ótica, os documentos de arquivo possuem grande vantagem, pois são altamente regulados, tanto em sua produção quanto em seus usos, por normas, inclusive do ordenamento jurídico, e rotinas administrativas.

Fish (1992) afirma que os sentidos costumam ser facilmente identificados e delimitados, a infinita pluralidade de significados só é temida caso se retire a palavra, a frase, o texto, o documento do seu contexto de uso. Se a situação é contextualizada, as convenções conseguem ser mapeadas e compreendidas, pois já estão institucionalizadas em sua comunidade discursiva.

Seria de fato perturbador se a norma flutuasse livre e indeterminada. Mas não existe semelhante estado; os enunciados emergem somente em determinadas situações e, dentro dessas situações, o significado normativo de um enunciado será sempre óbvio ou pelo menos acessível, embora numa outra situação esse mesmo enunciado, não sendo mais o mesmo, terá outro significado normativo que será não menos óbvio e acessível (FISH, 1992, p. 195)

Portanto, é importante que o processo de indexação discrimine os documentos, ou seja, considere os sentidos que eles adquirem em diferentes situações para representar seus assuntos. O sentido estará associado ao potencial de uso do documento em situações probabilisticamente presumidas. A indexação é um prognóstico realizado a partir de circunstâncias analisadas no presente, cuja eficácia é maior quando o seu processo de análise dos documentos para posterior representação baseia-se no mapeamento dos usos destes e nos contextos em que os mesmos estão inseridos dentro de sua comunidade discursiva. Certos contextos são mais frequentes do que outros, do mesmo modo que os usos dos documentos.

Ainda que a lei de acesso a informação, nº 12.527, tenha ampliado os usos dos documentos de arquivo, na fase de produção documental, quando da vigência legal desses, os usos dos documentos estão associados, em sua maioria, na realização das atividades de sua entidade produtora. Podemos associar ao que Fish (1992) chamou de “ambiência institucional”. Ainda que não seja normalizada a indexação em uma política de indexação ou em um manual da instituição, o indexador irá representar o documento a partir do contexto que ele é mais frequentemente encontrado. Isso acontece de forma intuitiva e automática. “É impossível pensar-se num enunciado independentemente de um contexto” (FISH, 1992, p. 197).

É importante considerar o processo de interpretação na indexação. A análise do documento envolve um processo duplo e simultâneo que é o de reconhecimento do contexto e produção de sentido. A representação envolve o julgamento e a escolha de conceitos e termos que costumam ser aplicados a contextos e usos mais frequentes. A indexação é um processo probabilístico, no entanto, marcado por caminhos familiares e previsíveis. Nesse estudo sobre os arquivos de universidades na

fase de produção documental, os documentos circulam, mediam relações, são usados em contextos e situações conhecidas por sua comunidade discursiva.

Indexar é fornecer caminhos para a recuperação do documento e da informação. No entanto, esse caminho não é completamente determinado, mas também não é completamente desconhecido. Há traçados delineados, ainda que não sejam fixos. Nesse processo de interpretação que faz parte do processo de indexação, a introdução de novas categorias de compreensão ou expansão das antigas sempre precisa se articular com as convenções já adotadas pela comunidade discursiva em que estão se sujeitando. “A mudança de uma estrutura de compreensão para outra não constitui uma ruptura, mas uma modificação dos interesses e preocupações já existentes; e porque já existem, eles forçam a direção da sua própria modificação” (FISH, 1992, p. 202).

Por esse motivo a indexação deve realizar a descrição e a discriminação. A descrição do conteúdo intelectual do documento é importante, baseada nas formas em que as expressões são usadas pela comunidade discursiva a qual o sistema faz parte. Essa operação visa representar o conteúdo intelectual do documento para facilitar a recuperação por parte dos interessados. Os termos de índice, portanto, não devem apenas descrever o conteúdo intelectual do documento, mas também discriminar o conteúdo intelectual de um determinado item de informação do conteúdo intelectual de outras informações no mesmo sistema que é semelhante a ele.

Descrição, então, é um processo que se preocupa apenas com um único item de informação, enquanto que a discriminação é um processo que deve considerar como um dado item de informação, como um documento ou uma imagem, deve ser distinguido de outros itens de informação no mesmo sistema. Em suma, a descrição lida com itens únicos de informação, enquanto a discriminação examina como os itens de informação individuais se relacionam com todo o sistema de informação (BLAIR, 2006, p. 298).

Normalmente, as pessoas buscam por documentos ou por informações não como um fim em si, mas como algo que irá viabilizar uma atividade ou prática. Assim, é a atividade ou prática que a busca atende e que definirá os critérios para uma busca bem-sucedida. O sentido e a determinação das descrições dos documentos não são propriedades das próprias descrições, mas dependem de como são usadas. Isso mostra que a determinação da representação é um aspecto profundo e complexo da linguagem que não pode ser esclarecido sem referência ao contexto e às circunstâncias de seu uso.

A indeterminação não pode ser “extirpada” da linguagem - não há uma clareza lógica subjacente de significado que só precisamos descobrir, como Frege e Russell pensavam. Se a indeterminação é algo com o qual teremos que conviver na recuperação de conteúdo, então precisamos entender quais são as consequências disso para o design e uso do sistema de informação (BLAIR, 2006, p. 339, tradução nossa).

Ao admitirmos que a indeterminação é um problema que não podemos eliminar no processo de recuperação da informação, precisamos pensar em maneiras para mitigá-la na representação dos documentos e da informação. Os sistemas de informação são, fundamentalmente, baseados em processos linguísticos. O usuário precisa formular uma pergunta acerca do documento ou informação que ele deseja, essa descrição precisa combinar com as representações disponíveis (essas representações disponíveis podem vir de várias fontes: indexação manual, indexação assistida por computador, indexação automática ou recuperação de texto completo). São todos processos linguísticos e, como tais, os sentidos das informações e dos documentos estão nas dinâmicas das interações entre esses objetos e os sujeitos que fazem parte dessa comunidade discursiva. Como afirma Wittgenstein (1999), os usos nos ensinam o sentido das palavras e, por conseguinte, inferimos que das informações e dos documentos também.

Portanto, é preciso aproximar a teoria da indexação às filosofias da linguagem para melhor compreender seus axiomas, princípios e critérios que têm orientado tal prática. Na terceira seção deste trabalho abordamos o realismo, o mentalismo e o pragmatismo como principais correntes que influenciaram a teoria da indexação. Apresentamos também a influência dessas filosofias nos estudos de linguagem e o impacto dos mesmos nos modelos de indexação. Nesta seção, além dos problemas identificados pela adoção naturalizada do formalismo e mentalismo no processo de indexação, propomos uma abordagem do processo a partir do pragmatismo, considerando a indexação como uma operação de análise e representação linguística dos documentos e de suas informações.

De acordo com Adersen e Christensen (1999), a literatura da Ciência da Informação, em seu histórico, trata a linguagem como algo secundário para a indexação, entretanto, trata-se da análise de enunciados que compõem o documento e sua representação, portanto há uma dimensão linguística importante. Para Frohmann (1992), a representação de um documento remete ao documento no presente e à realidade ou contexto social que ele pode representar ou refletir. Trata-se então de um processo de comunicação, ou seja, um processo de produção e intercâmbio de sentidos.

Não há processo de comunicação sem a linguagem (MARCONDES, 1992, p. 3). A informação não é transmitida diretamente, ela é comunicada por meio dos usos feitos das representações dos documentos. Essas representações dos documentos são constituídas pela linguagem. A comunicação da informação acontece na linguagem. Não há comunicação sem a produção de sentido. Portanto, é importante para a indexação compreender como os documentos e os sujeitos interagem entre si para produzir sentido.

A indexação é análise e representação de sentidos dos documentos. A teoria da indexação é basicamente e necessariamente uma questão de como as palavras, conceitos, expressões são usados em um contexto linguístico (jogos de linguagem) em que elas fazem parte. “Nós não começamos de

determinadas palavras, mas de certas ocasiões ou atividades” (WITTGENSTEIN, 1999). Então, o processo de indexação, como processo de comunicação, opera com produção de sentido e interpretação. São categorias fundamentais.

A tarefa central das pesquisas sobre recuperação da informação é entender como os documentos devem ser representados para uma recuperação efetiva. Isso é preliminarmente um problema de linguagem e sentido. Qualquer teoria de representação do documento e, por consequência, qualquer teoria da recuperação da informação deve ser baseada em uma clara teoria da linguagem e do sentido (BLAIR, 2003, p. vii-viii).

Nesse exercício de aproximarmos a teoria da indexação das filosofias da linguagem, a proposta é abandonarmos a concepção instrumentalista da linguagem em direção a sua dimensão social e performática. Como demonstra Wittgenstein (1999), falar uma língua é uma ação social. É um entendimento coletivo. O uso da linguagem é um fenômeno social baseado em regras e convenções acordadas, compreendidas, aceitas e seguidas, os chamados “jogos de linguagem”. Fazer parte do jogo é trabalhar os usos de tais regras que podem ser bem-sucedidos ou não, no sentido de que tais regras seguidas possibilitarem a execução das ações e atividades para as quais elas foram escolhidas ou não. O significado não é apriorístico, ele se dá no uso das palavras. Por isso não é possível construir sistemas com conceitos, definições, representações atemporais e universais. Eles são dinâmicos e, por conseguinte, precisam de constante atualização.

As convenções, regras, acordos seguidos por uma comunidade discursiva pertencem a ela. Isso não quer dizer que estes são efêmeros, tampouco queremos limitar o alcance e a influência das comunidades discursivas. Há usos que são históricos e semelhantes em diversas comunidades. Nós não estamos aptos para dar diretrizes claras e certas para a aplicação correta da palavra. Uma palavra pode ser utilizada em muitos contextos. Existe uma relação estreita entre o uso da palavra, a compreensão e a adesão das regras. A aplicação da palavra continua sendo um critério de compreensão da regra. “Seguir uma regra é uma prática social” (WITTGENSTEIN, 1999, § 146).

Assumindo que qualquer documento é parte de uma prática linguística, um jogo de linguagem, e que um documento em e por si mesmo não pode ser a base da indexação, a indexação então deve ser baseada no contexto e nos jogos de linguagem. A indexação não é meramente a sumarização do conteúdo de um documento, isto é apenas uma parte da representação de um documento. Se a representação é para ser significativa, então ela precisa refletir o jogo de linguagem a qual o documento faz parte. Sabendo que apenas o documento não é suficiente para entender seu significado, nós precisamos conhecer a linguagem e isso implica em como ela é usada.

A representação do documento deve refletir conceitualmente o seu jogo de linguagem. A análise de assunto de um documento precisa resultar na indexação de termos que reflitam o seu contexto linguístico (ANDERSEN; CHRISTENSEN, 1999, p. 16). As palavras não podem ter

sentidos ambíguos no seu uso e nos jogos correspondentes. A possibilidade de escolha das palavras depende do jogo de cada linguagem e o sentido delas depende também do jogo de linguagem em que será usado. O significado de uma palavra não pode ser dado a priori, mas o sentido se dá no uso da palavra, portanto a posteriori.

A indexação pode delimitar ou explicitar formuladamente o uso de um conceito em determinado contexto. Deve mostrar as relações entre vários jogos de linguagem, as palavras e os conceitos que queremos compreender, de modo a disponibilizar os usos potenciais que podem ser adotados. Não se deve definir assuntos por comportamentos, pois não são comportamentos e sim formas de uso das palavras em determinadas situações. O que precisamos estudar são os usos e as situações das comunidades discursivas. Seguir uma regra, no tocante aos jogos de linguagem, tem mais a ver com um fenômeno social do que com uma propriedade mental.

A teoria da indexação é uma teoria do sentido. O sentido é um pre-requisito para a comunicação e o processo de indexação é um processo de comunicação. Entendendo que sentido é uso, acreditamos que a teoria da indexação é uma teoria social da comunicação do sentido.

Seguindo os passos de Guedes (2010) que aproximou os estudos de linguagem do processo de indexação, partilhamos da concepção de Swales (1990) acerca de comunidade discursiva na qual o autor considera ser um conjunto de indivíduos com objetivos em comum, formalmente expressos ou não. Possui também mecanismos de intercomunicação variáveis, usados primeiramente para fornecer informação e feedback a seus membros, e um léxico específico que restringe e adéqua os documentos compartilhados por seus membros e ainda dificulta seu acesso por não-membros.

Para que um grupo de indivíduos possa ser reconhecido como uma comunidade discursiva, seus participantes devem estabelecer e compartilhar os procedimentos e as práticas a serem utilizados entre eles. A sobrevivência dessa comunidade depende de um equilíbrio entre o número de membros mais novos e mais antigos, para que sua manutenção e conservação sejam viáveis. Esses participantes não compartilham necessariamente características específicas como emprego, faixa etária e/ou classe sócio-econômica, mas devem compartilhar seus objetivos enquanto comunidade.

Swales (1990) aproxima o conceito de comunidade discursiva da sociolinguística ao relacionar o conceito às redes sociais e aos contextos que tais sujeitos partilham, não o fixando no tempo e no espaço. A principal necessidade de tal grupo é a socialização, suas ações e interações em grupo, o que acaba por refletir no comportamento linguístico. Para o autor, a comunidade partilha de objetivos específicos que a mantém enquanto grupo.

Em 1998, Swales reformula o conceito de comunidade discursiva dizendo que ela é o espaço de circulação responsável pela (re)produção de um grande número de gêneros discursivos, os quais têm como função social a validação das atividades interacionais fora das comunidades. Esses

gêneros são produzidos por agentes sociais situados numa rede de relações num contexto histórico e social específico. O consumo dos gêneros, ou seja, a sua interpretação, não se dá de modo aleatório, na medida em que “os textos que eu leio são quase totalmente previsíveis de acordo com o lugar que ocupo nas estruturas sociais e institucionais” (KRESS, 1989).

As formas linguísticas, as quais somos expostos e de que fazemos uso no nosso grupo social, são familiares e acessíveis a nós. Para outros indivíduos, pertencentes a grupos sociais diferentes (outras classes sociais, outros grupos étnicos, grupos de outra faixa etária ou de outro sexo), essas formas linguísticas podem não ser familiares ou acessíveis. Em outras palavras, diferentes agentes sociais podem fazer usos de diferentes formas linguísticas e produzir ou interpretar diferentes gêneros discursivos. “Uma das características que os membros estabelecidos dessas comunidades discursivas possuem é a familiaridade com os gêneros específicos que são usados na busca comunicativa destes conjuntos de objetivos” (SWALES, 1990, p.9)

Os diferentes agentes sociais de uma determinada sociedade “são membros de uma certa instituição social, com suas práticas, seus valores, seus significados, suas demandas, proibições e permissões” (KRESS, 1989, p.132) e, portanto, acabam formando diferentes comunidades discursivas. Swales (1990) enumera seis características que definem uma comunidade discursiva:

- Têm um conjunto de objetivos públicos em comum, que são compartilhados por seus membros;
- Estabelecem mecanismos de comunicação próprios entre os seus participantes;
- Têm como principal função a troca de informações entre os seus membros, o que possibilita a sua participação nas atividades da comunidade;
- Desenvolvem seu próprio elenco de gêneros, que inclui a seleção dos tópicos e dos elementos formais do discurso que são apropriados a ele;
- Selecionam, e até criam, itens lexicais que têm um significado específico e relevante para os membros daquelas comunidades, mas que geralmente pouco ou nada significam fora delas, e
- Possuem membros antigos, que são os detentores do conhecimento do discurso e do conteúdo destas comunidades, e membros novatos, que são apresentados a esse conhecimento pelos mais experientes.

No entanto, é importante destacar que as comunidades discursivas, por serem constituídas de agentes sociais, não são formações rígidas ou imutáveis. Há tensões e conflitos que levam a transições e novas configurações, inclusive com o abandono ou incorporação de novas convenções, acordos e práticas. As comunidades discursivas possuem caminhos previsíveis, gêneros discursivos conhecidos e regras compreendidas por seus membros, mas também estão em constante transformação e transição pela dinâmica da vida e de seus membros.

As comunidades discursivas podem ter alcance local como, por exemplo, seus agentes sociais trabalharem juntos em uma mesma instituição, ou mesmo global, quando os membros

partilham os mesmos compromissos, focam nas mesmas ações, possuem discursos semelhantes (HEMAIS; BIASI-RODRIGUES, 2005). A comunidade científica da Ciência da Informação pode ser considerada um exemplo de comunidade discursiva, como demonstrou Guedes em sua tese, **Nominalizações deverbais em artigos científicos** (2010).

Os valores e significados específicos das comunidades discursivas são articulados linguisticamente de modo sistemático. Essas articulações linguísticas são chamadas de discursos. Kress define discursos como “grupos de afirmações sistematicamente organizados que dão expressão aos significados e valores de uma instituição” (1989, p. 160). Os discursos não existem isoladamente, mas sim dentro de um sistema maior de discursos, com os quais estabelecem relações dinâmicas que garantem um contínuo movimento de expansão e retração causado pelo seu objetivo maior: a expansão e a colonização de áreas maiores do que as que eles ocupam numa dada instituição.

Desse modo, pode-se dizer que os textos produzidos pelas pessoas são “determinados pelas suas posições nas instituições, pelos seus lugares dentro de certos discursos e pelos seus lugares na interseção de certos grupos de discursos” (KRESS, 1989, p. 165), em outras palavras, os textos que produzimos são determinados pela história discursiva de cada um de nós.

Os textos surgem em contextos sociais específicos e são construídos de acordo com certos propósitos comunicativos. Os textos expressam significados que estão além deles, que estão nos discursos que os compõem e originam. “A língua sempre acontece como textos, e não como palavras e frases isoladas” (KRESS, 1989, p. 165). Nesse sentido, os textos são a realização material dos discursos. No entanto, os textos servem às ocasiões sociais das quais fazem parte. As características dessas situações, ou seja, seus contextos, seus participantes, seus propósitos comunicativos afetam diretamente os textos que são produzidos nessas situações. Essas ocasiões sociais são sempre “convencionalizadas”, apesar de seu grau de “convencionalização” poder variar dependendo do tipo de situação social. Algumas situações são totalmente “convencionalizadas” como, por exemplo, um casamento, enquanto outras são mais imprevisíveis, como uma conversa casual.

A “convencionalização” das situações e de suas formas linguísticas leva à convencionalização dos textos, ou seja, leva à criação dos gêneros discursivos. Nas palavras de Kress, “os gêneros têm formas e significados específicos, que derivam das funções, propósitos e significados das ocasiões sociais. Portanto, os gêneros se constituem num catálogo preciso das situações sociais relevantes de uma comunidade num dado momento” (KRESS, 1989, p. 168).

A comunidade desenvolve os gêneros, e a existência de gêneros configura grupos sociais como comunidades discursivas por compartilharem propósitos comunicativos efetivados através dos gêneros pertinentes a ela, portanto, um processo de “auto-alimentação” entre a comunidade

discursiva e os gêneros discursivos (ARANHA, 2007). Podemos dizer que os gêneros são sócio-historicamente construídos, não são somente objetos textuais mais ou menos semelhantes, mas eventos codificados inseridos em processos sociais comunicativos.

Um gênero compreende uma classe de eventos comunicativos, cujos exemplares compartilham os mesmos propósitos comunicativos. Esses propósitos são reconhecidos pelos membros mais experientes da comunidade discursiva original e constituem a razão do gênero. A razão subjacente dá o contorno da estrutura esquemática do discurso e influencia e restringe as escolhas de conteúdo e estilo. O propósito comunicativo é o critério que é privilegiado e que faz com que o escopo do gênero se mantenha focado estreitamente em determinada ação retórica compatível com o gênero. Além do propósito, os exemplares do gênero demonstram padrões semelhantes, mas com variações em termos de estrutura, estilo, conteúdo e público-alvo. Se forem realizadas todas as expectativas em relação àquilo que é altamente provável para o gênero, o exemplar será visto pela comunidade discursiva original como um protótipo. Os gêneros têm nomes herdados e produzidos pelas comunidades discursivas e importados por outras comunidades. Esses nomes constituem uma comunicação etnográfica valiosa, porém normalmente precisam de validação adicional.” (SWALES, 1990, p.58).

Apesar dos usos dos documentos de arquivo serem múltiplos e ubíquos, sobretudo após a aprovação da lei de acesso à informação, na fase de produção documental, majoritariamente os usos estão associados à realização das tarefas, atividades e funções das unidades de trabalho da entidade produtora. Os documentos e suas informações costumam ser recuperados para viabilizar os próprios processos de trabalho.

Diante dessa perspectiva, a indexação como processo de representação dos documentos que auxilia a sua recuperação, precisa considerar o mapeamento dos processos de trabalho para estabelecer suas diretrizes na fase de produção documental nos arquivos de universidade. Portanto, o mapeamento dos processos pode melhor esclarecer como os documentos são usados e para quais finalidades. Algumas universidades registraram as rotinas administrativas de suas unidades de trabalho em suas normativas como regimentos, regulamentos, portarias e instruções normativas como a Universidade de Brasília, a Universidade Federal do Paraná, a Universidade Federal de Santa Catarina. Na Universidade Federal Fluminense, nosso campo empírico, não há uma padronização no registro das funções e atividades da instituição. Pela autonomia das unidades de trabalho como as pró-reitorias, conselhos, faculdades, institutos, departamentos de ensino e coordenações de curso, tais unidades são responsáveis por normatizar suas atividades.

A UFF possui seu estatuto e regimento geral que normatizam as funções da instituição, descrevem as atividades de modo lato e apresenta a estrutura da universidade, dispõe inclusive sobre a organização administrativa. Ademais, o regimento geral estabelece normas gerais que disciplinam as atividades da instituição. Entretanto, as atividades específicas de cada unidade de trabalho devem ser reguladas pela própria, inclusive com regulamentos próprios como esclarecem as normas supracitadas. Algumas unidades possuem regimentos internos desde pró-reitorias como a

a de Gestão de Pessoas e a de Extensão; institutos como o de Letras, de Computação, inclusive de campi do interior do estado do Rio de Janeiro como o de Ciências da Sociedade, de Macaé, o Instituto de Ciências Humanas e Sociais, de Volta Redonda, e do Instituto de Ciência e Tecnologia, de Rio das Ostras. Ademais, alguns departamentos regularam suas atividades por meio de seus próprios regimentos internos como o de História, o de Química, e o de Física, este localizado em Volta Redonda.

Há outras disposições gerais como o regulamento dos cursos de graduação e o regulamento geral dos cursos de pós-graduação *stricto sensu* que definem processos de trabalho nesse âmbito para toda entidade produtora, perpassando unidades de trabalho envolvidas nessas funções como coordenações de curso e de programas de pós-graduação, departamentos, institutos e pró-reitorias.

Essas regulações são importantes, pois registram e padronizam os processos de trabalho, facilitando o mapeamento destes e, por conseguinte, os usos dos documentos na fase de produção documental. Não se deve cristalizar os usos dos documentos e ações das unidades de trabalho, pois eles fazem parte da dinâmica dos agentes produtores em um organismo vivo que é a própria entidade. No entanto, a razão da existência da instituição está vinculada a suas atividades-fim que possuem certa estabilidade. Ademais, há a cultura administrativa inserida em seu contexto sócio-histórico que tendem a se alterar lentamente.

Ao abordar a garantia literária, Mario Barité (2007) afirma que é preciso ampliar tal conceito em seu uso para a representação dos documentos, inclusive considerando o que o autor chamou de garantia organizacional. Retomando Beghtol (1995), Barité compartilha o conceito de que a garantia literária se trata de um conjunto de princípios ao redor dos quais uma literatura é estabelecida. “A garantia literária leva de forma implícita a apelação ao consenso científico, acadêmico e especializado, em razão do qual os documentos são elementos de registro, socialização e legitimação do novo conhecimento gerado” (BARITÉ, 2007, p. 2, tradução nossa). De certa forma, a garantia literária retoma os acordos, convenções, regras na produção de sentido abordadas anteriormente a partir de autores como Wittgenstein (1999), Fish (1992) e Swales (1990).

O uso da garantia literária para a representação dos documentos em processos como a classificação e a indexação não parte de uma organização do conhecimento apriorística, baseada em aspectos formais extraídos de conceitos abstratos e universais encontrados nos documentos, mas é estabelecida a partir de consensos, acordos e convenções de determinada comunidade. Barité (2007) identifica, na literatura da Ciência da Informação, pesquisas e relatos acerca da construção de vocabulários controlados construídos para satisfazer as necessidades de comunicação e intercâmbio de informações entre sujeitos que convivem em uma organização de grandes dimensões. Nessas entidades, a eficiência do trabalho e o cumprimento das atividades-fim estão sendo afetadas por

problemas não somente da gestão de documentos, mas pela falta de gestões eficazes do conhecimento e da informação técnica-administrativa e logística.

Há nas unidades de trabalho, dinâmicas, trânsitos, relações, chavões, usos particulares que refletem na própria forma de representar e recuperar os documentos. É necessário considerar as características e os contextos de cada organização para realizar a indexação dos documentos de arquivo na fase de produção.

Barité e Álvarez (1999) identificam como componentes da garantia literária para embasar as representações dos documentos a validação sociolinguística – quando os termos usados com certa regularidade tanto no ambiente de trabalho, quanto nos documentos; a validação terminológica – quando os termos já foram registrados em glossários, manuais e dicionários especializados, incluindo normas técnicas e revistas de referência; a validação linguística – quando um termo novo é registrado em um instrumento autorizado ou aceito por uma autoridade na língua como um dicionário de referência; a validação jurídica – quando a autorização é vinda pela legislação e ordenamento jurídico.

Essa ampliação do conceito de garantia literária vai ao encontro da proposta dessa tese de realizar o processo de indexação a partir da produção de sentido pelos usos dos documentos e assim realizar a representação dos documentos. As validações estão em consonância com a concepção de que os sentidos dos termos são baseados nas convenções, normas, consensos construídos nas relações entre os membros de dada comunidade.

A indexação de documentos de arquivo nas universidades deve ser baseada por uma consistente política de indexação. De acordo com Fujita (2012, p. 17), a política de indexação é um conjunto de critérios e decisões que refletem os interesses e objetivos de um sistema de informação, sobretudo, um sistema de recuperação da informação. Ao considerar o alcance, a abrangência, as necessidades de uso dos documentos, o histórico de demandas para a elaboração de uma política de indexação, a instituição pode usar o processo de indexação como estratégia gerencial em seu sistema de informação, sobretudo considerando a eficácia na recuperação de seus documentos, como aponta a autora.

A política de indexação pode ser determinada em um sistema de armazenagem e recuperação da informação pela seleção de tipos de documentos a serem indexados, procedimentos de análise e representação de assuntos, aspectos qualitativos da indexação como precisão, especificidade, exaustividade e revocação, instrumentos de controle de vocabulário tais como linguagens documentárias ou opção por trabalhar com linguagem natural, além da avaliação da indexação pela consistência e pela recuperação. Todos esses aspectos, entretanto, ganham significado quando aplicados ao contexto de um sistema de armazenagem e recuperação da informação que possui finalidades e objetivos e abriga condições em seu ambiente quanto à natureza da informação produzida e solicitada, bem como características da comunidade de usuários (FUJITA, 2012, p. 22).

Nessa perspectiva, a política de indexação para os arquivos de universidade, considerando os usos de seus documentos na fase de produção, precisa estar associada às políticas institucionais, à política de informação e à gestão documental. Carneiro (1985) afirma que uma política de indexação deve ser elaborada a partir da identificação da organização, das atividades, dos documentos, incluímos nesta lista o mapeamento dos processos de trabalho e o histórico das demandas de uso dos documentos. A partir da análise dessas variáveis é possível elaborar uma política de indexação que atenda ambos os eixos identificados por Guimarães (2004): o horizontal – associado às atividades de gestão – e o vertical – referente às atividades de organização e representação dos documentos. A seguir apresentamos um quadro com os principais requisitos apontados por autores da área para serem considerados na elaboração da política de indexação:

AUTORES	EIXO HORIZONTAL	EIXO VERTICAL
CARNEIRO (1985)	<ul style="list-style-type: none"> • SISTEMA DE ARMAZENAMENTO E RECUPERAÇÃO DA INFORMAÇÃO • CLIENTELA • RECURSOS FINANCEIROS, MATERIAIS, HUMANOS 	<ul style="list-style-type: none"> • COBERTURA DE ASSUNTOS • SELEÇÃO E AQUISIÇÃO DE DOCUMENTOS FONTE • PROCESSO DE INDEXAÇÃO • ESTRATÉGIA DE BUSCA • TEMPO DE RESPOSTA DO SISTEMAS • FORMA DE SAÍDA • AVALIAÇÃO DO SISTEMA
FUJITA (2012)	<ul style="list-style-type: none"> • ARMAZENAGEM E RECUPERAÇÃO DA INFORMAÇÃO • SELEÇÃO DOS DOCUMENTOS A SEREM INDEXADOS • FINALIDADES E OBJETIVOS DO SRI • AMBIENTE INFORMACIONAL • CARACTERÍSTICAS DA COMUNIDADE DE USUÁRIOS • INFORMAÇÃO PRODUZIDA E SOLICITADA 	<ul style="list-style-type: none"> • REVOCAÇÃO E PRECISÃO DA INDEXAÇÃO • EXAUSTIVIDADE E ESPECIFICIDADE • LINGUAGENS DOCUMENTÁRIAS • LINGUAGEM NATURAL • AVALIAÇÃO DA INDEXAÇÃO POR SUA CONSISTÊNCIA
GIL LEIVA (2008)		<ul style="list-style-type: none"> • EXAUSTIVIDADE E ESPECIFICIDADE • CORREÇÃO E CONSISTÊNCIA •
LANCASTER (1993)		<ul style="list-style-type: none"> • EXAUSTIVIDADE E ESPECIFICIDADE

Na realidade atual das universidades federais, sobretudo na fase de produção documental, com a adoção de sistemas em que o usuário é o próprio gestor do processo a exemplo do Sistema

Eletrônico de Informação (SEI), é importante o desenvolvimento de uma política de informação bem estruturada. A política de indexação expressa como norma a ser adotada pela comunidade acadêmica permite a uniformidade de condutas, processos e ações no tocante a análise e representações dos documentos dessas instituições, como afirma Guimarães (2004). Ademais, o registro da política de indexação permite à sua comunidade entender a própria forma com que os documentos são concebidos e organizados. De certo modo, é um reflexo de como os sujeitos veem seus processos de trabalho, seu contexto, suas ações, suas representações e suas maneiras de se comunicarem. Essa sistematização facilita inclusive a avaliação e aprimoramento tanto do processo, quanto da política de indexação.

Nos atuais sistemas de informação nos quais os documentos de arquivo fazem parte, as formas de comunicação e o imperativo tecnológico têm transformado os usuários em sujeitos responsáveis não somente pela maneira de recuperar os documentos demandados, no sentido de realizarem as perguntas corretas, compatíveis com a linguagem de indexação adotada, mas também em atuadores, classificadores, indexadores de seus próprios processos, caso sejam eles seus produtores. Como relatamos na seção anterior, em sistemas como o SEI, o requerente torna-se o responsável pelo registro, início do trâmite e seu acompanhamento até o fim do seu processo. Ademais, as formas de recuperação desse documento estão condicionadas à classificação e à entrada dos termos que o requerente fez na autuação do processo no momento de seu registro no Sistema Protocolo Integrado.

Portanto, um funcionário que almeja a sua progressão não precisa preencher um requerimento para dar entrada em seu pedido junto ao protocolo de sua unidade de trabalho. Ele precisa realizar as próprias funções do protocolo. No caso do SEI, registrar, autuar o processo, inclusive classificando o documento de acordo com os códigos de classificação do CONARQ, encaminhar para unidade responsável pela análise do pedido, além de escolher os termos de indexação para futura recuperação.

A política de indexação permite estabelecer diretrizes que padronizem esses sistemas não ficando a cargo do empirismo do requerente ou daqueles que desenvolvem a arquitetura desses sistemas. Esses critérios e diretrizes estabelecidos em uma política de indexação são frutos da observação de aspectos importantes da comunidade em questão como, no caso dos arquivos de universidade, das funções, atividades e – aqui traçando uma relação próxima com a classificação -; do mapeamento dos processos de trabalho que revelam os usos dos documentos e de suas informações; do volume e característica dos documentos; do histórico das demandas e as formas de recuperação destes documentos.

Há outros pontos, elencados por Rubi (2012), como importantes a serem observados para uma política de indexação bem elaborada como: o número e a qualidade de recursos humanos

envolvidos no desenvolvimento dos sistemas de informação; determinação dos recursos financeiros disponíveis para criação e manutenção do sistema; determinação dos equipamentos disponíveis.

A partir desses elementos analisados a política de indexação pode propor as formas que o processo de indexação deve ser conduzido, como o estabelecimento da cobertura de assuntos. A escolha dos assuntos centrais e periféricos deve está associada ao estudo dos usos dos documentos por sua comunidade. Na fase de produção documental nas universidades, nosso estudo revela que as demandas para a recuperação dos documentos estão associadas ao próprio desenvolvimento dos processos de trabalho de seus produtores. Aquele que demanda pelo documento é o próprio sujeito que o produziu. Por isso é importante o mapeamento dos processos de trabalho e seu registro, assim como a definição da cobertura dos assuntos a partir deles.

A exaustividade e a especificidade na indexação estão diretamente relacionadas com a capacidade de revocação e precisão do sistema de recuperação da informação. Como afirma Borko (1977), a exaustividade refere-se à quantidade de termos atribuídos a um documento, ou seja, à quantidade de assuntos atribuídos a um mesmo documento. Já a especificidade diz respeito ao nível de abrangência que o sistema e a linguagem de indexação permitem especificar os conceitos atribuídos ao documento.

A capacidade de revocação diz respeito ao número de documentos recuperados e pode ser mensurada por meio da relação entre o número de documentos relevantes sobre determinado assunto, recuperados pelo sistema de busca, e o número total de documentos sobre o tema, existentes nos registros do mesmo sistema. A capacidade de precisão, ou relevância, está relacionada ao número de documentos recuperados para atendimento das solicitações encaminhadas pelo usuário.

Portanto, a relação é inversamente proporcional, quanto mais exaustiva for a indexação, a precisão será menor, mas a revocação será maior. De outro lado, quanto mais específica for a forma de indexar, menor será a revocação, mas a capacidade de precisão aumenta. Então, a escolha deve estar baseada no tamanho do sistema, mas, sobretudo, nas formas em que os documentos são descritos e discriminados. Retomando Blair (2006), é preciso considerar a capacidade de descrever o documento, no sentido de representar as informações que, de fato, serão precisamente recuperadas, mas também considerar a capacidade de discriminar assuntos potencialmente de interesse para aqueles que usam tal sistema. Essas questões também estão relacionadas ao tempo de resposta do sistema e ao tempo gasto pelo usuário na varredura e análise do assunto de seu interesse.

Outro elemento importante para a política de indexação é a escolha da linguagem de indexação. De acordo com Rubi (2012), a linguagem adotada afeta o desempenho do sistema de recuperação da informação tanto na estratégia de busca, quanto o processo de indexação em si.

Portanto, a escolha pela linguagem livre, controlada, pré-coordenada, pós-coordenada deve ser baseada no estudo de como a comunidade recupera, usa e se relaciona por meio dos documentos. A estratégia de busca, que se escolhe por ser delegada ou não, também pode ser indicada pela política de indexação.

A forma de saída, ou seja, como os resultados da busca são apresentados e a avaliação do sistema são fatores apresentados também pela política de indexação.

Os documentos de arquivo têm diversos usos como registro, evidência, comunicação, informação. Como afirma Delmas (2010), os arquivos servem para lembrar, provar, compreender e identificar-se.

O uso dos documentos como evidência possui uma influência do ordenamento jurídico, como dito ao longo dessa tese, trata-se de conseguir provar direitos e deveres pela mediação de documentos, legalmente aceitos e socialmente validados por cumprirem os requisitos acordados pelos membros desta, ainda que coercitivamente pelos fatos sociais.

Como afirma Bearman (1995), a evidência é a fonte social e lega identidade e significância aos arquivos. A informação gerida pelos arquivistas são os documentos de arquivo. Diferem-se dos demais por sua necessidade de garantir confiabilidade enquanto evidências. Eles precisam ser preservados por serem evidências das transações, dos processos de trabalho. Os arquivistas precisam garantir a fidelidade na produção, a autenticidade na custódia para assim assegurar os usos dos documentos de arquivo como tais. Esse é a primeira responsabilidade dos arquivistas, eles precisam garantir o documento de arquivo como evidência.

Já o registro facilita a lembrança e lembrar-se não apenas facilita a construção da identidade dos sujeitos, o despertar da autonomia, a compreensão de seu contexto sócio-histórico, mas lembrar-se é uma utilidade para a gestão. As tomadas de decisão podem ser baseadas na análise de erros e acertos da própria história da entidade. “Lembrar é uma necessidade prática da vida cotidiana de qualquer pessoa ou instituição” (DELMAS, 2010, p. 26). Essa necessidade refere-se à continuidade dos sujeitos e instituições como organismos, ou seja, a continuidade de suas ações.

Os arquivos também são fontes de conhecimento e informação não somente pelos usos consagrados no século XIX com o positivismo francês e o historicismo alemão, ou como fontes para procedimentos administrativos, mas também como utilidade pública para reconhecer histórias coletivas, fiscalizar ações governamentais, referências morais e psicológicas, desenvolvimento científico e tecnológico, entre outros usos dos conhecimentos registrados em documentos de arquivo (DELMAS, 2010, p. 39).

Sobre os processos comunicacionais, os documentos de arquivo na fase de produção estão mediando relações a todo momento, operam diversos trânsitos entre sujeitos e instituições, instituições e instituições, sujeitos e sujeitos. Os sujeitos agem e interagem no domínio da

linguagem e, como foi arguido anteriormente, os documentos mostram a materialidade dessas ações e interações. Os documentos de arquivo são resultantes de atividades públicas ou privadas, mas que costuma ser regulados por legislações. São canais de comunicação, relativamente, padronizados com circuitos definidos. As formas de comunicação costumam ser sistemáticas por essas razões.

O segundo propósito dos arquivos, de acordo com Bearman (1995), é o da documentação. A documentação em dois sentidos: aquela que assegura o acesso ao documento por meio de sua representação e a documentação dos procedimentos adotados na organização e preservação dos mesmos. A documentação deve conter o conteúdo, a estrutura e o contexto associados aos processos de trabalho. A documentação serve ao uso. Ela é necessária por possibilitar a recuperação do documento para seus diversos usos.

Timothy Ericson (1988) argumenta que a meta dos documentos de arquivo não é a guarda, mas o uso. Documentos são identificados, classificados, indexados, preservados para atingir essa meta: o uso. Caso essas operações não tenham como princípio fundamental propiciar o acesso, tendem a serem fracassadas, pois serão acumulados e tratados documentos que não cumprem sua razão de existência.

A indexação nos arquivos de universidade, cujos métodos são pré-definidos na arquitetura dos sistemas de informação, sem análise das especificidades tanto dos arquivos, quanto das universidades, estimulam a adoção de termos genéricos que pouco se aproximam da produção de sentido dos membros dessas comunidades, tendo pouca aderência aos usos dos documentos, aos significados reconhecidos pelo grupo e às ações que os perpassam. “Estudar os usos dos documentos, os usuários e suas interações com os sistemas de informação é uma condição essencial aos arquivos” (BEARMAN, 1995, p. 404). É preciso saber como a comunidade pergunta, como formula suas questões, como acredita que uma resposta é constituída.

É preciso promover os arquivos como fontes de informação. Esta é uma grande demanda na fase de produção documental para as entidades, pois facilitam a execução dos processos de trabalho. O foco na análise organizacional e na modelagem de processos de trabalho facilita tanto a representação de documentos, quanto a sua busca, pela compreensão dos seus usos pelos próprios membros da comunidade.

A proposta da indexação é fornecer pontos de acesso que facilitem o “onde-sobre” de um assunto ou transação particular ou um grupo de assuntos ou transações. Todo documento de arquivo possui informações de um ou outro ou de ambos (BEARMAN, 1995, p. 411). A tarefa da indexação é decidir que tipo de informação será mais útil para o sistema de informação. Essa decisão, que pode estar indicada na política de indexação da instituição, precisa estar baseada no histórico de informações requeridas pelas unidades de trabalho, na fase de produção documental, associada aos processos de trabalho.

Bearman (1995, p. 412) aponta alguns problemas na representação dos documentos de arquivo a serem equacionados pela indexação dentre eles:

- A mesma informação em um documento pode ser útil em diferentes contextos de assunto e em diferentes tempos.
- Pode haver uma grande confusão sobre o sentido do assunto no documento. Essa confusão pode acontecer inclusive na escolha do termo ainda que se use um instrumento como o vocabulário controlado, pois há uma quantidade considerável de diferentes temas a serem abordados. Essa confusão é ainda maior na atribuição dos títulos dos documentos quando, geralmente, a linguagem utilizada é livre.
- Os assuntos correlatos podem ficar amplamente dispersos ainda que índices sejam utilizados por conta das infinitas possibilidades da linguagem, caso combinações especiais não sejam realizadas.

Nessa perspectiva, os instrumentos da indexação devem ser adotados nos arquivos como cabeçalhos de assunto, vocabulários controlados, tesouros, taxonomias, folksonomias, ontologias de domínio. A escolha do instrumento deve estar indicada na política de indexação, após a análise de como a comunidade usa seus documentos e realiza suas buscas. A elaboração do instrumento, inclusive, poderá ter maior eficácia se associada ao sistema de classificação da instituição, associando as funções e atividades realizadas pela entidade com os assuntos identificados nos processos de trabalho. Na fase de produção documental nos arquivos de universidade, a produção de sentido dos documentos se dá no desenvolvimento dos processos de trabalho.

A indexação na fase de produção documental dos arquivos de universidade pode facilitar o estabelecimento de relações entre os objetos dos documentos que estão visíveis apenas em seus conteúdos particulares e não foram clarificadas pelas outras formas de representação dos documentos como a própria classificação.

A padronização das entradas possibilitada pela adoção de uma linguagem de indexação na fase de produção documental dos arquivos deve considerar alguns aspectos tais como:

- O assunto atribuído está relacionado a qual ação e em que termos?
- Quais unidades de trabalho podem ser associadas aos usos daqueles documentos indexados?
- Que tipos documentais o assunto atribuído é frequentemente encontrado?
- Há limitações em termos de tempo espaço?

Essas questões podem facilitar a construção do instrumento em sua fase mais sensível que é a escolha dos termos autorizados. Esses devem ser checados em seu potencial de utilidade para os membros de sua comunidade, no caso da fase de produção documental dos arquivos de universidade, sobretudo, os sujeitos envolvidos com os processos nas unidades de trabalho. Assim,

as formas que os processos têm sido autuados nos protocolos tornam-se um recurso básico para a indexação. Os conjuntos de casos, as relações particular-instâncias, a recorrência de determinados assuntos pode nortear a construção da linguagem de indexação para o arquivo.

A aproximação da indexação à classificação nos arquivos é necessária também porque o contexto e as circunstâncias de produção são determinantes do significado do documento. Se os contextos das atividades e práticas no desenvolvimento dos processos de trabalho são importantes para compreender os assuntos dos documentos, eles devem estar no centro da análise de assunto realizada no processo de indexação. No caso dos arquivos, a indexação precisa considerar como os produtores costumam nomear seus documentos e suas atividades.

Se quisermos saber como as descrições costumam representar um significado do documento, nós precisamos examinar como essas descrições são usadas nas atividades e práticas daqueles que usam aquela informação, como as pessoas perguntam por tal informação e como elas falam sobre isso (BLAIR, 2003, p. 37)²⁸.

Portanto, o mapeamento dos processos de trabalho pode indicar as formas que os sujeitos nomeiam suas ações e como realizam as buscas pelos documentos que viabilizem tais processos. São acordos, regras, convenções tácitas e naturalizadas que podem passar a serem percebidas a partir do registro desse mapeamento.

A indexação é capaz de recuperar assuntos mais específicos que a classificação não represente. No entanto, quando a indexação é associada à classificação, aquele que busca o documento pode ter uma visão mais ampla do contexto que este está inserido. Quando é autuado no SEI um processo de capacitação de docente no exterior com ônus limitado no SEI, o código atribuído, de acordo com o Código de Classificação de Documentos de Arquivo relativos às atividade-fim das Instituições Federais de Ensino Superior, do Arquivo Nacional, é o 029.222 – que possui a seguinte notação “autorização de afastamento; diárias, inclusive compra de moeda estrangeira; lista de participantes, em caso de comitiva”. Esse código também é atribuído aos demais técnicos das IFES. A amplitude dos processos que podem ser classificados por este código como: o afastamento de docente para participação em eventos, curso de capacitação, realização de pós-graduação ou participar de uma comitiva – pode levar a alto número de documentos que podem ser recuperados.

Ademais, há diversos processos de trabalho envolvidos desde o requerimento feito pelo professor até a publicação da autorização do reitor no Diário Oficial da União que podem ser de interesse informacional aos agentes envolvidos no processo, cuja recuperação pelo código de classificação torna-se inviável. O plano de trabalho feito pelo professor, a carta aceite do orientador

²⁸Cabe destacar que Blair não está trabalhando com o conceito de descrição na perspectiva de função arquivística, mas na direção de nomear os documentos de forma precisa para aqueles que o querem recuperar. Conceito já trabalhado nesta seção ao diferenciarmos descrição de discriminação dos documentos.

no exterior, o parecer da comissão de afastamento, a ata departamental aprovando ou negando o pedido, o despacho do reitor são exemplos de documentos gerados durante o processo que estão invisíveis na classificação, mas podem ser representados na indexação. Outras entradas como o objetivo da viagem; o tempo de afastamento; o orientador responsável no exterior; a data de aprovação do pedido podem ser atribuídas de acordo com o estudo das demandas relacionadas a essa atividade na Universidade.

A padronização na entrada nos termos é importante e o controle pode ser customizado no sistema adotado. No entanto, a indexação terá mais eficácia, possibilitando a recuperação do documento, caso aquele que for indexar esteja familiarizado com os termos do sistema e partilhe das concepções de busca daquele que irá demandar pelo documento. Portanto, o sentido do documento precisa ser partilhado e esse sentido é produzido em seu uso.

A realidade dos sistemas de informação contemporâneos aponta muitas vezes para o usuário como próprio indexador de documentos, a exemplo do SEI que o próprio requerente se torna gestor do seu processo, inclusive realizando a representação de sua informação. Esses sujeitos não necessariamente conhecem todas as atividades desenvolvidas pela instituição da qual fazem parte, tampouco os processos de trabalho que geram documentos de arquivo. Todavia, conhecem a realidade das unidades de trabalho em que estão inseridos e as rotinas administrativas. Ademais, compartilham os usos dos documentos, ideais, valores, culturas, ou seja, operam os mesmos fatos sociais totais de sua comunidade discursiva. Assim, os sistemas poderiam ser configurados para direcionarem a indexação dos documentos somente para as unidades de trabalho que produzem tais documentos.

Todavia, é importante considerar e interrelacionar as formas de busca das unidades que participam do trâmite e/ou que usam tais documentos para a execução de seus processos de trabalho. Há consensos e muitas vezes as alterações de indexação nos documentos não são realizadas, pois há um entendimento naturalizado das formas de indexação e uma aceitação tácita destas realizadas na entrada dos documentos nos sistemas. Ainda assim, a indexação que fará parte do documento sempre será aquela feita ou aceita pelo último membro da comunidade discursiva que usou o documento.

8 CONCLUSÃO

Este trabalho procurou explorar o processo de indexação no universo dos arquivos de universidade, considerando a fase de produção documental e a atuação do protocolo como primeira unidade de arquivamento no sistema de informação das universidades federais.

O questionamento de como os documentos de arquivo estão sendo representados na perspectiva do tratamento temático da informação pelos sistemas de informação das universidades federais brasileiras surgiu da observação da realidade ao nosso redor: a Universidade Federal Fluminense. Percebemos alguns gargalos existentes na recuperação dos documentos e o quanto estes afetam as próprias atividades e funções da instituição.

O ponto de partida desdobrou-se em uma esfera complexa de sujeitos, entidades, agências, processos, trânsitos que procuramos abordar sem a pretensão de esgotá-los, mas tentando traçar caminhos que ofereçam alguns indícios do como se realiza o processo de indexação na fase de produção documental dos arquivos de universidade e o porquê ou porquês acontece de tal ou tais formas.

Destaca-se que o escopo dessa pesquisa se delimita à fase de produção documental. Essa circunscrição deve-se a alguns pressupostos que embasam essa pesquisa. Sobre a fase de produção documental, apresentamos diferentes perspectivas do campo arquivístico como a gestão de documentos, a arquivística integrada e os *records continuum* com intuito de ampliar nossa visão sobre o tratamento e organização dos documentos de arquivo. Cabe esclarecer que a delimitação na fase de produção documental feita por este trabalho é um procedimento metodológico com o objetivo de delimitar o campo de análise. Como foi demonstrado, partilhamos da compreensão de que os arquivos fazem parte de múltiplas dinâmicas com contextos variados de ações e relações sincrônicas e diacrônicas. Seus documentos atendem a diversos usos, não somente são produzidos, mas produzem atos e fatos em diferentes tempos e espaços.

Não entendemos as fases do ciclo vital como estanques, cristalizadas e teleologicamente sucessivas. Não há uma progressão linear dos arquivos porque são múltiplas as dimensões que seus documentos percorrem.

Ao analisarmos as universidades federais brasileiras, investigamos como elas configuram-se enquanto entidades produtoras, suas funções e atividades. Ademais, relacionamos os usos dos documentos aos seus processos de trabalho, pois percebemos que eles estão diretamente relacionados. Um dos pressupostos que amadureceu na trajetória dessa pesquisa foi que para compreender os usos dos documentos na fase de produção mister conhecer as unidades de trabalho, suas funções, suas atividades e suas rotinas. Nesta abordagem, consideramos as especificidades dos documentos de arquivo.

As unidades de trabalho na fase de produção documental possuem especificidades no uso da informação e dos documentos que refletem na própria busca e precisam ser consideradas como o fato dos usos dos documentos e as demandas por sua recuperação estarem diretamente relacionados aos processos de trabalho, essas questões foram observadas na quarta seção desta tese.

A universidade como entidade produtora, por sua autonomia, tem processos de trabalho muito diferentes da maioria das administrações das instituições que produzem documentos de arquivo. A indexação, apesar de ser amplamente utilizada no cotidiano dos protocolos dos arquivos, tem sido pouco problematizada tanto pela Ciência da Informação no tocante ao campo de análise, quanto pela Arquivologia em relação à operação em si.

Assim, investigamos o contexto sócio-histórico recente que as universidades federais passaram a partir da Reestruturação e Expansão das Universidades Federais (Reuni), instituído pelo decreto nº 6.096, de 24 de abril de 2007. Houve um grande impacto nas atividades dessas instituições que, por conseguinte, alteraram os processos de trabalho e a produção documental. Essa realidade apresentada na sexta seção ajuda-nos a dimensionar a complexidade da gestão de documentos dessas instituições.

No início da pesquisa, ao realizarmos o levantamento das fontes relacionadas à temática e a revisão de literatura acerca da indexação em arquivos, constatamos certa escassez de estudos dedicados a esse campo. No entanto, percebemos na análise empírica dos sistemas de informação das universidades federais brasileiras que o processo de indexação é uma operação bastante usada pelos arquivos de universidade para fazer a representação do conteúdo dos seus documentos.

No levantamento bibliográfico preliminar na base LISA acerca do termo "indexing" como assunto principal na pesquisa avançada, encontramos 21.419 ocorrências em periódicos acadêmicos que compõem esta base de dados. Sendo impossível a análise de todas as ocorrências, procuramos delimitar a pesquisa por títulos em periódicos, ainda assim encontramos 2.263 títulos dos quais ainda estamos analisando formas possíveis de delimitação deste escopo.

De forma semelhante, realizamos um levantamento bibliográfico na Base de Dados em Ciência da Informação (BRAPCI), base brasileira composta por 53 revistas científicas da área da Ciência da Informação. Buscamos pelo termo "indexação" em palavras-chave, encontramos 100 artigos sobre o tema, estamos revisando estas referências.

Com o objetivo de melhor qualificar nossa análise dos artigos procuramos delimitar nossa pesquisa considerando os periódicos com um maior número de publicações e mais acessados na área da Ciência da Informação e da Arquivologia. Realizamos então levantamentos acerca dos principais artigos sobre indexação e indexação em arquivos no *Journal of Documentation*, *Knowledge Organization*, *The Indexer*, *The American Archivist* e *Archivaria*. Os resultados obtidos na revisão que realizamos encontram-se nas seções apresentadas ao longo desse trabalho.

Essa etapa ajudou-nos a compreender a indexação como processo de análise conceitual do documento e representação dos seus assuntos que pode acontecer em linguagem livre ou em linguagens de indexação. Outra importante constatação foi que os assuntos dos documentos estão relacionados à produção de sentido que sua comunidade discursiva realiza. A representação dos documentos de arquivo para sua futura recuperação precisa estar associada aos usos que a comunidade discursiva em questão faz dos mesmos. Os sujeitos recuperam ou constroem as suas demandas pelos documentos para viabilizar suas ações. Os documentos mediam relações. Essa é a razão para sua criação, organização, arquivamento e recuperação.

O estudo da genealogia da indexação, a leitura de autores clássicos tanto no campo da classificação, quanto da indexação apontaram que, sobretudo na escola inglesa, os debates acerca da indexação foram bastante influenciados pelas teorias da classificação. Essa constatação foi apresentada na segunda seção desta qualificação.

No entanto, as teorias da indexação apontam diversos caminhos possíveis para a indexação dos documentos. Essas possibilidades que abordamos na segunda seção dessa tese mostraram o quanto as filosofias da linguagem são importantes no desenvolvimento do *corpus* teórico acerca da informação. A indexação como um processo informacional está inserido em um processo comunicacional e não há comunicação se não houver compartilhamento de signos e interpretações.

A problematização da produção de sentido pelas leituras realizadas no campo dos estudos da linguagem, conjuntamente com as teorias da indexação, tornou-se uma das principais questões abordadas, uma vez que se configurou a importância da produção de sentido para a atribuição do assunto. Ao indexar um documento, produz-se sentido acerca dele e a técnica para o estabelecimento do assunto dependerá da filosofia de linguagem que o indexador adota em sua abordagem ao documento. Diante da revisão literária que realizamos, percebemos a predominância de alguns paradigmas do campo da semântica na compreensão da indexação, dentre estes paradigmas problematizamos o realismo, o mentalismo, a semântica formal e cognitivismo social.

Nosso principal pressuposto é de que a indexação, entendida como um processo de análise de documentos e tradução seja para uma linguagem livre, seja para uma linguagem artificial ora denominada como documentária, ora denominada como linguagem de indexação propriamente dita, é realizada em larga escala pelos arquivos de universidade, sobretudo na fase de produção documental. No entanto, é um procedimento realizado de forma empírica e naturalizada, pouco problematizada, e que, por isso, tem tido pouca efetividade na recuperação dos documentos e/ou de suas informações. Outro pressuposto de grande relevância trata-se de entender a indexação como um processo de representação do documento baseado nos usos que os atores sociais de determinada comunidade discursiva fazem destes.

Há outras variáveis que foram desenvolvidas nas seções apresentadas neste trabalho como: a atribuição de assunto estar vinculada a produção de sentido. Assim, percebe-se a influência das teorias do sentido nos estudos sobre indexação. A forma de compreender a linguagem influencia no processo de indexação e, por conseguinte, na sua eficiência para os sistemas de recuperação da informação. Os resultados obtidos nesta investigação acerca das relações entre assunto, produção de sentido e filosofias da linguagem foram apresentados na terceira seção desta pesquisa.

Ao cotejarmos a literatura sobre indexação das filosofias da linguagem e da teoria da classificação, podemos perceber que são campos de grande interlocução e que tais filosofias influenciaram inúmeras abordagens desses processos de representação da informação e dos documentos. Ao pensar possibilidades de representação do conhecimento, os processos de classificar e indexar convergem na análise do universo documental. Suas particularidades encontram-se nos objetivos de suas representações. As classificações procuram possibilitar a compreensão de um todo, do universo de documentos classificado. Já a indexação procura fornecer pontos de acesso para que dada demanda por determinado assunto seja atendida, uma operação que trata mais das especificidades no tocante às demandas por informação e documentos.

Transpondo o processo de indexação aos arquivos, percebemos como a área trabalha a representação de seus documentos. Identificamos a descrição arquivística como função que possui tal compromisso. Pela especificidade do documento de arquivo, ou seja, de ser produzido ou recebido por uma entidade pública ou privada no exercício de suas atividades, ou de ser produzido ou recebido por um sujeito ao longo de sua trajetória de vida, a descrição possui o compromisso de manter o contexto de produção e uso desses documentos. O contexto é de extrema importância para os arquivos, pois evidencia os vínculos entre funções, ações, estruturas e os próprios documentos. A descrição está vinculada ao princípio da proveniência e da ordem original como foi abordado na quarta seção desta tese.

Revelou-se importante também analisar as diferenças entre as formas de classificação nos arquivos e a teoria da classificação que possui um maior diálogo com a teoria da indexação. São compreensões que partem de princípios bem distintos. Enquanto a teoria da classificação tem forte inspiração da tradição platônica-aristotélica cujo apagamento do contexto do documento é forçoso para a construção de classes, categorias, conceitos, assuntos que sejam semelhantes em aspectos considerados universais; a classificação arquivística necessita manter a organicidade entre os documentos e garantir a compreensão do seu contexto de produção. Essas questões abordadas ao longo da tese foram importantes para distinguir a indexação das demais formas de representação dos documentos de arquivo.

Outras análises foram trabalhadas conjuntamente com a perspectiva supracitada como: a indexação não possui o compromisso de manter o contexto de produção do documento, embora tam-

bém possa realizá-lo, este é dever da classificação arquivística, função de organização dos documentos de arquivo que tem como objetivo central garantir a organicidade e o vínculo arquivístico entre eles. O objetivo primeiro da indexação é representar os assuntos dos documentos.

O recorte temporal de nossa pesquisa encontra-se na fase de implementação do Sistema Protocolo Integrado pelas universidades federais, pois entendemos que este momento se configura como um imperativo para tais instituições repensarem seus processos de trabalho e a comunicação entre os sujeitos de sua comunidade discursiva, entendendo as mediações feitas por seus documentos. Esse imperativo não advém somente das normativas do governo federal, mas da própria contemporaneidade e da nossa sociedade, em um mundo interconectado por redes, cujos processos comunicacionais são intermediados por computadores e sistemas eletrônicos.

O protocolo como primeira unidade de tratamento documental, utilizado pela maioria das universidades federais, foi abordado na perspectiva de compreender seu papel estratégico no processo de indexação dos documentos de arquivo. Concluímos que se trata de uma importante instância de poder pelo seu papel central no processo comunicacional dessas instituições por meio da tramitação e controle de seus documentos. Essa posição estratégica, na maioria das vezes negligenciada pelos atores sociais das instituições, foi reforçada com a instituição do Sistema Protocolo Integrado pelo Ministério da Justiça e pelo Ministério do Planejamento, Orçamento e Gestão, através da Portaria Interministerial nº 2320 de 30 de dezembro de 2014.

A implementação SPI e a adesão das universidades ao SEI possibilitou as universidades repensarem seus processos de trabalho e as formas que representam seus documentos em seus sistemas de informação. Descrevemos como este processo está sendo realizado na Universidade Federal Fluminense, nosso campo empírico, a fim de pensarmos em novas abordagens sobre o processo de indexação nos arquivos de universidade, em especial na fase de produção documental, considerando o papel estratégico do protocolo.

Dessa forma, partimos de alguns pressupostos como a compreensão do arquivo como um sistema de recuperação da informação que precisa basear seus procedimentos e mecanismos a partir da análise da maneira em que os pesquisadores-administradores definem seus processos de trabalho em termos de assunto. Isto se dá porque os administradores têm usos particulares da informação, uma vez que, muitas vezes, eles são os próprios produtores das informações que demandam. Por esses usos informativos dos documentos de arquivo pelos administradores é que a gestão de documentos deve dar prioridade à representação e recuperação da informação, da mesma forma que faz com a avaliação dos documentos e sua organização.

Os processos de trabalho precisam dos documentos que viabilizem seu trâmite e muitas vezes o conteúdo informacional contido neles. Os usos administrativos e os usos informativos dos documentos são simultâneos e intrínsecos, não devem tratados de formas isoladas nos arquivos.

Por fim, oferecemos alguns apontamentos acerca do processo de indexação destacados na última seção desse trabalho, sem qualquer pretensão de esgotar tal problematização, mas com a intenção de apresentar possíveis vetores para novos estudos que reflitam sobre essas práticas sociais como os usos dos documentos de arquivo por sua comunidade discursiva.

REFERÊNCIAS

- 40 DOS 50 HOSPITAIS universitários federais estão sob gestão da EBSEH. **Correio Braziliense**, Brasília, 8 jul. 2018. Disponível em: <https://www.correiobraziliense.com.br/app/noticia/eu-estudante/tf_carreira/2018/07/08/tf_carreira_interna,693662/40-dos-50-hospitais-universitarios-federais-estao-sob-gestao-da-ebserh.shtml>. Acesso em: 20 nov. 2018.
- AI NOW RESEARCH INSTITUTE. **Algorithmic Impact Assessments: a practical framework for Public Agency Accountability**. Nova Iorque, 2018. Disponível em: <<https://ainowinstitute.org/aiareport2018.pdf>>. Acesso em: 23 out. 2018.
- ALBERCH I FUGUERAS. R. A gestión de documentos: entre la alianza interdisciplinar y la teoria del rompecabezas. **Acervo**, Rio de Janeiro, v. 28, n. 2, p. 110-120, jul./dez. 2015.
- _____. Los archivos, entre la memória histórica y la sociedad del conocimiento. Barcelona: Editorial UOC, 2004.
- _____. **Gestión integral de archivos**. Barcelona: Editorial UOC, 2001.
- ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. **The indexer**, v. 18, n. 4, p. 219-224, out. 1993.
- ALLDREDGE, E. Combining Archival and Records Management Terminology. **The American Archivist**, v. 33, n. 1, p. 61-65, jan. 1970.
- ALVES, I. et al. **Dicionário português de terminologia arquivística**. Lisboa: Instituto da Biblioteca Nacional e do Livro, 1993.
- ANDERSEN, J; CHRISTENSEN, F. S. Wittigenstein and Indexing Theory. In: ALBRECHTSEN, H; MAI, E.J. (Ed.). **Advances in classification research: Proceedings of the 10th ASIG SIG/CR Classification Research Workshop**, p. 1-26, 1999. Disponível em: <<https://journals.lib.washington.edu/index.php/acro/article/view/12479/11017>>. Acesso em: 18 dez. 2018.
- ANGUIANO PEÑA, G.; NAUMIS PEÑA. C. Method for Selecting Specialized Terms from a General Language Corpus. **Knowledge Organization**, v. 42, n. 3, p. 164-175, 2015.
- ARANHA, S. The search for more appropriate models for teaching academic writing. **Revista do GEL**, São Paulo, v. 4, n. 2, p. 97-115, 2007.
- ARAÚJO, C. A. Correntes teóricas da Arquivologia. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 18, n. 37, p. 61-82, maio/ago., 2013.
- ARENDT, H. **A condição humana**. Rio de Janeiro: Forense Universitária, 2007.
- ARISTÓTELES. **Metafísica**. Trad. Valentín García Yebra. Madrid: Editorial Gredos, 1990.
- ARQUIVO NACIONAL. (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005.
- ASSOCIAÇÃO DOS ARQUIVISTAS HOLANDESES. **Manual de arranjo e descrição de arquivos**. Rio de Janeiro: Arquivo Nacional, 1973.
- AUSTIN, J. L. **Quando dizer é fazer**. Porto Alegre: Artes Médicas, 1990.

BACHELARD, G. **O novo espírito científico**. Lisboa: Edições 70, 1986.

BANCO MUNDIAL. **Estratégia para o setor educacional**. Washington DC: Banco Mundial, 1999.

_____. **Relatório sobre o desenvolvimento mundial**. Washington DC: Banco Mundial, 1997. Disponível em: <<http://documents.worldbank.org/curated/en/634101468162557539/Relatorio-sobre-o-desenvolvimento-mundial-1997-o-estado-num-mundo-em-transformacao>>. Acesso em: 05 jan. 2019.

_____. **La enseñanza superior**. Las lecciones derivadas de la experiencia. Washington DC: Banco Mundial, 1994. Disponível em: <<http://www.bancomundial.org.br>>. Acesso em: 05 jan. 2019.

BANTIN, P. C.; BERNBOM, G. The Indiana University Electronic Records Project: analyzing functions, Identifying transactions, and evaluating recordkeeping systems – a report on methodology. **Archives and Museum Informatics**, v. 10, n.3, p. 246-266, 1996.

BARBOSA, A. **Teoria e prática dos sistemas de classificação bibliográfica**. Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação, 1969.

BARITÉ, M. La garantía literaria: vigencia y proyección teórico-metodológica. **Anais do VIII Enancib**, Salvador, outubro de 2007.

BARITÉ, M.; ÁLVAREZ, S. Terminología y Derecho: neología y bancos de datos terminológicos. **Informatio**, n. 3/4, p. 85-105, 1999.

BASSNETT-MCGUIRE, S. **Translation studies**. Londres: Routledge, 1991.

BAZERMAN, C. **Gêneros textuais, tipificação e interação**. São Paulo, Cortez, 2005.

BEARMAN, D. Archival strategies. **American Archivists**, v. 58, n. 4, p. 380-413, 1995.

_____. Diplomats, weberian bureaucracy and the management of electronic records in Europe and America. **American Archivist**, v. 55, n. 1, p. 168-181, 1992.

BEGHTOL, C. A proposed ethical *warrant* for global knowledge representation and organization systems. **Journal of Documentation**, v. 58, n. 5, p. 507-532, out. 2002.

_____. Domain analysis, literary warrant and consensus. The case of fiction studies. **Journal of the American society for Information Science**, p. 31-46, jan. 1995. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-4571%28199501%2946%3A1%3C30%3A%3AAID-ASI4%3E3.0.CO%3B2-F>. Acesso em: 18 ago. 2018.

_____. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. **Journal of Documentation**, v. 42, n. 2, p. 84-113, Jun., 1986.

BELKIN, N. J. Cognitive models and information transfer. **Social Science Information Studies**, v. 4, n. 2-3, p. 111-129, abri/jun. 1984.

- BELKIN *et al.* Ask for information retrieval: part I: backfround and theory. **Journal of Documentation**, v. 38, n. 2, p. 61-71, jun. 1982
- BELL, H. History of indexing societies: part I SI: first ten years. **The Indexer**, v. 20, n. 3, p. 160-164, Apr. 1997.
- BELLOTTO, H. **Diplomática e tipologia documental em arquivos**. Brasília, DF: Briquet de Lemos, 2008.
- _____. **Arquivos permanentes: tratamento documental**. Rio de Janeiro: FGV, 2006.
- _____. **Como fazer análise Diplomática e análise tipológica de documento de arquivo**. São Paulo: Arquivo do Estado, Imprensa Oficial, 2002.
- BELLOTTO, H. L.; CAMARGO, A. M. de A. Dicionário de terminologia arquivística. São Paulo: Secretaria de Estado da Cultura do Estado de São Paulo, 1996.
- BENJAMIN, W. Desempacotando minha biblioteca. In: _____. **Obras escolhidas**. São Paulo: Brasiliense, 1987. v. 2. p. 227-234. (Rua de mão única)
- BINDING, C.; TUDHOPE, D. **Knowledge Organization**, v. 37, n. 4, p. 287-298, 2010.
- BLAGDEN, J.F. How much noise in a role-free and link-free co-ordinate indexing system? *Journal of Documentation*, v. 22, n. 3, p.203-209, 1966.
- BLAIR, D. **Wittgenstein, language and information**. Back to the rough ground! Michigan: Springer, 2006.
- _____. Information retrieval and the philosophy of language. **Annual Review of Information Science and Technology**, v. 37, n. 1, p. 3-50, 2003.
- BOBBIO, N. **O positivismo jurídico: lições de filosofia do Direito**. São Paulo: Ícone, 1999.
- BORKO, H. Toward a theory of indexing. **Information Processing and Management**, v. 13, p. 355-365, 1977.
- _____. **Automated language processing**. New York: John Wiley & Sons. 1967.
- BORKO, H.; BERNIER, C. **Indexing concepts and methods**. New York: Academic Press, 1978.
- BOURDIEU, P. Estrutura, *habitus* e prática. In: _____. **A economia das trocas simbólicas**. São Paulo: Perspectiva, 2003, p. 337-361.
- _____. Social space and symbolic power. **Sociological theory**, v. 7, n. 1, p. 14-25, 1987.
- BRADFORD, S. C. **Documentação**. Rio de Janeiro: Fundo de Cultura, 1961.
- BRASCHER, M. Semantic Relations in Knowledge Organization Systems. **Knowledge Organization**, v. 41, n. 2, p. 175-180, 2014.
- BRASIL. Ministério da Educação. **Censo da Educação Superior 2017**. Brasília: MEC, 2018. Disponível em: <<http://emec.mec.gov.br>>. Acesso em: 25 jan. 2019.

BRASIL. Ministério da Educação. **Relatório de Gestão da Secretaria de Planejamento e Orçamento do Ministério da Educação**. Brasília: SPO/MEC, 2017. Disponível em: <<http://portal.mec.gov.br/secretaria-executiva/spo>>. Acesso em: 09 out. 2018.

BRASIL. Lei nº 13.467, de 13 de julho de 2017. Altera a Consolidação das Leis do Trabalho (CLT), aprovada pelo Decreto-Lei no 5.452, de 1º de maio de 1943, e as Leis nos 6.019, de 3 de janeiro de 1974, 8.036, de 11 de maio de 1990, e 8.212, de 24 de julho de 1991, a fim de adequar a legislação às novas relações de trabalho.

BRASIL. Ministério do Planejamento, Desenvolvimento e Gestão. **SEI Manual do Administrador**. Brasília: SETIC, 2015. Disponível em: <<https://softwarepublico.gov.br/social/sei/manuais/manual-do-administrador>>. Acesso: 18 mar. 2018a.

BRASIL. Ministério da Justiça. Ministério do Planejamento, Desenvolvimento e Gestão. Portaria nº 1.677, de 7 de outubro de 2015. Define os procedimentos gerais para o desenvolvimento das atividades de protocolo no âmbito dos órgãos e entidades da Administração Pública Federal.

BRASIL. Decreto-lei nº 8.539 de 8 de outubro de 2015. Regula o processo administrativo eletrônico nacional.

BRASIL. Ministério da Justiça; Ministério do Planejamento. Portaria interministerial nº 2.320 de 30 de dezembro de 2014. Institui o Sistema Protocolo Integrado no âmbito dos órgãos e entidades da Administração Pública Federal.

BRASIL. Lei nº 12.682, de 9 de julho de 2012. Dispõe sobre a elaboração e o arquivamento de documentos em meios eletromagnéticos.

BRASIL. Decreto-lei nº 7.724, de 16 de maio de 2012. Regulamenta a lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição.

BRASIL. Lei nº 12.550, de 15 de dezembro de 2011. Autoriza o Poder Executivo a criar a empresa pública denominada Empresa Brasileira de Serviços Hospitalares - EBSEH; acrescenta dispositivos ao Decreto-Lei nº 2.848, de 7 de dezembro de 1940 - Código Penal; e dá outras providências.

BRASIL. Decreto-lei de 15 de setembro de 2011. Institui o Plano de Ação Nacional sobre o Governo Aberto.

BRASIL. Lei 12.527, de 18 de novembro de 2011. Regula o acesso às informações.

BRASIL. Decreto-lei nº 4.915, de 12 de dezembro de 2003. Dispõe sobre o Sistema de Gestão de Documentos de Arquivo - SIGA, da Administração Pública Federal.

BRASIL. Decreto-lei nº 4.073, de 03 de janeiro de 2002. Regulamenta a lei nº 8.159, de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados.

BRASIL. Ministério da Fazenda. Portaria normativa nº 05 de 19 de dezembro de 2002. Dispõe sobre os procedimentos gerais para utilização dos serviços de protocolo, no âmbito da Administração Pública Federal, para os órgãos e entidades integrantes do Sistema de Serviços Gerais.

BRASIL. Lei nº 9.784, de 29 de janeiro de 1999. Regula o processo administrativo no âmbito da Administração Pública Federal.

BRASIL. Lei n.º 8.159, de 08 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados.

BRASIL. Arquivo Nacional. **Manual de identificação de acervos documentais para transferência e/ou recolhimento para arquivos públicos**. Rio de Janeiro: Arquivo Nacional, 1985. (Publicações 40)

BRASIL. Arquivo Nacional. **Manual de levantamento da produção documental**. Rio de Janeiro: Arquivo Nacional, 1985a. (Publicações 44)

BRESSER-PEREIRA, L.C. Estratégia nacional e desenvolvimento, **Revista de Economia Política**, v. 26, n. 2, p. 2023-230, abr.-jun. 2006.

BROOKS, P. C. The Selection of Records for Preservation. **American Archivist**, v. 3, n. 4, p. 221-234, 1940.

BROOKES, B. C. The foundations of information science: part I: Philosophical aspects. **Journal of Information Science**, v. 2, p. 125-133, 1980.

_____. The developing cognitive view in information science. **International Workshop on the Cognitive Viewpoint**, v. 77, p. 195-203, 1977.

BROUGHTON, V. Concepts and Terms in the Faceted Classification: the Case of UDC. **Knowledge Organization**, v. 37, n. 4, p. 270-279, 2010.

BUCKLAND, M. K. Documentation, Information Science and Library Science in the U.S.A. **Information Processing & Management**, v. 32, n. 1, p. 63-76, 1996.

_____. On the Nature of Records Management Theory. **The American Archivist**, v. 57, n. 2, p. 346-351, 1994.

_____. Information as thing. **Journal of the American Society for Information Science (JASIS)**, v.45, n.5, p.351-360, 1991.

BUDD, J. Organizing Acts and Objects: Metaphysical Foundations. **Knowledge Organization**, v. 41, n. 6, p. 419-428, jan. 2014.

BUREAU OF CANADIAN ARCHIVISTS. **Rules for archival description- RAD**. Ottawa, 1993.

CAMARGO, A. M. A. Sobre o valor histórico dos documentos. **Revista do Arquivo do Município de Rio Claro**, Rio Claro, n.1, p.11-17. 2003.

_____. Avaliação e destinação de documentos de arquivo; normas e procedimentos. (Relatório apresentado à ABNT, 2000)

CAMARGO, Ana Maria de Almeida & MACHADO, Helena Corrêa. **Como implantar arquivos municipais**. São Paulo: Arquivo do Estado, 2000. (Projeto Como Fazer, v.3)

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, v.14, n. 2, p. 221-241, set. 1985.

CARUCCI, P. **Il documento contemporaneo**: diplomatic e criteri di edizione. Roma: La Nuova Italia Scientifica, 1987. (Beni Culturali, 1).

CARVALHO FILHO, José dos Santos. **Manual de Direito Administrativo**. Rio de Janeiro: Lumen Juris, 2005.

CASTELLS, M. **A sociedade em rede**. Rio de Janeiro: Paz e Terra, 2009.

CAULLIRAUX, H. M.; CAMEIRA, R. Engenharia de processos de negócios: considerações metodológicas com vistas à análise e integração de processos. **Anais eletrônicos do III Simpósio de Administração da Produção**, São Paulo, 2000.

CENDÓN, B.V. Sistemas e redes de informação. In: OLIVEIRA, M. (Coord.). **Ciência da informação e biblioteconomia: novos conteúdos e espaços de atuação**. Belo Horizonte: UFMG, 2005. p.45-75.

CESARINO, M. A. N. B. Bibliotecas especializadas, centros de documentação, centro de análise da informação: apenas uma questão de terminologia? **Revista da Escola de Biblioteconomia da UFMG**, v. 7, n. 2, p. 218-241, 1978.

CHANTLAT, J. F. O gerencialismo e a ética do bem em comum: a questão da motivação para o trabalho nos serviços públicos. **VII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública**, Lisboa, Portugal, 8-11 Out. 2002 .

CHAUÍ, M. A universidade pública sob nova perspectiva. **Revista Brasileira de Educação**, n. 23, p.5-15, set.-dez. 2003. Disponível em: <<http://www.scielo.br/pdf/rbedu/n24/n24a02.pdf>>. Acesso em: 12 dez. 2018.

CHAUMIER, J. **Analyse et langages documentaires: le traitement linguistique de l'information documentaire**. Paris: Moderne d'Édition, 1982.

CHIERCHIA, G. **Semântica**. Campinas: Editora da Unicamp; Londrina: Eduel, 2003.

CHOMSKY, N. **Problemas do conhecimento e da liberdade**. São Paulo: Record, 2008.

_____. **Linguagem e mente**. Brasília: Editora UnB, 1999.

_____. Conhecimento da História e construção teórica na linguística moderna, **Delta**, v.13, special issue, São Paulo, 1997.

_____. **Language and problems of knowledge**. Cambridge: The MIT Press, 1988.

_____. **Studies on Semantics in Generative Grammar**. The Hague: Mouton & Co. N. V., Publishers. 1976.

_____. **Linguística Cartesiana: um capítulo da história do pensamento racionalista**. Trad. Francisco M. Guimarães. Petrópolis: Vozes; São Paulo: EdUSP, 1972.

_____. **Language and mind**. New York: Harcourt, Brace & World, Inc. 1968.

CLÉVE, C. M. LORENZETTO, B. M. **Governo democrático e jurisdição constitucional**. Belo Horizonte: Fórum, 2016.

CLEVERDON, C. W. Automation in indexing, **Aslib Proceedings**, v. 13, n. 4, p.107-109, 1961.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 4. **Journal of Documentation**, v. 14, n. 3, p. 136-143, 1958.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n.6. **Journal of Documentation**, v. 17, n. 2, p. 156-172, 1961.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 7. **Journal of Documentation**, v. 18, n. 2, p. 65-88, 1962.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 8. **Journal of Documentation**, v. 20, n. 3, p. 146-169, 1964.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 9. **Journal of Documentation**, v. 24, n. 4, p. 273-298, 1968.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n.10, . **Journal of Documentation**, v. 29, n. 1, p. 51-68, 1973.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 11. **Journal of Documentation**, v. 34, n. 1, p. 21-50, 1978.

CLASSIFICATION RESEARCH GROUP BULLETIN. Londres, n. 12. **Journal of Documentation**, v. 41, n. 2, p. 75-99, 1985.

COATES, E. J. Switching Languages For Indexing, **Journal of Documentation**, v.26, n.2, p. 102 - 110. 1970.

COBLANS, H. **Introdução ao estudo de documentação**. Rio de Janeiro: DASP, 1957. (Ensaio de Administração, n. 8).

CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO. 2.ed. São Paulo: Federação Brasileira de Associações de Bibliotecários, Cientistas da Informação e Instituições (FEBAB), 2004.

COHEN, M. C., MARCH, J. G. **Leadership and ambiguity: the American College President**. Boston: Harvard Business School Press, 1986.

COMTE, A. **Curso de filosofia positiva**. São Paulo: Nova Cultural, 1988.

CONDE VILLAVERDE, Maria Luisa. **Manual de tratamiento de archivos administrativos**. Madrid: Dirección de Archivos Estatales / Ministerio de Cultura, 1992.

CONSELHO INTERNACIONAL DE ARQUIVOS. **ISAD (G): norma geral internacional de descrição arquivística**. Rio de Janeiro: Arquivo Nacional, 2000.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). Considerações do Arquivo Nacional, órgão central do Sistema de Gestão de Documentos de Arquivo – SIGA, da Administração Pública Federal acerca do Sistema Eletrônico de Informações – SEI. Rio de Janeiro: CONARQ/GT SEI Avaliação, 2015.

CONSELHO NACIONAL DE ARQUIVOS (BRASIL). **Norma Brasileira de Descrição Arquivística**. Rio de Janeiro, 2005.

COOK, M.; PROCTER, M. **A manual of archival description**. Aldershot, Hants, England: GOWER, 1989.

COOK, T. What is past is prologue: a history of archival ideas since 1898, and the future paradigm shift. **Archivaria**, n. 43, p. 18-63, 1997.

CORDEIRO, R. I. N.; LA BARRE, K. Análise de facetas e obra filmica. **Informação e Informação**, Londrina, v. 16, n. 3, p. 180-201, jan./jun. 2011.

CORDEIRO, R. I. N. **Imagem e movimento**. Niterói: Eduff, 2000.

CORTÉS ALONSO, V. **Documentación y documentos**. Madrid: Anabad, 2011.

_____. **Archivos de España y América**: Materiales para un manual. Madrid: Fondo de Cultura Económica, 2008.

_____. **Manual de archivos municipales**. Madrid: Anabad, 1989.

COSTA FILHO, C. M. A.; SOUSA, R. T. B. Compreendendo o records continuum: contextualização, objetivos e reflexões. **Informação Arquivística**, v.6, n.2, p. 34-60, jul.dez. 2017.

Disponível em: <<http://www.aerj.org.br/ojs/index.php/informacaoarquivistica/article/view/208/114>>. Acesso em: 13 ago. 2018.

COYAUD, M. Introduction à l'étude des langages documentaires. Paris: Klincksieck, 1966.

CRETELLA JUNIOR, J. **Direito Administrativo**. Rio de Janeiro: Ed. Forense, 1998.

CRUZ MUNDET, J. R. **Manual de archivística**. Madrid: Fundación Germán Sanchez Ruipérez/Pirámide, 1994.

CUNNINGHAM, A.; MILLAR, L.; REED, B. Peter J. Scott and the Australian 'series' system: its origins, features, rationale, impact and continuing relevance. **Comma**, Liverpool, n.1, 2013. Disponível em: <<https://doi.org/10.3828/comma.2013.1.13>>. Acesso em: 11 ago. 2018.

DAHLBERG, I. Teoria do Conceito. **Ciência da Informação**, v. 7, n. 2, p. 122-131, 1978.

_____. Knowledge Organization. **Journal of Documentation**, v. 20, n. 4, p. 211-222, 1993.

DAY, R. Poststructuralism and Information Studies. **Annual Review of Information Science and Technology**, Nova Iorque, p. 575-609, 2005.

DELMAS, B. **Arquivos para quê?** Textos escolhidos. São Paulo: Instituto Fernando Henrique Cardoso, 2010.

_____. Origine et développement de l'enseignement de l'archivistique. **Archivium**, p. 61-73, 1988.

DRUCK, G. Precarização social do trabalho. In: IVO, A. *et. al.* (Orgs.). **Dicionário temático desenvolvimento e questão social**: 81 problemáticas contemporâneas. Rio de Janeiro: Annablume, 2013, v. 1, p. 373-381.

DUARTE, J. **Comunicação pública**: estado, mercado, sociedade e interesse público. São Paulo: Atlas, 2009.

DUARTE, L. F. D. Três ensaios sobre pessoa e modernidade. **Boletim do Museu Nacional**, Rio de Janeiro, n° 41, p. 1-69, ago. 1983.

DUCHARME, J.; ROUSSEAU, J. Y. L'interdépendance des archives et de la gestion des documents: un approche globale de l'Archivistique. **Archives**, v. 12. n. 1, p. 5-28, jun. 1980.

DUCHEIN, M. O respeito aos fundos em Arquivística: princípios teóricos e problemas práticos. **Arquivo & Administração**. Rio de Janeiro, v. 10-14, n. 1, p. 14- 33, abr. 1982/ago. 1986.

DURANTI, L. **I documenti archivisti: la gestione dell'archivio da parte dell'ente produttore**. Roma: Ministero per i beni culturali e ambientali, Ufficio centrale per i beni archivistici, 1997.

_____. The archival bond. **Archives and Museum Informatics**, v. 11, n.3-4, p. 213-218, set.1997a.

_____. Registros documentais contemporâneos como prova de ação. **Revista Estudos Históricos**, Rio de Janeiro, v. 7. n. 13, p. 50-64, 1994.

DURANTI, L., EASTWOOD, T.; MACNEIL, H. **The preservation of the Integrity of electronic records: UBC project report**. Disponível em: <<http://www.interpares.org/UBCProject/intro.htm#INTRODUCTION>>. Acesso em: 04 jan. de 2018.

ERICSON, T. L. Professional associations and archival education: a different role, or a different theater?. **American Archivist**, Chicago, v.51, n.3, p.298-311, 1988.

ESTEBAN NAVARRO, M.A.E. Elementos, actividades y critérios para la identificación, comprensión y selección de conceptos en la indización analítica. In: GARCIA MARCO, F.J.G.M. **Organización del conocimiento en sistemas de información y documentación**. Zaragoza: Universidad Carlos III de Madrid, 1999. p. 69-93.

FARRADANE, L.E.L. A scientific theory of classification and indexing: further considerations. **Journal of Documentation**, v. 8, n. 2, p. 73 - 92, 1952.

FARROW, J.F. A cognitive process model of document indexing, **Journal of Documentation**, v. 47, n. 2, p. 149-166, 1991.

FEIBLEMAN, J. K. Theory of integrative levels. **British British Journal for the Philosophy of Science**, p. 59-66, 1954. Disponível em: <<http://dx.doi.org/10.1093/bjps/V.17.59>>. Acesso em: 20 fev. 2019.

FERNÁNDEZ SAGREDO, F.; IZQUIERDO ARROYO, J. M. **Concepción lógico-lingüística de la documentación**. Madri: Ibercom, 1983.

FISH, S. Is there a text in this class? **Alfa** – Revista de Linguística, São Paulo, v. 36, p. 207-212, 1992.

_____. Interpreting the Variorum. **Critical Inquiry**, v. 2, n.3, p. 465-485, 1976.

FODOR, J. Semantics: an interview with Jerry Fodor. **Revel**, v. 5, n. 8, 2007. Disponível em: http://www.revel.inf.br/files/entrevistas/revel_8_interview_jerry_fodor.pdf. Acessado em: 08 de janeiro de 2018.

FOGL, J. Relations of the concepts 'information' and 'knowledge'. **The Hague**, v. 4, n. 2, p.21-24, 1979.

FOSCARINI, F. **Function-based records classification systems: an exploratory study of records management practices in central banks**. Vancouver: University British Columbia, 2009.

FOSKETT, D. J. **Abordagem temática da informação**. São Paulo: Polígono; Brasília: Ed. UnB, 1973.

_____. Classification and indexing in the Social Sciences, **Aslib Proceedings**, v. 22, n. 3, p.90-10, 1970.

_____. Two notes on indexing techniques. **Journal of Documentation**, v. 18, n. 2, p. 188-192, 1962.

FOUCAULT, M. **A palavra e as coisas**. São Paulo: Martins Fontes, 2007.

_____. **Em defesa da sociedade**. São Paulo: Martins Fontes, 2005.

_____. **Microfísica do poder**. Rio de Janeiro: Jorge Zahar, 2001.

_____. **A verdade e as formas jurídicas**. Rio de Janeiro: Nau Editora, 2001a.

_____. **A arqueologia do saber**. Rio de Janeiro: Forense Universitária, 1998.

_____. **A ordem do discurso**. São Paulo: Edições Loyola, 1996.

FREGÉ, G. Sobre sentido e a referência. In: **Lógica e filosofia da linguagem**. Trad. Paulo Alcoforado. São Paulo: Cultrix/USP, 1978, pp. 59-86.

FREITAG, B. **Escola, estado e sociedade**. São Paulo: Editora Moraes, 1986.

FREIXO, L.; TOUTAIN, L. B. (Org.). **Normas para a gestão de documentos arquivísticos**. Salvador: UFBA, 2014.

FREIXO, M. J. V., **Teorias e Modelos de Comunicação**. Lisboa: Instituto Piaget, 2006.

FROHMANN, B. Documentation Redux: prolegomenon to (another) philosophy of information. **Library Trends**, v. 52, n.3, p. 387-407, 2004.

_____. Discourse and documentation: some implications for pedagogy and research. **Journal of Education for Library and Information Science**, v.42, n.1, p. 13-26, 2001.

_____. FROHMANN, B. The ethics of Information Science theory. In: ANNUAL MEETING OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, 55., 1992, Pittsburgh. **ASIS...** Pittsburgh: JASIS, 1992. Disponível em: < <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.471.3667&rep=rep1&type=pdf>>. Acesso em: 27 out. 2018.

_____. Rules of indexing: a critique os metalism in information retrieval theory. **Journal of Documentation**, v. 46, n. 2, p. 81-101, 1990.

FUGMAN, R. The five axiom theory of indexing and information suply. **Journal of American Society of Information Science**, v. 36, n. 2, p. 116-129, 1985.

FUJITA, M S. L. A importância teórica e prática da indexação na fundamentação científica da organização e representação do conhecimento. **Complexidade e organização do conhecimento: desafios de nosso século**. DOBEDEI, V.; GUIMARÃES, J. A. C. (Org.) Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, p. 147-159, 2013.

_____. A política de indexação para representação e recuperação da informação. In: LEIVA, I. G.; FUJITA, M. S. L. (Ed.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 17-31.

_____. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **Datagrama zero – Revista de Ciência da Informação**, v. 5, n. 4, 2004. Disponível em: <http://www.dgz.org.br/ago04/F_I_art.htm> Acesso em: 08 de janeiro de 2018.

_____. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 1, n. 1, p. 60-90, 2003.

_____. Avaliação da eficácia de recuperação do sistema de indexação PRECIS. **Ciência da Informação**, v. 18, n. 2, p. 120-134, jul./dez. 1989.

GARCÍA GUTIÉRREZ, A. L. **Principios de lenguaje epistemográfico: la representación del conocimiento sobre patrimonio histórico andaluz**. Sevilla: Instituto Andaluz del Patrimonio Histórico, 1998.

GARDIN, J. C. Éléments d'un modèle pour la description des lexiques documentaires. **Bulletin des Bibliothèques de France (BBF)**, n. 5, p. 171-182, 1966.

_____. Analyse documentaire et théorie linguistique. In: **Les analyses de discours**. Neuchâtel: Delachaux et Niestlé, p. 120-168, 1974.

GERGEN, K. J. **An invitation to social construction**. Londres: Sage, 1999.

GIASSON, J. **A compreensão da leitura**. Porto: Edições ASA, 1993.

GIDDENS, A. Risco, confiança e reflexividade. In: U. Beck, A. Giddens, & S. Lash, **Modernização reflexiva**. São Paulo: Unesp, 1997.

GIL LEIVA, I. Aspectos conceituais da indexação. In: GIL LEIVA, I.; FUJITA, M. A. S. L. (eds.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012.

_____. **Manual de indiziación: teoría y práctica**. Gijón: Ediciones Trea, 2008.

GRAESSER, A. C.; ZWAAN, R. A. Inference generation and the construction of situation models. In C. A. Weaver III; S. Mannes & C. R. Fletcher (Eds.), **Discourse comprehension: essays in honor of Walter Kintsch**. Hillsdale: Lawrence Erlbaum, 1995, p.117-127.

GROSS, T. Naming and Reframing: a taxonomy of attacks on Knowledge Organization, **Knowledge Organization**, v. 42, n. 5, p. 263-268, 2015.

GROVER, V; KETTINGER, W. R. **Process think: winning perspectives for business process change in information ange**. Hershey: Idea Group Inc, 2000.

GUEDES, V. L. S. **Nominalizações deverbais em artigos científicos: uma contribuição para a análise e a indexação temática da informação**. 2010. 109f. Tese (Doutorado em Linguística) – Centro de Letras e Artes, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

GUIMARÃES, J. A. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação (RICI)**, v.1 n.1, p.77-99, jan./jun. 2008.

_____. As políticas de indexação como elemento para a gestão do conhecimento nas organizações. In: VIDOTTI, S. A. G. (Coord.) **Tecnologia e conteúdos informacionais: abordagens teóricas e práticas**. São Paulo: Polis, 2004. p.43- 52.

_____. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais”. In: RODRIGUES, G. M.; LOPES, I. L. (orgs.). **Organização e representação do conhecimento na perspectiva da ciência da informação**. Brasília: Thesaurus, p. 100-117, 2003.

HARVIAINEN, T.; HAMARI, J. Seek, share, or withhold: information trading in MMORPGs", **Journal of Documentation** , v. 71, n. 6, p.1119-1134, 2015.

HEMAIS, B.; BIASI-RODRIGUES, B. A proposta sócio-retórica de John M. Swales para o estudo de gêneros textuais. In: MEURER, J. L.; BONINI, A.; MOTTA-ROTH, D. (Org.). **Gêneros: teorias, métodos, debates**. São Paulo: Parábola, 2005. p. 108-129

HEREDIA HERRERA, A. **Manual de Archivística Básica: Gestión y Sistemas**. Madrid: Archivo Histórico Buap, 1991.

HERRE, H. Formal Ontology and the Foundation of Knowledge Organization. **Journal of Documentation**, v. 40, n. 1, p. 332-339, 2013.

HIDER, P.A survey of the coverage and methodologics of schemes and vocabularies use to describe information resources. **Journal of Documentation**, v. 42, n. 3, p. 154-163, 2015.

HJORLAND, B. Domain analysis in information science: Eleven approaches – traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422-462, 2002.

_____. The classification of psychology: a case study in the classification of a knowledge field. **Knowledge Organization**, v. 24, n. 4, p. 162-201, 1998.

_____. The concept of ‘subject’ in information science. **Journal of Documentation**, v. 48, n. 2, p.172-200, 1992.

HJORLAND, B; ALBRECHTSEN, H. An analysis of some trends in classification research. **Knowledge Organization**, v. 26, n. 3, p. 131-139, 1999.

_____. Toward a New Horizon in Information Science: Domain- Analysis. **Journal of American Society for Information Science**, v. 46, n. 6, p. 400-425, 1995.

HOWARTH, L. C. Clare Beghtol: exploring new approaches to the Organization of Knowledge. **Knowledge Organization**, v. 37, n. 2, p. 95-100, 2010.

IBEKWE-SANJUAN, F. Constructing and maintaining knowledge organization tools: a symbolic approach, **Journal of Documentation** , v. 62, n. 2,p. 229-250 , 2006.

INGWERSEN, P. **Retrieval interection**. Londres: Taylor Graham Publishing, 1992.

_____. Search procedures in the library—analysed from the cognitive point of view, **Journal of Documentation** , v. 38, n. 3, p.165-191, 1982.

INDOLFO, M.C. Gestão de documentos: uma renovação epistemológica no universo da arquivologia. **Arquivística.net**, Rio de Janeiro, v. 3, n. 2, p. 28-60, jul./dez. 2007.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 15489**: Information and documentation. Records management. Geneva, 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **UNE-ISO 30300**: Información y documentación. Sistemas de gestión para los documentos. Madrid: Aenor, 2011.

IZQUIERDO ARROYO, J. M. **Esquemas de lingüística documental**. Barcelona: Promociones y Publicaciones Universitarias, 1990. Tomo I,II e III.

JACOB, E. K.; SHAW, D. Sociocognitive perspectives on representation. **Annual Review of Information Science and technology**, v. 33, p. 131-185, 1998.

JAKOBSON, R. **Ensayos de lingüística general**. Madrid: Ariel, 1974.

JARDIM, J. M. Caminhos e perspectivas da gestão de documentos em cenários de transformações. **Acervo**, Rio de Janeiro, v. 28, n. 2, p. 19-50, jul./dez. 2015.

_____. **Sistemas e Políticas Públicas de Arquivos no Brasil**. Niterói: EDUFF, 1995.

_____. O conceito e a prática de gestão de documentos. **Acervo**, Rio de Janeiro, v. 2, n. 2, p. 35-42, jul./dez. 1987.

JENKINSON, H. **A manual of archive administration**. Londres: P. Lund, Humpries & co., 1937.

JONKER, F. **Indexing theory, indexing methods and search devices**. Nova Iorque: The Scarecrow Press, 1964.

JOUDREY, D. N. **Building puzzles and growing pearls: a qualitative exploration of determining aboutness**. 2005. 475p. Phd Thesis (School of Information Sciences), University of Pittsburgh, Pittsburgh.

KAISER, J. **Systematic Indexing**. Londres: Pitman, 1911.

KETELAAR, E. "Control though communication" in a comparative perspective. **Archivaria**, v. 60, p. 71-89, 2005.

KINTSCH, W. **Comprehension: paradigm for cognition**. Cambridge: Cambridge University Press, 1998.

KRESS, G. **Linguistic progress in socialcultural practice**. Oxford: Oxford University Press, 1989.

KYLE, B. Information Retrieval and subject indexing: Cranfield and after, **Journal of Documentation**, v. 20, n. 2, p. 55-69, 1964.

KWASNIK, B. H. Semantic warrant: a pivotal concept for our field. **Knowledge Organization**, v. 37, n. 2, p. 106-110, 2010.

LA BARRE, K. Facet analysis. **ARIST**, v. 44, p. 243-284, 2010.

LA TORRE MERINO , J. L.; MARTÍN-PALOMINO Y BENITO, M. **Metodología para la identificación y valoración de fondos documentales**. Madrid: Ministerio de Educación, Cultura y Deportes/S.G. de Información y Publicaciones, 2000.

LAKOFF, G.; JOHNSON, M. **Metáforas da vida cotidiana**. São Paulo: Mercado das Letras, 2002.

- LANCASTER, F. W. **Indexação e Resumos: teoria e prática**. Brasília: Brinquet de Lemos, 2003.
- LANDRY, B.C. **A theory of indexing: indexing theory as a model for information storage and retrieval**. Columbus: Ohio State University, 1971.
- LANGEDOEN, D.T; POSTAL, P. **The vastness of natural languages**. Oxford: Basil Blackwell, 1984.
- LANGRIDGE, D. **A classificação: abordagem para os estudantes de Biblioteconomia**. Rio de Janeiro: Interciência, 1977.
- LARA, M. L. G. Conceitos linguísticos fundamentais para organização e disseminação de informações. **Anais do V Encontro Enancib**, 2003. Disponível em: <http://enancib.ibict.br/index.php/enancib/venancib/paper/viewFile/1903/1044>. Acessado em 08 de dezembro de 2017.
- _____. Conceitos de organização e representação do conhecimento na ótica das reflexões do grupo Temma. **Informação & Informação**, Londrina, v. 16, n. esp., p. 92-121, jan./jun. 2011.
- _____. Informação, informatividade e linguística documentária: paralelos com as reflexões de Hjørland e Capurro. **Anais do VIII Enancib**, Salvador, outubro de 2007 Disponível em: <http://enancib.ibict.br/index.php/enancib/viiienancib/paper/view/2822/1950>. Acessado em 8 de janeiro de 2018.
- LE GOFF, J. Na Idade Média: tempo da Igreja e o tempo do mercador. In: **Por um conceito novo de Idade Média**. Lisboa: Editorial Estampa, 1995, p. 43-73.
- LEE, J. M. Wydham Hulme: a reconsideration. In: RAYWARD, W. B. (ed.). **The variety of librarianship**. Sidney: Library Association of Australia, p. 101-113, 1976.
- LIMA, K. O Banco Mundial e a educação superior brasileira na primeira década do novo século. **Revista Katálysis**, Florianópolis, v.14, n. 1, p.86-94, jan./jun. 2011.
- LLANSÓ I SANJUAN, J. Sistemas archivísticos y modelos de gestión de documentos en el ámbito internacional. **Revista Códice**, v. 2, p. 53-88, enero-junio 2006.
- _____. **Gestión de documentos: definición y analisis de modelos**. Bergara: IRARGI, Centro de Patrimonio Documental de Euskadi, 1993.
- LODOLINI, E. **Archivística**. Principios y problemas. Madrid: Anabad, 1993.
- LOPES, L. C. **A informação e os arquivos: teorias e práticas**. Niterói: Eduff, 1996.
- LOPEZ, A.P. A.; ÁVILA, R. F.; MELLO, R. G.; SANTOS, T. R. Blogs como ferramenta de ensino-aprendizagem de Diplomática e Tipologia Documental: uma estratégia didática para construção de conhecimento. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v.1, número especial, p. 86-99, out. 2011.
- LÓPEZ YEPES, J. **Linguística documental**. Aplicación a la comunicación social. Barcelona: Mitre, 1984.
- LÖWY, M. **Ideologias e Ciências Sociais: elementos para uma análise marxista**. São Paulo: Cortez, 2003.

- LYNCH, M. F. Subject indexes and automatic document retrieval: the structure of entries in Chemical Abstracts subject indexes. **Journal of Documentation**, v. 22, n.3, p.167-185, 1966.
- MAI, J.E. Analysis in indexing: document and domain centered approaches. **Information Processing & Management**, v. 41, n. 3, p. 599-611, 2005.
- MAIMONE, G. D.; SILVEIRA, N. C.; TÁLAMO, M. G. M. Reflexões acerca das representações temática e descritiva. **Inf. & Soc.:Est.**, João Pessoa, v.21, n.1, p. 27-35, jan./abr. 2011.
- MARCHESE, C.; SMIRAGLIA, R. P. Boundary Objects: CWA, an HR Firm, and Emergent Vocabulary. **Knowledge Organization**, v. 40, n. 3.p. 254-259, 2013.
- MARCONDES, D. **Filosofia, linguagem e comunicação**. São Paulo: Cortez, 1992.
- MARSDEN, P. When is the Future? Comparative Notes on the Electronic Record-Keeping Projects of the University of Pittsburgh and the University of British Columbia. **Archivaria**, v. 43, p.158-173, 1997.
- MARTIN, D. Use of Cataloging Techniques in Work With Records and Manuscripts. **The American Archivist**, v. 18, n. 4, p. 317-366, 1955.
- MARTÍN-PONZUELLO CAMPILLOS, M. P. **La construcción teórica em archivística: el principio de procedencia**. Madrid: Universidad Carlos III y BOE, 1996.
- MARTÍNEZ-ÁVILA, D.; SAN SEGUNDO, R. Reader-Interest Classification: concept and terminology historical overview. **Knowledge Organization**, v. 40, n. 2, p. 102-114, 2013.
- MARTINHO, N. O. **A dimensão teórica e metodológica da catalogação de assunto**. 2010. 189f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2010.
- MARTINS, H. Três caminhos na filosofia da linguagem. IN: MUSSALIM, F; BENTES, A.C. (orgs) **Introdução à linguística: fundamentos epistemológicos**. v. 3.São Paulo: Cortez, 2009. p. 439-474.
- MAUPERON, A. Adaptation to users' needs. **Aslib Proceedings**, v. 19, n. 7, p.232-240, 1967.
- MAUSS, M. Ensaio sobre a dádiva. In: _____. **Sociologia e antropologia**.Forma e razão da troca nas sociedades arcaicas. Trad. Paulo Neves. São Paulo: Cosac Naif, 2003.
- MCKEMMISH, S. Traces: Document, Record, Archive, Archives. In: MCKEMMISH, S. *et al.* (Ed.). **Archives: recordkeeping in society**. Wagga wagga: Centre of Information Studies, Charles Sturt University, 2005.
- _____. Placing records continuum theory and practice. **Archive Science**, v. 1, p. 333-359, 2001.
- _____. Yesterday, today, tomorrow. In: **Proceedings of the records management association of australia national convention**, v. 14, 1997, Perth: Records Management Association of Australia, 1997.
- MEIRELLES, H. L. **Direito administrativo brasileiro**. São Paulo: Malheiros, 2014.
- MENDO CARMONA, C. Consideraciones sobre el método en archivística. **Documenta e Instrumenta**. Madrid, v. I, 2004, p. 35-46. Disponível em: www.ucm.es/info/document/htm/default.htm. Acesso em: 02 de fevereiro de 2017.

_____. El largo camino de la Archivística: de práctica a ciência. **Revista de Historia de la Cultura Escrita**, Madrid, n. 2 p. 113-132, 1995.

MENNE-HARITZ, A. An archival system with old traditions in a time of change. **Archive Science**, v. 3, p. 321-327, 2003.

MELLO, C. A. B. **Curso de Direito Administrativo**. São Paulo: Malheiros, 2004.

MEDAUAR, O. **A processualidade no direito administrativo**. São Paulo: Revista dos Tribunais, 2008.

MERKL, A. **Teoria Geral do Direito Administrativo**. Granada: Comares, 2008.

MEY, E. **Introdução à catalogação**. Brasília: Brinquet de Lemos, 1987.

MILLER, T. The German registry: the evolution of a recordkeeping model. **Archivaria**, v. 3, p. 43-63, 2003.

MOSS, R. Categories and relations: Origins of two classification theories. **Journal of the Association for Information Science and Technology**, v. 15, n. 4, p. 296-301.

NEGREIROS, L. R. **Sistemas eletrônicos de gerenciamento de documentos arquivísticos: um questionário para seleção, aplicação e avaliação**. Belo Horizonte, 2007. 246 f. Tese (Doutorado) - Curso de Pós-Graduação em Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

NIETZSCHE, F. **Humano, demasiadamente humano**. São Paulo: Cia. das Letras, 2010.

_____. Sobre verdade e mentira no sentido extramoral. In: NIETZSCHE, F. **Obras Incompletas**. Trad. Rubens R. Torres Filho. São Paulo: Abril Cultural, 1999.

OLIVEIRA, R. P. A semântica formal. IN: MUSSALIM, F; BENTES, A.C. (orgs) **Introdução à linguística: domínios e fronteiras**. São Paulo: Cortez, v.2, 2009.

OLIVEIRA, L. M. V. **Modelagem e status científico da descrição arquivística nos arquivos pessoais**. 2010. Tese (Doutorado em História Social) – Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de História, Universidade de São Paulo, São Paulo, 2010.

ORTEGA, C. D. A noção de documento: de Otlet aos dias de hoje. **Datagramazero**, v. 11, n. 2, 2014. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000008400/cca9a49474077340b069f1222c313618>. Acessado em: 8 de janeiro de 2018.

PAULA, A. P. Administração Pública Brasileira entre o gerencialismo e a gestão social. **Revista de Administração de Empresas**, v. 45, n. 1, p. 36-49, jan.-mar. 2005.

PEIXOTO, J. P. **A revolução cultural e científica dos séculos XVII e XVIII e a gênese das academias**. Lisboa: A.C.C, 1997.

PIAGET, J. **A epistemologia genética**. Petrópolis: Vozes, 1985.

PIEIDADE, M. A. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1977.

PILLATTI, L.A.; CECHIN, M. R. Perfil das universidades brasileiras de e com potencial de classe mundial. **Avaliação: Revista da Avaliação do Ensino Superior**, Campinas, v. 23, n.1, p. 75-103, 2018.

PINKER, S. **Como a mente funciona**. Trad. Laura Teixeira Motta. São Paulo: Companhia das Letras, 1997.

PINTO MOLINA, M. Interdisciplinary approaches to the concept and practice of written text's documentary content analysis (WTDCA). **Journal of Documentation**, v. 50, n. 2, p. 111-133, 1994.

PLATÃO, **Mênnon**. Trad. Maura Iglésias. São Paulo: Ed. Loyola, 2007.

_____, **A República**. Trad. Maria Helena da Rocha Pereira. Lisboa: Fundação Calouste Gulbenkian, 2001.

_____, **Crátilo**. Trad. Carlos Alberto Nunes. Belém: Editora da Universidade Federal do Pará, 1988.

POMIAN, K. Coleção. **Enciclopédia Einaudi**. Porto: Imprensa Nacional/Casa da Moeda, 1984, p. 51-86.

POSNER, E. **Alguns aspectos do desenvolvimento arquivístico a partir da Revolução Francesa**. Rio de Janeiro: Arquivo Nacional, 1959.

RAFFERTY, P. Knowledge Representation in the Social Semantic Web. **Journal of Documentation**, v. 67, n.5, p. 896-899, 2011.

RANGANATHAN, S. R. **Colon Classification**. Bangalore: Sarada Ranganathan Endowment for Library Science, 1987.

_____, S. R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.

_____. Subject heading and facet analysis, **Journal of Documentation**, v. 20, n. 3 p. 109 – 119, 1964.

RAYWARD, W. Boyd. The origins of Information Science and the International Institute of Bibliography/ International Federation for Information and Documentation (FID). **Journal of the American Society for Information Science**, v. 48, n. 4, p. 289-300, 1997.

REED, B. Metadata: Core Record or Core Business? **Archives and Manuscript**, Londres, v.25, n. 2, p. 218-241, 1997.

RIBEIRO, C. V. dos S.; LEDA, D. B.; SILVA, E. P. e. A expansão da educação superior pública e suas implicações no trabalho docente. **Revista Educação em Questão**, Natal, v. 51, n. 37, p. 147-174, jan./abr. 2016.

RIOS, E. R.; CORDEIRO, R. I. N. Plano de Classificação de documentos arquivísticos e a teoria da classificação: uma interlocução entre domínios do conhecimento. **Perspectivas em Ciência da Informação**, v. 15, n. 2, p. 133-139, maio/ago 2010.

RHOADS, J. **The role of archives and records management in national information systems: a RAMP study**. Paris: General Information Programme and UNISIST, 1989.

ROBSON, A.; ROBINSON, L. The Information Seeking and Communication Model: A study of its practical application in healthcare, **Journal of Documentation**, v. 71, n. 5, p. 1043-1069, 2013.

ROCHA, C. L. A. Princípios constitucionais do processo administrativo no direito brasileiro. **Revista de Direito Administrativo**, Rio de Janeiro, v. 209, p. 189-222, jul./set. 1997.

RODRIGUES, A. C.; GARCIA, N. D. P. A FAMS e a política de gestão de documentos para a Prefeitura Municipal de Santos: estudo de caso sobre a elaboração da Tabela de Temporalidade de Documentos da Secretaria Municipal de Economia e Finanças (TTD SEFIN). In: XVIII Congresso Brasileiro de Arquivologia. Rio de Janeiro, RJ: Associação dos Arquivistas Brasileiros, 2012.

RODRIGUES, A. C. Identificação: uma metodologia de pesquisa para a Arquivística. In: **Estudos avançados em Arquivologia**. Marília, São Paulo: Oficina Universitária; Cultura Acadêmica, 2012a, p. 197-215. Disponível em: http://www.marilia.unesp.br/Home/Publicacoes/estudos_avancados_arquivologia.pdf. Acesso em: 10 fev. 2017.

_____. **Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos**. 2008. 258f. Tese (Doutorado em História Social) – Faculdade de Filosofia, Letras e Ciência Humanas da Universidade de São Paulo (USP), São Paulo.

RONDINELLI, R. C. **Gerenciamento arquivístico de documentos eletrônicos**: uma abordagem teórica da diplomática arquivística contemporânea. Rio de Janeiro: Editora FGV, 2005.

ROPER, M; MILLAR, L (eds). **The management of public sector records**: principles and context. Londres: International Records Management Trust, 1999.

ROSA, C. D. T. **Manual de padronização dos processos administrativos da UFSM**: divisão protocolo. Santa Maria: UFSM/PROAD, 2014.

ROSA, J. G. **Sagarana**. Rio de Janeiro: Nova Fronteira, 2001.

ROUSSEAU, J. ; COUTURE, C. **Os fundamentos da disciplina arquivística**. Lisboa: Publicações Dom Quixote, 1998.

RUBI, M. P. Política de Indexação. In: LEIVA, I. G.; FUJITA, M. S. L. (Eds.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 107-120.

RUMMLER, G. A.; BRACHE, A. P. **Melhores desempenhos das empresas**: ferramentas para a melhoria da qualidade e da competitividade. São Paulo: Makron Books do Brasil, 1992.

RUSSELL. B. **The Philosophy of Logical Atomism and other essays**: 1914-1919. Londres: John Slater, Allen & Unwin, 1986.

SABOURIN, P. Constructing a function-based classification system: business activity structure classification system. **Archivaria**, 51, 137-154, 2001.

SALERNO, M.S. **Projetos de organizações integradas e flexíveis**: processos, grupos e gestão democrática via espaços de comunicação-negociação. São Paulo: Atlas, 1999.

SALTON, G. A theory of indexing. **Regional Conference Series in Applied Mathematics**, Society for Industrial and Applied Mathematics, Filadélfia, 1975.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

_____. Tecnologia da informação, sistemas de informação e informação como utilidade pública. **Ciência da Informação**, Rio de Janeiro, v. 3, n. 1, p. 57-67, 1974.

SATIJA, M. P.; MADALLI, D. P.; DUTTA, B. Modes of Growth of Subjects. **Knowledge Organization**, v. 41, n. 3, p. 195-204, 2014.

SCHELLENBERG, T. R. **Arquivos modernos**: princípios e técnicas. Rio de Janeiro: FGV, 2006.

SCHMIDT, C.M.S. **Arquivologia e a construção do seu objeto científico**: concepções, trajetórias, contextualizações. 2012. 320f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes da Universidade de São Paulo (USP), São Paulo.

SEARLE, J. R. *Expression and meaning*. Cambridge: Cambridge University Press, 1979.

SEI GARANTIRÁ fim de processos em papel na UFF. **UFF notícias**, Niterói, 28 mar. 2017. Disponível em <www.uff.br/?q=noticias/28-03-2017/sei-garantira-fim-de-processos-em...uff>. Acesso em: 11 dez. 2018.

SERAFIM, M. O processo de mercantilização das instituições de educação superior: um panorama do debate nos EUA, na Europa e na América Latina. **Avaliação**, Campinas; Sorocaba, v. 16, n. 2, p. 241-265, 2011.

SLAUGHTER, S; LESLIE, L. **Academic capitalism**: politics, policies and the entrepreneurial university. Chicago: John Hopkins UP, 1997.

SMIRAGLIA, R. P.; HEUVEL, C. *Classifications and concepts: towards an elementary theory of knowledge interaction*, **Journal of Documentation**, v. 69, n. 3, p. 360-383, 2013.

SILVA, M. R.; FUJITA, M. S. L.A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, 2004, v.16, n.2, p.133-161.

SODRÉ, F.; LITTIKE, D.; DRAGO, L. M. B.; PERIM, M. C. M. Empresa Brasileira de Serviços Hospitalares: um novo modelo de gestão? **Serviço Social & Sociedade**, São Paulo, n. 114, p. 365-380, abr./jun. 2013.

SOERGEL, D. **Organization information**: principles of data base and retrieval system. Nova Iorque: Academic Press, 1985.

SOUSA, R. T. Alguns apontamentos sobre a classificação do documento de arquivo. **Brazilian Journal of Information Science**: research trends, v. 8, n. 1-2, 2014. Disponível em: <http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/4246/3085>. Acessado em 20 de outubro de 2017.

_____. **Classificação em Arquivística**: trajetória e apropriação de um conceito. 2005. 245f. Tese (Doutorado em História Social) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (USP), São Paulo.

_____. Os princípios arquivísticos e o conceito de classificação. **Organização e Representação do Conhecimento** – RODRIGUES, G. M. e LOPES, I. L. (Org.). Brasília: Thesaurus, 2003. p.240-271.

SOUZA, R. R. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectivas em Ciência da Informação**, v.11, n.2, p.161-173, 2006.

SOUZA, R. R.; TUDHOPE, D.; ALMEIDA, M. B. Towards a taxonomy of KOS: dimensions for classifying knowledge organization systems. **Knowledge organization**, v. 39, p. 179–192, 2012.

SPITERI, L. A simplified model for facet analysis: Ranganathan 101. **Canadian Journal of Information and Library Science**. v. 23, p. 1-30, 1998.

STEWART, S. **On longing**: narratives of the miniature, the gigantic, the souvenir, the collection. Chicago: The John Hoppinks University Press, 1984.

SUNDFELD, C. A. A importância do procedimento administrativo. **Revista de Direito Público**, São Paulo, n.º 4, p. 64-74, 2007.

SWALES, J. M. Textography: toward a contextualization of written academic discourse. **Research on Language and Social Interaction**, v. 31, n.1, p. 109-121, jan. 1998.

_____. **Genre analysis**: English in academic and research settings. Cambridge :Cambridge University Press, 1990.

SWASON, D. R. Information Retrieval as a Trial-and-Error Process. **Library Quarterly**, v. 47, n. 2, p. 128–148, Abr. 1977.

TALJIA, S; TUOMINEN, K; SAVOLAINEN, R. “Isms” in information science: constructivism, collectivism and constructionism. **Journal of Documentation**, v. 61, n. 1, p. 79-101, 2005.

TAUILE, J. R. Introdução à economia política da informação. **Ensaio FEE**, Porto Alegre, v. 2, n. 2, p. 89-108, 1981.

TAYLOR, H. A life in archives: retrospect and prospect. **Archives and Manuscripts**. n. 21, p. 222-236, nov.1993.

_____. Information ecology and the archives of the 1980's. **Archivaria**. n. 18, p. 25-37, 1984.

TENNIS, J. T. Conceptual Provenance in Indexing Languages. In: LEMIEUX V. (Ed.) **Building Trust in Information**: proceedings in Business and Economics. New York: Springer, 2016.

_____. The Weight of Language and Action: Epistemic Stance, Methodological Action, and Theoretical Perspective in Knowledge Organization. **Études de Communication**: organisation des connaissances - épistémologie, approches théoriques e méthodologiques, v. 39, p. 15-40, 2012.

_____. Epistemology, theory, and methodology in Knowledge Organization: toward a classification, metatheory, and research framework. **Knowledge Organization**, v.35, n.2/3, p. 102-112, 2008.

_____. Social tagging and the next steps for indexing. **17th Annual ASIS&T SIG/CR Classification Research Workshop**, p. 1-15, 2006.

_____. Two axes of domains for domain analysis. **Knowledge Organization**, v.30, n.3/4, p. 191-195, 2003.

TESSITORE, V. **Arranjo**: estrutura ou função? São Paulo: Arquivo do Estado de São Paulo, 1989.

THOMPSON, E.P. Tempo, disciplina de trabalho e capitalismo industrial. In: **Costumes em comum**. São Paulo: Cia das Letras, 2011.

TRINDADE, H. Universidade em perspectiva: sociedade, conhecimento e poder. **Revista Brasileira de Educação**, n. 10, p.5-15, jan.-abr. 1999. Disponível: <http://anped.tempsite.ws/novo_portal/rbe/rbedigital/RBDE10/RBDE10_03_HELGIO_TRINDAD E.pdf>. Acesso em: 10 dez. 2018.

UNIVERSIDADE DE BRASÍLIA. **Guia prático do SEI na UnB**. Brasília: UnB, 2016.

_____. **Manual de gestão de documentos da Universidade de Brasília**. Brasília: Cebraspe, 2015.

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO. **Manual de procedimentos e normas de protocolo e arquivo**. Vitória: UFES/SERPROG, 2014.

UNIVERSIDADE FEDERAL FLUMINENSE. **Programa de Expansão e Reestruturação da UFF (2006 a 2012)**. Niterói: UFF, 2010. (Caderno Técnico)

UNIVERSIDADE FEDERAL FLUMINENSE. **Censo UFF 2016**. Niterói: UFF, 2017. Disponível em: <www.uff.br/?q=alunos-censo-2016>. Acesso em: 12 out. 2018.

UNIVERSIDADE FEDERAL FLUMINENSE. **Relatório de implantação e acompanhamento do sistema eletrônico de informações na UFF**. Niterói, 2017a.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. **Manual de gestão de documentos – apoio à Gestão Técnica**. Belo Horizonte: DPFP/UFMG, 2013.

UNIVERSIDADE FEDERAL DO PARÁ. **Sistema de protocolo, tramitação e arquivamento de processos**. Belém: UFPA/Arquivo Central, 2000.

UNIVERSIDADE FEDERAL DE SÃO CARLOS. **Manual de gestão de documentos**. São Carlos: UFSCar, 2015.

UNISIST. Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Universidade Federal de Minas Gerais. Belo Horizonte, v.10, n.1, p.83-94, 1981.

UPWARD, F. Structuring the Records Continuum: Part One. **Archives and Manuscripts**, v. 24, n. 2, p. 268-285, nov. 1996. Disponível em: <http://arrow.monash.edu.au/vital/access/services/Download/monash:1553_86/DOC>. Acesso em: 23 jul. 2018.

VAN DIJK, T. A. Complex semantic information processing. In: WALKER, D. C.; KARLGREN, H. e KAY, M. (Ed.). **Natural language in information science**. Estocolmo: Skriptor, 1976, p. 127-163.

VAN DIJK, T.A; KINTSCH, W. **Strategies of discourse comprehension**. New York: Academic Press, 1983.

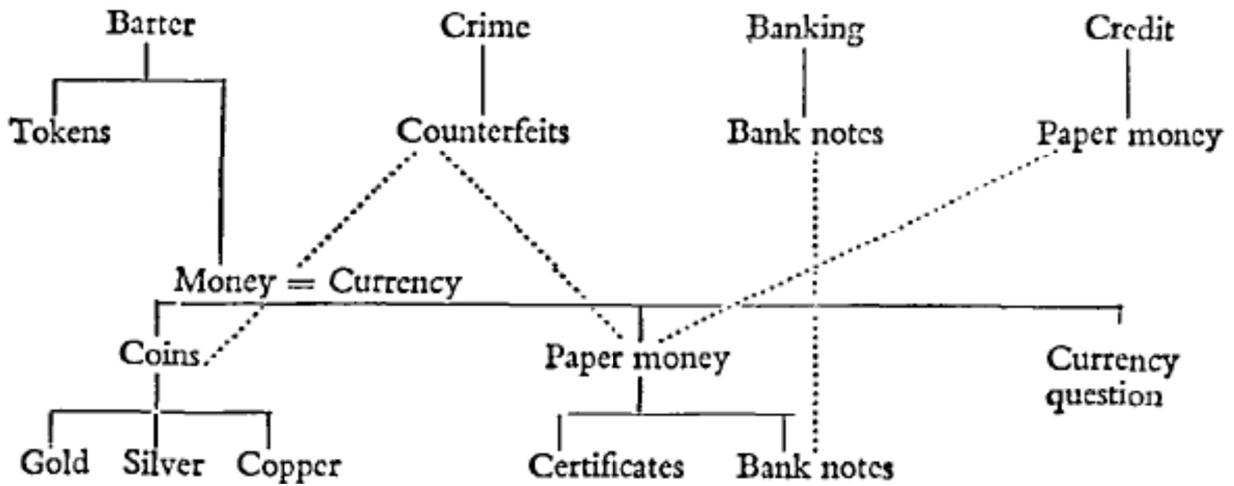
VICKERY, B. C. The structure of a connective index. **Journal of Documentation**, v. 6, n. 3, p. 140 - 151, 1950.

_____. Systematic subject indexing. **Journal of Documentation**, v. 9, n. 1, p. 48 - 57, 1953.

- _____. Developments in subject indexing, **Journal of Documentation**, v. 11, n. 1, p. 1 - 11, 1955.
- VYGOTSKY, L. S. **A formação social da mente**. São Paulo: Martins Fontes, 1991.
- WEINBERG, B. H. Indexing: history and theory. In: **ENCYCLOPEDIA OF LIBRARY AND INFORMATION SCIENCES**. Abingdon: Taylor & Francis, 2010, p. 2277-2290.
- WELLISCH, H. W. **Glossary of Terminology in Abstracting, Classification, Thesaurus Construction**. Nova Jersey: Medford, 2000.
- _____. Incunabula indexes. **The Indexer**, v. 19, n. 1, p. 3-12, 1994.
- WHEELER, J. Indexing: a current awareness bibliography. **Indexer**, v. 16, n.4, 1998.
- WHITE, S. A. **Business Process Modeling Notation**. Armonk: IBM, 2004.
- WILSON, I. E. Canadian University Archives. **Archivaria**, n. 3, p. 17-27, 1976.
- WITTGENSTEIN, L. **Investigações filosóficas**. Trad. J. C. Bruni. São Paulo: Nova Cultural, 1999.
- _____. **Tractatus logico-philosophicus**. São Paulo: Edusp, 1968.
- WITTY, F. J. The beginnings of indexing and abstracting: some notes towards a history of indexing and abstracting in antiquity and the Middle Ages. **Indexer**, v. 8, n. 4.1973, p. 193-198.
- YAKEL, E.; BOST, L. Understanding administrative use and users in university archives. **The American Archivist**, v. 56, n. 2, p. 593-615, 1994.
- YEO, G. Archival Description in the Era of Digital Abundance, **Comma**, V.2, p.15-25, 2013.
- _____. Debates em torno da descrição. In.: EASTWOOD, Terry; MACNEIL, Heather. (Org.). **Correntes atuais do pensamento arquivístico**. Belo Horizonte: UFMG, 2016. Cap.5
- ZENG, M. L.; GRACE, K. F.; ZUMER, M. *Using a semantic analysis tool to generate subject access points: a study using panofsky's theory and two research samples*. **Journal of Documentation**, v. 41, n. 6, p. 440-451, 2014.
- ZHANG, J; ZENG, M. L. A new similarity measure for subject hierarchical structures. **Journal of Documentation**, v. 70, n. 3, p. 364-391, 2014.

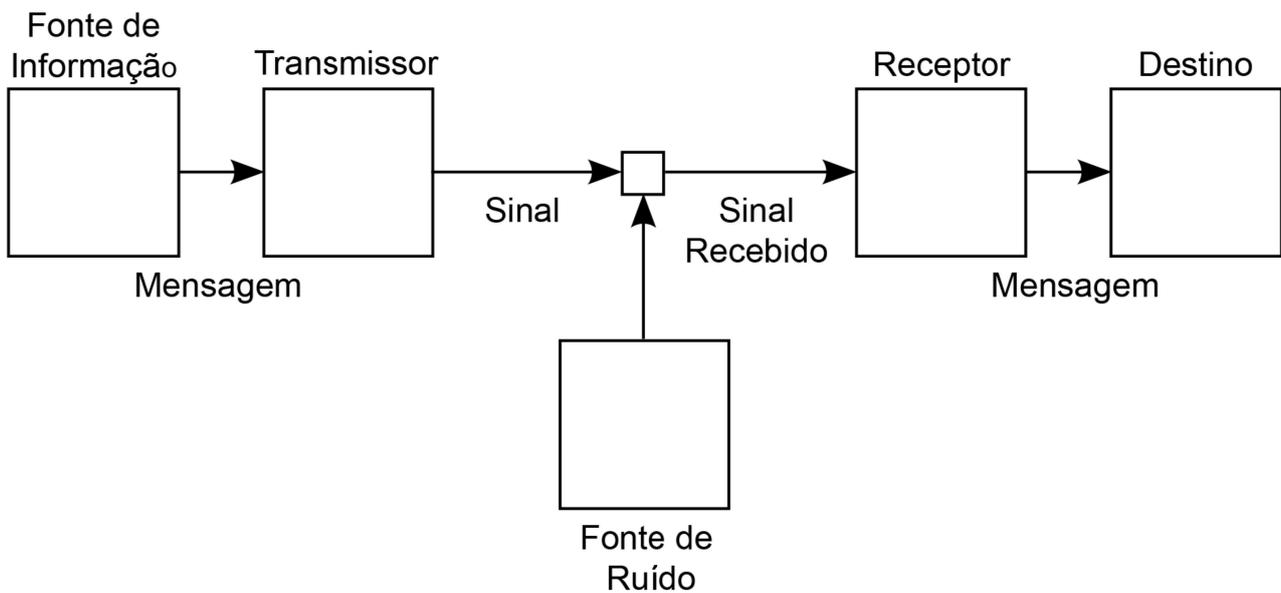
ANEXOS

ANEXO A – Diagrama conceitual em árvore



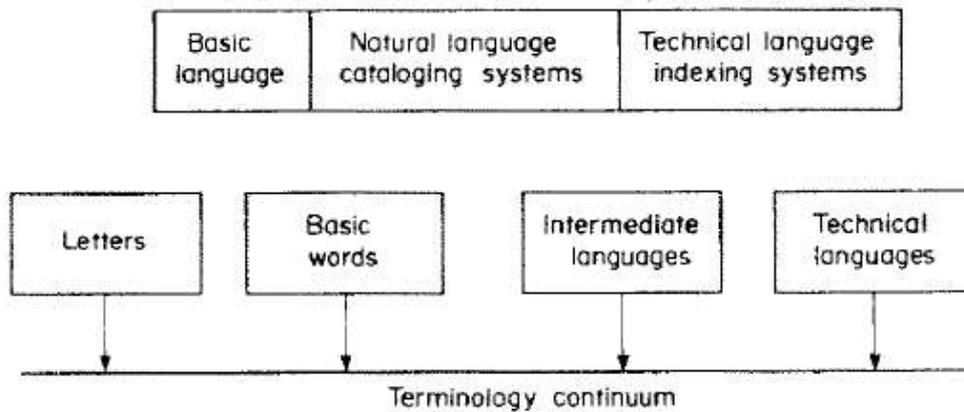
Fonte: Pettee (1946 apud VICKERY 1955, p. 2)

ANEXO B – Modelo de comunicação pela teoria matemática de Shannon e Weaver



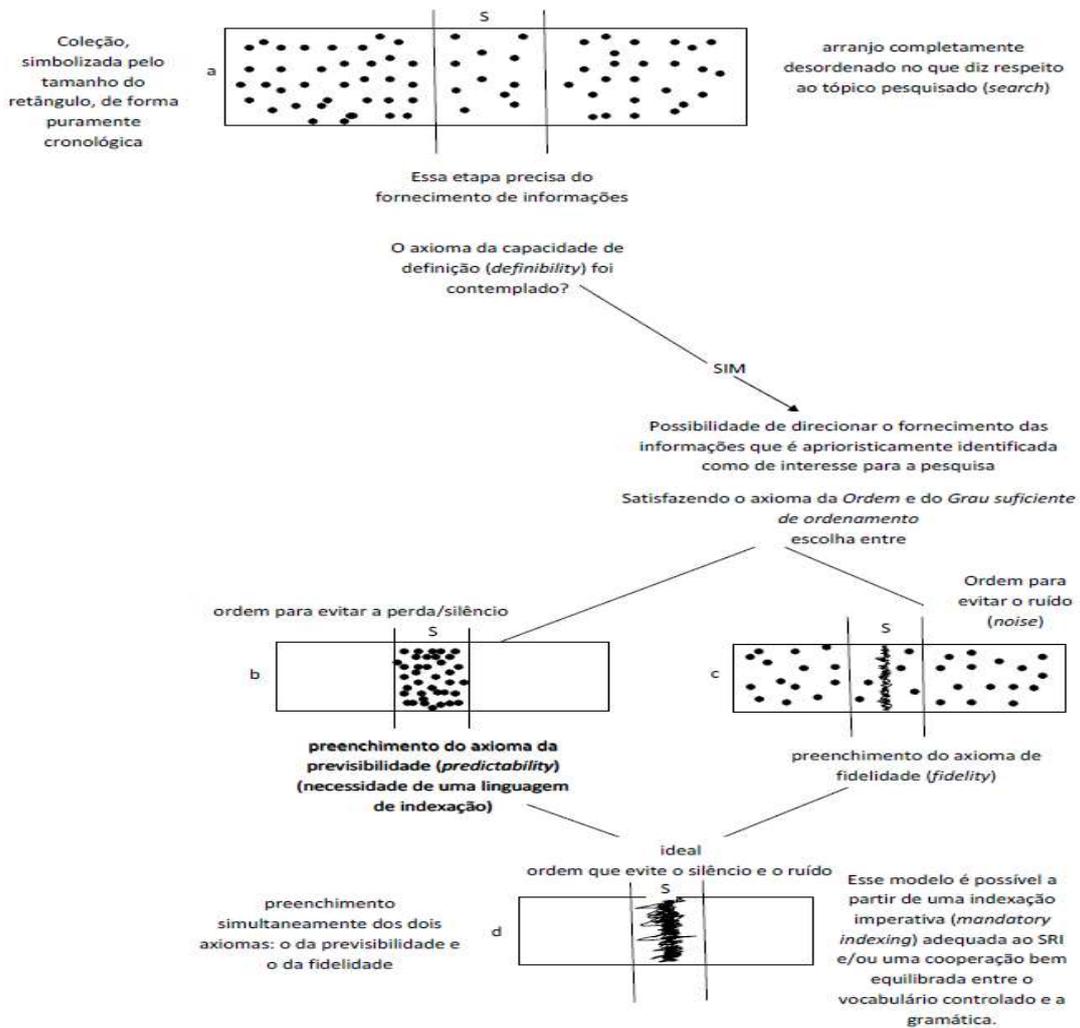
Fonte: Freixo (2006, p. 352)

ANEXO C – Fluxograma da teoria terminológica do *continuum* de Jonker



Fonte: Borko (1977, p. 357)

ANEXO D – Esquema de Fugman para os axiomas do processo de indexação



Fonte: Fugman (1985, p. 118, adaptação nossa)

ANEXO E – Modelo australiano do *records continuum*



Fonte: Upward (1996 apud COSTA FILHO; SOUSA 2017).

ANEXO F – Diagrama das etapas de gestão documental previstas pela ISO

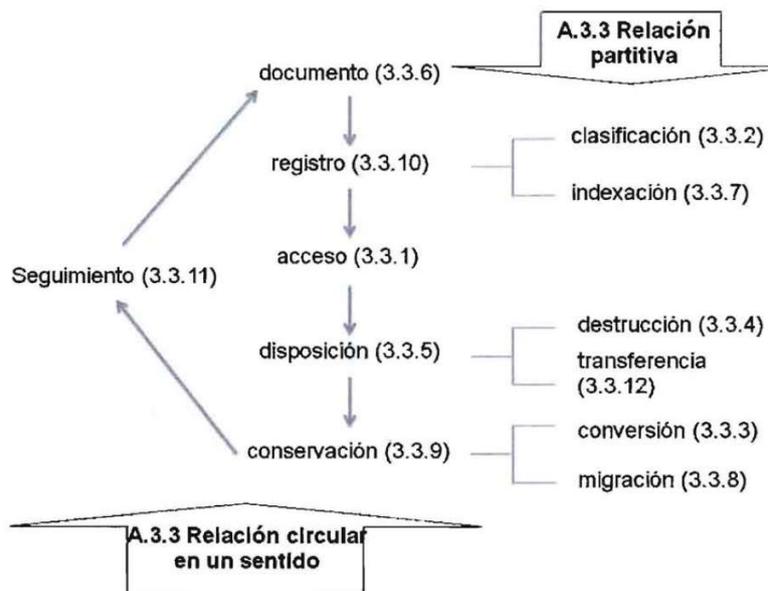
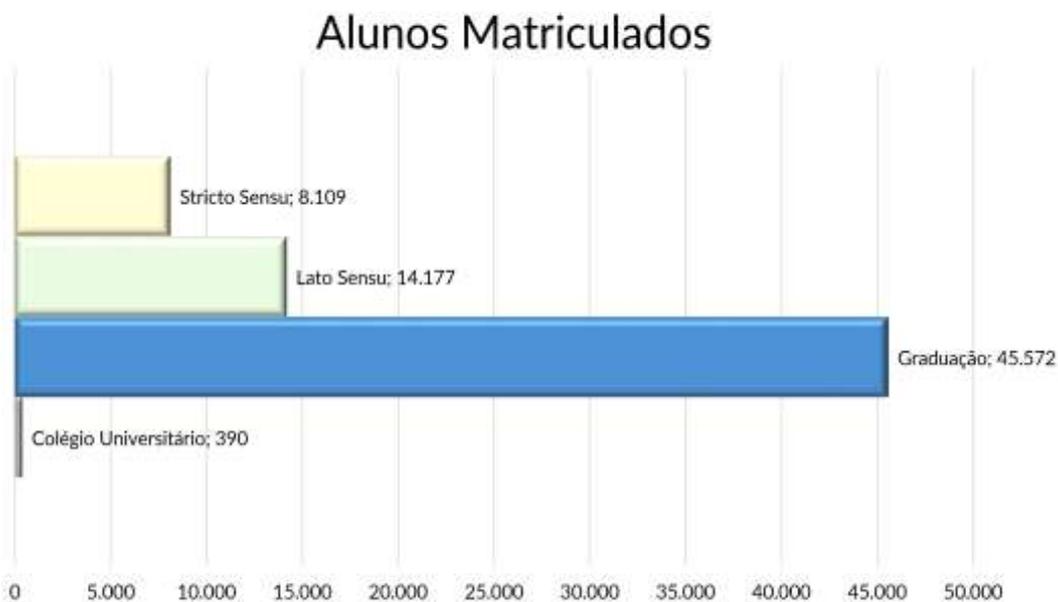


Figura A.3.3 – Conceptos relativos a los procesos de gestión documental (3.3)

Fonte: ISO 30300 (2011, p. 20)

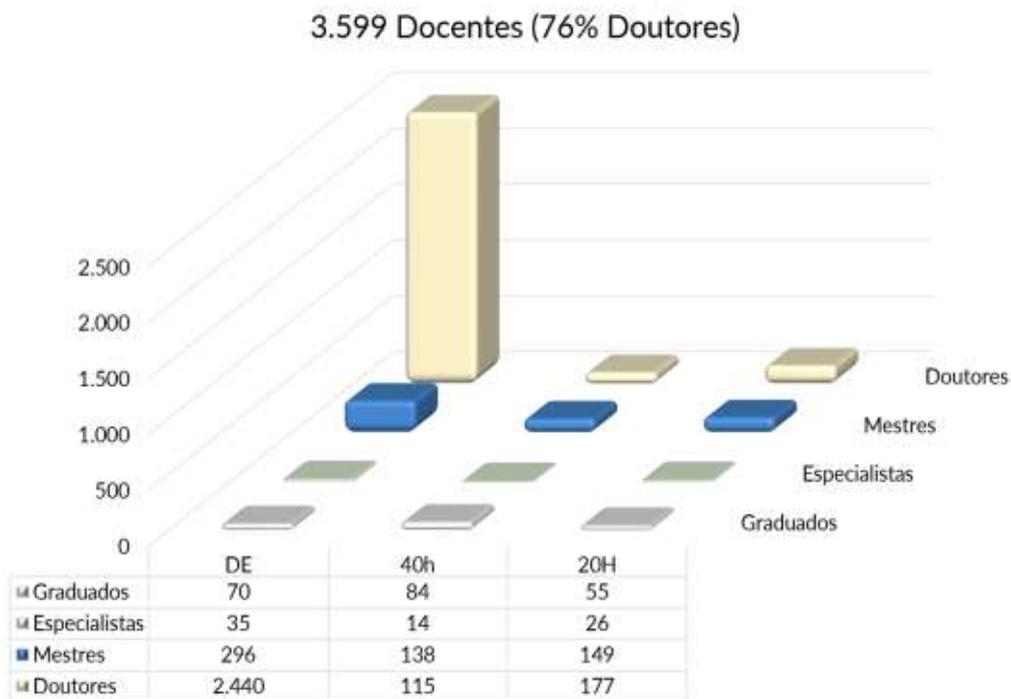
ANEXO G – Gráficos do Censo da UFF de 2016

CORPO DISCENTE



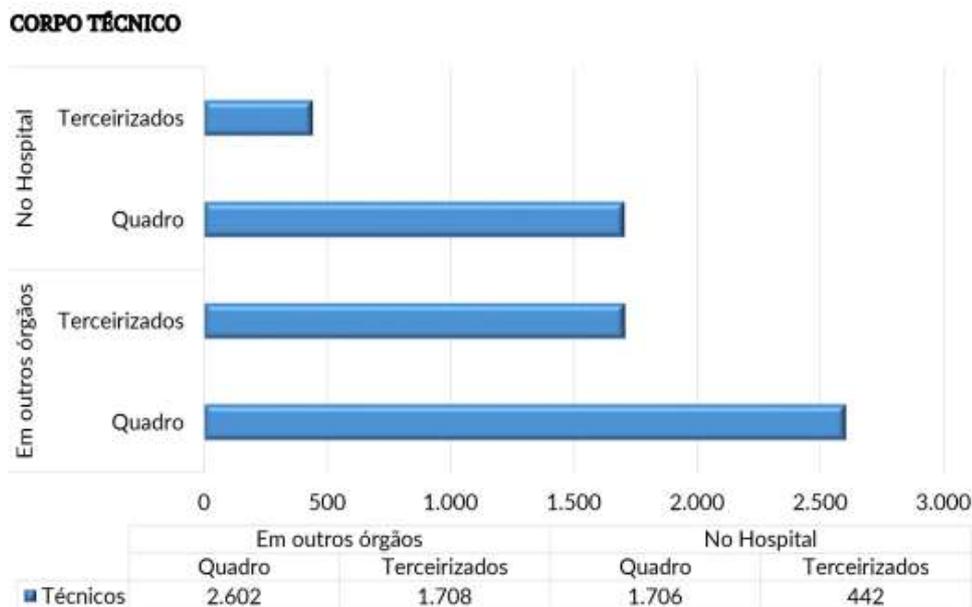
Fonte: UFF (2017).

CORPO DOCENTE



Fonte: Censo do Ensino Superior 2016 Mais Detalhes Série Histórica Censo 2017

Fonte: UFF (2017)



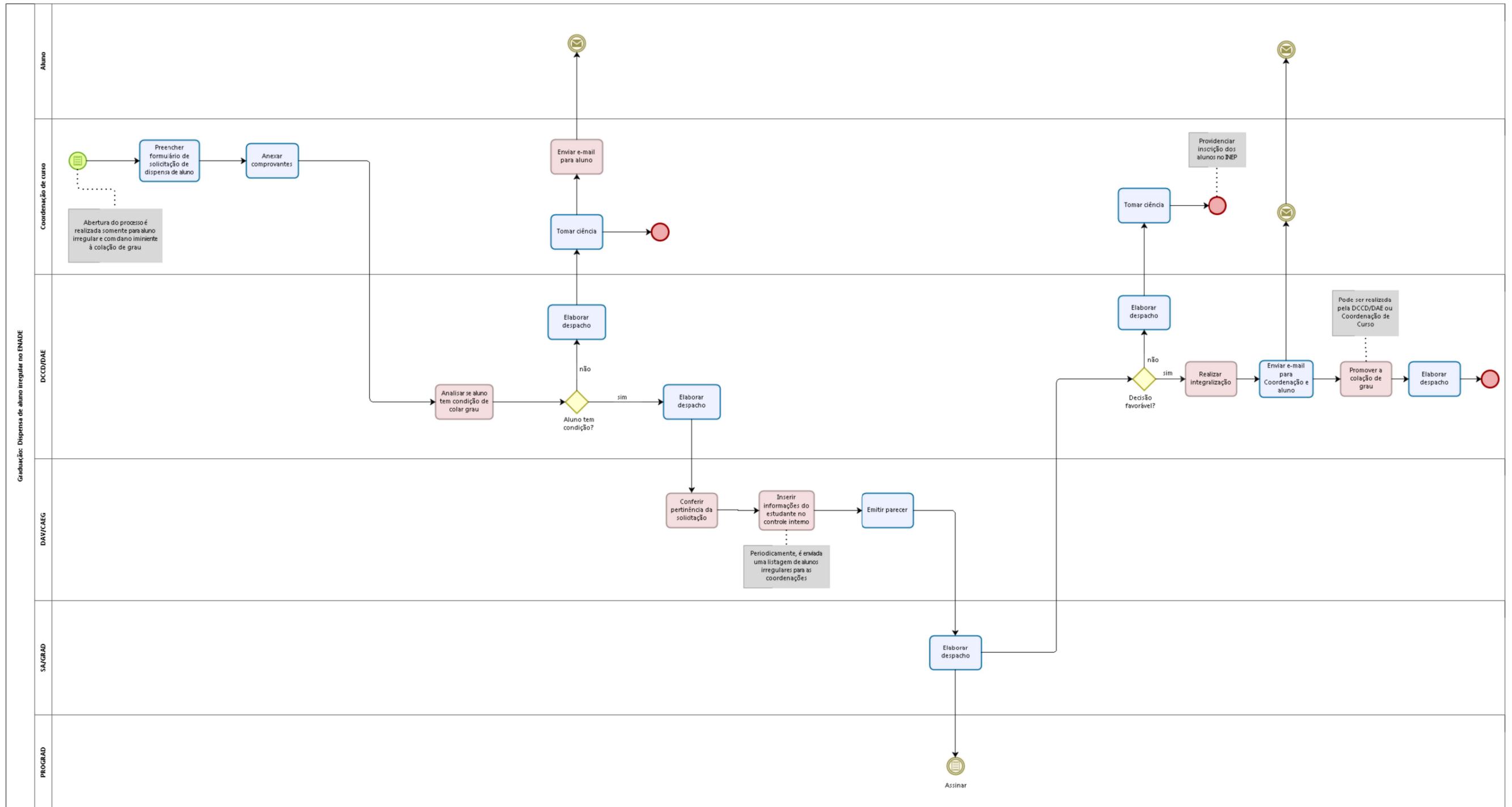
Fonte: UFF (2017)

ANEXO H – Tabela comparativa sobre os requisitos previstos no *E-Arq Brasil*

Seção do e-ARQ Brasil	Total de requisitos	Requisitos atendidos	Requisitos NÃO atendidos	Não foi possível avaliar
1 – Organização dos documentos arquivísticos	65	29%	57%	14%
2 - Tramitação e fluxo de trabalho	26	58%	19%	23%
3 - Captura	37	41%	54%	5%
4 - Avaliação e destinação	47	17%	79%	4%
5 - Pesquisa, localização e apresentação de documentos	41	44%	54%	2%
6 - Segurança	98	32%	5%	63%
7 - Armazenamento	20	0%	15%	85%
8 - Preservação	18	0%	67%	33%
9 - Funções administrativas e técnicas	3	67%	33%	0%
10 - Conformidade com a legislação e regulamentações	3	33%	33%	34%
11 - Usabilidade	25	56%	16%	28%
12 - Interoperabilidade	3	33%	33%	34%
13 - Disponibilidade	1	0%	0%	100%
14 - Desempenho e escalabilidade	5	60%	40%	0%

Fonte: Arquivo Nacional (2015)

ANEXO I – Mapa do processo de dispensa de aluno irregular do Enade – Modelo BPMN



ANEXO J – Tabela com as etapas de implementação e gestão do SEI na UFF

1. ADESÃO AO SISTEMA ELETRÔNICO DE INFORMAÇÕES NA UFF			
ETAPAS	DESCRIÇÃO	ÁREA(S)	DATA
Análise das especificidades do Sistema Eletrônico de Informação (SEI).	Consulta às IFES que já possuíam o SEI implantado, para informes de ordem técnica nas áreas de TI e de gestão de documentos. Neste último caso, para sanar dúvidas quanto à participação do protocolo na implantação do SEI e sua função pós- implantação.	Comitê de Gestão da Informação Comitê de Governança SDC - CAR STI GPCA-PROAD	Dez/16
Identificação de experiências na administração pública federal quanto ao uso do SEI	Experiências de outras instituições do Poder Executivo Federal que já adotaram o SEI, sendo aproximadamente 50% IFES; Questões técnicas quanto à instalação e infraestrutura; Ferramentas para o uso do SEI; Ilhas de digitação para dar conta do legado documental produzido fisicamente e que não pode passar ao largo da inserção do sistema; Redução de custos com suprimentos (tonner, aluguel de impressoras, etc.); Redução de tempo na tramitação de processos; Políticas de back up; Apresentação de uma versão de testes do SEI e os passos a serem seguidos para assinatura de adesão ao sistema.	Comitê de Gestão da Informação Comitê de Governança SDC-CAR STI GPCA-PROAD	Dez/16
Decisão sobre a adoção do Sistema eletrônico de documentos	Tomada de decisão pela alta administração da UFF em conjunto com suas Pró-Reitorias e Superintendências; Proposta de constituir Comissão para implantação e acompanhamento da Plataforma digital do PEN-SEI na UFF.	GAR SDC- CAR STI SRI SAEN PROPLAN GPCA-PROAD PROGRAD CASQ PROGEPE DCQ-EGGP Comitê de Gestão da Informação	Jan/17
Envio de Ofício ao Ministério do Planejamento (MP)	Ofício remetido ao Ministério do Planejamento, manifestando formalmente o desejo de participar do projeto Processo Eletrônico Nacional (PEN) e obter o direito de uso do Sistema Eletrônico de Informações (SEI).	GABINETE DO REITOR	Jan/17
Indicação de membros para composição da Comissão de Implantação do SEI na UFF	Representação de técnico-administrativos das principais áreas da universidade;	GABINETE DO REITOR SDC	Jan/17
Pré-definição de estratégias para atuação da Comissão de implantação e acompanhamento do SEI na UFF	Definição dos Projetos Pilotos ; Criação de email para facilitar comunicação do grupo e uso do Google drive para edição de arquivos encaminhados.	Comitê de Governança SDC STI PROPLAN	Fev/17
Montagem das Ilhas de Digitalização nas unidades protocolizadoras	O SEI é disponibilizado sem custos para a UFF. A aquisição dos equipamentos para montagem das ilhas de digitalização, ficou sob responsabilidade da PROAD	PROAD	

Fonte: UFF (2017a)

2. GESTÃO DO SEI NA UFF			
ETAPA(S)	DESCRIÇÃO	ÁREA(S)	DATA
Publicação da Portaria de criação da Comissão de Implantação do SEI na UFF	Portaria Nº 58.196 de 23 de fevereiro de 2017; Retificada pela Portaria Nº 58.198 em 06/03/17; Alterada pela Portaria Nº 60.083 de 30 outubro de 2017	GABINETE DO REITOR	Fev/17 Mar/17 Out/17
Elaboração do Plano de Trabalho	Elaborar as ações iniciais para implantação do SEI na UFF	SDC - CAR STI	Fev/2017
Envio do Plano de Trabalho ao MP	Definição do planejamento inicial para implantação do SEI na UFF e posterior formalização do ato da cessão de direito de uso do SEI.	SDC GABINETE DO REITOR	Fev/2017
Plano de Gerenciamento de Projeto (PGP)	Detalhamento das ações a serem desenvolvidas durante a vigência do projeto, propondo o monitoramento e controle das atividades atribuindo responsabilidades a partir da delimitação do escopo.	SDC STI Comissão SEI/UFF	Fev/2017
Plano de Riscos	Previsão de possíveis impactos com a implantação do sistema	SDC STI	Mar/17
Plano de qualidade	Identificar padrões de qualidade relevantes para o projeto, e determinar como alcançá-los.	SDC STI	Mar/17
Formalização do Acordo de Cooperação Técnica (ACT)	Assinatura do Acordo de Cooperação Técnica para cessão de uso do SEI na UFF com a presença do Secretário do Ministério do Planejamento, Desenvolvimento e Gestão e o Vice-Reitor da Universidade Federal Fluminense.	GABINETE DO REITOR SDC Comissão SEI-UFF	Abril/2017
Visitas às unidades dos Campi Fora de Sede, para divulgar o Projeto SEI/UFF, e o Plano de Trabalho, com o planejamento inicial de implantação	Foi feita a apresentação do Projeto SEI/UFF, abordando especialmente as vantagens do sistema eletrônico de informações, a estratégia de implantação na UFF, os avanços já ocorridos naquelas datas e a contribuição esperada de cada unidade.	Assessoria de Interiorização	Maio e Junho/17
Definição de reuniões semanais	A Comissão de Implantação e Acompanhamento do SEI na UFF passou a se reunir semanalmente para tratar de assuntos relacionados à gestão do Sistema	Comissão SEI/UFF	Fev a Dez/ 17
Implantação do processo piloto	Implantação do processo piloto: Reembolso de bilhetes de passagem e ativação do ambiente de produção do SEI	Comissão SEI/UFF	Set/17
Informe ao Ministério do Planejamento	Comunicado sobre o cumprimento do Decreto nº 8.539, de 8 de outubro de 2015.	Comissão SEI/UFF	Out/2017
Implantação sequencial de processos	Implantação de novos processos no SEI: <ul style="list-style-type: none"> • Adicional de Insalubridade; • Adicional de Periculosidade. 	Comissão SEI/UFF	Nov/17
Implantação sequencial de processos	Implantação de novos processos no SEI: <ul style="list-style-type: none"> • 2ª via de diploma de graduação; • 2ª via de histórico escolar de graduação; • Doação de bens móveis – FEC; • Autorização para dirigir veículo oficial 	Comissão SEI/UFF	Dez/2017

Fonte: UFF (2017a)

	a. DOCUMENTAÇÃO		
Definição de uma subcomissão de Documentação	Representantes da Superintendência de Documentação (Coordenação de Arquivos) e Gerência Plena de Comunicações Administrativas (PROAD), membros da Comissão do SEI	SDC GPCA/PROAD	Jan – Dez 17
Levantamento das tipologias documentais	Realizar o levantamento das tipologias documentais para fins de racionalização dos documentos a partir do mapeamento de processos.	CAR/SDC GPCA/PROAD	Jan – Dez 17
Padronização de documentos utilizados nos processos a serem implantados no SEI	Padronizar os documentos de acordo com os preceitos arquivísticos e administrativos;	CAR/SDC GPCA/PROAD	Jan – Dez 17
Orientação quanto à classificação de documentos	Atribuir à classificação de documentos de arquivo que serão inseridos no SEI;	CAR/SDC GPCA/PROAD	Jan – Dez 17
Produção de instruções para os processos implantados no SEI	Orientar quanto à classificação de acesso;	CAR/SDC GPCA/PROAD	Jan – Dez 17
Propor ações de melhoria quanto a racionalização de documentos a partir do mapeamento de processos	Assessorar nas questões relativas à preservação de documentos arquivísticos digitais.	CAR/SDC GPCA/PROAD	Jan – Dez 17
Autenticação de documentos	Orientar quantos aos procedimentos de autenticação de documentos	CAR/SDC GPCA/PROAD STI	Jan – Dez 17
Elaboração de normas	Elaboração de normativas internas com base nas determinações legais quanto à gestão de documentos no SEI	CAR/SDC GPCA/PROAD	Jan – Dez 17
Elaboração de manuais	Apoiar o uso do SEI	CAR/SDC GPCA/PROAD	Jan – Dez 17
Legislações vigentes	Levantamento e análise da legislação pertinentes aos processos	CAR/SDC GPCA/PROAD	Jan – Dez 17
Publicação e atualização das Normas Internas	Publicação e atualização das normas internas sobre os processos implantados no SEI	GPCA/PROAD GABR	Jan – Dez 17
Atualização de normas internas	Atualização das normas internas sobre os processos implantados no SEI	GPCA/PROAD	Jan – Dez 17
Assessoria às unidades protocolizadoras	Assessoria operacional nas unidades protocolizadoras, através de Treinamento interno	GPCA/PROAD	Jan – Dez 17
Acompanhamento dos trabalhos realizados no SEI nas unidades protocolizadoras	Monitoramento e orientação sobre os trabalhos realizados no SEI nas unidades protocolizadoras	GPCA/PROAD	Jan – Dez 17
	b. COMUNICAÇÃO		
Definição de uma subcomissão de Comunicação	Representantes da Superintendência de Comunicação Social, membros da Comissão do SEI	SCS	Jan – Dez 17
Elaboração do plano de comunicação	Abrange a proposta de mudança cultural no uso de documentos em o meio digital, buscando diversificar quanto aos instrumentos de marketing, principalmente interno (endomarketing).	SCS	Jan –Fev/17
Criação de logo para o projeto	Logo Institucional do SEI/UFF	SCS	Fev– Mar/17
Criação do site SEI/UFF	Site com informes para acesso e transparência das ações desenvolvidas pela Comissão	SCS	Fev – Mar/17
Confecção do material de divulgação	Informação e Marketing	SCS	Fev – Dez/17
Nótícias sobre implantação do SEI na UFF	Informação e Marketing	SCS	Fev – Dez/17
Envio de e-mails	Envio de e-mails sobre o SEI e curso ENAP Envio de e-mails para gestores com status do projeto	SCS	Fev – Dez/17
Alimentação do site SEI/UFF	Atualização dos informes	SCS	Fev – Dez/17

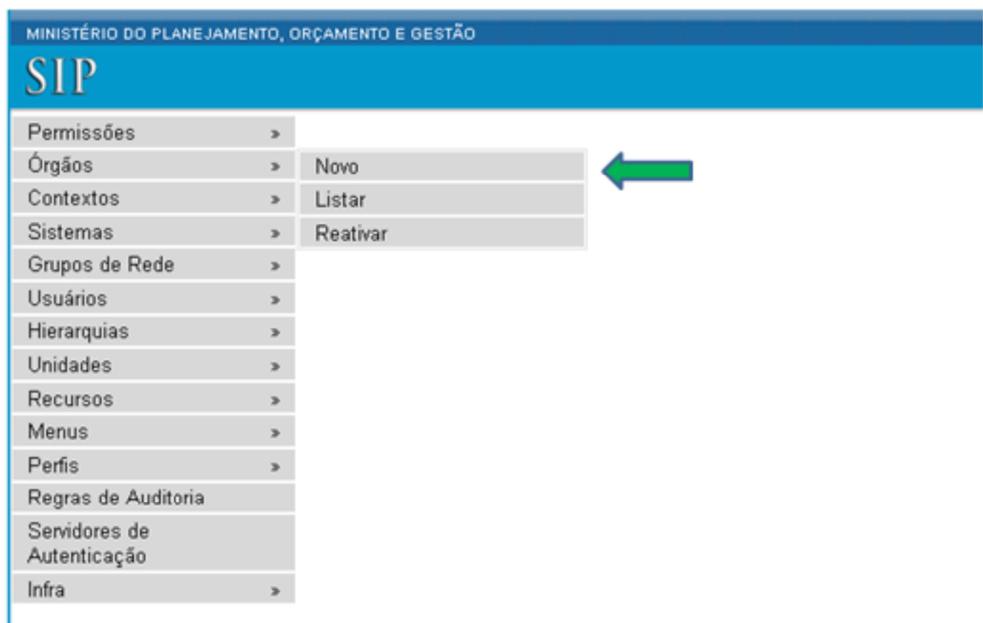
Fonte UFF (2017a).

ANEXO K – Folheto informativo sobre o SEI na UFF



Fonte: UFF (2018).

ANEXO L – Tela do software SEI, perfil “administrador” - Cadastros



Fonte: Brasil (2018a)



Fonte: BRASIL (2018a).

ANEXO M – Código atribuído para processo eletrônico



Fonte: Brasil (2018a)

ANEXO N – Tela inicial do SEI para usuários – Gestores de processos

UNIVERSIDADE FEDERAL FLUMINENSE

sei! Para saber+ Menu Pesquisa GCI   

- Controle de Processos
- Iniciar Processo
- Retorno Programado
- Pesquisa
- Base de Conhecimento
- Textos Padrão
- Modelos Favoritos
- Blocos de Assinatura
- Blocos de Reunião
- Blocos Internos
- Contatos
- Processos Sobrestados
- Acompanhamento Especial
- Marcadores
- Pontos de Controle
- Estatísticas 
- Grupos 

Controle de Processos

Ver processos atribuídos a mim Ver por marcadores Visualização detalhada

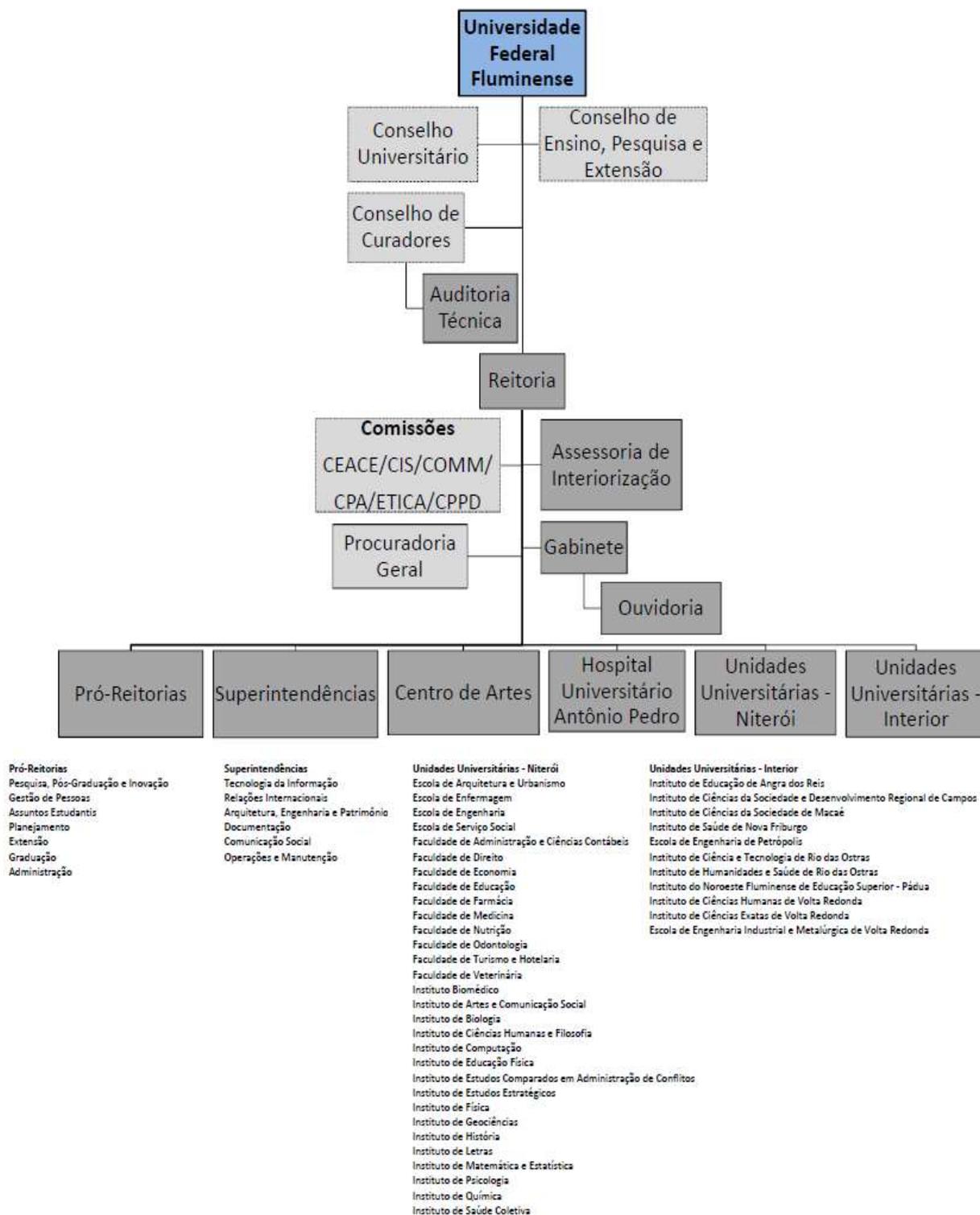
2 registros:		1 registro:	
<input checked="" type="checkbox"/>	Recebidos	<input checked="" type="checkbox"/>	Gerados
<input type="checkbox"/>	23069.155499/2018-30	<input type="checkbox"/>	23069.154289/2018-24 (38337029749)
<input type="checkbox"/>	23069.154888/2018-48		

Acesse as lojas App Store ou Google Play e instale o aplicativo do SEI! no seu celular.

Abra o aplicativo do SEI! e faça a leitura do código abaixo para sincronizá-lo com sua conta.

Fonte: <https://sei.uff.br/sei/controlador.php?acao=procedimento_controlar&acao_origem=principal&acao_retorno=principal&inicializando=1&infra_sistema=100000100&infra_unidade_atual=110000556&infra_hash=82d4db09a486fcbb5e8a8b5bb7f39ee49bafdb87a6c0d77ad3b26979dfb566b4>. Acesso em: 14 dez . 2018.

ANEXO O – Organograma da Universidade Federal Fluminense (UFF)



Fonte: <<http://www.uff.br/?q=organogramas>>. Acesso em: 12 dez 2018.

ANEXO P – Quadro de funções previstas pelo SEI

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

seii

Para saber+ Menu Pesquisa SE

Assinaturas das Unidades

Pesquisar Adicionar Excluir Imprimir Fechar

Cargo / Função:

Lista de Assinaturas (28 registros):

<input checked="" type="checkbox"/> Cargo / Função	Ações
<input type="checkbox"/> Analista	  
<input type="checkbox"/> Analista de Execução	  
<input type="checkbox"/> Analista de Infraestrutura	  
<input type="checkbox"/> Analista de Planejamento e Orçamento	  
<input type="checkbox"/> Analista de Tecnologia da Informação	  
<input type="checkbox"/> Arquivista	  
<input type="checkbox"/> Assessor	  
<input type="checkbox"/> Assessor Especial	  
<input type="checkbox"/> Assessor Técnico	  
<input type="checkbox"/> coordenador	  
<input type="checkbox"/> Coordenador(a)	  
<input type="checkbox"/> Coordenador(a)-Geral	  
<input type="checkbox"/> Coordenadora	  

Fonte: Brasil (2018a)

ANEXO Q – Código de Classificação das IFES no SEI

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

seii

Para saber+ Menu Pesquisa SE

Assuntos

Pesquisar Novo Desativar Excluir Imprimir Fechar

Palavras para Pesquisa:

« 2 »»

Lista de Assuntos (337 registros - 51 a 100):

<input checked="" type="checkbox"/> Código	Descrição	Ações
<input type="checkbox"/> 023.14	DESIGNAÇÃO. DISPONIBILIDADE. REDISTRIBUIÇÃO. SUBSTITUIÇÃO	  
<input type="checkbox"/> 023.15	REQUISICÃO. CESSÃO (Internas e/ou Externas de Servidores para a realização de serviços temporários)	  
<input type="checkbox"/> 024	DIREITOS, OBRIGAÇÕES E VANTAGENS	  
<input type="checkbox"/> 024.1	FOLHAS DE PAGAMENTO. FICHAS FINANCEIRAS	  
<input type="checkbox"/> 024.11	SALÁRIOS, VENCIMENTOS, PROVENTOS E REMUNERAÇÕES	  
<input type="checkbox"/> 024.111	SALÁRIO-FAMÍLIA	  
<input type="checkbox"/> 024.112.a	ABONO OU PROVENTO PROVISÓRIO	  
<input type="checkbox"/> 024.112.b	ABONO DE PERMANÊNCIA EM SERVIÇO	  
<input type="checkbox"/> 024.119	OUTROS SALÁRIOS, VENCIMENTOS, PROVENTOS E REMUNERAÇÕES	  
<input type="checkbox"/> 024.12	GRATIFICAÇÕES (inclusive Incorporações)	  

Fonte: Brasil (2018a)

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

sei Para saber+ Menu Pesquisa DELOG(2067) Nova Excluir Fechar

Seções

Modelo: Memorando

Lista de Seções (13 registros)

<input checked="" type="checkbox"/>	Nome	Ordem	Cabeçalho	Rodapé	Principal	Assinatura	Somente Leitura	Dinâmica	HTML	Ações
<input type="checkbox"/>	Cabeçalho	0	X				X	X	X	
<input type="checkbox"/>	Tipo e número	1					X	X		
<input type="checkbox"/>	Local e data	2					X	X		
<input type="checkbox"/>	Separador	3					X			
<input type="checkbox"/>	Interessado	4								
<input type="checkbox"/>	Separador	5					X			
<input type="checkbox"/>	Assunto	6								
<input type="checkbox"/>	Separador	7					X			
<input type="checkbox"/>	Vocativo	8								
<input type="checkbox"/>	Texto	9			X					
<input type="checkbox"/>	Separador	10					X			
<input type="checkbox"/>	Fecho	11								
<input type="checkbox"/>	Assinatura	12				X				

Fonte: Brasil (2018a)

ANEXO R – Classificação do grau de sigilo no SEI

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

sei Para saber+ Menu Pesquisa SE Salvar Cancelar

Nova Hipótese Legal

Nível de Restrição de Acesso

Sigiloso Restrito

Nome: Informação Pessoal

Base Legal: Lei 12.527/2011, Art. 31

Descrição: As informações pessoais relativas à intimidade, vida privada, honra e imagem terão seu acesso restrito, independentemente de classificação de sigilo e pelo prazo máximo de 100 anos a contar da sua data de produção, a agentes públicos legalmente autorizados e à pessoa que elas se referirem e poderão ter autorizada sua divulgação ou acesso por terceiros diante de previsão legal ou consentimento expresso da pessoa a quem elas se referirem.

Fonte: Brasil (2018a)

ANEXO S – Exemplo de registro da trilha de auditoria realizada pelo SEI

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

seil

Para saber+ Menu Pesquisa

SE

Salvar Escher

Configuração do Histórico

Lista de Tipos de Andamentos (93 registros):

<input checked="" type="checkbox"/> Completo	<input checked="" type="checkbox"/> Resumido	Nome
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	@DESCRICAÇÃO@
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	@MOTIVO@ do documento @DOCUMENTO@ no veículo @VEICULO@ de @DATA@ @TIPO@
<input type="checkbox"/>	<input type="checkbox"/>	Acesso automático ao processo
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Adicionado processo relacionado @PROCESSO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterada hipótese legal do documento @DOCUMENTO@ para @HIPOTESE_LEGAL@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterada hipótese legal do processo para @HIPOTESE_LEGAL@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterada ordem dos protocolos
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado grau de sigilo do documento @DOCUMENTO@ para @GRAU_SIGILO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado grau de sigilo do processo para @GRAU_SIGILO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado nível de acesso do documento @DOCUMENTO@ para @NIVEL_ACESSO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado nível de acesso do processo para @NIVEL_ACESSO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado nível de acesso geral para @NIVEL_ACESSO@V
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Alterado tipo de conferência do documento @DOCUMENTO@ para @TIPO_CONFERENCIA@
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Anexado ao processo @PROCESSO@
<input type="checkbox"/>	<input type="checkbox"/>	Anexo @ANEXO@ removido do documento @DOCUMENTO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Arquivado documento @DOCUMENTO@ no localizador @LOCALIZADOR@
<input type="checkbox"/>	<input type="checkbox"/>	Arquivo @ANEXO@ anexado no documento @DOCUMENTO@
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Assinado Documento @DOCUMENTO@ @USUARIO@

Fonte: Brasil (2018a)

ANEXO T – Autuação de processo no SEI

UNIVERSIDADE FEDERAL FLUMINENSE

seil

Para saber+ Menu Pesquisa

GCI

Salvar Voltar

Iniciar Processo

Tipo do Processo:
Graduação: Dispensa de aluno irregular no ENADE

Especificação:

Classificação por Assuntos:
110 - NORMALIZAÇÃO, REGULAMENTAÇÃO

Interessados:

Observações desta unidade:

Nível de Acesso

Sigiloso Restrito Público

Acesse as lojas App Store ou Google Play e instale o aplicativo do SEI no seu celular.

Abra o aplicativo do SEI e faça a leitura do código abaixo para sincronizá-lo com sua conta.



Fonte: Brasil (2018a)

ANEXO U – Árvore de processos e documentos no SEI



Fonte: Brasil (2018a)

ANEXO V – Resultado de busca no AtomM-UnB

Termo de Assuntos	Nota de âmbito	Resultados
Atividades		1
Tecnico de laboratorio e pesquisa		1
Ensino e pesquisa		1
Laboratório de Ensino e Pesquisa		1
Auxiliar de Laboratorio de Ensino e Pesquisa		1
Tecnico de laboratorio de Ensino e pesquisa		1
Membro da Câmara de Pesquisa e Pós-Graduação		1
Pesquisa Científica		1
Pesquisa Biomedicas		2
Atividades de planejamento		1
Projeto de pesquisa		1
Pesquisa		9

Fonte: <<https://atom.unb.br>>. Acesso em: 17 dez. 2018.