

INOVAÇÕES NA REVISÃO POR PARES: O PAPEL DO SOFTWARE

INNOVATIONS IN PEER REVIEW: SOFTWARE AGENTS' ROLE

Vinícius Medina Kern

vmkern@pq.cnpq.br

Universidade Federal de Santa Catarina

Resumo: Um grupo internacional de 33 cientistas, encabeçados por Tennant, divulgou recentemente sua perspectiva multidisciplinar sobre inovações futuras e emergentes para a revisão por pares editorial a partir de plataformas da Web 2.0: Amazon, Blockchain, GitHub, Hipotesis, Reddit, Stack Exchange e Wikipedia. Várias inovações propostas são ações a serem efetivadas por agentes artificiais, em sintonia com o conceito de sistema sociotecnológico de Fuchs, no qual as propriedades emergem da colaboração dinâmica humano-máquina. O objetivo da pesquisa foi identificar, entre as inovações descritas, as tarefas automatizáveis intensivas em conhecimento. Nesta pesquisa bibliográfica, o corpus foi o estudo de inovações emergentes de Tennant et al. De 56 trechos identificados que remetem a funções e tarefas da revisão por pares automatizáveis, 44 foram descartados por focar atividades burocráticas. Foram analisados os 12 que descrevem tarefas intensivas em conhecimento dos tipos alocação, avaliação, classificação, diagnóstico, modelagem, monitoramento e previsão. Há tarefas cuja automação agiliza o sistema ou o melhoram com serviços antes indisponíveis, bem como tarefas associadas à repressão da má conduta científica. Discutem-se essas inovações como sintoma da transição do papel do *software*, da informatização ao desempenho de tarefas em colaboração dinâmica com pessoas e do tecnocentrismo ao funcionamento sociotecnológico, em que o *software* apoia a tomada de decisão humana em vez de substituí-la.

Palavras-Chave: Revisão por Pares; Comunicação Científica; Sistemas Sociotecnológicos; Commonkads; Engenharia do Conhecimento.

Abstract: An international group of 33 scientists, led by Tennant, recently released their multidisciplinary perspective on future and emerging innovations for editorial peer review inspired in Web 2.0 platforms: Amazon, Blockchain, GitHub, Hypothesis, Reddit, Stack Exchange, and Wikipedia. Several innovations proposed are actions to be carried out by artificial agents, in line with Fuchs' concept of sociotechnological system, in which the properties emerge from human-machine dynamic collaboration. The research objective was to identify knowledge-intensive automatable tasks among the innovations described. What are these processes performed by software and what other actors in the system do they involve? The research objective was to identify these processes and actors involved. This is a bibliographic research using as corpus the study of emerging innovations by Tennant et al. Of 56 identified passages that refer to automatable peer review functions and tasks, 44 were discarded for focusing on bureaucratic activities. We analyzed the 12 excerpts that describe knowledge-intensive tasks, resulting in allocation, evaluation, classification, diagnosis, modeling, monitoring, and prediction tasks. Some tasks' automation speeds up the system or improve it with previously unavailable services. Some tasks are associated with the repression of scientific misconduct. The innovations are discussed as a symptom of software's transition from mere virtualization to dynamic collaboration with people, and from technocentrism to sociotechnological functioning, in which software supports human decision-making instead of replacing it.

Keywords: Peer Review; Scholarly Communication; Sociotechnological Systems; Commonkads; Knowledge Engineering.

1 INTRODUÇÃO

A revisão por pares usa sistematicamente árbitros ou revisores, usualmente voluntários, para avaliar a aceitabilidade de manuscritos submetidos para publicação (ZUCKERMAN; MERTON, 1971). Seu custo não-monetário global anual foi estimado em 1,9 bilhão de libras esterlinas (RESEARCH INFORMATION NETWORK, 2008), valor aproximado a R\$ 6,5 bilhões de 2008.

Apesar de amplamente disseminada e dominante como sistema de controle de qualidade na ciência, é inconsistente, sem transparência ou padronização (TENNANT et al., 2017), bem como “ocasionalmente corrupta, às vezes uma charada” (RENNIE, 2016, p. 31). A apropriação do trabalho gratuito dos autores e revisores por editoras que vendem seus periódicos (ou os artigos que os compõem) é apenas um dos fatos bizarros do sistema.

Este artigo trata da revisão por pares na modalidade editorial, isto é, aplicada à seleção de artigos enviados para publicação. Há outras formas de revisão por pares no âmbito acadêmico – para contratação, promoção e concessão de financiamento à pesquisa, por exemplo. Mesmo na modalidade editorial, há variações entre distintos sistemas, entre essas: a identidade de autor e revisor pode ou não ser revelada à outra parte, o editor pode ou não solicitar sugestões de potenciais revisores aos autores, as revisões podem permanecer opacas ou ser publicadas junto ao artigo.

Uma das fontes de variabilidade nos sistemas de revisão por pares é a inovação tecnológica e social típica da Web 2.0, que já permitiu a consolidação de algumas inovações na comunicação científica, como o depósito de *preprints* que permitem à comunidade conhecer avanços de pesquisa antes do usualmente demorado processo de revisão e publicação. O caso mais proeminente é o do arXiv, que, no domínio específico da Física de Altas Energias, leva a índices de publicação posterior em periódicos da ordem de 70% (ALVAREZ; CAREGNATO, 2017).

Tennant et al. (2017), um grupo multidisciplinar de 33 cientistas, examinaram modelos emergentes de seleção de conteúdos em plataformas da Web 2.0 aplicáveis à revisão por pares. Esses 33 autores também propõem um novo modelo de plataforma híbrida que, segundo postulam, resolve muitos de seus problemas técnicos e sociais.

Considerando que há inovações possíveis por meio de *software*, a pergunta de pesquisa é: Quais são as tarefas inteligentes ou intensivas em conhecimento (SCHREIBER et al., 2000) automatizáveis? Esta pesquisa tem o objetivo de identificar tarefas típicas de sistemas baseados em conhecimento aplicáveis à revisão por pares.

A pesquisa parte do estudo de Tennant et al. (2017), uma visão abrangente do panorama de inovações emergentes no sistema de revisão por pares no contexto da Web 2.0. Fundamenta-se, também, no conceito de sistema sociotecnológico de Fuchs (2005), no qual a emergência de propriedades decorre da colaboração dinâmica humano-máquina, bem como na tipologia de tarefas intensivas em conhecimento de Schreiber et al. (2000).

As próximas seções aprofundam a conceituação de sistema sociotecnológico, a descrição da tipologia de Schreiber et al. (2000) e uma breve discussão das inovações emergentes e futuras na revisão por pares estudadas por Tennant et al. (2017). Na sequência, apresentam-se os procedimentos metodológicos, os resultados da pesquisa e uma discussão sobre o caráter e as implicações das mudanças, seguidos por considerações finais.

2 A EMERGÊNCIA DE SISTEMAS COLABORATIVOS HUMANO-MÁQUINA

O reconhecimento das dificuldades de lidar com questões sociais no âmbito do desenvolvimento tecnológico levou ao conceito de “sistema sociotécnico”, na década de 1970. Churns (1976) propôs modelar as relações técnico-sociais e Bostrom e Heinen (1977) estudaram sistemas de informação para gestão (*management information systems*) sob uma perspectiva sociotécnica.

A abordagem sociotécnica representou avanço, mas permaneceu dualista. Sawyer e Crowston (2004) apontam o reconhecimento do caráter sociotécnico dos sistemas de informação na comunidade acadêmica, mas sem uma teorização efetiva sobre os arranjos, relações e elementos do relacionamento social-técnico.

Alter (2003) apontou a visão tecnocêntrica predominante, mesmo na comunidade “sociotécnica” de sistemas de informação. Propôs substituir o tecnocentrismo, o foco no artefato, por uma concepção de sistemas de informação como sistemas de trabalho dependentes de tecnologias de informação e comunicação.

Alguns cientistas reconheceram a insuficiência de apenas reconhecer a interdependência entre os subsistemas social e técnico e adotaram a designação “sociotecnológicos”, que enfatiza o caráter misto dos sistemas de informação. Para Eccles e

Groth (2006), sistemas sociotecnológicos são resolvedores de problemas compostos por pessoas e agentes tecnológicos em colaboração dinâmica. Esses agentes tecnológicos podem desempenhar tarefas simples, como ligar luzes conforme as condições ambientais, mas também tarefas críticas, como mudar uma rota aérea em função da aproximação de tráfego (ECCLES; GROTH, 2006).

O sentido do termo “sociotecnológico” usado por Eccles e Groth foi primeiramente mencionado por Fuchs (2005), que apresenta a internet como sistema sociotecnológico, já que só funciona por meio de ações humanas intensivas em conhecimento (*knowledgeable human activities*) que resultam em conhecimento humano objetificado (*objectified human knowledge*). Esta ideia de “funcionamento” complexo, não-mecanístico, parece compatível com a emergência da economia colaborativa típica da Web 2.0 (O’REILLY, 2007), em que pessoas produzem coletivamente, mediadas por *software*.

3 TAREFAS INTENSIVAS EM CONHECIMENTO DESEMPENHÁVEIS POR SOFTWARE

Existe uma “lacuna de comunicação proverbial” entre peritos em tecnologia e peritos em negócios (KILOV; SACK, 2009), o que restringe a apropriação da tecnologia nos negócios. Há pouca pesquisa sobre essa lacuna, mas parece verdadeiro que tecnólogos frequentemente repelem a comunicação ou são incapazes de documentar adequadamente, ao passo que outros profissionais raramente entendem as possibilidades da tecnologia em geral e, particularmente, os recursos da engenharia do conhecimento, dedicada a desenvolver “sistemas de informação nos quais o conhecimento e o raciocínio desempenham papéis essenciais” (SCHREIBER et al., 2000, p. ix).

O projeto europeu KADS (Aquisição de Conhecimento e Estruturação de Documentação, do inglês *Knowledge Acquisition and Documentation Structuring*) representou um avanço na superação da lacuna comunicacional e do tecnocentrismo na informática. Durou quinze anos e culminou na publicação do livro “*Knowledge engineering and management: the CommonKADS methodology*” (SCHREIBER et al., 2000). Apesar de um grande sucesso acadêmico em termos de citações¹, seus autores parecem haver-se desinteressado desse aspecto, possivelmente voltando-se para o mercado e a consultoria.

1 2.318 citações no Google Scholar em 12/08/2017, https://scholar.google.com.br/citations?view_op=view_citation&hl=en&user=IRZKdMYAAAAJ&citation_for_view=IRZKdMYAAAAJ:t6usbXjVLHcC.

Antes de CommonKADS, existia uma engenharia do conhecimento tecnocêntrica, focada em “transferir o conhecimento” da mente de um especialista humano para um sistema (*software*) especialista, usualmente na forma de regras de produção (*modus ponens*; regras na forma “se A então B”). CommonKADS inovou ao mover o foco da tecnologia para o negócio, o sistema, a organização. Assim, um projeto em CommonKADS pode resultar na decisão de **não** construir *software*.

Um projeto CommonKADS inicia por questionar o que é a organização que se pretende apoiar com técnicas computacionais inteligentes. Só após analisar o contexto da organização e do(s) problema(s) em foco é que se passa a elaborar modelos de conhecimento que possam apoiar a automação de tarefas intensivas em conhecimento. A expressão “intensivo em conhecimento” é “intencionalmente vaga” (SCHREIBER et al., 2000, p. ix), dada a dificuldade de distinguir inequivocamente tarefas ricas na aplicação de conhecimento de tarefas burocráticas mais simples.

Sistemas de *software* comuns, desenvolvidos a partir de métodos da engenharia de *software*, tratam de automatizar tarefas burocráticas que, mesmo que complexas, não demandam aplicação de conhecimento apurado. Os tipos de tarefas intensivas em conhecimento automatizáveis, descritos por Schreiber et al. (2000), são:

a) Tarefas de **análise**:

- a. **Classificação**: Analisar características e determinar a qual classe um objeto pertence dentre um conjunto (disjunto) de classes. É tarefa usual na taxonomia (biológica).
- b. **Diagnóstico**: Identificar um defeito no funcionamento de um sistema, considerando um padrão de funcionamento normal. É tarefa usual em sistemas técnicos.
- c. **Avaliação** (*assessment*): Caracterizar um caso em termos de uma classe de decisão. Um exemplo é a caracterização de um pedido de empréstimo como “sim/não/mais informação necessária”, tipicamente a partir de critérios dados.
- d. **Monitoramento**: Analisar um sistema dinâmico quanto à ocorrência de discrepâncias. Em cada ciclo de monitoramento, verifica-se o estado de funcionamento, como no caso de um projeto de *software*: as entregas estão em dia?

e. **Predição:** Analisar o comportamento de um sistema e prever seu estado em algum momento futuro. É considerada uma tarefa analítica com algumas características sintéticas. Um exemplo é a previsão do tempo.

b) Tarefas de **síntese:**

a. **Alocação** (*assignment*): Elaborar um mapeamento parcial entre dois conjuntos de objetos, respeitando certas restrições. A alocação de aeronaves a portões de embarque é um exemplo.

b. **Projeto** (*design*): Desenhar as características de um artefato projetado. Considera-se que a tecnologia atual ainda é precária quanto à criatividade requerida, portanto assume-se que todos os componentes do artefato estão predefinidos e a tarefa automatizável é um subtipo da tarefa genérica de projeto. Trata do projeto de configuração do sistema, como na escolha da configuração de um computador.

c. **Planejamento:** Desenhar as características de processos – atividades e suas dependências de tempo. De forma análoga ao projeto de artefatos, o planejamento de atividades só é possível com a tecnologia atual se os elementos básicos do plano estão predefinidos, como na automação do planejamento de viagens ou atividades de construção.

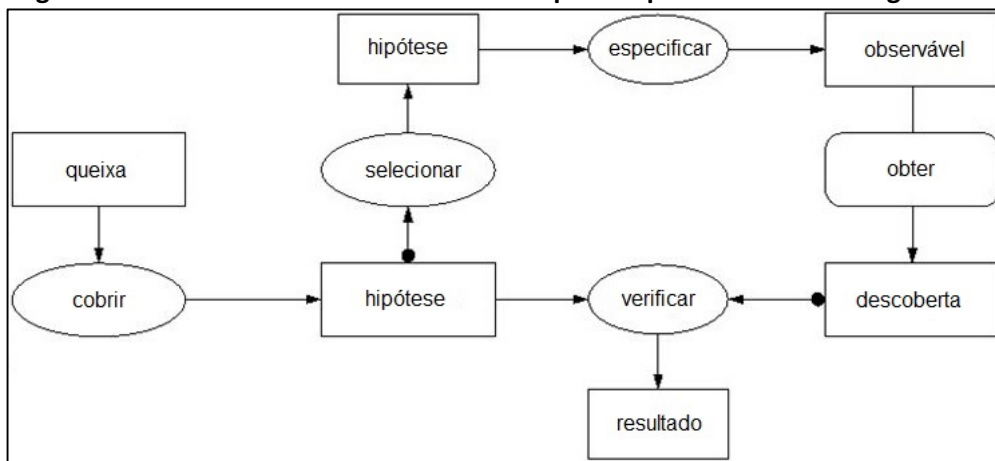
d. **Programação** (*scheduling*): Alocar atividades a recursos em certo intervalo de tempo. A programação frequentemente sucede o planejamento e resulta num mapeamento entre atividades e intervalos de tempo, como em “Cálculo I (semestre 1) precede Física II (semestre 2)” num currículo de curso de engenharia.

e. **Modelagem:** Construir uma descrição abstrata de um sistema. Os autores reconhecem que a modelagem é raramente automatizada, mas incluem esta categoria para completar o conjunto de tarefas intensivas em conhecimento (embora não discutam essa completude).

Schreiber et al. (2000) apresentam, ainda, um catálogo de modelos (*templates*) de tarefas intensivas em conhecimento com o objetivo de facilitar sua programação e automação. Cada modelo contém uma caracterização geral da tarefa definindo objetivo, entrada, saída e terminologia, um método-padrão que pode ser adaptado para situações particulares, um repertório de variações usuais do método-padrão e um esquema típico do domínio de conhecimento do problema.

O modelo de tarefa intensiva em conhecimento do tipo **Diagnóstico** detalhado por Schreiber et al. (2000, p. 138-143) é apresentado aqui como exemplo. A estrutura de inferência para o método-padrão de diagnóstico é ilustrada na Figura 1. As elipses correspondem a inferências, os retângulos de cantos vivos, a dados ou afirmações e o retângulo de cantos chanfrados, a uma função de transferência.

Figura 1: Estrutura de inferência do método padrão para a tarefa de Diagnóstico.



Fonte: Schreiber et al. (2000, p. 140, tradução nossa).

A inferência **cobrir** busca todas as potenciais causas ou hipóteses para a **queixa** através de uma cadeia causal, gerando um conjunto exaustivo de **hipóteses**. A inferência **selecionar** submete, uma por vez, cada **hipótese** à especificação (**especificar**) de alguma entidade **observável** cujo valor pode ser usado para limitar o número de hipóteses, podendo confirmar a presença desta hipótese ou eliminar outras hipóteses. A função **obter** transfere o valor **observável** usado para testar as falhas candidatas para um conjunto de **descobertas**. O método trata de **verificar** se a **hipótese** elencada entre as **descobertas** deve ser refutada ou mantida entre as potenciais causas. O laço **selecionar-especificar-obter-verificar** é aplicado a todas as hipóteses até chegar a um **resultado** que pode ser um conjunto de hipóteses vazio (as evidências são inconsistentes com as falhas conhecidas pelo sistema), único (solução única) ou múltiplo (o sistema não é capaz de diferenciar entre as hipóteses remanescentes).

Apesar das duas décadas desde o fim do projeto KADS, a abordagem metodológica e os modelos de tarefas intensivas em conhecimento da metodologia CommonKADS continuam atuais ou, pelo menos, não foram superados por outras técnicas. Dessa forma, parecem aplicáveis a novas demandas como as que vêm das inovações emergentes na revisão por pares editorial, abordadas a seguir.

4 INOVAÇÕES EMERGENTES DA WEB 2.0 PARA O SISTEMA DE REVISÃO POR PARES

Yarkoni (2012) examinou plataformas da Web 2.0 ou Web Social como Amazon, Google News, Grooveshark, Last.fm, Netflix, Reddit e Stack Exchange em busca de lições aplicáveis à revisão por pares. Apontou limitações da avaliação científica tradicional que poderiam ser melhoradas com o uso de técnicas dessas novas plataformas: vagareza e ineficiência, opacidade, baixa confiabilidade e falta de incentivo aos revisores. Concluiu recomendando o uso de práticas da Web social para uma revisão por pares modernizada: extração de dados desde repositórios e mecanismos de busca acadêmicos, filtragem colaborativa de conteúdos, comentários enfileirados (*threaded commenting*), sistema de reputação pontuada, categorização por tópicos usando etiquetas (*tags*), hierarquização dinâmica (*ranking*) de conteúdos e usuários, acesso via protocolos de interface de aplicação (API, do inglês *application protocol interface*).

Tennant et al. (2017) retomaram a ideia de Yarkoni analisando a funcionalidade de plataformas da Web social, com a expectativa de que, quaisquer que fossem, teriam desempenho superior ao do sistema de revisão por pares atual e evitariam muitos dos vieses existentes. Compararam as funcionalidades das plataformas da Web social – Amazon, Blockchain, GitHub, Hipotesis, Reddit, Stack Exchange e Wikipedia – com as necessidades da revisão por pares: controle de qualidade, métricas de desempenho de revisores quantificadas como incentivos ao engajamento, bem como sistemas de certificação e reputação.

Os autores consideravam que, em vez de propor uma inovação radical, com uma plataforma ou modelo de revisão por pares totalmente novo, seria melhor “considerar as vantagens e desvantagens de modelos e inovações existentes em serviços e tecnologias sociais” (TENNANT et al., 2017, p. 19). Sua análise encontrou as características favoráveis e desfavoráveis à aplicação na revisão por pares sumariadas na Tabela 1. É interessante observar que todas as características listadas na primeira coluna correspondem a tarefas burocráticas, não intensivas em conhecimento.

Tabela 1: Prós e contras potenciais das principais características dos modelos emergentes de revisão por pares.

Característica	Descrição	Prós	Contras/riscos	Modelos existentes
Votação ou classificação (<i>rating</i>)	Avaliação quantitativa de revisão (5 estrelas, pontos), incluindo votações para cima e para baixo	Filtro de qualidade orientado pela comunidade, simples e eficiente	Procedimento aleatório, autopromoção, manipulação (<i>gaming</i>), viés de popularidade, não estático	Reddit, Stack Exchange, Amazon

Característica	Descrição	Prós	Contras/riscos	Modelos existentes
Abertura (<i>openness</i>)	Visibilidade pública do conteúdo da revisão	Responsabilidade, responsabilidade, contexto, maior qualidade	Pressão dos pares, potencialmente menor qualidade, convida à retaliação	Todos
Reputação	Avaliação e classificação (<i>rating</i>) (pontos, estatísticas de revisões)	Filtro de qualidade, recompensa, motivação	Desequilíbrio baseado no status do usuário, encoraja manipulação, específico da plataforma	Stack Exchange, GitHub, Amazon
Comentário público	Comentários visíveis em artigo/revisão	Artigo vivo / orgânico, envolvimento da comunidade, progressivo, inclusivo	Propenso ao assédio, demorado, não interoperável, baixo reuso	Reddit, Stack Exchange, Hipotesis
Controle de versão	Versões e configurações gerenciadas	Objetos vivos / orgânicos, verificáveis, progressivos, bem organizados	Rastreamento de citações, demora, baixa confiança no conteúdo	GitHub, Wikipedia
Incentivos (<i>incentivization</i>)	Incentivo para se envolver com plataforma e processo através de "broches" (<i>badges</i>) / dinheiro ou reconhecimento	Motivação, retorno do investimento	Monetização de pesquisa, possível perversão por ganância, alto custo	Stack Exchange, Blockchain
Autenticação e certificação	Filtragem de contribuidores via processo de verificação	Controle de fraude, proteção ao autor, estabilidade	<i>Hacking</i> , difícil de gerenciar	Blockchain
Moderação	Filtragem de comportamento inadequado em comentários, classificação (<i>rating</i>)	Filtro de qualidade, orientado pela comunidade	Censura, discurso convencional (<i>mainstream</i>)	Reddit, Stack Exchange

Fonte: Tennant et al. (2017, p. 20, tradução nossa).

O objetivo declarado do estudo de Tennant et al. (2017, p. 4) foi “investigar a evolução histórica na teoria e na aplicação da revisão por pares num contexto sociotecnológico”. Os autores consideram que “tem havido inovação tecnológica considerável na revisão por pares na última década, o que está levando ao seu exame crítico como processo social” (p. 19).

O termo “sociotecnológico”, um neologismo, não foi definido, mas seu uso se ajusta às concepções de Fuchs (2005) e de Eccles e Groth (2006), de sistemas resolvidores de problemas a partir da colaboração dinâmica humano-máquina. Essa colaboração pode dar-se em tarefas burocráticas como as elencadas na Tabela 1, mas também em tarefas intensivas em conhecimento. Nesse sentido, a próxima seção traz os procedimentos metodológicos para identificar oportunidades de resolução de problemas de caráter intensivo em conhecimento por meio de *software* na revisão por pares.

5 PROCEDIMENTOS METODOLÓGICOS

Esta é uma pesquisa bibliográfica que tem como fonte (*corpus*) o estudo de Tennant et al. (2017) sobre inovações futuras e emergentes na revisão por pares editorial. É pressuposto da pesquisa que o material fonte é uma visão abrangente sobre o cenário de crise e mudança na revisão por pares, uma vez que os autores são um grupo de jovens líderes bastante envolvidos na discussão on-line sobre a comunicação científica.

O primeiro autor é paleontólogo e diretor de comunicação da ScienceOpen (www.scienceopen.com), uma plataforma *on-line* para busca e anotação de conteúdos voltada à **ciência aberta**, expressão que implica em acesso amplo e sem custos ou restrições à literatura científico-tecnológica (PINHEIRO, 2014). Outros autores são nomes conhecidos no ciberespaço científico, como o biomédico Cameron Neylon, editor do blog Science in the Open e o biofísico Daniel Mietchen, cocriador do periódico RIO (Research Ideas and Outcomes), dedicado a publicar todas as etapas do processo científico, incluindo ideias, projetos e resultados.

Os dados coletados são os trechos do texto-fonte que denotam funções ou tarefas desempenháveis por *software* no processo de revisão por pares. Foi feita uma primeira leitura do texto buscando identificar trechos que enfocam inovações efetivadas por agentes tecnológicos. Essa leitura foi apoiada por uma busca no texto pelos seguintes afixos, na expectativa de encontrar menção a tarefas automatizadas ou semiautomatizadas: ‘algorithm’, ‘autom’, ‘count’, ‘data’, ‘machine’, ‘measur’, ‘min’ (para termos relacionados a *mining*), ‘quant’, ‘software’, ‘system’, ‘techn’, ‘text’ e ‘tool’.

Os trechos identificados foram verificados quanto ao caráter das tarefas envolvidas: intensivas em conhecimento ou não-intensivas, ou seja, burocráticas, administrativas. Após encontrar os primeiros resultados, o estudo foi limitado às tarefas intensivas em conhecimento, por motivos de espaço.

A análise tratou de relacionar as tarefas encontradas às categorias de tarefas intensivas em conhecimento de Schreiber et al. (2000). Fica excluída do escopo a análise aprofundada das tarefas burocráticas, possivelmente usando a tipologia de funções de Tennant et al. (2017) representada na primeira coluna da Figura 1.

Também foram excluídos da análise os trechos que denotam inovações sem especificar precisamente de qual função ou tarefa se trata. Essas menções vagas, como “Foram levantadas questões [...] como a possibilidade de autores mudarem sua abordagem quando sabem que seu

artigo está sendo avaliado por uma máquina”² (TENNANT et al., 2017, p. 28), foram desconsideradas. A próxima seção apresenta os resultados.

6 RESULTADOS

Foram identificados 56 trechos do texto-fonte que denotam o desempenho de tarefas da revisão por pares automatizável por *software*. Desses, 44 referem-se a tarefas burocráticas como contagens, ordenações, acionamento automático de funções por gatilho e suporte ao acionamento humano de funções burocráticas. Dessa forma, restaram doze trechos sobre inovações emergentes na revisão por pares implementáveis em *software* como tarefas intensivas em conhecimento.

Cabe ressaltar que os tipos de tarefa intensiva em conhecimento têm os significados precisos definidos na Seção 3. Tarefas de “avaliação”, “diagnóstico” etc. foram identificadas considerando essas definições precisas. Os trechos de texto e as categorias de tarefa intensiva em conhecimento envolvidas são descritos no Quadro 1.

Quadro 1: Tarefas intensivas em conhecimento automatizáveis na revisão por pares

Inovações emergentes e potenciais indicadas no texto	Tarefa intensiva em conhecimento
p19 1 No Reddit, postagens altamente pontuadas são relativamente efêmeras, com um algoritmo automático de depreciação que as empurra mais abaixo nas listas à medida que novos conteúdos são adicionados.	Monitoramento.
p22 2 Gashler ... propôs um sistema descentralizado e protegido, onde cada revisor assinaria digitalmente cada artigo ... Essa rede de revisores e artigos poderia ser minerada (data mined) para revelar informação sobre a influência e conectividade de pesquisadores individuais na comunidade científica.	Modelagem.
p26 3 A tecnologia blockchain poderia ser alavancada para criar um sistema de revisão por pares tokenizado envolvendo penalidades para os membros que não respeitam os padrões adotados e vice-versa.	Diagnóstico.
p27 4 [A]prendizado de máquina (ML [do inglês machine learning]) e ferramentas de redes neurais ... A ML foi amplamente adotada para abordar vários desafios, incluindo reconhecimento de imagem , recomendação de conteúdo, detecção de fraude e otimização de energia ... em particular para detecção de plágio .	Classificação, diagnóstico.
p27 5 [O] uso de automação baseada em computador para revisão por pares ... três principais aplicações práticas. A primeira é determinar se um artigo em consideração atende aos requisitos mínimos ... A segunda abordagem é determinar automaticamente os revisores mais adequados para um manuscrito enviado ... Finalmente, ... ML [aprendizado de máquina] e redes neurais ... podem ter alguma forma de utilidade preditiva na identificação de resultados novos durante a revisão pelos pares. Nesse caso, o aprendizado da máquina seria usado para prever o impacto futuro de um determinado trabalho.	Avaliação, alocação, predição.

²Tradução nossa para “Questions have been raised [...] such as how authors may change their approach when they know their manuscript is being evaluated by a machine”.

Inovações emergentes e potenciais indicadas no texto	Tarefa intensiva em conhecimento
p27 6 [O] aprendizado de máquina seria usado para prever o impacto futuro de um determinado artigo (por exemplo, contagens futuras de citações) e, de fato, para fazer o trabalho de análise de impacto e tomada de decisão em vez de ou junto com um crítico humano.	Predição.
p27 7 [A editora] Elsevier usa um sistema chamado Evisse ... para verificar plágio, recomendar revisores e verificar informações de perfil de autor via conexão com [a base de dados] Scopus.	Diagnóstico, alocação.
p27 8 [O] [periódico] Journal of High Energy Physics usa a alocação automática de editores com base em um algoritmo orientado por palavras-chave.	Alocação.
p28 9 Foram levantadas questões ... tais como ... de que forma a avaliação por máquina poderia descobrir autoridade infundada em afirmações de autores através da análise de redes de citações.	Diagnóstico.
p28 10 Iniciativas como Meta, uma ferramenta de IA [inteligência artificial] que faz buscas em artigos científicos para prever a trajetória da pesquisa (meta.com), destaca a grande promessa de inteligência artificial na pesquisa e para aplicação na revisão por pares.	Predição.
p30 11 [A plataforma] ORCID tem um papel potencial adicional de criar a possibilidade de um arquivo público de informações e metadados de pesquisadores ... para combinar revisores potenciais com itens de interesse [artigos], evitando conflitos de interesse.	Alocação.
p30 12 [O] controle de qualidade é provido tendo um sistema de revisão por convite automatizado , mas gerenciado, interação pública e colaboração.	Alocação.

Fonte: Tennant et al. (2017, tradução e destaques nossos).

Há menções a todos os tipos de tarefa analítica e a dois tipos de tarefa sintética: alocação e modelagem – sem menções, portanto, às tarefas sintéticas de projeto, planejamento e programação. Verifica-se que alguns trechos distintos discutem uma mesma tarefa, bem como há um trecho que discute instâncias distintas do mesmo tipo de tarefa. Dessa forma, pode-se contabilizar os tipos de tarefas citados como alocação (5), avaliação (1), classificação (1), diagnóstico (4), modelagem (1), monitoramento (1) e predição (3), conforme a enumeração a seguir:

- a) **Alocação** (itens 5, 7, 11 e 12 no Quadro 1) de revisores ou (item 8) de editores a artigos.
- b) **Avaliação** (item 5) do cumprimento de requisitos em um artigo.
- c) **Classificação** (item 4) de imagens a partir de técnicas de aprendizado de máquina.
- d) **Diagnóstico** (item 3) do desrespeito de pesquisadores a padrões de conduta em uma rede ou plataforma, (item 4) da ocorrência de fraude ou (itens 4 e 7) de plágio em artigos, bem como (item 9) da falta de fundamento numa evidência baseada em citação.

- e) **Modelagem** (item 2) da rede de influência e conectividade de pesquisadores, o que requer uma concepção predeterminada de como representar esse tipo de modelo, com pesquisadores como nós e relações de influência e conectividade como arestas.
- f) **Monitoramento** (item 1) da entrada de novos artigos (no exemplo do Reddit, trata-se de postagens) de forma a corrigir discrepâncias de atualização que levam à concentração excessiva de atenção, dirigindo a atenção da comunidade aos novos itens e depreciando o destaque dado a itens mais antigos.
- g) **Predição** (item 5) da existência de resultados científicos novos em artigos por meio de aprendizado de máquina ou redes neurais, (item 6) da contagem futura de citações a serem recebidas por um artigo e (item 10) dessas e outras predições possíveis sobre o futuro da ciência.

7 DISCUSSÃO

Há uma “agenda positiva” em parte das tarefas intensivas em conhecimento identificadas: algumas meramente agilizam o processo, como a alocação semiautomática de revisores, enquanto outras representam novidades antes indisponíveis, como a predição do grau de novidade científica. Outra parte das tarefas identificadas corresponde a uma “agenda negativa”, ou tentativa de anular os malfeitos que parecem tornar-se cada vez mais comuns na publicação científica. Enquadram-se aí todos os casos de diagnóstico – do desrespeito a padrões, dificilmente produto de má-fé, às detecções de fraude, plágio ou uso negligente ou mal-intencionado, de qualquer forma errôneo, de citações para fundamentar afirmações.

Nem todos os erros levam ao recolhimento ou retratação de artigos publicados, mas Deniszczwicz e Kern (2016) mostraram evidências de que a fonte dos erros que levam à retratação de artigos publicados é majoritariamente o autor, com editores, revisores e publicadores desempenhando papel tipicamente complementar. Isso é exemplificado pelas revisões falsas (*fake reviews*), em que um autor fraudula a boa-fé do sistema ao simular (impersonar) a existência de um revisor. O editor é induzido a erro ao aceitar o parecer do “revisor”. O limite entre ser vítima ou negligente é tênue, por não verificar competentemente a existência do revisor, mas a disponibilidade de *software* para apoiar a detecção do problema seria muito bem-vinda.

Além da detecção de revisores falsos, não mencionada por Tennant et al. (2017), não é difícil imaginar outras tarefas intensivas em conhecimento aplicáveis à revisão por pares.

Crítérios de legibilidade como os índices de Flesch e de Kincaid (KINCAID et al., 1975), por exemplo, presentes até em ferramentas usuais como o editor MS-Word, poderiam ser usados na classificação de textos em níveis de legibilidade. Até tarefas burocráticas simples, como a verificação de uma lista de referências contra as citações feitas no texto, representariam avanço no precário ecossistema tecnológico da revisão por pares.

Quanto às transições observáveis no mundo do *software*, a primeira é a lenta evolução do papel das iniciativas de desenvolvimento: da mera virtualização, informatização, automação de escritórios na qual um **ambiente** físico e seus processos são transformados em *software*, para uma nova conformação de sistemas sociotecnológicos (FUCHS, 2005) na qual alguns *softwares* são **agentes** do sistema, desempenhando tarefas em colaboração dinâmica com pessoas. Essa transição tem sido muito lenta, talvez devido à já citada lacuna “proverbal” de comunicação entre peritos em tecnologia e nos negócios afetados por *software* (KILOV; SACK, 2009). Faltam métodos de desenvolvimento de *software* com essas características, mas CommonKADS (SCHREIBER et al., 2000) foi um primeiro avanço nesse sentido, ao preocupar-se com o contexto do “sistema” (i.e., *software*) em vez de apenas com a tecnologia.

A segunda transição observável é o arrefecimento do ânimo automatizador, tecnocêntrico, em benefício de uma visão de *software* como apoiador da tomada de decisão humana. Não se trata mais de substituir o trabalho humano, mas de apoiá-lo via automação sensata.

Essa mudança, plenamente compatível com a ideia de sistema sociotecnológico, é defendida no estudo de Tennant et al. (2017). Os autores afirmam que as novas plataformas de revisão por pares não podem simplesmente oferecer *software* e esperar que os pesquisadores o usem, mas devem priorizar o engajamento da comunidade científica e a compreensão das necessidades e processos culturais de diferentes disciplinas.

É preciso reconhecer duas classes de limitações nesta pesquisa: uma específica sobre os fundamentos adotados e outra genérica, sobre a abordagem defendida aqui, da automação não-indiscriminada, i.e., da construção de sistemas sociotecnológicos. Os fundamentos adotados são limitados porque o estudo de Tennant et al. (2017) não é exaustivo, já que esta mesma Discussão aponta oportunidades de automação de tarefas intensivas em conhecimento ausentes no artigo-fonte. Por outro lado, também a tipologia de tarefas intensivas em conhecimento de Schreiber et al. (2000) é criticável, já que os autores sugerem ser “completa”

mas não oferecem evidência nesse sentido. De qualquer forma, postula-se aqui a inexistência de outros fundamentos mais robustos para estudar o que foi proposto.

Quanto às limitações mais amplas, da abordagem da automação de tarefas da revisão por pares, é preciso evitar a aceitação acrítica da inovação tecnológica, típica do tecnocentrismo. Nesse sentido, Tennant et al. (2017) apresentam uma visão não-ingênua das inovações. Reconhecem que a aprendizagem de máquina e as redes neurais “são amplamente consideradas conformistas” (p. 28) e que os sistemas de *rankings* usados nas plataformas Web 2.0 que analisaram medem “popularidade e não qualidade” (p. 22), o que se opõe aos objetivos da ciência.

Reconhecem, também, que as métricas de reputação resumidas a uma fórmula expõem todo o sistema à manipulação (*gaming*), “como em quase todo processo que possa ser medido e quantificado” (p. 22). Implantar inovações acriticamente levaria a novos incentivos disfuncionais à publicação, que podem agravar os problemas atuais de carreirismo e imediatismo que degradam as possibilidades de progresso da ciência.

Este artigo concentrou-se nas inovações tecnológicas na revisão por pares na forma de tarefas intensivas em conhecimento. Uma investigação semelhante voltada às tarefas burocráticas, não-intensivas em conhecimento, parece pesquisa futura meritória. Dos 56 trechos de texto coletados do texto-fonte, 44 diziam respeito a tarefas burocráticas, algumas de bastante utilidade imediata, como a integração de informações por meio do identificador ORCID (<https://orcid.org/>) e a computação de métricas de desempenho do processo de revisão por pares como tempo mediano de avaliação. São métricas que poderiam ser públicas, beneficiando todo o ecossistema de avaliação científica.

8 CONSIDERAÇÕES FINAIS

O artigo contém uma análise do estudo de Tennant et al. (2017) sobre inovações emergentes e futuras na revisão por pares. O foco foi a identificação de tarefas intensivas em conhecimento implementáveis em *software*. As tarefas identificadas foram classificadas conforme a tipologia de Schreiber et al. (2000) para tarefas intensivas em conhecimento.

As principais conclusões são que a automação ou semiautomação de tarefas intensivas em conhecimento pode contribuir para evitar ou amenizar vários problemas atuais da revisão por pares editorial. A automação total ou parcial não pode ser acrítica. Deve considerar os potenciais impactos no sistema de comunicação científica, de acordo com o conceito de sistema

sociotecnológico (FUCHS, 2005), no qual agentes de *software* desempenham tarefas em colaboração dinâmica com pessoas.

Também foram discutidas duas transições no papel do *software* em geral e, em particular, na revisão por pares. Primeiramente, o *software* vem ganhando papel de agente resolvidor de problemas, em contraste com a mera virtualização ou informatização de ambientes e processos. Em segundo lugar, o ânimo tecnocêntrico usual de substituir o trabalho humano pela automação parece perder espaço para um novo papel de apoio ao trabalho humano.

REFERÊNCIAS

- ALTER, Steven. 18 reasons why IT-reliant work systems should replace “the IT artifact” as the core subject matter of the IS field. **Communications of the AIS**, v. 12, art. 23, p. 366-395, 2003.
- ALVAREZ, Gonzalo Rubén; CAREGNATO, Sônia Elisa. Preprints na comunicação científica da Física de Altas Energias: análise das submissões no repositório arXiv (2010-2015). **Perspectivas em Ciência da Informação**, v. 22, n. 2, p. 104-117, 2017.
- BOSTROM, Robert P.; HEINEN, J. Stephen. MIS problems and failures: A socio-technical perspective. Part I: The causes. **MIS Quarterly**, v. 1, n. 3, p. 17-32, 1977.
- CHERNS, Albert. The principles of sociotechnical design. **Human Relations**, v. 2, n. 9, p. 783-792, 1976.
- DENISZWICZ, Marta; KERN, Vinícius Medina. Fontes dos problemas na revisão por pares que levam à retratação de artigos divulgados no Retraction Watch. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais...** Salvador: UFBA, 2016. p. 3447-3466.
- ECCLES, D. W.; GROTH, P. T. Agent coordination and communication in sociotechnological systems: Design and measurement issues. **Interacting with Computers**, v. 18, n. 6, p. 1170-1185, 2006.
- FUCHS, Christian. The internet as a self-organizing socio-technological system. **Cybernetics and Human Knowing**, v. 12, n. 3, p. 57-81, 2005.
- KILOV, Haim; SACK, Ira. Mechanisms for communication between business and IT experts. **Computer Standards & Interfaces**, v. 31, n. 1, p. 98-109, 2009.
- KINCAID, J. Peter; FOSHBURNE, Robert P.; ROGERS, Richard L.; CHISSOM, Brad S. **Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel**. Naval Technical Training Command Millington TN Research Branch, 1975. 40 p. Disponível em: <<http://www.dtic.mil/get-tr-doc/pdf?AD=ADA006655>>. Acesso em: 13 ago 2017.

O'REILLY, Tim. What is Web 2.0: Design patterns and business models for the next generation of software. **Communications & Strategies**, n. 1, p. 17-37, 2007.

PINHEIRO, Lêna Vania Ribeiro. Free access to open science: concepts and implications for science communication. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 8, n. 2, 2014.

RENNIE, Drummond. Make peer review scientific: thirty years on from the first congress on peer review, Drummond Rennie reflects on the improvements brought about by research into the process--and calls for more. **Nature**, v. 535, n. 7610, p. 31-34, 2016.

RESEARCH INFORMATION NETWORK. **Activities, costs and funding flows in the scholarly communications system in the UK**: Report commissioned by the Research Information Network (RIN). 2008. 88 p. Disponível em: <http://www.rin.ac.uk/system/files/attachments/Activites-costs-flows-report.pdf>. Acesso em: 06 ago 2017.

SAWYER, Steve; CROWSTON, Kevin. Information systems in organizations and society: Speculating on the next 25 years of research. In: **Information systems research**. Boston: Springer, 2004. p. 35-52.

SCHREIBER, Guus; AKKERMANS, Hans; ANJEWIERDEN, Anjo; DE HOOG, Robert; SHADBOLT, Nigel; VAN DE WELDE, Walter; WIELINGA, Bob. **Knowledge engineering and management**: the CommonKADS methodology. Massachusetts: MIT Press, 2000. 455 p.

TENNANT, Jonathan P. et al. A multi-disciplinary perspective on emergent and future innovations in peer review [version 1; referees: awaiting peer review]. **F1000Research**, v. 6, n. 1151, 2017. Disponível em: <https://f1000research.com/articles/6-1151/v1>. Acesso em: 21 jul. 2017.

YARKONI, Tal. Designing next-generation platforms for evaluating scientific output: what scientists can learn from the social web. **Frontiers in Computational Neuroscience**, v. 6, article 72, 2012.

ZUCKERMAN, Harriet; MERTON, Robert K. Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. **Minerva**, v. 9, n. 1, p. 66-100, 1971.