

A background image showing two hands held palm up. The left hand holds a green apple, and the right hand holds a pink donut with white sprinkles. The background is a light grey, textured surface.

# Appraisal in the digital era

workshop

22 June 2017

BNP Paribas – Paris - France

*eabh* in cooperation with BNP Paribas and Banque Lombard Odier

***eabh***  
we do history

## **Appraisal strategies in central banks**

Decommissioning the network drives at the ECB:  
a pragmatic approach  
Maria Luisa Di Biagio (ECB)

Archival appraisal In the Croatian National Bank  
Egon Kraljević & Tvrtko Ujević (Croatian National Bank)

Document appraisal at Banco de Portugal  
António Gil Matos & Filipe Fernandes (Banco de Portugal)

## **Appraisal strategies in international organizations**

Archival appraisal in a digital context at the HAEU  
Samir Musa (Historical Archives of the European Union)

Application of the UNESCO/PERSIST guidelines for the  
selection of WBG social media content for long-term  
preservation  
Esther Y. Abangma & Paloma Beneito Arias (World Bank  
Group)

The e-Domec records management policy of the European  
Commission  
Seth van Hooland (Université libre de Bruxelles)

## **Appraisal strategies in commercial organisations**

Redefining the archival strategy at a multinational firm in the  
digital era  
Lourdes Fuentes Hashimoto (Total)

Achmea Holding appraisal policy from the perspective of the  
adopted maturity model  
Vincent Hoolt (Achmea Holding)

Creating a platform for digital appraisal at HSBC  
James Mortlock (HSBC)

## **Automatic appraisal in the era of GDPR**

Chris Winter (Hewlett Packard Enterprise)



**EUROPEAN CENTRAL BANK**

**EUROSYSTEM**

ECB-PUBLIC  
FINAL

## **Decommissioning the network drives at the ECB: A pragmatic approach**

**by Maria Luisa Di Biagio**

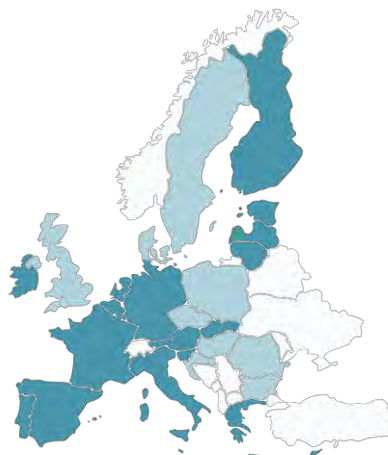
**EABH workshop „Appraisal in the  
digital age“ - Paris, 22 June 2017**



- **Migration of information from the network drives to the EDRMS**
  - **Background**
  - **Project organisation**
- **Criteria for the selection of the information to be migrated**
- **How we supported users**
- **Some reflections**
- **Conclusion**

**At the heart of the  
Eurosystem and the  
Single Supervisory  
Mechanism**

**~3.500 staff  
25 business areas**

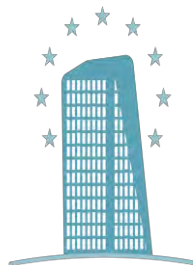


**Maintain the stability of the euro**

**Conduct monetary policy in the  
euro area**

**Conduct banking supervision  
on euro area banks**

**Ensure smooth functioning of  
payment systems**



## European Central Bank

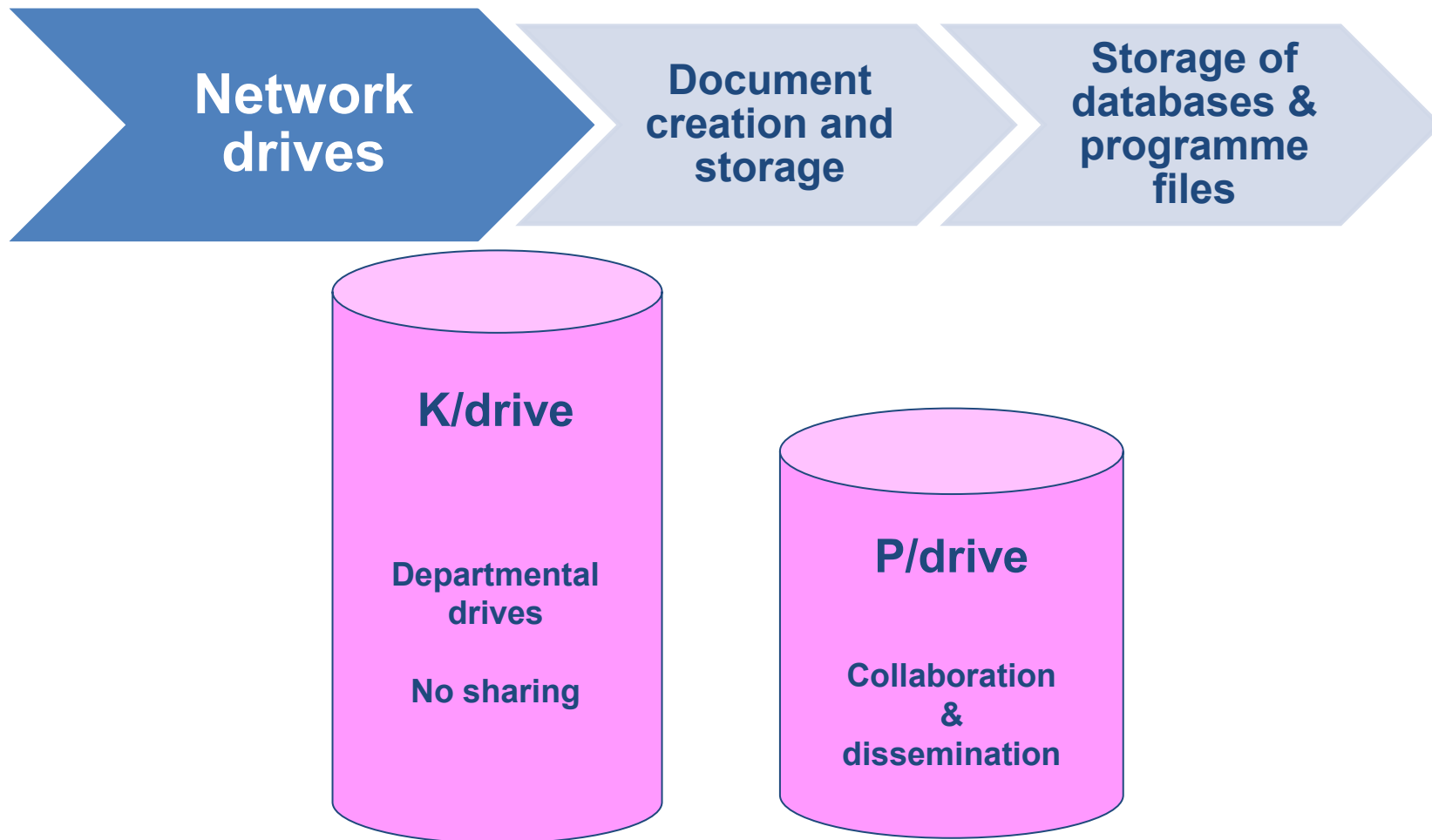
**Information is key to prepare  
decisions and perform our assigned  
tasks**

**A knowledge  
organization**



**EDRMS since 2007  
~10 million documents  
~18.000 users**

# Use of network drives before EDRMS (1998 – 2007/2009)



## Upon introduction of EDRMS

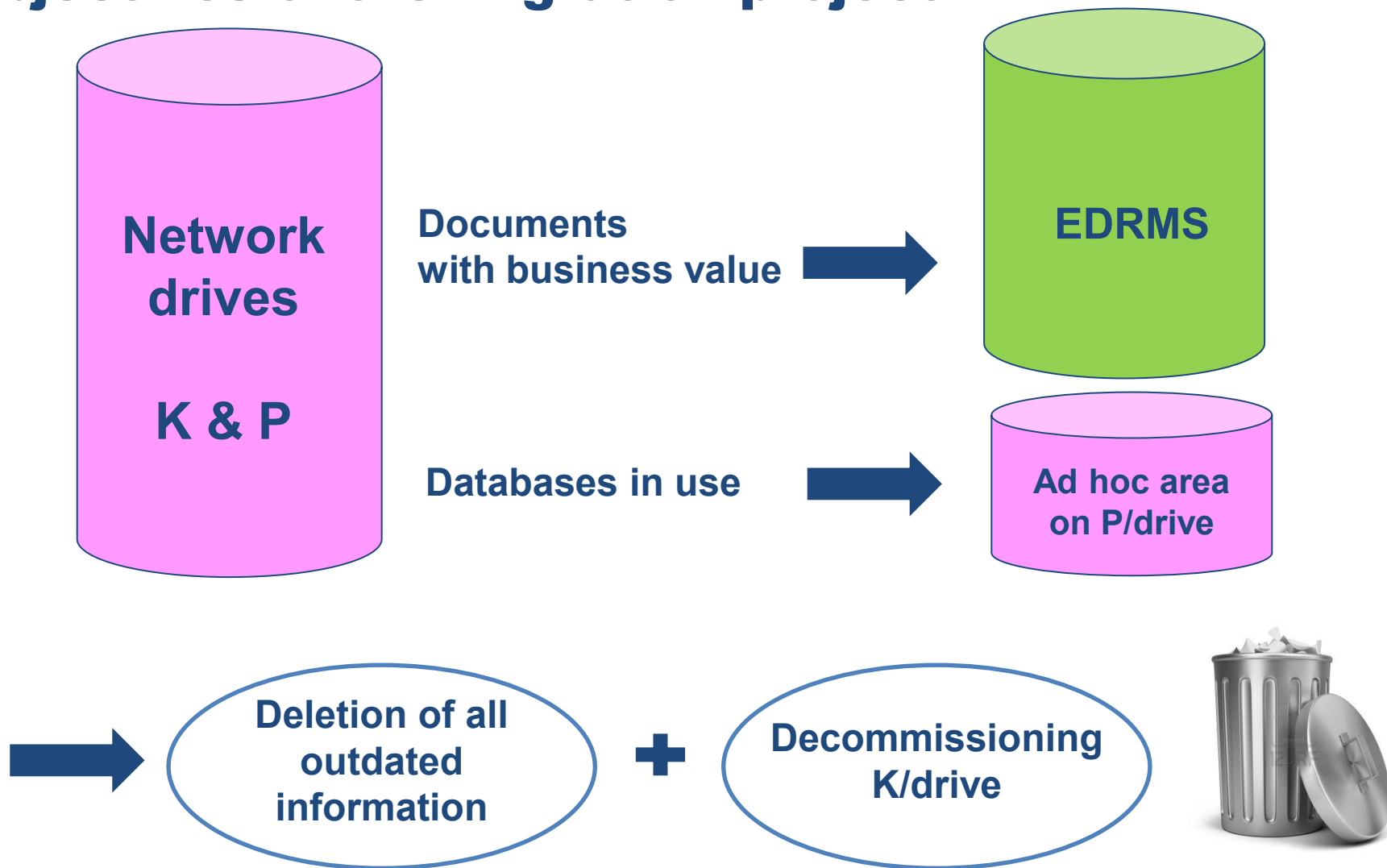
**Recommended freezing of K/drive to prevent further use**

- Not possible for all areas
- Some critical databases & programme files on K/drive

**Recommended migration to EDRMS of business information of relevance**

- Partially carried out for decision-making bodies, HR, accounting, legal documentation

# Objectives of the migration project





## Why migrating information from the network drives

**Impossibility to apply  
ECB IM policies**

- Retention
- Digital preservation

**Compliance risk**

- Increased time for public access requests
- Non compliance with data protection rules

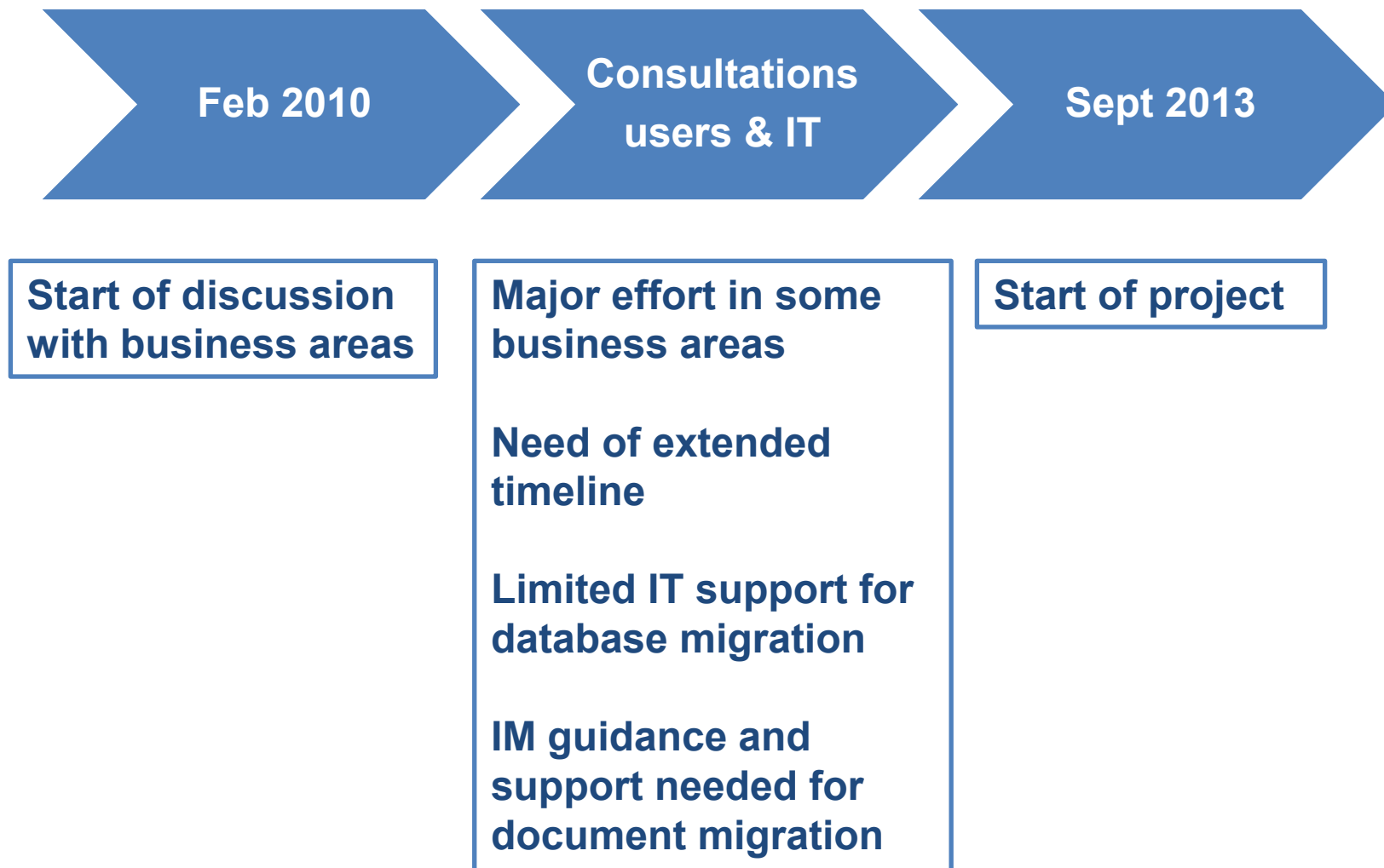
**Reduced information  
security**

- Non transparent access rights
- No audit trail
- Limited access reconciliation

**Poor records  
management**

- Unclear identification of final records
- No version control
- Limited use of metadata

## The project preparation



## Phase 1 - Migration of databases

Select  
databases for  
migration

Move  
databases to  
new  
destination

Freeze  
network  
drives

## Phase 2 - Migration of documents

Analyse  
content of  
network  
drives

Select  
documents  
to be  
migrated

Migrate  
documents

Authorise  
deletion of  
outdated  
documents

## Phase 3 – Decommissioning of K/drive

Remove user  
access from  
K/drive

Decommission  
K/drive

## What we migrated

### **DATABASES & PROGRAMME FILES IN USE**

**Databases**

**Programme files**

**Econometric packages**

**Calculation tools**

**Complex excel sheets with  
macros/links**

### **DOCUMENTS WITH NON EXPIRED RETENTION as per ECB retention policy**

**Documents with permanent retention**

**Documents with retention > 10-15yrs**

**Documents used on a regular basis  
and closely-related items**

**Recently created documents, if any**

## What we did NOT migrate

**Outdated databases  
& tools**

Content extraction (.pdf) in some cases

---

**Documents with  
short retention periods**

<10yrs, e.g. personal data

**Draft documents  
or superseded versions**

Only final versions in EDRMS

**Documents submitted to  
formal fora**

Migrated as part of the master files of these fora

**Password protected,  
encrypted documents, zipped  
files, un-supported formats**

Recommended un-protection,  
decryption, un-zipping before  
migration

## How we supported users

**Written guidance on what and how to migrate**

MIGRATION OF INFORMATION FROM NETWORK DRIVES – TOOLKIT FOR BUSINESS AREAS

**Tips for moving files in Windows Explorer and Microsoft Access 2007/2010**

**Guidance on migration of documents from the network drives to  
DARWIN**

**Templates to document migration activities**

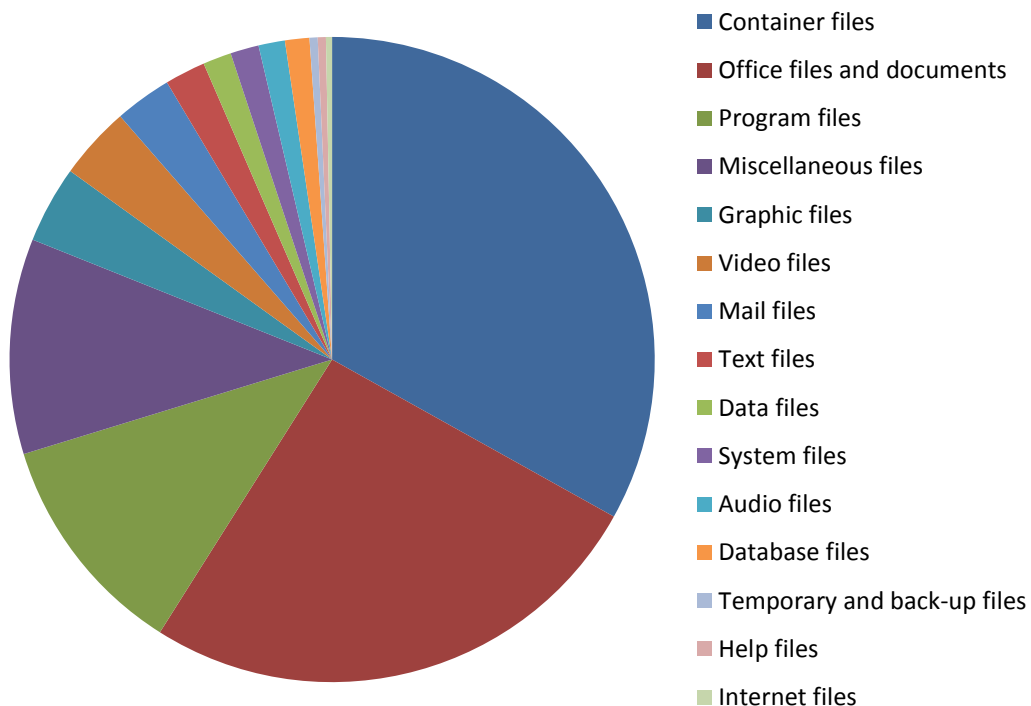
Phase 1 - IDENTIFICATION AND MIGRATION OF DATABASES / PROGRAMME FILES				
No.	Activity	Deadline	Responsible BA or unit (BA / unit acronym)	Status
1.1	Identify databases / programme files in use in the BA	Q2 2014		

**To guide users along the process**

List of databases / spreadsheets / programme files to be migrated to new business area folder on P\ECB business areas - [BA or business unit]					
Number	Name of database / programme file	Owner / responsible unit	Current location (path)	Linked files (Y/N)	Comments

## How we supported users

### Overview of size & content of drives by business area & file type



To assess effort for analysis & migration

### Overview of folders by business area

```

\\DIMECB01\LS-Data\LS-Common\01-LEGCO\
\\DIMECB01\LS-Data\LS-Common\02-ECB Consultations\
\\DIMECB01\LS-Data\LS-Common\03-DG-L\
\\DIMECB01\LS-Data\LS-Common\04-Member States\
\\DIMECB01\LS-Data\LS-Common\05-Third countries\
\\DIMECB01\LS-Data\LS-Common\06-Community Institutions - general\
\\DIMECB01\LS-Data\LS-Common\07-International Organisations - general\
    
```

## How we supported the selection of documents to migrate

**Records managers carried out initial high-level analysis of content of drives against the retention periods**

Indication of sets of documents for migration, content assessment or deletion

Based mostly on folder names

\\GIMECB01\Data\General\Training BdE\	3.6.5 - Training & development - 5 yrs retention
\\GIMECB01\Data\Project\ESCB\Finance Kit Projects	8.8.2 - 5 yrs after closure of project
\\GIMECB01\Data\Project\ESCB\MOP	8.8.2 - 5 yrs after closure of project
\\GIMECB01\Data\Committee\Market Operations Committee	1.4.9 Support to MOC - Permanent
\\GIMECB01\Data\Contribution	not clear which activity is referred to

**To guide & facilitate  
the selection of  
documents**



## How we supported the migration of documents to the EDRMS

### Define process for migration

Prepare documents for migration in network drives (delete drafts, duplications, password protected, etc.)

Assign appropriate metadata and permissions in destination folder in EDRMS

Move documents to EDRMS

Check completeness of migration

Adjust names & declare as records

### Provide import tools

Ad hoc bulk input of documents via workflows

Consistent metadata input

### Provide resources for migration

To help meet deadlines

To ensure accurate and smooth migration

## What went well

### Rigorous project management

- Tracking of activities
- Quarterly status reports

### Communication with users & management

- Regular reporting in different fora
- Liaising with users on status of activities

### Co-operation with IT function

- Timely reporting of technical issues
- Joint communication to all staff on project milestones

### Targeted migration where possible

- Focus on missing information with permanent retention in EDRMS (e.g. formal fora)
- Check after migration

### Delayed deletion of information

- ~1 year after completion of migration
- Possibility of retrieval via IT department

## Lessons learnt

**Long deadlines lead to loss of momentum**

- Requests of extensions caused project delays
- Last-minute uncontrolled migration close to final deadline could not be avoided

**Select staff with the right profile**

- In-house knowledge of databases & tools
- Historical knowledge of pre-EDRMS practices

**Exceptions must be kept to the minimum**

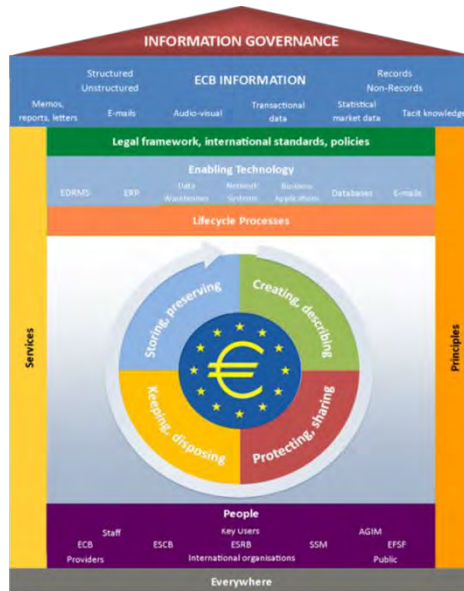
- Only upon justified business reasons
- Increased tracking and monitoring effort

**Use IT tools for analysis & migration**

- Automatic identification of duplications / unsupported formats
- Keeping of some original metadata

## To conclude...

- The ECB has achieved an important milestone in its information management journey
- The concerted effort of IM, IT and users has paid off
- A big step forward for the ECB information governance!



## Any questions?





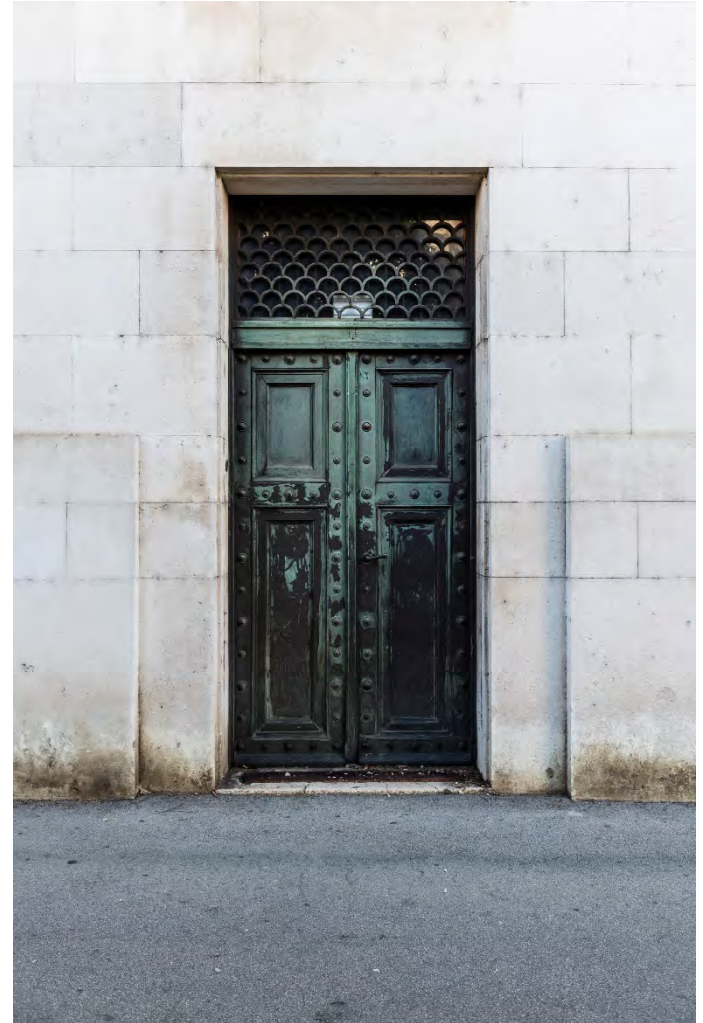
CROATIAN NATIONAL BANK

---

# Archival appraisal in the Croatian National Bank

Egon Kraljević and Tvrtko Ujević

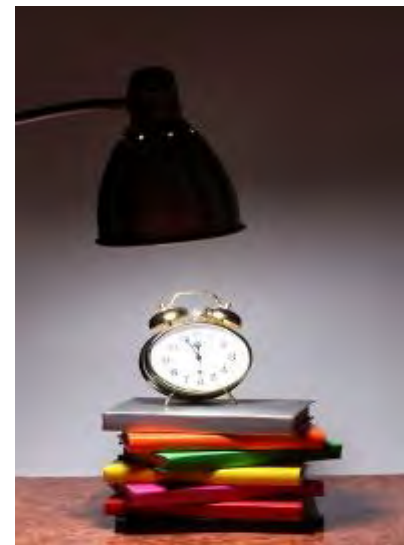
- ❑ Archives of public institutions are directly subordinated to the system of state archives
- ❑ every public institution is obligated to draw up a schedule of the type of documents/data originating from its work with a records retention schedule and submit it to the competent state archives.



# Creation of the new retention plan of the Croatian National Bank (CNB)

---

- The creation of the new schedule as an opportunity to reappraise the CNB's records





# Appraisal of records characteristic for the CNB

---

- Appraisal paradox: how to assess, at the present time, which documents, for cultural and historical reasons or on some other grounds, may be important for the future, and as such, worthy of preservation.

# Some appraisal guidelines

---

- Documents created as a result of the operations of the Bank's management are bound to have greater value than the documents created during the operations of subordinate departments. At lower business operation levels, documents recording operations that affect the scope of economic activity in the long-term, such as exchange rate setting, bear more relevance than those recording daily activities carried out on a large scale (e.g. daily payment transactions).

# Misunderstanding of the concept of permanent custody

---

- ❑ Permanent records are those that are relevant for business operations;
- ❑ All the data that employees assumed they might need for doing business in the medium and long term have permanent value.
- ❑ This interpretation, led to a surge in permanent retention designations, resulting in the overcrowding of our storage areas

# Some other things we learned from the appraisal process

---

- ❑ Interestingly, electronic data were not perceived as archive materials by CNB employees, which proves that the notion according to which archive materials comprise only documents in traditional paper form is still very common.

## Some other things we learned.....

---

- It is impossible to complete the work in just one cycle. An archivist must not give in to frustration if there is a lot of planned, but unfinished work left after appraisal that will have to be completed on another occasion. It is only after several cycles that personal expectations can be met, after which the following appraisal processes will only consist of smaller updates.

# Digitalisation and microfilming

---

- The permanent retention designation will serve as the main point of reference used to select materials for microfilming, which will mostly be directed at the protection of the most valuable materials in terms of business and cultural history.
- Since it is evident that the Croatian National Bank will digitalise more documents than will be recorded on microfilm, it is necessary to consider introducing retention periods for digitalised archive materials as well (which has not been the case so far). In our view, similar, if not the same criteria should be applied as in traditional appraisal.

Thank you for your attention!

---





BANCO DE PORTUGAL  
EUROSYSTEM

# Appraisal in Banco de Portugal

**A. Gil Matos** • Head of the Archive Unit

**Filipe Fernandes** • Record Manager

EABH Archival Workshop • Paris, 22 June 2017





## **Summary**

**Introduction**

**Appraisal role**

**EDRMS Implementation**

**Historical Archive digitalization**

**Future Developments**



## Banco de Portugal's Archive Unit

Created in 1959, in the Statistics Department, responding to its need to gather and organize a great deal of information for analytic purpose, it became three years later, an autonomous service. In 1976 it became a part of the Documentation, Editions and Museum Area;

It gathers the documental collections resulting from the over 170 years of the Bank's existence, as well as from extinct institutions related to the bank;





## Banco de Portugal's Archive Unit

Continuous role in the selection and classification of documents, working with the bank's departments to establish preservation schedules and selection tables, as well as managing long term and permanent storage of records.





## Banco de Portugal's Archive Unit

- In 1990 a committee with the purpose of establishing retention schedules and disposition for documental series, was created – *Comissão de Apreciação de Prazos de Guarda (CAPG)*. This committee is composed by members from the archive, the organizational and the juridical services, and gathers whenever is needed to decide about a record series classification, with it's issuing service.
- In 2012, a regulation for the recently implemented Electronic Documents and Record Management System was created. It was established that appraisal and selection of records would be managed in the same way for both physical and electronic documents.



## Before 2010

### Ingestion of documents to the Archive Unit

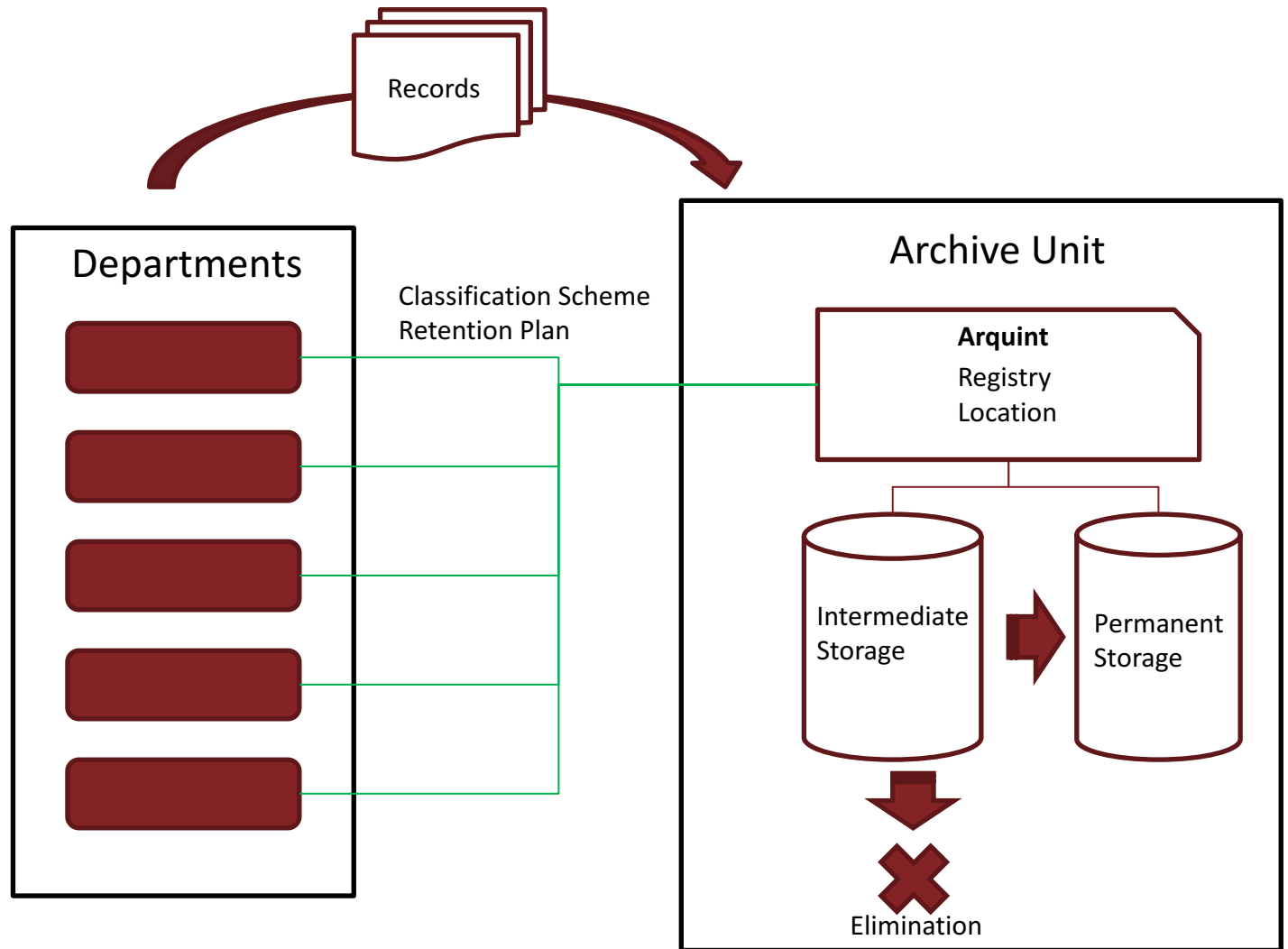
- Physical documents only (paper, tapes and CD /DVD);
- No systematic audit on records in the departments. Archive Unit and departments would approach due to the necessity to keep long term records, or the interest in permanent record keeping for the historical archives;
- Although very comprehensive, it didn't cover all the records produced in the bank.



## Before 2010

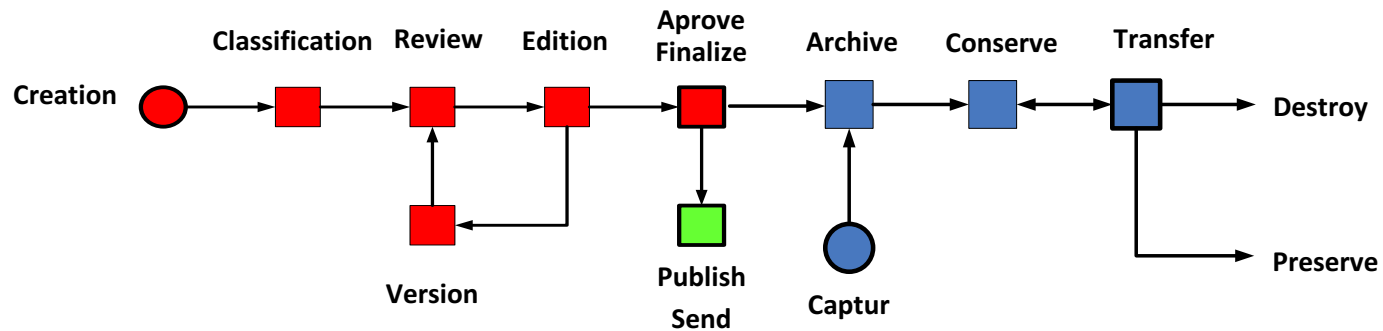
### Intermediate archive management (retention and disposal)

- Retention schedules established by CAPG;
- Record shipping for the deposits were officialised with a detailed consignment note, where later on the physical localization of volumes were also updated;
- Storage, retention and disposal managed by a Ms Access database (Arquint);





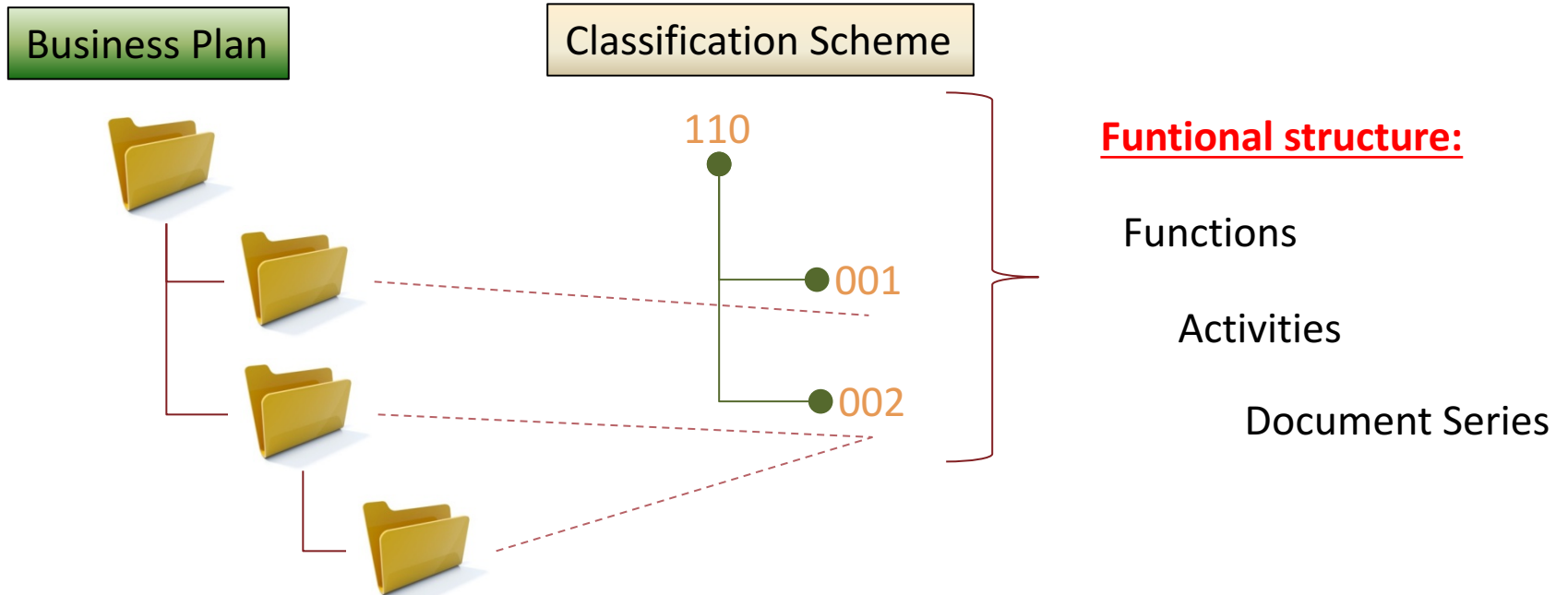
## Lifecycle management



## Document Management

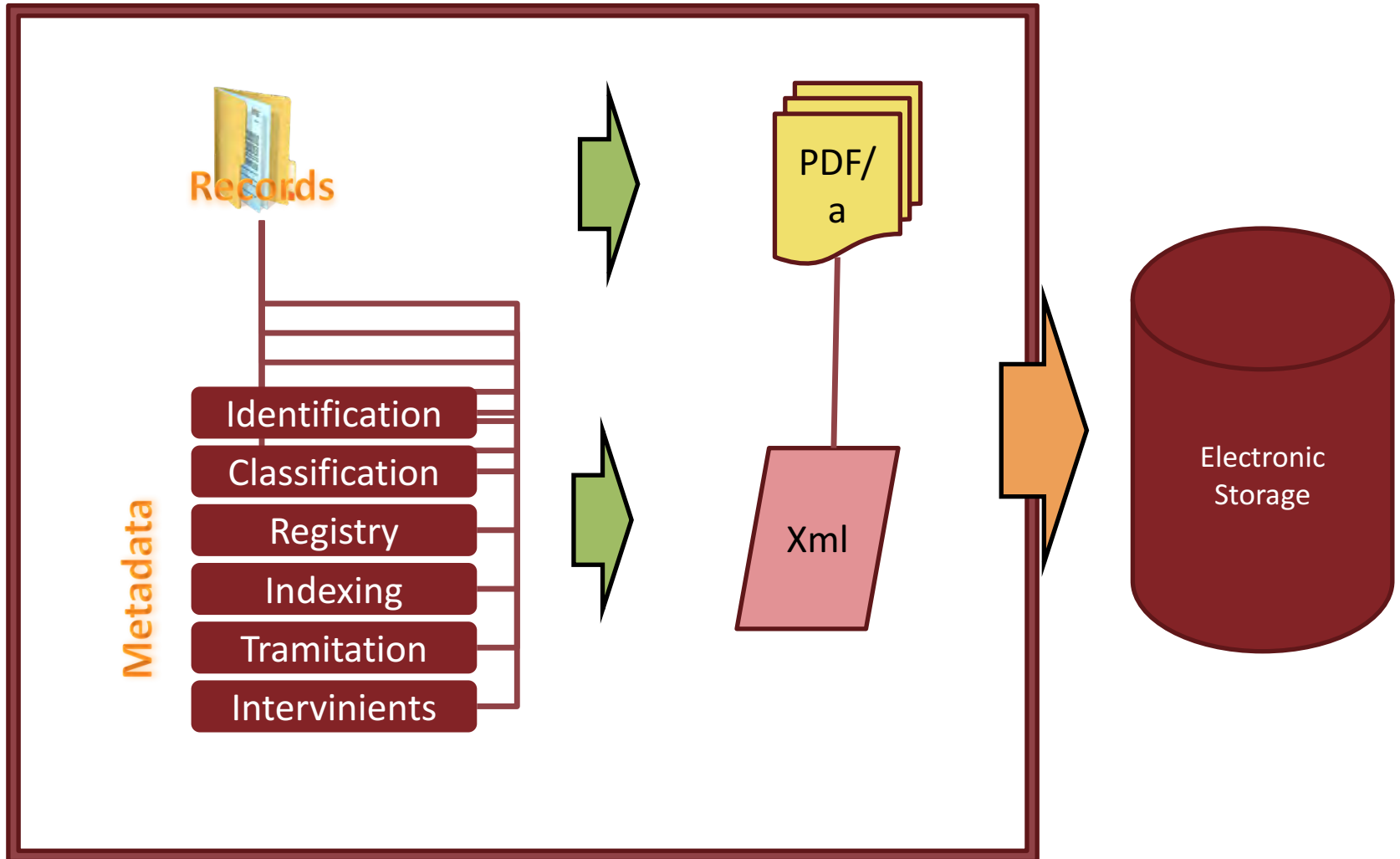
## Archive Management





**Classification scheme defines, for each folder:**

- archive classification code;
- archival designation;
- retention schedule after the folder is closed;
- disposal rule (destruction or permanent preservation on the historical archive).





- Creation of a network of interlocutors with the Archive Unit;
- Designing a workflow for physical documentation requests;
- Adaptation of Arquint to Livelink, and record registry migration;
- Creation of a correspondence digitalization system;
- Implementation of a electronic storage system (Centera).



- Conversion of the digitalized records on CD's to Tiff images (2010-2012);
- Generation of Pdf rendition of documents;
- Upgrading Winlib to Nyron and uploading the Pdf files (2013);
- Creation of database with the location of all documentation in deposit;
- Link to online search and digital record access on the Banco de Portugal's website.

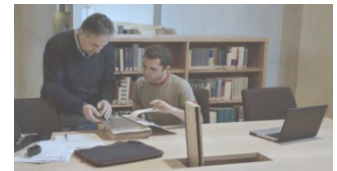
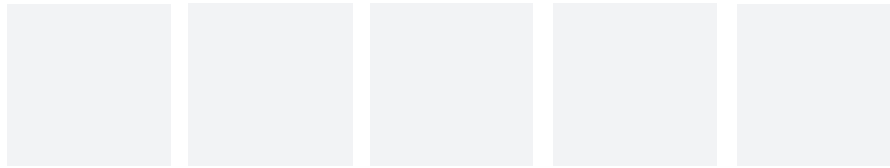


- Implementation of archive server and electronic storage system;
- Redacting our set of rules for record managing and archiving, and reviving the Retention Schedule Committee (in progress);
- Creating a support line with an intranet platform;
- Tutoring and presentation to newcomers;
- New historical archival system.

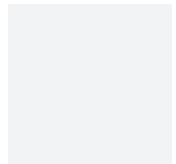
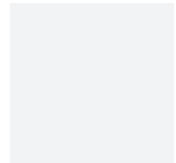
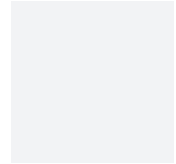


European  
University  
Institute

HISTORICAL  
ARCHIVES  
OF THE  
EUROPEAN  
UNION



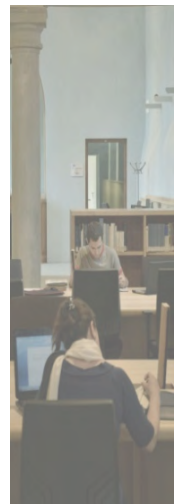
# Archival appraisal in a digital context at the Historical Archives of the European Union (HAEU)



Appraisal in the digital era

22 June 2017

BNP Paribas - 14 rue Bergère - 75009 Paris

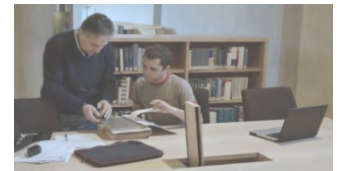
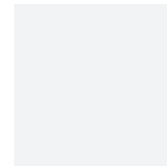
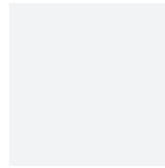
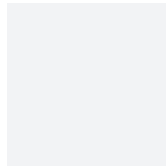
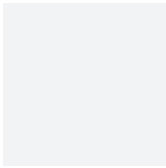
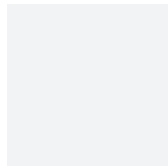


Samir Musa

Historical Archives of the European Union

[Samir.Musa@eui.eu](mailto:Samir.Musa@eui.eu)

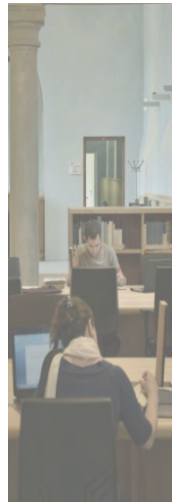
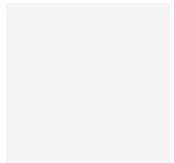
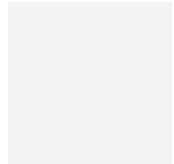
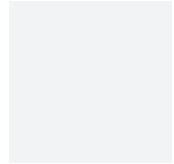
# INTRODUCTION



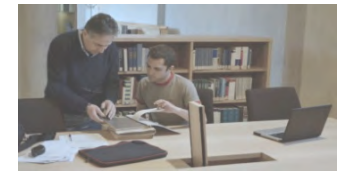
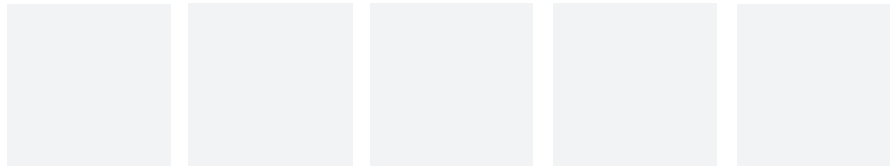
## The HAEU mission

---

- Specialized archives on **European Integration** and **cooperation**
- Collect, preserve and make accessible historical archives of the **EU Institutions**, of **private persons** and **organizations** with European scope
- Main objectives:
  - encourage research on the **history** of the EU Institutions;
  - promote the public interest in the process of **European integration**;
  - Increase **transparency** on EU institutions' work
- Established in 1984 at the **European University Institute**, an internationally renowned centre of academic learning



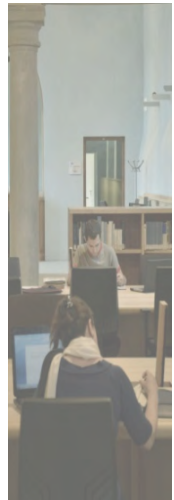
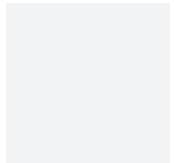
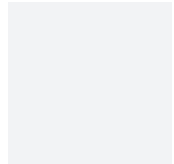
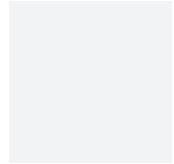


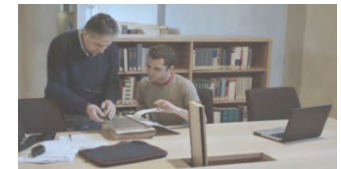
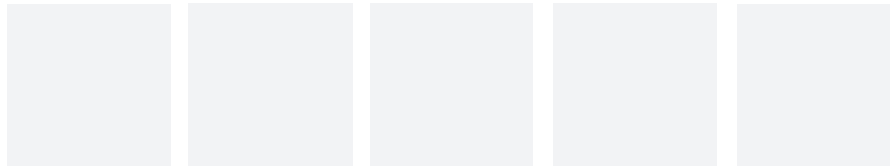


## The HAEU holdings

---

- EU historical archives declassified and open for consultation after **30 years**
- **5,500 linear meters** of paper files covering the years 1951 to 1983 (1999) with yearly transfers of **250 lm/10,000 files**
- Online-database contains **230 000 records** and **16,000 digital copies** of paper files
- **800** reading room **users** yearly produce 120 articles, monographs and doctoral theses
- **350.000 web site visitors** and **15,000 file downloads** per year





# The HAEU holdings

---

## EU Institutions:

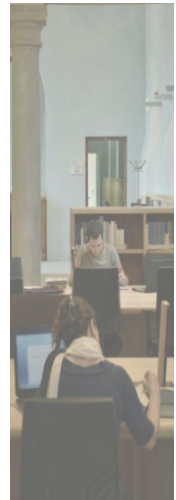
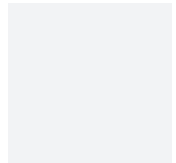
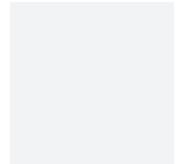
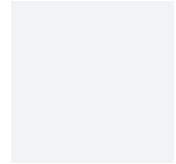
- European Parliament
- Council of the EU
- European Commission
- Court of Auditors
- Economic and Social Committee
- Investment Bank
- Centre for the Development of Vocational Training

## Private archives and collections

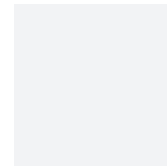
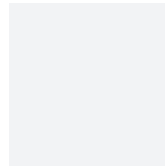
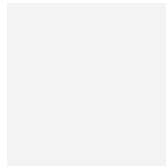
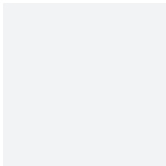
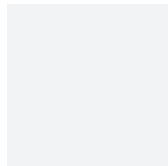
- Founders of the European Communities
- Pioneers and Federalists

## Private Archives of Groups

- Political Groups of the European Parliament
- European Federalists
- Non-EU Organisations



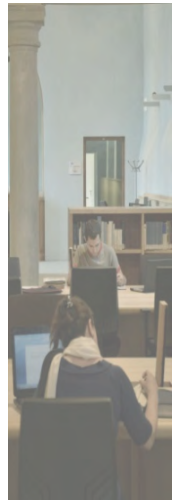
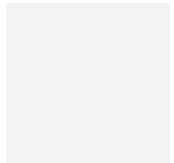
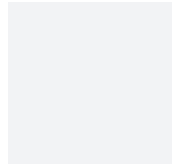
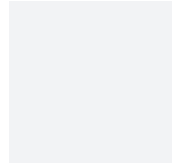
# DIGITAL SERVICES

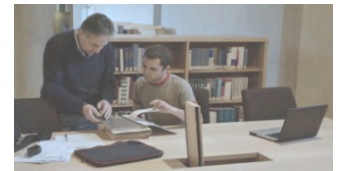
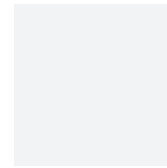
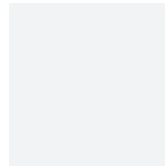
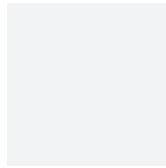
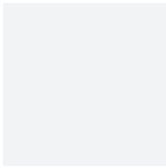
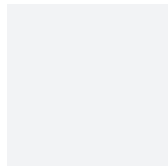


## The HAEU digital collections

---

- **Multilingual, multimedia and multiple provenances** collections
- Databases and digital collections:
  - Archival fonds including **textual documents, still images** and **ephemeral materials**
  - **Audio collection** (Oral history): interviews to EU personalities
  - **Audio-visual material** (not yet published due to copyrights)
  - **Digital-born archives** (not yet published due to thirty years rule)
  - **Websites** collections





## The HAEU digital services

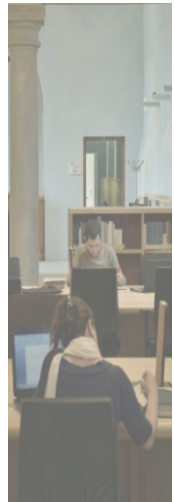
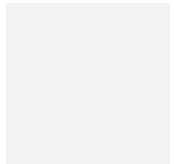
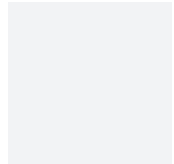
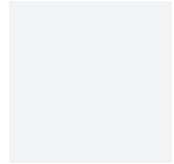
---

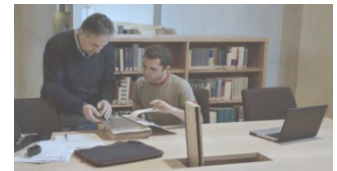
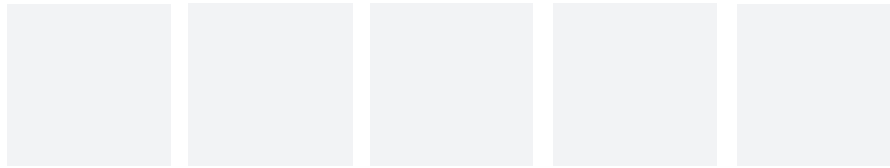
### Main goal:

- **Dissemination** of **digitised resources** (mainly paper-based documents), NOT (yet) born-digital records through **online inventory**

### Main services:

- **Digitisation** of analogue material: paper-based documents, videotapes (VHS), photographs (negatives and positives), slides, audiocassettes, minicassettes
- **Webarchiving project**: EU websites crawling
- **Digital preservation system**: born-digital records (in progress)





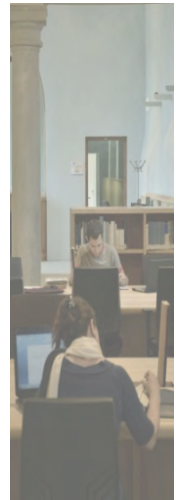
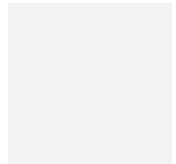
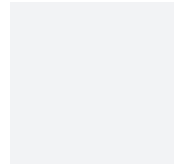
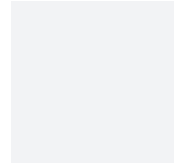
## The HAEU digital services

---

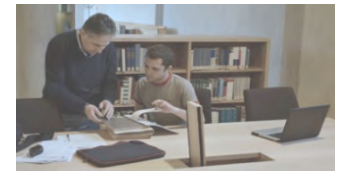
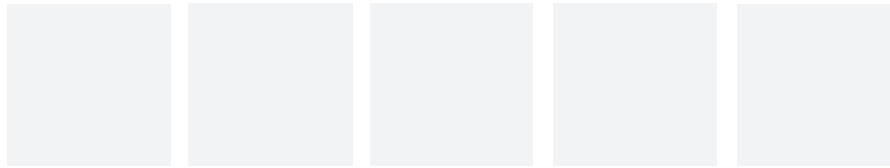
Acquisition of both digitised and born-digital material needs:

- Rules for description harmonisation (syntax and semantics)
- Rules for data transmission (exchange protocols)
- Rules for file formats
- Rules for authority files and file naming

The HAEU issued a series of **Policies** and **Guidelines**



# ARCHIVAL APPRAISAL POLICIES



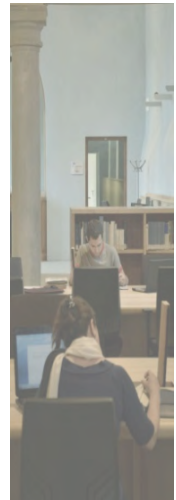
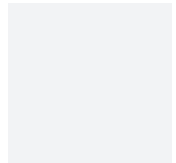
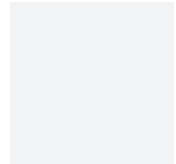
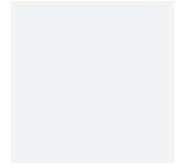
## Archival Appraisal at the HAEU

---

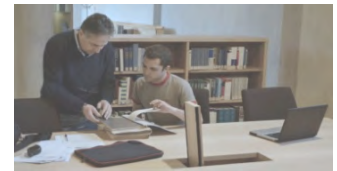
**Archival appraisal** has been defined by the HAEU practically as:

“The **selection of records** that are judged worthy of **permanent preservation as archives** and that will form a relatively small percentage of the EU Institution or Agency’s records”

(HAEU *Guidelines on the management of the Historical Archives of the European Union Agencies - Definitions*)



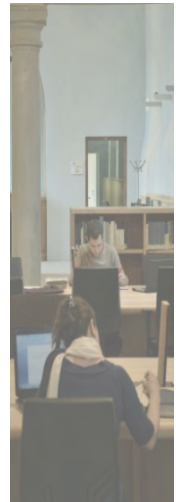


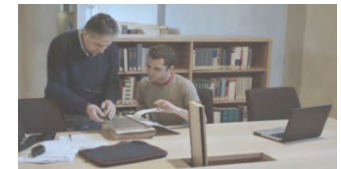
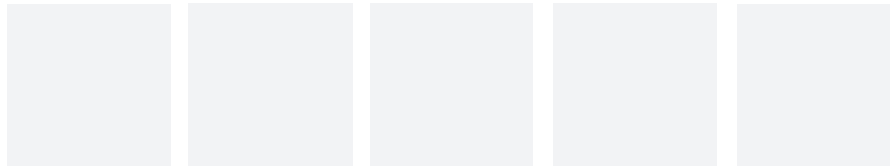


## Archival Appraisal at the HAEU

---

- Appraisal has always be based on a practical (and subjective) approach by archivists due to:
  - Heterogeneous fonds
  - Responsibility of description
  - Nature of records transferred to the HAEU
- **HAEU Guidelines on the management of the Historical Archives** of the European Union Agencies was submitted to the plenary session of the EU Agencies Working Group for definitive approval on **21 February 2017**.

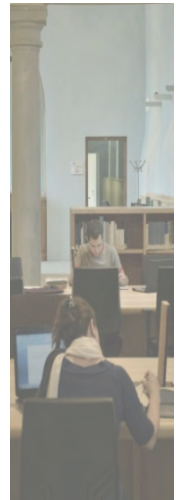
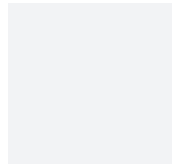
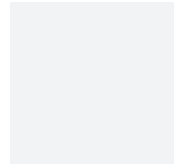
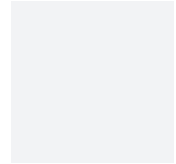


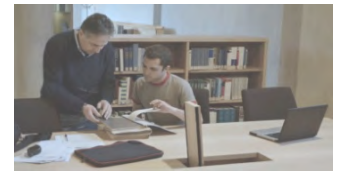
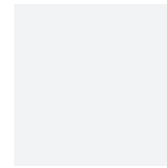
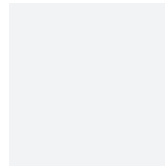
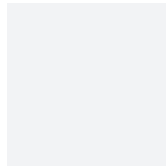
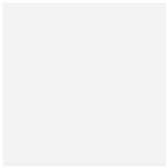
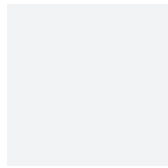


# HAEU Guidelines: methodology

---

- Roles and responsibilities (EU Agencies and HAEU)
- **Section 1** – Paper-Based Archives
- **Section 2** – Audio-Visual Archives
- **Section 3** – Digital Archives
- **Annexes:**
  1. Description metadata for Textual documents (file level)
  2. Description metadata for Textual documents (other level)
  3. Description metadata for Audio-Visual documents
  4. Preferred file formats

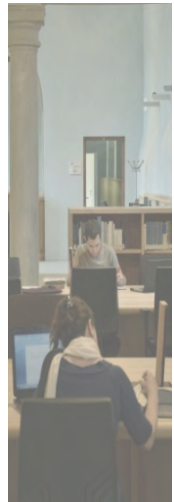
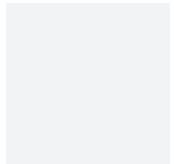
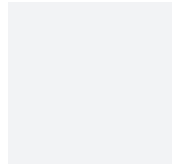
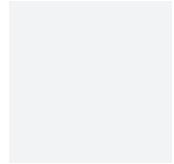


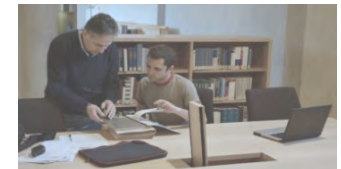
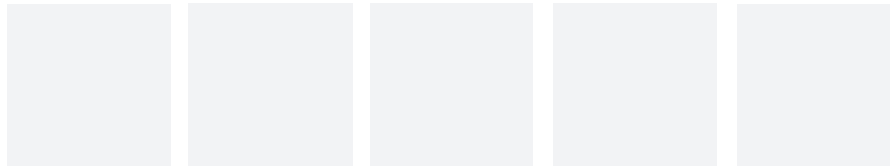


# Paper-Based Archives

---

- **Appraisal Policies and Guidelines**
  - Records which should be kept as EU Agency Archives
  - Records which should not be kept as EU Agency Archives
- **Destruction**
- **Transfer**
  - HAEU Guidelines on archival description for the Historical Archives of European Union Institutions XML/EAD
  - Annexes 1 and 2 – .xls or .csv
- **Conservation**

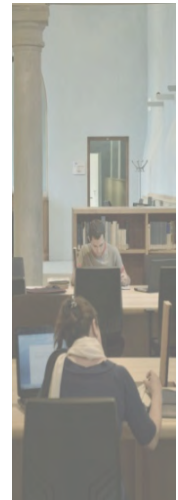
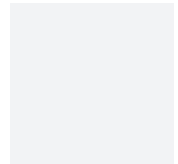
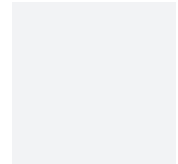
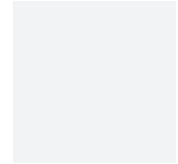


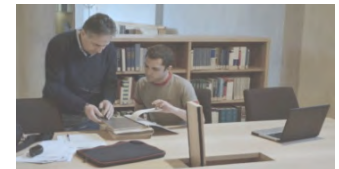
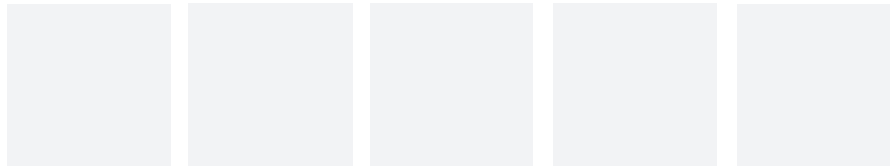


# Audio-Visual (analogue) Archives

---

- **Appraisal Policies and Guidelines**
  - Records which should be kept as EU Agency Archives
  - Records which should not be kept as EU Agency Archives
- **Destruction**
- **Transfer:** Annex 3 (.xls, .csv)
- **Conservation:** general recommendations

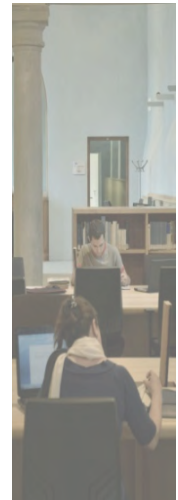
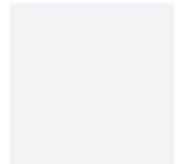
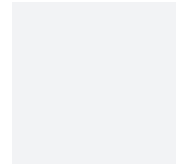
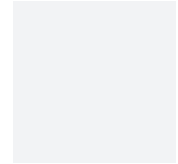




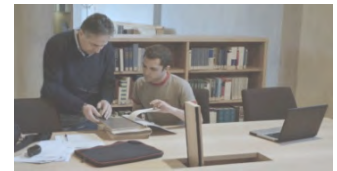
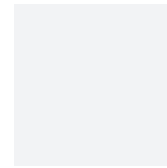
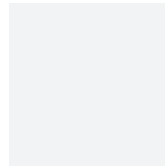
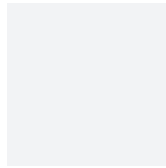
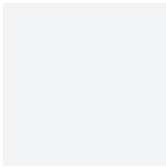
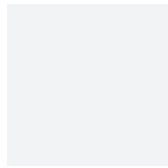
## Open issues

---

- **Intellectual Rights Management** for audio-visual material
- **Data protection**



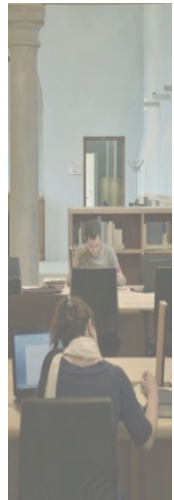
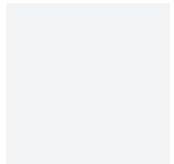
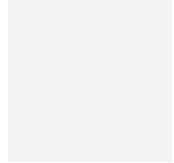
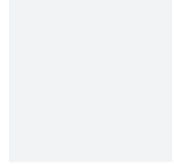
# APPRAISAL IN DIGITAL CONTEXT

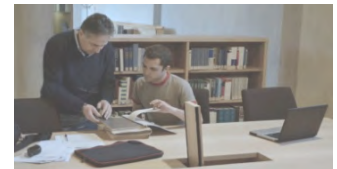
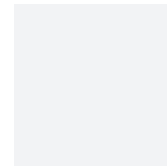
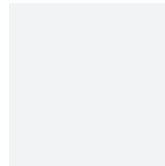
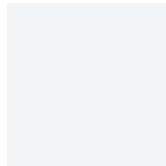
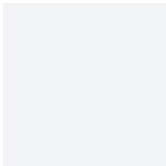
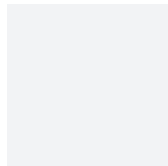


## Digital Archives: peculiarities

---

- **Long-term preservation** depends on
  - improved records management programs (best practices and policies)
  - a clear statement of the archives jurisdiction in the matter
- An integrated system for managing both **electronic** and **hard copy records** is needed (hybrid archives)
- **Schedules** have to be developed **at the time of design of systems**
- **Appraisal** needs to be anticipated at the creation stage

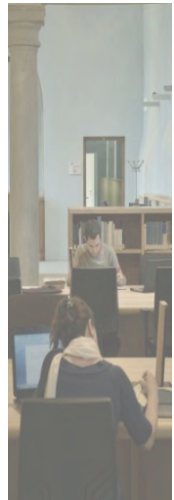
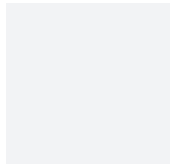
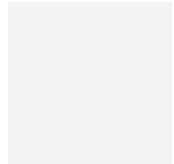
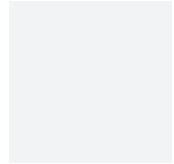




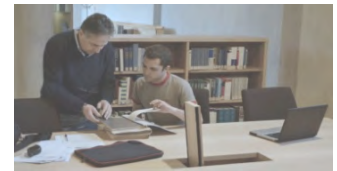
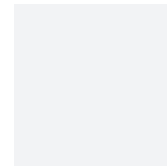
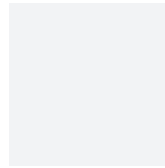
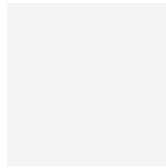
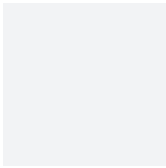
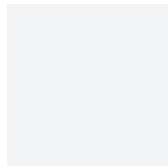
# Digital Archives: state of the art

---

- Permanent retention of digital records: **responsibilities**
  - In house vs. outsourcing
  - EU Institutions and Agencies ownership
  - HAEU digital services
- **Policies** and Procedures:
  - Preservation Plan (ISO 16363)
  - OAIS data and workflow model
- **Best Practices:**
  - Archiving Electronic Records
  - Integrity of Data
  - System Parameters

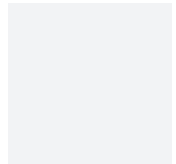
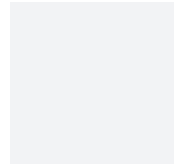
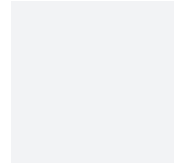




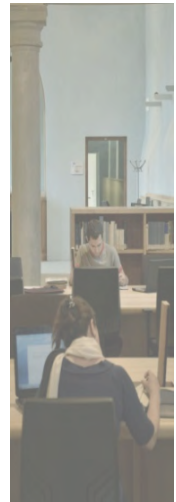


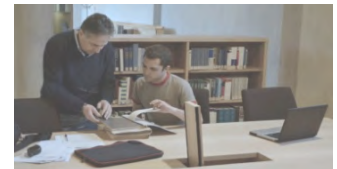
# HAEU Digital Preservation System (DPS)

---

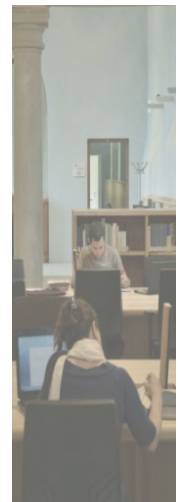
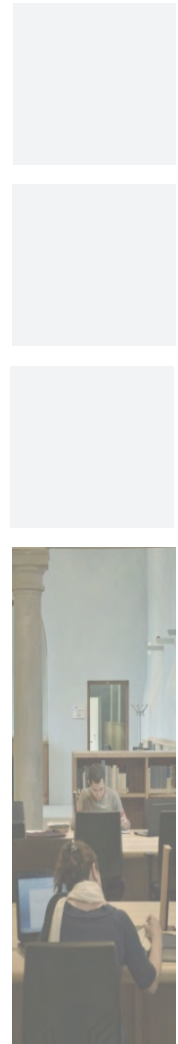
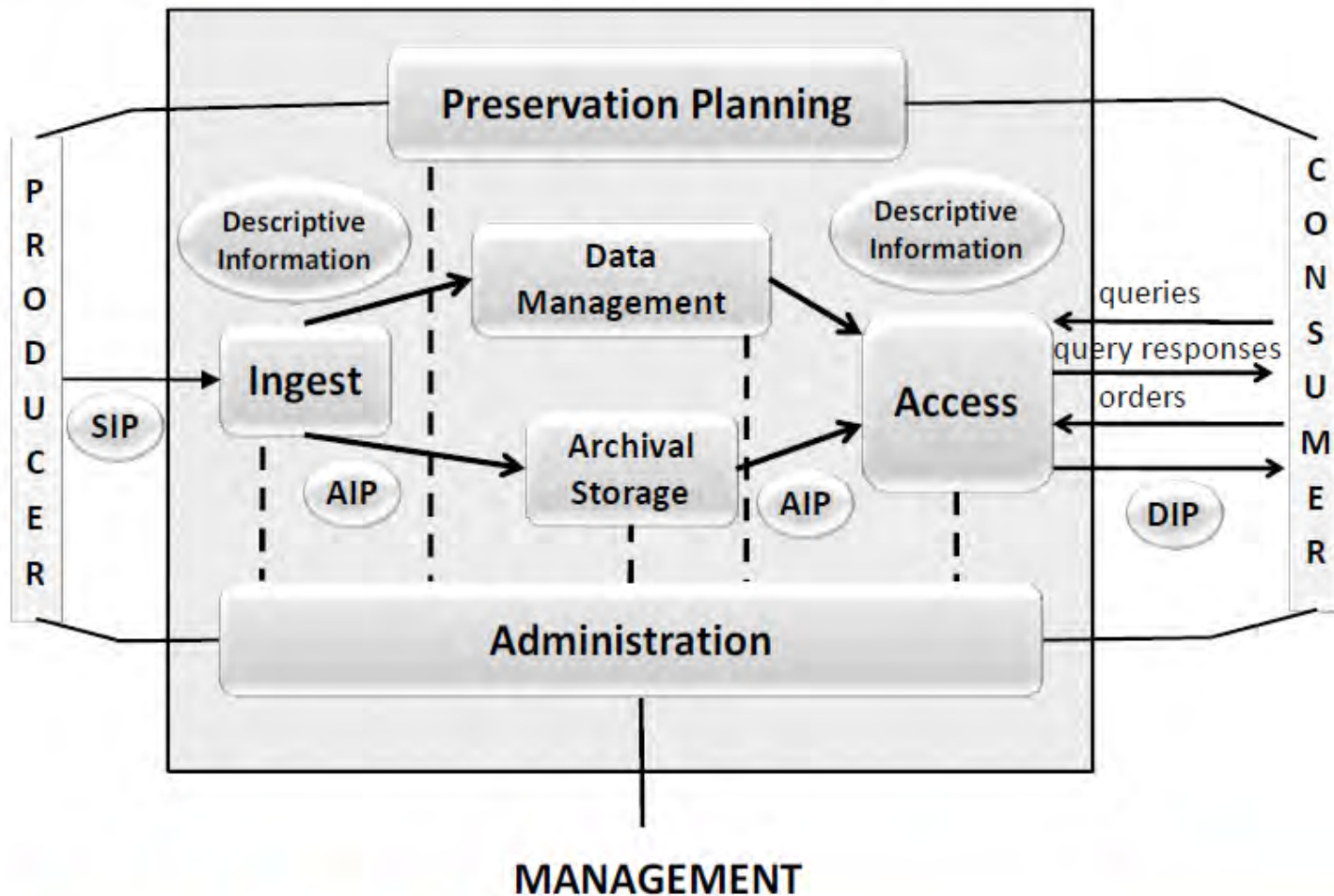


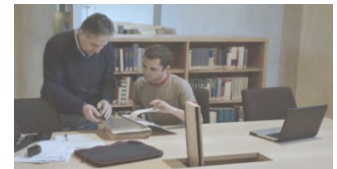
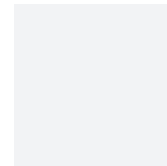
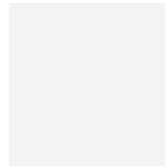
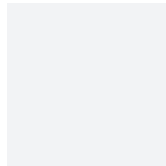
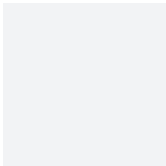
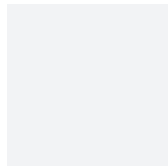
- Beginning 2016: launch of project and call for tender
- Four years timescale
- Recognized project management standards (PMI)
- Formal approach to documentation
- Use of standards and open source code





# OAIS Reference Model – Functional entities



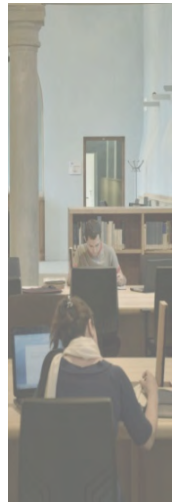
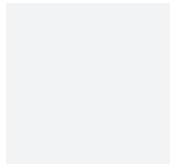
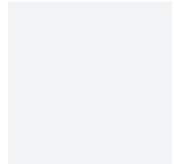
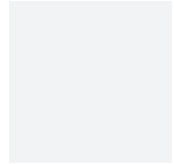


## Appraisal in the digital context

---

Appraisal in digital preservation system – focus points:

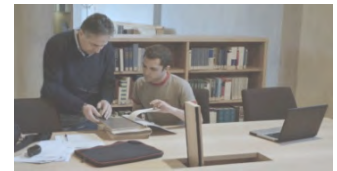
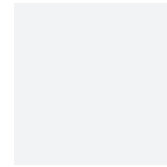
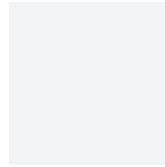
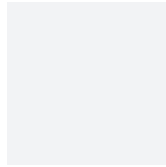
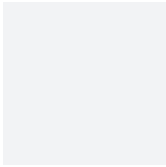
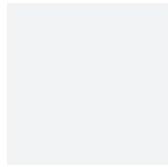
- Selection must be performed at Ingest stage (SIP) but also at Archiving stage (AIP)
- Submission agreements with External Producers: reference documentation
- Preservation plan: workflows
- Metadata control: scheduling automation



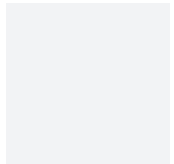
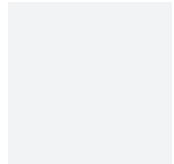
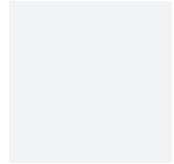


European  
University  
Institute

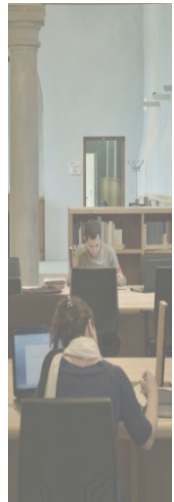
HISTORICAL  
ARCHIVES  
OF THE  
EUROPEAN  
UNION



Thank you for your attention



Samir Musa  
Historical Archives of the European Union  
[Samir.Musa@eui.eu](mailto:Samir.Musa@eui.eu)



# APPRAISAL AND SELECTION OF SOCIAL MEDIA CONTENT FOR LONG-TERM PRESERVATION



**WORLD BANK GROUP**

Esther Abangma  
Paloma Beneito

*This material is a product of the staff of the International Bank for Reconstruction and Development/The World Bank. The findings, interpretations, and conclusions expressed in this presentation do not necessarily reflect the views of the Executive Directors of The World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. This material should not be reproduced or distributed without The World Bank's prior consent.*

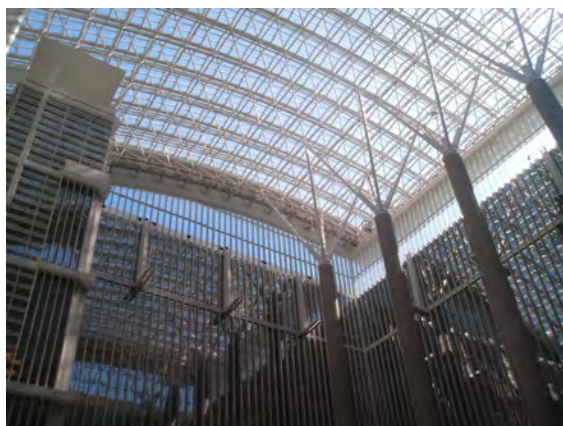
Paris, 22 June 2017

# The World Bank Group



- Established July 1944 to rebuild Europe after World War II
- Comprises 5 institutions: IDA, IBRD, MIGA, IFC and ICSID
- Provides finance and technical assistance
- Twin Goals:
  1. End extreme poverty by decreasing the percentage of people living on less than \$1.25/day to no more than 3%
  2. Promote shared prosperity by fostering the income growth of the bottom 40% for every country

# The World Bank Group



**10,000+**  
Staff from 174  
countries

**189**  
Member Countries

**25**  
Executive Directors

**+120**  
Offices Worldwide

- Strong functional macro-appraisal tradition and multi-faceted appraisal environment: Assessment of records values (operational, legal, cultural..) primarily based on comprehensive analysis of functions executed by the business. Context of records creation and use & records nature and characteristics guide appraisal decisions.
- *Big Bucket* approach to retention
- Appraisal decisions documented in *General Records Schedules* (GRS) for common and administrative functions and *Records Retention and Disposition Schedules* (RRDS) for specific functions.
- GRS and RRDS instruments applied to paper records, records captured in official record keeping system (e.g. EDRSMS, Documentum based)
- For highly disorganized, technically complex and/or very large digital collections, functional classification and application of retention instruments too resource intensive and costly, so complementary selection criteria identified and applied.



# Appraisal at the WBG


What

Who

How

Results

Collection	Appraisal approach	Retention rules	Examples
Digital records captured in official record keeping systems	Macro-functional appraisal	RRDS / GRS	WBdocs / IFCdocs / Image Bank / ADAM records.
Paper records created and managed in accordance with RM policies	Macro-functional appraisal	RRDS / GRS	Paper records managed by business units
Low volume / organized / well described collections (not captured in official systems)	Reclassification (macro-functional appraisal)	RRDS / GRS	Orphan records
High volume / metadata	Selection (based on criteria such as significance - incl. capstone-like approaches - sustainability, etc.)	Selected content retained permanently	Spark social media Lotus Notes databases, audio-visual material non captured in official RM systems, etc.

-  (Jive Software) used for collaboration / social media since 2011
- Bank's RM policy: *"Records created using social media are electronic records and should be captured in a WBG records management system"*.
- Spark place owners informed of this requirement upon site request
- BUT, do they capture?
- Not systematically, so WBG Archives decided to select and capture most valuable content for long-term preservation.
- Given very large volume of non-traditional social media records (non classified functionally), complementary section criteria needed to be identified.



## Deletion of inactive Spark places

This effort is part of a larger Intranet cleanup and revamp that will improve the search functionality.

[Find out if your group or space will be affected!](#)



In 2016's content clean up initiative, **3824** Spark sites identified as inactive

## 3 objectives

Appraisal criteria to select most valuable sites for archiving

Improve governance of social media sites

Select technical solution to capture sites as Bank records

# Multi-expertise team

What

Who

How

Results

- 1 senior IT officer
- 2 appraisal archivists
- 1 web archivist
- 1 web developer



## Timeline

Nov-Dec  
2016

- Appraisal criteria identified and applied
- 373 sites selected for capture
- Site owners informed

January  
2017

- Technical solution selected
- Governance improvements introduced

April-May  
2017

- 2<sup>nd</sup> appraisal performed
- 244 Spark sites captured as WBG archives

# 1. Definition of appraisal criteria

Team considered:

- Large volume of content to be evaluated: 3824 sites with over 300000 items (documents, blogs, discussions, wikis, etc.)
- Limited resources and short time available: appraisal at individual site or item level in order to apply Bank's official retention rules impossible
- Policy and technical issues related to long-term accessibility and preservation
- Inspired by 2016 [UNESCO/PERSIST Guidelines](#) for the selection of digital heritage for long-term preservation

## Significance



- Only sites with 10 pieces of content or more
- Sites shedding light on the activities, influence, style, and personality of significant WBG staff members, such as blogs created by WBG Managers, Vice-Presidents and the President
- Sites containing obvious significant evidence of the Bank's policies, business, mission and decisions, e.g. Definition of the Nigeria Corporate Governance Banking Program



- Empty sites
- Sites that never received any traffic
- Sites created for social, sport or testing purposes

## Sustainability



- Sites for which the Archives have the required access rights to migrate to different systems.
- Sites that the Archives can make accessible in the long term for research, exhibition or other uses to meet its public's expectations



- Private, restricted, confidential sites

\* Availability and legal obligations criteria not applied as not relevant in this case

- ✓ Only business related sites containing 10 pieces of content or more and highly relevant ones with less content, classified as “public”, “open” and “members only” were archived.
- ✓ Application of criteria reduced number of sites to be archived from 3824 to 373:

#### Application of appraisal criteria for the selection of Spark content to be archived

Summary	All	No content/ no traffic	10 pieces of content + very significant / VIP	10 pieces of content Open , public and member only access	Minus non business / test sites Minus sites with active children
Inactive spaces	1217	464 (38%)	314 (25%)	153	143
Inactive groups	1297	182 (14%)	500(38%)	135	125
Inactive projects	1310	500 (38%)	200 (15%)	105	105
<b>Totals</b>	<b>3824</b>	<b>1147 (30%)</b>	<b>1015 (26%)</b>	<b>393</b>	<b>373</b>

- ✓ All site owners informed and asked to take action to capture those that we were not archiving, if needed.

## FAQ - Closing Inactive Spark Places

Version 6

created by Robert Malloy on Jan 6, 2017 4:39 PM, last modified by Robert Malloy on Feb 7, 2017 11:04 AM

Visibility: Open to anyone

*This reminder is being sent to ALL inactive place owners (If you previously contacted us to keep your site active there is no need to contact us a second time).*

We are extending two more weeks the closure of inactive Spark places. Although the identified inactive sites are correct the owners of "Projects" was misaligned in the spreadsheet. This has been corrected. Sites in the attached spreadsheet will be either Deleted or Archived – please search the three tabs for your email address to see if any of your Spark communities are on the list.

- **WHY** – This effort is part of a larger Intranet cleanup and revamp that will improve the new intranets search as well as improve search within Spark. This will be repeated annually.
- **HOW** - ITSKI will archive the business related public, open, and member only Spark places identified for closure that contain 10 pieces of content or more. If your place was restricted, private or secret, of a social or cultural nature, and/or did not exceed 10 pieces of content, it will be deleted. If you wish to keep any of the documents from places marked for deletion, you'll need to download and save them. You can find guidance here: <http://file>. For questions and support, contact [ITSKI IM STAFF](#).
- **WHEN** – On February 10th inactive sites will no longer be discoverable in Spark.
- **HELP** - For questions, support or to ask for your place to remain active, contact [ITSKI IM STAFF](#) or check out the [FAQs in Spark](#).



## 2. Governance improvements

Site request form updated to include information about RM policy, where records should be captured, etc.

Site request form informs requesters upfront of selection criteria for the future

Sites are only created if there is a strong business purpose

Users directed elsewhere if other applications better meet their purposes

## Request a Place

Please include the following fields in your request:

1. Place name:
2. Place description:
3. Access level: Open to all, Members only, Private, or Secret
4. Information that will be shared within the group: WBG Internal Use Only, Official Use Only (OUO) - WB, OUO - IFC, OUO - MIGA, OUO - ICSID, or Confidential
5. Place administrators:
6. What results you are trying to achieve and how you will measure them:
7. Key activities:
8. VPU:
9. Allocated resources (staff/STC time of a core team) - if applicable
10. Sponsoring Unit-Department - if applicable

## CLICK HERE TO SUBMIT A GROUP REQUEST

### GROUP TYPES

	OPEN	MEMBERS ONLY	PRIVATE	SECRET
VIEW	Anyone	Anyone	Members (Invited or Approved) (membership request is open to anyone)	Members (Invite Only)
CONTRIBUTE	Anyone	Members	Members (Invited or Approved) (membership request is open to anyone)	Members (Invite Only)

**Please be advised that for technical and policy related reasons we will not be able to archive restricted, private and secret places once they become inactive.**

BEFORE REQUESTING A NEW PLACE OR SUPPORT, PLEASE REVIEW THE RESOURCES PROVIDED BELOW:

- Have you already searched in Spark for a similar place? As much as possible we would like to avoid duplicating places.
- Are you familiar with the Spark platform? If not, good resources are the [Tipsheets](#), [Spark 101 Handbook](#), [Technical Guide for Place Administrators](#).
- Are you familiar with the WBG information classification and records management rules? If not, check the [Note on Information Classification](#)
- Do you need any training? If you do please consider taking the [self-paced course](#) or take a look at the resources provided in the [Help and Training Center](#)
- Do you need any technical help? If you do please ask a question in the [Spark Help Desk](#)
- Do you need information/records management support? If so, contact [archivesandRMguidelines@worldbankgroup.org](mailto:archivesandRMguidelines@worldbankgroup.org)

### A NOTE ON RECORDS MANAGEMENT

As established in AMS 10.11 [Management of Records: Procedures](#), "records created using social media are electronic records and should be captured in an appropriate form and entered in a WBG records management system". Accordingly, if your Spark place will contain [records](#), please make sure that:

- Any [records](#) shared in Spark are also captured in an official records management system, such as WBdocs or IFCdocs [1].
- To avoid duplications, do not upload records to Spark. Instead, publish links to the records already captured in an official records management system.

[1] To capture blogs and discussions as records you need to convert them to PDF first by clicking on "Actions">"View as PDF". Once converted you may save it in the relevant official repository.

### 3. Selection of technical solution

- Extensive market research to identify best tool to capture sites
- Several larger and smaller software providers considered and consulted
- Ultimately, decision to use open source software  
<https://www.httrack.com>

Tool	e.g.
Open Source web crawlers	HTTrack Heritrix WGet
Commercial web capture tools	Hanzo Archives (privately developed web crawler)
APIs to pull content	Pagefreezer Proof Point Actiance Smarsh
Aggregators or manual methods to capture content	RRS feeds Converting to PDF

**HTTrack WEBSITE COPIER**  
Free software offline browser

About Download Manual Forum Blog Information Français

**Version 3.49-2 (05/20/2017)**  
Engine fixes (keep-alive, redirects, new hash/tables, unit tests)

Installing HTTrack: Go to the [download section](#) now!

For help and questions: [Visit the forum](#), [Read the documentation](#), [Read the FAQs](#), [Browse the sources](#)

**Welcome**

HTTrack is a [free \(GPL\)](#), libre/free software) and easy-to-use offline browser utility.

It allows you to download a World Wide Web site from the Internet to a local directory, building recursively all directories, getting HTML, images, and other files from the server to your computer. HTTrack arranges the original site's relative link structure. Simply open a page of the "mirrored" website in your browser, and you can browse the site from link to link, as if you were viewing it online. HTTrack can also update an existing mirrored site, and resume interrupted downloads. HTTrack is fully configurable, and has an integrated help system.

WinHTTrack is the Windows 2000/XP/Vista/Seven release of HTTrack, and WebHTTrack the Linux/Unix/BSD release. See the [download page](#).

© 2017 Xavier Roche & other contributors - Web Design: Leticia Kauler...

What

Who

How

Results



SPARK

**ARCHIVED WEBSITE:** SBA Donor Relations

**Year Archived:** 2017 ( [View Website Profile](#) )

[Back to Intranet Web Archives Home Page](#)

The screenshot shows the SPARK website interface. At the top, there is a navigation bar with the SPARK logo and links for Home, Browse, My Groups, and My Content. A search bar and a 'Create' button are also visible. Below the navigation bar, there is a large banner for the 'SBA Donor Relations' group. The banner includes the group name, a 'Follow' button, and a 'Join this group' button. Below the banner, there are tabs for Overview, Content, People, Projects, Reports, and Calendar. At the bottom of the banner area, there are links for Actions, About, Share, and Manage. Below the banner, there is a section for 'Access: Open - Membership is open and non-members can view content and participate.' and a warning icon for 'Data Classification: WBG Internal Use Only - Content may be viewed by all WBG staff but may not be shared with external audiences or the public.'

Persist criteria helpful to facilitate selection of large amounts of digital content if meaningful metadata (at record, system, level, etc.) is available

Some risks of missing out significant records

Httrack fully adequate to capture Jive sites, with some additional programming

A second appraisal at site level was required as 25% of sites were deemed not relevant as Bank records

- reduced sites from 373 to 269.
- next exercise will only consider sites with + 20 pieces of content

Not all sites qualified to be archived as websites.

in 20% of cases just the documents were extracted (also with Htrack) and captured in EDRMs.

# Questions?

**Contact us:**

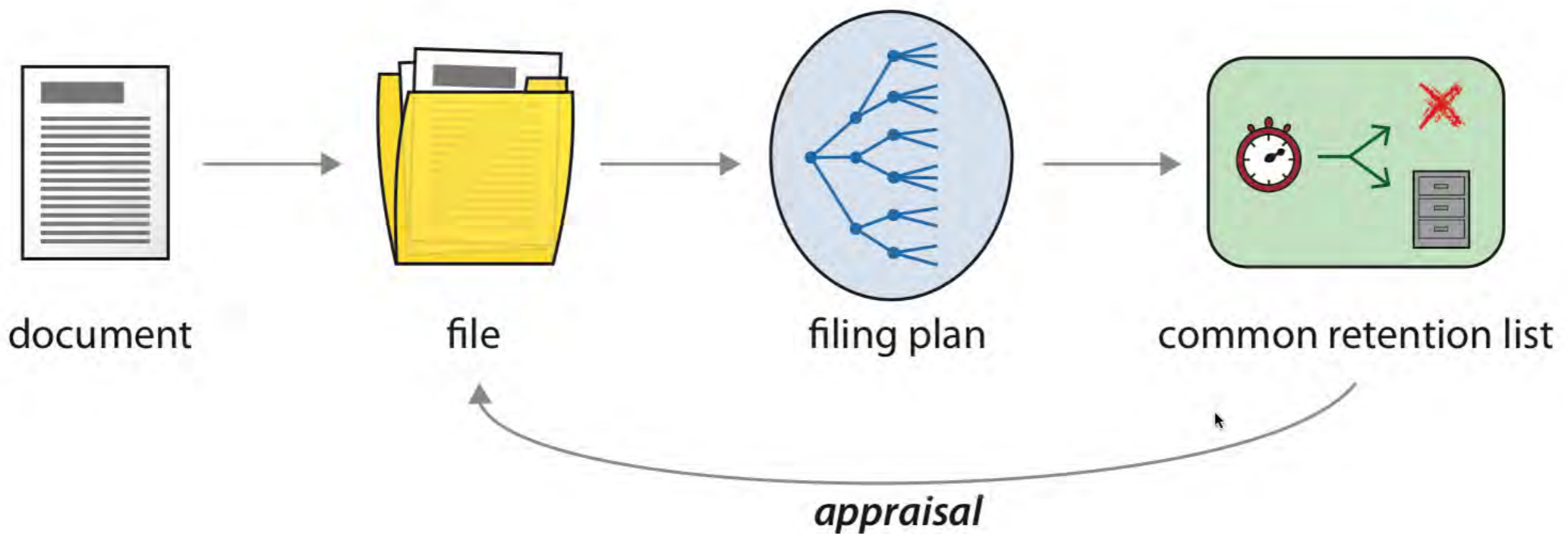
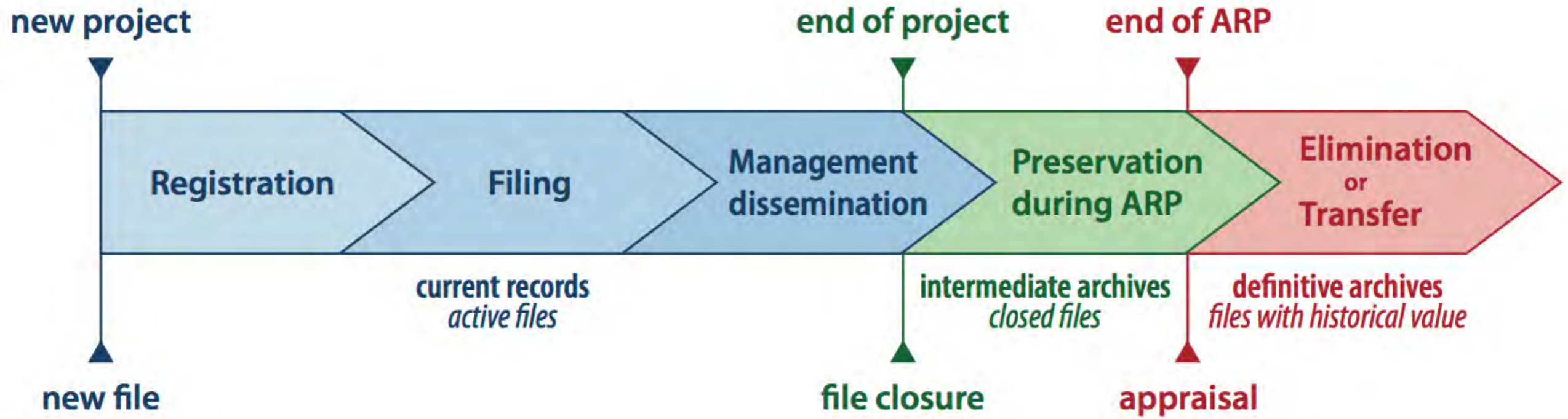
[eabangma@worldbankgroup.org](mailto:eabangma@worldbankgroup.org)

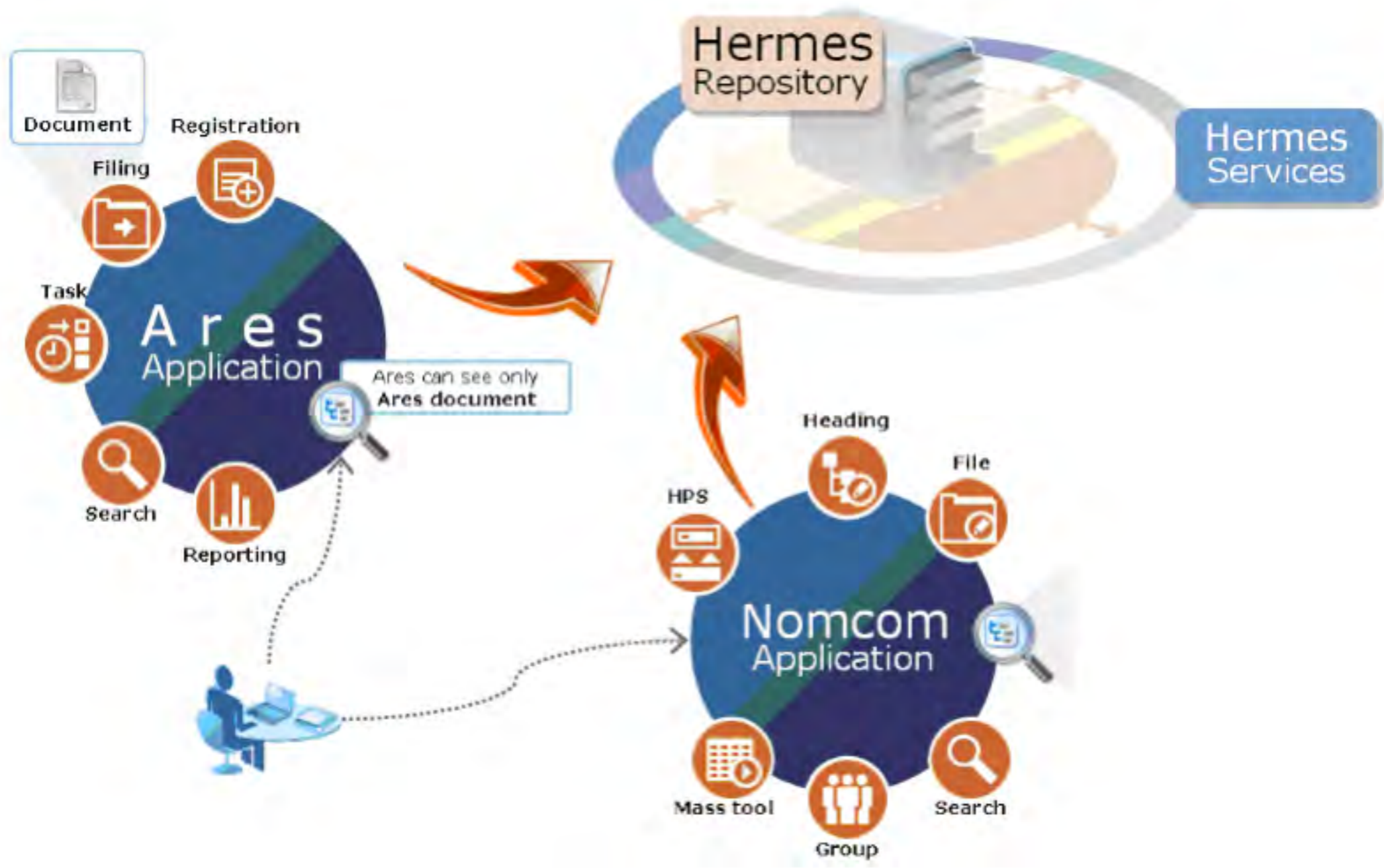
[pbeneitoarias@worldbankgroup.org](mailto:pbeneitoarias@worldbankgroup.org)

# **e-Domec records management policy of the European Commission**









Source : eDomec EC training materials

# Filing plan

Common  
Nomenclature

**04.04.02**

**Public health**

**04.04.03**

**Food security**

**04.04.03.050**

**Food legislation**

**04.04.03.050.020**

**Packaging of food products**

**04.04.03.050.020.010**

**Information for consumers on  
food packaging**

**04.04.03.050.020.020**

**Nutritional information**

**04.04.03.050.020.030**

**National notifications**

C  
-  
F  
-  
C  
E  
S  
-  
S  
P  
E  
C  
I  
F  
I  
C  
A  
T  
I  
O  
N  
S

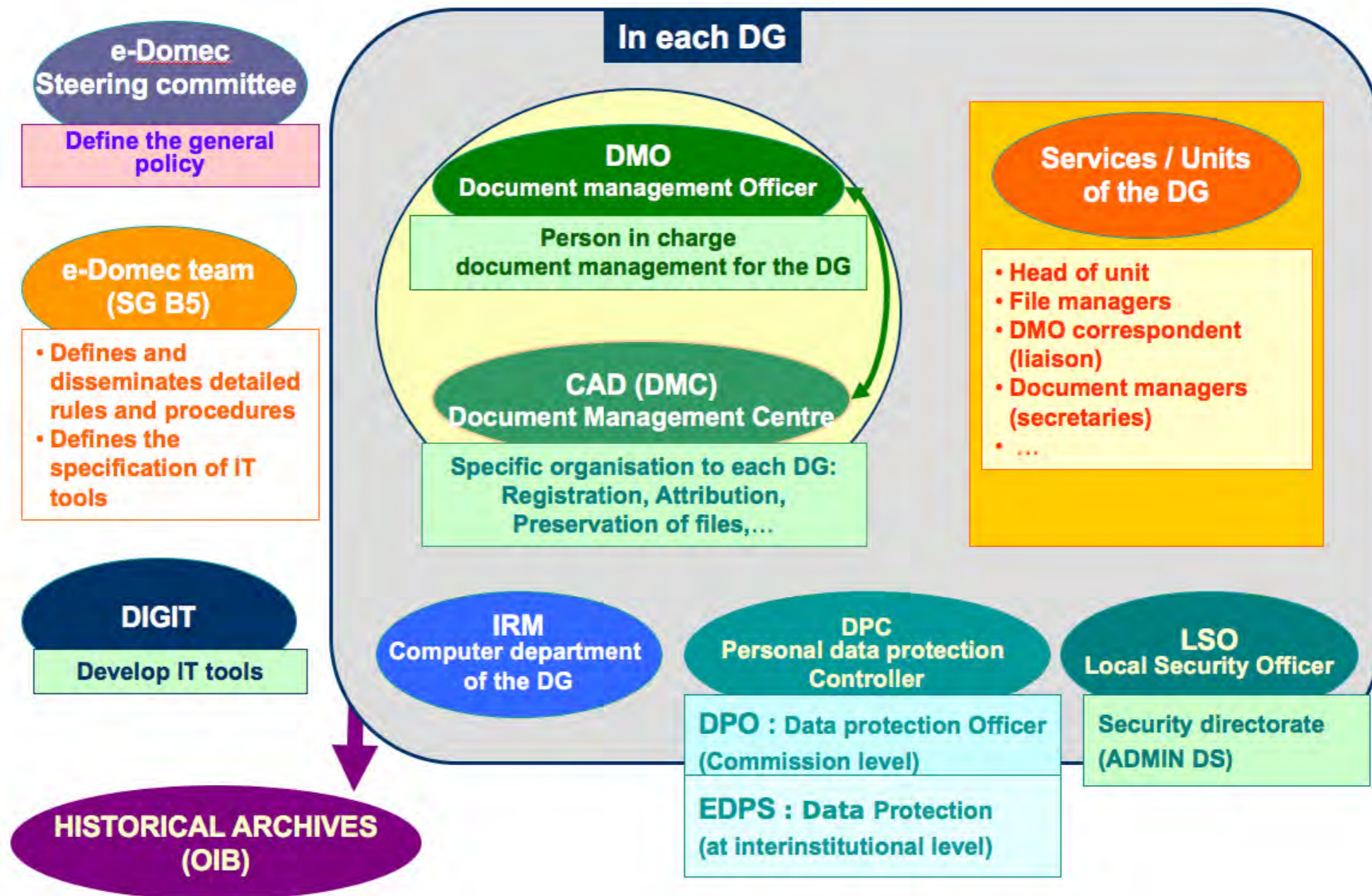
# Common Retention List

- Institutional and governance issues
- Strategy and coordination
- Commission decision-making
- EU financial system and the budget
- Policy development
- Implementation of policies and legislation
- Programmes, projects and grants
- International activities and enlargement
- Transparency, information, communication and relations with citizens
- Monitoring the application of Community law
- Legal proceedings
- Administration

# Common Retention List

h_code	h_name_en	h_retention_perio	post_arp_action
7.1.1	Work programmes, annual action programmes, planning and reporting, including where appropri	5	THA
7.1.2	Procedures for award of grants	10	SAM_SEL
7.1.3	Management of grant agreements and decisions (financed projects)	10	SAM_SEL
7.1.4	Procurement procedures	10	SAM_SEL
7.1.5	Contract management (following award of contracts)	10	SAM_SEL
7.2.1	Implementation of agricultural aid (intervention, programmes and direct payments)	10	SAM_SEL
7.2.2	Implementation of aid programmes under the Structural Funds (ERDF, ESF, FIFG), the EAFRD a	10	SAM_SEL
7.2.3	Implementation of other Community aid programmes	10	SAM_SEL
7.3	Decentralised management of programmes, projects and grants	10	SAM_SEL
7.4	Indirect centralised management of programmes, financial instruments and pilot projects	10	SAM_SEL
7.5	Evaluation of expenditure programmes	6	THA

# Division of responsibilities



# Change management

- Importance of investing in training across the hierarchy :
  - Head of Units
  - Policy officers
  - Secretaries

# Contact

- [svhoolan@ulb.ac.be](mailto:svhoolan@ulb.ac.be)
- <https://twitter.com/sethvanhooland>
- Looking for trainees / employees ?
- Collaborate on case-studies ?



# Redefining the archival strategy at a multinational firm in the digital era

---

INFORMATION DIVISION

Archives & Records management Departement

Lourdes FUENTES HASHIMOTO

Head of the Project Management Unit

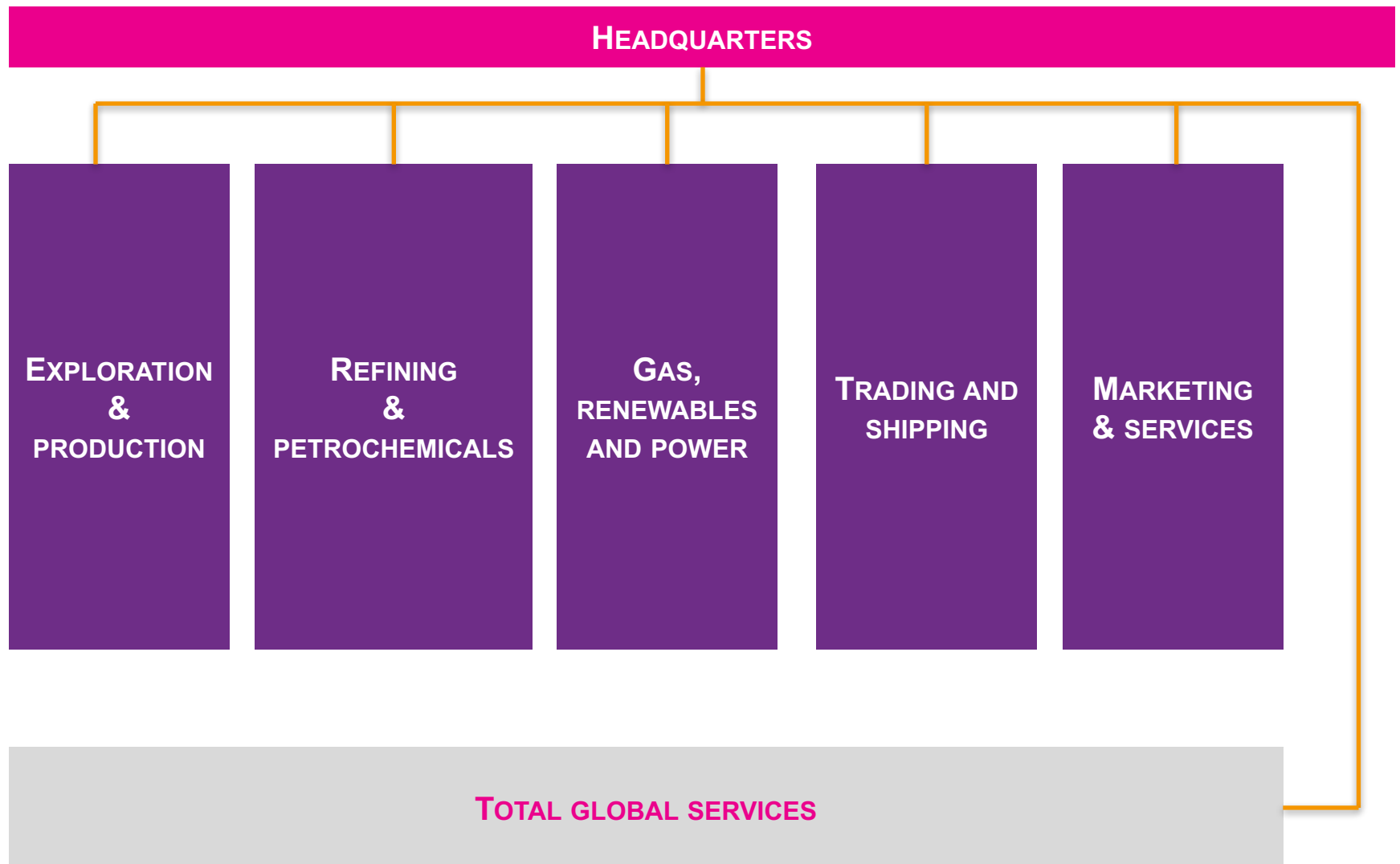


# Agenda

---

1. Total's organizational structure: developing information services at a multinational firm.
2. Total's records management policy and archival strategy : are we dealing with the right information produced by the company?

# 1) Total's organizational structure



**Secretary General**

**Total Global Information  
Technology Services**

**Total Global  
Financial Services**

**TOTAL GLOBAL  
SERVICES**

**7 ACTIVITIES**

**~ 1 600 employees**

**SOLUTION *Makers***

**Total Global  
Procurement**

**Total Global Human Resources Services**

**Total Consulting**

**Total Facilities  
Management  
Services**

**Total Learning  
Solutions**

# Information Services, Competitive Intelligence and Records & Archives Management at Total

---

- The Information Division belongs to Total Facilities Management Services.
- We are located in Paris.
- Our services should cover all the activities of the Group.
- We work with all of the branches and subsidiaries all over the world.

# Total at a glance : a major energy operator

---

**2nd**

FRENCH COMPANY  
BASED ON ITS  
ANNUAL TURNOVER

**4th**

LARGEST  
INTERNATIONAL  
COMPANY IN THE  
ENERGY SECTOR

PRESENT IN

**130**

COUNTRIES ON THE  
**5** CONTINENTS

**100 000**

EMPLOYEES

A WIDE RANGE OF  
ACTIVITIES AND  
PROFESSIONS  
COVERED

A VERY  
RESTRICTIVE LEGAL  
FRAMEWORK TO  
IDENTIFY AND APPLY

A RECORDS &  
ARCHIVES  
MANAGEMENT  
GENERAL STRATEGY  
SINCE  
**1980**

**3**  
INFORMATION  
GOVERNANCE  
POLICIES DEPLOYED

# TOTAL'S ACTIVITIES

## EXPLORE AND PRODUCE

- 1 OIL AND GAS
- 2 SOLAR
- 3 BIOMASS

## TRANSFORM AND DEVELOP

- 4 SPECIALTY CHEMICALS
- 5 POLYMERS
- 6 REFINING - PETROCHEMICALS

## SHIP AND MARKET

- 7 TRADING - SHIPPING
- 8 PRODUCTS AND SERVICES



# What kind of information does a multinational firm produce?



Oil and Gas



Solar



Bioenergies



Refining - Petrochemicals



Developing, Producing and Marketing polymers



Specialty chemicals



Trading and Shipping



Products and Services



## 2) Total's records management policy and archival strategy

---



1. Deploying one general retention schedule for the entire company : a challenging approach in a decentralized Group.
2. A specific strategy for vital records : identifying the most important documents within the company.
3. A new system for managing digital/non digital records & archives : how to collect and preserve digital materials?

# One framework, many different legal entities



## A specific strategy for « vital records » since 2005

- The company has never tried to control all the information produced, it will be impossible.
- Total launched a vital records program in 2005, the aim of this program was to identify within the company, and its affiliates the most important information that is necessary to prove our rights and obligations.

# How to collect and preserve digital archives?

---

- Total decided in 2014 to stop making the difference between digital/non digital records and archives.
- We have a new policy and a new system launched in 2016 : e-Traces.
- E-Traces is the one single point of entry for the users. If an entity has a project and asks for archival services, the Archives & Records management Departement provides assistance without thinking in terms of format.

# Avec TRACES toutes vos archives à portée de clic



e-Traces, l'application de gestion des archives physiques et numériques du Groupe Total est désormais ouverte aux utilisateurs.

Merci

[lourdes.fuentes-hashimoto@total.com](mailto:lourdes.fuentes-hashimoto@total.com)

Tél. : (+33) 1 41 35 56 55



# Appraisal policy and RM maturity

22-June 2017 | Vincent Hooft | Paris

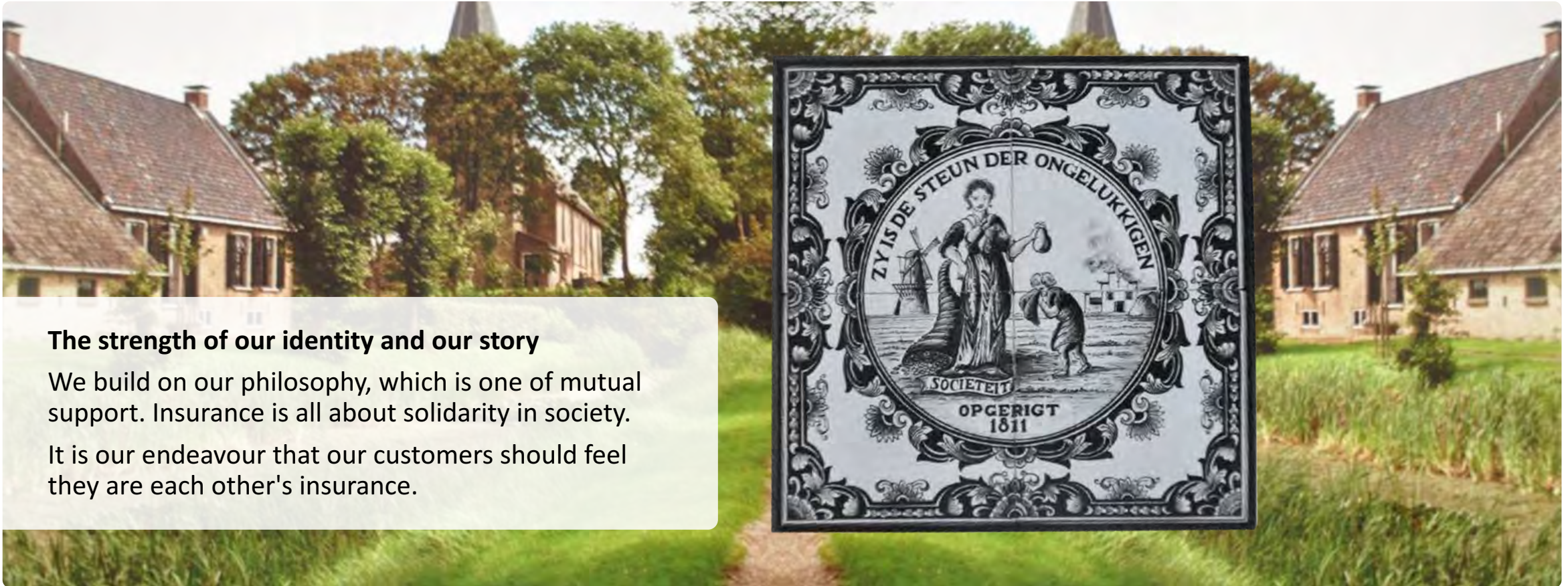
## Contents

- Achmea; history
- Records management policy
  - Appraisal
- Records management implementation
  - Maturity model
- Summary





**Est. 1811, today Achmea is the largest insurance provider in the Netherlands.**



**The strength of our identity and our story**

We build on our philosophy, which is one of mutual support. Insurance is all about solidarity in society.

It is our endeavour that our customers should feel they are each other's insurance.

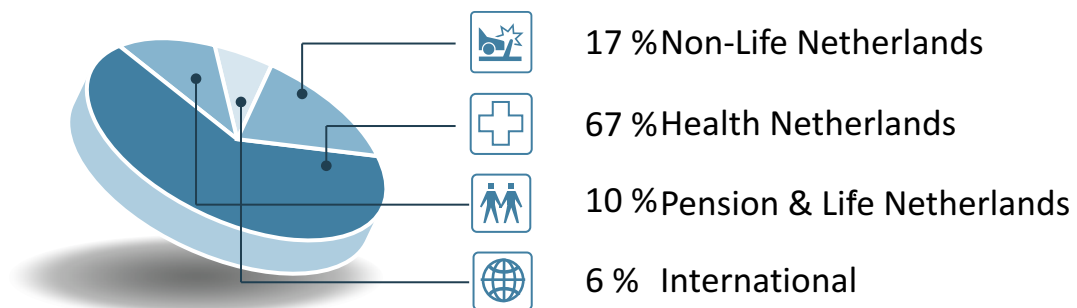
# Achmea's core activities: Non-Life, Health, Pension & Life and Asset Management

Non-life  
€3,2 billion

Health  
€13,1 billion

Pension & Life  
€1,8 billion

Assets under  
Management  
€116 billion



Distribution of Gross Written Premiums in 2016

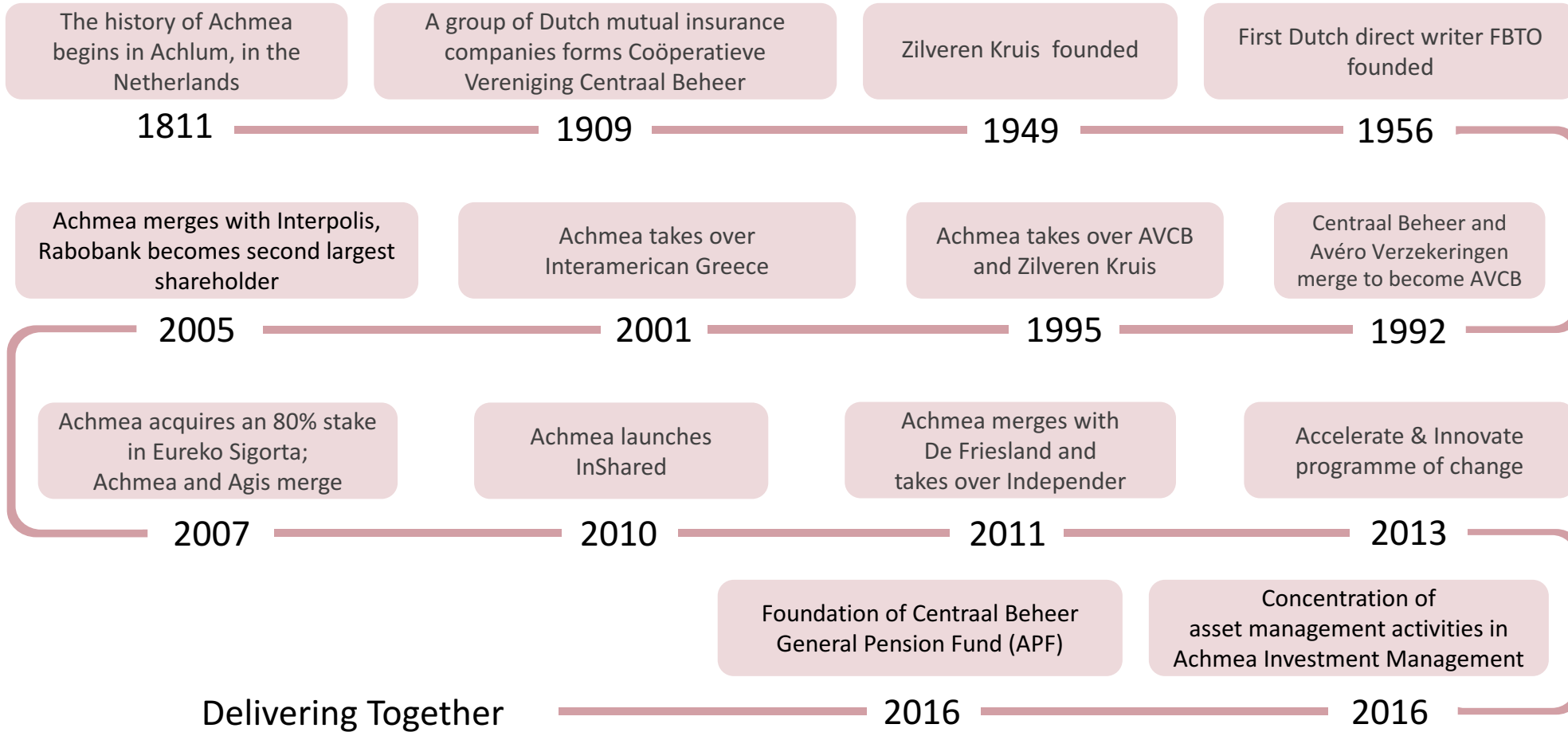
★★★★☆  
Dutch Financial Market  
Regulator (AFM) Customer  
Centricity Score

Avg. Customer  
Satisfaction Score  
Non-Life  
7,4

Avg. Customer  
Satisfaction Score  
Health  
8,0

Avg. Customer  
Satisfaction Score  
Life  
6,9

# Our history, from Achlum to where we are today



## Current state of the historical Archives of Achmea

- Recently established as part of the Records Management policy
- Haphazard collection of records:
  - Backlog of appr. 1.000 metres of paper records
  - Part of the collection at a provincial Archive
  - Some collections of electronic records of archival value exist, mostly of marketing material.
  - Appraisal? We're lucky to have some records
- However this is all part of a larger issue;
- **No Records Management policy and well-established RM practices at Achmea**



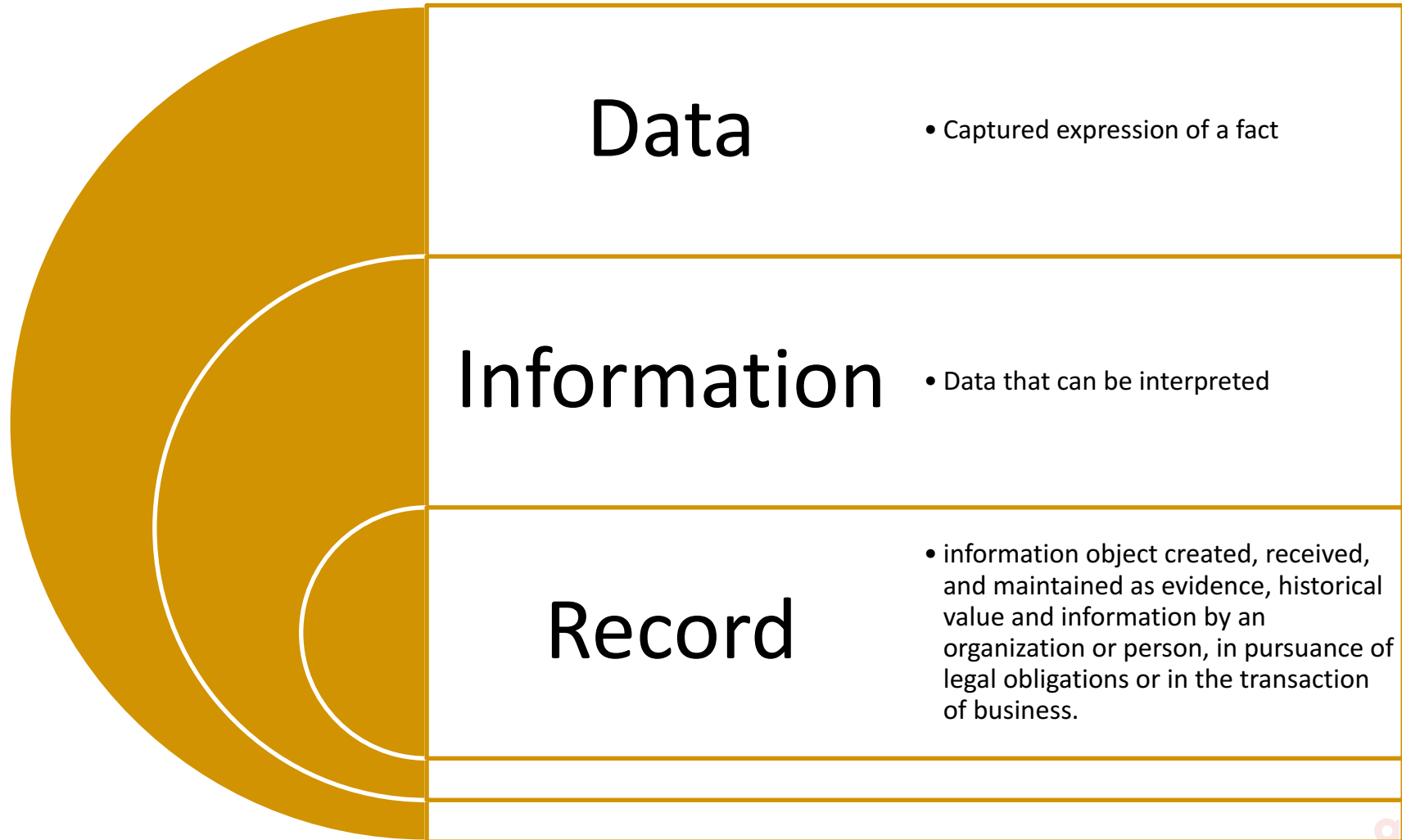
## Records management at Achmea

Records Management Policy approved in November 2016. This states:

- Definition of records
  - Historical value one of the criteria
  - Format independent
    - Meaning; paper, electronic, in databases, on social media etc.
- Relation to Data Governance
- Place of historical Archives as part of Records Management
- Governance model
- Monitoring and measuring state of Records Management
- Use of the ARMA Generally Accepted Recordkeeping Principles (GARP) maturity model
- Establishment of a Records Management programme for each Business Area of Achmea running since 2012.



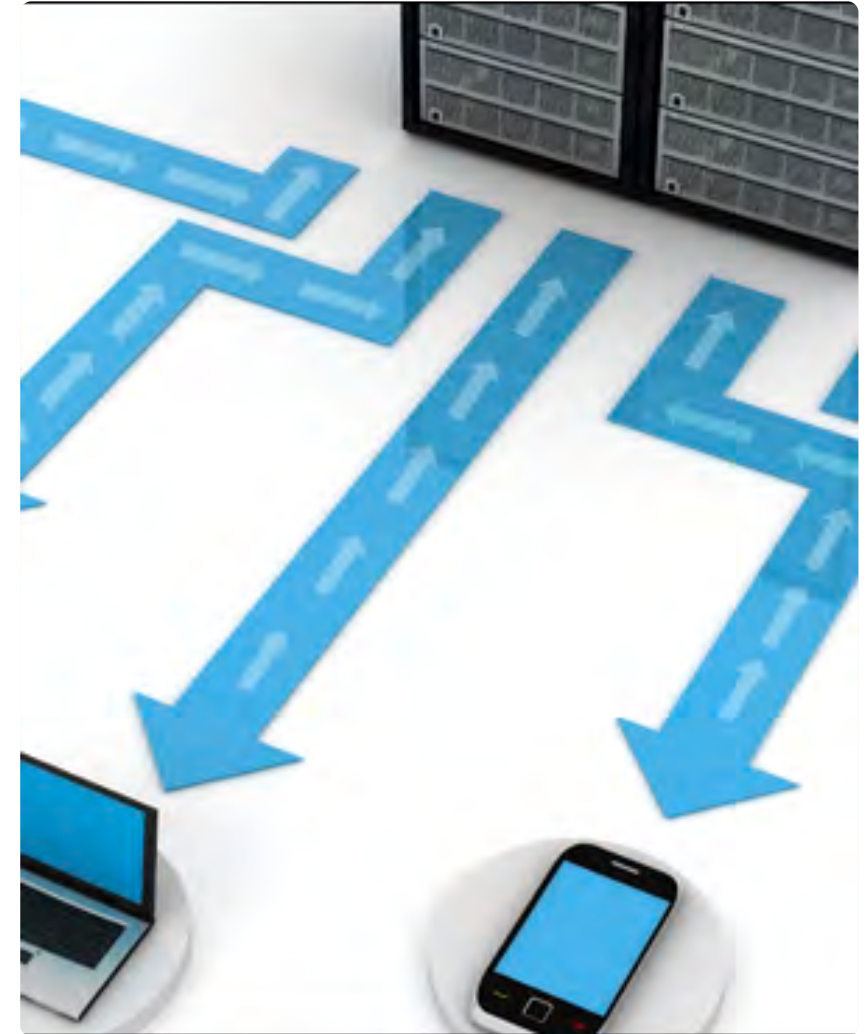
## From Data to Records



# Appraisal of Records as part of Records Management

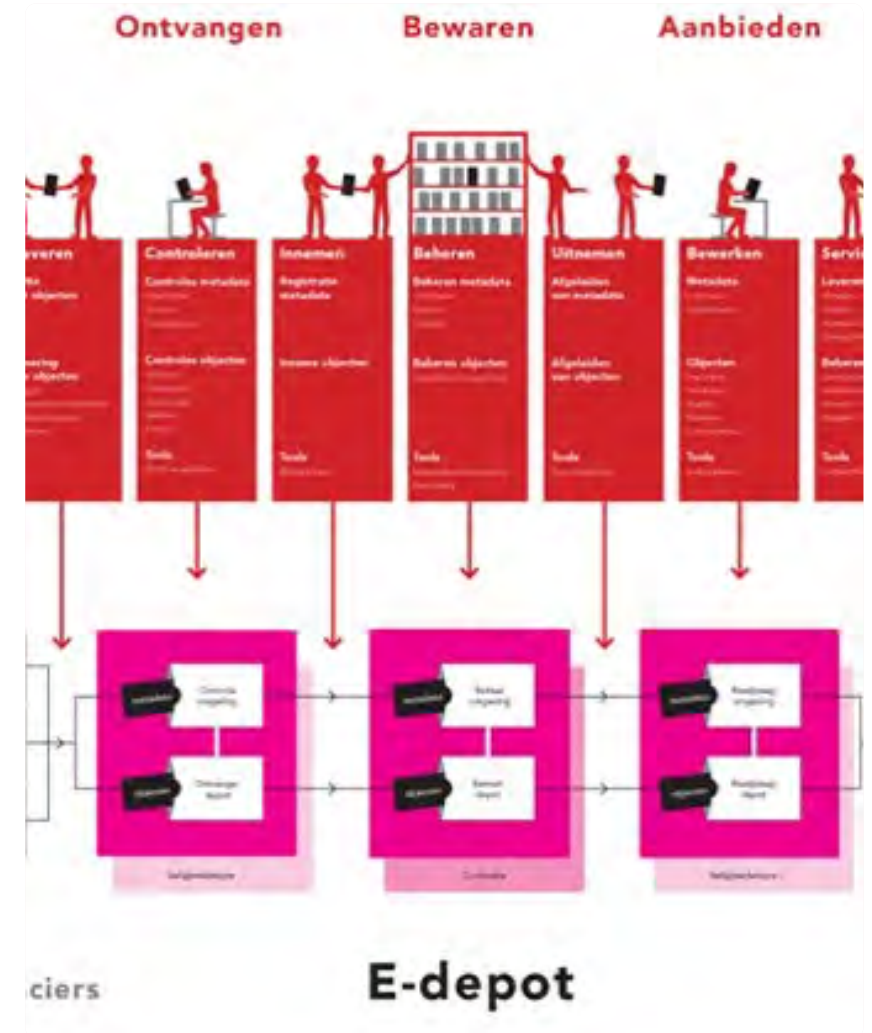
At Achmea we realised:

- Digital born records are the main source, also for historical archives.
- Example; replacement of employee magazine by vlogs on the Intranet
- Example; Facebook campaigns
- It's a flood, not a stream of information
- Without proper Records Management in place we'll never be able to ensure we retain the information of historical value
- We need the business to be pro-active



# Appraisal of Records as part of Records Management

- Appraisal is part of our records retention strategy
- Determining value of information in a process upfront ( i.e. before creation)
- Managing the lifecycle from the moment of creation
- Challenge: exceptional occurrences in standard processes with a finite life.
- Establishing Records Management programmes key in retaining electronic records
- Corporate project forces development of file and retention plan





# Implementation of Records management programme

## Standardisation of steps for each Business Area

- First step:
  - Gap analysis; 8 principles; current level vs. Desired level
    - Interviews with the business
    - Check with Records management policy
- Second step:
  - Project plan
  - Assignment of project responsibility within the business
  - Establishment of a Records Official role within the business.
- Third step:
  - Develop policies and procedures
  - Create inventory of all records and record holding systems
  - Create a file and retention plan
- Fourth step:
  - Implementation and/or improvements to record holding systems
  - Implementation of procedures
- Final step:
  - Monitoring and reporting established
  - New gap analysis

# Generally Accepted Recordkeeping Principles (ARMA)

8 principles:

1. Accountability
2. Transparency
3. Integrity
4. Protection
5. Compliance
6. Availability
7. Retention
8. Disposition

Arma based these on ISO 15489 and business.

Maturity model consisting of five levels:

1. Substandard
2. In Development
3. Essential
4. Proactive
5. Transformational

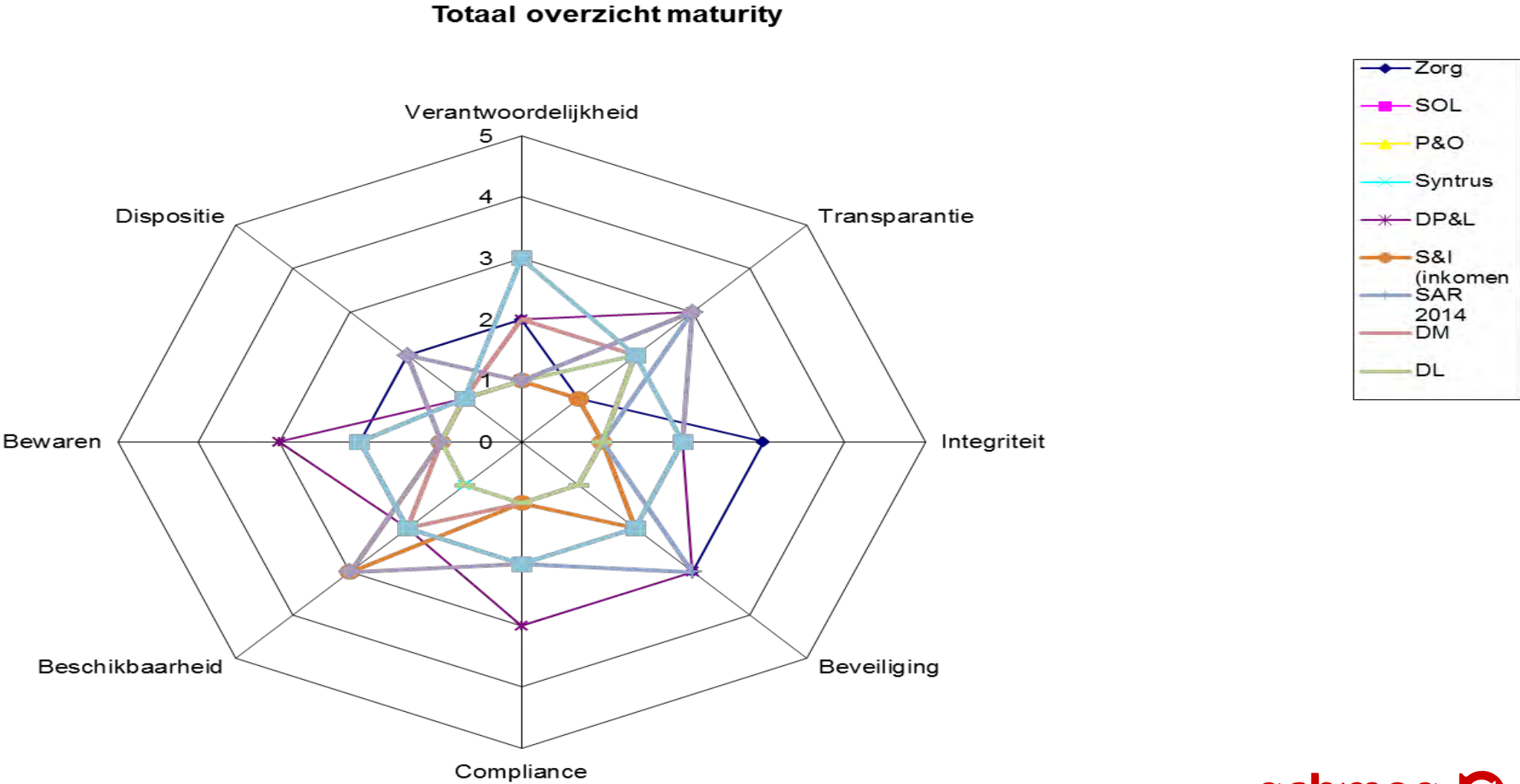
The Records Management Policy mandates level 3 for all Business Areas of Achmea.

## Reasons for using a maturity model

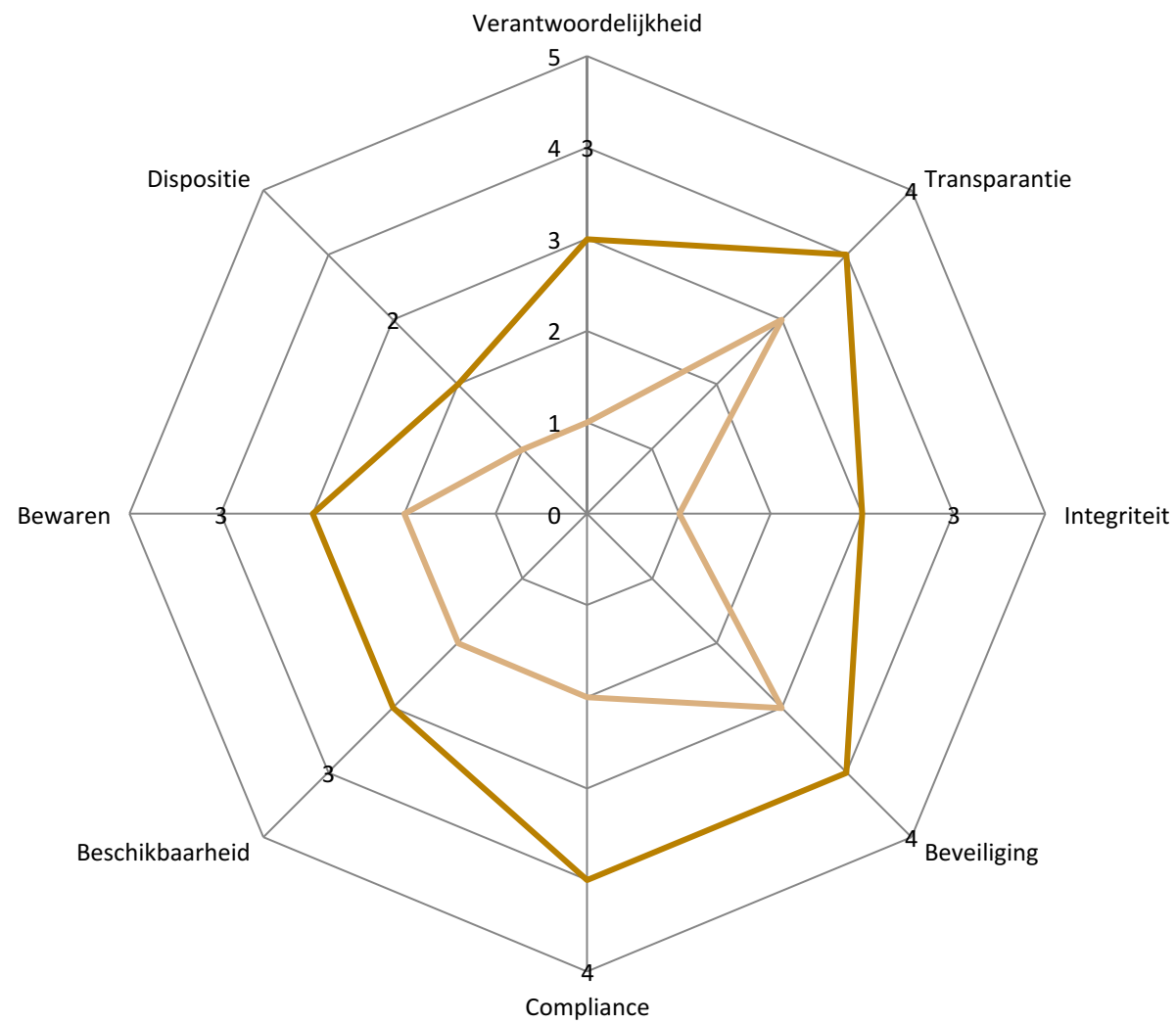
- Achmea is familiar with the concept of a maturity model to improve
- The Data Management Maturity Model is the core of the implementation of the data governance policy
- Achmea desires the use of standards and established models
- Creating a sense of urgency; call to action
- It provides a clear path for improvement.



# Results of the approach; overview status



# Results of the approach; before and after a business project



## Results of the approach; lessons learned

- Approach works
- Establish goals upfront
- It's slow going
- Reuse elements approach of Data Management Maturity;
- Establish corporate programme
- Work together with Data Governance and Compliance
  - They're in the same boat and want the same thing
- It's all about managing risk



## Summary

The Achmea Approach:

- Core is establishing Records Management practices in the business
- Electronic records, in all forms, can be of historical value
- Appraisal up front as part of the Records Management practices ensures we retain at least what we know we require
- Using an established maturity model is effective in a commercial environment
- You will always require a policy as the foundation.

```
1 <tr><td background="images/image01.jpg" width="845" height="10"></td></tr>
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43 <table border="0" width="845" height="371">
44 <tr><td width="125" valign="top" height="367">
45 <table border="0" width="123" height="302">
```

# Creating a Platform for Digital Appraisal at HSBC

James Mortlock, Digital Archives Manager, HSBC  
22 June 2017

**#HSBCArchives**

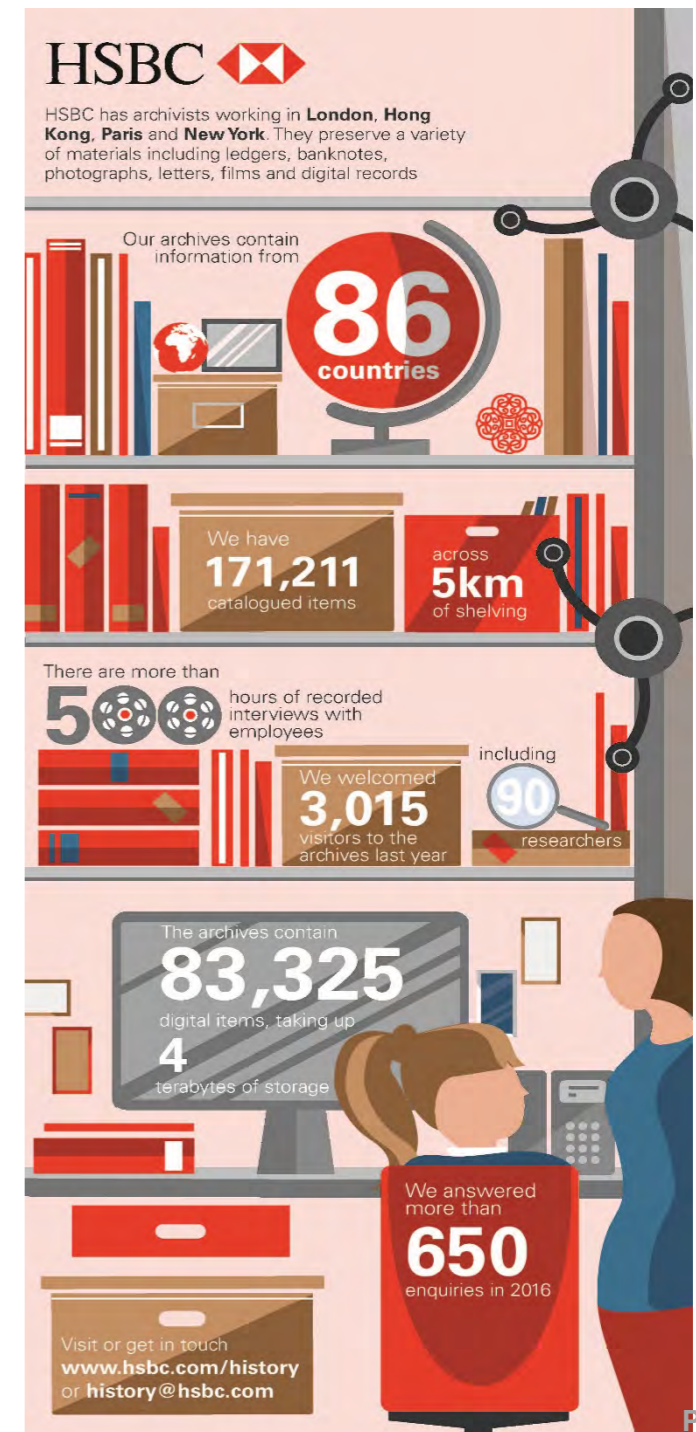
PUBLIC





## HSBC Archives – a brief overview

- Digital Archives system in operation since 2015
- Active collecting programmes > encountering born-digital records on a day-to-day basis
- Large volume of digital copies through project work; but no wholesale digitisation plans
- Close collaboration with Records Management, IT, Information Security, Legal and Compliance
- Physical records still play a key role; strong focus on enhancement of descriptive and management metadata



## HSBC's global digital archives system



**Preservica's Enterprise Edition software** provides secure ingest, storage and preservation strategies to safeguard born-digital records and digital copies of physical records. It provides sophisticated search and access functionality.



**Axiell's CALM software** provides regional Archive catalogues for both physical and digital content. It stores descriptive metadata and provides search functionality. It does not store the actual content; one-click links take the user to digital records.

## System design

- 4 regional Calm catalogues covering our four archival regions (HQ/UK/US/HK) with FR to be added this year. Customised API communicates with the digital repository
- 1 instance of Preservica Enterprise Edition (customised for HSBC use) which creates and links to records in Calm.
  - Preservica creates records in Calm upon ingest
  - All catalogue metadata from Calm synchronised to Preservica on a nightly basis (including physical records)
- Digital repository and associated cataloguing system have been designed to follow OAIS (where possible)
- Hosted in the UK
- Hourly snapshot back ups, we also have Disaster Recovery and Testing systems in place

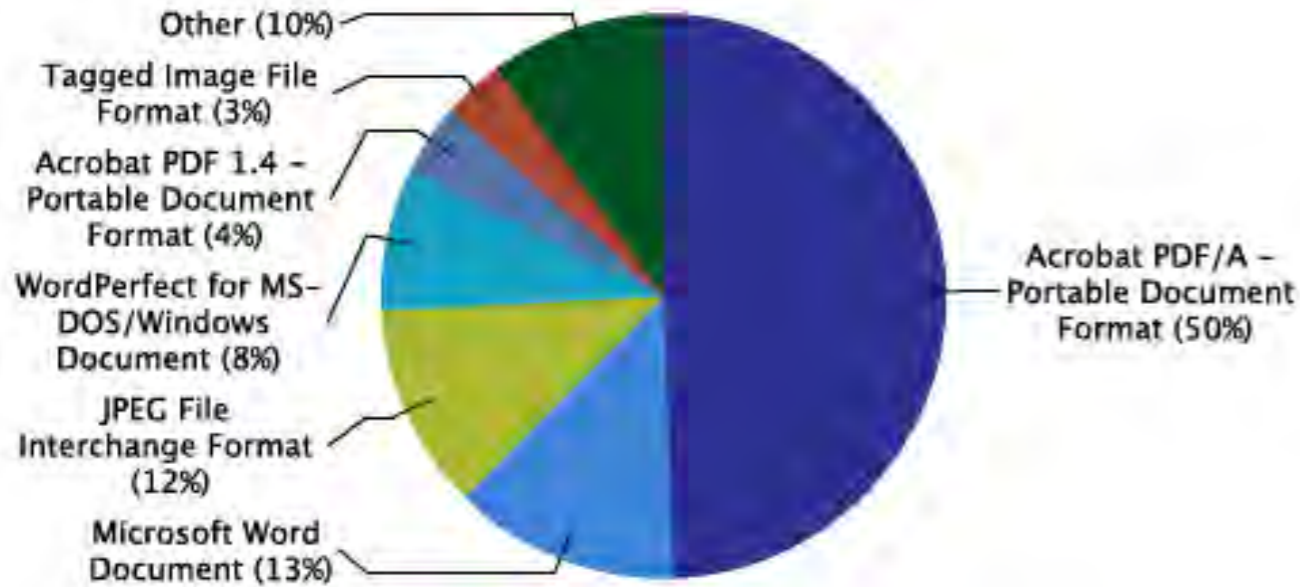
## The HSBC journey...

- Peer research and initial scoping of potential solutions
- Robust business case presented and a detailed requirements document produced
- Formal RFI/RFP process carried out; valuable input from business analysis to reassess all procedures (physical and digital)
- Vendors selected and contract signed
- Development work started with vendor; extensive collaboration required with IT, ISR and Legal
- Testing; migration of data/records; staff training
- System goes live at the start of 2015
- Ongoing research and training requirements
- Regular software upgrades and occasional hardware upgrades; disaster recovery testing



# A growing platform

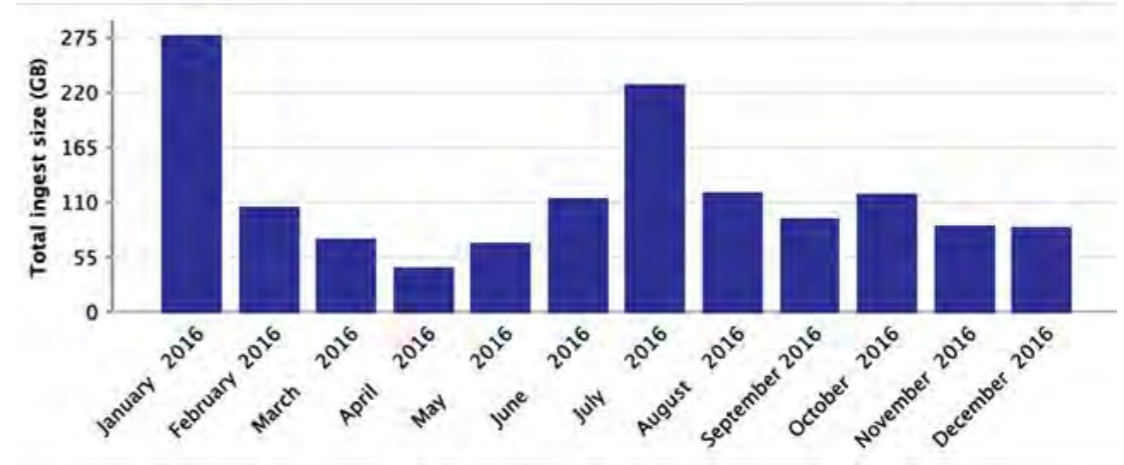
## File formats in GDA



## Ingests by size in 2016

Start date: 01 Jan 2016

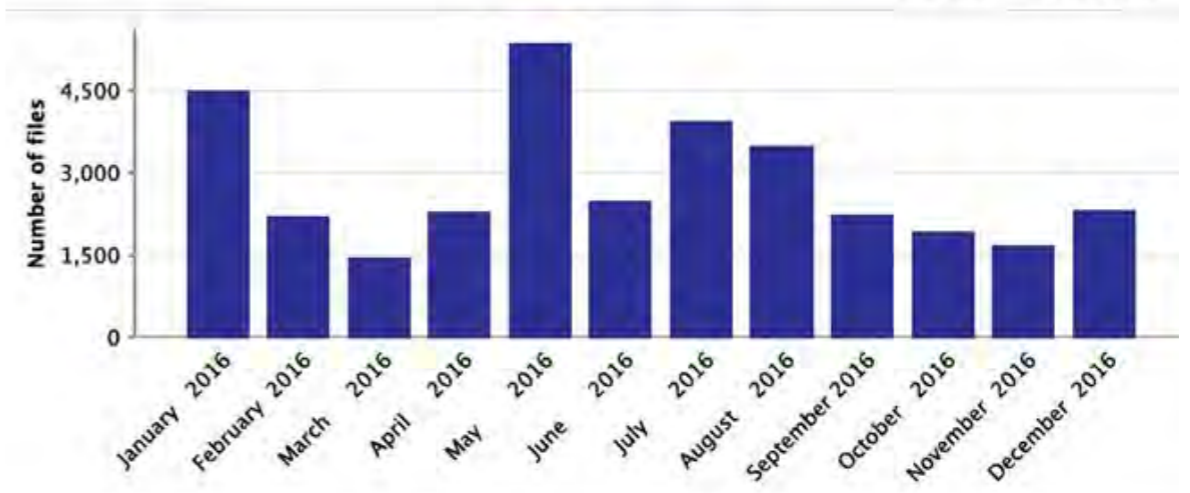
End date: 31 Dec 2016



## Ingests by file submission in 2016

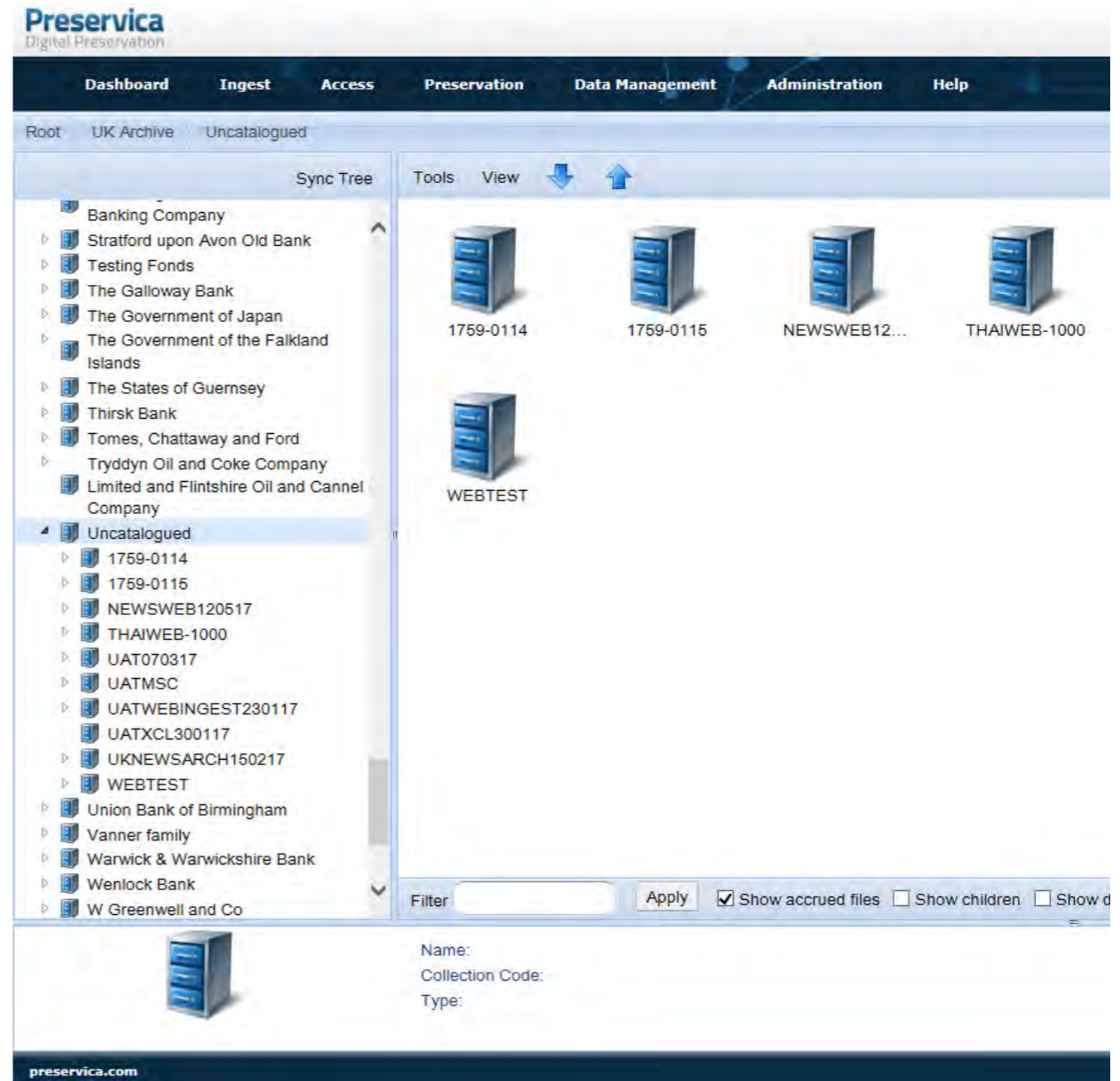
Start date: 01 Jan 2016

End date: 31 Dec 2016



## Why post-ingest appraisal?

- Security (Integration with information governance rules)
- Compliance checks (ISR file format agreements)
- Depositor expectations
- Audit trail
- Preservation
- Storage costs
- Ability to review (e.g. searching through contents)
- Better understanding of the accession (e.g. characterisation of file formats)
- Indexing and technical metadata profile
- Provision of robust disaster recovery support



## Post-ingest appraisal in practice

- Accession reference generated (regardless of whether collection is accepted)
- Sensitivity review (security tagging)
- Automated or manual SIP ingest
- Technical metadata harvesting
- Searching/discovering of deposit
- Removal of unwanted material (audited, but reasons documented in accession record)

The screenshot displays a digital ingest workflow interface. At the top, a 'Workflow Details' box provides key information:

Property	Value
Workflow Context	HQ Born Digital Ingest
Workflow Definition	HSBC Born Digital Ingest
Workflow ID	4049104
Workflow State	Completed
Date Started	09.03.15 09:04:44
Date Finished	09.03.15 09:55:42
Number of Files	252
Total Size	4 GB
Collection Code	HQ 4GB
Submission name	Images in introduction +
Top Level Record	Images in introduction +

Below the details is a 'Back' button. The main section is titled 'Step Progress' and contains a table with the following columns: State, Name, Progress, Started, Finished, and Messages.

State	Name	Progress	Started	Finished	Messages
	Select		09.03.15 09:04:44	09.03.15 09:04:58	
	Copy XIP Package		09.03.15 09:04:58	09.03.15 09:10:33	
	Validate CALM Mapping Document		09.03.15 09:10:34	09.03.15 09:10:36	
	Locate CALM Instance from Region within XIP		09.03.15 09:10:37	09.03.15 09:10:39	
	Metadata Integrity		09.03.15 09:10:39	09.03.15 09:10:42	
	Content Integrity		09.03.15 09:10:42	09.03.15 09:10:45	
	Fixity Check		09.03.15 09:10:45	09.03.15 09:13:39	
	Characterise		09.03.15 09:13:39	09.03.15 09:17:12	<a href="#">View</a>
	Transfer Agreement Compliance Check		09.03.15 09:17:15	09.03.15 09:17:24	
	Create CALM Skeletal Record		09.03.15 09:17:24	09.03.15 09:18:21	
	Store Files		09.03.15 09:18:21	09.03.15 09:39:30	
	Store Metadata		09.03.15 09:39:30	09.03.15 09:40:39	
	Store Metadata File		09.03.15 09:40:39	09.03.15 09:41:24	
	Update Search Index		09.03.15 09:41:24	09.03.15 09:45:33	
	Thumbnail Creation		09.03.15 09:45:33	09.03.15 09:55:39	
	Delete Source XIP		09.03.15 09:55:39	09.03.15 09:55:42	

## Future appraisal plans

- Automated depositing
- Records Management integration
  - Historical records collecting framework
- Depositor access
- Improved discovery tools
- Systems integration
- Open to new ideas from the sector...



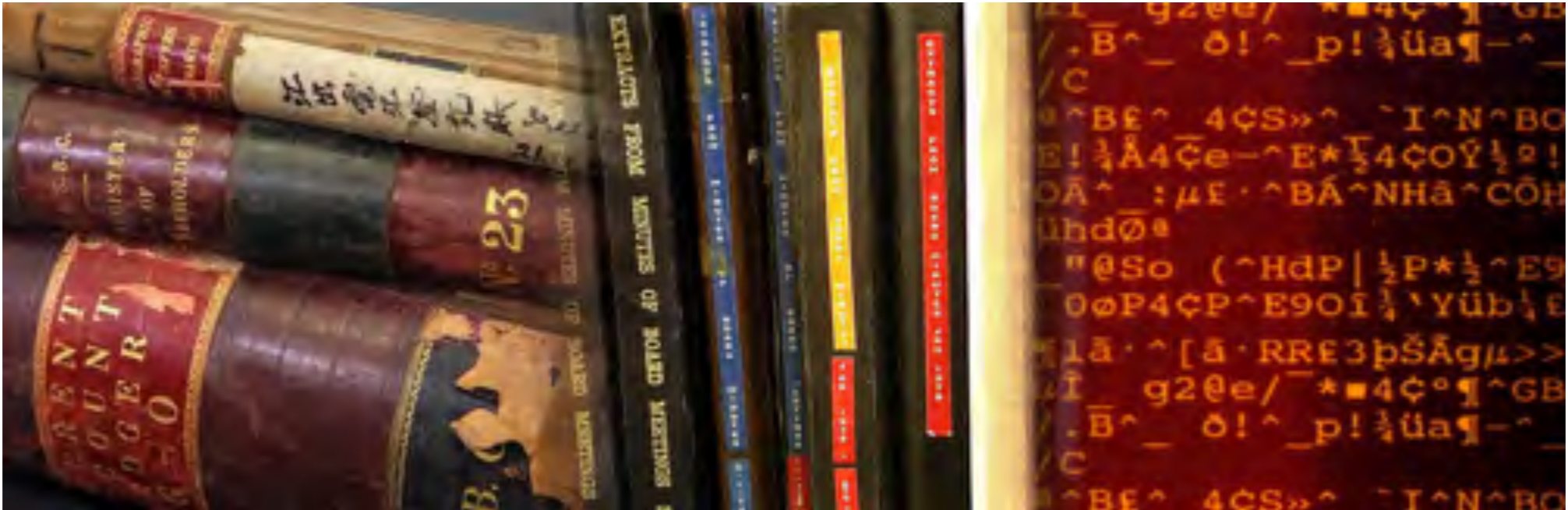


## Should we retain the original file structure?

- Can we be sure of what the 'original' file structure was?
- Does the folder structure contain contextual value?
  - Valuable information on how the records were used
  - ...but also masking records
- How do we judge a creator's intent? Can assessing an existing file structure lead to subjective judgements by the archivist
- For the corporate archivist working with a strong records governance framework, do we have access to the original file structure?
- Should we have to make case-by-case judgements on large digital accessions?
- If it is a matter of access to records, do new discovery techniques remove the need for structure?
- The role of technical metadata – does this make the maintenance of an original structure unnecessary?

## Moving beyond functional macro-appraisal

- The same benefits for large organisations and businesses in the digital world as currently exist for physical records
- Greater level of digital record creation and maintenance means the modern archivist is faced with more records to appraise
- Requires a greater reliance on collecting frameworks for automated collection and appraisal
- Functional context easier to appraise in a large organisation than individual record types
- Reconciling discovery techniques with parsimonious appraisal





Thank you

➤ [www.HSBC.com/history](http://www.HSBC.com/history)

➤ [history@hsbc.com](mailto:history@hsbc.com)

➤ [#HSBCArchives](https://twitter.com/HSBCArchives)



# Automatic Appraisal in the era of GDPR

# Challenges of the digital information age

*The Internet of Things, streaming video, machine-to-machine data, emails, texts, and social media/mobile/web data + shrinking budgets is forcing organizations to rethink how they manage and govern their data*



# And the trends that affect our ability to manage information



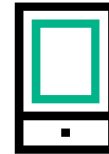
## Evolving Services

- Online 24/7
- Single view
- Timeliness & accuracy
- Efficiency of service
- Productivity & delivery



## Digital Transformation

- Cost to manage paper
- Manual vs Automation
- Volume & variety
- Streamlined processes
- Long term accessibility



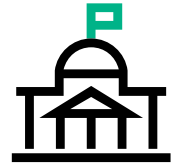
## Mobility

- Access anywhere
- Range of devices
- Support field workers
- Asset management



## Cloud

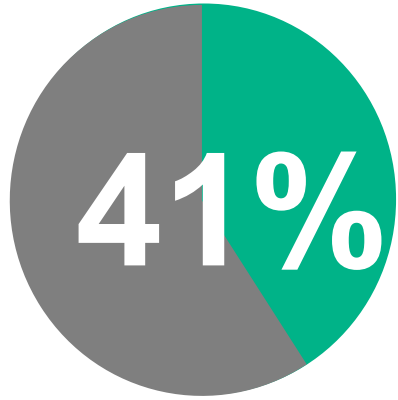
- Cost of legacy systems
- Scalability & flexibility
- Data sovereignty
- Private, Public or Hybrid



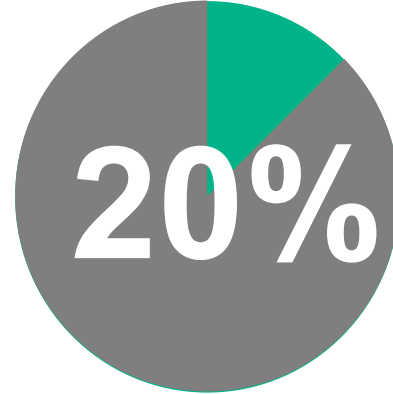
## Global Transparency

- Statutory compliance
- Privacy & security
- Safe Harbor, GDPR
- Auditable records/archive

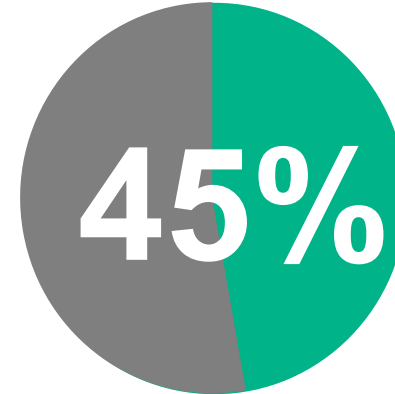
# How mature are our Policies, People and Systems?



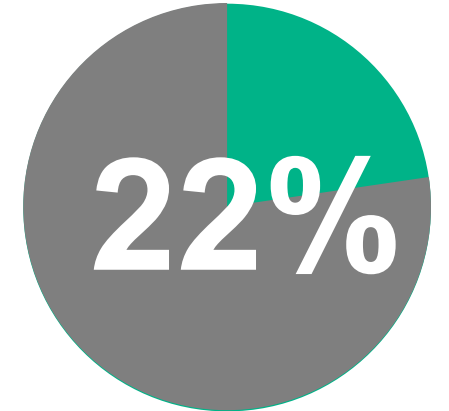
of organizations describe their email as **"chaotic"**



of organizations say **bad IG** practices is the most likely cause of **data loss**



of organizations say a lack of IG leaves them open to litigation risks

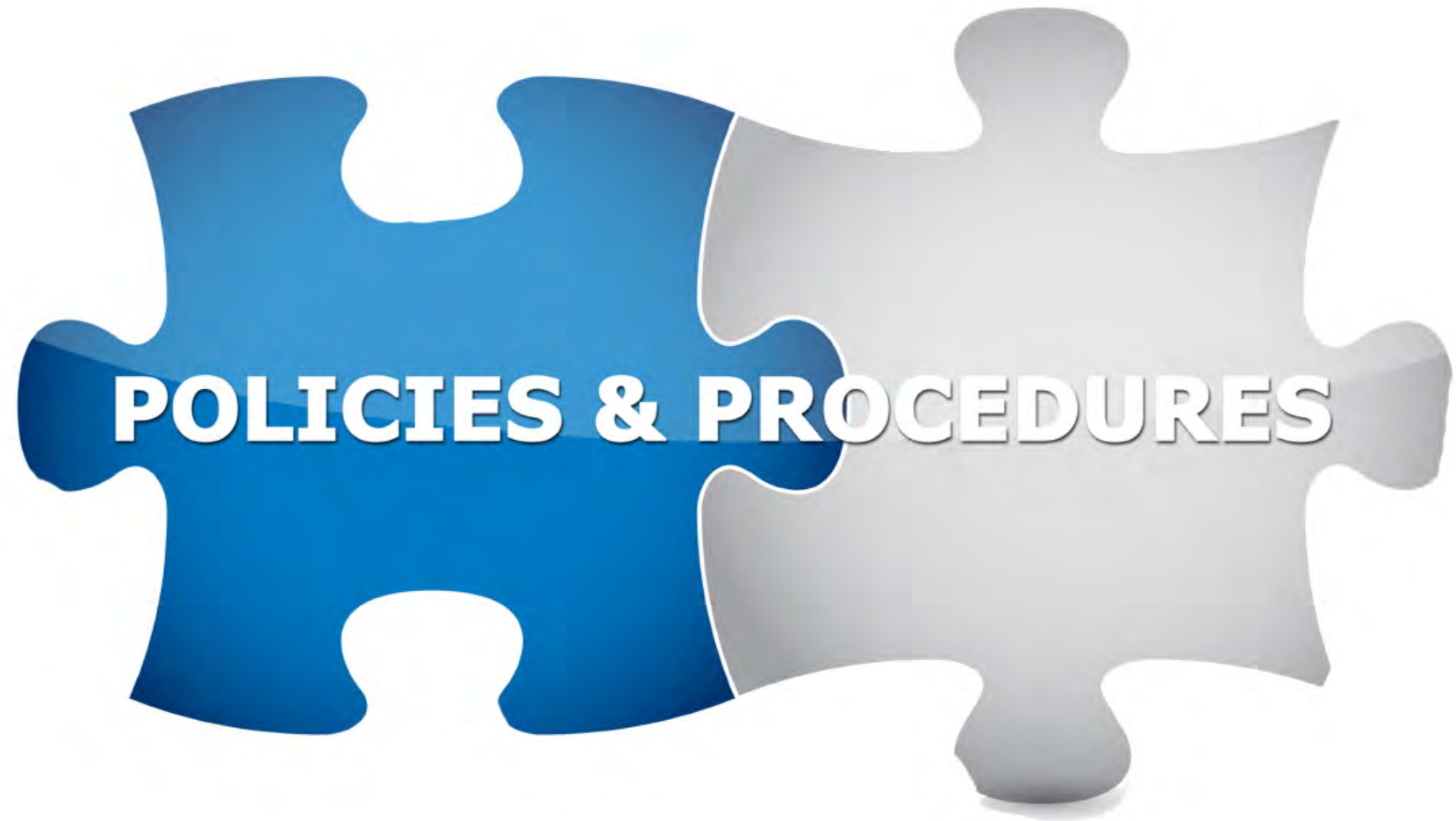


reported **negative financial impact** from cases involving electronic records

---

# Information Governance

For internal or organisational reasons





**“We have real issues in successfully capturing and selecting the right digital information so as to avoid non-appraisal”**

Head of Archiving at a European National Bank.

---

# ...and Information Governance

For external reasons.



---

# An Example: GDPR

- GDPR replaces the previous Data Protection Directive
- Data Protection Directive created to regulate control of Personal Data
- GDPR seeks to harmonize data protection laws across 27 EU member states
- Clearer rules for data transfer across borders
- Better control over individuals' Personal Data



**“The level of risk associated with the GDPR has catapulted data protection into the boardroom.”**

Jane Finlayson-Brown, PARTNER, Allen & Overy

---

# What Information is affected?



- Any information by which a individual can be identified
- No distinction between a person in private, public or work roles
- Obvious type:
  - Name, email, national identification number, address
- Characteristic data
  - Medical, physical, genetic, mental data
- Inferred data
  - IP Address, cookies, profiling data

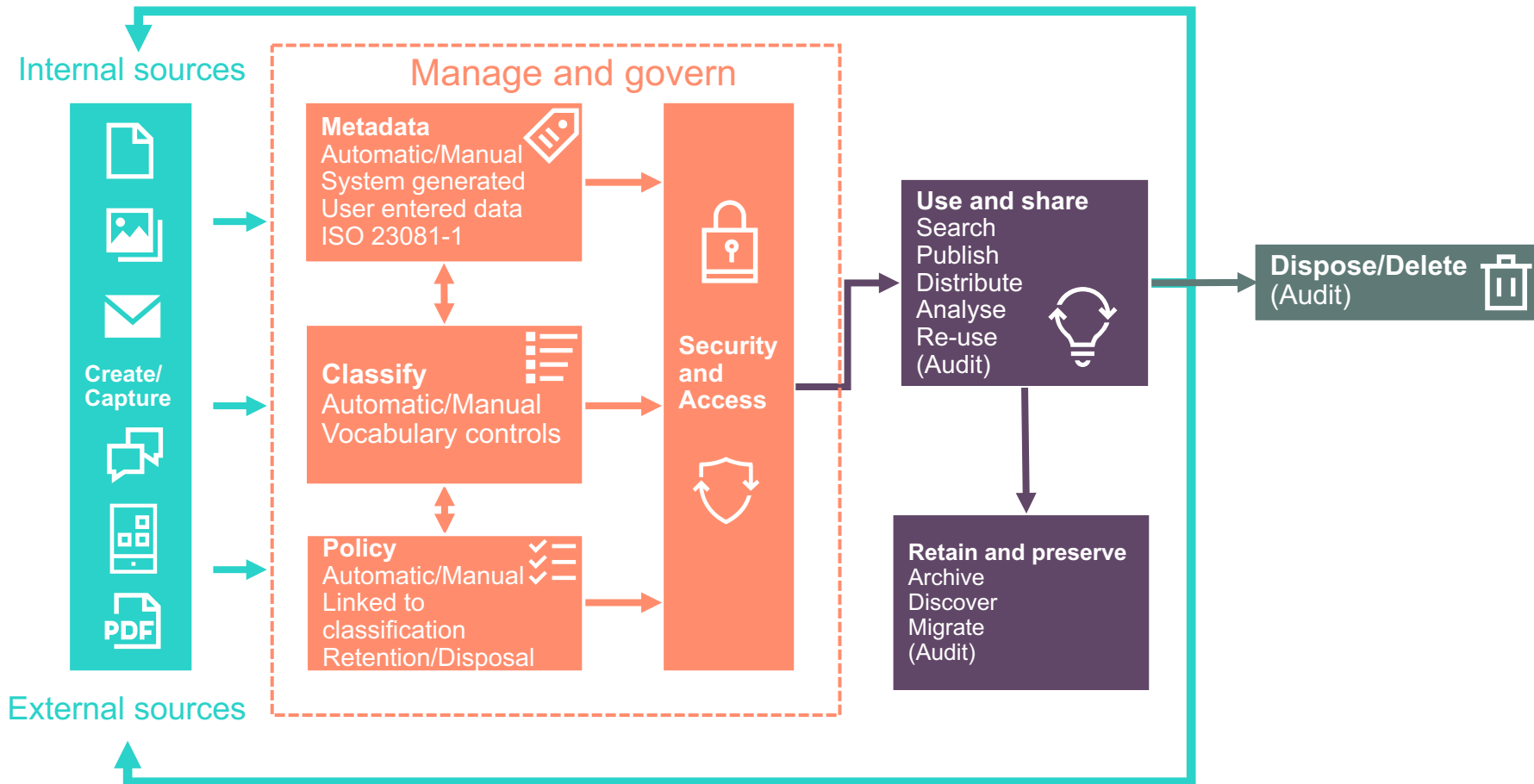
**Anything that be linked directly to a person**

---

# What actions need to be taken under GDPR?

- **Tracking and Management of Consent**
- **Data Protection by Design**
- **Formulating Retention Decisions on Personal Data**
- **Implement Data Breach reporting**
- **Appointment of a Data Protection Officer**

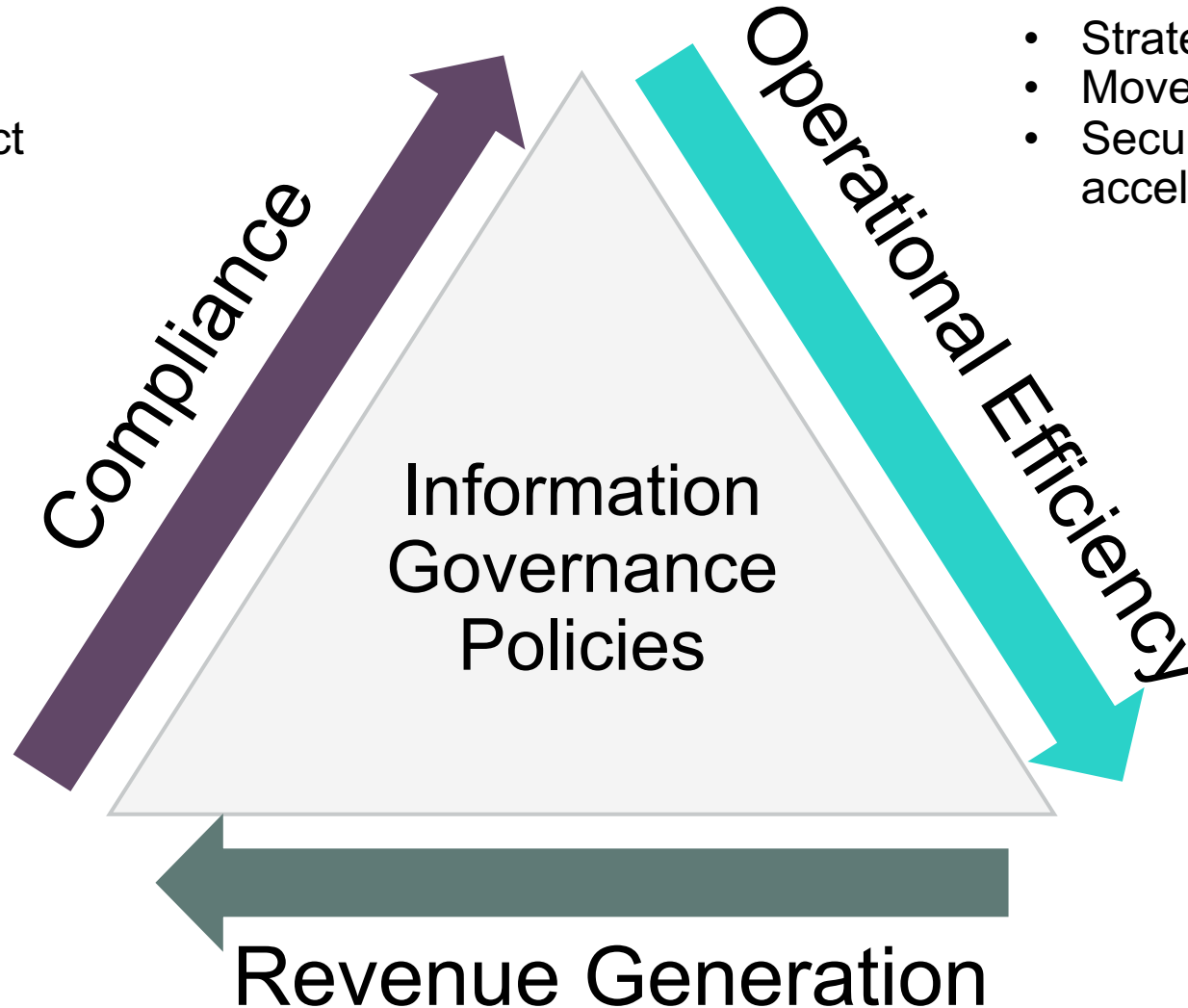
# Compliance hinges on Lifecycle Management



# Compelling business logic for sound appraisal strategies

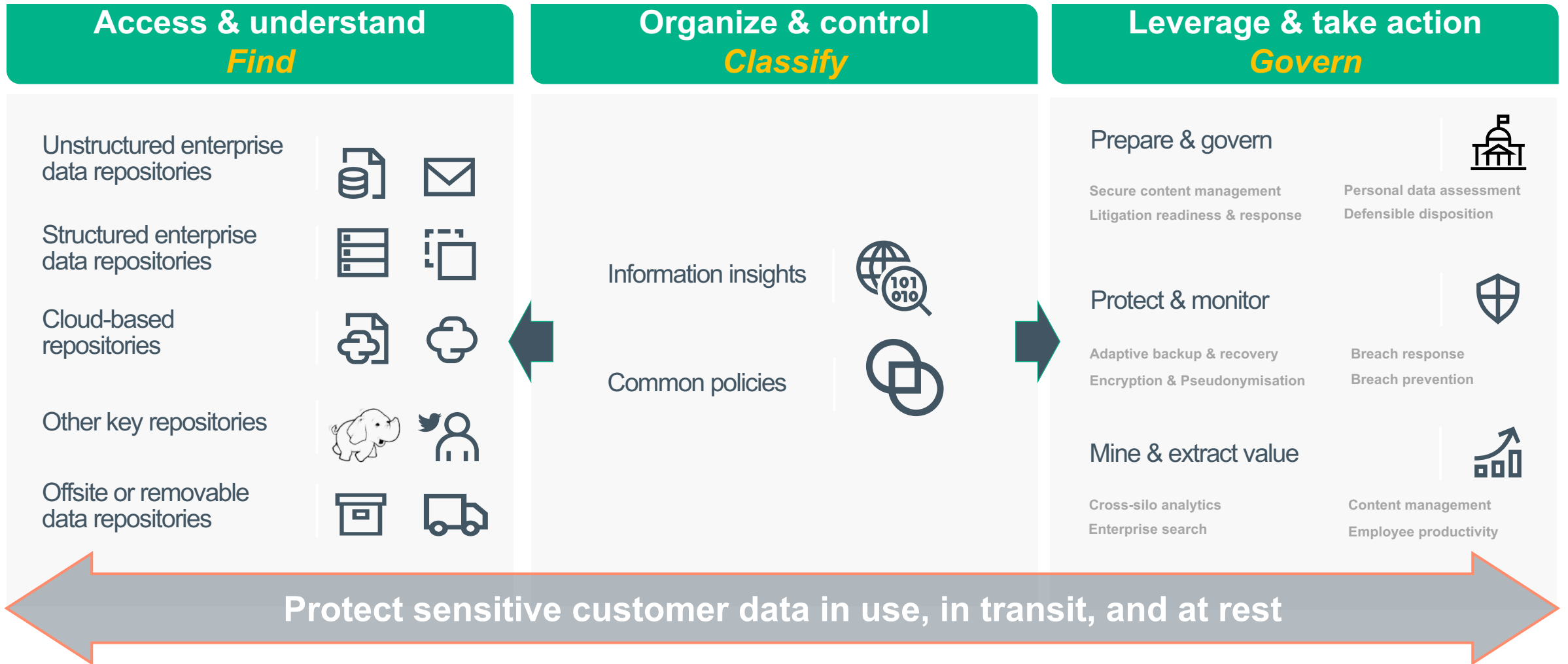
- Fine
- Reputation damage
- Government contract pre-requisite
- Enforcement action

- Strategic records management
- Move to cloud accelerator
- Security and compliance accelerator





# HPE's long-term Information Governance vision

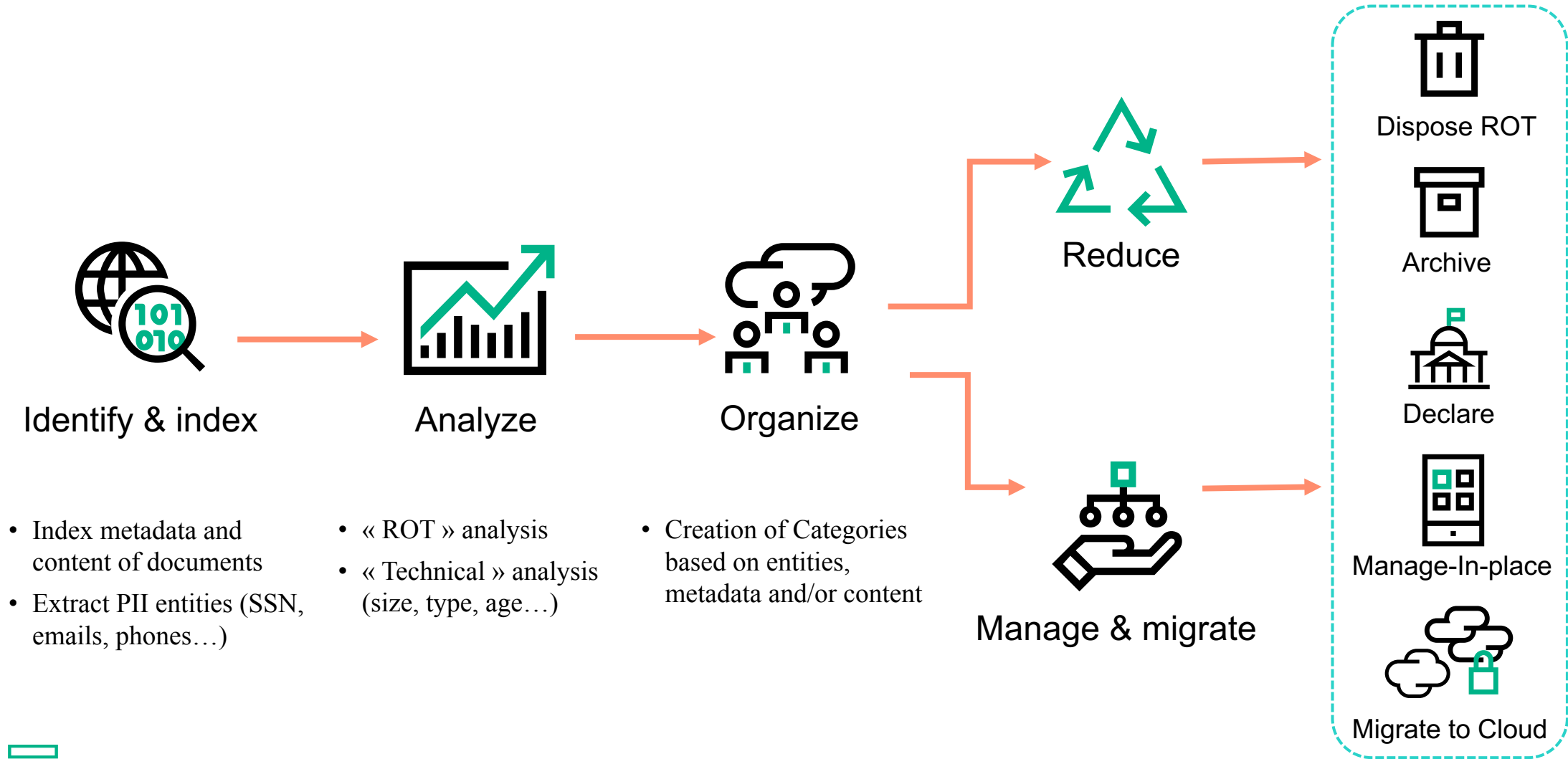




# Automatic Appraisal of Unstructured Data

# Unstructured Data Analysis in five steps

With HPE ControlPoint



# Content curation Dashboards

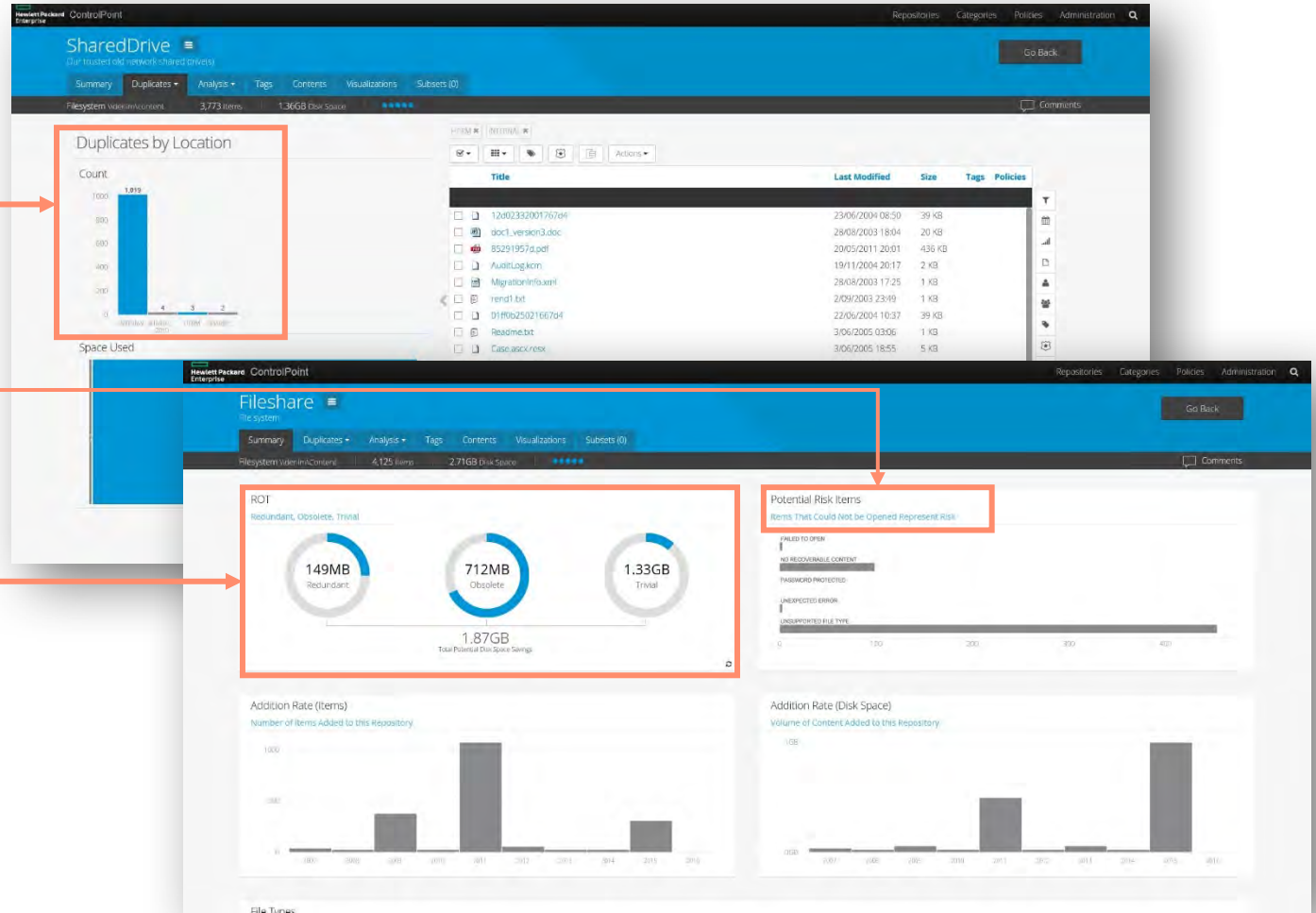
Policy to drive content clean-up, management or migration

## Identify

- ROT data
- Items of interest to protect
  - Sensitive information
  - Records and critical content
  - Intellectual property

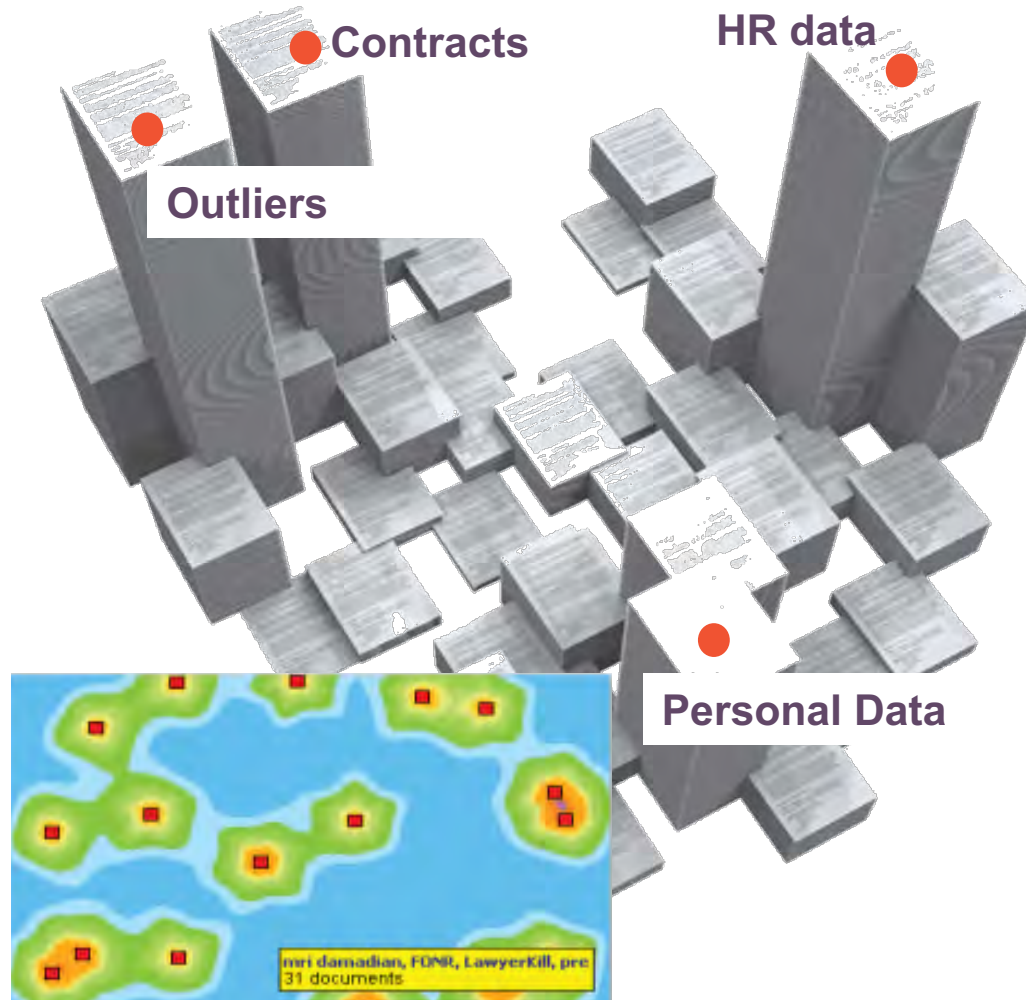
## Apply policy to drive action

- Dispose
- Migrate
- Manage-in-place



# Automatic Data Classification and Clustering

- Group data by topic automatically
- Build categories using training documents, Booleans, keywords and confidence degrees
- ...and Grammars. Add structure to fundamentally unstructured information



- Categories
  - TRAINED CATEGORIES
  - LIFECYCLE CATEGORIES
    - 1 - Transient Content
      - eTrash
    - 2 - Work in Progress
      - Email - Content Disposal
      - Filesystem - Content Disposal
      - Filesystem - Content Migration
      - SharePoint - File Accounting Conte...
      - SharePoint - File Contract Drafts
      - SharePoint Retention
    - 3 - Business Records
      - ACC - Accounting Statements
      - CON - Contracts**
      - FIN - Financial Reports
      - HR - Job Descriptions
      - MED - Medical Reports
      - MKT - Marketing Brochures

# Visualisation and Sample Data

The screenshot displays the HP ControlPoint Filesystem Visualizations interface. The top navigation bar includes 'Repositories', 'Categories', 'Policies', and 'Administration'. The main content area is titled 'Content Visualization' and features a 'Cluster Map' on the left and a file list on the right.

**Cluster Map:** A heatmap visualization showing several distinct clusters of data points, primarily in shades of green and yellow, set against a blue background. Below the map, the current mode is identified as 'Sockeye Salmon, Subsistence Sockeye Salmon, Fisheries, AYK'.

**File List:** A table listing 20 PDF files with columns for Title, Last Modified, Size, and Policies. The files are sorted by last modified date.

Title	Last Modified	Size	Policies
214076619.pdf_4d018cd6f38fc3cfa4490f6135135d8e.pdf	20/05/2011 13:53	245 KB	
57409803.pdf	20/05/2011 12:15	358 KB	
57240425.pdf	20/05/2011 12:15	270 KB	
57240225.pdf	20/05/2011 12:15	564 KB	
57239851.pdf	20/05/2011 12:15	727 KB	
57239745.pdf	20/05/2011 12:14	707 KB	
57207264.pdf	20/05/2011 12:14	518 KB	
57207211.pdf	20/05/2011 12:14	316 KB	
57207166.pdf	20/05/2011 12:14	573 KB	
57174410.pdf	20/05/2011 12:14	379 KB	
57174118.pdf	20/05/2011 12:14	596 KB	
56471157.pdf	20/05/2011 12:13	399 KB	
54475697.pdf	20/05/2011 12:08	470 KB	
54404175.pdf	20/05/2011 12:08	776 KB	
50307639.pdf	20/05/2011 12:06	118 KB	
41860778.pdf	20/05/2011 12:00	488 KB	
177297049.pdf	20/05/2011 13:33	247 KB	
124077818.pdf	20/05/2011 13:13	175 KB	
123199291.pdf	20/05/2011 13:13	493 KB	

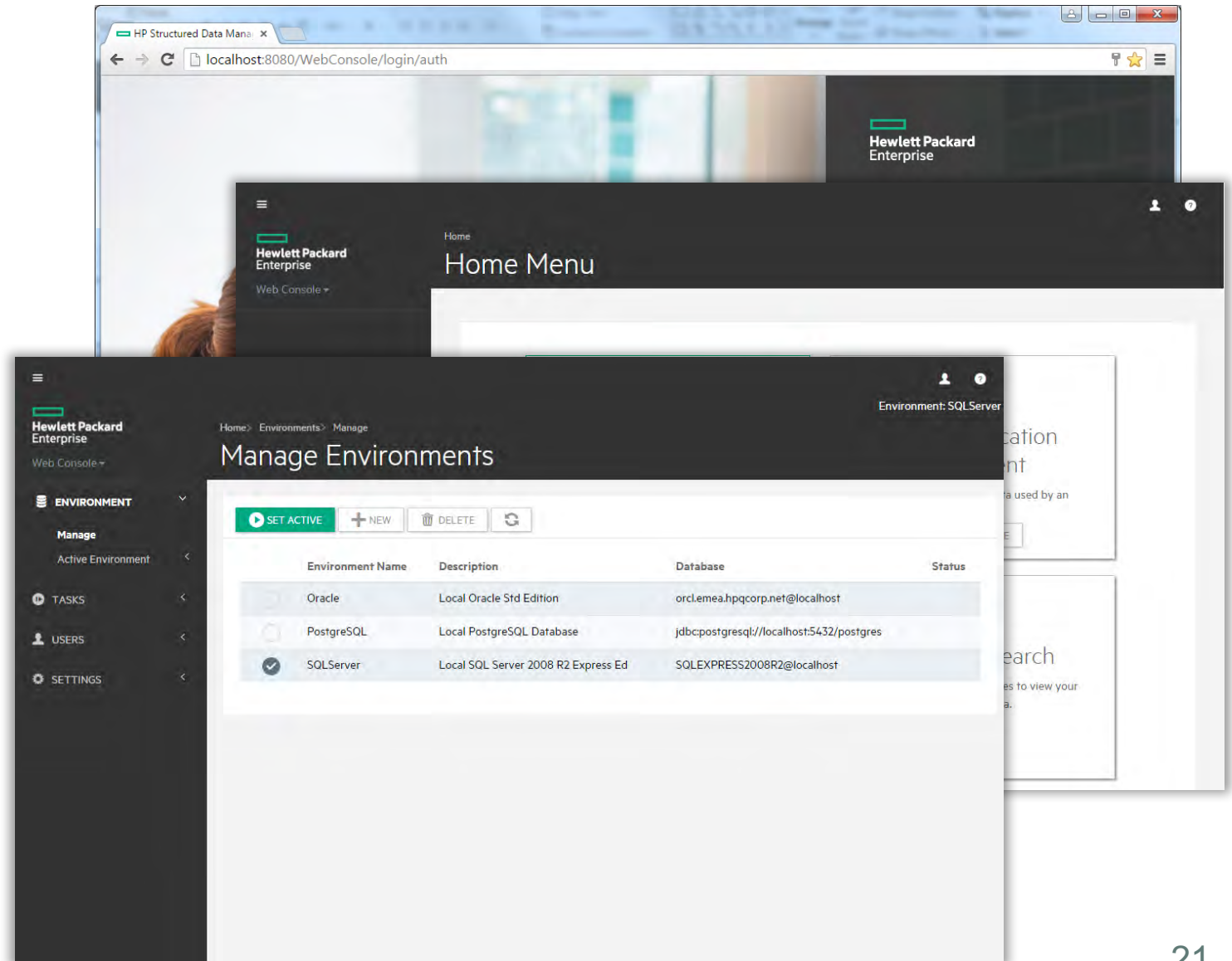
Items 1 - 25 of 40



# Automatic Appraisal of Structured Data

# Analyze your Line Of Business applications

- Browser based management console
- Application Retirement “Wizard”
- Reporting Interface for Archived Data
- Free Text Search
- Custom Interfaces (Extensions)



The screenshot shows the Hewlett Packard Enterprise Web Console interface. The main content area is titled 'Manage Environments' and displays a table of environments. The table has the following data:

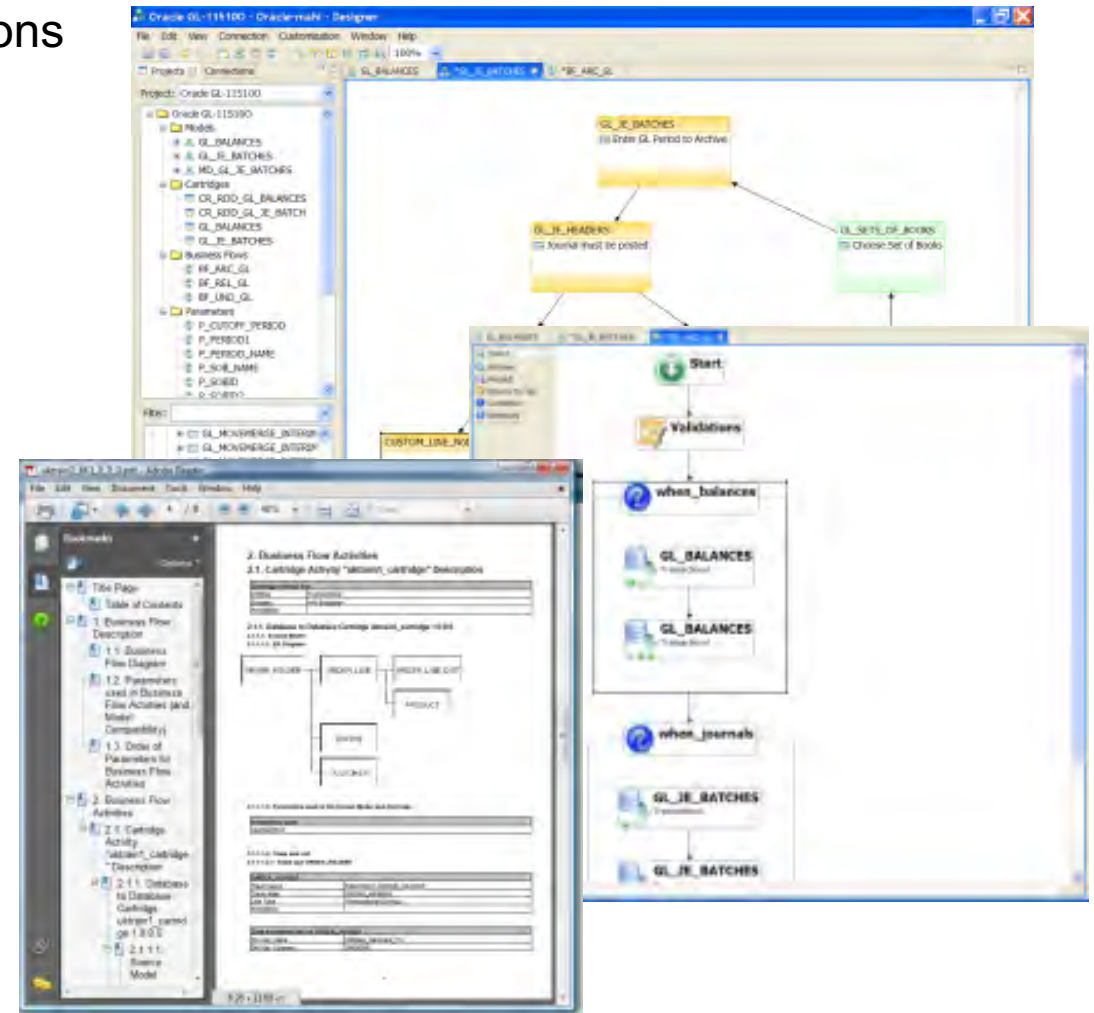
Environment Name	Description	Database	Status
Oracle	Local Oracle Std Edition	orcl.emea.hpqcorp.net@localhost	
PostgreSQL	Local PostgreSQL Database	jdbcpostgresql://localhost:5432/postgres	
SQLServer	Local SQL Server 2008 R2 Express Ed	SQLEXPRESS2008R2@localhost	

The 'SQLServer' environment is selected, indicated by a checkmark in the status column. The interface also includes a sidebar with navigation options: ENVIRONMENT, Manage, Active Environment, TASKS, USERS, and SETTINGS. The top right corner shows the user's environment as 'SQLServer'.



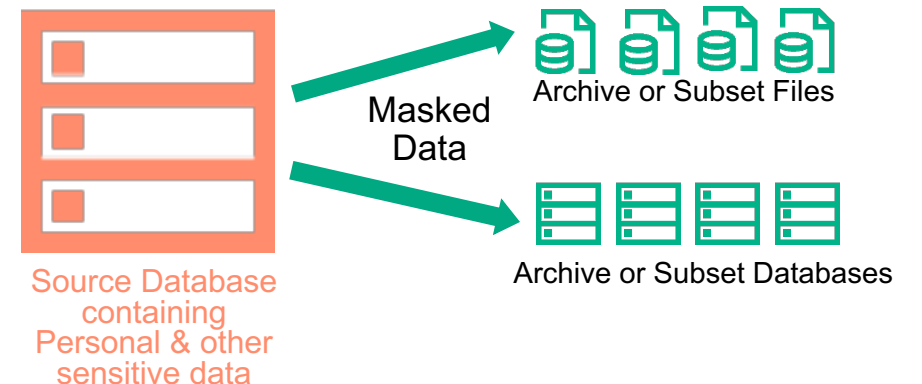
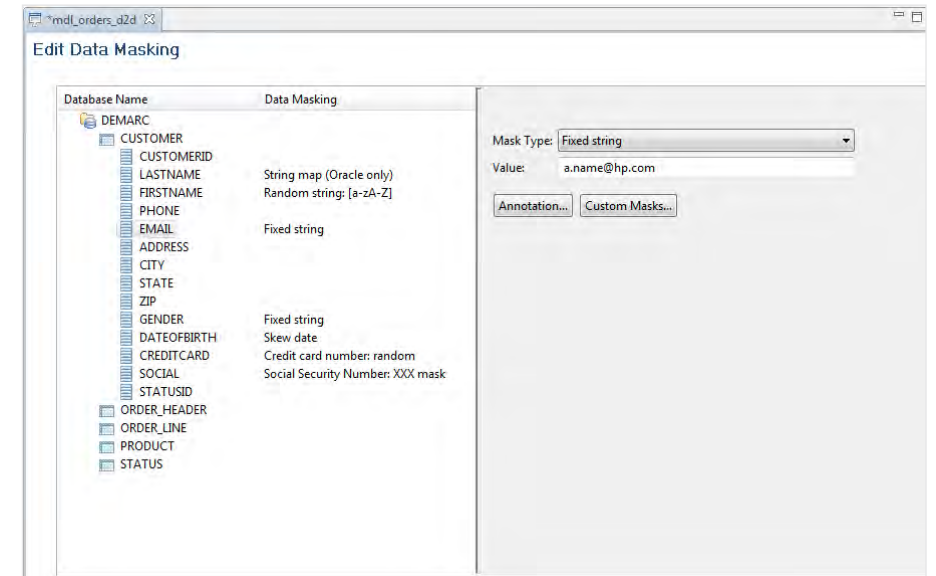
# Design an Archive flow per LOB Application

- Visual modeling environment to define business transactions
- Edit and Enhance pre-packaged modules
- Create Modules for:
  - Data Movement
  - Data Extraction
  - Data Indexing
  - Reporting
- Build bespoke modules to integrate into almost any application
- Define tasks as part of a workflow
- Produce process documentation



# Employ Data Masking

- Data Masking routines can be added to any source object at design time
- Can be applied to archiving and/or subsetting
- Data is masked using native DB routines or scripted functions.
- Supports both pre-supplied, custom developed and third-party routines
- Data is masked during the movement process before writing to the source to ensure data security

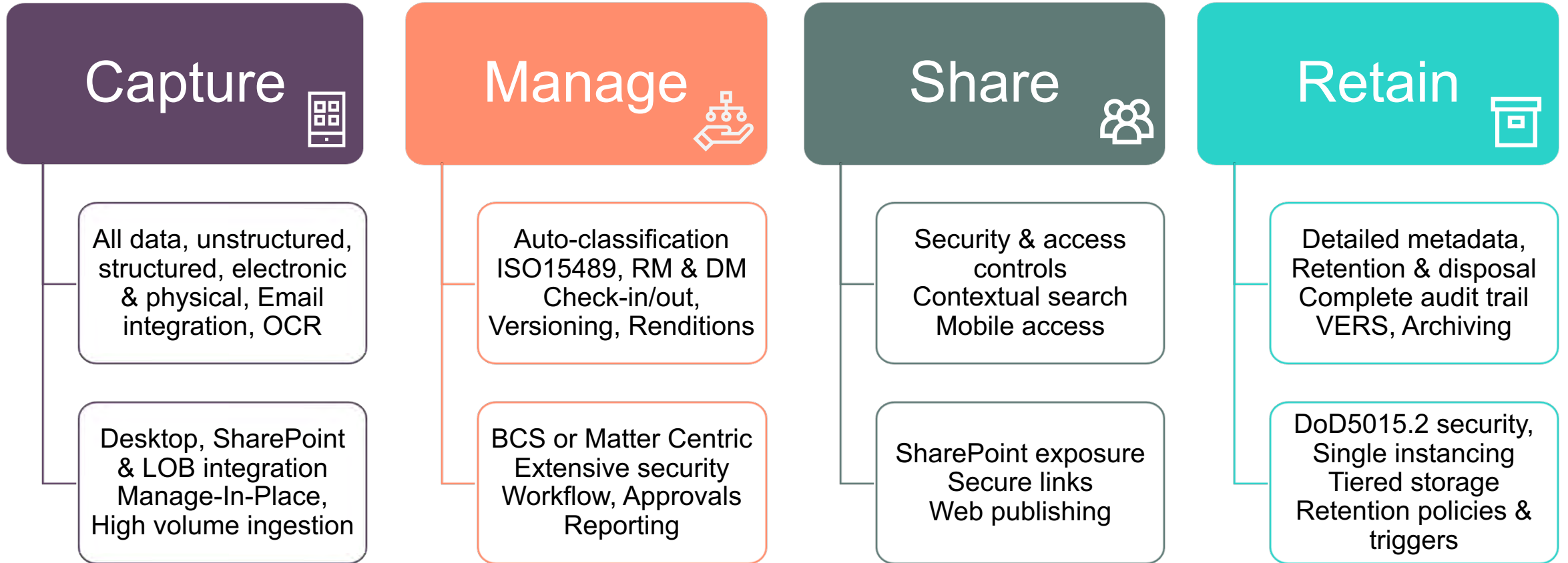




# Automatic Classification of Archived Records

# Securely and Intelligently Manage data across the lifecycle

with HPE Software's Secure Content Management Suite



# Your Classifications

## The key to Records Management

- File Plans follow your processes
- They enable Retention decisions
- This facilitates Disposition
  - Document deletion
  - Document transfer to another archive section or solution, either internal or external to the corporation



# Auto-Classification

- Train Categories from Classifications

- Using notes
- Using historic records
- or both

- Holding Classifications

- Where records get stored “to be classified”

- Target Classifications

- Linked to IDOL categories
- Confidence threshold
- Automatic folder creation

Classification - Cartography / Mapping - Digital

SAP Business Object Links | Notes | Access Controls | Active

General | Numbering | Auto-Classification

This Classification is not suitable for the auto-classification process

Use this Classification to hold records waiting to be auto-classified

Make this Classification available for selection by the auto-classification process

Minimum confidence level required

70

Create containers when importing documents

Record Type for container

File Folder

Limit container sizes

Maximum contents

250

Create containers for different date created ranges

Create new date range container for every:

Month

Create containers for each owner location

Create containers for each Additional Field value

Additional Field name

Country Code

Category Training

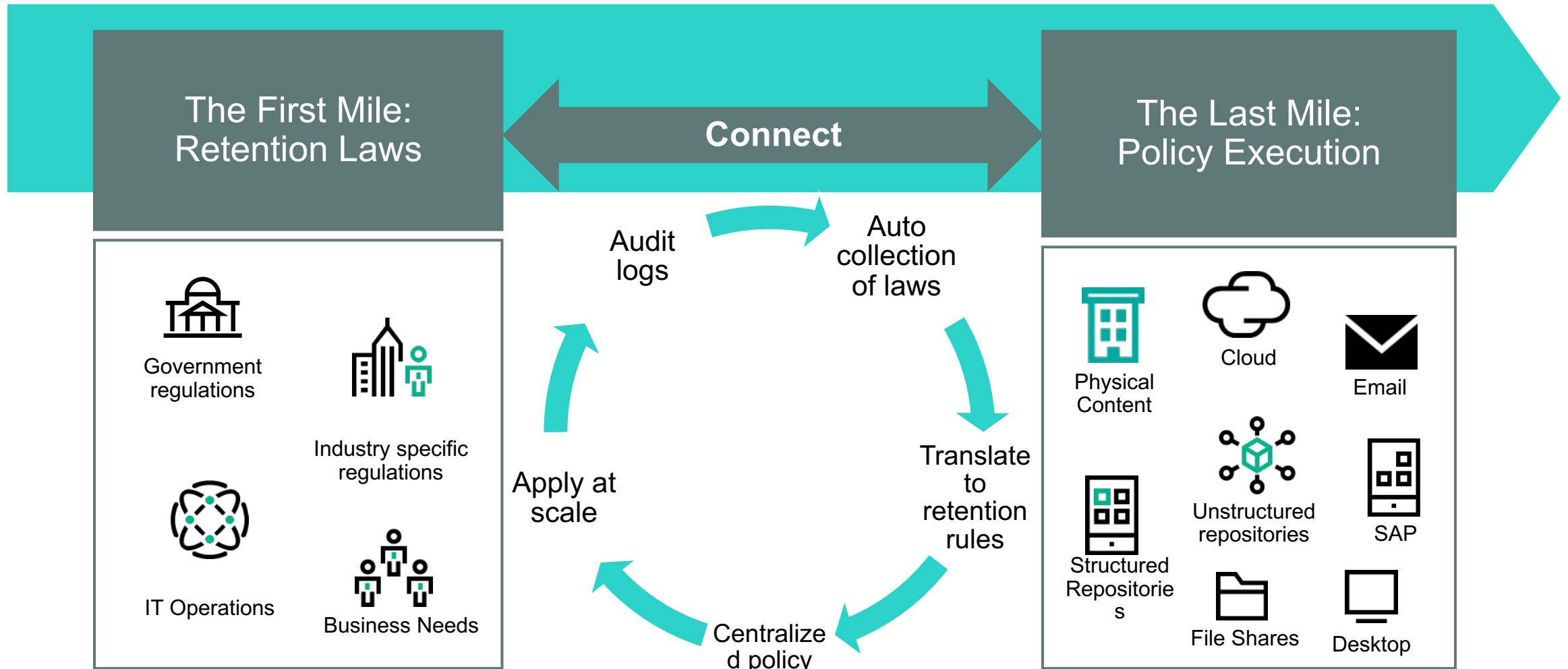
OK Cancel Help



# Automatic Updating of Classifications and Retention Decisions

# How do we better connect legal regulations to our content?

The first and last mile of retention





# Policy Center Data Model

90 Jurisdictions

35.000 Legal Citations

Retention  
Minimums

Retention  
Maximums

Production

Administrative

Limitation of  
action

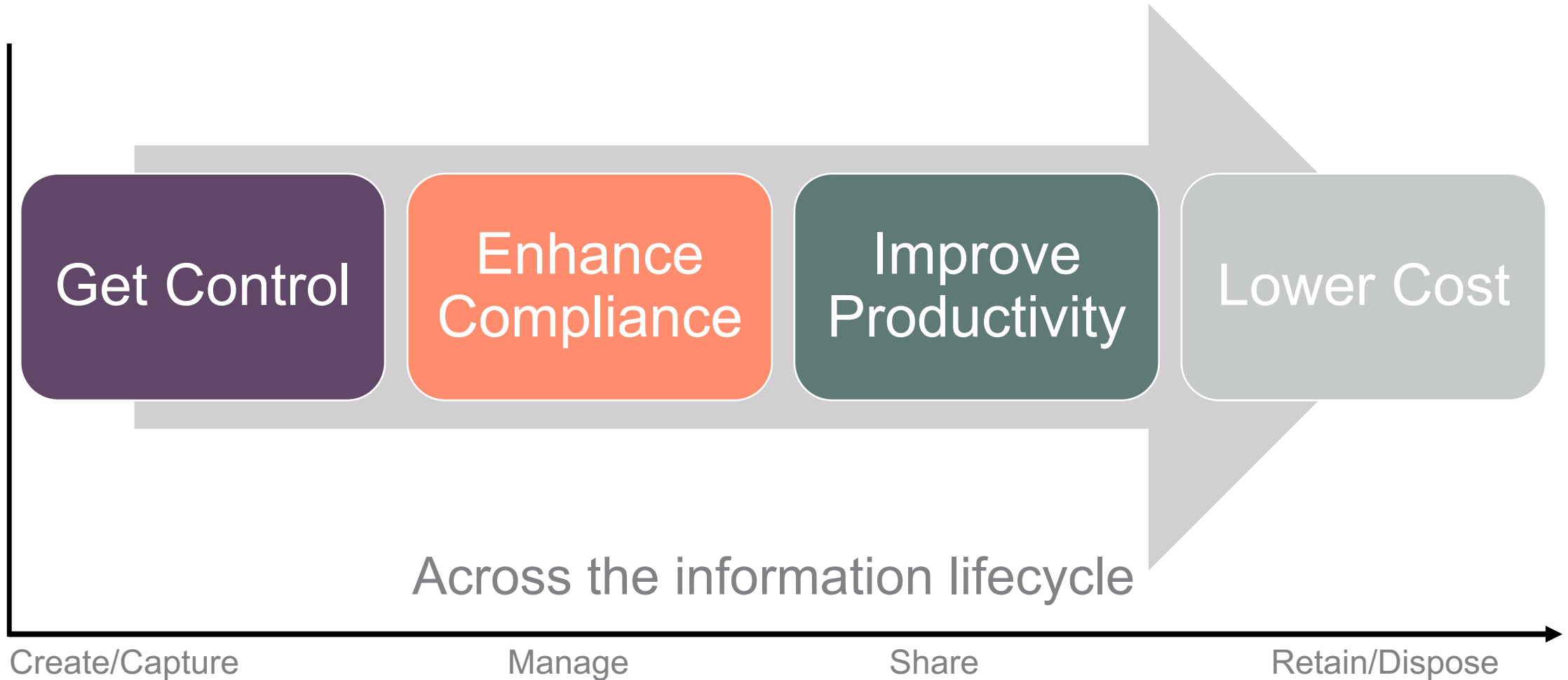
Limitation of  
assessment

Verticals and Industries



# Summary

# Automatic Appraisal and Classification allow you to





# Q&A



**Thank you!**

Thank you for your attention!

*eabn*

[bankinghistory.org](http://bankinghistory.org)

@bankinghistory

*we do history*