45th SIO annual meeting, Brussels

DIGITAL ARCHIVING STRATEGY

For Your Emails Only # 22.05.2019

Sébastien SOYEZ
Head of works < Digital Archiving Unit >
STATE ARCHIVES BELGIUM
OVERVIEW

I. THE STATE ARCHIVES BELGIUM

II. DIGITAL STRATEGY

III. E-MAILS
   a. Extern: Supervision of Federal Public Institutions
   b. Intern: Records Management Policy + Digital Archiving

IV. CHALLENGES
   a. Organisational / Technical
   b. Legal

V. SOLUTIONS
   a. Intern: State Archives
   b. Extern: Belgian Foreign Office, FelixArchief, Capstone

VI. WRAP UP
I. STATE ARCHIVES BELGIUM

- Federal Scientific Institution (specific status for archivists)
- Depending on the Federal Science Policy Ministry
- HR: approx. 240 FTE (ca. 90 archivists)
- 19 deposits
- Around 320 km paper archives (from 8th century till now)
- More than 350 TB of digital archives (25 millions digital files)
II. DIGITAL STRATEGY ARCH.BE

INSPECTION / CONTROL (digital aspects)
- Advices & Recommendations to civil servants / archive producers
- Trainings (intern/extern)
- Support the Archiving-as-a Service / Government Cloud Service
- Records Management Policy (extern)

e-TRANSFER
- Select & Collect data/documents (SIP)
- Organization & Management of Policy & Procedure
- Development & Maintenance of technical tools

MANAGEMENT + ARCHIVING
- Records Management Policy (intern)
- SAM – State Archives Managmt (Policy + Guidelines) (AIP)
- Create & manage metadata (EAD, EAC, EAG, METS, …) – (AIP)
- Manage sustainable formats & storage media’s
- Storage (intern + Long Term Platform)

ACCESS by PUBLIC
- Designing a new policy
- Create new search engines + online visualization tool
- Online Data Management (DIP)
- Online Data Storage
- (Inter)nationals collaborations
- Crowdsourcing/funding activities

RESEARCH PROJECTS:
- HECTOR = digital transition
- PROMISE = web archiving
- SODA = data archiving

Tasks out of the unit since 02/2019
III. EMAILS (management & archiving)


III. EMAILS (management & archiving)


IV. CHALLENGES

A) Organisational & technical aspects

--> **Bad uses** : no rules, mix privacy/professional
--> **Roles & Responsibilities** (RM, IT, Lawyer, HoU, User)
--> **Individual** mailbox versus **Functional** mailbox
--> **Selection / Deletion**
--> **Classification**
--> **Archiving** of messages + attachments
--> **Authenticity** (content & context)
--> **Technical interoperability** (formats)
IV. CHALLENGES

B) Legal issues

--> Privacy: access & use of recorded/archived emails

--> Authenticity + Probative value: metadata < Sender, Dates, Headers, Message, ...

MUST BE COMPLIANT

→ GDPR & Archives Law
→ Electronic Communications Law
→ Work Convention (ex. CyberControl of workers)
→ e-IDAS (EU) & Digital Act (BE) = Digital Archiving Trust Service
V. SOLUTIONS (State Archives Belgium)

1) Classification

– Temporary solution (from 2009-2016)

--> classify e-mails in the mailbox: use a global classification of the institution
V. SOLUTIONS (State Archives Belgium)

1) Classification

– Ideal solution (since 2017)
  --> **classify** emails centrally on the RM or on the Secured File System of the institution

  --> **export** emails into the file related to the task (context)

  --> if necessary, apply the **Capstone approach** (full archiving of mailboxes from HoU, Dir and GenDir level)
V. SOLUTIONS (State Archives Belgium)
V. SOLUTIONS (State Archives Belgium)

1) Classification

– TRICK

--> try to automate the classification of e-mails

I.A. / machine learning tools that can facilitate the identification of subjects, sender/receiver, ... and then classify automatically the e-mails
V. SOLUTIONS (State Archives Belgium)

1) Classification

   -- TRICK

   --> try to automated the classification of e-mails
V. SOLUTIONS (State Archives Belgium)

2) Selection

– During the current phase, final responsibility to the user
  --> delete informative e-mails (newsletters, mailing-lists, ...) &
  private e-mails (non-professional!) / personal e-mails (professional)

  --> preserve e-mails with probative value (administratively/legally)

– TRICK
  --> when a user leaves the institution, he/she is responsible to clean the
  mailbox(es)
V. SOLUTIONS (State Archives Belgium)

3) Record / Archive

– Content & descriptive elements
  --> Header informations

  --> Message content

  --> Attachment(s)
### V. SOLUTIONS (State Archives Belgium)

#### 3) Record / Archive

- Format(s)

<table>
<thead>
<tr>
<th>EMAILS</th>
<th>Retention time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&gt;10 years</td>
</tr>
<tr>
<td><strong>Message</strong></td>
<td></td>
</tr>
<tr>
<td>Attachment(s)</td>
<td>EML, MSG, XML</td>
</tr>
<tr>
<td>Text Doc</td>
<td>DOCX, ODT, TXT, PDF</td>
</tr>
<tr>
<td>Presentation</td>
<td>PPT, ODP</td>
</tr>
<tr>
<td>Table / Stylesheet</td>
<td>XLS, ODF</td>
</tr>
<tr>
<td>Database</td>
<td>MDB, ACCDB, FMJ</td>
</tr>
<tr>
<td>Image</td>
<td>JPEG, JPEG2000, PNG</td>
</tr>
<tr>
<td>Video</td>
<td>AVI</td>
</tr>
<tr>
<td>Sound</td>
<td>MP3, WAV</td>
</tr>
</tbody>
</table>
V. OTHER SOLUTIONS (Belgian Foreign Office)

- Diplomatic E-Mails (BE, Ministry of Foreign Office, 2010-2014)

  - Distinction between 2 Emails-flows
    - Normal
    - Diplomatic

  - (Almost) Automatically Archived
    - RM-System
      (Sharepoint 2013)
V. OTHER SOLUTIONS (FelixArchief Antwerpen)

• E-David (BE, Stadsarchief Antwerpen, 2007)

– MAIN COMPONENTS
  • Email = Content, Structure, Transmission data

– CONTEXT
  • Classified with the other records

– MIGRATION
  • Preserved in an « archiving-format » (XML, PDF/A?)
V. OTHER SOLUTIONS (NARA)

- CAPSTONE approach (USA-NARA, 2011 + update 2017)
VI. WRAP UP

• A sustainable and ideal solution for managing and/or archiving e-mails is related to the implementation of a **Records Management** or a **Digital Archiving System**

  – **integrated**: all information systems

  – **global**: all the documents with probative value (at least)

  – as **automated** as possible

"Capture e-mail, but use Records Management" (Jason Baron, NARA, 2010)
VI. WRAP UP

• Classifying
  – Transversal: must concern the entire institution
  – Functional: must reflect tasks & activities

• Selecting & Recording
  – Basic uniform rules (first selection)
    1) Message **formally** created/received?
    2) Implies possible action, follow-up, response (**engaging**)?
    3) Contents **substantial** & **non-ephemeral** informations?
      If answer=yes to those 3 questions → e-mail must be recorded!
  – Unique identifier
VI. WRAP UP

• Roles distribution
  – RM/Archivist
    • Global & Functional classification
    • Metadata based on legal obligations & administrative usefulness
    • File Formats identification
  – IT
    • Technical implementation (automated classification, format migration,...)
  – LAWYER (DPO)
    • Verifies if laws are proportionaly applied
    • Sustains the RM/IT
  – User
    • Sets business criteria for selection
    • Applies the rules (classification, record)
THE DEVIL IS IN THE DETAILS

E-MAILS ARCHIVING might seem simple at a first look but will take more time and effort to complete than expected
Questions?

Sébastien SOYEZ
Head of works < Digital Archiving Unit >
STATE ARCHIVES BELGIUM
sebastien.soyez@arch.be
References


- J. Lappin, Preserving e-mails - records management perspective, 2012. ([http://thinkingrecords.co.uk](http://thinkingrecords.co.uk))


- M. Pennock, Curating e-mails, Digital Curation Center, University of Bath, 2006. ([http://www.dcc.ac.uk](http://www.dcc.ac.uk))

- C. Prom, Preserving e-mail, Digital Preservation Coalition, 2011. ([http://www.dpconline.org](http://www.dpconline.org))
