Review of the SPAR Architecture

Scalable Preservation & Archiving Repository
Système de Préservation et d’Archivage Réparti (SPAR)

{BnF}
Agenda

• SPAR Overview
• SPAR in real
• Next steps
How to manage digital archives

• Build a central repository to diminish the diversity (media, formats, departments …)

• To take inspiration from good practices and standards

⇒ Key requirements:
  – OAIS compliance
  – modularity and distributivity
  – abstraction
  – use of well known formats and standards
  – use of open-source technical building blocks
Digital archiving at BnF

- Preservation digitization
- Audiovisual
- WEB Archiving

SPAR - Infrastructure
- Ingest
- Administration
- Storage
- Rights management
- Preservation planning
- Data management
- Access
- Storage Abstraction Service (SAS)

SPAR - Realization

SPAR

Production applications

Dissemination applications

2010/07/05-PASIG
Initial deployment of SPAR in the overall digital system

In this first phase, the archive packages in SPAR won’t be disseminated in an intense way => focus on ingesting and auditing
Agenda

• SPAR Overview
• SPAR in real
• Next steps
Overall ingestion workflow
(preliminary QA)
Overall ingestion workflow
(preparation)
Overall ingestion workflow (archiving & dissemination)
Demonstration with one package

<table>
<thead>
<tr>
<th>Information on the package</th>
<th>Master</th>
<th>Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of files</td>
<td>36</td>
<td>1092</td>
</tr>
<tr>
<td>Weight</td>
<td>4.77 Mb</td>
<td>20 Mb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time for the package 5841888</th>
<th>Time</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>QA</td>
<td>16:45:08</td>
<td></td>
</tr>
<tr>
<td>Preparation</td>
<td>17:09:35</td>
<td>00:24:27</td>
</tr>
<tr>
<td>Begin dissemination</td>
<td>17:15:19</td>
<td>00:05:44</td>
</tr>
<tr>
<td>End dissemination</td>
<td>17:40:06</td>
<td>00:24:47</td>
</tr>
<tr>
<td>Begin archiving</td>
<td>17:44:46</td>
<td>00:04:40</td>
</tr>
<tr>
<td>End archiving</td>
<td>17:51:00</td>
<td>00:06:14</td>
</tr>
<tr>
<td>Total</td>
<td>01:05:52</td>
<td></td>
</tr>
</tbody>
</table>
Agenda

• SPAR Overview
• SPAR in real

• Next steps
Overview of an ingest

Pre Ingest

Ingest

Preservation Planning

Administration

Rights management

Data management

Storage

Access

SAS

Pre Ingest

Ingest

Storage

Access
Production so far

Ingest FIL_NUM_CONS_A

- Packages
- Sets
- Groups
- Monographs
- Periodicals
- MPV
- Images

2010/07/05-PASIG
Iterative development

- May 2010: SPAR in production for the digitization channel
- September 2010: third-party storage and better administration
- End of 2010: audiovisual material
- 2011: Web archiving, third-party archiving, …
Thanks for your attention

Questions?

Thomas Ledoux
thomas.ledoux_AT_bnf.fr