Eric Dussert
Printed Books Digitization Project Manager
Direction des Collections/Affaires scientifiques et techniques
Bibliothèque nationale de France
Goals and Rules

Toward the digitizing of the most remarkable and peculiar collections…
• An Historical and Multimedia Offer…
• A concerted digitization with national and international partners…
• Circulating digital collections : a new and necessary task

• Main Directions :
  – Highlighting Master Pieces and Treasures from our collections
  – Preserving fragile and unmanageable documents

Prospect and matters : new challenges

• Technology
• Quality
• Dissemination, Communication…
• Management…

A work in progress

29 novembre 2012
Gallica by the numbers

Vous trouverez sur Gallica
1 912 411 documents
379 574 Livres,
50 271 Cartes,
28 446 Manuscrits,
531 669 Images,
910 322 Presse et Revues,
10 022 Partitions,
2 107 Paroles et musiques,
dont
89 408 de E-distributeurs,
129 809 d’autres bibliothèques

Port de Marseille

BnF
29 novembre 2012
Since 1992, we digitize:

- Printed books (1992)
- Microforms
- Newspapers (2000)
- Rare books
- Manuscripts
- Drawings and photographs
- Audiovisual collections
- Objects (performing arts, etc.)
- 78 tours records
- Animated images

We offer:
- files on Gallica website
- files on Gallica IntraMuros (intranet)
- « nomad » books (e-pub)
Materials
Programs

**First « Mass contract » (2008-2010) : 330.000 documents**

**Selected Topics for binded books collections**
French History (enlarged to Colonial History)
French Litterature (novel, poetry, theater…)
Law and Jurisprudence
Sciences (medecine and natural sciences)

**Limited collections of thematic works based on reference bibliographies :**
Sinica (France-China)
European Philosophy and Social Sciences
Periodicals of Art and Litterature
French Translations of European Master Works

**Second « Mass contract » is running : 2011-2014**

« corpus » about topics or regional studies :
1914-1918
Japan
Poland

BnF
29 novembre 2012
External Libraries
30% digitized by BnF Contractants
provided by French Libraries

Orphan and under copyrights works
(test with 10,000 books during 2013)

Projects in cooperation with Referent
Institutions like
Law (Cujas Library)
Art History (INHA)
Sciences (Biodiversity Heritage Library ; BIUM)
Youth Litterature (associated specialists)

Extended Programs
Ex. : French speaking Press (Haïti, Mali, Tunisie…)

Sponsorship : « Adopt a book »

Private Partnerships

BnF
29 novembre 2012
A library…
- based on a catalog (legal deposit)
- bibliotheconomical uses (choosing original material, preservation, etc.)
- various materials
- different kind of readers

« A New kind of library » means a constant effort
- evolution of the contents and appearance of new objects
- evolution of the flows/contracts/programs
- evolution of tools (host, web display, zoom, etc.)
New General superiority of Law

Music Rights, author Rights of writers, creators dead during war, etc.
Complex cases of pictures (ex. review Jazz)

Technology

- 400 dpi, colours, etc.
- Quality management
- Upgrading classical texts
Conditions of Digitizing

Scientific choices
BnF Stacks/International Projects/Private founding

Metadata

Physical Characteristics
What technic can be used? What digitizing frequency? OCR?
Specific and precious documents

Others Programs of digitizing

Law
Quality Progress

- Variety of technical sectors, operations, specific contracts (with their own characteristics)
- Catalog corrections (metadata)
- Example: The « Recueils » case (compendium)
  Félix Fénéon, Les Impressionnistes (Paris, La Vogue, 1886)
  (http://gallica.intramuros.bnf.fr/ark:/12148/bpt6k850398w)
- What OCR?
- What OCR for what mobile files?
- New mobile (nomad) files: new rules for a new component
  How to select books who are to become products for new market?
  Are Dumas or Zola good candidates?
  What kinds of books are eligible for this purpose?
- A commercial interest is appearing in the libraries, even in patrimonial ones
Numerous copies
Numerous copies

- Political will to promote historical archives
  - A large digital library for French and European People and (and french-speaking populations)
  - A wish of Independance

- Influence of different institutions (local ones too)
- Densification/Massification of contracts
- Technical requirement, standards evolution
- Or an « expectation horizon » of readers
- Last but not least, the internal logic of fine « corpus »

**legitimate multiplied copies**
**but double entry is often the consequence of**

Incompatibility between rhythms of different programs or contracts;

Even if they are sized and built on different logics

A digital library needs to own full data (images, OCR, bibliographical data and summary), not the only metadatas

**How to manage gaps where as some documents are already present in multiple copies ?**

Ex. Voltaire’s reference edition (Garnier frères, 1877-1885) : 4 unities missing
New Ways of Reading?

- Expectation horizon
  - « Audimat »: digitisation of frequently read books
  - « Audit Grands Texts »: digitization of French classical authors with OCR HQ
  - « Fresh offer » and reference texts

- Studies about Webusers

- « Good Neighbour » Law, persistence of traditional uses (serendipity)

- Cursory reading
  (many opened files, no reading)
Conclusion

Twenty years after the start of digitizing…

- A new principle: « The more you digitize, the more you create gap »
- Avoiding multipled copies: a new deal?
- Raising general quality level
- Mass programs lead to a more complex selection process
Conclusion

- Digitizing, between wish and frustration...
- A great satisfaction: distant readers are more and more satisfied
- Researchers have never been frightened by abundance
Thank you for your attention