

Collaboration between Wikimédia France and the Bibliothèque nationale de France

Jean-Frédéric Berthelot & Sébastien Beyou

1er décembre 2010



Rights

• Public domain material : no copyright claims

Rights

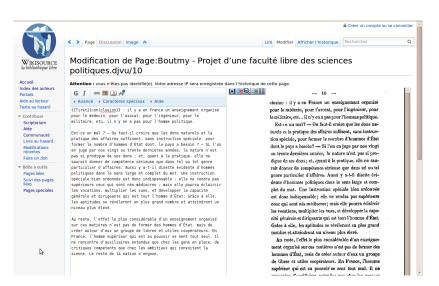
- Public domain material : no copyright claims
- 1978 law : public data

Rights

- Public domain material : no copyright claims
- 1978 law: public data
 - « La réutilisation non commerciale de ces contenus est libre et gratuite dans le respect de la législation en vigueur et notamment du maintien de la mention de source. »
 - « La réutilisation commerciale de ces contenus est payante et fait l'objet d'une licence. Est entendue par réutilisation commerciale la revente de contenus sous forme de produits élaborés ou de fourniture de service. »







Partnership

- gift of 1400 books to Wikisource with images and OCR
- the Wikisourcians have to proofread the texts

Partnership

- gift of 1400 books to Wikisource with images and OCR
- the Wikisourcians have to proofread the texts
- another side the BnF allows the Wikimedia community to use its autority data

Milestones

- October 2007 Conference of Wikimédia France targeting the academic world
- March 2008 End 2009 Discussions with the BnF
- January 2010 March 2010 Finalization, communication, beginning of the technical work
- July 2010 Release of the books on Wikisource
- October 2010 First statistics

Difficulties

Difficulties

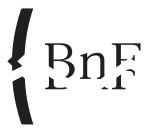
Knowledge spreading

Income

Difficulties

Knowledge spreading

Income



Community

- an informal group of volunteers with many cultures and opinions
- a partnership has to respect the existing communities
- a community lives before and after a partnership
- the community has a common background and culture

Technique

- learned the specific format for the OCR (XML ALTO developed by the Library of Congress)
- learned the specific format of books of Wikisource (DjVu)
- transformed the OCR into pure text, and then to DjVu
- extracted the images in the books

Technique

- intensive work of 3 engineers during 2 months (part-time because volonteers)
- a total of 6 months
- easier because all 3 were at the same place (Rennes, France)

Statistics

Visualize the evolution of the work done on the books

Statistics

- 169 books have been begun on the 1416 books
- 114 contributors
- 10,000 contributions by humans, 19,500 if we count bots
- an average of 376 modifications of characters by page

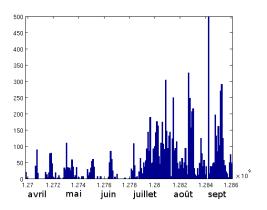


Figure: Activity of the contributors

Statistics

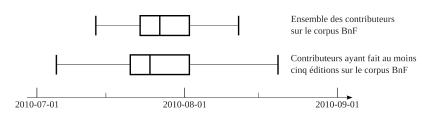


Figure: Temporal activity of the contributors

Outcomes

- a form of acknowledgement by the BnF
- an experiment with the French Wikimedia communities
- the first big partnership with the French GLAMs; one of the first in the world

Contacts

```
Discussions (WMFr)
            Rémi Mathis
            Alexandre Moatti
            Valérie Chansigaud
Discussions (BnF)
            Dominique Stutzmann
  Technique
            Sébastien Beyou
            Jean-Frédéric Berthelot
            Vincent Juhel
Advisor Wikisource
```

Nicolas Vigneron