Archiving 125 Years of "Het Belang van Limburg"

Filip Beckers

Concentra Media
Herckenrodesingel 10 – B3500 Hasselt - Belgium
fbeckers@concentra.be

Abstract

Concentra Media and its newspaper Het Belang van Limburg celebrated their 125th anniversary in 2004. As Het Belang van Limburg was never archived on microfilm, the company decided to digitalize the editions of 125 years that were available in 620 bound books and counted for 520,000 newspaper-pages going back to 1879, to avoid a further deterioration of the unique copies while using them for searches.

Concentra Media found a technical partner in its former 100% subsidiary Host-It [now Cegeka], together with a Dutch partner X-Cago.

1 Scope

- Preserve 125 years of publications of Het Belang van Limburg
  620 books
  520,000 pages

2 Objectives

2.1 Primary objectives

- Reproduce original size and colour
- Metadata: date – edition - page
- Searchable
  OCR
  Indexing

2.2 Secondary objectives

- Develop search-interface
- Develop consultation interface

3 Technical directives

- Use or follow industry-standards
- Open system
  Able to implement new technologies
  Complement content with recent newspapers (digital newspaper platform)

4 Live Demo

- Interface
- Search
- Search-results
- Quality OCR

5 Challenges

5.1 General challenges

- Inventory & search missing editions
- Commercialisation
- Limited target group
- Copyright- & royalty-issues
5.2 Technical challenges
- Security (PDF)
- Performance
- Powerful interfaces

6 Partnerships
- Academic partners
- Local government
- Cultural heritage
- Business-to-business

7 Applications
- Journalists
- Projects
- FrontPage web shop
- Digital newspaper

References

http://1934.hbvlonline.be—Website with live demo with content of 1934
http://www.hbvl.be/service/webshop.asp—Web shop to order historic FrontPages