Spain’s DILVE experience: A BIP repository as data adapter for publishers

32nd International Supply Chain Seminar
Frankfurt, 5 October 2010

Jesús Peraita - Neturity SL

Promoted by
What is DILVE?

- DILVE (www.dilve.es): the Spanish Books-in-print repository
- Managed by the Spanish Publisher's Association (Federación de Gremios de Editores de España - FGEE)
- *Free of charge*: restricted to Spanish publishers, but data *available to international professionals* of the book chain.
- Bibliographical data directly *managed by publishers*
- More than: 650 publishers, 160 distributors, 600 bookstores
- Over 270,000 records (Oct 2010)

**DILVE acts as data adaptor and format transformer for publishers**
DILVE as data adaptor

• Besides *data extraction managed by end-users*, DILVE allows publishers to **push data** to particular environments requiring specific data formats and transmission means:
  – ENCLAVE (Spanish National Library): xml-MARC21 / FTP
  – Libranda (eDistributor): ONIX 3.0 / FTPES
  – SINLI (Spanish Commercial interchange format): plain text / e-Mail

• DILVE carries out specific *business consistency checks* for each environment before sending data

• All the procedures are *transparent for the publisher* and require one single upload process to DILVE and one push of a **Send** button
enclave editores-bne (Dec 2008 - Nov 2010)

Integration of copyrighted material in digital public libraries
http://bdh.bne.es/bnesearch/

Promoted by:
• Biblioteca Nacional de España (BNE)
• Federación de Gremios de Editores de España (FGEE)

Partially funded by:
• Ministerio de Industria, Comercio y Turismo

Bibliographic data provided by DILVE

Technical co-ordination by Neturity SL
Project objectives

• Include books under rights in the Biblioteca Digital Hispánica (BDH) (Primary objective !)

+  

• Detect generic problems & possible solutions
• Define models & infrastructure for other public libraries
• Stimulate publishers to define their digital plans
• Foster the creation of new agents in the e-book chain
• Promote media awareness

• Break the chicken & egg situation for digital books:
  – "There's no offer because there's no demand because there's no offer...", but a small was enough to start the ball rolling
Data requirements

• BDH already has the *complete basic bibliographic record* of the printed version

• *"Enriched" data is provided by the publisher*

• Required data for Enclave:
  – **Main description** - plain text format (ONIX L33 01)
  – **TOC** - plain text format (ONIX L33 04)
  – **Cover image** - jpg, png (ONIX L38 04)
  – **Excerpt** for on-site browsing - pdf (ONIX L33 23)
  – One **URL** per book, to in e-distributor / e-retailer environment
  – **Full text** - pdf (ONIX L33 34) *Internal BNE use only for indexing in Autonomy*

>>> All data transfers to BNE come *from DILVE* <<<
User navigation at BDH (1)

1. User carries out a search in BDH. Full text search is possible.

2. A listing of matching records is displayed. Enclave records are identified.

3. User selects an Enclave record for detailed view. Up to 20% can be browsed on-site. Full document can be consulted/bought at e-distributor’s site.

On-site browsing (20%)  External access to e-Distributor
## Basic bibliographic data: BNE

<table>
<thead>
<tr>
<th>Título</th>
<th>La ciudad y los derechos humanos - : una modesta proposición sobre derechos humanos y práctica urbanística</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autor</td>
<td>Caz Enjuto, Rosario del Gigosos Pérez, Pablo Saravia Madrigal, Manuel</td>
</tr>
<tr>
<td>Datos de edición</td>
<td>Madrid Talasa</td>
</tr>
<tr>
<td>Fecha</td>
<td>2002</td>
</tr>
<tr>
<td>Tipo de Documento</td>
<td>2152892</td>
</tr>
<tr>
<td>Materia</td>
<td>Urbanismo - Aspectos sociales Derechos sociales y económicos</td>
</tr>
<tr>
<td>Descripción física</td>
<td>133 p.</td>
</tr>
<tr>
<td>PID</td>
<td>711.4:316.34342.72/.73</td>
</tr>
<tr>
<td>CDU</td>
<td>84-88119-90-9</td>
</tr>
</tbody>
</table>

### Book description: Publisher

Se impone recuperar una visión más global del urbanismo sostenible; recuperar el urbanismo social, ahora que la desigualdad de nuestro mundo global se agudiza en las ciudades. Y el horizonte hacia el que dirigir la renovación del pensamiento y la acción urbanística no puede ser otro que el de los derechos humanos; el de las garantías de dignidad existir armonía ciudadana, mejoras medioambientales ni una adecuada adaptación a las nuevas tecnologías sin ese mayor respeto (menos complaciente) a los derechos humanos de todas las personas: de las que están, de las que llegan a la ciudad.

### Book TOC: Publisher

1. El lugar del derecho
2. Entender derechos y obligaciones
Data flows (1)

Publisher

DILVE

BNE

Publisher data
(Enriched data)

Basic
TOC
Description
Cover image

e-Distributor URL

Partial Content
(Extract)
File (PDF)

Full-text file
for indexing
(PDF)

e-Book files

If the e-Distributor has generated these files they will already be resident in its systems

e-Distributor / e-Retailer

Publisher

If the e-Distributor has generated these files they will already be resident in its systems
Data flows (2)

All users have access to all of this data in DILVE

Basic metadata + Enriched data + 20% extract PDF (Free formats)

Publisher

BNE

Publisher data (Enriched data)

Basic

TOC

Description

Cover image

e-Distributor URL

Partial Content (Extract) File (PDF)

Full-text file for indexing (PDF)

e-Distributor / e-Retailer
Each publisher triggers in DILVE the upload process to ENCLAVE

- A MARC-XML metadata record is generated
- Files are renamed as needed
- The package is zipped and FTP'd to BNE
Enclave’s Current status

• Started December 2008 - Ends November 2010

• Officially presented in LIBER (Madrid) in October 2009

• Open to the public since November 2009

• Two calls for participation: July 2009 & January 2010

• 142 publishers and over 3200 books selected for inclusion
Enclave’s Conclusions

• As expected, Enclave has made a big “push” to the digital publishing chain in Spain

• Great public success for BDH

• Overall data model very convenient for publishers
  – However: Full text indexing via PDF file (BNE requirement) problematic for books in ePub format

• Record retransmissions due to problems in metadata down from 14% in first call to 6% in second call

• URLs for e-distributors are very volatile and an alternative updating mechanism has been set up (Excel files)
Further Information

General Documentation
www.dilve.es

Support / Information
dilve@dilve.es

Enclave
- Website: http://bdh.bne.es/bnesearch/
- Docs: www.bne.es/es/enclave - www.dilve.es

Thank you!