CBS and WorldCat: Strategic Directions

Hellenic Academic Libraries Link
30th September 2007

Janifer Gatenby, Strategic Research
Agenda

• CBS Currently
  – Role of CBS
  – Performance and Reliability
  – Implementations; partners; implementation projects
  – Data import , export, formats
  – Character encodings, UNICODE
  – Standards
  – Models

• Integration with WorldCat and VDX

• Future Strategy
  – Role of partners – Job management
  – Grid services; WorldCat API – Web 2.0

• Summary
CBS software

Centraal Bibliotheek System

- Designed for union catalogues
  - Group level; importance of defining data standards, scripts

- Generic

- Evolving
  - GUM project (2000-2002) conversion to UNIX (CBS4)
    - Recently re-engineered
  - UNICODE project (2003-2004)
Robust

- All new versions implemented on Dutch union catalogue first
- Dutch union catalogue and Unity UK – constantly monitored
  - Peak 40 transactions per second
- NLA: only 2 instances of being down in 2 years; never slow
**Speed**

- av. <0.1 - 0.4 secs
- Index re-build: 2 hours for 12 million records
- Highly tuned batch load
  - 10,000 records per hour (full matching and merging)
  - 30,000 records per hour (identifier matching and merging)
  - 50,000 records per hour (no matching and merging)
- Real time export: 80 records a second
# CBS union catalogues

<table>
<thead>
<tr>
<th>Country</th>
<th>Organisation</th>
<th>Name</th>
<th>Records in M</th>
<th>Copies in M</th>
<th>Sim. Users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OCLC PICA</td>
<td>GGC-NCC</td>
<td>14</td>
<td>25</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>GBV</td>
<td>GVK</td>
<td>25</td>
<td>50</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>DDB</td>
<td>Iltis</td>
<td>9</td>
<td>16</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>HeBIS</td>
<td>HeBIS</td>
<td>11</td>
<td>16</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>BSZ</td>
<td>SWB</td>
<td>13</td>
<td>30</td>
<td>750</td>
</tr>
<tr>
<td></td>
<td>Abes</td>
<td>SUDOC</td>
<td>8</td>
<td>19</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>NLA</td>
<td>NBU</td>
<td>17</td>
<td>40</td>
<td>450</td>
</tr>
<tr>
<td></td>
<td>OCLC PICA</td>
<td>UnityUK</td>
<td>5</td>
<td>40</td>
<td>?</td>
</tr>
</tbody>
</table>
Flexibility

• Multiple format
  – Cataloguing format – MARC21, UNIMARC, PICAMARC, MAB

• FCV = File Format Converter
  – Input conversions
  – Output conversions
  – Presentations
    • Labelled formatting; ISBD; Brief presentations
  – Tool in indexing
  – Centre of UNIMARC for OCLC worldwide
# CBS – Integration into Local Environment

<table>
<thead>
<tr>
<th></th>
<th>🇳🇱</th>
<th>🇩🇪</th>
<th>🇫🇷</th>
<th>🇦🇺</th>
<th>🇬🇧</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cataloguing database</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
</tr>
<tr>
<td>Public version</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>Teratext Lucene</td>
<td>ZPortal CBS/PSI</td>
</tr>
<tr>
<td>Authentication</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>CBS / PSI</td>
<td>Via LDAP</td>
<td>CBS / PSI</td>
</tr>
<tr>
<td>Resource Delivery</td>
<td>CBS</td>
<td>CBS</td>
<td>CBS</td>
<td>VDX</td>
<td>VDX</td>
</tr>
<tr>
<td>Extras</td>
<td><strong>DAI</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
CBS – ILS interoperation

- GGC / NCC, GBV, DNB, HeBIS – LBS
- BSZ – SunRise, Dynix, Aleph ++
- France, Australia – Innovative, Dynix, Aleph, Infor ++
- UK – Dynix, Talis, Infor, ++

Any system supporting the standard interfaces – significant potential for open source ILS systems
Online Contribution to CBS

- Windows client WinIBW
- Web client
  - Holdings
  - Record capture
  - Simple new record input
- Z39.50 / UCP
  - From Aleph systems in Germany
- SRU update
  - Real time update from an alternative interface
  - Q1 2006
## Adding a Holding

### National Library of Australia

**Title:** Η εποχή: ανεξοδές / Stelios Dēmas  
**Author:** Stelios Dēmas  
**Year:** 2003  
**ISBN:** 0-9751607-7-X

### New Holding

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call no.</td>
<td></td>
</tr>
<tr>
<td>Local system number</td>
<td></td>
</tr>
<tr>
<td>Serial volume no.</td>
<td></td>
</tr>
<tr>
<td>Serial dates</td>
<td></td>
</tr>
<tr>
<td>Completeness note</td>
<td></td>
</tr>
<tr>
<td>Referral note</td>
<td></td>
</tr>
<tr>
<td>URI</td>
<td></td>
</tr>
</tbody>
</table>

For this record: Save | Close
Adding a record
Batch Contribution to CBS

- FTP
  - Files sent to FTP server associated with CBS (PUSH)

- OAI PMH
  - CBS OAI server enacts transactions to retrieve pre-defined sets of data from another database (PULL)

- Email / SMTP
External Output

Catalogue ↚ Mirror OPAC

Documents ↚ Repository

Catalogue ↛ Union catalogue

Push

OAI

Pull

SRU

Update

Push

SRU: Search / Retrieve via URL

SRU Version 1.1 13th February 2004
NCC

WorldCat

GGC

Active participation and commitment

PUSH

SRU UPDATE M21

WorldCat identifiers, FRBR work id

SRU UPDATE
CBS – Enquiry and Resource Delivery

- OpenURL

Active participation and commitment
Dear Colleagues,

Some good news - the NLA has successfully tested the Thai script in CBS, including full round trip testing from input to output, indexing, searching and export to the Teratext database, and are ready to go into production with it. This is the first of many new scripts that they are considering introducing throughout this year. From a national service viewpoint, this is a most significant improvement over the previous service where the main database was Roman script with a separate database for CJK only. Bemal is quite delighted and those of you who have worked with her will know how high her standards are.

Janifer
Dēmas, Stelios.

Δήμας, Στέλιος.

Ἡ ζοή : anekdota / Stelios Dēmas.
Η ζωή ανεκδοτα / Στελιος Δήμας.

Also Titled: Zoē mou
Zωή μου

Published: East kew, Vic. : Tsonis Publishing, c.

Description: 77 p. ; 21 cm.

ISBN: 097516077X

Dewey Number: 889.734

Subjects: Greek wit and humor. Modern.

Language: Greek, Modern (1453-)

Libraries that have this item:
National Library of Australia (ANL) 3510537 N 889.734 D372zo
State Library of Victoria (VSL) 1263698 LD held
UNICODE - Transliteration

The cataloguer has typed:

700 #1$s6$a01$sПушкин$bАлександр Сергеевич$f1799-1837
700 #1$s6$a01$s7$ba

The $7 in the second tag 700 is an instruction for generating the transliteration; the system should return:

700 #1$s6$a01$sПушкин$bАлександр Сергеевич$f1799-1837
700 #1$s6$a01$s7$ba$sPuškin$bAleksandr Sergeevič$f1799-1837
CBS and WorldCat

- Increase exposure
  - Presenting library resources
- Access to enriched data
  - Linked manifestations, works, identities
  - Mined data – audience level, holdings count (rarity)
  - Lib 2.0 data – facets, social network data
- Analysis services
  - Collection analysis
  - Locate related collections around the world
- Delivery services
  - Library Policy registry, Resolver & OPAC registry
  - International resources
Open WorldCat

Limitations on loading imposed by harvesters – Google only interested in titles with many holdings

Exposure
Worldcat.org statistics

- 14 million hits a month
  - 88 million records, 30 million articles
- 1 million hits a month from Google Books
- 800,000 links to OPAC a month (up 360% in one year)
CBS and WorldCat

Real time update

Local system A
Local system B
Local system C

FTP
SMTP
OAI

CBS Union catalogue

WorldCat

Google, Yahoo
Worldcat.org

WinIBW
browser

Real time update

OCLC PICA
Maria Callas: sacred monster
by Stelios Galatopoulos

Type: Book
Language: English
ISBN: 0684850858 9780684850859
OCLC: 40967710

Related Subjects: Callas, Maria, | Sopranos (Singers) -- Biography.
Editions: 3 Editions
Citations: Cite this Item | Export to EndNote | Export to RefWorks
# FRBR – Alternative editions

<table>
<thead>
<tr>
<th>Libraries</th>
<th>Details</th>
<th>Subjects</th>
<th>Editions</th>
<th>Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Other Editions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Displaying Editions 1 - 3 out of 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. <a href="#">Maria Callas: sacred monster</a> by Stelios Galatopoulos</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Language: English  Type: Book</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. <a href="#">Maria Callas: sacred monster</a> by Stelios Galatopoulos</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Language: English  Type: Book</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. <a href="#">Maria Callas</a> by Stelios Galatopoulos</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Language: English  Type: Book</td>
<td></td>
</tr>
</tbody>
</table>

Displaying Editions 1 - 3 out of 3
FRBR in CBS

181. The unknown Callas: the Greek years
Petsalis-Diomidis, Nicholas / Amadeus Press / 2001

182. Maria Callas: diaries of a friendship
Sutherland, Robert / Constable / 1999

183. Maria Callas à l’Opéra de Paris
Chœurs et orchestre de l’Opéra / 1987

184. Callas: portrait of a Prima Donna
Jellinek, George / Dover / 1986

185. Donizetti, Lucia di Lammermoor, Calla
Donizetti, Gaetano / s.n. / 1986

186. Maria Callas: Rarities - interviews - rel
Callas, Maria / s.n. / 19XX

187. Maria Callas: een leven als een Griekse tragedie: biografie
Zoggel, Karl H. van / Tirion / 2003

results 1

identical titles of Callas | 2 hits

1. Callas: portrait of a Prima Donna
Jellinek, George / Dover / 1986

2. Callas: portrait of a Prima Donna
Jellinek, George / Ziff-Davis / 1960

534667 [2]
Faceted Searching Sorting

Find in a Library

Search results for 'abby williams hill'

Results 1-10 of about 56 (0.24 seconds)

1. Abby Williams Hill and the lure of the West
   by Ronald Fields; Abby Williams Hill; Washington State Historical Society. Museum.
   Language: English Type: Book
   Publisher: [Tacoma, WA]: Washington State Historical Society, 1939.

2. Abby Williams Hill: progressive and painter
   by Mary Sudar
   Language: English Type: Book : Thesis/dissertation/manuscript : Archival Material
   Publisher: [n.p.?, 1983].

   by Abby Williams Hill; University of Puget Sound.
   Language: English Type: Book
   Publisher: [Tacoma, Wash : University of Puget Sound, 1979].

4. The Abby Hill collection
   by Abby Williams Hill; University of Puget Sound.
   Language: English Type: Internet Resource : 2-D image : Computer File
   Publisher: Tacoma, WA : The University of Puget Sound.

5. Niguel Canyon: wild mustard
   In: independent spirits.
   Language: No Linguistic Content Type: Visual Material : Slide
   Publisher: 2000.

# Audience Level and Rareness

<table>
<thead>
<tr>
<th>Workset ID</th>
<th>Title</th>
<th>Author</th>
<th>Number of Manifestations</th>
<th>Holdings In WorldCat</th>
<th>Holdings used for Audience Level</th>
<th>Audience Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>11574134</td>
<td>Basic research methods for librarians /</td>
<td>Powell, Ronald R.</td>
<td>6</td>
<td>1505</td>
<td>1302</td>
<td>0.7</td>
</tr>
</tbody>
</table>

**Relative Audience Level:**

- Scholar
  - 43% of all works in WorldCat have a higher Audience Level.

### Hide Manifestations

<table>
<thead>
<tr>
<th>OCLC #</th>
<th>Language</th>
<th>Publication Date</th>
<th>Number of Holdings</th>
<th>Holdings used for Audience Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>11574134</td>
<td>English</td>
<td>1985</td>
<td>442</td>
<td>414</td>
</tr>
<tr>
<td>55846881</td>
<td>English</td>
<td>2004</td>
<td>370</td>
<td>346</td>
</tr>
<tr>
<td>37437347</td>
<td>English</td>
<td>1997</td>
<td>344</td>
<td>307</td>
</tr>
<tr>
<td>23132502</td>
<td>English</td>
<td>1991</td>
<td>343</td>
<td>312</td>
</tr>
<tr>
<td>56630747</td>
<td>English</td>
<td>2004</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>389450949</td>
<td>English</td>
<td>1991</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Constantine P. Cavafy

From Wikipedia, the free encyclopedia

Constantine P. Cavafy, also known as Konstantin or Konstantinos Petrou Kavafis, or Kavaphes (Greek Κωνσταντίνος Π. Καβάφης) (April 29, 1863 – April 29, 1933) was a major Greek poet who worked as a journalist and civil servant. He has been called a skeptic and a neo-pagan. In his poetry he examines critically some aspects of Christianity, patriotism, and homosexuality, though he was not always comfortable with his role as a nonconformist. He published 154 poems; dozens more remained incomplete or in sketch form. His most important poetry was written after his fortieth birthday.

Contents

1 Biography
2 Work
   2.1 Historical poems
   2.2 Sensual poems
   2.3 Philosophical poems
3 Bibliography
4 Notes
5 External links

Biography

Cavafy was born in 1863 in Alexandria, Egypt to Greek parents, and was baptized into the Greek Orthodox Church. His father was a prosperous importer-exporter. He spent some time in England in earlier years and acquired British nationality. After his
**Collection Analysis: Gestalt effect**

The screenshot shows a data analysis interface for analyzing library collections. The interface is designed to display collection analysis results, focusing on unique items across various categories such as Library, Division, Language, Format, and Audience. The data is organized in a table format with columns for totals, unique items, and shared items across different categories. The interface also includes options for displaying results as numbers only and various analysis filters.
Delivery Services

Enterprise product architecture

• WorldCat Resource Sharing
• International Fee Management
• International Registry
• Super Resolver

Global

• Regional Delivery Systems
• Accept Request Transfer Message
• Use Registry & Fee Management

Local Holdings
• Update global registry
• Link to OPAC for availability
• URL resolver

OCLC Delivery Strategy
Registry use: links to Full Text (via OpenURL)
Registry Machine Level Services

APIs for External Delivery Systems
WorldCat sends Request Transfer Message to User’s Library’s Delivery System.
Current and Future Strategy

- Help libraries worldwide to provide world class service that are cohesive
  - Flexibility in integration of regional data and services

- Use of WorldCat Grid Services
  - Data access via download and web services
    - E.g. Worldcat API enriched data, mined data, linked data, authority data
  - Software components

- Partner involvement
  - E.g. Job Management

- Identify standards gaps and help develop them
  - Promote existing standards and interoperability
CBS – Heal Context

- Evolution of work done to date: Capabilities
  - load data from the **Advance** system
  - determine **best format** for the union catalogue (MARC21, UNIMARC, other)
  - interoperate with **existing ILS systems** (ADVANCE, Horizon, Aleph, Virtua and AVEKT)
  - interoperate with emerging **open source** systems
  - Interoperate with **content management systems**, **electronic resource** management systems and **repositories**
Virtual Union Catalogue

• **Advantages:**
  - Light weight organization
  - Relatively easy to implement; easy to get agreements
    - Need metasearch engine and portal
  - Ad-hoc basis: easy to extend – easy to disconnect

• **Disadvantages:**
  - Slow
  - Limited common searching accesses
  - Limited sorting / relevance
  - Duplicates
  - Searching inefficiencies
  - Varying quality
  - **Availability of all systems**
  - Configuration maintenance

---

**Search / Retrieve Standards**
- Z39.50
- SRU
- OpenSearch

**Maintenance Standards**
- none
Update protocols compared

• **SRU Update**
  - Two *databases* are *not identical* – e.g. A union catalogue is not a slave of the contributing catalogues

• **immediate availability** of records
  - Scalability, no scheduling & batch maintenance

• interactive *feedback* – alignment, enrichment
  - Identifiers
  - Inter-record links – authorities,

• Diagnostics – diffused *error resolution*

• **OAI & FTP**
  - Suited for batch loading
  - OAI is used where the data source exposes the metadata to be harvested for discovery purposes
  - Where the data source does not expect the data to be changed by the harvesting site or does not care
  - Hence one way – no mechanism for response or future synchronisation
Local clients

External data bases

WinIBW

Union catalogue

Local systems

browser

ports

Tolk

SRU

SRW

MIS

ILL

Import

Export

PSI

CAT

CBS components