

Consortium of European Research Libraries & UNIMARC

Anthony G. Curwen

Formerly Consultant, CERL, &
Corresponding Member, PUC

March 2008

CERL & UNIMARC

CERL

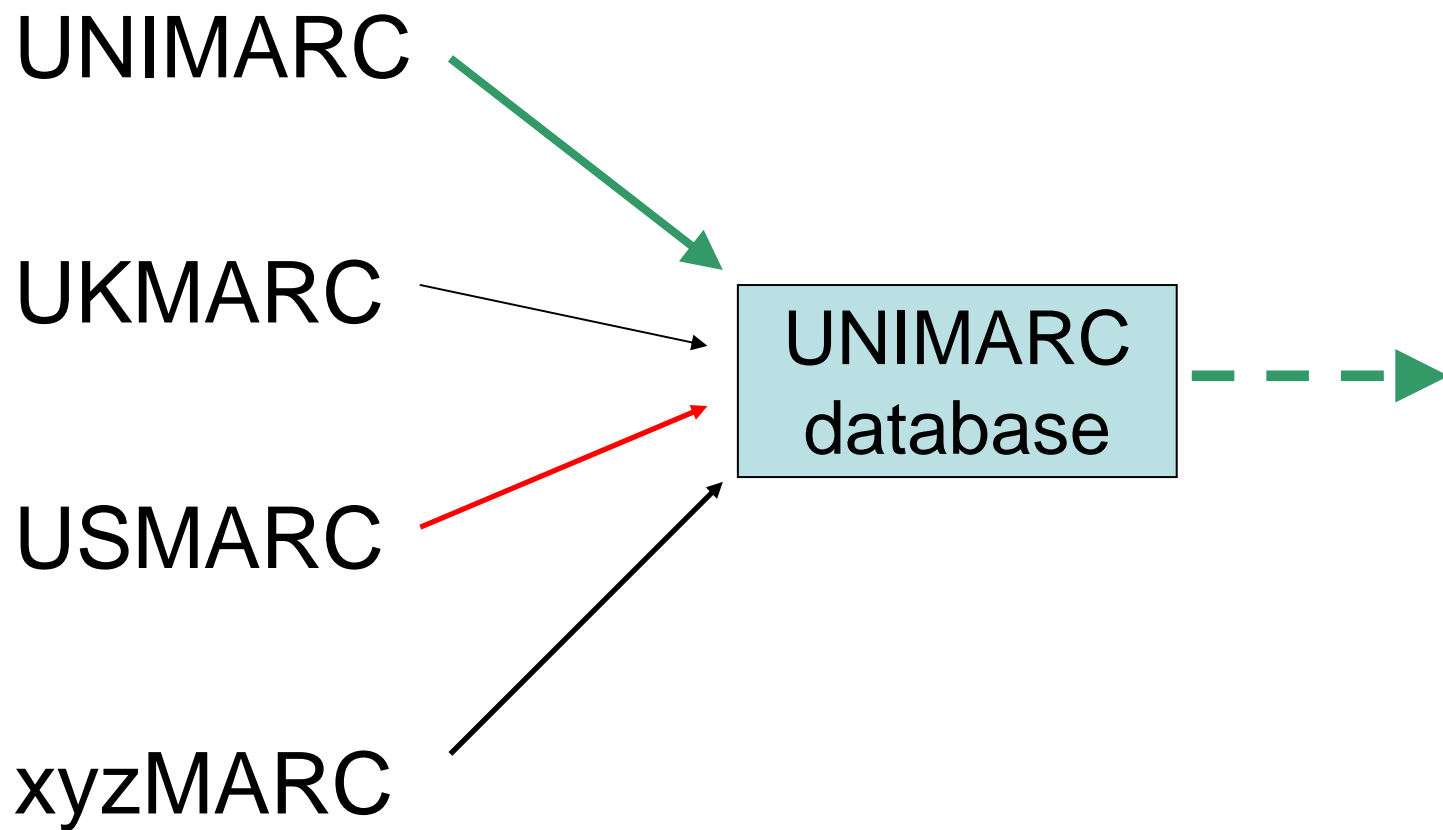
A consortium to create a union catalogue of books of the hand press period, 1450 – ca. 1830, to be a resource for

- scholars conducting research into the period through its printed heritage
- cataloguers and bibliographers

CERL & UNIMARC

- *The thinking in 1993/94 :*
- UNIMARC as standard format for the European consortium?
- Database provider required to handle UNIMARC (and other formats!), particularly multi-volume records on 2+ hierarchical levels
- As by-product, a large central source of good UNIMARC records, both original and converted

CERL & UNIMARC



CERL & UNIMARC

- Several bids for contract to be CERL's database host received from organizations both inside and outside Europe.
- *IF* a European organization could have met all CERL's requirements at that time, UNIMARC history might have been different!

CERL & UNIMARC

... but Research Libraries Group, inc. (RLG) won the contract. It had wide experience with major libraries holding important collections of early books.

RLG accepted records in all formats and converted them to RLINMARC (internal USMARC) for the Hand Press Books (**HPB**) database.

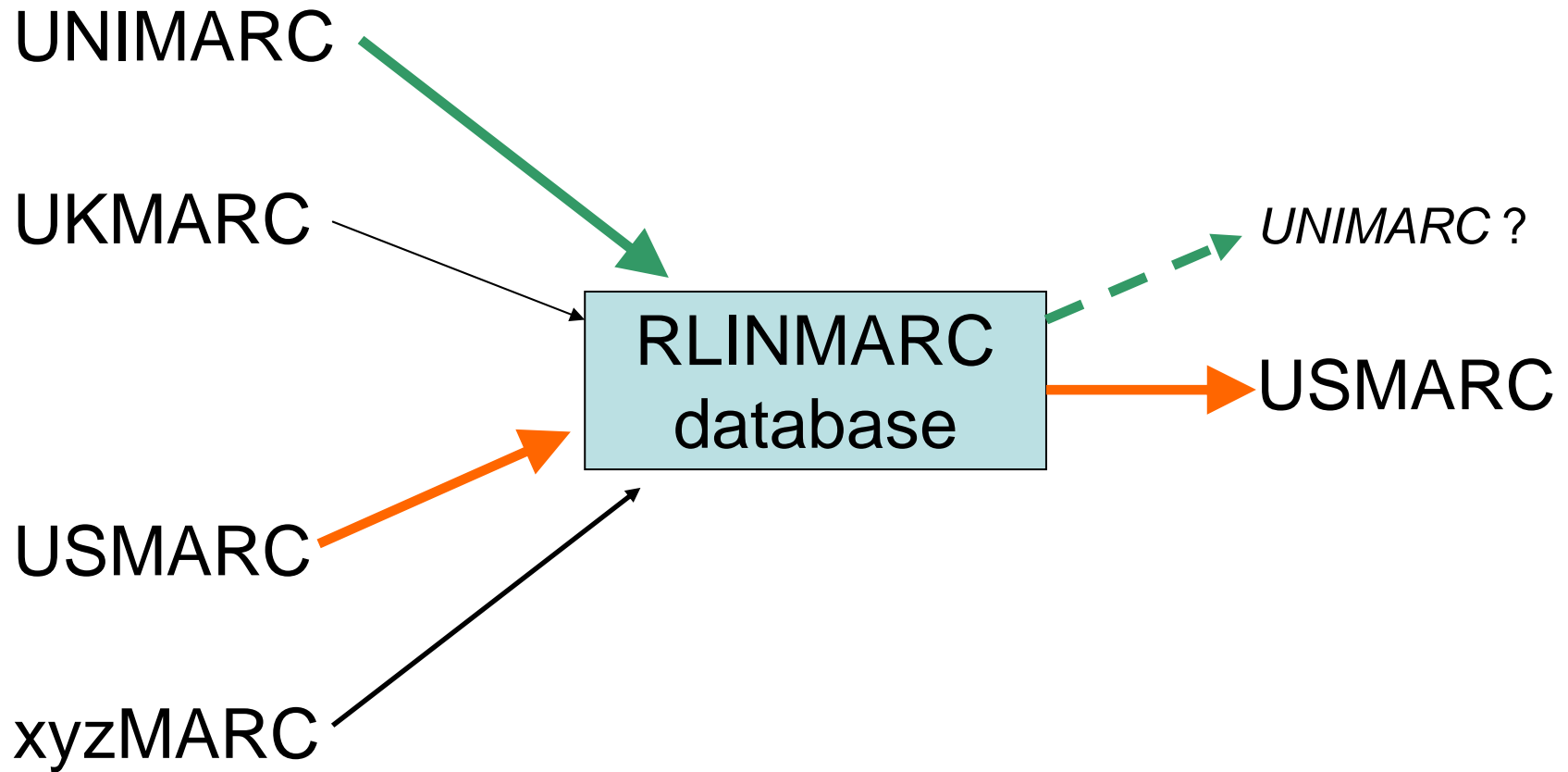
CERL & UNIMARC

RLG (a) met several CERL requirements for special indexes (access points) for old books, and other features;

(b) preserved in RLIN 886 fields any fields & subfields not fully converted, and

(c) prepared a detailed USMARC-to-UNIMARC conversion specification for CERL members who wanted to download records in UNIMARC (rather than the usual USMARC) from the RLIN databases.

CERL & UNIMARC



CERL & UNIMARC

For UNIMARC records in the HPB there was some loss of data (for example, coded fields not matched fully or at all in USMARC),

and ***multi-level UNIMARC records*** were a particular problem.

USMARC was not designed to display hierarchical structures!

A solution was devised to link such records...

CERL & UNIMARC

- *Record control numbers were used as links between screens for 'parent' and 'child' records, as in this imaginary example ...*

Record ID PQRS1234567-E
Author Newruc, G. A.
Title The true historie of Spuriosa.
Imprint London : at the Sign of the Red Herring,
 1632-1637.
Phys. Desc. 3 v. (8o)

For details of vols. in this set, enter Related Record IDs:
PQRS12345678-E PQRS12345679-E
PQRS12345680-E

CERL & UNIMARC

However doing this retrieved only the converted lower-level record (which may have had no distinctive title) and no information about the parent work:

Record ID PQRS1234569-E

Title 2. [Vol.] II.

Imprint 1635.

Phys. Desc. xviii, 326 p.

For details of the complete work, enter Related record ID
PQRS1234567-E

... which took the puzzled user back to the beginning again.

CERL & UNIMARC

2006 : RLG merged into OCLC.

- HPB went with it, as a distinct database.
- Enhancements already agreed with RLG were incorporated
- Display of converted UNIMARC multilevel records much improved by adding short titles of 'parent' records to the 'children', etc.
- *But it's still a 'flat-record' compromise!*

CERL & UNIMARC

UNIMARC export

Export of records from HPB and other files in UNIMARC format on demand?

No! — Little evidence that it had been used much (or at all?) in RLG days; and no CERL libraries using UNIMARC objected to removal of this requirement after joining OCLC.

Other techniques available now.

CERL & UNIMARC

Sources of CERL's UNIMARC files:

Croatia

Netherlands

France

Poland

Germany

Russia

Italy

Slovenia

Lithuania

Sweden

CERL & UNIMARC

Requirements for rare books:

Fields for rare books are essential for CERL member libraries, which have directly or indirectly proposed several of them to the Permanent UNIMARC Committee.

CERL & UNIMARC

Requirements for rare books:

Copy-specific data is very important

Edition 3 should see further provisions, e.g. use of subfield \$5 for library + shelfmark in 6XX subject fields:

6XX Subject field(s) for the book as published

6XX Subject field(s) for added copy-specific matter

\$5 Location + shelfmark for the copy

CERL & UNIMARC

Requirements for rare books:

Provenance is increasingly important: CERL is taking a lead in promoting the recording of provenance data.

Traditionally, provenance = ownership, but CERL's wider view = the history of individual copies of books following their publication – “*the itinerary of the book*”.

CERL & UNIMARC

Requirements for rare books: Provenance

“The itinerary of the book”

Who sold, auctioned, bought, presented... the item?

Where did this happen? *And when?* *What* else happened to it?

Can these data be systematically indexed and searched in the system?

CERL & UNIMARC

Requirements for rare books: Provenance

Edition 3 of UNIMARC Bibliographic will contain 620 (expanded) for the item when published, & also 621 for Place and Date of Provenance –

where was the individual copy after publication, and *when*?

621 will allow for this and even the recording of partial data – provenance marks are frequently incomplete.

Requirements for rare books: Provenance

Important for European cultural history –the transmission of knowledge, the development of the book trade, the building and dispersal of collections, etc.

Also important for the *security* of collections and the restitution of illegally obtained works of art, including rare books, etc. : carefully recorded manuscript initials and dates and other marks can establish ownership.

CERL & UNIMARC

Requirements for rare books: Provenance

A problem: even if new fields and subfields are provided for provenance data and libraries start using them, has every library

- (a) the *time* to investigate it thoroughly?
- (b) the *expertise* for it, either in-house, or readily available?

Recording provenance in detail = a *lot* of work!

CERL & UNIMARC

Requirements for rare books: Provenance

Gunilla Jonsson's suggestion :

Enlist the help of our expert users – when they request books, give them forms and invite them to note their suggestions about unidentified provenance marks, etc.

When verified, these can be added to the record.

Cf. CERL's "Scholars' notepad" in HPB

CERL & UNIMARC

Assisted Searching: The CERL Thesaurus

Publications in XV – XVIII centuries show many variant names and variant spellings, personal, corporate and geographic.

For instance, how can one search for Firenze OR Fierenze OR Florenz OR Florence OR Fiorenze OR Fiorentia ... and be sure that all are included?

CERL proposed an on-line Thesaurus ...

CERL & UNIMARC

*Assisted Searching:
The CERL Thesaurus*

*Devised and maintained for CERL by the
Data Conversion Group (DCG), at
Niedersächsische Staats- und Universitätsbibliothek,
Göttingen*

*Assisted Searching:
The CERL Thesaurus*

- Derived principally from member libraries' authority files, *with, progressively*
- References to standard sources
- Links to digitised printers' devices
- Linking of place names + printers
- Limiting to HPB records which have provenance information
- Users' comments

*Assisted Searching:
The CERL Thesaurus*

CT contains (2008-02-05)

	<i>Number of records</i>
Imprint places	4 186
Imprint names (personal + corporate)	29 732
Other personal names	641 182
Other corporate names	7 803
Reference works cited	<u>2 107</u>
<i>Total</i>	<i>685 010</i>

*Assisted Searching:
The CERL Thesaurus*

which are subdivided into

- Non-fictional names
- Fictional names
- Identification uncertain
- Names used for more than one place, imprint, person or body

*Assisted Searching:
The CERL Thesaurus*

CERL members can use the Thesaurus for *Assisted searching* in the HPB database.

A simple search (e.g., for Firenze as place of publication) can be seamlessly converted to a multiple 'OR' search by using the Thesaurus on its database in Göttingen directly from the OCLC screen...

Assisted Searching: The CERL Thesaurus

FirstSearch

HandPressBook Basic Search

- Enter search terms in one or more boxes and click on **Search**.

Searching [Stat](#)

Basic Search **Advanced Search** **Expert Search** **Previous Searches** ▼

i **News** **?** Current database: **HandPressBook**

Intro **News** **Help**

	<input type="button" value="Search"/> <input type="button" value="Clear"/>
Search in database:	HandPressBook i Union catalog of European printing from the 15th century to the 1830s
Keyword	<input type="text" value="PL: FIRENZE"/> ?
Author	<input type="text"/> ?
Publisher	<input type="text" value="Title: Enter terms appropriate for the box (see Help)"/>
Title	<input type="text"/> ?
Year	<input type="text"/> (format: YYYY-YYYY) ?
	<input type="button" value="Search"/> <input type="button" value="Clear"/>

Assisted Searching: The CERL Thesaurus

PL: FIRENZE retrieved. Click on 'Find variations...'

The screenshot shows a web browser window with the URL <http://firstsearch.oclc.org/WebZ/FSQUERY?format=BI:next=html/records.html:bad=html/records.html:numrecs=10:sessionid>. The page title is "FirstSearch: Hand Press Book List of Records". The main heading is "Hand Press Book List of Records". Below the heading, there are instructions: "Click on a title to see the detailed record." and "Click on a checkbox to mark a record to be e-mailed or printed in Marked Records." A green link says "Refine your search results: 8,903 (Show me how)". The navigation bar includes "Searching" and "Results" tabs, with "Results" selected. There are links for "Staff View", "My Account", "Options", and "Comments". Below the navigation bar, there are tabs for "List of Records", "Detailed Record", and "Marked Records", with "List of Records" selected. A "Go to page" dropdown menu is visible. A toolbar contains icons for "Related Subjects", "Related Authors", "Limit", "E-mail", "Print", "Export", and "Help". The search results section shows "Hand Press Book results for: pl: FIRENZE. (Save Search)" and "Records found: 8,903 Find variations of this Publisher Location (FIRENZE)". A red arrow points from this link to a callout box. Below the search results, there is a "Prev" and "Next" navigation bar with a page number "1" in the center. The first search result is "1. **Tragedie.** Author: Alfieri, Vittorio. Publication: Firenze, 1814. Document: Book".

Records found: **8,903** [Find variations of this Publisher Location](#)

Assisted Searching: The CERL Thesaurus

CERL Thesaurus matches for "FIRENZE": 2

Click a **Search** button to search for variants of your term.

[No Title Available](#) (1) [Firenze](#) (31)

... And a click on 'Search' takes the user back to the OCLC screen

Firenze

1. **Firenze**
2. **Florenz**
3. Florence (Italy)
4. Fierenze
5. Firenze
6. Fiorentia
7. Fiorentina
8. Fiorenz
9. Fiorenza
10. Fiorenze
11. Fiorenzia
12. Fiorneza

Assisted Searching: The CERL Thesaurus

where OCLC has now performed a multiple 'OR' search:

The screenshot shows a web browser window displaying the FirstSearch interface. The search query is: `pl= "Firenze" OR pl= "Florenz" OR pl= "Florence Italy" OR pl= "Fierenze" OR pl= "Florence" OR pl= "Fiorentia"....`. The results show 13,555 records found. A green box highlights the text "Refine your search results: 13,555" with a green arrow pointing to the search results area.

Hand Press Book List of Records

• Click on a title to see the detailed record.
• Click on a checkbox to mark a record to be e-mailed or printed in Marked Records.
• **Refine your search results: 13,555 (Show me how)**

Searching Results Staff View My Account Options Comments Exit Hide tips

List of Records Detailed Record Marked Records Go to page

Hand Press Book results for: pl= "Firenze" OR pl= "Florenz" OR pl= "Florence Italy" OR pl= "Fierenze" OR pl= "Florence" OR pl= "Fiorentia".... (Save Search)
Records found: 13,555

Prev Next

1. **Discorso dell'eccellentiss. filosofo m. Francesco de' Vieri cognominato il Secondo Verino. Intorno a' demonii, volgarmente chiamati spiriti.**
Author: Vieri, Francesco de, fl. 1547-1590.; Sermartelli, Bartolommeo,
Publication: In Firenze, : Appresso Bartolomeo Sermartelli,, MDLXXVI. [1576]
Document: Italian : Book.
[See more details for locating this item](#)

2. **Painters, Italian.**
Author: Campiglia, Giovanni Domenico,; Pazzi, Pier Antonio,

Refine your search results:
13,555

*Assisted Searching:
The CERL Thesaurus*

— but while the Thesaurus is a great help for CERL members using the HPB Database, *it is also free for everyone* to use as a research tool in its own right.

Assisted Searching: The CERL Thesaurus

CERL Thesaurus

Click a **Search** button to search

[No Title Available](#)

Cancel

CERL Thesaurus

Clicking on this button brings up the Thesaurus data in structured form ...

No Title Available

1. Firenze

Search

Search Firenze

1. **Firenze**
2. **Florenz**
3. Florence (Italy)
4. Fierenze
5. Firenze
6. Fiorentia
7. Fiorentina
8. Fiorenz
9. Fiorenza
10. Fiorenze
11. Fiorenzia
12. Fiorneza
- ...

CERL & UNIMARC : *CERL Thesaurus*

...Like this:

Record Identifier	cnl00005908
Headings	Firenze Florenz
Found in	[Diercke] [Geo-Duden] [LGB(2),2,617] [Orbis Latinus, sv. 2, str. 78]
Geographical Notes	Italija
Variant Forms	Florence (Italy) [CLC] Fierenze Fiorenze Fiorentia <i>[+ 25 more]</i> Phlorentia
Fictional Forms	Amsterdam Gelopoli Londra

CERL & UNIMARC : *CERL Thesaurus*

Am I spending too much time advertising the CERL Thesaurus when I ought to be talking about UNIMARC?

Above that structured display on the previous screen there are five icons for various functions. Click on the middle one, 'Display internal format', and you will see...

CERL & UNIMARC : *CERL Thesaurus Format*

0RL 00000nz####22000003##45##
001 cnl00005098
035 ##\$zcnl00000225
035 ##\$zcnl00004609
100 ##\$a20050322xmulv50#####ba
110 ##\$a0
215 #1\$aFirenze**\$c**HR**\$5**CiZaNSK
215 #1\$aFlorenz**\$c**DE**\$5**GYMG**\$c**BSB-Ink
290 #1\$aDiercke**\$a**Geo-Duden**\$a**LGB(2),2,617
290 #1\$aOrbis Latinus, sv. 2, str. 78
356 #1\$8scr\$aItalija
415 00\$aFlorence (Italy)**\$c**CLC
415 01\$aFierenze
[+ 27 more]
415 01\$aPhlorentia
415 11\$aAmsterdam
415 11\$aGelopoli
415 11\$aLondra
801 21\$aDE**\$b**GYMG**\$c**20010822**\$n**SAUR-2:00001002
801 21\$aHR**\$b**NSK**\$n**210706107**\$g**PPIAK

CERL & UNIMARC : *CERL Thesaurus Format*

I consider this to be a brilliant adaptation and application of UNIMARC.

The CT format is one third pure UNIMARC Authorities, one third very slightly adjusted U/A, and one third additional fields and subfields for the needs of the thesaurus.

This 'child of UNIMARC' is thriving in Göttingen!

CERL & UNIMARC

For general information about CERL :

<http://www.cerl.org>

and to connect to the Thesaurus click on

CERL Thesaurus Search

For general information about CT:

http://www.cerl.org/web/en/resources/cerl_thesaurus/main

CERL & UNIMARC

Thank you!

tony.curwen@zen.co.uk