LIBER QUARTERLY 19 (3/4), February 2010 — ISSN: 1435-5205. P324–335
http://liber.library.uu.nl/
Igitur, Utrecht Publishing & Archiving Services

Some ISIS-Software History and Technical Background on the new FOSS Integrated Library System ABCD

Egbert de Smet

University of Antwerp, Prinsstraat 13, B-2000 Antwerp, Belgium,
egbert.desmet@ua.ac.be

Abstract

While free and open source software is gaining momentum in library automation, it is still not a trivial event to launch a full-blown integrated library system (ILS), given today's high demands on such library systems. In this article we describe the ISIS-software origins of the newly launched ABCD-software for which there is a more-than-average interest in many parts of the libraries and information communities in the developing world. Many — especially younger — librarians in the West have never heard about ISIS, its relevance and its technological concepts, so we will try to shed some light on this in order for librarians to better judge the meaning of the ABCD software for international librarianship and — why not — perhaps even for their own purposes.

Key Words: ISIS; CDS/ISIS; ABCD; integrated library system

The Origins of the ISIS Software and the ISIS Family

For those who know about ISIS (or, with its full name: CDS/ISIS), there is often a partly correct but also partly incorrect association with 'old' (and outdated?) software. Indeed the origins of ISIS go back to the 1970s — almost pre-history in computer science — when within the International Labour Organisation, one of the many United Nations member organisations, two earlier developed documentation systems were combined into one: the 'Central Documentation System' (CDS) merged with the 'Integrated Set of Information Services' software and became CDS/ISIS. This software, running on the obvious platforms of the time (VMS and large computer systems), ran

as a series of specialised functions and commands to store or retrieve records in a database (with its own idiosyncratic database format) and to apply B-Tree based indexing to them, including 'word-indexing', which was far from obvious at the time (and is known nowadays as the very popular 'full-text' indexing). CDS/ISIS was one of the first software packages implementing the new principles of 'ISAM' (the Index-Sequential Access Method) on which many database technologies were based later on. The main developer, Gianpaulo Del Bigio, was not only a specialist programmer but he was also familiar with library science concepts such as UNISIST (another old system on which many modern bibliographic systems were based).

UNESCO became highly interested in CDS/ISIS as a tool, because it could perhaps be offered to its less-developed country partners for advanced and professional bibliographic control. Mr. Del Bigio moved to UNESCO and migrated the software in 1985 to the then new 'micro-computer' platform of DOS to present 'Micro CDS/ISIS' as one of the first database software packages that not only could run on two floppy disks (A and B) but also offered powerful storage and indexing capacity.

This Micro CDS/ISIS for DOS became a huge success in its world-wide distribution and use, mostly in developing countries. UNESCO received and responded to tens of thousands of requests for the software and sent it out on floppy disks (later also on CDs) to all hidden corners of the South. By this time the software (as from version 2.3) was fully menu-driven and in its version 3 it offered full networking capability with multi-user database, record and index protection. The software also incorporated — again because of the genius of Del Bigio — a built-in programming language (CDS/ISIS Pascal), which not only allowed for the development of local functions for data manipulation, but even for the creation of independent interfaces. In addition the menu structure was not only language-independent (allowing for the creation of versions in any language, including the Chinese version which I saw still widely used there back in 1995), but also for changes to the menu structures and sequences. I described this very 'open' architecture of the software elsewhere to illustrate how CDS/ISIS in fact represented a forerunner of 'open' software (de Smet, 2008).

This enormous success later on also became a handicap: for a long time the software continued to run in its typical (DOS-based non-graphical) 'black window' where the keyboard had still to be used to invoke functions with-

out the use of a 'mouse', while especially in Western libraries GUI (graphical user interface) techniques — Windows/Icons/Mouse/Pull-down menu's — became ubiquitous. Slower access to more powerful PCs and lower replacement cycles in the South caused this situation to last much longer in the developing world and contributed to the 'old-fashioned' (and hence outdated) image of the software.

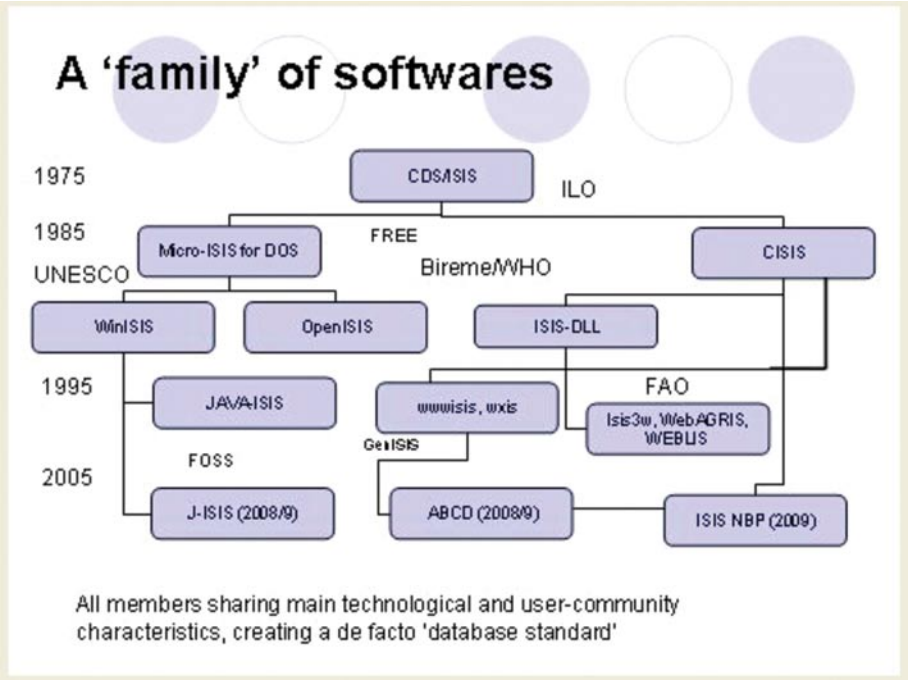
From 1995 on UNESCO and Del Bigio worked on a full migration of the software to Windows: 'CDS/ISIS for Windows' was presented in 1997 (at the first World Congress on CDS/ISIS in Bogota, Colombia), and became very popular world-wide as 'WinISIS'. The power of the software, mostly based on a powerful 'formatting language' (instructions to exactly define how data from the fields in the records are extracted and presented to several software functions such as display, sorting, indexing, exporting, etc.) remained in tact while lots of graphical features became available: hyperlinks, multi-media, run-time relational links to other databases, etc.

Again ten years later, in 2005, several 'WWW-capable' server versions were developed (ISIS3W from Poland, OpenISIS in Germany, and most importantly 'wwwisis' from Brazil), showing not only that the software kept closely in touch with new technological developments, but also that development centres and expertise in many countries out of the UNESCO-context were getting involved. For instance, the Polish technology was implemented by another UN organisation, FAO, for use in its AGRIS-system for agricultural information and later also in ASFISIS for fishery and aquatic sciences. The Brazilian institute 'BIREME', which is the Latin-American information centre of the World Health Organisation, now runs huge databases (16 million records) with abstracts and online full text based on ISIS technology developed in-house; BIREME has become the main development centre of ISIS.

In 2005 UNESCO opted fully for 'free and open source software' and shifted its focus to wider support for this type of software packages, decreasing its direct input in software development of CDS/ISIS and handing this over to a 'FOSS ISIS community' (see www.iccisis.org). Elsewhere I discuss the fact that ISIS in fact always was 'open' and 'free' software in view of its built-in openness (de Smet, 2008). Still in 2009 a complete overhaul based on a client-server Java architecture was presented as 'J-ISIS', now based on Berkely DB unlimited capacity and Lucene (full-text) indexing.

A family picture of the many members of the ISIS family, all reflecting these technological concepts, looks as shown in Figure 1 (with of course some necessary simplifications and omissions):

Fig. 1: The ISIS family of software packages.



In Figure 1 a general overview is given, reducing — for simplicity reasons — the family to more or less four generations, each after a time span of about ten years. The involvement of some UN organisations is indicated with two main streams: the UNESCO stream leading to the new J-ISIS and the BIREME stream leading to ABCD and the new 'Network Based Platform'.

The Specific Technological Concepts of ISIS

The brief historical sketch above shows how (CDS/)ISIS not only became widely used in libraries in many, mostly Southern, parts of the world, but

also how others took up the technology and began programming software based on the same technological concepts. These concepts, which constitute the common 'family characteristics' of the ISIS software family, can be summarised as follows:

- 'scheme-less' records (i.e., no fixed structure is imposed on the records, but each record carries its own 'identity card' as an ISO-2709 header describing structure and contents) with variable length and structure (i.e., fields of any length can be present 0 up to x times in any record).
- All data are stored in fields and subfields which are simply identified by their numerical tag followed by the string of the field data; momentarily, this basic 'tag value' approach is again gaining interest in computer science thanks to a new 'NO-SQL' movement (see e.g., NoSQL, 2009; Lai, 2009).
- All data from the fields go through a 'filter' mechanism steered by a detailed and powerful 'formatting language'. This grammar enables librarians to define the use of their data in very detailed and advanced ways (including relational use of data from other databases) without having to become real programmers, and in our view this is the basis of ISIS's rather unique capability to allow for sophisticated local development and management in situations where computer science skills and knowledge are often very limited.
- Based on the ISIS Formatting Language values in the databases can be indexed in many ways into an 'Inverted File' where all extracted strings are accessible in an almost volume-independent fast (B-tree) structure for immediate retrieval of the records (with exact information on field occurrence and word positions) linked to that extracted value.
- The Formatting Language also allows for importing and exporting the data according to very specific instructions, largely taking away the need for conversion expertise when migration is needed.

Throughout its history the software has used various capacity parameters, from an 8 Kb record limit and 512 Mb maximum database size in the original DOS-version via a 32 Kb limit in Windows up to 1 Mb records today in special editions of the CISIS-library with a 4 Gb max. database size. Currently the newest generation of ISIS (ISIS4G) limites records to 2 Gb while the size of the data base is only limited by the hardware and OS. This will allow for

complete handling of full-text documents within ISIS without having to split the documents into a number of records beforehand. Practice has shown that limits as described above have seldom been major hindrances to real applications, which mostly concerned bibliographic or descriptive factual information.

As one can see in the family picture above (Figure 1), 2009 was a crucial year with three quite different new developments, one of them being ABCD which we will describe later after we briefly discuss some non-technical characteristics of the software. The ISIS 'Network Based Platform' mentioned is BIREME's effort to introduce ISIS into the most recent software technologies of unlimited storage capabilities, full-text indexing with vector-space and relevance ranking concepts based on Python, but also providing a strong abstraction layer (through the 'ISIS cell' concept) allowing for links with any ISIS environment.

The Specific Social Reality of ISIS

In addition to its rather unique technological concepts, ISIS as a software family also represents a specific social reality. Software packages nowadays are much more than just technical artifacts, they have a social dimension, including strong 'believers' and 'non-believers'; it is indeed remarkable to note how even political interventions and 'rising temperatures' guide decisions on software selection — also when it comes to library automation projects.

Due to its development within the UN, ISIS has both its defenders (believers) and opponents. It never incorporated 'SQL' or relational database concepts, which are still very dominant in modern computer science; most computer science students in the developing countries will only have to read a textbook on SQL when it comes to databases. Therefore, ISIS has never gained strong support in computer science environments.

However, ISIS did raise interest among library and information experts who developed into computer scientists without losing sight of the very specific demands of bibliographic and documentary information systems. Standards such as ISBD, MARC and ISO2709, but also UNISIST (with the concept of bibliographic levels) and thesaurus technology remained at the

core of the database technology itself. When the WWW became ubiquitous, ISIS again did not need input from computer scientists who thought they had re-invented the information world: by simply using the long-standing Formatting Language, ISIS records could be presented as HTML-pages. The same story applies to XML ...

Apart from its image as 'computer science by and for information workers', ISIS has another telling characteristic which also originates from its UN basis, i.e., the image of the typical developing world user environment. Just like in car-production Volkswagen was said to be 'Mr. Nobody's Porsche' (but we know how this story ended recently ...) ISIS was sometimes seen as a 'poor man's library system'. Foregoing the fact that ISIS in fact is not even a library system, but only a general-purpose tool allowing poor librarians to develop their own sustainable databases to semi-automate their libraries, in academic publications ISIS compared unfavourably with, e.g., KOHA — a comparison which now needs to be made again with the arrival of ABCD! — and was seen as a poor substitute for the 'real thing', when means or skills were not available or insufficient to allow for the use of professional software packages. Even from a simply technological point of view we think this judgment is unfair, but the social reality and image of ISIS as a tool for poor librarians in underdeveloped environments will be very difficult to eradicate.

Of course UNESCO focused on promoting the software in developing countries — according to its own specific mandate and for very good reasons. Not only does ISIS offer its professional capabilities without the need for state-of-the-art software skills, but it also runs on moderate or even low-level hardware — I recall many situations where African libraries were still using old Windows 3-based or even non-graphical-capability PCs to run ISIS.

In addition, users of ISIS in the Western world were typically to be found in non-profit environments — the free distribution of ISIS by UNESCO was even explicitly based on the non-profit requirement —, again strengthening the image of ISIS of being linked with economically less powerful users. In Flanders (Belgium), e.g., typical ISIS users used to be the NGO network on development cooperation, smaller medical (and school) libraries and cultural work, although the latter have now been merged into larger units with more economical power to purchase commercial solutions.

However, the use of ISIS in state-of-the-art systems such as the ScIELO information system of BIREME (which offers huge amounts of documents on medical and social information in open access full-text electronic journals) and its 'Virtual Health Library' proves that UNESCO's policy to focus on developing countries does not reflect a technological weakness, but rather an ambition to provide tools for opening up the information society to everybody. This is indeed UNESCO's basic mandate when it comes to information, and that is also the reason why in our view UNESCO cannot ever fully withdraw from the ISIS development scene.

ABCD as a New ISIS Family Member

This article will not fully describe ABCD as an integrated library system itself — for a full description we refer to other literature (de Smet, 2009).

We do wish to emphasize the basic statement that — for the first time in the history of ISIS — a full-blown dedicated and integrated library system is available based on ISIS technology. Other attempts to offer such integrated library solutions based on ISIS, e.g., Open MarcoPolo and Weblis, were very meritorious efforts, but never (or not yet?) reached the level of ABCD in its attempt to offer state-of-the-art technology without breaking with the typical ISIS community needs as sketched above.

The ABCD acronym, however, illustrates an ever further-reaching ambition: not only to automate libraries ('Automatisación de Bibliotécas', 'Automatisation de Bibliothèques' or AB in many languages) but also to provide a tool for automating Documentation Centres ('Centros de Documentación' etc. or simply 'CD').

What does this mean ? We think this is the most unique and special feature of the software, so it merits some elaboration here. By being fully independent from the bibliographic structures used — but already offering some widely used standards such as MARC (21 and UNI) or CEPAL (a widely used bibliographic format in Latin America) ABCD facilitates *any* structure to be created by the software itself and subsequently being managed (cataloguing, OPAC and circulation). This is because ABCD offers a full ISIS interface in a web environment (using PHP-programming) including all typical ISIS defini-

tions such as 'Field Definition Table', 'Field Selection Table' and of course the crucial Formatting Language.

This means that not only libraries, but documentation centres, archives or museums can use the software as well while retaining their own dedicated information structures. Or even better: organisations can use ABCD to integrate such related but still different functions into one web-based system. As it is fully web-based, ABCD of course can offer multi-media capability (important, e.g., for museums), but also full electronic document handling, paving the way for digital libraries.

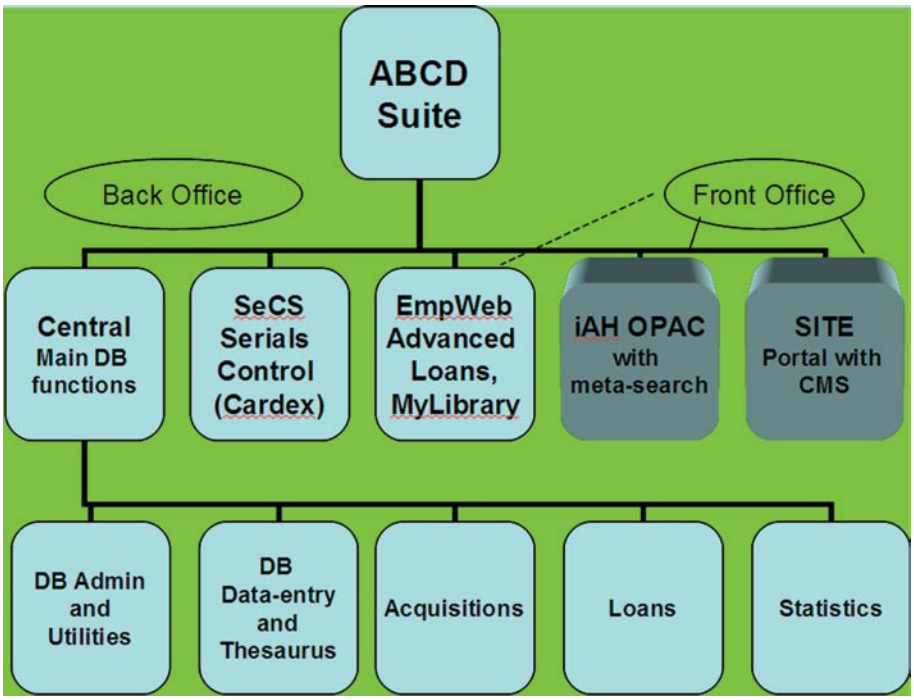
This capability, maybe not unique but still very special and powerful, in its own right offers something extra to the typical ISIS users, eradicating the need to develop skills (and/or maintenance resources) for running different systems for different purposes with a common 'information service' functionality. The users often cannot afford to develop such different support mechanisms ...

Still ABCD is presented as software which incorporates all major and important current standards in information services. When the Flemish Interuniversity Council's Development Co-operation section (VLIR/UOS) decided to support ABCD development through its project 'DOCBIBLAS' (Development Of and Capacity Building in ISIS-Based Library Automation Systems) in 2007–2010, it explicitly included requirements to allow for usage by university libraries, such as MARC, Z39.50, MODS/METS etc. In addition, ABCD comes with its own OAI-HP server allowing harvesting of records by other non-ISIS based web services.

In more than one way ABCD represents the culmination and integration of many mostly BIREME-developed ISIS tools, e.g., the meta-search capability (i.e., including any number of internal and external databases into the search), the CMS-based library 'portal', advanced Serials Management (for local and union catalogues of serial, also electronic, publications) and finally the Advanced Loans module, which allows for linking to non-ISIS user databases and multiple loans policy implementations. The ABCD software therefore is best presented as a 'suite' of cooperating but also independent software packages (Figure 2):

In this scheme the 'central' module contains the 'heart' of the software with the main crucial library automation functions, enabling smaller libraries

Fig. 2:



with insufficient technological skills to fully automate without leaving the familiar ISIS environment (as most ABCD users are expected to be current ISIS users) and building upon that often significant experience: locally developed database structures and print formats, or indeed the available skills in using the ISIS Formatting Language, can still be re-used while migration to more professional or more widely used standards is still easily within reach.

The main menu of this central module showing the multitude of functions (including, e.g., statistics) is shown in Figure 3.

Conclusion

With the newly born ISIS family member 'ABCD' a long-standing tradition of ISIS developments is continued and fully brought to bear in modern database-

Fig. 3:


ABCD
BIREME - Centro Latino Americano e do Caribe de Informação em Ciências da Saúde
System Administrator, adm | [logout]
Home -
Language: English
Module
help edit help file Script: homepage.php
Database
Select:
Marc
Statistics
Update database definitions
Configure Z39.50 client
Data entry
Reports
Utilities
Administration
Create Database
Reset last inventory number (copies database)
Users administration
Translate messages and help pages
ABCD 1.0.2
2009 BIREME - Centro Latino Americano e do Caribe de Informação em Ciências da Saúde
http://www.bireme.br
BIREME - OPAS - OMS

driven web applications. Combining the old but proven technological concepts of the historical ISIS with both web technology and integrated library and documentation centres' automation functions, ABCD brings a hopefully welcome solution and continuation for the many existing ISIS users all over the world. Perhaps even it will attract renewed interest from young, modern librarians who are open to alternatives for the commercial ILS providers who are by (economical) necessity making their globalised users' market more dependent rather than more self-supporting and self-empowered.

References

de Smet, E. (2008): 'The ISIS-Software Family: From "Free and Open" to "Free and Open Source Software"', *Innovation, Journal of Appropriate Librarianship and Information Work*, June.

de Smet, E. (2009): 'ABCD: A New FOSS Library Automation Solution based on ISIS', *Information Development* 25: 61–67.

Lai, Eric (2009): 'No to SQL? Anti-Database Movement Gains Steam', *Computerworld Software*, July 1, http://www.computerworld.com/s/article/9135086/No_to_SQL_Anti_database_movement_gains_steam.

NoSQL (2009): Wikipedia, http://en.wikipedia.org/wiki/NoSQL.