The ISIS-software origins and technological principles as framework for the new FOSS integrated library system 'ABCD'.

Egbert de Smet, University of Antwerp

Abstract
While Free and Open Source Software is still gaining momentum also in library automation, it is still not a trivial event to launch a full-blown ILS given today's high demands on such library systems. In this article we explain the ISIS-software origins of the newly launched ABCD-software for which there is a more-than-average interest in many parts of the Southern libraries and information communities. Many – especially younger – librarians in the West have never heard about ISIS and its relevance and technological concepts, so we will try to shed some light on this in order for librarians to better judge on the meaning of the ABCD software for international libarianship and – why not ? - perhaps even their own purposes.

Overview :

The origins of ISIS

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with its full name : CDS/ISIS), there is often a partly correct but also partly incorrect association with 'old' (and outdated?)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 stem back to the 1970's – almost pre-history in Computer Science – when in the International Labour Organisation, one of the many United Nations member organizations, 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 then obvious platforms (VMS and large computer systems), ran as a series of specialized functions and commands to store or retrieve records in a database (with its own idiosyncratic database format) and apply B-Tree based indexing on them, including the then far from obvious 'word-indexing' (nowadays known and very popular 'full-text'). CDS/ISIS was one of the first softwares implementing the then new principles of 'ISAM' (Index-Sequential Access Method) on which later many database-technologies became grounded. The main developer, Gianpaulo Del Bigio, not only was a specialist programmer but also close to library science concepts like UNISIST (another old system on which later on many modern bibliographic systems were based).

UNESCO got highly interested in CDS/ISIS as a tool which perhaps could be offered to its less-developed country partners for advanced and professional bibliographic control. Mr. Del Bigio moved to UNESCO and migrated the software to the then (1985) new 'Micro-computer' platform of DOS to present 'Micro CDS/ISIS' as one of the first database softwares which not only could run on two floppy disks (A and B) but still 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 request for the software and sent it out on floppy disks (later also CD's) to all hidden corners of the South. By this time

the software (as from version 2.3) was fully menu-driven and got (in its version 3) full networking capability with multi-user database-, record- and index-protecion. The software also got – again from the genious of Del Bigio – a built-in programming language (CDS/ISIS Pascal) which allowed developing not only local functions for data-manipulation but even the creation of independent interfaces. In addition the menu-structure was not only language-independent (allowing creating of versions in any language, including the Chinese version which I saw still widely used there back in 1995), but also changing the menu-structures and sequences. I described this very 'open' architecture of the software elsewhere to illustrate how CDS/ISIS in fact was a fore-runner in the concept of 'Open' software (see [reference]).

This enormous success later on became also a handicap : way too long the software kept running in its typical (DOS-based non-graphical) 'black window' where still the keyboard had to be used to invoke functions without the use of a 'mouse', while especially in the Western libraries GUI (graphical user interface) techniques – Windows/Icons/Mouse/Pull-down menu's – became ubiquitous. Slower access to more powerful PC's and lower replacement cycli in the South caused this to last much longer in the Developing World and created the somehow 'old-fashioned' (and hence out-dated) image of the software.

Still as from 1995 UNESCO and Del Bigio worked on a full migration of the software into Windows as 'CDS/ISIS for Windows', presented in 1997 (at the first World Congress on CDS/ISIS in Bogota, Colombia), which again 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 like display, sorting, indexing, exporting..) remained while now lots of graphical features became available : hyperlinks, multi-media, run-time relational links to other databases etc.

Again 10 years later – 2005 – the software got several 'WWW-capable' server-versions (ISIS3W from Poland, OpenISIS in Germany, and most importantly 'wwwisis' from Brazil), proving not only that the software kept closely in touch with new technological developments but also got development centers and expertise in many countries out of the UNESCO-context. E.g. the Polish technology was hired by another UN organization : FAO for use in its AGRIS-system for agricultural information (later also ASFISIS for fishery and aquatic sciences) and the Brazilian institute 'BIREME' is the Latin-American information center of the World Health Organization, nowadays running huge databases (16 million records) with abstracts and full-text online based on ISIS-technology developed in-house and meanwhile having become the main development center of ISIS.

UNESCO opted in 2005 fully for 'Free and Open Source Software' and shifted its focus to a wider support for this type of softwares, lowering its direct input in software development of CDS/ISIS and handing this over to a 'FOSS ISIS community' (see www.iccisis.org). 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.

ISIS as a software family

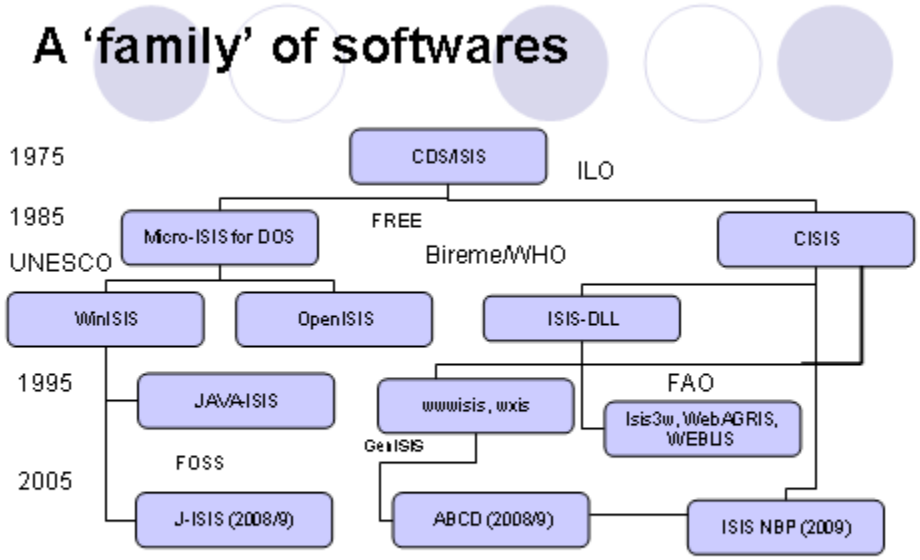
The brief historical sketch above shows how (CDS/)ISIS not only became widely used in libraries in many, mostly southern, parts of the world but also how others took up the technology and started programming software based on the same technological concepts. These concepts, which form the common 'family characteristics' of the ISIS-software family, can be summarized as follows :

- 'scheme-less' records (i.e. no fixed structure is implied onto 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 with their numerical tag followed by the string of the field-data; this basic 'tag-value' approach nowadays is gaining interested in computer science again thanks to a new 'NO-SQL' movement (see [reference]).
- All data coming from the fields go through a 'filter' mechanism steered by a detailed and powerful 'formatting language'; this grammar allows librarians, without a need to become real programmers, to define the use of their data in very much detailed and advanced ways (including relational use of data from other databases) and in our view is the basis for ISIS' rather unique capability to allow sophisticated local developments and management with still often very limited availability of computer science skills and knowledge.
- 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 allow importing and exporting the data according to very specific instructions, taking away mostly the need for conversion expertise when migration is needed.

A family-picture of the many members, all reflecting these technology concepts, looks as follows (with of course some necessary simplifications and omissions) :



All members sharing main technological and user-community characteristics, creating a de facto 'database standard'

As one can see, last year 2009 was a crucial year with 3 quite different new developments having been presented, one of them being ABCD on which we will focus after we also briefly discussed 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 links with any ISIS-

environment.

The specific social reality of ISIS

Regardless of its rather unique technological concepts, ISIS as a software family also represents a specific social reality. Softwares nowadays are much more than just technical artifacts, often they include phenomena like strong 'believers' and 'non-believers' – it is indeed remarkable to note how even political interventions and 'raising temperatures' guide decisions on software selection, also when it comes to library automation projects.

Due to its background in the UN, as illustrated above, ISIS has both its defenders (believers) and opposition. It never used '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 had to read a textbook on SQL when it comes to databases – and therefore also never got a strong interest in computer science environments.

The interest however came from library and information experts who converted them to computer scientists without losing sight on the very specific demands of bibliographic and documentary information systems. Standards like 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 : simply using the long-standing Formatting Language ISIS records could be presented as HTML-pages. Same story for XML...

In addition to this aspect 'computer science by and for information workers' ISIS – again from its international UN-based bias – carries a second characteristic : the 'developing world' typical users-environments. 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 “poor man's library system'. Forgetting that ISIS in fact is not even a library system, only a general-purpose tool allowing poor librarians to develop their own sustainable databases to semi-automate their libraries, ISIS compared unfavorably in scientific publications against e.g. KOHA – a comparison which now needs to be fully re-done with the arrival of ABCD ! - and became seen as a poor substitute for the 'real thing' when means or skills were non-existing or insufficient to allow use of professional softwares. Even from a simply technological point of view we think this is wrong judgement, 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 emphasized and focused on promotion of the software, according to its own specific mandate, in developing countries for very good reasons. Not only ISIS offers its professional capabilities without a need for time-and-again the very latest software skills, 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 PC's running ISIS.

Also users of ISIS in the Western world typically were 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 in this sense of being linked with economically less-powerful users. In Flanders (Belgium) e.g. typical ISIS-users were the the network of NGO on Development Co-operation but also in Cultural Work and – long ago because nowadays all incorporated into larger units with more economical power to purchase commercial solutions – smaller medical (and school) libraries.

The use of ISIS however in state-of-the-art systems like the ScIELO information system of BIREME – offering huge amounts of documents on medical and social information in Open Access full-text

electronic journals – proves that this option to focus on developing countries is not a technological weakness but rather an arbitrarily chosen ambition : to provide tools for joining the Information Society to everybody. This is indeed UNESCO's basic mandate when it comes to information, and that is also a 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 present ABCD as an integrated library system itself – we can refer to some other publications elaborating on this (see [reference]).

The basic statement however here is that – for the first time in the history of ISIS – a full-blown dedicated and integrated library system is provided based on ISIS-technology. Other attempts to offering such integrated library solutions based on ISIS, e.g. Open MarcoPolo and Weblis, were very meritorious efforts, but never (or not yet ?) reached to 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 Centers ('Centros de Documentación' etc. or simply 'CD').

What does this mean ? We think it is even the most unique and special feature of the software, so this needs some brief comment here.

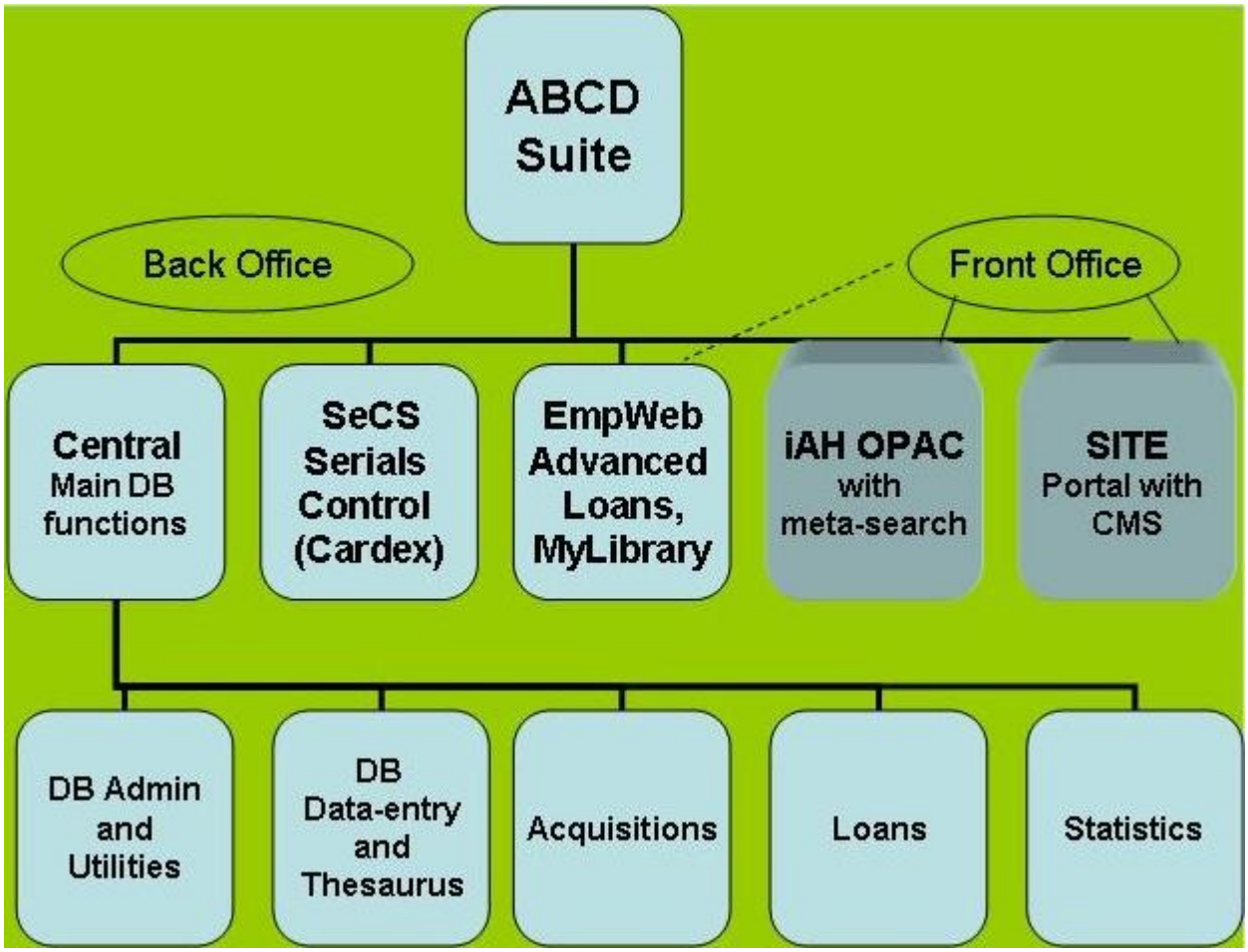
By being fully independent from the bibliographic structures used – but already offering some widely used standards like MARC (21 and UNI) or CEPAL (a widely used bibliographic format in Latin America) – ABCD allows *any* structure to be created by the software itself and subsequently being managed (cataloging, OPAC and circulation). This is indeed based on the fact that ABCD offers a full ISIS-interface in a web-environment (using PHP-programming) including all typical ISIS-definitions such as 'Field Definition Table', 'Field Selection Table' and of course the crucial Formatting Language.

This means that not only libraries, but equally documentation centers, archives or musea can use the software while sticking to their own dedicated information structures, or even better : organizations can use ABCD to integrate such related but still different functions into one web-based system. By being fully web-based, ABCD of course can also offer multi-media capability (important e.g. for musea) but also offer 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 again offers something extra to the typical ISIS-users, taking away the need to develop skills (and/or maintenance resources) for running different systems for different purposes with a common 'information-service' functionality. They often cannot even afford to develop such different support mechanisms...

Still ABCD is presented as a software incorporating all major and important current standards in information services. When the Flemish Interuniversity Council's Development Co-operation section (VLIR/UOS, see http: //www.vliruos.be) decide to support the ABCD-development throught its project 'DOCBIBLAS' (Development Of and Capacity Building in ISIS-Based Library Automation Systems) in 2007-2010, such professional criteria allowing usage by university libraries, were explicitly required : 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 webservices.

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



In this scheme the 'Central' module contains the 'heart' of the software with the main crucial library-automation functions, allowing smaller libraries with less technology skills to fully automate without leaving the familiar (as most ABCD-users are expected to be the current ISIS-users) ISIS environment 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 widely-used standards is still easily within reach.

Conclusion

With the new-born ISIS-family member 'ABCD' an old long-standing tradition of ISIS-developments is continued and fully brought into the modern database-driven web-applications. Combining the old but proven-technology concepts of the historical ISIS with both web-technology and integrated library and documentation centers automation functions, ABCD brings a hopefully welcome solution and

continuation for the many existing ISIS-users all over the world, but maybe even 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 globalized users market more dependent rather than more self-supporting and self-empowered.