Informatics for Meeting the Challenges to Public Health

Editorial

ABCD: Sustaining and Extending the Capabilities of ISIS Software by Applications
Egbert de Smet and Ernesto Spinak

Dynamics of Fisheries Information Transfer in India
M Jayapradeep

Information-Seeking Behavior of Oncology Professionals
Dr. V. Vishnu

Ritu Gupta and B.M. Gupta

Cover: Devapalan

Centre for Informatics Research and Development

Informatics Studies
Gayathri, A4, Kundappakunnu, Trivandrum – 695043
URL: informaticsstudies.org, E-Mail: editor@informaticsstudies.org
ABCD: Sustaining and Extending the Capabilities of ISIS Software by Applications

Egbert de Smet and Ernesto Spinak

Abstract

The paper presents the aims, objectives and capabilities of ABCD, the last in the ISIS-Software family meant for automation of libraries and documentation centres. ABCD software repackages different modules of BIREME documentation technology, as a suite. It is a DBMS with capabilities to manage all database functions in an information handling scenario similar to what (Win)ISIS did on earlier computing systems. It provides for definition of database structures, indexing, searching and formatting output and covers library housekeeping operations like acquisition and circulation. So there is no need for existing ISIS-users to migrate to non-ISIS solutions. Being ‘free’ and open source, ABCD is expected to become the most used package in the knowledge management scenario.

The paper also discuss major functional upgrades in ABCD 2.0 facilitating use of numerous features of the underlying CISIS-technology, the power of ABCD as a digital library solution to retrieve specific piece of information instead of a document, and its UNICODE (UTF-8) compliance that will enable use in non-Latin alphabet environments such as Arab, Sanskrit, Amharic, Persian etc. The severe challenges that projects like ABCD which work for the target group of small organisations in very-low-resources environments may face; both in finances and skills are also outlined.

Introduction

ABCD is a package in the line of ISIS family of software meant for automation of libraries and documentation centres. It was released in 2009. During its three years of existence it has been published in 4 languages (English, Spanish, Portuguese and French). It continues the tradition of the popular WinISIS software, which has not been scaled up for the recent versions of operating systems. In the new environment, ABCD becomes the main solution for ISIS-software users worldwide especially for those in the developing countries. ABCD software is being re-packaged as a suite comprising different modules of BIREME documentation technology. ABCD provides possibilities to create and manage all database functions as in (Win) ISIS in the modern WWW-environment. It provides for database definition, indexing, searching and formatting options. In addition, it can cover library housekeeping operations like acquisition and circulation. So there is no need for existing ISIS-users to migrate to non-ISIS solutions in order to cope up with fast changing technologies and solutions. Being ‘free’ and ‘Open Source’, ABCD is expected to become the most used package like ISIS in the library and knowledge
management scenario, although it will be
difficult to assess the popularity. But
comparative studies show that with the
leading FOSS library solution KOHA,
ABCD scored favourably especially when
flexibility is considered. Not only huge
university libraries but also consortia have
started to opt for ABCD. In a survey among
African institutes ABCD came second to
KOHA.

**ISIS Software Family**

The ISIS family of information storage and
retrieval software, since its creation under the
name of ‘Integrated Set of Information
Subroutines’ by International Labour
Organization (ILO) in 1960, has evolved in
the last 30-40 years through different
implementations, under different operating
systems, programming languages and
functionalities, and in numerous
organizations like BIRME. Initially it was
known as Micro ISIS running on a
mainframe computer -IBM 360. In 1975
UNESCO rewrote it as CDS/ISIS to develop
a database of UNESCO publications and to
impart regular training on computerized
information retrieval systems. Later
International Development Research Centre
(IDRC) incorporated modifications to Micro
ISIS using it with minicomputers (HP 3000)
and the modified version was called
MINISIS. In 1986 UNESCO issued an ISIS
version for use on Microcomputers that is
IBM PC compatibles. Members of all ISS
software family are functionally compatible;
which means they all have been designed as
generalized information management
systems with which similar operations can
be performed and data interchanged (Raman

The ISIS implementations like OIT IBM
DOS, IDRC HP MINISIS, UNESCO CDS
ISIS OS/CICS, UNESCO CDS/ISIS Mini-
micro, UNESCO WINISIS and BIREME
CISIS, WWWISIS and ISISDLL have had
historical importance and were distributed
worldwide. BIREME has cooperated with
UNESCO in the development of the
WINISIS – the Windows version, whose
primitives are CISIS based. The development
of the Windows version was the result of
strategic development policy decision aiming
at maintaining it on the new operating system
environment with the leading role CDS/
ISIS has played in international scene. CDS/
ISIS Mini-Micro version is the standalone
implementation highly used worldwide
while WWWISIS remains very popular for
web-based applications. Based on BIRME’s
WWWISIS, IBISCUS Association from
France has developed for UNESCO the
package Gen ISIS (Rajasekharan 2010). Gen
ISIS is an authoring software (for Win 32)
for visually producing web forms to query
CDS/ISIS databases. Two versions of the
tool were developed - GenISIS Web for Web
publishing and GenISIS CD for developing
CDROM interfaces for CDSISIS databases.
They were written in Microsoft Visual Basic
by Pierre Chabert. Its source code was freely
available. Gen ISIS implemented the web
query form, display of query results and
display of particular records with details. It
required the use of a local web server. It was
possible to export the generated application
to real Internet/Intranet server (Rajasekharan
2012).

The Founding Father : Mr. Giampaolo
Del Bigio

The ISIS family has brought remarkable
contribution worldwide over the years to the
operation of computer based information
storage and retrieval systems (ISRS)
particularly for developing countries. It is
probably one of the most important
international contributions to the early digital
inclusion of an entire cultural and scientific
sector, in this case, libraries and information
centers around the world. ISIS has been an
efficient solution to empower those
organizations to be able to streamline their
information processing activities by using
modern and inexpensive technologies.

The ISIS family of software evolved mainly
due to institutional commitments. But Mr. Giampaolo Del Bigio has dedicated his life to ISIS. His commitment to the dissemination of ISIS technology around the world, as an UNESCO international servant has no parallel. In the same vein the contribution of BIREME for sustaining the ISIS family of software is also unique.

**BIREME’s Role in Developing and Sustaining ISIS Solutions**

BIREME is the Latin American and Caribbean Center on Health Sciences Information, a specialized center of PAHO/WHO, located in São Paulo, Brazil. It contributes to the development of education, research and health in Latin America and the Caribbean through the democratization of access, publication and use of information, knowledge and scientific evidence. To achieve its goals BIREME has developed numerous systems and solutions for documentation and information dissemination. BRME is the organization, which has contributed to the worldwide use of ISIS software solutions launched by UNESCO.

Initially it led the dissemination of ISIS in Latin American and Caribbean, and with a team of system analysts and programmers led by Abel Packer, Adalberto Tardelli and others developed new implementations of the software, pioneered the use of ISIS in CD-ROM and implemented the most complex online applications which can handle high volume of records and number of access. Some of the applications developed by BIRME are discussed below based on information available in ABCD site (Bibliotheca 2013a).

**CISIS in RedDes:** It is a library of functions developed by BIREME in C language to allow the manipulation of ISIS databases without using or installing CDS/ISIS - MicroISIS/WinISIS (UNESCO).

**CISIS Utilities:** It consists of a set of executable programs developed with CISIS library that extends and makes it significantly easier to generate, manipulate and transform data within ISIS database model, empowering it with functions not yet available in CDS/ISIS.

**iAH - Interface for Access to Health Information:** It was designed for information retrieval from ISIS databases over the Internet. It is written in the IsisScript language using WWWISIS engine, which gives the IsisScript code multi-user access to the ISIS databases across the Common Gateway Interface (CGI).

**iAHx - Integrated Research System VHL:** is integrated Virtual Health Library Developed by BIREME using Lucene Solr in order to perfect the mechanism of presentation of search results of the Virtual Health Library and its collection of information sources, enabling to view them in an integrated, individualized, typed and sorted sequences by various criteria.

**Continuity of ISIS Software**

ABCD is the most important and latest project of BIREME intended to support the documentation/knowledge management systems worldwide.

ABCD stands for ‘**Automatización de Bibliotecas y Centros de Documentación**’ (Spanish), which means: **Library and Documentation Centers Automation**. Its development is promoted and coordinated by BIREME, with the support of Flemish Interuniversity Council (VLIR), which is the agency of the Belgian Government for university cooperation for development (VLIR).

**ABCD v1**

ABCD system was first officially released on December 3rd 2009 at a BIREME-Symposium in Sao Paulo, Brazil (de Smet, 2010a and 2010b). Soon after its initial release it became clear, mostly from the ‘ISIS discussion list’, still the main channel of communication of the large ‘community of ISIS users’, that ABCD would quickly replace
the popular WinISIS which in turn replaced the Micro CDS/ISIS software as the dominant application of still growing and alive ISIS-family (de Smet, 2009).

There are many reasons for Isis users turning to ABCD. It is a solution for managing library system with e.g. acquisition and circulation modules fully prepared. WinISIS and CDS/ISIS for DOS provided only the environment in which such functions as well as numerous other applications could be developed. Many organizations including BIRME had done that earlier. But in addition to all that being a ‘textual database’ software for general purposes ABCD still preserves all the ‘ISIS-philosophy’.

But ABCD is a web-based integrated library management software comprising the main basic library functions. Such an application was a long felt need for the ISIS community, since the first MS-DOS version came out in 60s. Several library automation systems were developed since then and are still in use worldwide. BIRME also developed earlier a system which was limited to the circulation services. But the main characteristics of the present package are the coverage of the main library functions, its web centrality and its development and maintenance under the methodology of Free and Open Source Software.

Main Functions

- Definition of any number of new databases (similar to WinISIS), which includes: FDT, PFT, FST, and worksheets directly on the Web, or copying from existing ones either from the Web or from WinISIS on a local hard disk,
- Cataloguing of books and serials, independently of the format: MARC, LILACS, AGRIS, etc.
- End-user searching (OPAC),
- Loans circulation,
- Acquisitions,
- Statistics,
- Library services like SDI, barcode printing, quality control, etc.
- Compatible with CDS/ISIS database technology for the bibliographic databases, i.e. reading ISIS-databases and making use of ISIS Formatting Language for producing output and indexing of records;
- Run on both Windows and Linux platforms;
- Use of MARC-21 cataloguing formats and other current standards or protocols (Dublin Core, METS, Z39.50..);
- Published as Free and Open Source Software (FOSS) with the accompanying tools for the developer community;
- Multi-lingual.

ABCD is aligned with the CISIS/1660 version 5.2 platform, and will eventually be made compatible with later versions. This means that the inverted file entries are 60 characters long, and will increase in length in the ISIS-NBP based version. ABCD is compatible with programming languages accepted by the GNU licences, i.e. PHP, Java, Javascript, Python, etc. The current version of ABCD is written in PHP v.5 and IsisScript. The system is totally language independent. It is available in Spanish, English, French and Portuguese and can be translated into other languages like other CDS/ISIS applications.

Smart Header

In the history of evolution of database management systems ABCD was the earliest to promote a ‘No-SQL’ approach, foregoing the strict ACID-requirements of relational databases. But it also provides more flexibility and scalability in No-SQL database idea, mainly based on the very simple ‘key-value’ concept of storing objects (like documents) in a ‘flat table’ with an identifier (key) for fast retrieval. The structure of the ‘value’ (or stored object) is left open to the application in this philosophy. In the case
of ISIS also there was a No-SQL ‘avant-la-lettre’ since it was developed in the mid-seventies (!). That structure was and still is largely based on the ISO2709 format for bibliographic information, which is actually the basis of the MARC-standard(s). This ISO-format uses a smart ‘header’ which describes the exact structure with which fields exist and the length they have, for each individual record. So each record is a ‘document’ and can indeed contain information of any structure without imposing the fixed-field structure of relational tables. In addition, by analysing this header having only numerical information - something that the computers are particularly good at - the whole document or record can be understood. It allows detailed formatting and processing without really parsing the whole document as is necessary in the case of XML or HTML files (de Smet 2005).

This approach of being a ‘flat file’ has the important advantage of being simple as far as creating a structure is concerned, unlike relational structures. This in turn allows the software to offer structure-creating functions and that is what makes ISIS, and also ABCD, ‘general purpose’ packages.

Another strong feature from the early versions is the word-indexing facility. Each word in a field can be used individually, even from within the text. This means full-text indexing ‘avant-la-lettre’ again which has become so ubiquitous and apparent in the Google-era of the WWW we are now living in.

**WWW Compatibility**

The versatile ‘formatting language’ is another feature of ABCD, which is very important now. It was easy to make ISIS, WWW-compatible since HTML- (or XML-) tagging could be easily added as ‘literals’ into the formats.

Since ABCD is fully based on ISIS, it kept all the above features and is easily replacing WinISIS as the real ‘ISIS for WWW’ application. A new, not unimportant characteristic, which ABCD can claim, is that ISIS has changed itself from ‘free’ software to ‘Free and Open Source Software’ (de Smet, 2008).

**Library Modules**

Since 2010 many users of ISIS started experimenting with ABCD. This has put forward many challenges to ABCD. Beyond just extending the existing ISIS-concepts to the web it has to take care of numerous aspects.

ABCD is a suite of different modules. The ‘WinISIS’-like management module (Central) is accompanied by a separate but fixed-format Serials Control module using the ISSN standard. It has also a versatile CMS to create ABCD-sites. The Site of the BIREME-lead ‘Virtual Health Library’ network, is based on this solution (de Smet 2012c). There is also an advanced search interface ‘iAH’. While Acquisition, Circulation and Statistics are three other modules within the Central module, a more advanced module for loans management (EmpWeb) has also been added to the suite. It adds more complexity with different technologies like Java, Jetty and a relational database for the management of the transactions and of users if so desired (de Smet 2012a). A detailed description of the modules and functions are dealt in an earlier study by Egbert de Smet (de Smet 2010b).

**Features and Modules of ABCD**

Descriptions of the basic modules of ABCD provided in ABCD web site (Bibliothecca 2012) are the following:

Like any similar advanced and complicated software ABCD was forced to release some smaller updates (and ‘patches’), combined in an upgrade v1.1 and 1.2b over the years 2011 and 2012 respectively. Presently there is a version 1.2t (transitional), which adds many new but smaller features and at the same
<table>
<thead>
<tr>
<th>Administrator</th>
<th>Management of User accounts</th>
<th>Full database creation process</th>
<th>Backup/Restore</th>
<th>Inverted file generation/update</th>
<th>Database reinitialization</th>
<th>Record/database unlocking</th>
<th>Maintenance of Messages database and translation to other languages</th>
<th>Creation of Help-in-context pages</th>
<th>Management of z39.50 (protocol for searching and retrieving information from remote databases)</th>
<th>Management of OAI (the Open Archives Initiative Protocol for Metadata Harvesting)</th>
<th>Quality and consistency controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical processes</td>
<td>Data entry</td>
<td>Creation/editing of worksheets</td>
<td>Creation/editing of formats for quality control</td>
<td>Linking to Authority files</td>
<td>Management of the Acquisitions module</td>
<td>Import/Export (Isis, MARC, XML, etc.)</td>
<td>Global changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics</td>
<td>Pre-defined basic package with the most frequently used statistics</td>
<td>Ad-hoc reports generator, with direct output to graphs and export to Excel-like spreadsheets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Services</td>
<td>Print generation (on demand or scheduled)</td>
<td>Printing of barcodes and spine labels</td>
<td>Selective dissemination of information through profiles</td>
<td>Online services by demand (photocopies, reference, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans/Circulation</td>
<td>Databases for identification of loan objects and users</td>
<td>Definition of service policies or rules</td>
<td>Multiple display formats</td>
<td>Pre-defined calendar for calculating suspensions, reservations and loans</td>
<td>Assigning of different functions to different operators</td>
<td>Definition of lending slips and other prints</td>
<td>Detection of inconsistencies in transactions</td>
<td>Statistics on all types of transactions</td>
<td>Loans</td>
<td>Devolutions</td>
<td>Reservations</td>
</tr>
<tr>
<td>OPAC(iAH)</td>
<td>Simple editor for ABCD-Site administration, modelled on BVS SiteEditor</td>
<td>for parameter definition (.def files)</td>
<td>Editor for display formats</td>
<td>Editor for indexing techniques (.FST)</td>
<td>Web service connecting the OPAC with the loans system</td>
<td>Editor for shortcuts (calls to external applets through .pft)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KardexPeriodical Control(SeCS)</td>
<td>Title creation according to the ISSN standard</td>
<td>Import of titles from ISO MARC/ISSN formats or native SeCS</td>
<td>Creation of frequency masks for automatic generation of new issues</td>
<td>Generation of collective catalogues with multiple collections</td>
<td>Entry of new issues, manually or automatically through barcode reading</td>
<td>Export of full or selective collections</td>
<td>Control listings of missing or duplicated issues</td>
<td>Hypertext navigation between different versions of families of titles</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
time prepares the transition to v2.0, which will be discussed later in this paper.

**Latin America Africa and Asia**

The ‘classical territory’ of ISIS being Latin America, the use of ABCD is without doubt more visible there. The Cuban Ministry of Higher Education has selected ABCD for all universities in the country. ABCD has become the 2nd best-known FOSS for library automation in a survey on FOSS-library systems in Africa. In Africa ABCD is used in university libraries of Mozambique, Zimbabwe, Tanzania, Kenya and Ethiopia. The ‘Open University’ network of Zimbabwe also uses ABCD. The Kenyan Agricultural Research Institute recently selected ABCD after a thorough comparison of all available packages from the FOSS scene like Weblis and KOHA. The most important positive factor is again the flexibility of the package. ABCD can be used not only to automate their library but also for many documentation functions and for developing digital archives or museums. In many Asian countries ABCD is now generating interest among earlier ISIS software users who were forced to change over to different packages to come back to their beloved stream of solutions.

In Asia the use of ISIS has declined significantly over the last decade, partly due to the fact that more alternatives in FOSS became available (e.g. KOHA). But it was also due to the lamentable neglect until recently of accepting Unicode as a standard to deal with non-Latin scripts. E.g. in India there is still a healthy interest for ISIS but if the package is to take care of the numerous languages and scripts of India Unicode compatibility is a must. Hopefully the Unicode feature is now becoming available with ISIS and ABCD (see infra) and so we can expect that in India where thousands of rural libraries, that came up and are sustained by local initiatives with meagre funds exist, interest in a package like ABCD can go up again.

The slow withdrawal of the support to programmes like ISIS by UNESCO - due to limited funding and also new policies regarding software development, which is now ‘generic’ for all FOSS in education has really affected documentation projects worldwide. The ISIS-software was the one, which suffered a major decline in development and dynamics due to this policy change. Many specialized information systems of great value to humanity have been developed in numerous parts of the world using ISIS. But still with the UNESCO-born J-ISIS (Kenai 2013) and the ABCD now there is new life to projects using ISIS solutions. Both software, ISIS and ABCD support Unicode. Hence their usage will be very high outside the Western world.

But now the ‘visibility’ of ABCD is not that much high. It is not observed or mentioned in the survey by Marshall Breeding, which is continuously updated in the web (Breeding 2013). It is mainly due to the fact that many users - as has always been the case for ISIS solutions (de Smet 2005) - are smaller institutes with poor connectivity and lower professional profile, operating mostly in a silent way and under the radar of the big international scenario.

In the meantime ABCD has got its first ‘clone’ or ‘offspring’ in ‘SuiteSaber’, a software produced and now used in Brazil. In fact it is mostly an exercise of adjusting style-sheets for the presentation of the interface, but preserving most of the underlying codes.

**ABCD as a Document Manager**

One concept already mentioned above is the feature of allowing different textual databases to be created in ABCD on top of the books catalogue(s). In that way ABCD is not only a reasonably advanced ‘library management system’ like KOHA, but a software that goes further in scope at a deeper level by including more Web 2.0 features like tagging, showing book covers etc. But in fact it is also a ‘general
purpose’ document manager. It can as well serve a centre for documentation (what the letters CD’ in the ABCD - denotes). It can be used to develop archives and museums. For archives the new ABCD version will offer an implementation of the ISAD/G-standard.

An example of such approach is illustrated below from an application in a medical documents system in Uruguay.

**Migration to ABCD**

The Medical Union of Uruguay (Sindicato Médico del Uruguay - SMU) is the professional association that brings together more than 20,000 doctors in the country. The SMU managed administrative files from the late 90's using a relational software, called Docupact, which left the market long back. This software, besides being downright obsolete, has very poor indexing mechanisms and had no quality control for data entry. Therefore SMU decided to migrate to ABCD. But at the time of migration (November 2012) there were nearly 90,000 digitized documents in TIF and PDF spanning over more than 11 years.

The migration process took about two months of work where they could recover and convert 100% of valid records with their associated digital files. The migration and conversion was done using the CISIS utilities for BIREME on which ABCD itself runs. This is an important experiment and it proved that system managers of ABCD with knowledge of the (command-line driven) CISIS-tools have an endless series of possibilities for very powerful database-operations and tricks at their disposal.

**Medical Documentation System**

The Archive and Documentation Section of SMU manages more than 10 different types of documents corresponding to the different Sections of SMU.

- Acts of the SMU General Secretariat, which meets on a weekly basis, where each meeting discus and resolve dozens of issues pertaining to the medical corporation.
- Resolutions of Specialized Medical Boards
- Proceedings of Arbitration Committee on Ethics
- Files which are personal files of associating members
- Minutes of General, Ordinary or Extraordinary Meetings
- Accounting department paperwork: daily cash books, bank statements, bill payments, collection receipts, payments of taxes on personal income of physicians
- Miscellaneous documents, such as letters, memos, contracts

The system includes several controlled vocabularies, on the medical and administrative issues, authority files on personal and institutional names, etc. The ‘quality control’ is mainly based on these, but also post-processing of the records is done with quality control scripts based on the powerful CISIS-utilities. Among database experts, ISIS is known as a database technology. This is very true when it comes to the enormous power that database-managers get while using the command-line tools. But ABCD is much more than its utility.

We provide below screen shots from an example - the Minutes of the General Secretariat of SMU with links to the main and supporting documents presented to the weekly meeting.

The system operating in the background is using ABCD in the Archive and Documentation Section, and all documents can be consulted via Intranet throughout the institution/network using the interface iAH.

**ABCD Version 2.0 : New Features**

By the end of 2013 the ABCD version 1.2 is expected to be replaced by a real full-blown
ABCDD: Sustaining and Extending the Capabilities of ISIS Software by Applications

Figure - 1: Worksheet for Data Entry

Figure - 2: Presentation in ABCD
new release version 2 of ABCD with major upgrades.

The major step (from 1 to 2 instead of 1.2 to 1.3) is justified by the introduction of some more basic innovations, on top of many smaller improvements already present in the actual transitional version such as Unicode and Digital Library.

Before discussing those major innovations we need to review briefly some relevant changes or additions of the transitional version currently available (version 1.2):

- Compatibility with the newer versions of Apache (up to Apache 2.4) and PHP (up to 5.4)
- Configuration: several basic folders for databases can be defined and selected from the login screen on Central; this allows more versatile set-ups e.g. with test databases, different versions of the same databases etc.
- The http-protocol 'get' (to communicate between the client and the server) is changed to 'post', allowing much longer strings to be included into the URL's (which was necessary for Digital Library applications)
- Additional formats : ISAD/G, Unimarc
- Small additional features in data-entry like a counter for fixed-length fields, a print format and pick list editor button in the cataloguing module (to allow more practical experimenting and ‘previewing’ of changes in display formats or editing authority control files)
- Simplified MARC demo worksheets by request from many librarians who deem the full MARC-implementation too complicated and unnecessary
- Loaning of multiple items at once, and from copy information not only in a dedicated ‘loan objects’ database but also from within the catalogue - more practical for smaller libraries.
- Upgrade of the BVS Site, incorporated into ABCD, from v4 to v5, and re-styled Site-CMS
- Features, e.g. comparing the message tables for different languages of the interface, easier creation of links for new searches in display-formats etc
- Compatibility with PHP 5.3 or higher, Change all calls to IsisScript by direct

Figure - 3: OPAC Using IAH for Retrieval
ABCD: Sustaining and Extending the Capabilities of ISIS Software by Applications

Figure - 4: Search Result Using LAH OPAC
CISIS-calls, Thesaurus management module, New OAI module (replacing the old one which had problems), Add a ‘traffic-control’ function for the Update Inverted File function to avoid locks, replacement of ABCD Site v.4.1 with version BVS Site 5.1.11, Add SDI function, Updates for the help on documentation and translations, Update SeCS-Web and finish documentation, Mirror sites and servers for download and distribution of ABCD and EmpWeb - updates

Unicode Compatibility

Unicode is the standard for extension of the old ASCII-table (which was limited to 256 possibilities to represent text-characters into binary values, e.g. A=65) to a much bigger memory space for holding 65Kb of values, in principle allowing all known alphabets to be incorporated.

The introduction of Unicode means that the basic mechanisms of a database storage and retrieval engine no longer can assume that all characters are single-byte (8 bits) as was the case with ASCII, but also multi-byte characters can be correctly dealt with. E.g. for Amharic the decimal values take 3 bytes as is the case for Devanagari (Indian) scripts. Until now ISIS did not allow this. UNESCO not developing ISIS in this aspect is in fact against UNESCO mandate - to serve the poorer countries and cultures of the developing world, mostly in the South and East where non-Latin scripts are used.

Before Unicode in ABCD non-Latin characters could only be used with their HTML-encoding, e.g. ‘&1234;’. While this worked to some extent it is far from ideal, e.g. 6 ASCII-characters are used to represent in fact only one character. Indexing and searching then was performed on these HTML-coded sequences, which is not the right way to do things.

This feature should not be mixed up with the language of the interface or even the script used for that interface. An Amharic interface e.g. exists in Ethiopia but is not based on the Unicode standard but on HTML-encoding of non-Latin characters.

As part of an effort, again supported (as was initially the case for the development of ABCD) by the Flemish Interuniversity Council ‘VLIR’, in the framework of ABCD a special version of CISIS, being the underlying core of ABCD, was developed. BIREME had already expanded into several versions of CISIS over the last few years, in order to cope with new demands and technologies, e.g. versions allowing longer indexing keys (60 instead of 30, now also 256 being available), larger records (now up to 1Mb instead of the typical 32Kb of WinISIS) and larger databases (up from 512Mb to 4Gb). E.g. the ‘FFI’-version of CISIS allows records up to 1Mb and 4Gb databases. With the aim to provide faster indexing for full-text databases also the indexing engine of ISIS got alternatives, dropping word-positioning data (rarely used in modern text-retrieval) in the ‘postings’ of the Inverted File for (much) higher performance, i.e. speed. This is the ‘Lind’ version of CISIS.

There are some good technical reasons why one version could not just simply replace another version but rather co-existent versions of CISIS (30/60 index keys, FFI, Lind...) remain available. Each version is optimized for specific purposes, e.g. the classic ISIS is still a very powerful and efficient database engine for bibliographic records, while other versions, amongst them now also Unicode compliant versions, each have their specific use - and this specificity can be assigned to any database in ABCD.

For above reasons ABCD incorporates a new feature of allowing the CGI-executables to reside in sub-folders of the cgi-bin folder, each sub-folder containing a different version. For each database one can define which version is to be used. Since these executables have a very low ‘footprint’ (in the server's
memory and storage), e.g. the only one needed to run normal ABCD-tasks is ‘wxis’ with less than 500 Kb), one can easily combine several such versions. This is the basic new technical feature of ABCD which not only makes Unicode possible (Unicode being one of the available CISIS versions now) but also allowing the Digital Library feature - which had to be developed in 2011 but postponed during all of 2012 also until this technique was fully implemented and tested.

So, ABCD Unicode implies:

- A Unicode-subfolder of the cgi-bin (e.g. /ABCD/www/cgi—bin/uni) to contain the Unicode new version executables of CISIS
- An entry into the definition file ‘abcd.def’ (in the bases-folder) mentioning the name of the database and its CISIS-version, e.g. ‘MyArabDB=unicode’
- A new table (with new format) for the definition of ‘alphabetic’ characters and of uppercase translation. ABCD in the table below shows the beginning of such a Unicode table starting from the Latin alphabet ‘A’ but further on continuing with (initial sections of) the definition of resp. Cyrillic, Arab, Devanagari (for e.g. Hindi, Sanskrit, Marathi) and Amharic (Ethiopic).

<table>
<thead>
<tr>
<th>Code</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>065</td>
<td>A</td>
</tr>
<tr>
<td>066</td>
<td>B</td>
</tr>
<tr>
<td>067</td>
<td>C</td>
</tr>
<tr>
<td>208 128</td>
<td>CYRILLIC CAPITAL LETTER IE WITH GRAVE</td>
</tr>
<tr>
<td>216 128</td>
<td>ARABIC NUMBER SIGN</td>
</tr>
<tr>
<td>216 129</td>
<td>ARABIC SIGN SANAH</td>
</tr>
<tr>
<td>216 130</td>
<td>ARABIC FOOTNOTE MARKER</td>
</tr>
<tr>
<td>216 131</td>
<td>ARABIC SIGN SAFHA</td>
</tr>
<tr>
<td>224 164 132</td>
<td>DEVANAGARI LETTER SHORT A</td>
</tr>
</tbody>
</table>

Likewise the uppercase translation is defined in the file ISISUCTAB.TAB:

```
097=065 # a -> A
225 136 128=225 136 128 # ETHIOPIC SYLLABLE HA
(no substitution in Amharic or Arab)
```

Some sample screenshots illustrating the use of non-Latin scripts in ABCD are provided below (Figures 5-7). The first two shows records with multiple different non-Latin scripts (with meaningless contents), the third shows the mixed-script index listing in ABCD.

**Digital Library Solution**

Expectations towards library services have changed a lot in the last two decades, mostly due to the success of the WWW and the Internet in more general, which made people familiar with electronic documents, and information of all sorts. Now libraries deal not only with printed materials. They have to develop and operate ‘digital libraries’ manage and avail electronic documents, which may be local ‘repositories’ of theses and research papers or real collections of e-books and other documents in electronic formats. More and more old printed materials are also getting ‘digitized’ usually scanned and processed with OCR in-house to make them suitable for electronic access and preservation.

Many libraries, also in the South, currently are busy setting up ‘digital libraries’ with software such as Greenstone or DSpace. For smaller libraries this often means having to create capacity within its own staff on such new additional software in addition to running the ‘classical’ library system. By
Figure - 5: Multiple Non Latin Scripts in ABCD

Figure - 6: Non Latin Scripts in ABCD

Figure - 7: Mixed-Script Index Listing in ABCD
including a ‘digital library’ feature into ABCD it expects to serve such smaller organisations with the ‘all-in’ solution ABCD wants to be.

Of course the ambitions of ABCD as a digital library system are limited, but basically what most librarians want to do is to store their electronic documents (whether scanned or ‘digital-born’) on their computers and make them available to end-users through WWW-technology, allowing access e.g. also from homes or workplaces, making the library even more ‘virtual’.

The ‘Digital Library’ feature of ABCD encompasses the following functions:

- Fields can be characterized as either ‘URL’ or ‘Doc’ fields (in the data-entry type list of field-types)
- When such fields exist in a database, an extra icon for ‘PDF’ or ‘document upload’ will be shown in both toolbars (the main one and in the record-toolbar) of the Central module.
- Clicking on that icon serves a file-selection window; the selected file will be uploaded in the folder of the database dedicated to store ‘attachments’ to records (e.g. pictures but now also PDF and other documents)
- All (but only) text-contents of the document will be extracted by the powerful Tika-tool of Apache Software Foundation. For Windows also a free tool ‘pdftohtml.exe’ can be used with much smaller footprint
- The text will be loaded in the ‘document’ field defined for that database and the link to the uploaded original will be put into the URL-field with hyperlink formatting, allowing immediate and direct access to the full text in the original document format.
- After indexing the record(s) all words can be searched using the normal search interfaces in ABCD Central and iAH.

But when we consider the present requirements of documentation centres the above facilities are not sufficient to make ABCD a perfect tool for digital libraries. But it can go to a good extent to fulfil the requirements. The limitation is in the maximum record size of the new CISIS-versions, i.e. 1Mb, which means that some very large textual documents will need to be split over several records. But tests we did on many documents showed that after text-extraction the file-size rarely exceeds 1Mb. The Digital Library feature in ABCD is presented with some screenshots below.

Figure - 8: DigLib Record in ABCD Central

In Informatics Studies 3(1), January – March, 2016
Figure - 9: Full-text Index in ABCD

Figure - 10: The End-User Search (iAH) in a DL Database
The Figure – 8 shows a DigLib record (partly) in the ABCD management interface (Central) as the librarians will see it after importing the document (which will typically take just a few seconds).

The screen shot given below (Figure – 9) shows a segment of a full-text index for searching in the librarians interface.

In the Figure- 9 the end-user search (iAH) in a Digital Library database, followed by another one displaying the part of the results is presented.

Sample screenshot showing the search results is provided in Figure – 11.

**The Limitations of ABCD**

The limitations of ABCD as a Digital Library tool can be summarized as follows:

- Very large documents cannot be loaded in one record and would need to be split over several records (which can be linked however)
- No highlighting of search words is available (yet)
- Since the ‘large records’ versions of CISIS (Lind and FFI) do not allow incremental indexing, the librarian has to index the database with a script after adding new records, but indexing is fast.
- In the coming ABCD version 2 also, editing records with larger texts inside is not possible yet due to memory limitations of the IsisScript language, which hopefully can be extended later on.
- More tests needs to be conducted before
finalizing ABCD ver 2 with larger sets of documents (which is a resource and time-intensive task) to see how performance evolves. This is e.g. a known problem of GSDL (version 2), which does not use a database to store the full texts.

It is also to be mentioned here that the ‘digital library concept’ as implemented in ABCD is something more than a ‘collection’ of electronic documents. Because of the versatility of the underlying ISIS-technology and the integrated ‘meta-search’ approach of the ABCD Site, the library can offer its bibliographic records (the catalogue of the books and other printed materials) along with many other databases of electronic information: scientific abstract databases, the local repository, an expert-database, an events-database and a real ‘digital library’ with e.g. the institute’s scientific papers. This way the digital library becomes, like a real library, a resource for many different types of information, but now in an electronic digital format (de Smet, 2012b).

**Challenges for ABCD**

Does the future of ABCD look bright with this strong list of features and innovations? It depends. Unfortunately there are factors, which we cannot keep under control. The paper has already mentioned the new ‘restricted’ role of UNESCO. Also BIREME (as a WHO branch centre) suffers from the same type of problems of budget cuts within the UN system and needs to focus more and more on its core business of ‘document delivery’ - now mostly in digital format - rather than software development.

The ABCD remains as an effort by relatively very few people - a small team of programmers in BIREME with a major input from e.g. G. Ascencio (Venezuela) and the authors of this paper – who make it possible by relatively small VLIR-funding. But now some larger institutions using ABCD (e.g. Catholic University in Santiago de Chili) and the Cuban ‘Universidad de las Ciencias de Informatica’ (UCI) also promised resource support and free programming input by their staff and students. Especially this last type of input is most welcome but has proved to be very difficult to channelize. Many promises were made by enthusiastic computer scientists but they have not yet materialized into real development input. The ISIS community is still large and therefore many volunteers are available for testing and adding end-user tuning (e.g. display formats, worksheets, translation) but few such volunteers with coding and programming skills are showing up. Currently ABCD is seeking support from the Cuban Computer Science University.

Even with a limited development team it has become clear that co-ordinating an Open Source software development in a somehow ‘anarchistic’ environment of the often isolated ISIS-experts is quite a big challenge. ABCD needs a better-organized software management both for development and distribution, using better tools for such purposes (SourceForge, GIT, Symphony) but then again it requires time, know-how and support by people. For the users ABCD needs to be developed with services like a Wiki, FAQ’s, online documentation etc. After the release of v2.0 ABCD needs to focus on the type of infrastructure.

ABCD 2.0 will run under new versions of PHP. As discussed above, it will facilitate use of various features that the implementation of the underlying CISIS-technology will make it possible. In addition to the standard 32K records with 60-character index keys now it can be used dynamically for larger record sizes (up to 1 Mb), larger maximum database sizes (up to 32Gb) and longer indexing keys (up to 256 characters). This will allow using ABCD for digital library management with the document texts loaded automatically inside the database. This will enable indexing full-text and retrieval of the specific page of section...
containing required information instead of retrieving e-document. UNICODE compatibility (allowing multi-byte characters in the UTF-8 standard) is another feature, which will have major impact on the use of the system in non Latin alphabet environments such as Arab, Sanskrit, Amharic, Persian etc.

The main challenges of ABCD, which are typical for FOSS are: keeping the whole effort co-ordinated and moving on without large funding. In case of ABCD the challenges are severe as it works mostly for the ISIS-target group of small organisations in very-low-resource environments (both in finances and skills). What the group behind ABCD is undertaking is a very important task meant for conserving and disseminating world's knowledge heritage, major part of which is held even now locally by communities, which cannot afford costly technologies to record, document or conserve what they possess. So a project like ABCD needs to be fully backed up by International Organizations like UN or UNESCO.

**UNESCO’s Mandate and ABCD**

The power of software system like ABCD for knowledge management and dissemination should reach the resource-poor communities in every corner of the globe. It needs a backing by some International organization. The apt machinery is UNESCO which is known as the ‘intellectual’ agency of the United Nations. It has supported the development of the earlier solutions of the ISIS family. Especially when ABCD is Unicode-compatible it would definitely support UNESCO’s mandate, better than ever before, with other ISIS-solutions. ABCD has achieved all the objectives of the ISIS related projects which continued development during the last half century under UNESCO and other UN-organizations like BIREME.

UNESCO was created in order to respond to the firm belief of nations, that peace must be established on the basis of humanity’s moral and intellectual solidarity. So it strives to build networks among nations that enable such solidarity, by building intercultural understanding; through protection of heritage and support for cultural diversity. The future of nations depends not only on their economic capital or natural resources, but on their collective ability to understand and anticipate changes in the environment - through education, scientific research and the sharing of knowledge (UNESCO 2012). In a connected world - with the emergence of the creative economy and knowledge societies, along with the dominance of the Internet and WWW, the full participation of the whole population in the new global public space is a prerequisite for peace and development. They have to rely on the power of intelligence to innovate, expand their horizons and sustain the hope of a new humanism. UNESCO exists to bring this creative intelligence to life; for it is in the minds of men and women that the defence of peace and the conditions for sustainable development must be built (UNESCO 2012).

ABCD is a package that can infiltrate into the majority of populations of the world which should be considered economically poor groups, to give them the power to conserve and distribute their knowledge heritage as well as access knowledge and information existing outside their communities at less cost. It can help to narrow down the digital divide.

So in the present networked world ABCD is an important tool UNESCO can promote to conserve knowledge heritage, and disseminate it for the sake of the world; to achieve its mission to contribute to the building of peace, the eradication of poverty, sustainable development and intercultural dialogue through education, the sciences, culture, communication and information. So it will be a great service to humanity if
UNESCO takes up the full commitment to further develop and sustain ABCD project.

If this paper can generate some interest on ABCD among the programmers who believe in the unique ‘ISIS’-mission and -technologies, experts on FOSS-management and coding tools for complicated applications and decision makers of international organizations responsible to collect and conserve knowledge heritage, it may reach its goal.

References


de Smet, Egbert (2012a). Special features of the advanced loans module of the ABCD integrated library system. Program: Electronic Library and Information Systems, 45(3), 323 -332


Egbert de Smet: has PhD in Information Science. Worked for INASP, UK for the promotion of electronic information use and skills. In association with BIREME (WHO-Brazil) developed ABCD System. Conducted numerous training sessions on ABCD System in Southern hemisphere. Authored papers on history and philosophy of the ISIS-technology. Is presently with University of Antwerp, Belgium.

Ernesto Luis Spinak Fontan: took Masters in Information Society from Universidad Oberta de Catalunya, Barcelona. Has served as Computer Science Engineer; Professional Librarian, Engineer in Land Surveying and Professor of Library Science under Universidad de la República, Uruguay. Taught Computer Networks, Data Processing, and Statistics. Provided consultancy to and conducted numerous workshops on library automation and reengineering of information services in 19 countries.

Gupta, B M: was Emeritus Scientist in the National Institute of Science, Technology and Development Studies (NISTADS), New Delhi. Has Associateship in Documentation from INSDOC and PhD from Karnataka University. Is a recipient of the Full bright Fellowship and is an elected the Fellow of the Society for Information Science. Has more than 100 research papers, national and international conference presentations, technical reports and books. Edited ‘Handbook of Libraries, Archives & Information Centers in India’ (16 Volumes) and ‘South Asia Bibliography and Documentation’ (8 Volumes). Has been Guest Editor of special issues of ‘Scientometrics’ and ‘DESIDOC Journal of LIT’.

M Jayapradeep, M has MA in Political Science, MLISc and MPhil in Library and Information Science. He obtained PhD for his work on Fisheries Informatics from Karpagam University. He has served Indian Air Force, Biotechnology Information Centre (Government of India), Kerala Agricultural University, Kerala University of Fisheries and Ocean Sciences, and is presently Librarian, Sri Vellappalli Natesan College of Engineering, Mavelikara, Kerala. His areas of involvement are ICT applications, library automation, knowledge management and extension activities. He has published/presented numerous research papers in journals and conferences.

Ritu Gupta: has Masters in Library and Information science from Annamalai University, Chidambram and is presently working for her PhD at Sri Venkateshwara University, Meerut, India. She has numerous published research papers and seminar presentations to her credit. She is a school librarian and her areas of interest are information support to health care systems, library management, bibliometrics and scientometrics.

Vishnu, V: obtained Masters’ in Library and Information Science with First Rank and PhD in Medical Informatics from University of Kerala. Joined service at Kerala Agricultural University and is presently Head of Information Division, Regional Cancer Research Centre, Trivandrum. He was instrumental in developing Oncology Knowledge Resource Centre; the most sought after information source for medical research in India.