J-ISIS

the power of ISIS combined with the power of Java

15 May 2013

Egbert de Smet,
University of Antwerp
Overview

- Some ISIS-background information
- J-ISIS: basic technology
- J-ISIS demo’s:
  - Creation of a database and importing ISO-records
  - Editing features
  - Browsing features
  - Searching
  - Formatting and Groovy format exits
- Digital Library with J-ISIS
- Web-JISIS:
  - Technology
  - Demo: browsing/editing/searching records
Some ISIS background info (1)

- The ISIS-software family

A ‘family’ of softwares based on the same technology: FDT, FST, FL, QL
Some ISIS background info (2)

- **Crucial elements:**
  - **FDT:** fields with 0-n occurrences and variable length, subfields
  - **PFT:** the ‘formatting language’ is the heart of ISIS
    * Exact definition of what comes out, for presentation, sorting, indexing, exporting
    * Allows semi-relational (‘runtime’) setups
  - **FST:** exact definition of indexes using PFT
- **Limitations:** 32Kb/1Mb records, 512Mb/512Gb master-files
J-ISIS basic technology

- ISIS: basic technology preserved (FDT, PFT, FST) but with different storage engine (BdB)
- Berkeley DB: no-SQL (‘key-value’) FOSS database owned by Oracle, mostly embedded, with as limits:
  - $2^{31}$ records, fields and indexable-entries
  - 2Gb max. record- and field-size
  - no max. database size
- Lucene indexing: FOSS full-text indexer from ASF with ranking
- Java (server-client) with rich set of libraries and features (e.g. Tika, solr, IDE’s etc.)
- Web-JISIS: embedded Jetty, Struts2, AJAX for fast/high interactivity (e.g. Search-term suggestions)
- Embedded WWW-browser: JavaFX with HTML5, CSS, JavaScript, DOM
Installation and starting

- Installation: download package and unzip (same in Linux and OS/X), but needs Java pre-installed
- Starting: double-click in bin-folder
- Open connection (port 1111)
- Open database from list
J-ISIS database creation

- A new database can either be:
  - Created from scratch
  - Imported from existing ISIS-database
  - Imported from MARC ISO2709, MARCXML, MODS, etc.

- For each database one has to create/edit:
  - FDT
  - WKS(‘s)
  - FST
  - PFT(‘s)
J-ISIS import wizard

- Formats: ISO, MARCXML, MODS, CSV
- 2 options:
  - Only data-import ISO
  - Copy existing FDT and FST
- MARC import with leader-fields in 30xx
- Export: idem
J-ISIS create database-structure

- FDT = a fields-’repository’ defining each fields’
  - Tag (1-999)
  - Name,
  - Type
  - Repeatable or not
  - Subfields
  - Indicators (in case of MARC)
Worksheet definition

- Any (or all) field(s) of the FDT can be used in any sequence
- For each field can be defined: default value, help message, validation, picklist
Index definition: FST

Exact definition of extraction of database-values to act as input for search, sort, export (and validation).

Each definition defines ID, Name, Technique (0-9) and the PFT

Field Selection Table

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Technique</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 - lines extracted from the Pft</td>
<td>TRN=1v1^b</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0 - lines extracted from the Pft</td>
<td>EN=1v2</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0 - lines extracted from the Pft</td>
<td>mhl (IAU=1v100%)</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>0 - lines extracted from the Pft</td>
<td>mhl</td>
<td>TI=1v120</td>
</tr>
<tr>
<td>120</td>
<td>8 - Prefixed terms tech 4</td>
<td>'/TW=/, V120</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>8 - Prefixed terms tech 4</td>
<td>'/BI=/, V120</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0 - lines extracted from the Pft</td>
<td>(IAU=1v200%)</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>0 - lines extracted from the Pft</td>
<td>mhl</td>
<td>TI=1v220</td>
</tr>
</tbody>
</table>
Data entry

- Values can be entered in each field or subfield
- Picklists can be defined as in WinISIS (no limits)
- Validation PFT’s can be applied upon saving
- Automatical indexing will be applied on each new record
(Web-)JISIS searching

- Based no longer on ISIS QL but Lucene QL:
  - Index-listing support
  - Boolean (AND/OR/NOT)
  - Proximity and range
  - Truncation and escaping characters
  - Ranking result

- Guided search and Expert Search (=no support in QL building)
Printing and sorting

- Fully follows the powerful multi-level sorting of WinISIS
- Up to 4 levels
- Printing in HTML or plain text with pre-defined PFT
- By range, search or marked list
Z39.50

- Z39.50 servers (e.g. LC) can be defined
- Searches to retrieve MARC-records can be done
- [not fully implemented yet]
Groovy programming plug-in

- Groovy = object-oriented programming ‘superset’ of Java, general purpose scripting
- Groovy can be embedded as ‘format exit’ into PFT’s (like the old ISIS-Pascal did before) as : &myProgram(‘...’) or used to produce reports, change database records, etc..
- E.g. putting date into output : &ActualDate()
  ```groovy
  package jisisgroovie;
  import java.text.SimpleDateFormat;
  import java.util.Date;
  def ActualDate() {
    def date = new Date();
    def sdf = new SimpleDateFormat("yyyyMMddhhmmss");
    return sdf.format(date);
  }
  ActualDate()
  ```
J-ISIS Digital Library

• Using TIIKA, all types of documents can be selected and then:
  ▫ Text extracted and loaded into a(ny) dedicated field of any J-ISIS database
  ▫ Automatic creation of URL hyperlink to original document, stored in J-ISIS repository (idocs-subfolder)
  ▫ Full-text indexed and searchable
  ▫ With highlighting of search-keys

• Very fast processing, economical storage
Digital Library example record

<<Reference Manual
(Version 1.5)
UNESCO
Information Society Division
Sector of Communication and Information
© UNESCO, June 2004
(This manual refers to Winisis 1.5 build 3)
Revised by: Ben Winnubst, New Zealand (June 2004)

Foreword
CDS/ISIS is a menu-driven generalized Information Storage and Retrieval system designed specifically for the computerized management of structured non-numerical data bases. One of the major advantages offered by the generalized design of the system is that CDS/ISIS is able to manipulate an unlimited number of data bases each of which may consist of completely different data elements. Although some features of CDS/ISIS require knowledge of and experience with computerized information systems, once an application has been designed the system may be used by persons having had little or no prior computer experience. For advanced users, CDS/ISIS offers a wide range of programming facility allowing the development of specialized applications through the use of its powerful print formats. For real computer programmers, an external
Web-JISIS

- Prototype not yet finished, but functional
- Based on external Servlet server (Tomcat or Jetty) or the embedded Jetty-server (work in progress)
- Uses J-ISIS server part as database-server
- Functions implemented: database-selection, browsing, editing and searching
- Very fast and interactive
Web-JISIS (1)

• Put the web-jisis3.war file into the webapps-folder of your jsp-server (TomCat/Jetty)
• Start the server
• Open the URL: http://localhost:8585/web-jisis3/
• Login with Admin/Admin
• Lay-out is CSS-based, so can be fully changed
Web-JISIS: browse records

- After database selection the database can be browse with the normal navigation buttons
- Browse can work with any defined PFT
Web-JISIS (2)

- Editing records with one of the defined worksheets
Web-JISIS with Digital Library

Search supported by dynamic suggestion lists

Example Digital Library result
Some concluding remarks/evaluation

+
- 

• Builds upon WinISIS
• Adds WWW-technology/XHTML
• Proven technologies : Java, Lucene, Berkely DB...
• Fully multi-platform
• No more limits re capacity
• Both desktop and WWW-client are available
• ISIS core elements (concepts, Formatting Language) preserved

=> All an ISIS-user can wish ?

• Interfaces for both Desktop and WWW not fully polished/decorated
• Real library functions (e.g. Loans, acquisitions) not yet implemented (Louvre?)
• No direct compatibility with old ISIS-technology
• Cold-water fear of existing ISIS-users
• Limited developer’s team (J-C Dauphin + ?)
Practical information

• J-ISIS website:
  http://www.kenai.com/projects/j-isis
  for downloads, forum and support and documentation
• Full manuals:
• Author: Jean-Claude Dauphin, previously ISIS-officer at UNESCO headquarters (jc.dauphin@gmail.com)