SIARD Archive Browser – The Components⋆

Arif Ur Rahman[1,2], Gabriel David[1,2], and Cristina Ribeiro[1,2]

[1] Departamento de Engenharia Informática–Faculdade de Engenharia,
Universidade do Porto
[2] INESC TEC
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
{badwanpk,gtd,mcr}@fe.up.pt

Abstract. The Software-Independent Archival of Relational Databases
(SIARD) project developed a tool known as the "SIARD Suite" for pre-
serving relational databases. The tool converts a relational database
to a XML format. This paper presents the components of the SIARD
Archive Browser which is a simple to use and platform-independent tool
for browsing a SIARD Archive. This may be helpful for users interested
in using the software. Moreover, it may be useful for people who want
to re-use the code and develop software for browsing a SIARD archive
with more functionality.

Keywords: SIARD, database dissemination.

1 Browser Components

The SIARD Archive Browser[1] is a free, platform-independent and easy to use
tool developed in Java. It offers functionality including viewing table data, sort-
ing a table by a column, searching a specific value in a table and joining tables
(left join) [1].

The architecture of the SIARD Archive Browser is presented in Figure 1. The
SIARD archive is generated by the SIARD Suite. It contains the archived data
and metadata. The browser has a very simple user interface for users to retrieve
data. The control component includes different parsers for retrieving data from
the archive based on some criteria. Users interact with the user interface, the
control component mediates the input, converts it to commands for retrieving
data from the SIARD archive and processes data before displaying it on screen.

The control component has sub-components which interact with each other
to process the data. The sub-components include the following.

- Path: This component manages the file paths in the archive. Different file
  paths are set including the metadata file path and the table files paths when
  a user opens an archive for browsing.

⋆ This work is supported by FCT grant reference number SFRH/BD/45731/2008.
[1] SIARD Archive Browser can be downloaded from http://goo.gl/RWCUI
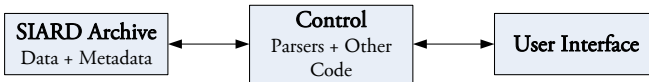
Fig. 1. Basic Architecture

- Parsing Metadata File: The `metadata.xml` file is parsed using the XML DOM parser. There are two reasons for using the DOM parser for parsing this file. Firstly, the size of this file is small and it does not require a lot of memory. Secondly, the browsing process requires the extraction of pieces of metadata which can be done easily using the DOM parser.
- Parsing a Table for Displaying: Users may choose to access a table and search the data in it. A SAX XML parser is used for loading all the data in a table which requires lesser memory for the operation than to DOM parser.
- Parsing a table to identify the single matching row: When a base table is displayed on screen, a user may click on any of the foreign keys in the table. Clicking a value in a foreign key column displays the single matching row in the referenced table. A SAX XML parser is used for this functionality which parses the referenced table till the matching row is found.
- Parsing a table to identify all referenced rows: When a table referenced by the base table is displayed on screen, a user may click on the key column used as a reference in the base table. The SAX XML parser for this functionality parses all the base table. The matching rows are identified and displayed on screen.
- Joining tables: The software supports joining (left join) the base table with the tables it references. The result includes all the rows from the base table joining them with the tables it references, one at a time.

These components may be re-used in other software with more functionality for disseminating a SIARD archive. For example they may be re-used to develop a software for browsing SIARD archives stored in a DSpace[2] repository.

2   Future Work

This is a work in progress and in the future support for browsing large objects in an archive will be included. Furthermore, the functionality of joining tables in which users may include some criteria (SQL 'where' clause) will be included.

Reference

1. Rahman, A.U., David, G., Ribeiro, C.: SIARD Archive Browser. In: Zaphiris, P., Buchanan, G., Rasmussen, E., Loizides, F. (eds.) TPDL 2012. LNCS, vol. 7489, pp. 496–499. Springer, Heidelberg (2012)

---

[2] `http://www.dspace.org/`