Java CDS/ISIS: A Rigorous Tool For Effective And Efficient Data Sorting

saravanan T

Follow this and additional works at: https://digitalcommons.unl.edu/libphilprac
Part of the Library and Information Science Commons
Java CDS/ISIS: A Rigorous Tool for Effective and Efficient Data Sorting

Dr. Saravanan, T.

Associate Professor, Department of Library and Information Science, Annamalai University, Tamilnadu, India. E-mail: sarlib@live.in / tsarlib@gmail.com
ORCID ID: https://orcid.org/0000-0003-0123-6210

Abstract

An attempt has been triggered in this paper to slice the features of data sorting capabilities of UNESCO’s Java CDS/ISIS application (Saravanan, T. 2018b & 2018c). J-ISIS was initiated by UNESCO and further developed by Jean-claude Dauphin (2016). J-ISIS clears maximum WinISIS limitations and restrictions. J-ISIS includes the CDS/ISIS concepts to keep the assets and experience of the users. The effectiveness and efficiency of J-ISIS data sorting technique is evaluated in this paper. A bibliographic database comprising 1641 records is opted to test the J-ISIS sorting features. The database contains 17 fields, subfields and repeatable field. J-ISIS data sorting technique is very fast and more effective than the other techniques as J-ISIS adopts Berkeley DB to store all the records of a database and Lucene for creating Inverted files. J-ISIS has an embedded Web Browser, which helps to display the records in variety of the formats. J-ISIS time consumption to sort 1641 records with the 17 fields, subfields and repeatable field in text format is 17188 milliseconds for name, 31724 milliseconds for department, 30204 milliseconds for department & yr, and 29503 milliseconds for department & yr & name. To sort the records in html format, J-ISIS consumed 17448 milliseconds for name, 35024 milliseconds for department, 35486 milliseconds for department & yr, and 36143 milliseconds for department & yr & name. The results prove that the sorting techniques of J-ISIS are excellent in information processing and retrieval. This paper explores the sorting features of J-ISIS only, and no other applications/features are discussed.

Keywords: Library, FOSS, application, UNESCO, Java CDS/ISIS, J-ISIS, Jean-Claude Dauphin, JC Dauphin, Database, Sorting.

Introduction:

Data Sorting is an important process in an information retrieval system. Sorting is the process of arranging data into meaningful order so that data can be analyzed more effectively. Sorting demands more depth work while comparing to the normal indexing method. In early days, indexing had been done manually by the skilled people, who were experts in those fields. Nowadays, sorting could be done in a sophisticated way using the advanced technology. Many powerful applications/databases offer these sorting techniques to handle the data. Like as WinISIS, the J-ISIS has its own ability to sort the data in a desired way so that the user’s demands would be satisfied. The effectiveness and efficiency of sorting techniques of J-ISIS is evaluated in different ways and the results are presented here for better capture. J-ISIS and other ISIS family products are different from other applications as they let the designers to design their own databases, which may work well in standalone or local network or web server. This paper may help the advanced users of CDS/ISIS or WinISIS. Depth knowledge in the said ISIS application is an important thing to handle the sorting in J-ISIS.
J-ISIS: An overview

J-ISIS is a new multiplatform Free and Open Source Software (FOSS) ISIS that provides the same successful concepts and functionalities as the actual UNESCO CDS/ISIS FOR WINDOWS (WinISIS) software. J-ISIS removes many of WinISIS limitations and restrictions, uses a Client/Server architecture, is fully UNICODE, and benefits of the latest software developments. J-ISIS follows the CDS/ISIS concepts to keep the assets and experience of the users, such that users familiar with the CDS/ISIS software family will also be familiar with J-ISIS and will retrieve the same concepts. J-ISIS is a rich client application written to communicate with a specifically designed database server that is accessible as a stand-alone application and over some local network. It can be used on a single host machine, over a small local network, without Internet access or a Web server installed. According to Jean-Claude Dauphin (2016), J-ISIS is not an Integrated Library System (ILS) as ABCD, it’s a non relational (No SQL) database management system that uses the ISIS concepts and that is particularly well suited for the storage and retrieval of bibliographic information. While it is possible to use J-ISIS for publishing an OPAC, Managing Acquisitions, Loan and Patrons/Users as it is done with WinISIS there are some specific ILS modules under development. As CDS/ISIS, J-ISIS is a flexible Information Storage and Retrieval system designed specifically for the computerized management of structured non-numerical databases. One of the major advantages offered by the generalized design of the system is that J-ISIS is able to manipulate an unlimited number of data bases each of which may consist of completely different data elements. Furthermore J-ISIS uses TCP/IP protocol to communicate between computers and is a database server that follows the Client/Server architecture. For advanced users, J-ISIS offers a wide range of programming facility allowing the development of specialized applications through the use of its powerful print formats. The J-ISIS embedded Web browser and Web server offer the possibility to use the new Web technologies such as HTML5, CSS3, and JavaScript inside ISIS print formats. The Groovy programming language has replaced ISIS Pascal offering the same functionalities and much more. For real computer programmers, the ISIS_DLL is replaced by the J-ISIS core library (jisis-core.jar), which is interoperable between different platform and which provide all necessary tools for developing J-ISIS based applications. J-ISIS keeps the assets of CDS/ISIS and WinISIS such as FDT, Worksheet, PFT and FST.

Objectives: The following objectives are vouched in this paper.

1. To build an indexing system for theses bibliographic data using Java CDS/ISIS.
2. To explore the effectiveness and efficiency of sorting techniques of Java CDS/ISIS.

Research Design:

A bibliographic theses database consisting of 17 fields, subfields, and repeatable field, has been built using J-ISIS application. 1651 records were added into the database. The bibliographic details have been sorted in multiple ways as shown below. The opted fields name for the purpose of sorting are;

- Name (Researcher)
- department
- department & year
- department, year & name
Invoking Java CDS/ISIS:

At the time of invoking J-ISIS the screen will look like as shown in the figure-1. The database connection needs to be set first for further actions.

Figure-1: Loading J-ISIS

J-ISIS remote database details offers are explored via the figure-2 where users need to enter the id & password to move to the next stage. The sample database in this paper has been designed using the CCF standard format. The figure-3 explores the available databases in which the database named as ‘autbdb’ is selected.

Figure-2: Login
Figure-3: Database Selection

The Figure-4 explores the FDT screen for the database AUTBDB where the required fields are identified.

Figure-4: FDT

The Figure-5 explores the Worksheet screen for the database AUTBDB. The required fields can be selected and sequenced in a desired way. Further, the advanced data entry worksheet editor of J-ISIS lets the users to access more features.
The Figure-6 shows the FST screen for the database AUTBDB where a few fields have been included for the purpose of indexing along with the indexing technique and format.

The Figure-7 shows the PFT screen for the database AUTBDB where the format has been written to generate the required output.
Figure 7: PFT

The Figure 8 shows the data viewer that reflects the bibliographic details for the MFN: 83. The maximum record status i.e.: 1641 has been highlighted as shown in the screen shot.

Figure 8: MFN-Maximum Record

Let us see the sorting methods of J-ISIS towards the given sample outputs. The print feature of J-ISIS shows the sort option where the conditions need to be fulfilled.
The *Figure*-9 shows the researcher name wise print and sort option. The print options namely database, records, parameters, output etc., must be filled with appropriate data. The sorting option let us to view the *Figure*-9.1 where the required options need to be completed.
The Figure-9.2 shows the output in text format where the researcher name and other few details are sequenced. For illustration purpose, throughout this paper the record with the MFN: 83 has been selected.

The researcher department wise print and sort option is retrieved towards the Figures-10 to 10.2.
Figure-10.1: Researcher Department wise sort

Figure-10.2: Researcher Department wise sort-output

The Figure-10.2 displays the record for the MFN: 0083 as per the department wise sort. Observe the next MFN i.e.:0091.

The researcher department & Year wise print and sort option is retrieved towards the Figures-11 to 11.2.
Figure-11: Researcher Department & Year wise print

Figure-11.1: Researcher Department & Year wise sort
Figure 11.2: Researcher Department & Year wise sort-output

The Figure 11.2 displays the record for the MFN: 0083 as per the department & year wise sort. Observe the next MFN i.e.:0082, which can be compared with the output available in the Figure 10.2 where the next MFN was 0091.

The researcher Department, Year & Name wise print and sort option is retrieved towards the Figures 11 to 11.2.
The Figure-12.2 displays the record for the MFN: 0083 as per the department, year & name wise sort. Observe the next MFN i.e.:0084, which can be compared with the output available in the Figure-10.2 (next MFN was 0091) & Figure-11.2 (next MFN was 0082).
Figure-13: Time consumption for printing & sorting 1641 records (department wise sort)

<table>
<thead>
<tr>
<th>Time Consumption (in milliseconds)</th>
<th>Name (Figure-9.2)</th>
<th>Department (Figure-10.2)</th>
<th>Dept_Yr (Figure-11.2)</th>
<th>Dept_Yr_Name (Figure-12.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print Records</td>
<td>113</td>
<td>119</td>
<td>101</td>
<td>104</td>
</tr>
<tr>
<td>Read DB</td>
<td>5275</td>
<td>5833</td>
<td>5491</td>
<td>4406</td>
</tr>
<tr>
<td>Interpret FST</td>
<td>5275</td>
<td>6093</td>
<td>5802</td>
<td>4655</td>
</tr>
<tr>
<td>Write Temp</td>
<td>55</td>
<td>56</td>
<td>52</td>
<td>26</td>
</tr>
<tr>
<td>Prepare file to Sort</td>
<td>5717</td>
<td>6298</td>
<td>6028</td>
<td>4791</td>
</tr>
<tr>
<td>Sorting</td>
<td>17188</td>
<td>31724</td>
<td>30204</td>
<td>29503</td>
</tr>
</tbody>
</table>

Figure-14: Time consumption for printing & sorting 1641 records (Text output)

<table>
<thead>
<tr>
<th>Time Consumption (in milliseconds)</th>
<th>Name (Figure-i)</th>
<th>Department (Figure-ii)</th>
<th>Dept_Yr (Figure-iii)</th>
<th>Dept_Yr_Name (Figure-iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print Records</td>
<td>169</td>
<td>72</td>
<td>84</td>
<td>95</td>
</tr>
<tr>
<td>Read DB</td>
<td>5563</td>
<td>4318</td>
<td>4545</td>
<td>4767</td>
</tr>
<tr>
<td>Interpret FST</td>
<td>5844</td>
<td>4525</td>
<td>4817</td>
<td>5049</td>
</tr>
<tr>
<td>Write Temp</td>
<td>49</td>
<td>42</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>Prepare file to Sort</td>
<td>6068</td>
<td>4699</td>
<td>4997</td>
<td>5232</td>
</tr>
<tr>
<td>Sorting</td>
<td>17448</td>
<td>35024</td>
<td>35486</td>
<td>36143</td>
</tr>
</tbody>
</table>

Figure-15: Time consumption for printing & sorting 1641 records (html output)
Conclusion:

The results reflect the effectiveness and efficiency of J-ISIS while printing & sorting the data. A bibliographic database named as AUTBDB was accessed using J-ISIS for information processing and retrieval. J-ISIS maintains the accuracy and reliability in data processing. The figure-13 shows the actual time consumed by J-ISIS to process all the records for printing and sorting in Text format as per the department. The figure-14 indicates the time consumption of J-ISIS for printing and sorting the total 1641 records in text format as per the name, department, department & year, department & year & name. The figure-15 reflects the time consumption of J-ISIS for printing and sorting the same 1641 records in html format as per the name, department, department & year, department & year & name. The graphical representation in figures-14 & 15 show crystal clear information about the sorting capabilities of J-ISIS. The database contains 17 fields, subfields and repeatable field as explained in the Figure-4. J-ISIS time consumption to sort the total records in to text file is 17188 milliseconds for name, 31724 milliseconds for department, 30204 milliseconds for department & yr, and 29503 milliseconds for department & yr & name. To sort the same records in html format, J-ISIS consumed 17448 milliseconds for name, 35024 milliseconds for department, 35486 milliseconds for department & yr, and 36143 milliseconds for department & yr & name. The reports generated in html format are given in the annexure (Figures i-iv) for better understanding. J-ISIS supports various document formats such as Text, docx, Rtf, Pdf, html, mp3, Flv etc., Librarians, who have previous experience in CDS/ISIS applications may think about J-ISIS to overcome the OS compatibility related issues. Librarians may use this application to control their library functions without complications. Further, there would be no need to spend money to purchase the commercial applications. Hope, the institutions/companies/organizations/research centres etc., always expect the reduced budget proposals from the Librarians desk. Don’t they? Librarians, time knocks your door steps. Make a right choice.

Acknowledgement:

I would like to thank Mr. Jean-Claude Dauphin (Retired from UNESCO, Paris), who is behind the development of J-ISIS (Java-Integrated Set for Information Services) application and offered valuable suggestions.

***

References:


Saravanan, T. (2011). WinISIS: The information storage and retrieval application for faculty member’s research performances. *Humanities, 47*(20), 393-400.


Annexure

Html output

Figure-i: Name wise sort

Figure-ii: Department wise sort
Figure-iii: Department & Year wise sort

Figure-iv: Department, Year & Name wise sort

***