



CENTERIS 2012 - Conference on ENTERprise Information Systems / HCIST 2012 - International Conference on Health and Social Care Information Systems and Technologies

EwayDataCollect: Web Service-based Data Collection Tool for Education System

Ruchika^{*a} and Anita Goel^b

^a Department of Computer Science, University of Delhi, Delhi-110007, India

^b Department of Computer Science, University of Delhi, Dyal Singh College, New Delhi –110003, India

Abstract

In education system, there is a need to collect data from the education providers, to facilitate informative decision making by the policy makers. Data collection is handled using tools like questionnaires, online surveys, online forms submission, and Electronic Data Interchange. This paper presents, EwayDataCollect, a tool for online data collection in education system. EwayDataCollect leverages the web service technology for data collection. It uses web service interface which is interoperable and supports communication between heterogeneous environments. The tool enables computer to computer data collection of both, statistical and detailed data. It facilitates collection of data from any number of educational providers. EwayDataCollect consists of three components: (1) Eway-Invocation, (2) DCF-Dissemination, and (3) Eway-Processing. The first component is responsible for registering the data requestor, accepting Data Capture Format (DCF) and details like URLs and email-ids of data providers. The task of second component is to disseminate DCF to the data providers. The third component of tool collects filled DCF from the data provider's registered URL. The collected data is parsed and stored in a database. EwayDataCollect is faster and easier to use. EwayDataCollect architecture, implementation and application are presented in this paper.

© 2012 Published by Elsevier Ltd. Selection and/or peer review under responsibility of CENTERIS/SCIKA - Association for Promotion and Dissemination of Scientific Knowledge. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Web Service; XML; Data Capture Format; Education Management Information System(EMIS); Data Consumers; Data Providers.

* Corresponding author. Tel.: +919999398936;
E-mail address: ruchikathukral1980@gmail.com

1. Introduction

Education Management Information System (EMIS) requires collecting of data from education providers and stakeholders, for decision making and long term planning. The data needs to be updated regularly to provide latest information to policy makers. Several methods exist – online and offline; for the collection of data. Generally, data is collected from vast network of education providers like schools, colleges, universities, institutes etc.

During data collection, there is a need to focus on quality, accuracy and reliability of data, since policies and decisions are based on the collected data. Traditionally, mail-based methods, telephone-based methods, personal visit methods, respondent tapes and electronic methods have been used for data collection [10][18]. Different tools like Fax machines, Electronic Data Interchange (EDI) etc were used for communication between homogeneous environments [5]. ICT tools have resulted in a shift in the way data is collected from the data providers. Data may be collected online, using online surveys, online forms, email etc. or offline using CD-ROM, floppy diskettes etc. in addition to the traditional methods of data collection. Though, data collection has become faster with the use of ICT; still there are many issues to be handled, like, fast Internet connection, cost, time and heterogeneity.

To collect data, there is a need to provide a questionnaire like format to the data provider for filling the data. Data Collection Format (DCF) is a template designed by the data consumer, which contains questions for which the data is to be filled by the data providers. Generally, DCF requires the data provider to fill statistical data. DCF ensures uniformity of data collected from the different data providers.

Data collection is a time consuming process. Collecting data from different geographical locations requires leveraging the technology for data collection. Though computers are used, but computers in hundreds and thousands of education providers are different with different operating systems, software, versions, networking platforms etc. Computer to computer data transfer is a challenging task in heterogeneous environments. Due to heterogeneity of computer systems, online data collection becomes difficult.

Web Service is based on a set of standards that allows it to be interoperable and technically provide possibilities to interact between heterogeneous environments. Web service support XML, a standard language used for communication between applications over the World Wide Web.

This paper describes EwayDataCollect, a tool for online collection of data for education system. EwayDataCollect exploits web services for data collection. The tool enables platform-independent, computer to computer collection of, both, detailed and statistical data. EwayDataCollect interacts with two target groups – (1) *Data consumers* who want data to be collected in DCF, and (2) *Data providers* who perform data entry in DCF. The tool provides an interface for:

- accepting DCF from data consumers
- sending DCF to data providers
- collecting filled DCF from data providers, and
- storing the collected data

EwayDataCollect consists of three components - (1) Eway-Invocation, (2) DCF-Dissemination, and (3) Eway-Processing. The *Eway-Invocation* component defines an interface for data consumer. The data consumer interacts with this interface to provide email addresses and URL of data providers and DCF. The task of *DCF-Dissemination* is disseminating DCF to the data providers. *Eway-Processing* component handles collection of filled DCF, parsing of data and its storage.

Since DCF is sent to data providers in a portable format, it is independent of hardware and software used. It eliminates the need of latest computers or special infrastructure for data entry. Moreover, collecting data directly from its origin - the data provider, reduces chances of errors in the collected data. The electronic

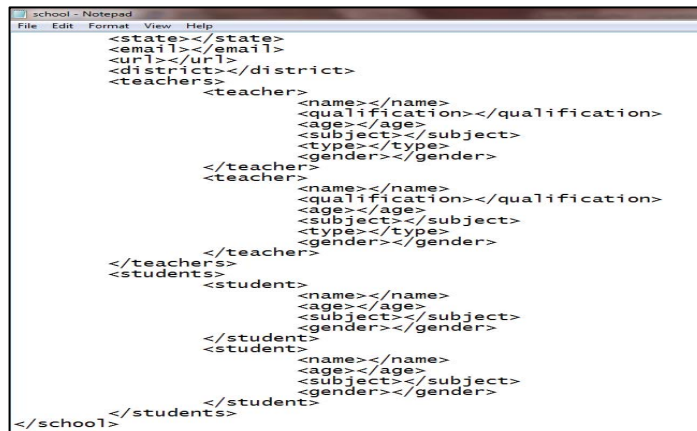
collection of data does away with CDs, hand-filled or printed forms for data collection, thus supporting *green computing*. It also speeds up the process, resulting in timely collection of data.

Here, we describe the design and implementation of EwayDataCollect. The tool application is developed in Java Eclipse environment. We present results based on applying EwayDataCollect to collect data from several schools.

In this paper, Section 2 is about XML based Document Exchange. Section 3 describes data collection using Web Service-based Interface. Section 4 illustrates the architecture of EwayDataCollect. Section 5 focuses on the case study on data collection for EMIS from schools. Section 6 shows results of implementation of the tool. Section 7 is about related work and section 8 states the conclusion.

2. XML Based Document Exchange

Data collected is required to be relevant, reliable and on time. DCF must be designed in a manner so that it is self-explanatory, easy to understand and easy to fill. DCF should not be too long that people avoid providing the data [17]. Here, the questionnaire has been designed to be small, easy to understand and fill. XML schema is used for DCF. DCF is designed to accept both statistical data, like, total number of classes, teachers, students, and, detailed data, like, details of all teachers. XML schema takes single line entry for statistical data and multiples lines for detailed data. Fig 1 shows a DCF schema.



```

school - Notepad
File Edit Format View Help
<state></state>
<email></email>
<url></url>
<district></district>
<teachers>
  <teacher>
    <name></name>
    <qualification></qualification>
    <age></age>
    <subject></subject>
    <type></type>
    <gender></gender>
  </teacher>
  <teacher>
    <name></name>
    <qualification></qualification>
    <age></age>
    <subject></subject>
    <type></type>
    <gender></gender>
  </teacher>
</teachers>
<students>
  <student>
    <name></name>
    <age></age>
    <subject></subject>
    <gender></gender>
  </student>
  <student>
    <name></name>
    <age></age>
    <subject></subject>
    <gender></gender>
  </student>
</students>
</school>

```

Fig. 1 XML Based Document

3. Data Collection using Web Service-based Interface

Computers spread across hundreds and thousands of data providers cannot be same – different operating systems, software and versions. Computer to computer data transfer in heterogeneous environment requires use of web services. Web services are platform-independent and can access data stored in the XML format from any platform. Web service based interface is used for collecting data from data providers. For ensuring platform-independent mobility, DCF is sent and received in XML format. The data providers need to import XML schema and export data in XML format.

Data collection requires interaction among the Education data consumers and Education data providers. Education department, national education agencies etc. act as data consumers. Educational institutes like schools, colleges, universities etc are data providers. Since data providers need not improvise their compute facility for data filling, data can be entered directly by the data providers, as against a centralized system of

data entry. De-centralizing data entry results in fewer errors and timely data collection. Its use is best exploited when collection of detailed data needs to be done.

4. EwayDataCollect Architecture

EwayDataCollect is a web-based service for data collection from education providers. Since hundreds and thousands of education data providers have different computing environments, computer to computer data transfer is not possible. EwayDataCollect exploits web services for data collection. Web services are platform-independent and can access data stored in the XML format from any platform. EwayDataCollect sends DCF to data providers in XML format and also retrieves the filled data form in XML format. This enables computer to computer data transfer. Since data providers need not improvise their compute facility for data filling, data can be entered by data providers, as against a centralized system of data entry. De-centralizing data entry results in fewer errors and timely data collection. Its use is best exploited when collection of detailed data needs to be done.

The deliverables provided by EwayDataCollect to education data consumers are as follows:

- A service which sends DCF to data providers via e-mail
- A service which collects data from data providers in DCF, over the Internet, electronically
- Collected Portable data
- Database files populated with data

The deliverables provided by EwayDataCollect to education data providers are as follows -

- DCF form for entering data, regardless of hardware and software of computer
- Easy-to-fill data entry format

The tool architecture is composed of three main components, namely, *Eway-Invocation*, *DCF-Dissemination*, and *Eway-Processing*, as shown in Fig 2. The purpose of *Eway-Invocation* is to accept DCF list of data provider’s email-ids and URLs from the data collectors. *DCF-Dissemination* sends DCF to the data providers. *Eway-Processing* collects filled DCF from the registered URLs, parses XML data and stores it in the database.

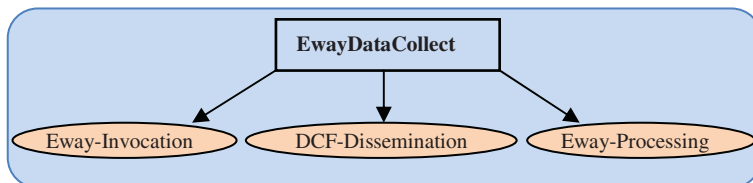


Fig. 2 Components of the Tool

4.1. Eway-Invocation

Fig. 3 shows the structure of *Eway-Invocation*. The education data consumer interacts with this component. Eway-Invocation provides three interfaces to the education data consumer for input, as follows:

- *Register_DCF()* accepts XML schema of DCF and converts it to DCF.xml
- *Register_email()* accepts emailFile having email-id of each education data providers. The email-ids provided in this file are referred to as registered email-ids
- *Register_URL()* accepts URLFile having URL of each education data providers. The URLs provided in this file are referred to as registered URL

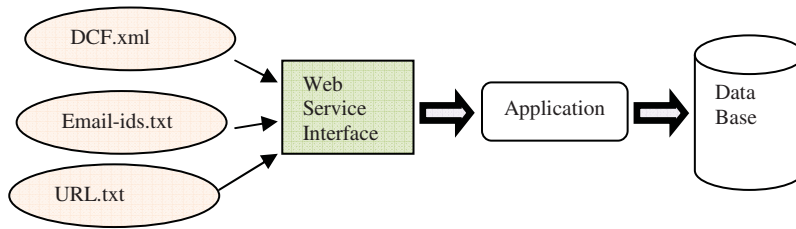


Fig. 3 Eway-Invocation Component

The files are uploaded by the data consumers via the web interface and are stored in the data base. Fig 4 shows the interface for education data consumers of EwayDataCollect.

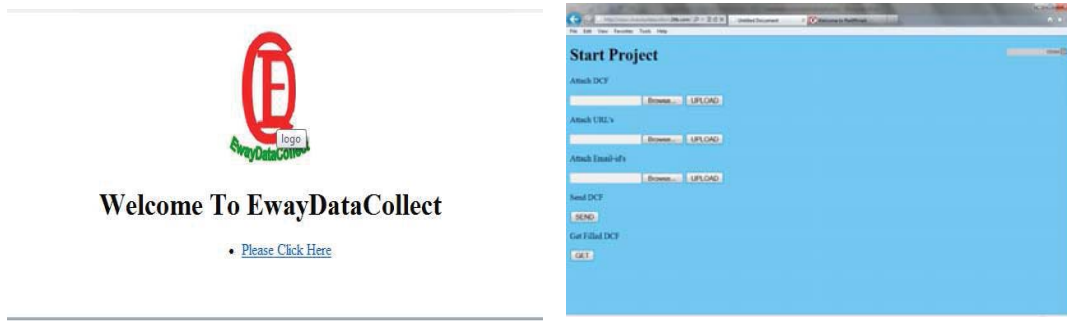


Fig. 4 User Interface of EwayDataCollect

4.2. DCF- Dissemination

This component of the tool is responsible for sending DCF to education data providers. For this, DCF in the XML format is sent to all registered email-ids, as shown in Fig. 5. Education data providers receive the DCF as XML schema, which needs to be stored as DCF.xml file.

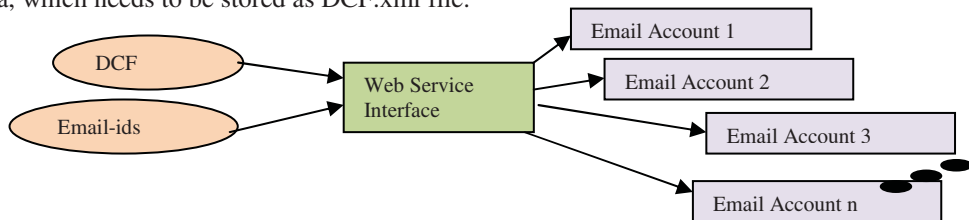


Fig. 5 DCF-Dissemination Component

4.3. Eway-Processing

The Eway-Processing component comprises of three steps. It requires education data providers to fill the DCF with data and upload the DCF at their registered URL. To enter data in DCF, the data provider is instructed to copy the XML schema received in their email and save as DCF.xml file. The data provider can import the file in Excel and fill the relevant data. This way the data is filled at the data provider's convenience

and saved as *filledDCF.xml* file [12]. The *filledDCF.xml* file is then uploaded by the data provider at its registered URL, as shown in Fig. 6.

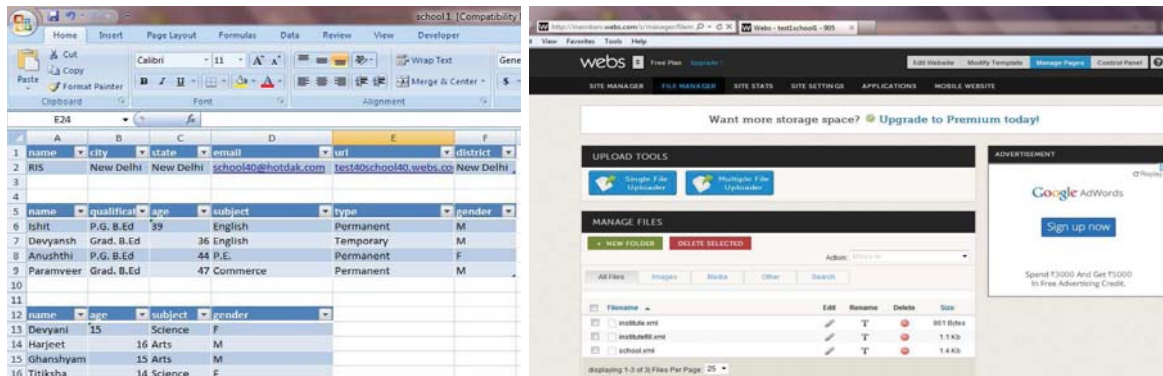


Fig.6 Eway-Processing (Filled DCF and Uploaded on URL through Control Panel)

The tool provides an interface *get()* which picks the filled DCF from the registered URL. The data is then parsed and stored in the database. Fig. 7 shows the structure of *Eway-Processing*.

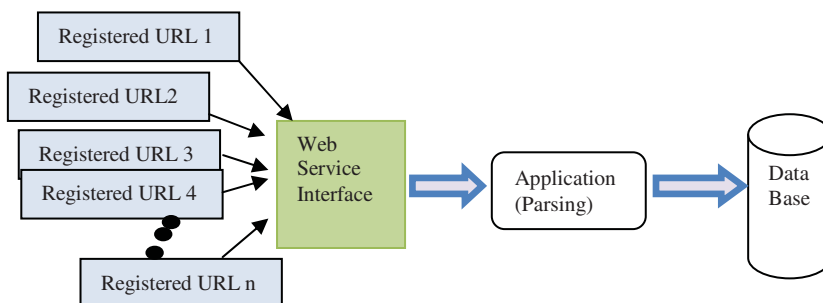


Fig.7 Eway-Processing Component

This component of the tool retrieves the *filledDCF.xml* files from the registered URLs. The data in XML format is then parsed and stored in the database. The tool saves the data in database under different tables like teachers, students, school etc.

A data collection cycle using EwayDataCollect is summarized here:

- Education data consumer invokes:
 - *Register_DCF()* with XML schema of DCF.
 - *Register_email()* with file containing email-id of data providers.
 - *Register_URL()* with file containing URL of each data provider, to upload the filled form
- *send()* disseminates DCF as attachment to email-ids.
- Education data providers:
 - Import DCF in Excel. A tabular format for data entry is displayed. The data is filled in the spreadsheet form.
 - Export filled DCF in XML format.
 - Upload DCF at its URL.

- *get()* picks up DCF from the URLs.
- Filled DCF is parsed and data is stored in database.

An overall interaction of EwayDataCollect with the external environment is shown in Fig 8.

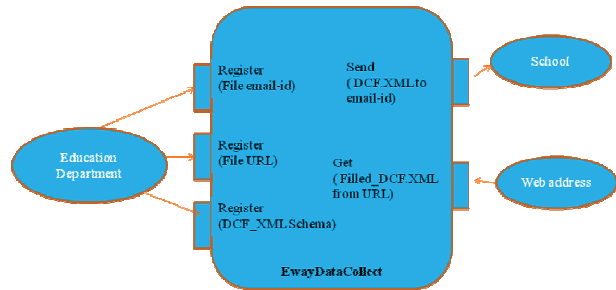


Fig. 8 Interface to the environment

5. Case study: Data Collection for EMIS from schools

EMIS is created to utilize data collected from the different institutions for processing and analysis by decision makers. Integrated data and information systems are at the very core of EMIS development in their support of the educational management functions throughout the education system. National Ministry of any country manages educational data to form policies [4][5][7][8][9][13][22][23].

The tool has been used to collect data from about fifty senior secondary schools. As a data collector, we collected the email-ids for correspondence, and also the web addresses (URLs) and created two separate files *email.txt* and *URL.txt*, respectively. DCF was created wherein the schools were to provide the data. The DCF was designed to include statistical data like school information, and detailed data like teacher information - teacher name, age, subject, sex etc. Such kind of entries is considered as detailed data. The XML schema for DCF was saved in *DCF.xml* file. As a data collector, we have initialized first component of the tool with three files - *DCF.xml*, *email.txt* and *URL.txt*.

In the second phase, web service interface sent the DCF to registered email-ids. The schools received the *DCF.xml* file and import it in Excel. After filling the DCF, schools export the data, saved it as *filledDCF.xml*, and uploaded it at their registered URL. Our tool then collected the file from the registered URL. Collected data was in XML format. Data in XML format was parsed and stored in database.

Advantages of EwayDataCollect over existing system are:

- Less cost of setting up infrastructure at different level of administration.
- No new software or hardware was required to complete the whole process.
- Basic knowledge of computer would help the representative of school to do all needful.
- Rectifications can be done easily at school level only.
- Printing cost of paper in addition to cost of paper has also been reduced.
- This is much faster than the existing system of data collection
- DCF takes both statistical data and detailed information about students or teachers like name, age, subject, experience etc.

6. Implementations and results

We have made observation for each phase of the tool. For data collectors, designing of DCF is relatively easy in comparison to multiple choice questionnaire or open ended and close ended questions. DCF includes both statistical and detailed data collection fields. No special training had been provided to any school for filling the DCF which in addition had cut training expenses.

In second phase of our tool, DCF is disseminated to 50 schools within 6-7 minutes as shown in chart 1. In the current system of EMIS, it takes around a month to disseminate DCF in printed form or copied in CD's etc. to the schools.

In third phase, where we collect data from schools, data had been collected by our tool in very less time as compared to the current system as shown in Chart 2. In existing system, online data entry at central level, submitting forms, Internet interruptions etc are many reasons for delays in data entry.

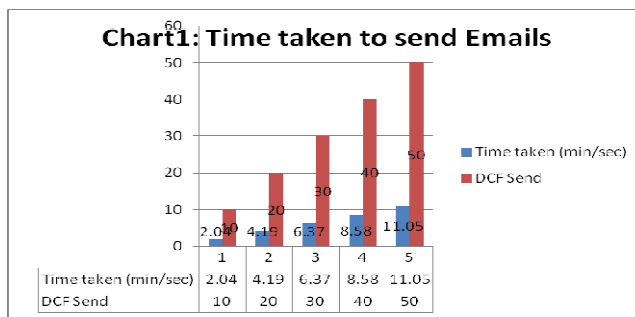


Chart 1: Time Taken to Send Emails

From chart 2, we see that the time taken to collect the filled DCF from the schools is very less. In existing system, process of verification and rectifying errors is long. In comparison, chances of errors are very less in our tool, but if any, can be conveyed to the school by email. After rectifying *filledDCF.xml*, school authorities are required to upload it on their registered URL.

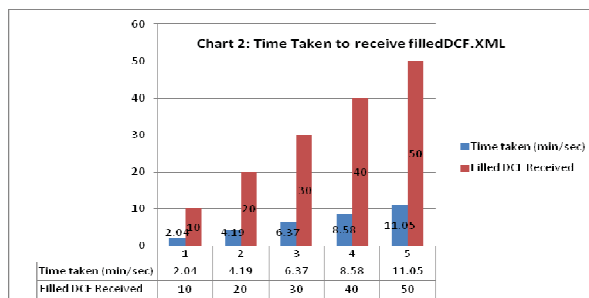


Chart 2: Time Taken to Receive DCF

7. Related Work

Education information system provides information to policy makers to form policies to improve educational standards. Data collection is most important task for any Management Information System. There are many educational providers, like, universities, colleges, schools, private institutes etc. Data collected from education providers is stored in EMIS [3][4][5][6][11][13][21][23]. Many online methods are being used for data collection for offline and online centralized data entry. Decentralized data entry or direct collection of data from its origin has not been implemented by education system due to different computing machines at the education providers. We have proposed a tool to collect data directly from the education providers without any modification to their computing environment.

Education data consumers design DCF to collect data, uniformly, from education providers. DCF aims to capture only statistical data. The lack of detailed data may hamper the decision-making process. DCF used by education system is normally long multi pages format to be filled by data providers and collected by data consumers for centralized data entry after verification [3][4]. Due to interrupted Internet connection and power supply, online form may be resubmitted many times causing delays in integration of data. We use DCF based on XML schema for capturing both statistical and detailed data from education providers. It is easy to understand and quick to fill. We use XML documents in data collection from education providers, since data

can be easily exported in XML format which is transferable through web service interface. XML schema can be easily imported and exported in Excel worksheet, as done by Zoonoses system [12].

Web services were used by Australian government to collect data for EMIS about the performance of students [7]. But education management still follows the way of data collection via online surveys or offline printed forms. Web services have been used for integrating data, which also include context mediation, quality at source etc [16]. We use web service interface for collection of data from education providers [19][20] for its qualities such as interoperability and data exchange over networks in heterogeneous environments [1][15][17][18]. We focus on data collection element of EMIS which is still following the online and offline data collection methods with centralized entry of data after verification of data providers filled DCF.

In this paper, we are proposing web service based data collection tool which uses XML document for exchange of information over the net irrespective of any computing platforms, programming languages used for applications, networks etc. Our tool is collecting data directly from its origin hence minimizing the chances of errors. The tool enables collecting of relevant data on time without any updating of software and hardware component of computing system at any level of administration in education system.

8. Conclusion

In this paper, we present a Web Service-based tool to collect data online from educational data providers. Application to application data exchange needs homogeneous computing environment. Many educational providers such as universities, colleges and school are located as distant places across a country and setting up homogeneous computing system tends to be very costly. Data exchange between heterogeneous computing systems is possible by using web service interface. We have proposed a tool for collection of data directly from educational data providers without incurring any expense for setting up homogeneous system. The tool has been used successfully for collecting data from different education data providers. XML based schema has also proved to be easy to understand. As there was no need to update the hardware or software infrastructure to fill in the data, it is also a cost efficient method. Once the data is uploaded on web address, data consumers can collect the data at their own convenience. Collected data is relevant and more reliable as well as on time. Data received on time definitely helps decision makers to form policies for improvement of the education system.

References

- [1] Ana-Ramona Lupu, Razvan Bologa, Gheorghe Sabau, Mihaela Muntean "Integrated Information Systems in Higher Education" WSEAS Transaction on Computers ISSN: 1109-2750 may 2008.
- [2] Ali Arsanjani, Brent Hailpern, Joanne Martin, Peri L. Tarr "IBM Research Report Web Services: Promises and Compromises" RC22494 (W0206-107) June 20, 2002.
- [3] Arun C. Mehta "DISE II: HE-MIS Higher Education Management Information System" ACM/HE-MIS/Sept 14, 2007.
- [4] Arun C. Mehta "Status of Educational Management Information System in Iraq and Suggestions for Improvement" Workshop on EMIS and Educational Statistics for the Officers of Iraq, Dead Sea, Amman, Jordan, 03-15 April, 2006.
- [5] Asefeh Asemi, Ali Safari, Adeleh Asemi Zavareh "The Role of Management Information System (MIS) and Decision Support System (DSS) for Manager's Decision Making Process" International Journal of Business and Management Vol. 6, No. 7, July 2011.
- [6] Asil A., Almonaies Manar, H. Alalfi James, R. Cordy Thomas, R. Dean "Towards a Framework for Migrating Web Applications to Web Services" Proceedings of the 2011 Conference of the Center for Advanced Studies on Collaborative Research, 2011.
- [7] Australian Government "Higher education information management system (heims) heims web services interface technical specification functions", Commonwealth of Australia, 2008.
- [8] Azlinah Mohamed, Nik Abdullah Nik Abdul Kadir, Yap May-Lin, Shuzlina Abdul Rahman, and Noor Habibah Arshad, "Data completeness analysis in the Malaysian Educational Management Information System" (IJEDICT), Vol. 5, Issue 2, pp. 106-122, 2009.

- [9] California Department of Education “California Special Education Management Information System (Casemis)” Technical Assistance Guide, Dec. 2010.
- [10] Darcy Haag Granello, Joe E Wheaton “Online Data Collection: Strategies for Research” JCD. Alexandria: Vol.82, Iss. 4; pg. 387-394, 2004
- [11] “Educational Management Information System in India” <http://www.educationforallindia.com/page3.html>
- [12] European Food Safety Authority “Technical report on the use of Excel/XML files for submission of data to the Zoonoses system” European Food Safety Authority, 2012:EN-233. [33 pp.], <http://www.efsa.europa.eu/en/publications>, 2012.
- [13] Haiyan Hua and Jon Herstein “Education Management Information System (EMIS): Integrated Data and Information Systems and Their Implications In Educational Management” Annual Conference of Comparative and International Education Society March 2003.
- [14] Hamid R., Motahari Nezhad, Boualem Benatallah “Web Services Interoperability Specifications” 0018-9162/06, IEEE, 2006.
- [15] Jameela Al-Jaroodi, Nader Mohamed, and Junaid Aziz “Service Oriented Middleware: Trends and Challenges” Seventh International Conference on Information Technology, IEEE, 978-0-7695-3984-3/10, 2010
- [16] ark Hansen, Stuart Madnick, Michael Siegel “Data Integration using Web Services” Working Paper 4406-02 CISL 2002-14 May 2002.
- [17] Marcus Powell, “Rethinking Education Management Information Systems: Lessons from and Options for Less Developed Countries” Working Paper No. 6, 2006.
- [18] McDonald, Heath and Adam, Stewart “A comparison of online and postal data collection methods in marketing research, Marketing intelligence & planning” vol. 21, no. 2, pp. 85-95, 2003.
- [19] Ruchika, Anita Goel “Web Services for Education Management”, In Proceedings of International Conference on Data Management, ICDM, India, Nov. 2011.
- [20] Ruchika, Anita Goel “Using Web Service Interface for Data Collection”, In International Journal of Computer Science Issues, IJCSI-2012-9-2-1757, 2012 (in press).
- [21] Shooebridge, JInfoDev study #1253 “Education Management Information System: Case Studies and Lessons Learnt, Case Study of Nigeria”, 2006.
- [22] “The data collection method using the Internet”, The 11th Meeting of the Heads of NSOs of East Asian Countries, Tokyo, Nov. 2006.
- [23] Tom Cassidy, “Education Management Information System (EMIS) Development in Latin America and the Caribbean: Lessons and Challenges”, Jan. 2005.