# Spreadsheet Data Transformation into Web Database using XSLT

Aung San Htoo
*University of Computer Studies, (Mandalay)*
*aungsanhtoo@gmail.com*

## Abstract

*In the system of my paper, XML (Extensible Makeup Language) and XSLT (Extensible Stylesheet Language Transformation) technologies are used with intention of storing and retrieving the large amount of business finance data to the web database, the SQL server through the webpage. The emergence of XML is a significant advance in area of Business to Business (B2B) integration. User can store or transmit structured data using the highly flexible open standard, XML format. The input file must be the Microsoft Excel and transformed into XML file with schema so that the user can understand this on the webpage. As result of doing this paper, the achievement is that the integrated data could have been put into the web database for the Diamond Palace Company Limited. This is a good experience and knowledge as to conversion from Excel to XML. It also is a usage enterprise for the benefit of transferring large amount of data on the web database. This is a way of a system-structure as well as Client-Server Application Architecture.*

## 1. Introduction

With the repaid development of web technology more and more accessible information are available from the web based database. Most of the companies worldwide access the data from the remove sites as remote clients to the servers by storing and retrieving the data [5]. The need to remove redundant process, to reduce software/hardware and staff costs, and to develop their data query processing have become be come key requirements for managing business performance from terminal. The need for the data transformation to produce accurate and secure information is a requirement for the data enquiry [4]. Before, the communication and management control is used through online (e-mail/internet) such as transferring the daily financial and report in the internal company office.

In this paper present the data transformation solution and design interface on the website. What mainly to be approved is data transformation on the web as input from specific Microsoft Excel file to XML file to store the large amount of data on the web database. In design interface, user can access data to the web database from the remote client according to their access role and authentication permission for accessing data enquiry. As design interface solution users can access data and get the information from web page about the company's service and product for customer.

## 2. Background Information

### 2.1 Client-Server Database Architecture

Client-server database design offer many potential advantages to growing and developing companies which are seeking efficient and low-maintenance business solution. The web browser has become the software application of the internet browser. A typical browser (on the client side) interacts with a server to request and retrieve information over a network [3].

Client-Server application is generally divided into three components [5]. The presentation logic of the application most often resides the client, application logic is typically divided between the client and server, while the data management layer usually reside on the server [3].

### 2.2 Web Based Application

Web based application provide with the opportunity of save time and money, and improve the way user interacts with clients, suppliers and business partners [3]. It also provides the power of desktop and server application with a flexibility and accessibility of the Web. Using Web browser, users can securely access applications from anywhere within the each of the company intranet or internet. Benefits of Web-Based application are [5, 6]:

- Existing application can work as both a desktop and Web solution.
- Employee has data access from anywhere through internet.
- Applications can be maintained remotely, reducing overhead costs for system up deep.

- A Web-based application can host parts of an application in different locations; multiple machines can host multiple instances of the application to prevent server overload.
- Application security could be centrally managed.

Providing user with the opportunity to save time and money, and improve the way user interact with clients, suppliers and business partners.

## 3. Transformation

Most of the web application need the manipulating the data to get the target output of source file [4]. We use target XML file as in put of Microsoft Excel file because large amount of data transferring to web database is more flexible and faster, can change user interface design and presenting on web form as XML style sheet schema as user likes [2,1].

There is also meaning data conversion to meet the target output such their system value changes. The structure of XML is deigned as user define tab, it is not concern on their data type to store data into database table, data must be converted into a separate data type value from XML file as data table.

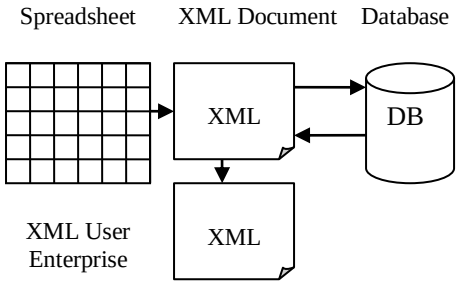


**Figure-1** XML Transformations for usage enterprise

### 3.1. XSLT Tool and Technology

The core task of an Extensible Stylesheet Language Transformation (XSLT) is to apply a stylesheet to a source and produce a result document Figure -2

Regarding first approximation think of the resource document, the stylesheet and the result document as each being an XML document. XSLT perform a transformation process because the output (The result document) is the same kind of object as the input (the source document) this has immediate benefit [1].

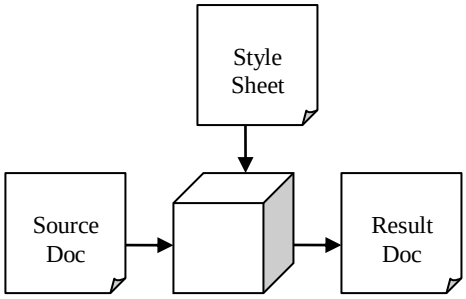


**Figure 2.** Transformation process

## 4. System Contribution & Implementation

### 4.1. System Overview

In this system demonstrate a simple scenario that emphasizes the importance of implementing simple spreadsheets. In the transformation step spreadsheet Excel file to XML file conversion in each client must use excel file conversion process and then result output of XML document upload to the web database server by changing data table format of XML document in the window base or web base by using data table schema with XSLT. When user needs to view data from the web database, user can browse through web browser from website by passing User schema as user enterprise format. If authenticated user wishes to delete data table from database, user can delete existing data from the web database server.

### 4.2. System Design

In this system, user input is Microsoft Excel Spreadsheet file by uploading the file to web application server to convert or transform into XML format from user clients. Transformation process may be web application or window application from client user to store data table to the web database [3]. When user needs to view data from the web page, query interface can be made on the web browser.

These are the functional components of the architecture shown in Figure (3):

- Transformation of spreadsheet data to XML format
- Transformation of XML Data schema to XML User schema
- Mapping the XML Data schema table to the relational tables in SQL Database.
- Publishing data from the SQL database back to user schema XML format
- Transformation of XML Data schema to database Server.

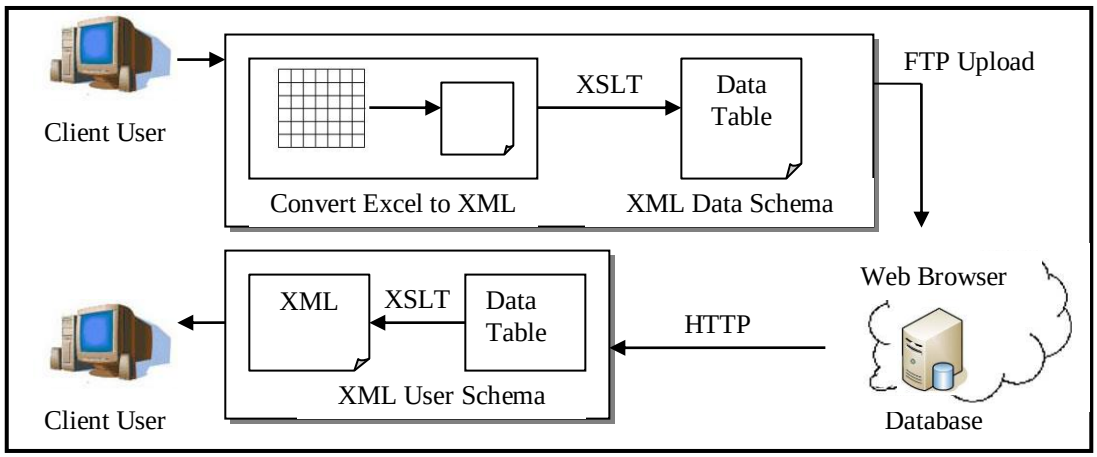


**Figure 3.** System design

- Transformation of XML User schema from Database Server.

### 4.3. Schema Matching

For data entry need to match the data as schema matching one for input data, from different data format to another one destination format, are used between two schemas [2]. As different data format Excel file to XML file data transformation, in storing and retrieving data from database need one schema (data table schema) to convert data type for data set matching, one schema (User Schema) to represent the data to user, having various forms of metadata, of which schema are the most important[4]. Discovering semantic corresponding between different schemas is therefore a critical step in many such applications; A) correct the data transformations from the source schemas to a single target schema; B) to find similar structures across multiple schema for e-business or scientific workflow, to identify semantically correct mappings of messages, often in XML format, from one service or step to the next [2, 1].

### 4.4. Schema Matching Problem

The schema matching problem takes as input two schemas and possibly some auxiliary information, and returns a mapping that identified elements that correspond semantically. This can be very challenging: even schemas for the same entities form different source may have very different structural and naming conventions, and may also use different data models. Similar or even the same labels may be used for schema element having different meanings or having subtly different semantics, due to different in the unit, precision, resolution, aggregation, measurement protocol, etc,;

Given two Schema, Data Table Schema and User Schema create mapping between the schemas;

- Mapping might be directional or symmetric
- Mapping might be the form of a query or it might be a set of data type expression between items in each schema [2].

### 4.5. Schema Matching Example

In schema matching, input data are concentrated on the Excel files which entry by user for data storing and retrieving or representing to user across data table schema and user schema. Firstly, coming input data are matched each field of input record, to determine for their record set what data type are used in data table schema as Figure 4

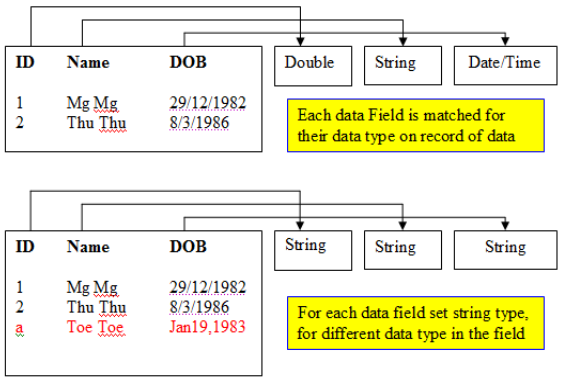


**Figure 4.** Schema matching each field

### 4.6. Data Matching Decision

When data are matched, data table schema set decision on the record of the field that what type of data may be; if one field is used as heterogeneous

data schema set the field of the data into string data type. When matching may be the following items

- Data value
  - May fine common patterns or phrases in data values
- Elements name
- Constraint information
- Structural information
- Domain knowledge : synonym, related term, etc
- Cardinality relationship between elements

For the user schema will not concern for their data type, in this schema all the elements and attributes tab are generated into XML code form on the web or data presenting to the user [2].

### 4.7. Form Design Interface

User can make the data conversion from specific excel (spreadsheet) file to XML file by using XML conversion program at the client and then transform to data table with XSLT and upload the XML data table to the web server to store data to the web database.

In put spreadsheet data Excel file are chosen from the file system then convert file into the XML format, after XML transformation completed data must convert with their data type value into data table to store data into web data base.

After the XML converted into data table then XML file can be stored into database by transferring data from window base or web base application program into web database.

The another servicing is data enquiry on the website which mention the company‘s services and products about the company profile when user needs to data enquiry from browser user need to go to Data Enquiry page with authenticated user name and password.

User can do input the process on the web browser, when user need to do data conversion or data viewing as presentation as data table or XML tab view user can continue what’s user want to do servicing on the web browser such as new data input or data enquiry.

User could be available the data from the web database as data table view with their data type in two list boxes. When user needs to view or generate as XML code, user can get data as XML code from data table to XML view state by passing user schema and data table schema.

## 5. Conclusion

The web technologies, XSLT technologies, UML diagram, IIS Web Servers and ASP.net, XML/ SQL Server and JavaScript programming are being used. The system is implemented as XML converter for client user, built web server and database server for data query process on the web. C# programming is used for excel file to XML file transformation. This system can view and study the overall of the operation of business management.

### 5.1. Evaluation and Further Extension

This system is only one way data access method to the database by uploading the data to web database. Publishing data from database backed to client is XML spreadsheet file. This system could be used in large business organization in transferring business data and financial report of the Microsoft Excel file which most enterprise in social business organization by converting the XML file to the Web Database. As data input Excel spreadsheet must be simple data table, the system will not access complex data set table which mean the excel spreadsheet table in two or more cells or fields would not merge into one cell or field.

The system can further implement as further extension, two way data access the excel file for client users and data set can do as additional extended merging complex data table and cleaning each data table into small one for normalization or large one for consolidation as data mart by manipulating data procedure, reducing or extracting repeated data table or data set.

## 6. References


[1] C Von See, N Keskar, *XSLT Web Developer’s Guide*, Copyright 2002

[2] G Wang, J Goguen, Young-Kwang Nam, Kai Lin, Chris Von See, N Keskar, *Interactive Schema Matching With Semantic Function.*

[3] Ivan Voras *A High Performance Memory Database for Web Application Caches.*

[4] J.D.D.Daniel, *Data Transformation Services: Creating Data Mart by Consolidating Multi-Source Enterprise Operational Data*.

[5] Lawrence Chung ,*Client Server Architecture.*

[6] M Kolsek, *Session Fixation Vulnerability in Web-based Applications.*

[7] R Mohamed, *No Bricks in the Wall – Delivering, Continuing Professional Development Through Web technology*.

[8] T Johnson, *Datasheet Infomatica B2B Data transformation.*

[9] Wei Liu, *Web Data Base Integration*, Sept 2006.