

Automating Exchange of Educational Certificates Using DRESS



By

Umair Anwar

2011-NUST-MS PhD-IT-049

Supervisor

Dr. Sharifullah Khan

Department of Computing

A thesis submitted in partial fulfillment of the requirements for the degree
of Masters of Science in Information Technology (MS IT)

In

School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.

(June 2015)

Approval

It is certified that the contents and form of the thesis entitled “**Automating Exchange of Educational Certificates Using DRESS**” submitted by **Umair Anwar** have been found satisfactory for the requirement of the degree.

Advisor: **Dr. Sharifullah Khan**

Signature: _____

Date: _____

Committee Member 1: **Dr. Hamid Mukhtar**

Signature: _____

Date: _____

Committee Member 2: **Dr. Sarah Shafiq Khan**

Signature: _____

Date: _____

Committee Member 3: **Mr. Mujtaba Haider**

Signature: _____

Date: _____

Abstract

Bologna process aimed to ease the mobility of students across Europe. Previously, some efforts were made in related domains but these were not focused on mobility of the students between institutes. Hence, efforts were put developing standards and proposing suitable architectures that fit all across Europe under the umbrella of Bologna process. Inspired by this, the mobility project tried to gather and reuse all the work previously done in related areas. Mobility reused the vocabularies, ideas and focused on the mobility of students. However, it created an opportunity for a semi-automated mapping tool for mapping proposed standard to the heterogeneous schema of different institutes. It lacks enough vocabulary to cover exchange of information for some educational certificates in Pakistan.

We suggest a prototype infrastructure that provides more control and improves this exchange of information between partaking institutes by covering more detailed vocabulary. The term infrastructure includes both the architecture and our proposed standard (DRESS).

We avail the opportunity that The Mobility Project provided and suggest a mapping tool that semi-automatically create mappings between our standard and the heterogeneous data-sets of different partaking institutes.

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: **Umair Anwar**

Signature: _____

Acknowledgment

All gratefulness and wonderfulness to **ALLAH** and with His blessings, i am in good health and he gave me the ability to complete this work. I pray for more of Allah's blessings. It would not be possible without the efforts that were made by my **parents** for supporting me during this duration and throughout my life.

I am truly thankful to my supervisor **Dr. Sharifullah Khan** for his guidance and who had been a great coach. I want to express my gratitude towards him for believing in me and facilitating me for the completion of this work. I appreciate **Dr. Hamid Mukhtar**, **Dr. Sarah Shafiq Khan** and **Mr. Mujtaba Haider** for suggesting me and for their support.

I am thankful to all my colleagues, co-workers and contributors who were supportive in completing this research.

Umair Anwar

Table of Contents

1	Introduction	1
1.1	Problem Statement	3
1.2	Contributions	3
1.2.1	Motivation	3
1.2.2	Goals	4
1.2.3	Challenges	4
1.3	Thesis Organization	4
2	Literature Review	6
2.1	Bologna Process	6
2.2	Qualifications Exchange Standards	7
2.2.1	European Qualifications Framework	7
2.2.2	Europass	8
2.2.3	Schema for Academia	10
2.2.4	Dublin Core	12
2.3	European Learner Mobility	12
2.3.1	The Mobility Project	12
2.3.2	The REST Mobility	13
2.4	Information Manifold	13
2.5	MAPQFTOOL	14

3	Requirements Analysis	15
3.1	Definitions	15
3.2	Business Process	16
3.2.1	Make Exchange Agreement	16
3.2.2	Find Student Data	17
3.2.3	Exchange a Document Details	17
3.3	Mapping and Web Service Challenge	17
3.4	Software Specifications	18
3.4.1	Functional Requirements	18
3.4.2	Non-functional Requirements	18
4	Architecture & Design	19
5	Implementation	22
5.1	Implementation Stages	22
5.1.1	Defining the Schema	23
5.1.2	Map Generation	23
5.1.3	Deploying DRESS based Web-Service	23
5.1.4	Client	23
5.2	Technologies	23
6	Conclusions	24
6.1	This is the End	24

List of Abbreviations

Abbreviations	Descriptions
DRESS	Document Record Exchange Standard of Students
SCHAC	Schema for Academia
LDAP	Lightweight Directory Access Protocol
FVUSPEC	Finnish Virtual University Specifications
MLO	Meta-data for Learning Opportunities
WSDL	Web Service Description Language
EHEA	European Higher Education Area
ECTS	European Credit Transfer and Accumulation System
EA	Exchange Agreement
NQF	National Qualifications Framework
EQF	European Qualifications Framework

List of Figures

2.1	EQFs against NQFs [3]	8
3.1	Making Exchange Agreement	16
3.2	Finding Student Data	17
3.3	Exchanging a Document Details	18
4.1	Multiple Nodes Exchanging Information Independently in Dis- tributed Architecture [4]	19
4.2	Multiple Nodes Exchanging Information Through a Central Control Authority in Distributed Architecture	20
4.3	Architecture Diagram	21

List of Tables

2.1	Example of SCHAC attribute: schacGender.	11
-----	--	----

Chapter 1

Introduction

With time passing, the exchange of official and legal documents digitally is getting more and more importance. In near future, it will be inevitable to transform our systems to facilitate this change. Same is the case with the exchange of educational certificates between different partaking institutes and authorities. Its importance can be imagined by the number of students and institutes involved.

In 1999, it was decided in Bologna declaration to create European Higher Education Area which facilitate to standardize the exchange of information across Europe. Alone in Europe, more than four thousands partaking institutes with more than two million students were involved in this process in academic year 2009-2010.

Enter Universities and students statistics from Pakistan.

It is a well established fact that more and more systems are digitized every year. The educational institutes are also making their record digital

and this trend is increasing. For example, more and more universities are implementing electronic student information systems to keep student courses and credit record. However, the documents often called degrees/certificates issued by these autonomous bodies are still exchanged in paper form.

This manual approach of exchanging physical documents and re-enter the data again manually in digital or physical form by people is error prone and exhaustive. As a result the local data-sets of these autonomous bodies has different record for the same individual.

This creates an opportunity for a common data-exchange standard which bring a common ground for the exchange of information. To resolve the issues facing in manual exchange of documents, the institutional systems must talk to each other using a common standard. This would make the entire process more dependable and less error prone.

This research digs into the currently adopted solutions, standards and the new research in the area of student mobility and exchange of official documents related to students. It suggests a prototype infrastructure including data format and the architecture to exchange degree records digitally.

Each institutional system is an autonomous body maintaining its data separately. The schema of the data is different in different institutes. For our proposed architecture and standard to work, each institute must implement web-services based on DRESS for the exchange of documents. This approach is very time consuming, inefficient, and error prone. This provides us another opportunity to suggest a tool that maps these different types of

databases schema from different institutes to DRESS.

To handle the different levels of heterogeneity, we came up with a mapping tool which semi-automatically maps institutes data-sets to our proposed standard and creates web-services to access their data.

The rest of the chapter is organized as follows. In Section 1.1, the problem statement is stated. In Section 1.2, thesis contributions are stated. In Section 1.3, we conclude the chapter with an outline for the rest of the thesis.

The remaining chapter is sorted out as this; In Section 1.1, we explain the problem Statement. In Section 1.2, we discuss what this research has contributed. In Section 1.3, we explain how this research is organized into chapters.

1.1 Problem Statement

” ”

1.2 Contributions

Our research work contributes in areas.

All contributions of are summarizing as follows:

1.2.1 Motivation

-
- Finding 2.

1.2.2 Goals

- Propose a generic and extendable standard for the exchange of educational certificates with well defined vocabulary.
- Suggesting a suitable prototype architecture for the exchange of educational certificates such that every data owning entity is the owner of its own data to build trust and have built-in control for authenticity.
- Implementation of a mapping tool to ease mapping of heterogeneous schema with our proposed standard DRESS.

1.2.3 Challenges

- Finding 1.
- Finding 2.

1.3 Thesis Organization

Let us now discuss how we have organized the remaining research in chapters:

In Chapter 2, we go through the literature review. We discuss the work that has already been done for exchanging student information. We look into the already existing standards, a few implementations and related projects. We will go deep in details in this section, to understand the ideas and technologies that are already in use.

In Chapter 3, we discuss the business process in detail using sequence diagrams and define the functional and non-functional requirements.

In Chapter 4, the proposed architecture and design are discussed in detail.

In Chapter 5, our proposed standard "DRESS" is presented. We discuss its vocabulary and the possible values.

In Chapter 6, we discuss the implementation and the test-bed environment we created. In this chapter, we also evaluate whether the goals are achieved.

In Chapter 7, the research work is concluded and possible future work is presented.

Chapter 2

Literature Review

There already exists a few standards and practices identified with exchanging degree or courses record. It is important to go through these, before proceeding onward to the new standard and the architecture we are proposing. We will review what these standards cover and what we can reuse.

2.1 Bologna Process

It intends to make European educational framework of standards engaging different countries in Europe to compare, contrast and make compatible their educational systems. [2]

To improve the mutual recognition of degrees and programs, education ministers from 29 countries signed bologna declaration in 1999. Other partaking countries joined the program later. [8] Bologna process is quite often named as European Higher Education Area (EHEA). EHEA focuses on transferal and convergence adaption by 46 countries. This process benefits Europeans and it has its significance for other educational institutes and

communities. a) The leading role of European institutes, b) the lessons that are learned in the implementation of the framework of standards, and c) the practices adopted guide the educational communities around the world. 2010 was marked as the deadline across Europe for implementing the agreed specifications. [?]

To meet the 2010 deadline, Spain started to implement the convergence of undergraduate engineering degrees that conformed EHEA in 2008. This standardization provided some opportunities for mobility and unified measurements. [?]

2.2 Qualifications Exchange Standards

2.2.1 European Qualifications Framework

EQF is an agreed reference framework that helps participating countries to compare national qualifications and make them more clear, readable and understandable across Europe. The point is to advance mobility of workers and learners. This was settled upon by European universities in 2008 to relate their national qualifications to EQF. The new qualifications from 2012 carry a reference to suitable EQF level.

EQF comprises of eight reference levels, each showing what a learner knows and has the capacity to understand it. National qualifications of the partaking countries identify and relate with these eight levels ranging from basic (level 1) to advanced (level 8) as shown in figure 2.1. This simplifies qualification comparison in partaking countries supporting mobility of learners and empowering them to not repeat what they have already learned.

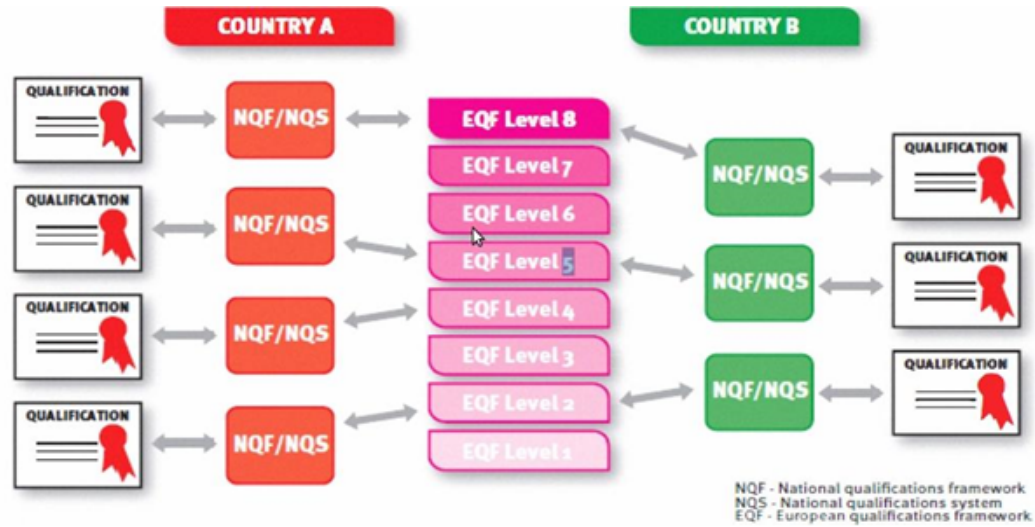


Figure 2.1: EQFs against NQFs [3]

EQF concentrates on learning results as opposed to concentrating on learning inputs. It covers all types of education including professional, vocational and school education. It tries to validate formal and in addition informal education.

2.2.2 Europass

Collection of five documents which intend to ease mobility when seeking employment across Europe. These include the Curriculum Vitae, the Language Passport, the Mobility, the Diploma Supplement, and the Certificate Supplement. One can fill himself the Curriculum Vitae, and the Language Passport but the rest of the documents are issued by the related authorities. It follows a standard template format system, a layout. Same format helps to achieve neutrality and transparency while presenting one's skills.

The motto as mentioned on the Europass website's homepage is as follows; "Five documents to make your skills and qualifications clearly and easily understood in Europe"

Europass has defined XML schemas for CV and Language Passport. The documents can be exported in XML format when created on Europass. These exported XML documents can be imported to Europass and converted to HTML, PDF, Microsoft Word or ODT templates.

Europass specifies JSON schema according to Internet Engineering Task Force's JSON specifications draft. The europass JSON vocabulary is close and similar to europass XML schema. The JSON objects for europass documents (CV and Language Passport) can be validated using Europass JSON validator.

All these documents have some common XML schema attributes which describe document type, printed preferences.

Europass does not explain details related to degrees or educational certificates in XML certificate.

Europass Curriculum Vitae

Europass Curriculum Vitae (ECV) is a template which one can create online and it can be exported in xml format. The ECV XML schema contains vocabularies related to document type, printing preferences, personal details, contact details, skills, and educational degrees and institutes. The XML vocabulary related to degree details is very little only to cover the scope of

a CV.

Europass Language Passport

Europass Language Passport (ELP) is a template. One can create it online and export it in europass xml format. It contains XML vocabulary related to language skills and the scale of six values to score proficiency.

2.2.3 Schema for Academia

The need for the inter-exchange of information between institutes across Europe has highlighted the importance of common attributes for this exchange to take place. Schema for Academia (SCHAC) is the result of the attribute coordination between different institutes. It plans to define and advance common attributes in the field of higher education to encourage data-exchange between institutions. It does not plan to supplement different the national schemas, rather it suggests a coordinated framework on top of the different national schemas.

Schema for Academia (SCHAC) describes vocabulary related degrees and courses. The schema is not technology dependent and written for LDAP (Lightweight Directory Access Protocol). It aims at promoting a common framework to inter-exchange data between educational institutes. It defines attributes that describe individuals and their LDAP profile.

It is a collection of schemas which can be classified into following categories;

- Personal Characteristics

- Location Information
- Student Information
- Employee Information
- Linkage Identifiers
- Administration Information
- Confidentiality (Visibility)
- Authorization, Entitlements
- Group-related Attributes

SCHAC has a clearly defined meta-information for defining an attribute. To discuss the ideas used by SCHAC, let us look in detail the schacGender attribute as an example shown in Table 2.1. It uses ISO-standards wherever possible. In schacGender attribute, it is using values from ISO-5218. However, it lacks hierarchy which prevents from reuse-ability and it can be considered its disadvantage.

Table 2.1: Example of SCHAC attribute: schacGender.

Name	schacGender
Description	Male or Female, specify the legal gender
Format	0 - Not Known, 1 - Male, 2 - Female, 9 - Not Specified
Values	Single
References	ISO-5218
Example	schacGender = 1

Surname issue.

It is also worth mentioning that SCHAC has a category student information for curriculum, major and degree but no attribute is defined. This is because SCHAC is not completed till now and it is in progress.

2.2.4 Dublin Core

The Dublin core is a simple meta-data standard consisting of set of elements to describe information resources on the network. There are two type of elements; simple and qualifiers. It has 15 simple elements and qualifiers which have additional three elements namely Audience, Provenance and RightsHolder. Qualifiers help in resource discovery.

2.3 European Learner Mobility

Some related work has been done recently and systems have been proposed based on the above mentioned standards. These are "The Mobility Project" and "The REST Mobility" projects.

2.3.1 The Mobility Project

It aimed to provide a platform and infrastructure for exchange of electronic data exchange between educational institutes. Infrastructure includes data format, architecture and the prototype software. The system will be called The Mobility later in this paper.

The Mobility is peer to peer like architecture. Nodes exchange data using SOAP base web service. Other web services like XML-RPC and REST were not used due to their limitations. XML-RPC not have developer defined data-types and character set. REST does not imposes a standard specification,

instead it follows set of rules and is used for speedy development of web service interface.

The nodes represented the universities, and their number tends to change. So there was a need for system to maintain this record and UDDI was used. He did not recommend the central or delegated private registry instead gave advantages and disadvantages of both. Central single registry has all information at one place but also it a single point of failure.

The software has two transport modules and each have web interface.

Nagrozki proposed a new standard, defined its vocabulary re-using ideas taken from SCHAC to leverage ISO and RFC rules. Some like grade, credits were taken in inspiration from Eropass Mobility.

Although The Mobility project was started by MUCI and CINECA, two European Higher Education Consortia. Many universities consortia, individual universities and companies joined in later on.

2.3.2 The REST Mobility

This is alternative implementation of The Mobility. Nagrozki's system used SOAP web service for data exchange. Karol created a RESTful implementation of the Mobility. The Mobility lacked data model. In The REST Mobility a data model is proposed since REST is resourceful. The model proposed not represents or intends to be a standard.

2.4 Information Manifold

Providing a uniform interface for querying data from many sources is the aim of Information Manifold. It enables a simple user to not worry about locating sources and manually combining results. This leads to concept of

Deep Web. Data integration systems give users a common global schema called mediated schema for posing queries. To answer these queries semantic relationships called mappings are needed between mediated schema and the sources schema.

2.5 MAPQFTOOL

This tool helps comparing National Qualification Frameworks against European Qualifications Framework in Europe. This automates the process of creating mappings between these frameworks and stores the mappings in the database.

Chapter 3

Requirements Analysis

3.1 Definitions

We define the basic terms that are used in exchange of documents. It is necessary to understand these before we go through the requirements.

1. Exchange Agreement:

It is an understanding between partaking institutes, between the requester and the provider, for exchanging the educational certificates. This agreement comprises of;

- web-service access point
- authentication credentials

2. Requester:

The partaking institute asking for the exchange.

3. Provider:

The partaking institute providing the details.

4. Coordinator:

A person responsible for signing and exchange of agreements between institutes.

3.2 Business Process

This section describes the business that is involved in the execution of exchange of different educational institutes. The process is explained using the sequence diagrams.

3.2.1 Make Exchange Agreement

For two universities to exchange data, they have to create an exchange agreement first. The agreement will have the web-service access point and exchange secrets. These details are used for requesting exchange and for authenticating the requester.

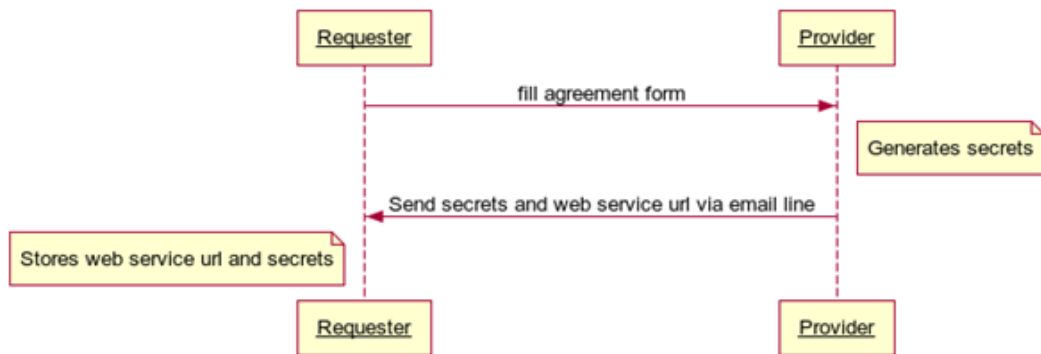


Figure 3.1: Making Exchange Agreement

After the implementation of web-service by the provider. The requester fills a form and asks for the credentials. The emails the access point and exchange secrets to the Coordinator.

3.2.2 Find Student Data

To find a student record, the requester asks a provider from the agreed providers list for a student record. The provider sends back list of documents associated with the student.

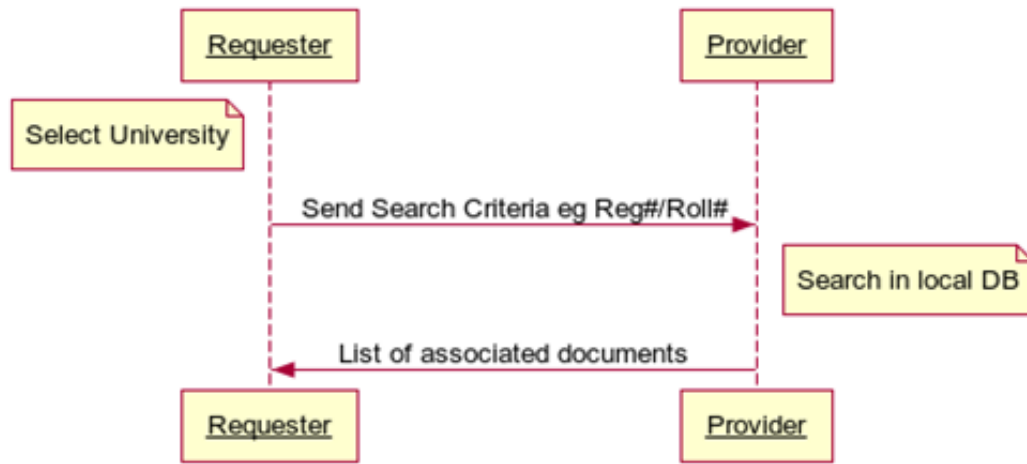


Figure 3.2: Finding Student Data

3.2.3 Exchange a Document Details

To exchange a document details, the requester asks a provider with search criteria and document type. The provider sends back the document details using DRESS.

3.3 Mapping and Web Service Challenge

It is very important to understand that to implement the scenarios mentioned in section 3.2, every provider must implement a web-service based on DRESS. This is a challenging task since every partaking institute (provider) has different schemas and different database management systems. This schema is

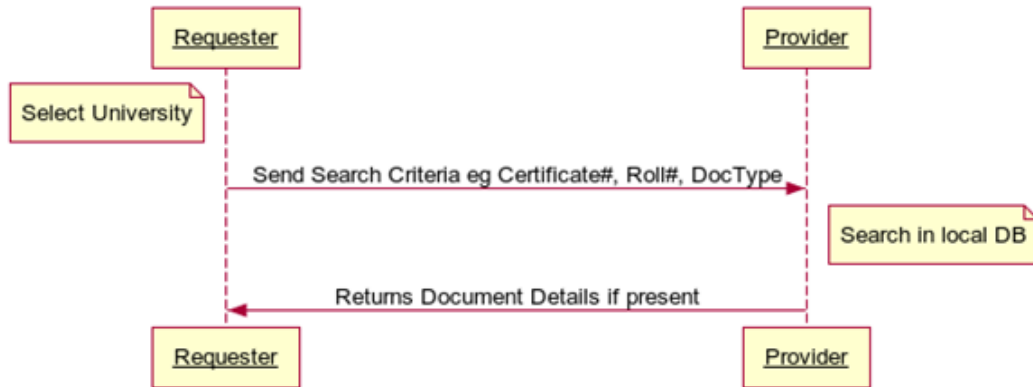


Figure 3.3: Exchanging a Document Details

required to be mapped to DRESS. This must be served to the requester using a web-service. This creates an opportunity for a semi-automated tool which maps the schemas to DRESS and generate a web-service automatically.

3.4 Software Specifications

Based on the business process and the challenges we discussed, functional and non-functional requirements are;

3.4.1 Functional Requirements

3.4.2 Non-functional Requirements

Chapter 4

Architecture & Design

From the requirements analysis, we suggest student exchange system should have distributed architecture. As partaking institutes are autonomous and themselves maintain the its own data. It signs agreements independently for exchanging data with other universities. Each can be a requester plus a provider of data. The circles/nodes in the figure below represent universities. The arrows represent exchange of data.

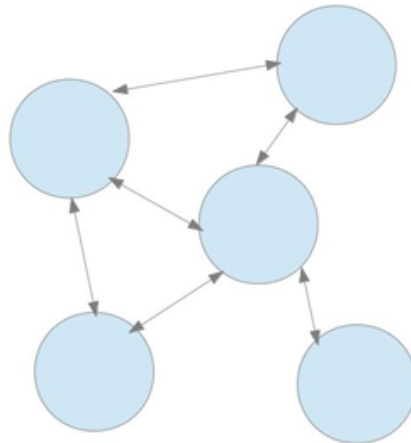


Figure 4.1: Multiple Nodes Exchanging Information Independently in Distributed Architecture [4]

This peer to peer like distribute architecture has benefits over adding a middle agent or central server in the system.

1. Avoidance from single point of failure.
2. Lesser load.
3. Each university having control over its own data and thus building trust in the system.

We suggest student exchange system will have distributed architecture. Each university has its own data and signs agreements with HEC for exchanging data with other universities. Each can be a requester plus a provider of data. The circles/nodes in the figure below represent universities

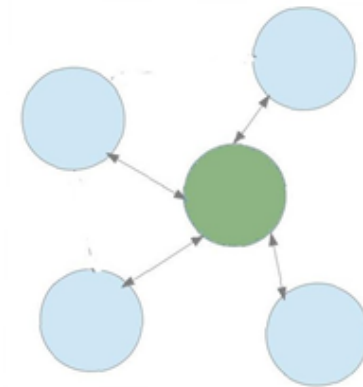


Figure 4.2: Multiple Nodes Exchanging Information Through a Central Control Authority in Distributed Architecture

There are some choices to be made at this point. We will be using web services for exchanging data as they provide a high abstraction from network issues and use well known standards like XML over HTTP. There are some XML based data exchange protocols on web. These are XML-RPC, SOAP, and REST.

The nodes will exchange data using SOAP based web service in our system. We chose SOAP as it forces to follow a formal standard and supports developer defined data types.

The number of universities can increase when agreements are signed with new universities for exchange data. The web service URLs need to be saved so that requester can retrieve this URL and request that university. This can be achieved by developing a custom system "URL registry" for saving web services URLs. Now we have to make a choice. URL registry can be global or each requesting node can have its own private URL registry. We will use private registry to avoid single point of failure and to minimize load.

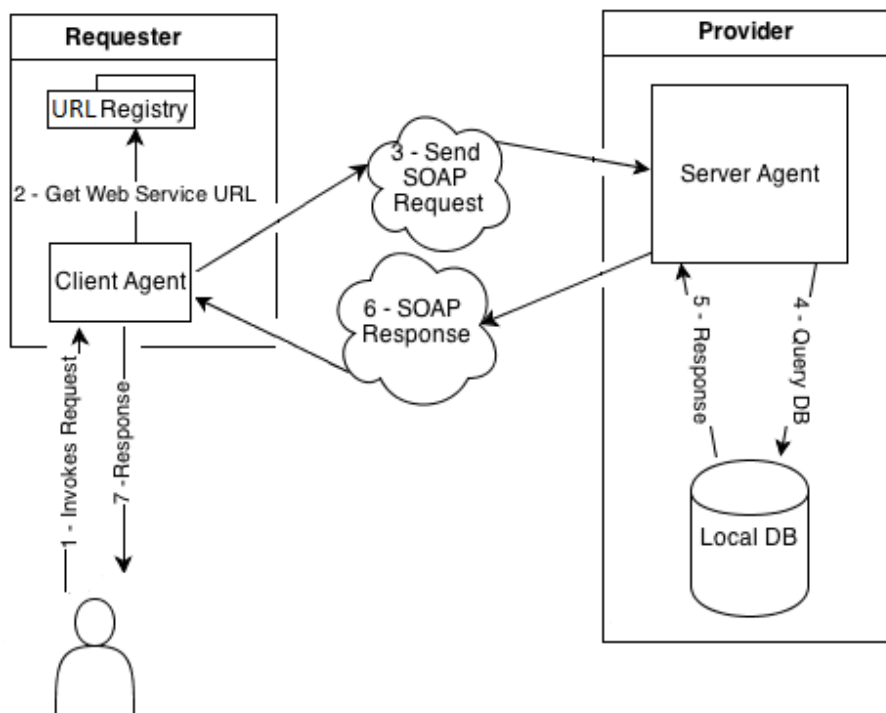


Figure 4.3: Architecture Diagram

Chapter 5

Implementation

Automating exchange of educational certificates using DRESS can be divided into four stages of implementation. The very first stage begins with defining the schema, that is our proposed standard and writing an xsd. The second stage involves implementation of a mapping tool that generates mapping and filters over heterogeneous schemas. The third stage is the development of a web-service that uses the generated mappings and filters to serve requests using DRESS. In this stage we implement a soap service describing the possible operations in a WSDL that satisfy the business processes described in chapter. The fifth stage is the last stage which includes a GUI interface for a Client which is a central authority to send and retrieve information from different institutes/nodes.

5.1 Implementation Stages

Automating exchange of educational certificates using DRESS can be divided into four stages of implementation. The very first stage begins with defining the schema, that is our proposed standard and writing an xsd. The second

stage involves implementation of a mapping tool that generates mapping and filters over heterogeneous schemas. The third stage is the development of a web-service that uses the generated mappings and filters to serve requests using DRESS. In this stage we implement a soap service describing the possible operations in a WSDL that satisfy the business processes described in chapter. The fifth stage is the last stage which includes a GUI interface for a Client which is a central authority to send and retrieve information from different institutes/nodes.

5.1.1 Defining the Schema

5.1.2 Map Generation

5.1.3 Deploying DRESS based Web-Service

5.1.4 Client

5.2 Technologies

Chapter 6

Conclusions

6.1 This is the End

In this research we reported the design and implementation of.

Bibliography

- [1] M. Ali, H. Qureshi, and M. S. Akhtar (2013), *Analysis of growth in Students Intake and Degree Awarding Contribution: A Comparison of Stanford and MIT*, MLDM 2013: International Conference on Machine Learning and Data Mining, in press.
- [2] <http://www.ehea.info/Uploads/Irina/Bologna>
- [3] P. Pouyioutas, H. Gjermundrod, and M. Michael, "MAPQFTOOL: A software tool to support national qualifications frameworks," Information Society (i-Society), 2011 International Conference on, pp. 198-203, Jun. 2011.
- [4] Rafal Nagrodzki, "The Mobility Project," Institute of Informatics, University of Warsaw, Warsaw, Master's Thesis 2009.
- [5] A. Duran, Y.B. Moon, and E. Giraldo, "Work in progress - the European Higher Education Area ("Bologna process") in Engineering Education in Spain," Frontiers in Education Conference, 2009. FIE '09. 39th IEEE, pp. 1,2, Oct. 2009.
- [6] Karol Kanski, "Integration of Services in the Mobility Project," Institute of Informatics, University of Warsaw, Warsaw, Master's Thesis 2011.

- [7] Alon Halevy, Anand Rajaraman, and Joann Ordille, "Data integration: the teenage years," In Proceedings of the 32nd international conference on Very large data bases (VLDB '06), pp. 9-16, 2006.
- [8] Szentirmai, L.; Radacs, L., "Improvement of academic and research standards of higher engineering education in light of Bologna process," Emerging eLearning Technologies & Applications (ICETA), 2012 IEEE 10th International Conference on , vol., no., pp.381,386, 8-9 Nov. 2012
- [9] <https://wiki.refeds.org/download/attachments/1606048/SCHAC%2B1.5.0.pdf?version=3&modificationDate=1429195142624&api=v2>