OPEN ACCESS
EURASIA Journal of Mathematics Science and Technology Education
ISSN: 1305-8223 (online) 1305-8215 (print)
2017 13(7):3969-3991
DOI10.12973/eurasia.2017.00767a

E-Assessment Data Compatibility Resolution Methodology with Bidirectional Data Transformation

Kaleem Razzaq Malik
COMSATS Institute of Information Technology, & University of Engineering and Technology, Pakistan
Tauqir Ahmad
University of Engineering and Technology, Lahore, Pakistan

Received 11 November 2016 ▪ Revised 28 December 2016 ▪ Accepted 11 May 2017

ABSTRACT
Electronic Assessment (E-Assessment) also known as computer aided assessment for the purposes involving diagnostic, formative or summative examining using data analysis. Digital assessments come commonly from social, academic, and adaptive learning in machine readable forms to deliver the machine scoring function. To achieve real-time and smart e-assessment, data modeling needs dramatic improvements at the level of representation which will improve examinees to gain prompt response instantly after attempting exams. Whereas, computer based inference to gain intelligence in assessing results through computations is becoming a useful feature in todays' testing systems. Induction of rule base linked data is desired to be reformed from the old tradition data model found either in spread sheet or relational database used for data storage. These data forms are essential to be converted into semantical annotated form to support Artificial Intelligence. This can be done with the use of Semantic Web data model Resource Description Framework (RDF) built-up using hierarchal and linked data representation. Updating assessment source data later for results is one of the hardest problem of all viabilities in traditional and semantically augmented systems when combined for evaluating. This study purposes a methodology of bidirectional data transformation back and forth from Relational Database (RDB) and RDF. A case study representing qualitative analysis of transforming student's results information into RDF store reforming data as ready to be analyzed. At the end of this study outcomes show how data updating becomes feasible by following proposed data transformation procedure.

Keywords: e-assessment, semantic annotation, knowledge management

Correspondence: Kaleem Razzaq Malik, *University of Engineering and Technology, Lahore, Pakistan & COMSATS Institute of Information Technology, Sahiwal, Pakistan*. Phone: +92-333-761-6116
Email: krmalik@gmail.com

State of the literature

- A few studies focused on data transformation and its integrated representation in scientific teaching and learning contexts.
- There are few suitable studies to assess the performance in problems like student grades, comparing different formats of presentation of data for the assessment.
- Qualitative small case studies show that the diverse data formats can be used for the effective and automated analysis of students' assessment.

Contribution of this paper to the literature

- E-assessment data compatibility resolution methodology
- This paper presents a novel approach to measure and analyze the students' results using semantically annotated data and transformation
- An algorithmic approach for the patterns finding from the assessment data for the student and finding the patterns for passed and failed students

INTRODUCTION

Data modeling and analysis is the most active field of research in current era of information sciences and technologies under the umbrella of data sciences. Data is further formulated to gain Computer Assisted Assessment (CAA) for diagnostic data influence, cumulative results, and innovative outcomes. Nowadays, web is used as common interface for computational and distributed system services for assessment based utilization of scoring data. Assessment using web as a resource for communication requires powerful and supportive representation of data. Well known and commonly used data representations of data available on web are RDB, XML and RDF (Das, Sundara, & Cyganiak, 2012). RDB is the mechanism used for backend storage and querying of data following relational data model standards, constraints and rules. XML is the common tags' data representation of XML document concerned with data objects. And RDF has embedded semantics and hierarchal attachment in the form of data linkage (AlObaidi, Mahmood, & Sabra, 2016; Musen, 2015).

Combined with information technology growth e-assessment made its way towards maturing and extensive acceptance in the field of health and education (Umair, Björklund, & Petersen, 2015). Assessment based on computers provides common benefit in reduction of paper use with fast data collection and evaluation. CAA comes with scoring mechanism, a function which helps examinees to get fast feedback and result as soon as they attempt questionnaire. This not only provides fast feedback but also provides shortening teacher's efforts, e-learning assistance, and a mechanism for self-assessment (Mettiäinen, 2015). In other benefits include utilizing video oriented simulations for representing testing items. These videos based simulations are used in improving e-Learning objectives in assisting teachers' teaching and students' learning goals and assessment which relates this electronic teaching closer to conversional teaching context and circumstances.

E-Assessment

E-assessment also known with other names like computer assisted assessment, online assessment, and computer aided assessment, help formalizing and summarizing the outcomes using electronic devices built-up for computation in the fields of health, education, psychiatric, and psychologies. Such assessments can support educational system in many ways. E-assessment is considered a huge change in the field of educational assessment when considering traditional assessment methods (Kardan, Sani, & Modaberi, 2016). At large scale enabling devices and hardware for conducting electronic examination by huge number of students seems hectic due to its security threats towards educational dishonesty or plagiarism issues. Many methods and techniques are introduced including virtual exam conduction, virtual e-learning, blended e-learning, and blended e-assessment etc. to overcome majority of the issues concerning assessment based e-learning. E-assessment comes from computations on the data collected for scoring and analyzing results and outcomes using computers (Umair et al., 2015). Data modeling and data fetching rational as ready to be computed is a considerably wanted feature embedded into e-assessment systems.

Role of Distributed Data

Emerging information with different data patterns and forms decreases devices and databases capacity to prepare for information investigation because of lesser information similarity. A data component, not to be disregarded, is that dominant part of the system is yet taking a shot at customary devices and databases. Change in the information, system, and preparing models are craved to be focused on conquering these similarity issues by adopting data transformation mechanism. Solutions do not cover to overcome all issues related to smooth utilization of data transformation features. This study proposes bidirectional data transformation methodology with common language data representation to overcome issues of update concerning data and metadata (Stantchev, Prieto-González, & Tamm, 2015).

Unorganized Web Data and Understandability Issues

It again feels like an intimidating task to clarify why an association ought to grasp semantic data representation as innovation instead of keep relying on relational databases data representation (Chung, Niemi, & Bewley, 2003). In this part, formal establishments of different representations, and need to assess them against five criteria: understandability of the models, interoperability, accessibility of data for recovery, capacity to derive new data which is provably right, and the capacity of the information to address psyches and machines alike with a mutual semantics. The idea of machine-reasonable reports does not infer some mystical manmade brainpower which permits machines to grasp human mumblings (Agus, Penna, Peró-Cebollero, & Guàrdia-Olmos, 2016). It just demonstrates a machine's capacity to take care of an all-around characterized issue by performing very much characterized operations on existing all around characterized information (Kardan et al., 2016). Rather than requesting that machines comprehend individuals' dialect; it includes requesting that individuals attempt (Agus et al., 2016).

Unorganized to Organized Web Using Semantic Web

Semantic Web introduced RDF as a data model having capability of capturing hierarchal relationship among different resources (Musen, 2015). This made system to cover the gap of intelligence which was long missing when using data for analysis and results. Systems enhanced to handle data modeled in RDF became able to capture reasoning with the help of available data. A data comprised of multiple triples, when combined built-up a statement. A triple is combination of Subject, Predicate and Object. Even though it's easy to characterize, RDF at the level with the force of a semantic web will be finished dialect, fit for communicating conundrum and repetition, and in which it will be conceivable to expression addresses whose answers would to a machine requires a pursuit of the whole web and an incomprehensible measure of time to determine. This ought not dissuade us from making the dialect finish. Each mechanical RDF application will utilize a construction to confine its utilization of RDF to an intentionally restricted dialect. In any case, when connections are made between the RDF webs, the outcome will be a declaration of a tremendous measure of data. Plainly because the Semantic Web must have the capacity to incorporate a wide range of information to speak to the world, that the dialect itself must be completely expressive (Das et al., 2012).

Role of Semantic Web in Assessment

Semantic Web based platform are digging in their place increasingly for producing, sorting out and marking e-learning contents for e-assessment (Chang, 2001; Liu & Khine, 2016). Through this research paper we introduce and extend the way to work with automated creation of CAA using Semantic Web based data model (Queirós & Leal, 2012). We improve the work previously developed in the way into two vital orders: to start with, we add new RDF elements (comments) as composite identifiers, to the metadata used for analyzing data; secondly, we enhanced semantic interpretation for mapping between the relational ontology and the targeted information. The semantic reading is based on the scientific classification of educational objects, further can be used similarly with other pedagogical principle related to data content and design (Liu & Khine, 2016).

At the point when a student experiences trouble in capturing topic theme that requires guidance of teacher. It will also help to cope with the week areas to overcome such problems. Student's learning curve using results for each subject can help to produce statistical assessment over teaching methodology and student attentiveness. It can be used to view and observe the learning problems, which is behaving as obstacles towards better education. A trained system can help in fast and real time assessment more effectively. Our proposed mechanism built on semantic annotation can provide new ways to overcome the data compatibility issue for assessment. This will help to design better software and Machine Learning strategies to be used to encourage the learning procedure and make it viable for teachers and students.

This paper is further divided into section like literature review representing current trend towards assessments using Semantic Web and different data mapping tools techniques available for transformation. Further in coming section is of presenting the proposed methodology for transformation with algorithms and mathematical modeling. Afterwards is a case study of students' assessment transformed into semantical data linkage where with examples obtained from functions in practice. The improves application area for automated assessment for e-learning, and specifically, development of intelligent CAA systems, yet the concepts can get generalized in the relation to ontology creation and evaluation.

LITERATURE REVIEW

The increasing interest in e-Learning by the development of formalization of ontologies based on learning objects, processes, and designs (Knight, Gasevic, & Richards, 2006; Sicilia & Barriocanal, 2005). Ontologies development for learning assessment is less researched area and only recently, few introduced techniques for ontology oriented assessment as; user specific assessment for authors with the value embedded for assessment ontologies (Chung et al., 2003). To represent adaptive learning on the bases of slide shows and objective tests (Holohan, Melia, McMullen, & Pahl, 2005) has worked on semi-automated learning objects. This work was further extended (Holohan, Melia, McMullen, & Pahl, 2006) by introducing domain specific querying based on relational databases for dynamic problem generation. A technique for creating feedback using Semantic Web based approach for generating question oriented test assessment (del Mar Sánchez-Vera, Fernández-Breis, Castellanos-Nieves, Frutos-Morales, & Prendes-Espinosa, 2012). Other assessments methods recently introduced include personalized self-assessment on feedback (Belcadhi, 2016), implicit leaners assessment on the bases of relevance (Kardan et al., 2016), and automated distributed knowledge assessment (Stantchev et al., 2015), these are built-up on Semantic Web technologies. Moreover, work on Heterogenous data appearing in for assessment is still lacks to its full extents (Musen, 2015). This study has been focused on provides mapping and transformation mechanism to solve many issues involving real-time assessment and inference occur.

History of mapping language timeline wise starts from 2003 to 2012 including start of the art languages and platforms for data transformation between RDB and RDF (as shown in Figure 1). These mapping languages and platforms are like D2R, R_2O, D2RQ, Relational.OWL (de Laborda & Conrad, 2005), Virtuoso RDF Views, DB2OWL, RDBtoOnto, Triplify, Ultrawrap, R3M, D2RQ/Update and R2RML. They have introduced mapping in the form of direct, indirect and language based approaches for RDB and RDF, and even some have shown partial bidirectional data transformation using query oriented approach. Now to see which language or platform provided better solution for data transformation without losing any necessary information about data is done using firstly by defining them separately and briefly. And then by comparing their supported features and capabilities for data transformation process.

Among mapping languages given in Table 1 brief introduction of each start with Direct Mapping (Arenas, Bertails, Prud'hommeaux, & Sequeda, 2013) which provides a direct mechanism to transform RDBs into Semantic Web by mapping table as class and field to a properties. Whereas, URIs are generated automatically following RDB schema and data.

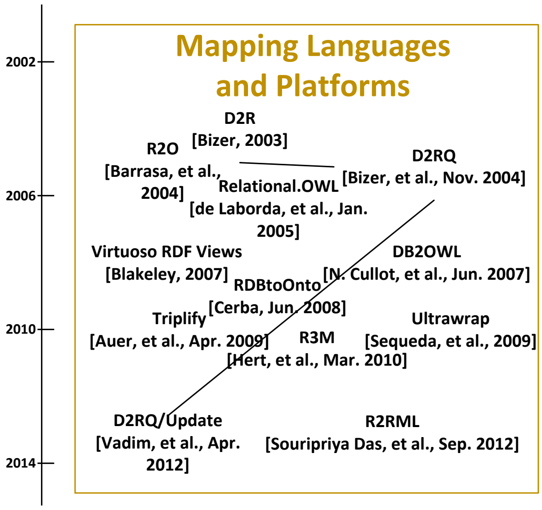


Figure 1. Mapping languages evolution history timeline oriented tree graph

R_2O (Barrasa Rodríguez, Corcho, & Gómez-Pérez, 2004) is aimed to cope complex mapping and low similarities among RDB to ontologies with schematic implementation either found in RDFS or OWL. In Relational.OWL (de Laborda & Conrad, 2005), OWL Full based ontology representation to describe the schema and data of a RDB. Openlink Software a server named Virtuoso Universal Server provides RDF views (Blakeley, 2007) to represent relational data on the Semantic Web. A SQL SELECT query is used to translate dataset found in database into a set of triples. Whereas, SQL DDL forms a syntax level aspect of view. D2RQ (Bizer & Seaborne, 2004; Cyganiak, Bizer, Garbers, Maresch, & Becker, 2012) is used to transform RDB based data into virtual RDF graphs. Where access to this this Semantic Web data is through SPARQL queries and Linked Data. It is the descendant to the XML oriented D2R mapping. Triplify (Auer, Dietzold, Lehmann, Hellmann, & Aumueller, 2009) is a query oriented transformation of RDB into RDF statements to distribute Linked Data from RDBs. Triplify transformation is developed using PHP scripts/code. R2RML (Das et al., 2012) a mapping language made a recommendation by W3C to make a standardized approach for RDB to RDF transformation. OntoAccess mediation platform based transformation language known as R3M (Hert, Reif, & Gall, 2010). As an update, attentive transformation language, it enables providing partial bidirectional query oriented RDF oriented contact to the RDB.

In Table 1, features like relation to class, update, record URI, data reuse, datatypes, integrity constraints, write support, data transformation, query base transformation, and bidirectional transformation are mapped. In the given table, different symbols like tick, and cross marks with or without box are used where tick mark represents supported feature, cross represents not support feature, tick mark within a box represents partially supporting feature, and cross within box represents unknown. It clear that bidirectional transformation, update, and write support features necessarily required for updating a data and schema of either data model of RDB or RDF are in R3M only which is again query oriented and partially supported (Michel, Montagnat, & Faron-Zucker, 2014). No other approach for transformation provides required skill set to accomplish bidirectional data transformation with improved capability and capacity to solve update issue. Where update issue is about a change introduced in data either available at RDF or RDB should also be updated only at the point where it appears in RDF if data is changed in RDB whereas in RDB if data is changed in RDF. This study works its way in resolving the issue of update by introducing a mapping mechanism in intermediate common form of data gained through bidirectional data transformation.

MATERIAL AND METHODS

Obviously, since the issue of taking care of and dissecting vast scale information has been around for a considerable length of time, it is not shocking that few conventional but rather proficient techniques displayed in the past might be utilized to comprehend or alleviate the issues of taking care of the huge information issue. These techniques can be found in some past information mining concentrates, for example, arbitrary examining, information buildup, isolate and vanquish, and incremental learning. Among them, a conceivable approach to take care of the huge information issue of assessment is to have data procure just

Table 1. Feature wise Comparison between mapping languages and platforms

Features	R2O (2002)	D2RQ (2004)	Relational.OWL (2005)	Virtuoso (2007)	Triplify (2009)	R3M (2010)	R2RML (2012)	Direct Mapping (2012)	D2RQ/Update (2012)
Relation to Class	✓	✓	☑	✓	✓	☑	✓	☑	✓
Update	✗	✗	✗	✗	✗	☑	✗	✗	✗
Record URI	✓	✓	✗	✓	✓	✓	✓	✗	✓
Data Reuse	✗	✗	✗	✗	✗	✗	✗	✗	✗
Datatypes	✓	✓	✓	✓	✓	✓	✓	✗	✓
Integrity Constraints	☑	☑	☑	☑	☑	✓	☑	✗	☑
Write Support	✗	✗	✓	✗	✗	✓	✗	✓	☑
Data Transformation	✓	✓	☒	☒	✓	✓	✓	✓	✓
Query base Transformation	✓	✓	✗	✓	✓	✓	✓	✗	✓
Bidirectional Transformation	✗	✗	✗	✗	✗	☑	✗	✗	✗

Note. Symbol ✓ *represents supported features, symbol* ✗ *represents not supported features, symbol* ☑ *represents partially supported features, symbol* ☒ *represents unknown features.*

the fascinating information rather than all the information. One of the agent research patterns has been on diminishing the many-sided quality of information. One instinctive system is to utilize the important segment examination or other measurement decrease strategies to lessen the quantity of elements of the information. As of late, another promising pattern called design decrease, which depends on an alternate thought. Contrasted this innovation is gone for decreasing the number of examples rather than the extent of measurements amid the union procedure. Therefore, it can likewise be utilized to diminish the multifaceted nature of information. Unique in relation to these strategies, some encouraging headings for taking care of the enormous information issue as of late have been highlight choice, dispersed figuring, and distributed computing.

Assessment of data coming from any data providers specially from Web either centralized or distributed can be transformed into common data form to become machine readable. Whereas, enriching data to have semantics embedded into it using RDF data model (AlObaidi et al., 2016; Dafli et al., 2015). Architecture of the proposed methodology where data coming from users and passing through bidirectional data transformation process making data viable for assessment and translation for all system back and forth for real time computational capabilities as shown in Figure 2.

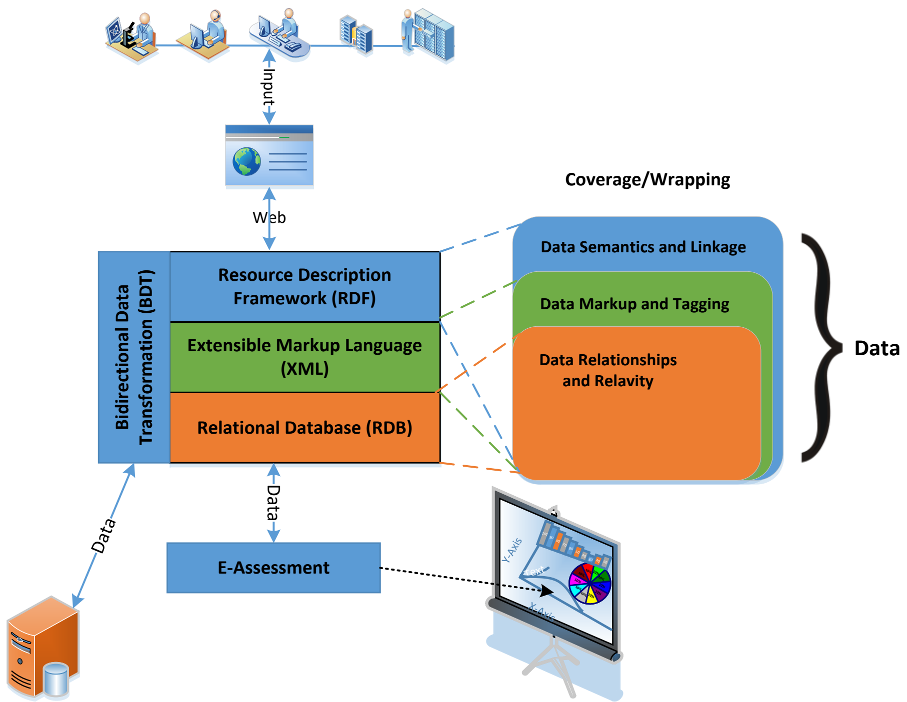


Figure 2. E-Assessment bidirectional data transformation architecture

Now by looking at the transformation process working in detail (as shown in Figure 3) represents that web is used as source of data input for assessment. Then each form of data is sent for the assessment classification by passing through analyzing the type of assessment either being monitored or calculated. Then generated data is sent to RDB for storage and querying. Which is still not capable to support intelligence oriented data assessment. Which requires data transformation into semantical annotation format. For such transformation to happen data is firstly transformed into XML data form which supports full customization and reusability of data in other applications. This data is now becoming highly feasible to be get transformed into RDF. Whereas, due to XML format the data is translated containing information of data and metadata making bidirectional data transformation possible through mapping.

In Table 2, Semantic Annotations Alternative used in RDB, XML and RDF. These comparisons can be useful for better understanding of the algorithms for transformation procedure as presented in this section of the paper.

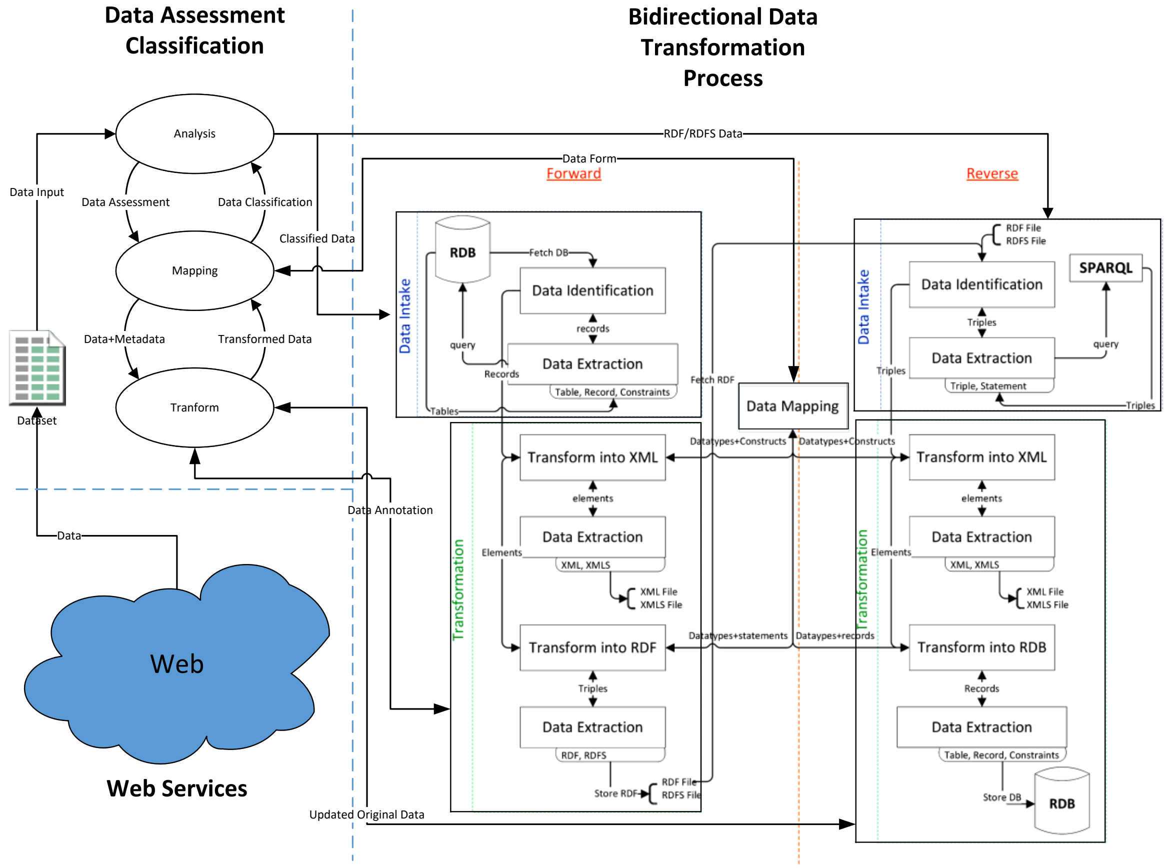


Figure 3. Complete Bidirectional data transformation procedure for e-assessment

Table 2. Semantic Annotations Alternative used in RDB, XML and RDF

Concepts	Relational Schema	XML Schema	RDFS
Table	Table_Name	Complex type element	rdfs:Class
Field	Field_Name	Simple Element	rdf:Property
Cardinality	Min	xs:restriction xs:minLength or xs:minInclusive	owl:restriction owl:minCardinality
Cardinality	Max	xs:restriction xs:minLength or xs:minInclusive	owl:restriction owl:maxCardinality
Referencing	Ref_Key_Field Ref_Key_Table	xs:keyref xs:selector and xs:field	rdfs:domain rdfs:isDefinedBy
Primary Key	Key_Field	xs:key xs:selector and xs:field	rdfs:subPropertyOf rdfs:isDefinedBy
Composite key	Key_Fields	xs:key xs:selector and xs:field	(P) Embedded in rdfs:subPropertyOf
Data type	Datatype	xs:restriction base	rdfs:range

Data Transformation Algorithms

Given first algorithm represents an effective method to resolve transformation from RDBS into XMLS. Initially, this algorithm take input from given DB and scans for available relations and meta data on relations. Beginning from taking input from RDB computation starts and eventually continues by producing resulting schema for XML to be built on. The transition from relations to tags keeps on happening depending on input. Typically, by incorporation from W3C rules for XML based tagging along with mapping rules defined in this study output is reform from each tuple. Algorithm can be divided into sections like undertaking of RDB relation, defining element tag for each relation, grabbing and keeping track of each information of all tuples concerning, and tracking keys and reference keys used for unique identification for integrity constraints in a relation and then recoding them in separate tags.

Algorithm: Transformation from RDB to XML Schema
Input: RDB file
Output: XML Document (XML Schema)
Begin

  Select *data file* from the document
  Make XML Document.name as RDB.name
  Suppose bigdata files has total n file data schemas in it
  Loop For i = 1 to n **do**

    Select $file_i$*.name* from RDBS

    Make $file_i$*.name* as *elementi.name* under XML document

    Build Tag <xs:element name=" $file_i$*.name*">

    Suppose $file_i$*.name* has total m fields in it

```
                Build Tag <xs:complextype>
                Build Tag <xs:sequence>
                Inner Loop For j = 1 to m do
                        Select fieldj of filei
                        Condition IF fieldj is primary key
                                Make fieldj equal to attribute having attribute.type equally mapped to
                                fieldj.datatype
                                Build Temporary Variable save values of mapped attribute.type and
                                fieldj.name in one iteration of temp1 array variable
                                /*here temporary variable is use to contain information which will be used
                                afterwards*/
                        Condition Else IF fieldj is foreign index
                                Make fieldj equal to simple element having element.type equally mapped to
                                fieldj.datatype
                                Build Temporary Variable save values of mapped attribute.type and
                                fieldj.name in one iteration of temp2 array variable
                        Condition Else
                                Make fieldj equal to elementj having elementj.type equally mapped to
                                fieldj.datatype
                                Build Tag <xs:element name=" fieldj.name" type=" elementj.type" />
                        End IF
                End Inner Loop
                Build Tag </xs:sequence>
                Loop For k = 1 to temp1.length
                        Build Tag <xs:attribute name="temp1k.name" type=" temp1k.type" use="required" />
                End Loop
                Loop For k = 1 to temp2.length
                        Build Tag <xs:attribute name="temp2k.name" type=" temp2k.type" />
                End Loop

                Build Tag </xs:complextype>

        End Loop
End
```

Next algorithm concerns with transformation from XML Schema to RDFS, where each element tag representing a relation made a class in RDF and a field in the algorithm is made a property as an alternative for the concepts. In this algorithm, document name is used to represent a root class and then through a looping mechanism a Class for each table is found to be in complex element. For each simple element makes it as its property and assigns data type against each type of that property. And now through this way we can generate triples as a representative to XML Schema and indirectly our source files.

Algorithm: Transformation from XML Schema to RDFS
Input: XML Document (XML Schema)
Output: RDFS Triples

Begin
Select XML Document.name from the document
Build Triple XML Document.name rdfs:Class rdf:resource
/* here dot symbol shows property of the document selected*/
Suppose XML Document has total n complex elements in it
Loop For i = 1 to n **do**
Select $element_i$ from XML Document
Selected Tag <xs:$element_i$ *name*=" $element_i$*.name*">
/*i-th element of complex type*/
Make Triple $element_i$*.name* rdfs:Class XML Document.name
Suppose $element_i$ has total m sub elements in its sequence tag
Inner Loop For j = 1 to m **do**
Select *sub-*$element_j$ of $element_i$
Make Triple $element_i$*.name* rdf:Property *sub-*$element_j$*.name*
Make Triple *sub-*$element_j$*.name* rdfs:DataType *sub-*$element_j$*.type*
End Inner Loop
Suppose $element_i$ has total p attributes in it
Inner Loop For k = 1 to p **do**
Condition IF $attribute_k$*.use* is as *required*
/*equivalent to the tag <xs: $attribute_k$ name = "$attribute_k$*.name*" type =
"xs:$attribute_k$*.type*" use="required">*/
Make Triple $element_i$*.name* rdf:Property $attribute_k$*.name*
Make Triple $attribute_k$*.name* rdfs:range $element_i$*.name*
Make Triple $attribute_k$*.name* rdfs:DataType $attribute_k$*.type*
Condition Else
/*equivalent to the tag <xs: $attribute_k$ name = "$attribute_k$*.name*" type =
"xs:$attribute_k$*.type*" >*/
Make Triple $element_i$*.name* rdf:Property $attribute_k$*.name*
Make Triple $attribute_k$*.name* rdfs:domain $element_i$*.name*
Make Triple $attribute_k$*.name* rdfs:DataType $attribute_k$*.type*
End IF
End Inner Loop
End Loop
End

Assessments Mathematical Modeling

Assessment will be focusing on time-oriented classification of data by matching mechanism of ordering of each instant occurrence. Matching mechanism among data is given in definition 1 as follows:

Definition 1 (Matching Mechanism)

A Matching Representations between $m \in XML\ profiles$ at Eq. (1) and $c \in Current\ Captures$ at Eq. (2), at time $t_n \in Time$, as X_m and $X_c^{t_m}$, forming sequence $\left(x_c^{t_1}, x_c^{t_2}, x_c^{t_3}, \cdots\right)$ with $t_1 \le t_2 \le \cdots \le t_n$, where:

For each individual j,

$$x_c^i[j] \in \left[\min_j^c, \max_j^c\right] \qquad \text{............(1)}$$

$$x_m[j] \in \left[\min_j^m, \max_j^m\right] \qquad \text{............(2)}$$

with $i = 1, 2, 3, \cdots$

Definition 2 (Pattern Recognition)

Given a resource *r* (could be any category of data) and its scoring functions S^r, *r*'s pattern recognition at time t' of individual's match value x^t for a time $t < t'$, is defined as:

$$P^r\left(t', x_c^t\right) = \begin{cases} reject & \textbf{if } t' > t^r \textbf{ or } S\left(x^{t'}\right) < 1 \\ add\ one & \textbf{if } S\left(x_c^{t'}\right) \approx S^r\left(x_m^t\right) \\ x^t & otherwise \end{cases} \qquad \text{............(3)}$$

Data Mapping Algorithms

Algorithms *TranslateTag()*, *GenerateTriple()* and *Transformation()* here are representing a mechanism of transforming data for computation and mapping with the help of equations defined earlier.

Algorithm *TranslateTag()* Translates corresponding tag into RDF tag

---

Input: number of items (n); value recorded x^t; type of each item (type); recoded instant at time *t*
Output: list of annotated (<tags>) XML element
1. Collect data generated from the session *S*
2. **Loop** decision iterator i:=1 to n **do**
3.   **If** $P^r\left(t', x_c^t\right)$ == 1 according to Formula (3)
4.   extract each row and tag it as an element
5.   close the each corresponding tag
6.   List *L* := add element
7.   **End If**
8. **End Loop**
9. **Return** list *L*

Algorithm *TranslateTag()* is taking values generated by an assessment session and running it through the process contained by agents on the bases of Eq. (3) formula to see its fitness to be translated into a tag. Then each tag based schema is along with its value is generated. And according to the w3c standards for XML tags each tag is closed accordingly in statement 4 and 5. All generated tags are further stored in a list to be returned to the calling function for concerned *TranslateTag()* Algorithm.

Algorithm *GenerateTriple()* Generation of corresponding list of triples for given XML tag

Input: number of tags (n); type of each tag (type)
Output: list of annotated (<tags>) XML/RDF

1. Expand each tag
2. **Loop** decision iterator i:=1 to n **do**
   //number of tags contained within an element
3. **If** tag[i].isElement() == 1
   //isElement() returns 1 when current tag is element
4. extract each tag and annotate it as triple
5. generate unique id for new resource
6. List *T* := add triples
7. **End If**
8. **End Loop**
9. **Return** list *T*

Here algorithm *GenerateTriple()* further transforms XML tags into RDF triples. Statement 3 is test made to check input tag being complete tag by returning 1 under the method named isElement(). Each resource generated during this transformation process new id is given based on previously does not exist. All generated triples are further stored in a list to be returned to the calling function for concerned *GenerateTriple()*.

Algorithm *Transformation()* from Assessment Data into Annotated RDF Format

Input: Data File to annotate; type of each data item (type)
Output: annotated (<tags>) item into triple reduced from original

1. Collect data generated from the session *S*
2. **Repeat** until EOF //End of File (EOF)
3. annotated List *L* := *GenerateTriple(TranslateTag())*
   //according to Algo 1 and 2
4. **End Loop**

Algorithm *Tranformation()* is using both previously developed algorithms undertaken for updated list of triples at statement 3. Complete input is read at statement 2 under a loop storing in annotated list *L*. Now we have to calculate the time complexity of the algorithms in the discussion. Let's consider, Time complexity be represented by T_j for algorithm j against statement i is represented as Ti resulting in the form of upper and lower bound by the use of ϴ (theta) sign where j=1,2,3 and 1≤ i ≤ 9.

SEMANTICAL ANNOTATION OF STUDENT DATASET ASSESSMENTS

To prove the methology validity experiment was done on around 243 students of 8^{th} graders in Pakistan specifics by examining their skills on the subject like Mathematics (Mat), English Language (Eng), Urdu Language (Urd), Islamiat (Isl), Social Science (So), General Science (Sci), Arabic Language (Arab), and Elective Subject (Eltv). Chosen candidates were picked randomly to perform online assessment tests and their results are shown in **Figure**Figure 4.

Then whole assessed data was further made available in relational database table, shown in Figure 5, enabled to be queried. Now make it more intelligence oriented data form transformation process is executed which is being developed in Java platform.

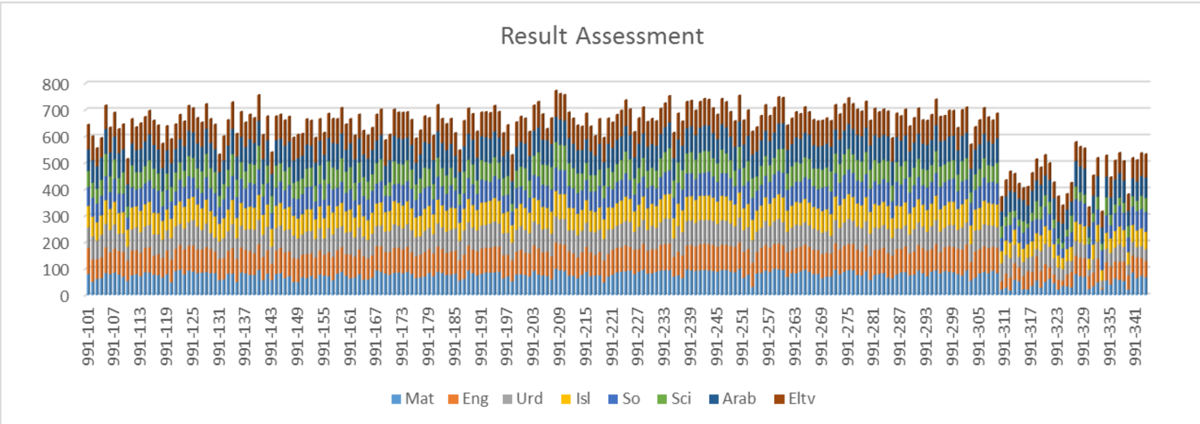


Figure 4. Students assessment clarifying Subjective Impact on Results' Total

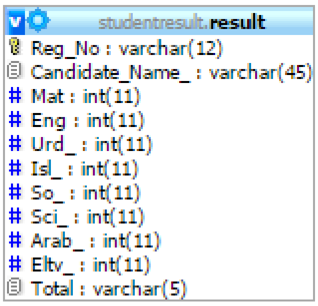


Figure 5. Relation Database table schema representation of Result dataset

```xml
<?xml version="1.0" encoding="UTF-8"?>
<studentresult xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3schools.com/XSD_output_studentresult.xsd">
    <result>
        <Reg_No>991-101</Reg_No>
        <Candidate_Name_>Abdul Aleem</Candidate_Name_>
        <Mat>80</Mat>
        <Eng>98</Eng>
        <Urd_>78</Urd_>
        <Isl_>80</Isl_>
        <So_>72</So_>
        <Sci_>60</Sci_>
        <Arab_>82</Arab_>
        <Eltv_>93</Eltv_>
        <Total>643</Total>
    </result>
    <result>
        <Reg_No>991-102</Reg_No>
        <Candidate_Name_>Abdul Hanan</Candidate_Name_>
        <Mat>50</Mat>
        <Eng>85</Eng>
        <Urd_>86</Urd_>
        <Isl_>75</Isl_>
        <So_>72</So_>
        <Sci_>56</Sci_>
        <Arab_>85</Arab_>
        <Eltv_>91</Eltv_>
        <Total>600</Total>
    </result>
```

(a)

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified" attributeFormDefault="unqualified">
<xs:element name="studentresult">
<xs:complexType>
<xs:sequence>
<xs:element ref="result" minOccurs = "0" maxOccurs = "unbounded" />
</xs:sequence>
</xs:complexType>
</xs:element>
    <xs:element name="result">
        <xs:complexType>
            <xs:sequence>
                <xs:element name="Reg_No">
                    <xs:simpleType>
                        <xs:restriction base="xs:string">
                            <xs:maxLength value="12"/>
                            <xs:minLength value="1"/>
                        </xs:restriction>
                    </xs:simpleType>
                </xs:element>
                <xs:element name="Candidate_Name_">
                    <xs:simpleType>
                        <xs:restriction base="xs:string">
                            <xs:maxLength value="45"/>
                            <xs:minLength value="1"/>
                        </xs:restriction>
                    </xs:simpleType>
                </xs:element>
```

(b)

Figure 6. XML level data transformation (a) XML data sample after transforming (b) XMLS data sample after transforming

This transformation takes data and metadata both and transforms them into XML format for data (as shown in Figure 6-a) and XSD format for Schema or metadata (as shown in Figure 6-b) of the given database. In Figure 6 (a) translated tag like result is representing a single instance or record of the student placed in equivalent sub tags named after each field available to represent a data classification for subjects' resultant value for specific student. Whereas Figure 6 (b) is the translation of data structure based on metadata (data about data) achieved from table of RDB clearly defining data type, limitations, and customizations.

Data translated in the form of XML is useful and machine readable for all the application supporting XML. As todays' era is of web founded mostly, which makes such transformation even more useful and distributable. By passing it further into the proposed bidirectional data transformation mechanism, next comes the RDF form including hierarchal and linkage embedded in the data model supporting rule oriented transactions possible on the data (AlObaidi et al., 2016). RDF level data transformation represented in Figure 7(a) as for RDF data sample after transforming and Figure 7(b) for RDFS data sample after transforming. This make data more into the supported form of Artificial Intelligence (AI) and Web oriented.

When this transformed RDF data is stored into the RDF store (shown in Figure 8) then total generated triples out of data and metadata comprises of 2,826 triples for 243 students stored records. This RDF store is capable to be utilized for applying AI assessment algorithms for training a system for future assessments and their semantics.

```xml
<?xml version='1.0' encoding='ISO-8859-1'?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:studentresult="http://www.studentresult.tms/studentresult/">

<rdf:Description rdf:about="http://www.studentresult.tms/studentresult#result">
<studentresult:Reg_No>991-101</studentresult:Reg_No>
<studentresult:Candidate_Name_>Abdul Aleem</studentresult:Candidate_Name_>
<studentresult:Mat>80</studentresult:Mat>
<studentresult:Eng>98</studentresult:Eng>
<studentresult:Urd_>78</studentresult:Urd_>
<studentresult:Isl_>80</studentresult:Isl_>
<studentresult:So_>72</studentresult:So_>
<studentresult:Sci_>60</studentresult:Sci_>
<studentresult:Arab_>82</studentresult:Arab_>
<studentresult:Eltv_>93</studentresult:Eltv_>
<studentresult:Total>643</studentresult:Total>
</rdf:Description>
```

(a)

```xml
<?xml version='1.0' encoding='ISO-8859-1'?>
<!DOCTYPE rdf:RDF [
<!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
<!ENTITY rdfs 'http://www.w3.org/TR/WD-rdf-schema#'>
<!ENTITY owl 'http://www.w3.org/2002/07/owl#'>
<!ENTITY xsd 'http://www.w3.org/2001/XMLSchema#'>
]>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#" >
    <rdfs:Class rdf:about="http://www.studentresult.tms/studentresult#studentresult">
        <rdfs:label>studentresult</rdfs:label>
        <rdfs:comment>Table</rdfs:comment>
        <rdfs:subClassOf rdf:resource="http://www.studentresult.tms/studentresult#null"/>
    </rdfs:Class>
    <rdfs:Class rdf:about="http://www.studentresult.tms/studentresult#result">
        <rdfs:label>result</rdfs:label>
        <rdfs:comment>Table</rdfs:comment>
        <rdfs:subClassOf rdf:resource="http://www.studentresult.tms/studentresult#studentresult"/>
        <rdfs:subClassOf>
            <owl:restriction>
                <owl:onProperty rdf:resource="http://www.studentresult.tms/studentresult/Reg_No" />
                <owl:maxCardinality rdf:datatype="&xsd;nonNegativeInteger">12</owl:maxCardinality>
            </owl:restriction>
        </rdfs:subClassOf>
        <rdfs:subClassOf>
            <owl:restriction>
                <owl:onProperty rdf:resource="http://www.studentresult.tms/studentresult/Reg_No" />
                <owl:minCardinality rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
            </owl:restriction>
        </rdfs:subClassOf>
```

(b)

Figure 7. RDF level data transformation (a) RDF data sample after transforming (b) RDFS data sample after transforming

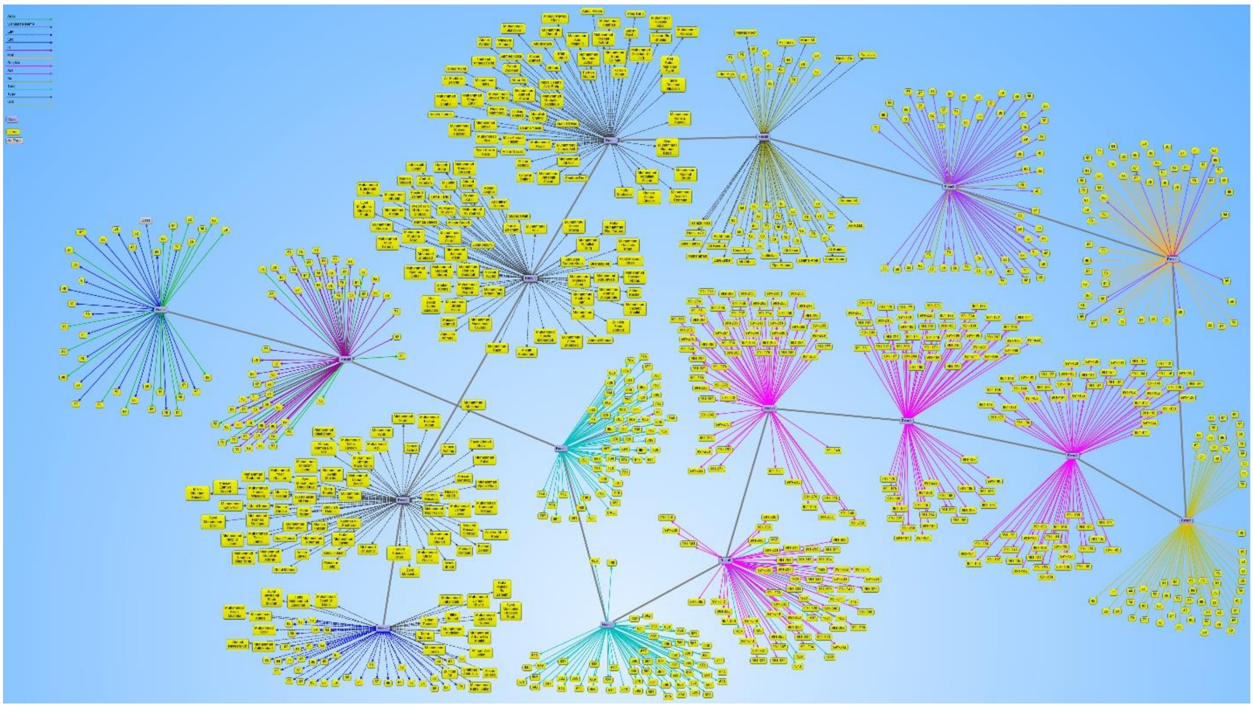


Figure 8. RDF Graph representation of Result Assessment dataset

RESULTS AND DISCUSSION

As in Web applications, generally data is stored in the form of Relational Database (RDB) or Resource Description Framework (RDF). When these datasets need to transform from RDB to Semantic Web (SW) based system, there is no known way to do this transformation without data loss due to compatibility issues. The literature describes several rigid techniques that do transformation from RDB to RDF with limited customization, but failed to present an intermediate way that helps to avoid compatibility issue. In this paper, we have represented a new methodology that allow us to do data mapping that can be used to understand their differences at the level of data types. This mapping is done using Extensible Markup Language (XML) based data structure as their intermediate data presenter. We performed control experiment to investigate whether Document Type Definition (DTD) or Extensible Markup Language Schema (XMLS) works better for performing transformation from RDB to RDF, and shows XMLS give better mapping results for process of transformation. This approach will allow data transformation from RDB to RDF without data loss and compatibility issue and thus traditional systems can easily be transformed to Semantic Web based system (Musen, 2015).

Finding paths of improving assessment calculation by formulizing the concepts in reforming data reachable and compatible with other areas in Computer Assistant Assessment (CAA). For example, through the data translated in XML and used to calculate heat circular graph presented in Figure 9 which depicts student versus subjects' marks variations. Focus on subject-wise variations in students' result are clearly visible and grey area concerning specific students' range are highlighted as lines that represent score reaching to zero. Whereas Figure 10 brings weak students, which is represented as a slice taken out of the Figure 9. In Figure 10, students' Reg. No ranging from 311 to 341, start dramatically decline in securing marks under different subjects up to passing level. This assessment was calculated through Eq. (1) and (2) by mapping XML values gained from different subject oriented tags per available dataset after passing through data transformation process. This can be viewed and observed to overcome different factors like teachers' teaching methodology, students' learning curve, students' attentiveness, and attendance etc. by answering 3 W's (what, when, and why) by making intelligent survey.

Total 243 students were clustered into two groups of pass and fail.

Cluster 0	Pass	233	95.88477	%
Cluster 1	Fail	10	4.115226	%

Patterns defined on the failed group of students as shown in Figure 11, where red shade represents danger zone as more darker means near to failure. Whereas, green shade is representing the safe state.

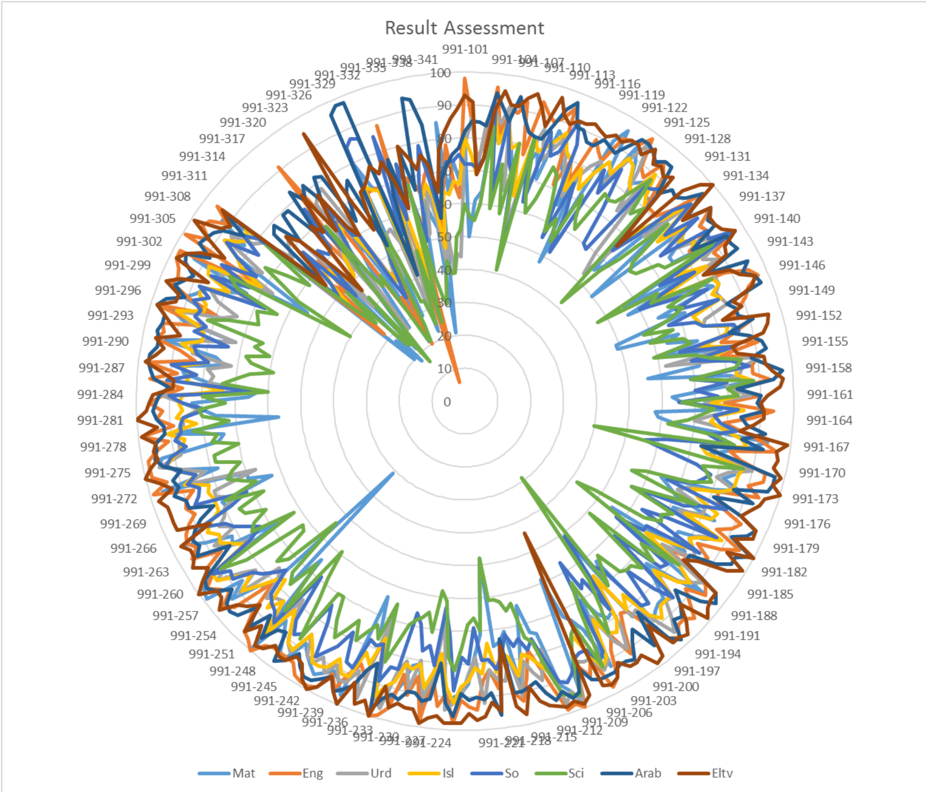


Figure 9. Each Student-Subject (colored) Assessment

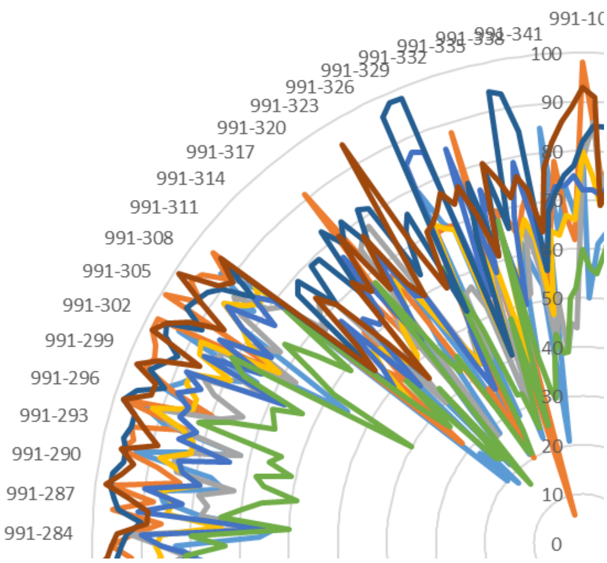
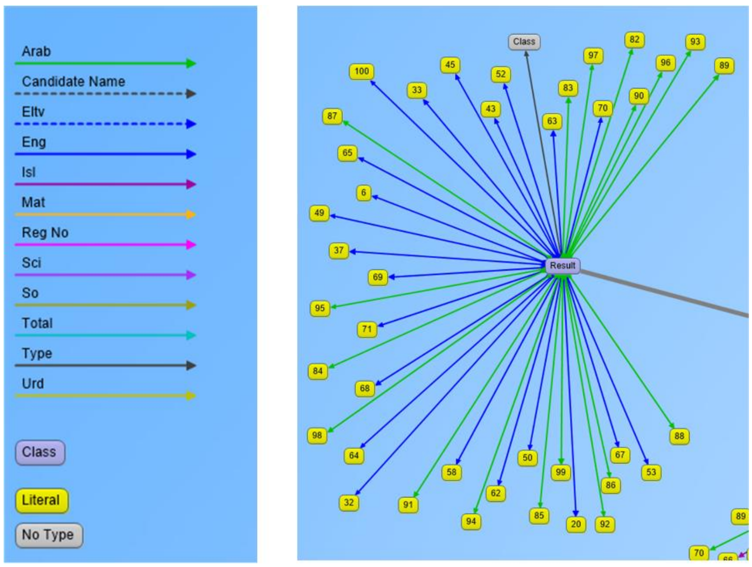


Figure 10. Results problem spectrum where students fail

Mat	Eng	Urd	Isl	So	Sci	Arab	Eltv	
20	32	68	46	50	44	56	55	FAIL
18	49	73	58	67	49	79	74	FAIL
52	53	40	52	64	24	79	57	FAIL
22	50	53	50	61	43	79	46	FAIL
21	69	51	51	62	16	68	71	FAIL
21	37	46	46	42	22	64	95	FAIL
35	20	43	42	36	21	73	69	FAIL
23	49	25	41	54	25	41	74	FAIL
20	6	29	39	75	25	61	61	FAIL
21	53	41	47	56	43	56	64	FAIL

Figure 11. Pattern defined on the cluster of failed students

Furthermore, RDF store represented in Figure 8 by taking a closer look in Figure 12 (a) gives arrow linkages representing predicates between of source and object resources. In Figure 12 (b), a node of result as source is being explored against some of the literal values nodes for English and Arabic subject predicates also showing result as a class of RDFS. This information is ready to be used for assessment with exploring and querying using SPARQLE (a language for navigating and querying RDF).



(a) (b)

Figure 12. RDF Graph Close-up (a) Legends of the Graph (b) Values belonging to Arab and Eng Subjects

Random Forest is a standout amongst the most prominent machine learning algorithms. It is a kind of collective machine learning algorithm that is also known as bagging. The bagging method is used for developing multiple diverse models from a single dataset. The focus is to understand the results and their validation for analytical data modeling related issues. Decision tree are used as random forest based bagging classifier. Following is the cross-validation using bagging classifier on the results of 243 students.

Correlation coefficient	0.9964
Mean absolute error	5.0265
Root mean squared error	8.1159
Relative absolute error	7.7072 %
Root relative squared error	9.1744 %
Total Number of Instances	243

CONCLUSION

Research on currently available tools and methodologies with their frameworks can help as a state of art being used for transformation considering assessment calculations. By finding weak areas and providing alternative mechanisms in recovering and providing deep analysis having semantics. Even for bidirectional transformation to work properly we need to map different data models for transformations. A case study is used to show handling of mappings, implementations, and updates to further improve data assessment compatibility with Artificial Intelligence. Assessment made in the case study explores the range and subjects in which specific students lack to produce high impact by scoring which may be used to analyze different factors effective on producing issues towards learning. This methodology will help in opening new opportunities as projects and research work for Asian countries at both academic and professional level (Umair et al., 2015). This will help in improving with the enhanced utilization and compatibility among system concerning semantic web by updating the part of system, not necessary requires updates, as for industries, economics, and governance on data oriented assessments. This will help inducing large scale loss free transformation of traditional systems into semantically enriched systems back and forth. This will also improve return on cost investments. Collaboration of systems with outer World remaining within the domain of system or extending access to other domain based systems (Umair et al., 2015). Improved compatibility among systems using either semantic or relational data model will benefit us to move one step forward in upcoming trend of web based assessment orient systems.

ACKNOWLEDGEMENT

I would like to acknowledge my father, brother, friends and my colleagues for their support, suggestions and reviews, while conducting this study. I am also thankful to them on successful completion of this part of the project.

REFERENCES

Agus, M., Penna, M. P., Peró-Cebollero, M., & Guàrdia-Olmos, J. (2016). Assessing Probabilistic Reasoning in Verbal-Numerical and Graphical-Pictorial Formats: An Evaluation of the Psychometric Properties of an Instrument. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(8), 2013-2038.

AlObaidi, M., Mahmood, K., & Sabra, S. (2016). *Semantic Enrichment for Local Search Engine using Linked Open Data.* Paper presented at the Proceedings of the 25th International Conference Companion on World Wide Web.

Arenas, M., Bertails, A., Prud'hommeaux, E., & Sequeda, J. (2013). A Direct Mapping of Relational Data to RDF, W3C Recommendation 27 September 2012.

Auer, S., Dietzold, S., Lehmann, J., Hellmann, S., & Aumueller, D. (2009). *Triplify: light-weight linked data publication from relational databases*. Paper presented at the Proceedings of the 18th international conference on World wide web.

Barrasa Rodríguez, J., Corcho, Ó., & Gómez-Pérez, A. (2004). R2O, an extensible and semantically based database-to-ontology mapping language *Proceedings of the Second Workshop on Semantic Web and Databases, SWDB 2004* (Vol. 3372, pp. 1069-1070). Verlag, Berlín, Alemania: Springer.

Belcadhi, L. C. (2016). Personalized feedback for self assessment in lifelong learning environments based on semantic web. *Computers in Human Behavior, 55*, 562-570.

Bizer, C., & Seaborne, A. (2004). *D2RQ-treating non-RDF databases as virtual RDF graphs*. Paper presented at the Proceedings of the 3rd international semantic web conference (ISWC2004).

Blakeley, C. (2007). Virtuoso RDF views getting started guide. *OpenLink Software*.

Chang, C.-C. (2001). Construction and evaluation of a web-based learning portfolio system: An electronic assessment tool. *Innovations in Education and Teaching International, 38*(2), 144-155.

Chung, G. K., Niemi, D., & Bewley, W. L. (2003). Assessment Applications of Ontologies.

Cyganiak, R., Bizer, C., Garbers, J., Maresch, O., & Becker, C. (2012). The d2rq mapping language. *Online]* http://d2rq. *org/d2rq-language*.

Dafli, E., Antoniou, P., Ioannidis, L., Dombros, N., Topps, D., & Bamidis, P. D. (2015). Virtual patients on the semantic Web: a proof-of-application study. *Journal of medical Internet research, 17*(1), e16.

Das, S., Sundara, S., & Cyganiak, R. (2012). R2RML: RDB to RDF Mapping Language.

de Laborda, C. P., & Conrad, S. (2005). *Relational. OWL: a data and schema representation format based on OWL.* Paper presented at the Proceedings of the 2nd Asia-Pacific conference on Conceptual modelling-Volume 43.

del Mar Sánchez-Vera, M., Fernández-Breis, J. T., Castellanos-Nieves, D., Frutos-Morales, F., & Prendes-Espinosa, M. P. (2012). Semantic Web technologies for generating feedback in online assessment environments. *Knowledge-Based Systems, 33*, 152-165.

Hert, M., Reif, G., & Gall, H. C. (2010). *Updating relational data via SPARQL/update.* Paper presented at the Proceedings of the 2010 EDBT/ICDT Workshops.

Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2005). *Adaptive e-learning content generation based on semantic web technology*. Paper presented at the 12th International Conference on Artificial Intelligence in Education AIED 2005, Amsterdam. Netherlands.

Holohan, E., Melia, M., McMullen, D., & Pahl, C. (2006). *The generation of e-learning exercise problems from subject ontologies.* Paper presented at the Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06).

Kardan, A. A., Sani, M. F., & Modaberi, S. (2016). Implicit learner assessment based on semantic relevance of tags. *Computers in Human Behavior, 55*, 743-749.

Knight, C., Gasevic, D., & Richards, G. (2006). An ontology-based framework for bridging learning design and learning content. *Educational Technology & Society, 9*(1), 23-37.

Liu, Y., & Khine, M. S. (2016). Content Analysis of The Diagrammatic Representations of Primary Science Textbooks. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(8), 1937-1951.

Mettiäinen, S. (2015). Electronic Assessment and Feedback Tool in Supervision of Nursing Students during Clinical Training. *Electronic Journal of e-Learning, 13*(1), 42-56.

Michel, F., Montagnat, J., & Faron-Zucker, C. (2014). *A survey of RDB to RDF translation approaches and tools.* (Research Report), Université Nice Sophia Antipolis.

Musen, M. A. (2015). Semantic Web Technology for Mapping and Applying Clinical Functional Assessment Information: DTIC Document.

Queirós, R., & Leal, J. P. (2012). Orchestration of E-Learning Services for Automatic Evaluation of Programming Exercises. *J. UCS, 18*(11), 1454-1482.

Sicilia, M.-Á., & Barriocanal, E. G. (2005). On the convergence of formal ontologies and standardized e-learning. *International Journal of Distance Education Technologies, 3*(2), 13.

Stantchev, V., Prieto-González, L., & Tamm, G. (2015). Cloud computing service for knowledge assessment and studies recommendation in crowdsourcing and collaborative learning environments based on social network analysis. *Computers in Human Behavior, 51*, 762-770.

Umair, S., Björklund, A., & Petersen, E. E. (2015). Social impact assessment of informal recycling of electronic ICT waste in Pakistan using UNEP SETAC guidelines. *Resources, Conservation and Recycling, 95*, 46-57.

http://www.ejmste.com