AN ELECTRONIC DATA INTERCHANGE FRAMEWORK FOR EDUCATIONAL INSTITUTES

ATTA-UR-RAHMAN[1,*] AND FAHD ABDULSALAM ALHAIDARI[2]

[1]Department of Computer Science
[2]Department of Computer Information System
College of Computer Science and Information Technology (CCSIT)
Imam Abdulrahman Bin Faisal University (IAU)
P.O. Box 1982, Dammam 31441, Saudi Arabia
*Corresponding author: aaurrahman@iau.edu.sa; faalhaidari@iau.edu.sa

ABSTRACT. *Electronic Data Interchange (EDI) is a valuable concept which facilitates the exchange of data amongst various related entities in a particular domain. In the educational sector an efficient and well-defined EDI system is a necessity. In Saudi Arabia there are over 770 colleges and the current manner of communication amongst them is primitive, depending mostly on paper-based interactions. This is evidently tedious and time consuming. Thus, in this study a new EDI grammar dedicated for educational facilities in Saudi Arabia, is proposed. First, the framework of the educational facilities is analyzed. Then, the grammar of the proposed EDI language is defined accordingly w.r.t the Entity Relationship Diagram (ERD). A thorough description on the message fields is provided covering several sample messages, the import/export mechanism and implementation detail.*
Keywords: EDI, HL7, Machine readable format, Grammar, Language

1. **Introduction.** Electronic Data Interchange (EDI) is an application to exchange electronic business documents between two different business entities or between business entity departments and divisions. Exchangeable business documents must be transferred in the standard transactions sets which is related to specific business documents [1]. A Business Document is any related business document that must be sent between entities in the same domain such as purchase orders or customer invoices [2]. Using EDIs in the educational domain will ease the interchange of current and historical information and records of the students to employers, other educational institutes, agencies, or when the student is transferred from one institute to another [3,4]. In Saudi Arabia, there are more than 700 colleges, and educational deanships within the higher education universities and institutions [5]. The current situation of interchanging the students' records is manual and tedious. Thus, this paper proposes an educational EDI to be applied in Saudi Arabia's educational entities to replace all human readable, and paper-based documents with machine readable, and electronically coded documents [2].

Motivation, Contribution and Significance: According to Saudi Arabia's Ministry of Education statistics in 2016, there are 70 public and private universities, colleges, and technical institutes, with a total number of 770 colleges, and educational deanships within the higher education institutions [5,6]. These institutes provide different types of certificates, diplomas, fellowship and degree programs from Bachelor to Ph.D. The total enrolled students in all these institutes are 1,622,441 [6]. With these large numbers of institutes and universities in Saudi Arabia, there is an acute need for educational EDI to exchange students' records. The current situation of transmitting students' records and

transcripts is a manual process, where the student will go to his/her current educational entity, take a stamped hard-copy of the transcript and go to the new educational entity and give it to them for a manual verification. This process is tedious especially when the entities are in different regions. So, having such an EDI will ease the students' records transmission between educational entities. In this paper, an EDI system serving the higher educational sector is proposed to be applied within Saudi Arabia's institutes. The rest of this paper is organized as follows. Section 2 presents an overview on education facilities in Saudi Arabia and discusses the EDI in educational domain. In Section 3 the proposed EDI is presented. A comparison between the proposed EDI and HL7 is provided in Section 4. Section 5 concludes the paper.

2. **Electronic Data Interchange (EDI).** EDI is an application to exchange electronic business documents between two different business entities or between business entity departments and divisions [1]. Exchangeable business documents must be transferred in the standard transaction sets. Transaction sets are related to specific business documents, combining more than one transaction set together provides the business with a series of standardized transaction formats that cover a wide range of business communication needs [1]. Business Document is any related business document that must be sent between entities in the same domain such as purchase orders or customer invoices [2]. EDIs function without human intervention thus decreasing error rate and assuring loss prevention. EDIs are machine-readable, domain specific and schema oriented. EDI has been introduced 30 years ago, and there are many EDI standards in use nowadays such as X12, EDIFACT, and HL7, and for each standard there are many different versions such as HL7 version 2, 2.5.1, 3 [7]. EDIs have replaced the human readable and paper-based documents with machine readable, and electronically coded documents [2]. Using EDI, the systems can communicate even if it is utilizing different or incompatible format [7].

2.1. **EDI in education sector.** Several EDI standards have been proposed to ease data exchange for educational agencies.

X12-130 transaction set contains the standard format of establishing a data Student Educational interchange [8]. This transaction standard is used by schools and post-secondary educational institutions to interchange the current and historical information and records of the students. The student's records can be sent using transaction set 130 to employers, other educational institutes, agencies, or when the student is transferred among the institutes [3,4].

EDexchange is an Educational EDI sponsored by the Postsecondary-Electronic-Standards-Council (PESC). The EDexchange project was initiated in 2011 by California Community College and Technology Centre (CCCTC), which is considered as one of the largest educational institutes in the United States with over 112 colleges serving more than 2 million students in the USA. EDexchange aims to facilitate the exchange of educational data amongst facilities within the United States. It is currently in trial-phase and is operated in 6 institutions in the US, including, University of Phoenix and California Community Colleges. Moreover, it is a web-based open source project built on PESC standards. It is emphasized that EDexchange is vendor-independent, no third-parties are required. This vendor independence is believed to encourage trust within the educational administration and more importantly guaranty full-integration and compatibility on a larger scale. Furthermore, it must be noted that it merely facilitates the interchange of educational data and does not store any exchanged data or files in compliance to legal requirements [9]. One of the PESC approved standards is the College Transcript standard, defined for exchanging student transcripts between educational institutes. According to the implementation guide version 1.1.1 [10], the standard consists of four components as transmission data, student, note message and user-defined

extensions. The PESC College Transcript Standard is formatted in the following way.

```
<HSTrn:CollegeTranscriptXMLns:ColTrn="urn:org:pesc:message:CollegeTranscript:v1.1.0"
XMLns:AcRec="urn:org:pesc:sector:AcademicRecord:v1.3.0" XMLns:core="urn:org:pesc:core:CoreMain:v1.4.0"
XMLns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:org:pesc:message:CollegeTranscript:v1.1.0
/CollegeTranscript_v1.1.0.xsd"><TransmissionData></TransmissionData><Student></Student><NoteMessage
```

The data elements can be further extended (i.e., complex elements) such as the Transmission Data and Student elements.

2.2. **EDI in health sector.** Health Level Seven (HL7) is a comprehensive EDI dedicated for exchanging information within the medical community. The first version was established in the late 1980's, shortly followed by versions 2 and 3. HL7 is accredited by ANSI and is well recognized and adopted in more than 50 countries worldwide [11-15]. HL7 is a way of automated information exchange among the healthcare stakeholders. It supports various activities such as, patient admission, appointment scheduling, pharmacy prescriptions, equipment inventory, insurance handling, and financial recording. Many systems are generic in text processing [16-25].

3. **Proposed EDI for Universities in Saudi Arabia.** The scope of the proposed system is to provide an educational EDI that can be implemented as a framework for exchanging students' records among universities in Saudi Arabia. In this section, the proposed system framework including its users, entities and relationships are presented in an Entity Relationship Diagram (ERD) in Figure 1.
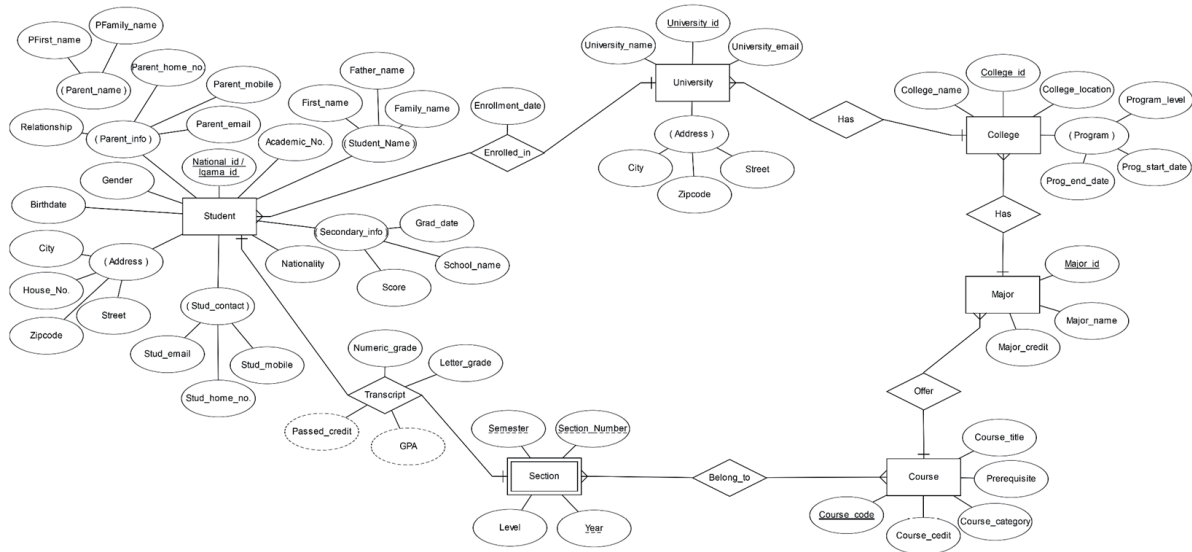

FIGURE 1. ERD of the proposed system

3.1. **System framework.** The main users of the proposed system are universities and any educational institutes that require students' records. The internal stakeholders are the students, the faculty and the administration staff, etc. The proposed system has 6 entities, 6 relationships and 38 attributes. The entities are: student, university, college, department, course and section.

Enrolled_in: a one-to-many relationship between "Student" and "University" entities in which many students can be enrolled in one university at a time. This relationship has one attribute which is the enrollment date.

Has: a one-to-many relationship between "University" and "College" entities and between "College" and "Major" entities in which one university has many colleges and one college has many departments. This relationship does not have any attributes.

Offer: a one-to-many relationship between "Major" and "Course" entities in which one department offers many courses. This relationship does not have any attributes.

Belong_to: a many-to-many relationship between "Course" and "Section" entities in which many courses belong to many sections. This relationship does not have any attributes.

Transcript: a one-to-one relationship between "Student" and "Section" entities in which one student in one section can generate a grade report. It has 4 attributes including letter grade, numeric grade, passed credits and GPA. The mapping of the ERD is shown in Figure 2.
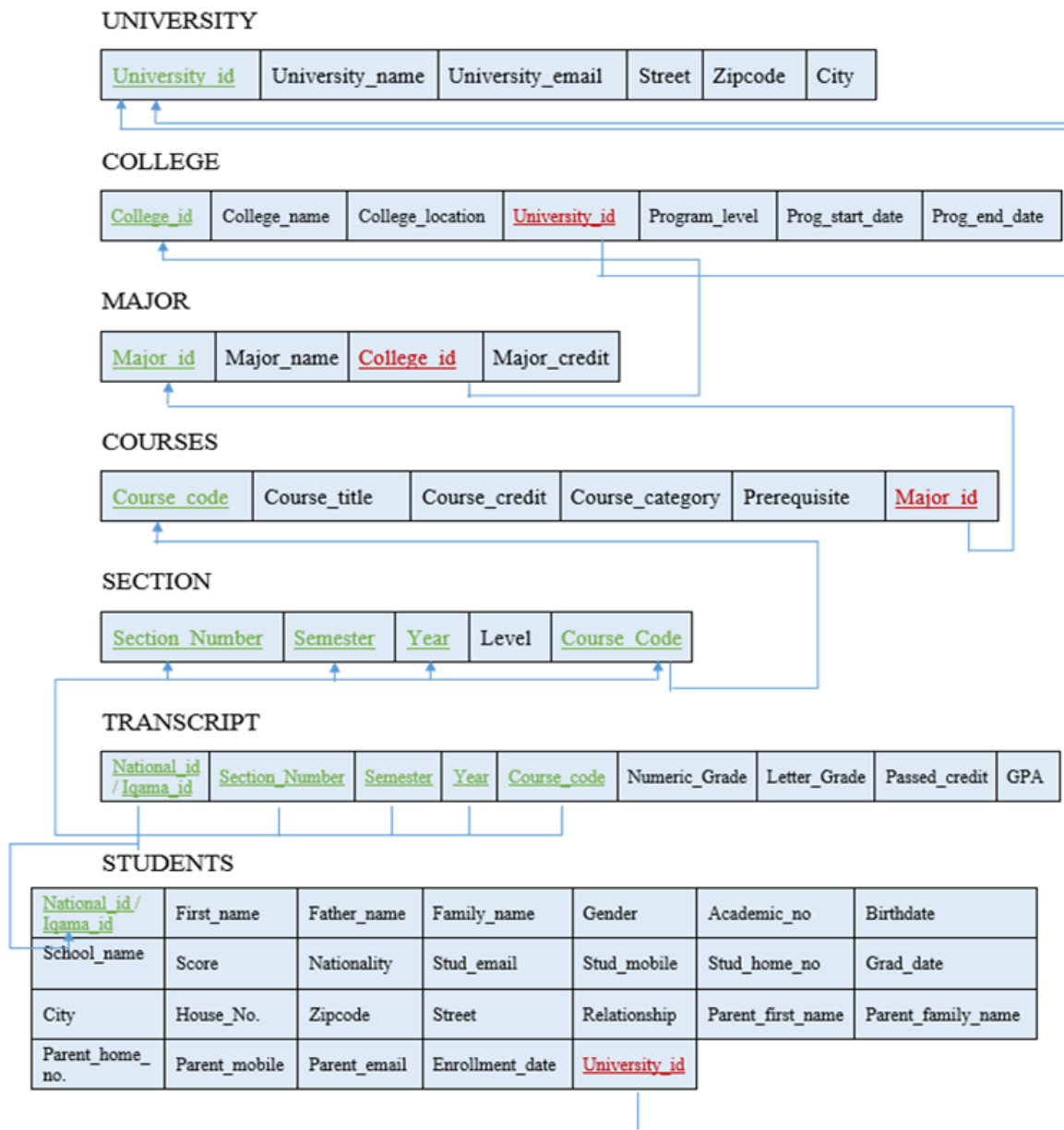

FIGURE 2. Relational database schema

3.2. **Message format.** To fully support the existing framework for universities, well defined standards need to be asserted. All existing operations and communications need to be represented by structured electronic messages. Each message should hold the information that would normally be exchanged via paper document transactions. These messages should be formatted in a way that allows other parties to interpret and understand its purpose and contents without any human intervention. For instance, a student

can request to transfer from one university to another without the need to physically visit either party. In the proposed EDI, each message consists of segments. A segment groups multiple relating field together. A segment may be repeated within the same message or may not occur at all, depending on each message's definition presented later. For instance, the course_prerequisite field in the Add_course message is both optional and repeatable, seeing that a course can have multiple pre-requisites or none. A field can be simple or complex, i.e., contains other fields. The data within the message will be in ASCII format. More importantly, encryption of sensitive data within the message will be done to ensure data privacy whilst exchanging messages in the proposed EDI. Sensitive information like payment data will be concealed to prevent data exposure. Delimiters are used to distinguish each message's components from each other. Figure 3 depicts the delimiters used in the proposed EDI. Each segment is written in a new line. Fields are separated using a ^ character. If a field is repeated, a semi-colon is used to separate the repeating fields. If a field consists of other fields (complex fields), a # is used to separate each of the subcomponents. If the subcomponents themselves consist of other fields, then a $ is used to separate them. A backslash is used as an escape character.

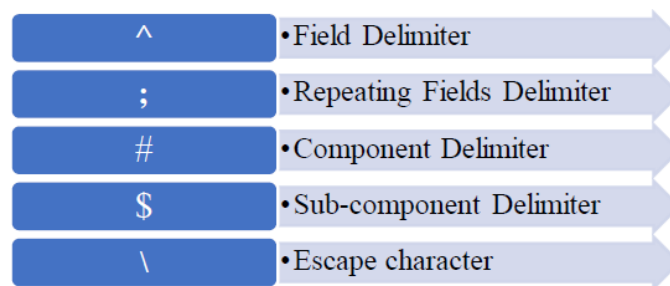| ^ | • Field Delimiter |
;	• Repeating Fields Delimiter
#	• Component Delimiter
$	• Sub-component Delimiter
\	• Escape character

FIGURE 3. Field delimiters

Messages: Messages are the primary means of communication in the proposed EDI. All interactions and communications within the framework should be represented by equivalent messages. In this paper we only present sample messages that are commonly exchanged in a university framework. The complete EDI guide should cover all framework messages. The sample messages that are defined are Register_student, Create_course, Get_student_transcript, Student_change_major and Drop_course. Sample of generated *Register_student* Message is:

```
HDR^application1^application2^RS^20180329^072600^ STD^4522855601^^929040786^Ali#Ahmed#Al-shaikh^Saudi
Arabia^22103991^Male^0580918962;0580056212#013111000 #alig@yahoo.com^Raakah #31441#Main Street#9211^Ahmed$Al-
shaikh#Father#0582829345##^Orbit Secondary School^20120811 ^99.05%^ PRG^Bachelors^College of Computer
```

4. **Testing.** In this section we will compare between the proposed EDI and HL7 in terms of message format, datatype/entities, integration, import/export, complexity, compilation, usability/reusability and standard matching.

4.1. **Message format.** Like HL7 the proposed EDI messages consist of segments and fields. Naturally, some differences exist in the proposed message format. For instance, the name of segments in HL7 must be exactly three characters; however, in the proposed EDI segment names can contain up to five characters. Additionally, the message delimiters differ between HL7 and the proposed EDI as depicted in Table 1 which shows the different message delimiters.

4.2. **Entities and datatypes.** The format and datatype of some common message fields between the proposed EDI and HL7 differ. This is due to the nature and requirements of the two different domains each serve. Table 2 illustrates a comparison between some common fields in the *register_patient* message in HL7 and the *register_student* message

TABLE 1. Delimiters comparison

Delimiter	HL7	Proposed EDI
Field Delimiter	^	|
Repeating Fields Delimiter	;	∼
Component Delimiter	#	^
Sub-component Delimiter	$	&
Escape character	/	/

TABLE 2. Common fields in HL7 and proposed EDI

Field Purpose: Patient/Student Identification					
	Field Name	**Datatype**	**R/O**	**Max. length**	**Notes**
HL7	patient identifier list	CX	—	—	
	patient identifier list. ID Number	String	R	15	
Proposed EDI	student_national_id	String	R	10	Required only with Saudi nationals. Contain only digits and must start with 1
	student_iqama_id	String	R	10	Required only with non-Saudi nationals. Contain only digits and must start with 2
	student_academic_number	String	O	10	Contain only digits and must start with 2
Field Purpose: Patient/Student Name					
	Field Name	**Datatype**	**R/O**	**Max. length**	**Notes**
HL7	Patient Name	PN	—	—	
	Patient Name.Family Name	String	O	194	
	Patient Name.Given Name	String	O	30	
	Patient Name.Second Name	String	O	30	
Proposed EDI	sudent_name	Complex	—	—	
	student_name.first_name	String	R	50	
	student_name.father_name	String	R	50	
	student_name.family_name	String	R	50	
Field Purpose: Patient/Student Address					
	Field Name	**Datatype**	**R/O**	**Max. length**	**Notes**
HL7	Patient Address	XAD	—	—	
	Patient Address.Street	SAD	O	184	
	Patient Address.City	String	O	50	
	Patient Address.State	String	O	50	
	Patient Address.Zipcode	String	O	12	
Proposed EDI	address	Complex	—	—	
	address.city	String	O	30	Formatted as defined in the Saudi National Postal Organization
	address.zipcode	Digits	O	5	
	address.street	String	O	30	
	address.house_number	String	O	10	

in the proposed EDI. Three common fields are compared: patient/student identification, name and address. The purpose of this comparison is to illustrate the robustness and capability of the proposed EDI.

The *register_patient* in HL7 contains 12 segments and the *Register_student* message contains 3 segments in the proposed EDI. When registering a patient in HL7, a patient can be identified by a composite field named patient identifier list PID-3 which is an internal patient ID that should be unique within a facility. The PID-3 field is of type CX, a datatype defined in HL7 used in defining an identifier along with its details. PID-3 has a maximum of 250-character length, which includes the identifier itself of type string and maximum length 15 characters. In field Identifier Type Code PID-3.5 within the composite field PID-3, HL7 allows to specify the type of ID, such as Social Security number, birth registry number, and organization ID. In the proposed university EDI, a student can be identified by using three fields: student_national_id, student_iqama_id, student_academic_number. The 1st is required for all Saudi nationals, which is a string of length 10 digits starting with the digit 1 and is unique for all Saudi nationals. The 2nd ID is used for non-Saudi residents, which is a unique string for all non-Saudi residents in Saudi Arabia. It has a maximum length of 10 digits starting with number 2. The 3rd identifier is the student_academic_number. It is also a string of digits starting with the number 2 and has a maximum length of 10 digits. The student academic number is only unique within the university. Another similar field is the patient/student name. In HL7 this is held in the patient_name PID-5 field, which is a composite field that contains 14 fields such as family name, given name, and second name. In the proposed EDI, this is held in the student_name complex field which contains the student's first name, father name and family name. All name sub-fields are required as opposed to HL7 which permits to register a patient without name. In HL7, this field contains 14 subfields such as street name, city, and state. In the proposed EDI the student address consists of 4 subfields: city, zip code, street and house number. The address is optional in both EDIs.

4.3. **Data integration.** Like HL7, the proposed EDI allows complete data integration. The defined standards in this study will facilitate the collaboration between different educational agencies across the country. Using the proposed EDI, data are integrated to compose fields and segments. The composed messages can be exchanged between various relating agencies. Figure 4 illustrates the EDI transaction process and data integration amongst various educational agencies. The following steps are taken to send or receive data between different related educational agencies.

1) The educational agency that wishes to send data fetches the data from internal database.
2) The fetched data is processed by the EDI translator. The processing includes formatting the data according to EDI standards (i.e., data ordering, sequencing, escape characters, etc.)
3) The EDI translator produces an EDI message.
4) The constructed EDI message is sent to the other agency. Within the same organization, point-to-point communication is used. Alternatively, communicating to external agencies is done via an EDI network.
5) The receiving agency is also connected to the sending agency.
6) The receiving agency receives the EDI message.
7) The message is processed by the EDI translator. The type of message is interpreted, and data is extracted.
8) The extracted data can be stored at the receiving agency's internal database.
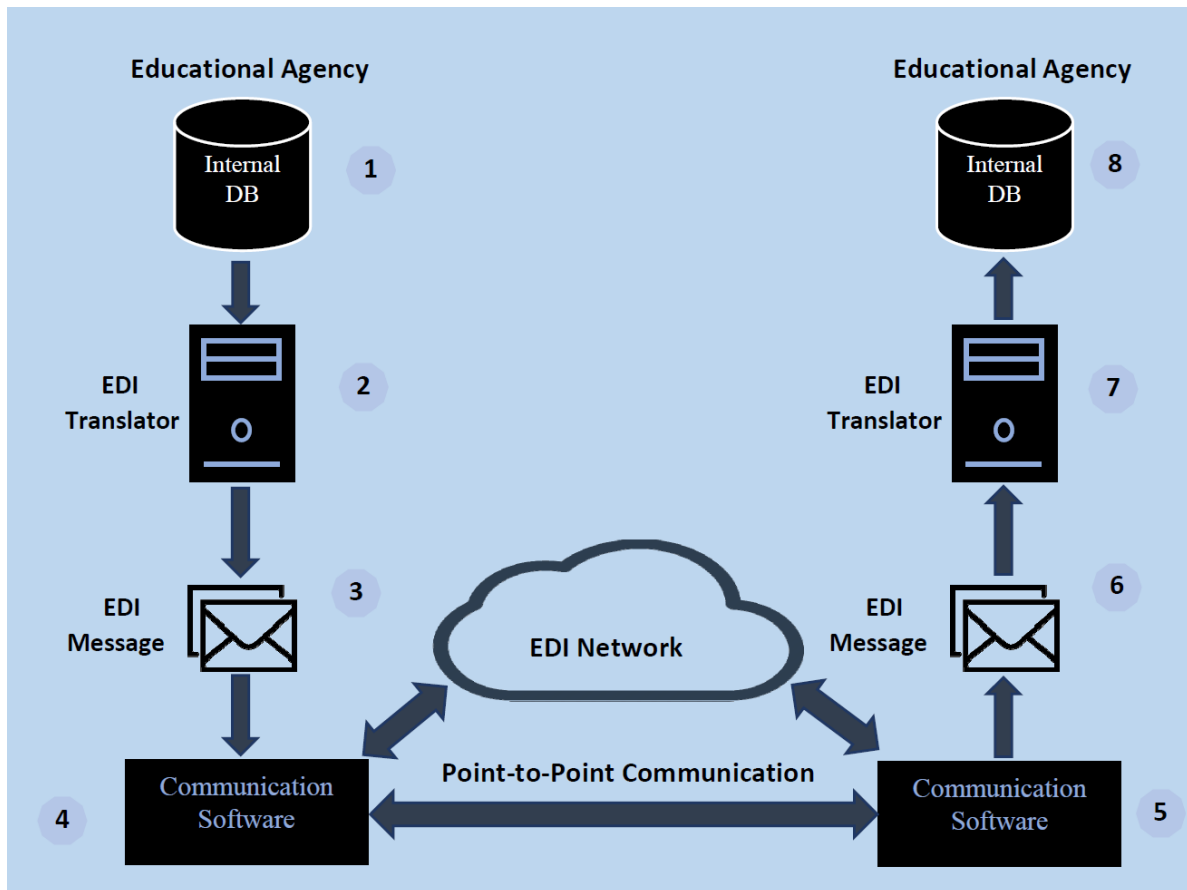
FIGURE 4. EDI integration

4.4. **Import/export.** The proposed EDI allows the import and export of data. This functionality allows transporting data from and to the EDI message interfaces. For instance, the *student_academic_transcript* message can be exported to a PDF file when needed. The process of exporting/importing data is handled by the EDI translator. When creating a new message, the data can be fetched from the internal database. Segments are created one by one by fetching their consisting field data. Delimiters are added to separate message contents.

4.5. **Complexity.** The complexity of the proposed EDI system depends on several factors, such as the message length, datatypes, message processing, data encryption and decryption, message validation and error handling. Although the EDI implementation is relatively complex, the overall benefit is significant. The complexity of the EDI system can be reduced by using efficient and simple algorithms and techniques in the implementation.

4.6. **Compilation.** The EDI messages are translated and interpreted by the EDI compiler. Based on grammar rules the compiler can distinguish the type of a message, parse and interpret the contents. Furthermore, validation is done to determine the correctness of the message format. The compiler can report errors within a received message. For instance, a missing segment or field is identified and reported. Also, a mechanism is adopted to ensure the integrity of the data, and checksums can be used to guarantee that the data is intact after data communication.

4.7. **Usability/reusability.** EDIs are used in different domains. The magnitude of their need contributes greatly to the success and acceptance of EDIs. Furthermore, the ability to reuse and update EDIs is an important factor in EDI implementation. Seeing that

requirements in organizations are continually changing, an efficient EDI should possess the ability to easily adapt and adjust accordingly.

4.8. **Standard matching.** In the proposed EDI the used standard is declared in the message header. This allows the message translator to identify the message type, and then according to the EDI standard, the message is parsed and interpreted. The well-defined message format and built grammar rules enable the standard matching efficiently. The proposed EDI matches the HL7 standard.

5. **Conclusion.** The need to electronic data exchange across educational facilities in Saudi Arabia is crucial. In this study, this problem is addressed by defining new EDI standards. The proposed EDI standards facilitate the exchange of data and files in universities, colleges, institutes, etc. The system framework was first analyzed to specify system users, entities, relationships and data fields and the data is transmitted via proposed EDI messages. The format, structure and contents of these messages were defined according to stakeholders' needs. For sake of clarification, four messages were defined. In addition, language/grammar rules were constructed to enable the complete data integration and compilation by the EDI translator and enable efficient data import/export capabilities. Illustrative demo-interfaces are provided to demonstrate the system. Finally, the proposed EDI system is tested based on the well-known HL7 standards to match its capabilities and robustness. In future, the proposed grammar can be further improved and standardized.

REFERENCES

[1] The Gale Group Inc., *Electronic Data Interchange (EDI)*, Gale Encyclopedia of E-Commerce, Encyclopedia, 2018.

[2] S. Shahzad, *What Is EDI and How Does It Work? Certificate of Declaration*, 2012.

[3] 1EDISOURCE, *EDI 130 Student Educational Record (Transcript) – X12 Transaction Set*, https://www.1edisource.com/resources/edi-transactions-sets/edi-130/, accessed on 13-Mar-2018.

[4] American National Standards Institute, *X12*, http://www.x12.org/, accessed on 13-Mar-2018.

[5] *The Total Number of Colleges, Institutes and Educational Deanships in Higher Education Institutions*, https://departments.moe.gov.sa/PlanningInformation/RelatedDepartments/Educationstatisticscenter/EducationDetailedReports/Docs/Table9-01_36-37.html, accessed on 13-Mar-2018.

[6] Saudi Arabia Ministry of Education, *The Higher Education Statistical Center*, https://www.moe.gov.sa/en/TheMinistry/Education/Institutions/Pages/TheHigherEducationStatisticalCenter.aspx, accessed on 13-Mar-2018.

[7] *What Is EDI (Electronic Data Interchange)? | EDI Basics*, https://www.edibasics.com/what-is-edi/, accessed on 13-Mar-2018.

[8] EDI BASICS, *ANSI – American National Standards Institute*, https://www.edibasics.com/ediresources/document-standards/ansi/, accessed on 13-Mar-2018.

[9] *PESC | EDEXCHANGE*, http://www.pesc.org/edexchange.html, accessed on 15-Mar-2018.

[10] P. E. S. Council, *Implementation Guide to the Postsecondary Electronic Standards Council XML Standard Format for the College Transcript*, 2008.

[11] *Health Level Seven International*, http://www.hl7.org/index.cfm, accessed on 16-Mar-2018.

[12] Atta-ur-Rahman, M. H. Salam and S. Jamil, Virtual clinic: A telemedicine proposal for remote areas of Pakistan, *The 3rd World Congress on Information and Communication Technologies (WICT'13)*, Vietnam, pp.46-50, 2013.

[13] Atta-ur-Rahman, A. Bakry, K. Sultan, M. A. A. Khan, M. Farooqui and D. Musleh, Clinical decision support system in virtual clinic, *Journal of Computational and Theoretical Nanoscience*, vol.15, no.6, pp.1795-1804, 2018.

[14] Atta-ur-Rahman and J. Alhiyafi, Health level seven generic web interface, *J. Comput. Theor. Nanosci.*, vol.15, pp.1261-1274, 2018.

[15] Atta-ur-Rahman and M. I. B. Ahmed, Virtual clinic: A CDSS assisted telemedicine framework, in *Telemedicine Technologies*, Elsevier, 2019.

[16] H. M. Faisal, M. Ahmad, S. Asghar and Atta-ur-Rahman, Intelligent quranic story builder, *International Journal of Hybrid Intelligent Systems*, vol.14, pp.41-48, 2017.

[17] Atta-ur-Rahman and F. A. Alhaidari, The digital library and the archiving system for educational institutes, *Pakistan Journal of Information Management and Libraries (PJIM&L)*, vol.20, no.1, pp.94-117, 2019.

[18] Atta-ur-Rahman and F. A. Alhaidari, Querying RDF data, *Journal of Theoretical and Applied Information Technology*, vol.26, no.22, pp.7599-7614, 2018.

[19] N. Aldhafferi, A. Alqahtani, Atta-ur-Rahman and M. Azam, Constraint based rule mining in patient claim data, *Journal of Computational and Theoretical Nanoscience*, vol.15, no.2, pp.1064-1071, 2018.

[20] N. Shahzadi, Atta-ur-Rahman and A. Shaheen, Semantic network based semantic search of religious repository, *International Journal of Computer Applications (IJCA)*, vol.36, no.9, pp.1-5, 2011.

[21] N. Shahzadi, Atta-ur-Rahman and M. J. Sawar, Semantic network based classifier of Holy Quran, *International Journal of Computer Applications (IJCA)*, vol.39, no.5, pp.43-47, 2012.

[22] Atta-ur-Rahman, K.Sultan, N. Aldhafferi and A. Alqahtani, Educational data mining for enhanced teaching and learning, *Journal of Theoretical and Applied Information Technology*, vol.96, no.14, pp.4417-4427, 2018.

[23] Atta-ur-Rahman and S. Das, Big data analysis for teacher recommendation using data mining techniques, *International Journal of Control Theory and Applications*, vol.10, no.18, pp.95-105, 2017.

[24] Atta-ur-Rahman and S. Das, Data mining for students' trends analysis using apriori algorithm, *International Journal of Control Theory and Applications*, vol.10, no.18, pp.107-115, 2017.

[25] Atta-ur-Rahman, Teacher assessment and profiling using fuzzy rule based system and apriori algorithm, *International Journal of Computer Applications (IJCA)*, vol.65, no.5, pp.22-28, 2013.