A Semantically Enriched E-Tendering Mechanism

Jingzhi Guo and Ka-Ieong Chan

Department of Computer and Information Science, University of Macau Av. Padre Tomás, Pereira, S.J., Taipa, Macau {jzguo, ma86522}@umac.mo

Abstract—E-tendering is widely used in corporate and government purchasing in e-business practice. Existing e-tendering platforms cannot solve semantic interoperability problems between heterogeneous e-tendering systems. This paper proposes a novel semantic-enriched e-tendering (SEET) approach, which enables heterogeneous e-tendering systems to be semantically connected and interoperable by applying collaborative conceptualization theory. Based on this approach, a SEET platform is implemented and guarantees that e-tender inviters and e-tender bidders can exchange their e-tendering documents in a semantically consistent manner.

Keywords-electronic tenderring; tender; inviter; bid, bidder; semantic interoperability; semantic heterogeneity; semantic consistency; e-tendering document; document engineering

I. INTRODUCTION

Tendering is an important means of procurement and is intensively researched in traditional business [14]. The paper-based method of traditional tendering has been common place within the industry for a significant number of years. Electronic tendering (or e-tendering), "in its simplest form, is described as the electronic publishing, communicating, accessing, receiving and submitting of all tender related information and documentation via the internet, replacing the traditional paper-based tender processes, and achieving a more efficient and effective business process for all parties involved" [2]. It brings the convenience, saves both time and money, eliminates transcription errors, and increases speed of bid analysis to the people.

However, despite the apparent benefits of e-tendering, many companies have been slow to adopt electronic construction of tendering, and tender industry has not kept up with other industries clearly. The BCIS e-tendering survey [1] indicated only 13% participants adopt the web-based portal in e-tender while the 60% of them still use disks or other physical media to distribute tender documents. This evidenced that e-tendering is worth studying seriously.

In e-tender industry, there are many inviters and bidders that seek cooperation opportunities in various tendering platforms, where the inviters and bidders need to use tender documents to exchange tendering information. E-tender documents are often ad hoc formatted by different e-tender platforms in terms of Request for Information (RFI) [16], Request for Tender (RFT), Request for Proposal (RFP), Request for Quotation (RFQ) [10], Expression of Interest (EOI) [15], Advanced Tender Notice (ATN) [3], and Sale Tender (ST). By observation, we found that the exchanged

documents are often semantically heterogeneous and highly need manual processing. Firstly, most documents used in etender are simply monolithic and not computer-interpretable, for example, e-mail attachment in Microsoft Word, scanned PDF or drawings in picture formats, which are only designed for human comprehension. Secondly, e-tender documents come from different platforms are not semantically consistent in concepts, because they are created in different contexts of companies and can only be locally correctly interpreted in their individual e-tender platforms. Semantic interoperability such as computer readability and computer understandability between different bidders and inviters of various e-tendering systems is key challenging issues of designing e-tendering systems. It require a feasible solution to processing inbound and outbound e-tender information.

Despite semantic interoperability problem of existing e-tendering platforms, many bidders still use heterogeneous e-tender documents to send their bids to inviters. In the situation of unavailable semantically-enriched e-tendering platforms, e-tender inviters have to accept those semantically heterogeneous bids that cannot be digitally and automatically interpreted. They have to pay high cost to manually read the bidding documents or create new forms or schema to store the received tender information that are heterogeneous, so that they can understand the received documents. These are the barriers of using existing e-tendering platforms and urgently ask for a feasible solution to the problem.

This paper aims to solve the above mentioned etendering problem by proposing a novel SEET approach. This approach defines a new e-tender information structure and apply collaborative concepts to construct e-tender documents, which allow e-tendering users to exchange tendering information in a semantically consistent way.

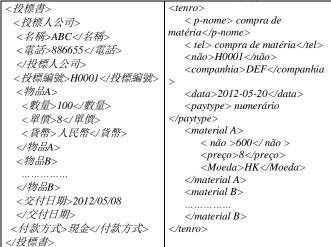
The rest of the paper is organized as follows: Section II describes the challenging research issue. Section III proposes a novel semantically-enriched e-tendering approach. Section IV discusses the related work, and finally the conclusion is made.

II. DESCRIPTION OF CHALLENGING ISSUE

Currently in industrial practice, there are at least two types of e-tender documents: one is the type of monolithic documents in plaintexts or drawings, and the other is the type of structured documents that are computer-readable but semantically heterogeneous. While believing that monolithic documents can be transformed to structured documents, a most important research issue is how an e-tendering platform can integrate heterogeneous e-tender document information

of different e-tendering systems in a semantically consistent manner such that semantically heterogeneous e-tender documents can be semantically interoperated.

TABLE I. BIDDER DOCUMENT XML (I) (J)



Consider an example illustrated in Table I where Bidder A and Bidder B use e-tender XML schemas (I) and (J) to devise their e-tender documents. Respectively, an Inviter uses e-tender database schema (K), as shown in Table II, to store received bidding information. It obvious that all of them use different languages, different terms and different schema structure. Thus, the inviter cannot understand the concepts and structures received from the bidders and it is hard to compare the bidders' e-tender documents.

TABLE II. INVITER DATABASE SCHEMA (K)

tender no	company	phone	name	spec	unit	price
H0001	ABC	886655	glass		8	100
H0002	DEF					
H0003	XXX					
H0004						

This example shows that many bidders sending many etender documents with different schemas to the inviter, and these e-tender document schemas and used terms are unknown to the inviter. It means if bidders and inviter have no prior collaboration for schema mapping and term integration, the received bidding information cannot be interpreted and shall have the following particular problems:

- The language-different bidding information in English, Chinese and Portuguese, etc. are not computer understandable.
- The meanings of different terms, such as "名稱", "投標編號", "物品 A", "數量", "單價" and "交付日期" from Bidder A and "p-nome", "tel", "não", "companhia", "data", "paytype", "preço" from Bidder B, are not understandable by the inviter though many terms are actually equivalent to "Company", "phone", "tenderNO", "name", "unit", " price", "Paydate" in the inviter's database.

The above problems can be summarized as a generic research problem of semantic interoperability of e-tendering systems or e-tender semantic consistency, that is, how a newly-designed e-tendering platform can semantically integrate heterogeneous e-tender information in a semantically consistent manner.

III. APPROACH TO SEMANTIC-ENRICHED E-TENDERING

Aiming at solving semantic interoperability problem of heterogeneous e-tendering systems, this paper proposes a novel SEET approach, which is a technical solution based on CONEX technology [6][7]. It provides novel methods of e-tender document representation and transfer by applying a set of semantically consistent common vocabularies of some natural languages. It is flexible for multiple bidders to work with any inviter without meaning ambiguity in interpreting heterogeneously formed e-tender documents.

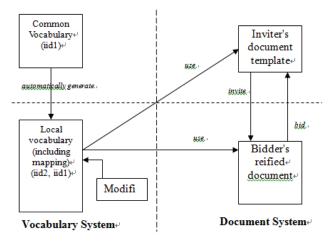


Figure 1. Seet Overview Design

Figure 1 provides an overview of SEET design. This approach divides e-tender inviters and bidders into concept designers and concept users. In vocabulary system side, designers of inviters and bidders first define their local vocabularies and then map them onto common vocabularies, which are supplied by a special vocabulary provider. In document system side, designers of inviters design their own e-tender bid templates and publish them in SEET platform. When bidders bid a tendering project, they must download corresponding e-tender bid templates and then reify them as instance bidding documents for bid submission to inviter.

SEET approach consists of the method development of SEET document representation and SEET document transfer. The former is used to create and use semantic consistent etender documents by applying common vocabulary and LCMap (locIid, comIid) of CONEX [6][7]. The latter is a protocol of how an e-tender document is sent, received and transformed. In the rest of this section, the methods of SEET document representation and SEET document transfer will be elaborated. Meanwhile, its theoretical foundation of semantic consistency will also be given.

A. SEET document representation method

SEET document is a key media to exchange e-tender information. In general, many systems are designed only for human comprehension, but lack of computer understanding. In order to achieve computer understanding, SEET develops a document representation method that separates syntax, metadata and instance data in different layers by adopting XML Product Map (XPM) [7] to develop e-tender document syntax, e-tender document templates and e-tender instance documents. It skillfully applies the hierarchical structure of XPM as document syntax and the collaborative common terms (sign of semiotics) [6] as metadata of document and individual terms for document instantiation.

XPM [7] is a sign representation language being composed of a set of terms (i.e., signs) in structure and concept (i.e., signifier and signified in semiotics [13]), which organizes the terms (or signs) in a hierarchy, such that:

$$complex \ sign = sign_0(sign_1, sign_2, ..., sign_n)$$
 (3)

XPM language can be used to directly describe a document or transforms a monolithic and/or non-standard document into a semantically readable and understandable document. SEET document representation by XPM can enable e-tender bid documents to be computer-readable and computer-understandable. Particularly, an XPM SEET document is represented in three separate layers: structure layer, pattern (template) layer and instance layer. By this representation separation, each layer can be arbitrarily designed and work independently.

Structure layer. It is a document structure (i.e., syntax) layer, which establishes a syntax foundation of designing and creating personalized templates for each e-tender inviter. The syntax of any personalized e-tender document template is as the same syntax of any complex sign of Formula (2), in which each atomic sign of Formula (1) is a tuple such that atomic sign = (tid, term, [interpretation], [context]), where tid is a structural sign identifier, term is a word or phrase, [interpretation] is a definition or annotation of a term, and [context] is a term reference to specify a context where a word or phrase is defined.

Pattern layer. It is a document template layer, which is based on the syntax layer and use the collaborative concepts (i.e., defined terms as atomic signs) of CONEX vocabulary [6][7] to develop personalized templates as abstract complex signs (Formula 3) according to the required e-tender information. E-tender inviters of e-tendering systems are responsible for defining different patterns (i.e., document templates) used in e-tendering systems.

Instance layer. It is a document reification layer, which instantiates document templates to reified documents as concrete/particular complex signs (Formula 3). During document instantiation, users (mostly e-tender bidders) fill in e-tender templates with local terms that are mapped onto common terms of CONEX vocabularies [6][7]. Since local terms are different in natural languages, local reified e-tender documents are semantically interoperable via common terms between different natural languages and dialects, thus

maintaining semantic consistency between heterogeneous systems.

With the above-mentioned three layers, an e-tender document can be displayed in two modes: computer display mode and human display mode. The former is a mode of computer-understandable for computer to use, utilizing internal identifiers (IID) of atomic signs (Formula 2). It mainly handles computer processes when e-tender documents are transformed and stored. The latter is a mode of visual and human-readable. It displays e-tender documents in human-readable and human-understandable natural languages, utilizing terms of atomic signs (Formula 2). Table III is an example of a simplified XPM file in computer display mode and Table IV is an example of the simplified XPM file in human display mode.

TABLE III. COMPUTER DISPLAY MODE (A SIMPLIFIED XPM FILE)

TABLE IV. HUMAN DISPLAY MODE (A SIMPLIFIED XPM FILE)

```
<xpm:sign xpm:term="Requirement">
<xpm:sign xpm:term="Object_A">
<xpm:sign xpm:term="Quantity"/>
<xpm:sign xpm:term="UnitPrice" />
</xpm:sign>
<xpm:sign>
pm:term="Payment Date"/>
```

B. SEET document transfer method

SEET document transfer method describes how an etender document is transferred from one local e-tendering system to another local e-tendering system, yet maintaining semantic consistency in the meaning of e-tender documents.

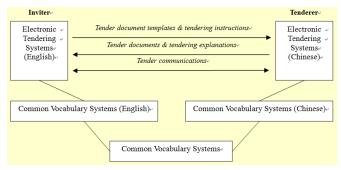


Figure 2. Seet Document Transfer Method

Figure 2 illustrates the e-tender document transfer protocol, which guarantees that the exchanged e-tender documents are semantically consistent without meaning ambiguity. It shows the exchange process of e-tender document templates and reified documents between an inviter and possibly many bidders in different natural languages.

SEET document transfer method provides an automatic transparent mechanism. It maintains semantic consistency of document information through mapping local terms of etendering documents onto terms of common vocabularies with same internal identifiers (IID) [6]. During document sending in document exchange, all terms of a local e-tender document in local natural language will first be transformed to local term IDs (locIids) based on local vocabulary, and then transformed to common term IIDs (comIids) of common vocabulary based on local-to-common term mapping datasets of LCMAP(locIid, comIid) at sending company. During document receiving in document exchange, all terms of a common e-tender document will first be transformed from comlids to loclids based on LCMAP (locIid, comIid) at receiving company, and then transformed to local terms in local natural language for human interpretation based on local vocabularies.

For example, given two vocabularies and a mapping LV_1 (L_IID: $10017 \rightarrow name$), CV_1 (C_IID: $20012 \rightarrow name$), and $LCMAP_1$ (10017, 20012), and two vocabularies and a mapping LV_2 (L_IID: $10011 \rightarrow 4$), CV_2 (C_IID: $20012 \rightarrow 4$), and $LCMAP_2$ (10011, 10012), then based on a document transfer process of $LV_1 \leftrightarrow LCMAP_1 \leftrightarrow LCMAP_2 \leftrightarrow LV_2$, we have (10017, 10017, 10017, 10017, 10011, 10017, 10011, 10017, 10011

C. Theoretical foundation of SEET approach

SEET approach is built on a proven theoretical foundation of semantic consistency model of collaborative conceptualization theory [6], that is, structure mapability, concept equivalence and context commonality.

First, common terms of common vocabularies in different natural languages are collaboratively designed, where all common terms in different natural languages share a same common identifier (comIid). This guarantees that any common vocabulary is identical in meaning interpretation and can be unambiguously used by local vocabulary designers in meaning when making local-to-common term mapping. It further guarantees semantic consistency between local terms and common terms, that it, concept equivalence.

Second, semantic consistency is not only maintained in single local-to-common term mapping but also in e-tendering document templates. This owes to the generic XPM language that is used to design e-tendering document templates such that all e-tendering document templates are designed by e-tendering inviters. E-tendering bidders only apply the e-tendering document templates to reify the templates to produce instance e-tendering documents. This guarantees that all e-tendering document processing structures are the same, that is, structure mapability onto a same hierarchy of sign IDs for executables to have same execution effects in computing.

Third, collaborative common term design, local-tocommon term mapping and common e-tendering document templates provide a common context for all natural language-different e-tendering systems. Such common context guarantees that all meaning interpretations on all local receiving e-tendering documents by the receivers are exactly the same as the meaning interpretations on all local sending e-tendering documents by the senders.

Following semantic consistency model of CONEX, any local e-tendering document of one e-tendering system can be semantically represented and transferred to another e-tendering system for local interpretation without the loss of the original meaning.

IV. SEET COLLABORATION PLATFORM

Based on the SEET approach designed above, this section implements the SEET platform illustrated in Figure 3, which consists of a layer of user web-interface and a layer of concept collaboration.

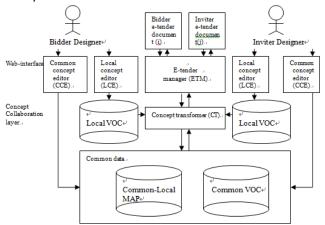


Figure 3. SEET Collaboration platform

A. Layer of User Web-Interface

The layer of user web-interface consists of concept editor and e-tender manager.

Concept Editor. It consists of a common concept editor (CCE) and a local concept editor (LCE). The former is engaged in designing common terms of the corresponding common vocabulary (common VOC). The latter is responsible for designing local terms for local vocabulary (local VOC) and builds local-to-common maps (Common-Local MAP).

E-tender Manager (ETM). It consists of document manager and tender project manager. The former controls the document representation and e-tender process. The latter manages tender project for users to be easy to read e-tendering document and observe e-tender project status.

B. Layer of Concept Collaboration

The layer of concept collaboration consists of many local vocabularies, many common vocabularies, local-to-common mapping repository and concept transformer.

Local Vocabulary (local VOC or LV). It is designed for a company to record its local term information. It is composed of local concepts written in XPM and structured in a triple, such that:

$$LV_{XPM} = (locIid, AN, FC)$$
 (4)

where "locIid" is a unique identifier of a local concept. "AN" is the definition of the locIid-ed concept. "FC" is a formal concept in the form of a word or a phrase that is defined by AN and is readable and understandable by human.

Common Vocabulary (common VOC or CV): It is designed for a natural language to record its common term information universal to all natural languages. It is formatted in XPM and structured in a triple, such that:

$$CV_{XPM} = (comIid, AN, FC)^{LANG}$$
 (5)

where "comIid" is a unique internal identifier of a common concept. "LANG" indicates that AN and FC are written in natural language of LANG.

Common-Local Map repository (Common-Local MAP or CLM): It stores local-to-common mapping results that is created by local users on local concept editor. It is formatted in XPM and structured in a couple, such that:

$$CLM_{XPM} = (LocIid, ComIid)$$
 (6)

where LocIid is semantically equivalent to ComIid.

Concept transformer (CT): It implements the SEET document transfer method that transforms one local e-tender document of one e-tendering system to another local e-tender document of another e-tendering system without losing semantic consistency.

The implementation of SEET platform assures that any local e-tender bids can be submitted to a corresponding e-tender inviter for semantically consistent interpretation without semantic ambiguity.

V. FEATURES OF SEET PLATFORM

SEET platform presents some important features comparing with traditional paper-based tendering systems and existing e-tender websites. These features are information integration, flexibility, availability, semantic unambiguity, compatibility, extensibility and maintainability, which are shown in Table V.

TABLE V. COMPARISON OF PAPER BASE, E-TENDER BOX AND SEET

Features	Paper base	e-tender Box [4]	SEET	
Main Technology	Paper base	Doc, xml, pdf	XPM	
Information Integration	none	part	whole	
Flexibility	Low	Medium	High	
Cost	Low	High	Medium	
Semantic unambiguity	Low	Medium	High	
Compatibility	Low	Medium	High	
Expansibility	Low	Medium	High	
Maintainability	Low	Medium	High	

SEET platform can provide these features because SEET approach is designed on a generic method of document structuring and term collaborative editing. Firstly, SEET focus on the capability of maintaining semantic consistency above e-tendering platform. It bases on CONEX technology to effectively solve two specific problems. (a) Natural language problem: many bidders and many inviters are located in different countries and regions and adopt different

natural languages for bidding and invitation in e-tenders. (b) Inconsistent interpretation problem: inviters and bidders interpret incoming tendering documents differently from the interpretation of the document originators.

Second, SEET platform emphasizes on semantic unambiguity and flexibility features for both bidders and inviters such that they share same meaning for a same yet differently represented document. For a concept-same etendering document, it can be feasibly represented as heterogeneous documents in both forms and terms. Architecturally speaking in design, it clearly separates document syntax from document templates and separates document templates from document instance. By this separation, e-tendering documents are easy to be maintained, extended, and be compatible with existing language-different documents.

Third, SEET platform pays high attention on contexts. It represents e-tendering documents in two modes of computer display and human display. This allows contextual interpretation of e-tendering documents in heterogeneous contexts or system environments and increases compatibility of e-tendering documents in different companies.

VI. RELATED WORK

A. Electronic Tendering

Tendering has been gradually transformed from paperbased tendering into electronic tendering. Companies of many countries have already made many efforts of building a of e-tendering platforms on Web (e.g., tenders.nsw.gov.au, unitender.com, and electronictender.com). However, existing e-tender platforms are far more than perfect. Most e-tendering systems combines drawings and PDF files with computer-readable documents. Some provide online tendering information to bidders together with terms and conditions. For example, e-Tender Box (ETB) [4] provides online services for suppliers to view tender notices, standard terms and conditions for tenders arranged by Government Logistics Department (GLD) as well as contract award notices of the bureau and departments of the Hong Kong Special Administrative Region Government.

Research on e-tendering focuses on interoperability among heterogeneous distributed e-tendering systems. Interoperability issue can be divided into sub-issues of mapping different e-tender database schemas, intelligent integration of e-tendering information, and electronic document exchange among different parties [9]. In another classification, e-tender interoperability problem can be discussed by the aspects of technical interoperability (refers to the ability to connect system by defining standard protocols and data formats), semantic interoperability (refers to the exchange of information in an understandable way within and across organizational borders), and organizational interoperability (refers to enabling processes to cooperate) [11]. Amongst these aspects, semantic interoperability is most challenging.

B. Semantic Interoperability

Semantic interoperability between distributed heterogeneous systems is one of the most important research problems in research areas of database, semantic web and electronic commerce. Earlier researches focus on the work in heterogeneous database systems (Kashyap and Sheth [8]), workflows and service oriented architectures (Nagarajan et al [12]). They explore approaches to heterogeneous database schemas integration and techniques of XML schema mediation (Gomadam et al [5]). Nevertheless, since information sources for mediation are mostly created and run in different contexts, predefined mediation rules cannot cover all heterogeneous source information, particularly the information created after the making of mediation rules. This leads to semantic conflicts between context-different distributed systems.

Collaborative concept exchange (CONEX) [6][7] is an approach to collaborative conceptualization of heterogeneous concepts of disparate information sources for consistent semantic interoperation. It introduces collaborative editing of heterogeneous concepts of different systems as the core methodology of resolving semantic conflicts. There are advantages of applying CONEX in designing e-tendering platform. It can inherit the key feature of semantic consistency between documents from CONEX approach.

VII. CONCLUSION

This paper has proposed a novel SEET approach to solving a challenging issue of semantic interoperability for etendering document exchange between semantically heterogeneous e-tendering systems, which consist of e-tender inviters and e-tender bidders who are located in different countries and regions and write e-tendering documents in different natural languages. SEET approach enables semantic interoperability by integrating heterogeneous e-tendering systems of e-tender inviters and e-tender bidders through novel SEET methods of e-tendering document representation and transfer based on the theory of collaborative conceptualization [6]. E-tendering document representation method has applied XML Product Map (XPM) to develop etender document universal syntax and utilize collaboratively designed common terms of CONEX [7] to build semantically consistent mapping between common terms and local terms and to construct e-tender document templates. Etendering document transfer method has built a document transfer protocol that guarantees that any exchanged document is semantically consistent between document sender and document receiver. Based on SEET approach, the implemented SEET platform has several good features, such as semantic unambiguity, flexibility and compatibility, comparing with paper-based tendering systems and existing Web-based e-tendering platforms.

Contributions made in this paper are: (1) applied collaborative conceptualization theory to design semantically consistent e-tendering systems, (2) proposed a SEET design

approach, (3) implemented a SEET platform of e-tendering systems.

SEET approach in this paper is still evolving. In future, more stringent work will be provided in the aspects of business model development of e-tendering, technical implementation of various patterns of e-tendering process, and evaluation matrix development on various e-tendering routine processes.

ACKNOWLEDGMENT

The research work reported in this paper has been partially supported by University of Macau Research Committee research grant number MYRG069(Y2-L2)-FST12-GJZ.

REFERENCES

- [1] BCIS, "2009 e-Tendering Survey Report," RICS, 2009.
- [2] S. Christensen and W. Duncan, "Maintaining the integraity of electronic tendering - reflections on the capacity of the Australian legal framework to meet this challenge," eLaw Journal, Murdoch University Electronic Journal of Law, 13(2), 2006, pp. 8-36.
- [3] Deirdre, "MIA Consulting Services. Using Advanced Tender Notices," MIA Consulting Services. 20/10/2011. http://www.miaconsultingservices.com.au/using-advanced-tendernotices.
- [4] E-tender box, "Government Logistics Department (GLD)," Hong Kong. http://www.gldpcms.gov.hk. Access on 3rd Jun 2012.
- [5] K. Gomadam, A. Ranabahu, L. Ramaswamy, A. Sheth and K. Verma, "Mediatability: Estimating the degree of human involvement in xml schema mediation," in: Proc. ICSC, 2008, pp. 394–401.
- [6] J. Guo, "Collaborative Conceptualization: Towards a Conceptual Foundation of Interoperable Electronic Product Catalogue System Design," Enterprise Information Systems 3(1), 2009, pp.423-438.
- [7] J. Guo, "Collaborative Concept Exchange," VDM Publishing, Germany, 2008.
- [8] V. Kashyap and A. Sheth, "Semantic and schematic similarities between database objects: A context-based approach," VLDB J., 5(4), 1996, pp. 276–304.
- [9] A. Kayed and R. Colomb, Infrastructure for Electronic Tendering Interoperability, 1999, doi:10.1.1.22.4237.
- [10] S. Mhay and C. Coburn, "Request for...Procurement Processes (RFT RFQ RFP RFI)," The Negotiation Experts. http://www.negotiations.com/articles/procurement-terms/.
- [11] A. Mondorf and M. Wimmer, "Interoperability in e-Tendering: The Case of the Virtual Company Dossier," In: Proc. ICEGOV 2008, ACM Press, 2008, pp. 110-115.
- [12] M. Nagarajan, K. Verma, A. Sheth, and J. Miller, "Ontology driven data mediation in web services," Int. J. Web Service Res., 4(4), 2007, pp. 104–126.
- [13] F. Saussure, "Course in General Linguistics", Open Court Publishing Co., Chicago, 1986.
- [14] C. Thorpe and J. Bailey, "Commercial contracts, A practical guide to deals, contracts, agreements and promises", Woodhead, Cambridge England, 1996.
- [15] UN Procurement Division, "Expression of Interest". http://www.un.org/Depts/ptd/eoi.htm.
- [16] Walmart, "Request for Information". http://media.npr.org/assets/blogs/health/images/2011/11/Walmartheal thpartnerships.pdf.