PERSONALIZABLE MODELS FOR BUSINESS-TO-BUSINESS INFORMATION INTERCHANGE IN COMMUNICATION LAYERS

TING TIN TIN

UNIVERSITI SAINS MALAYSIA

2011

PERSONALIZABLE MODELS FOR B2B INFORMATION INTERCHANGE IN COMMUNICATION LAYERS

by

TING TIN TIN

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

November 2011

ACKNOWLEDGEMENTS

In the first place, I would like to record my heartfelt gratitude to my supervisor, Dr. Vincent Khoo Kay Teong, and co-supervisor, Dr. Cheah Yu-N for their guidance and advice in helping me to get this research done. My appreciation and gratitude to my faculty dean, Prof. Rosni Abdullah for her moral support extended to me during this research.

I am also thankful to the Ministry of Science, Technology and Innovation for the financial support for this research through the National Science Fellowship. My appreciation also goes to the School of Computer Sciences and the Institute of Postgraduate Studies for providing me with various assistance and facilities.

I would like to extend many thanks to my examiners, Dr. Shahida Sulaiman, Associate Professor Dr. Chan Huah Yong, and Professor David Ngo, for their kindness and patience in reviewing my thesis. Thanks to the RosettaNet Engineering Team and Marketing Team in Malaysia for granting me the comprehensive materials, seminars and training. I would like to gratefully thank Mr. William Tay from Intel, Mr. SY Lee, Mr. Peter Oh, Dr. Chuah from Polar Twin Advance, Dr. Kong Voon Pang, and Mr. Danny Co Cheng, for their kind help in offering industrial experiences and valuable comments. I want to express my deepest appreciation to my colleagues, Mr. Johnson Foong and Mr. Kor Chan Hock for their assistance.

I would also like to acknowledge my dear parents, Mr. Ting Lik Kiew and Mdm. Wong Yik Mei for their love throughout these years without which I would not be what I am today.

Finally, glory be to God for His love, mercy and grace.

DECLARATION

I hereby declare that I have conducted, completed the research work and written the dissertation entitled “Personalizable Models for B2B Information Interchange in Communication Layers”. I also declare that it has not been previously submitted for the award of any degree or diploma or other similar title of this for any other examining body or University.

Name of Student: Ting Tin Tin Signature:

Date:

Witnessed by

Supervisor: Dr. Vincent Khoo Kay Teong Signature:

Date:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AS2	Applicability Statement 2
ASP	Application Service Provider
B2B	Business to Business
B2C	Business to Consumer
BPEL	Business Process Execution Language
DB	Database
ebXML	Electronic Business eXtensible Markup Language
EDI	Electronic Data Interchange
FTP	File Transfer Protocol
HTML	Hypertext Markup Language
ICT	Information and Communication Technology
IP	Internet Protocol
IS	Information Systems
ISDN	Integrated Services Digital Network
ISE	Information Systems Engineering
ISO	International Organization for Standardization
MNC	Multi National Company
OASIS	Organization for the Advancement of Structured Information Standards
OFTP	Odette File Transfer Protocol
PESC	Postsecondary Electronic Standards Council
PI	Personalization Index
PIP	Partner Interface Process
PO	Purchase Order
R&D	Research and Development
RAE	RosettaNet Automated Enablement
RNIF	RosettaNet Implementation Framework

RSS	Really Simple Syndication
SME	Small and Medium-sized Enterprise
SOA	Service Oriented Architecture
SQL	Structured Query Language
TCP/IP	Transmission Control Protocol/Internet Protocol
TPA	Trading Partner Agreement
UN/CEFACT	United Nations Centre for Trade Facilitation and Electronic Business
W3C	World Wide Web consortium
WS	Web service
XML	eXtensible Markup Language
XOT	X.25 over TCP

MODEL-MODEL YANG BOLEHDIPERIBADIKAN UNTUK PERTUKARAN MAKLUMAT *BUSINESS-TO-BUSINESS* DALAM LAPISAN-LAPISAN KOMUNIKASI

ABSTRAK

Piawaian *Business-to-Business* terbukti dapat meningkatkan kecekapan prosedur penukaran dokumen dengan menghapuskan proses-proses manual. Akan tetapi, didapati piawaian integrasi dalam lapisan-lapisan komunikasi itu bersifat tidak fleksibel. Tesis ini telah mengambil inisiatif untuk meninjau piawaian-piawaian yang sedia ada dengan mempromosi dan menyepadukan konsep-konsep permodelan yang bolehdiperibadikan itu dengan prosedur penukaran dokumen yang telah distandardisasikan. Penyelidikan ini bertujuan untuk menghasilkan suatu rangka kerja konsep untuk permodelan yang diperibadikan. Pemerolehan dan penjanaan model-model *push-pull* dapat divisualisasikan dengan gambar rajah Venn tiga lingkaran dan diwakili secara formal melalui Teori *Set*. Cabaran utama ialah penghasilan satu set peraturan yang dapat mengelakkan percanggahan dalam hubungan antara pengguna dan watak dalam setiap model. Untuk menyediakan suatu sarana konsepsual yang membandingkan model, satu set notasi dicadangkan untuk mengenalpasti kekuatan dan kelemahan ciri-ciri pemperibadian. Notasi ini digunakan dalam formula-formula khas untuk menganggar Indeks Pemperibadian setiap model. Cabaran dalam proses penganggaran ini adalah dari segi pengenalpatian ciri-ciri pemperibadian dalam pelbagai struktur, kandungan dan tujuan, serta kaedah pengimbangan antara kekuatan dan kelemahan ciri-ciri pemperibadian. Dengan penggabungan Indeks Pemperibadian ke dalam piawaian *Business-to-Business*, penyelidik-penyelidik dapat membandingkan pelbagai model secara kuantitatif. Didapati bahawa model-model *push-pull* yang dicadangkan mempunyai Indeks Pemperibadian 0.6, berbanding dengan DINET 0.5 dan RosettaNet 0.4. Hal ini telah membuktikan bahawa model *push-pull* yang disarankan itu mempunyai aras kebolehperibadian yang terbaik, dengan rujukan kebolehperibadian tertinggi bertahap 1.0.

PERSONALIZABLE MODELS FOR BUSINESS-TO-BUSINESS INFORMATION INTERCHANGE IN COMMUNICATION LAYERS

ABSTRACT

Business-to-Business standards have been proven in increasing the efficiency of the document interchange procedures by eliminating human interventions. However, the integration standards contain inherent inflexibilities in the communication layers. This thesis has attempted to review the existing standards by promoting and integrating possible personalizable modelling concepts into the standardized document interchange procedures. A conceptual framework for personalization modelling is the main research outcome. In particular, the push-pull models derivation and generation can be visualized by 3-circle Venn Diagrams and more formally expressed using the Set Theory. The main challenge is to produce a set of rules that could avoid user-role relationships contradiction within each model. In order to provide a conceptual means for comparing models, from different sources, a set of notations for specifying the strengths and weaknesses of the personalization features is proposed. The notations are used in a specially derived formula to estimate each model's Personalization Index. The challenges encountered in the estimation process are the identification of personalization features from diverse structures, contents and intentions, and the way to balance between the strengths and weaknesses. With the incorporation of personalization indices into B2B standards, researchers would then be able to quantitatively compare different types of models. It was found that the proposed set of push-pull models has a Personalization Index of 0.6, compared to 0.5 for DINET and 0.4 for RosettaNet. This shows that the proposed push-pull model has the best personalizability level, with 1.0 being the highest possible personalization level.

CHAPTER 1

INTRODUCTION

1.1 Overview

In this competitive world of business, profit and revenue are the vital and critical driving forces for determining the direction and survival of an organization. Technologies brought about by the Information and Communications Technology (ICT) have been tangibly proven to bring attractive returns of investment regardless of short term or long term returns and thus, they have been widely adopted by the large enterprises. This adoption has brought about remarkable results especially in the Business-to-Business (B2B) electronic commerce (e-commerce) industry. Today, the parties in B2B e-commerce exchange the business transactions electronically utilising the Internet, Intranet, Extranet or private networks as its media (Turban et al., 2010). The salient features of the B2B e-commerce is that it automates the processing of business transactions. The processed business transactions have become the key asset for any organization especially in making critical and vital decisions for formulating effective business planning and strategy. Therefore, B2B information interchange among organizations is an essential and critical process in their day-to-day operations.

Electronic Data Interchange (EDI) is an established technology in B2B information interchange. However, one major drawback inherited from EDI is the great variety of different protocols adopted by different groups of users. Consequently, the complexity of systems integration increases exponentially when the number of trading partners grows (Strong, 2005; Channabasavaiah et al., 2003). However, all is not lost when the standardized process (B2B standards) come to its rescue to simplify the integration process. With this development, B2B information interchange sees a new beginning in its adoption among the

organizations but even with this breakthrough, the standardization system still contains and inherits certain weaknesses.

RosettaNet is one of the standards organization formed by a group of worldwide companies who directly and indirectly involve in the B2B information interchange in the high-end semiconductor industries (RosettaNet, 2007b). They have found that enterprises investing in B2B standards are mostly Multi-National Companies (MNCs) (RosettaNet, 2007b; Fuks & Wieczerzycki, 2006). However, some eighty percent of their trading partners are the Small and Medium-sized Enterprises (SMEs), who are reluctant or do not have the financial and human resources to invest in B2B standards (Fuks & Wieczerzycki, 2006). The main reason for this phenomenon is that B2B standards have restricted the implementation (Fuks & Wieczerzycki, 2006; JBoss RedHat, 2009). This thesis presents the research into those restrictions and proposes a personalization modelling conceptual framework as a possible solution to mitigate those restrictions.

1.2 Statements of Research Problem

In order to overcome the problem of inflexibility in the B2B integration standards, the statements of research problem are:

(i) How to incorporate personalizability into B2B standards through the three communication layers?

(ii) How to possibly formalize personalizability in B2B standards?

(iii) How to conceptualize personalizable models?

Personalizability in this research is defined as the degree of flexibility to control, alter, modify, or reuse certain communication layers in B2B standards based on the user (both the sender and receiver) requirements. Thereafter, personalizability and flexibility are used interchangeably. Communication Subject refers to the business documents interchanged

between two organizations. The infrastructure, system, framework or technology used by the organization to interchange the business documents is called Communication Channel. Communication Attributes refer to the configuration of the B2B information interchange process such as the Internet connection, and time to response. A solution hereby denotes a B2B information interchange technology which includes any communication layers. The main research problem is how to formalize the incorporation of certain level of personalization into three communication layers of standardized B2B information interchange.

1.3 Research Objectives

This research focuses on personalization modeling. Hence, the identified objectives of this research are as follows:

(i) To research into the role of personalizability in standardized B2B information interchange.

(ii) To define a more formal means for personalization in B2B information interchange models.

(iii) To establish possible components of a personalization modeling framework.

The first objective of this research is to clarify the role of personalizability in enhancing standardized B2B information interchange. In order to fulfil the objective, this research is seeking to answer the questions such as challenges in standardization adoption, the relationship between personalization versus standardization, and research challenges in personalizing B2B standards. The second objective of this research is to define a more formal means for personalization in B2B information interchange. This requires input from questions such as various personalization strategies, personalizable characteristics, and the measurement of personalizability level. This leads to the third objective which is to establish possible components of a personalization modeling framework. The framework must

encompass the outcome of the second objective which are the strategies to generate personalizable models, identification of personalization features, and the formula to measure personalizability level. In the long run, it is hoped that the personalization modelling framework will eventually produce personalization models that are more easily comparable.

1.4 Research Methodology

This thesis adopts the research methodology from Vaishnavi & Kuechler (2004). It includes five steps: awareness of a research problem (proposal), suggestion (tentative design), design and development (artefact), evaluation (performance measures) and conclusion (results).

1.4.1 Identification of Research Problem

The first step to carry out this research is to identify the research problem. Different B2B standards are studied and RosettaNet is chosen due to its worldwide adoption and availability in Malaysia. Similarly, related publications on RosettaNet (both academia and industry) are studied. Besides using Google Scholar search engine, databases such as Scopus, ACM portal, IEEE Explore, Elsevier, ISI Web of Knowledge, and Springer are utilized to access the published works. Based on the literature review, related works are classified into B2B standards, B2B information interchange, B2B e-commerce, personalization, push-pull modelling, software agent, Service Oriented Architecture (SOA), performance evaluation, and implementation.

1.4.2 Proposing a Solution

This research has identified the techniques/approaches to be adopted/adapted in order to realize the personalizable standardized B2B information interchange as shown in Figure 1.1. The business process of standardized B2B information interchange (RosettaNet) is

examined to observe the relationships between the users and their roles (Yan et al., 2008; RosettaNet, 2007b). A Set Theory is adopted to represent the different push-pull models. Venn Diagrams are used to generate and visualize different models. Different B2B information interchange research works are then reviewed to extract and compile the relevant personalization features to form a list of feature notations for three communication layers (Cixing & Zhu, 2009; Yan et al., 2008; Kotinurmi, 2007; Fuks & Wieczerzycki, 2006; Wang & Song, 2006; Papazoglou & Ribbers, 2006; Simmons, 2004; Masud, 2003; Willaert, 2001; Tikkala, 2004; DINET, 2010; Surfcontrol, 2007; Damodaran, 2004; Söderström, 2003; Schönberger, 2006; Shi & Chen, 2007; Riemer & Totz, 2001, 2003). Finally, from the list of strengths and weaknesses, an index-based evaluation is proposed to evaluate each model (Kumar & Stecke, 2008).

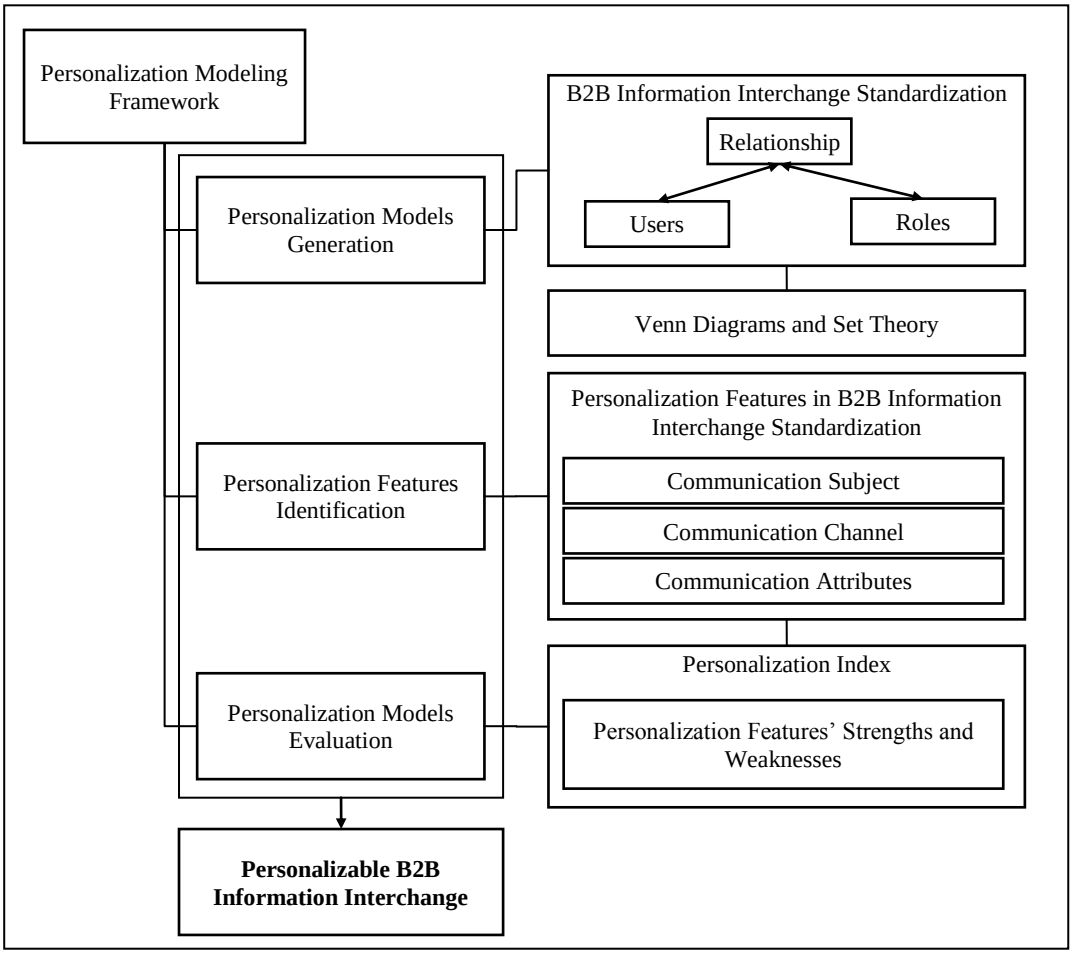

Figure 1.1: Theoretical framework of the proposed solution.

1.4.3 Prototype Development

A conceptual prototype is developed based on the push-pull models generated from the framework. Since the research team is working with an educational institution (Universiti Sains Malaysia) and there is an existing educational prototype, the case study is carried out in that environment. The outcome of the prototype aims to provide a solution adaptable by the business entities.

1.4.4 Evaluation

In this thesis, the generated push-pull models are evaluated by the computation of a Personalization Index. The proposed Personalization Index is modified from Kumar & Stecke (2008) by taking into account both the strengths and weaknesses of personalization features. Considering both the strengths and weaknesses in the Personalization Index is vital in order to enable the researchers to examine the different effects of personalization in each solution. Without the discussion on weaknesses it will result in a 'too good to be true' situation. A solution that offers many features and options is certainly offering a larger number of variants and hence performing better in terms of personalization (Kumar & Stecke, 2008). Thus the measurement of personalization is to compute the number of options offered by a solution. The algorithms also allow the user to assign weights to the features. The Personalization Index of a model with more features is potentially to be decreased by the weakness index if the solution could not overcome the challenges aroused in the adopted personalization features. Therefore, the formula will help the researchers to consider rationally the extent of personalizability to be adopted in a model. Besides, the response time of a model is also recorded to assess its performance after incorporating the personalization features.

1.4.5 Conclusion

The result of the computation of the Personalization Index is summarized into a personalization cube to highlight the comparison between different models. Now personalization modeling is formally captured into the framework and the outcomes (different push-pull models) are evaluated and compared.

1.5 Research Contribution

The main contribution of the research is a personalization modelling framework. The framework consists of the following:

(i) Possible personalizable push-pull models through role-based 3-circle Venn Diagrams and Set Theory, with the appropriate rules.

(ii) Notation for personalization features and inherent weaknesses.

(iii) Personalization Index formula to estimate personalizability level of a model.

(iv) Categorization of research works into a three-dimensional cube with Personalization Index as the unit of measurement.

1.6 Thesis Outline

Chapter 1 summarizes this research work by firstly providing an overview of the research background and issues followed by identifying the research problems and objectives. The research methodology focusing on evaluation is presented. Finally the contributions of this research to the academia and industry are summarized.

Chapter 2 covers the literature survey on the existing B2B information interchange Research and Development (R&D) works. This chapter includes the background and evolution of B2B interaction, emerging technologies, research challenges, related works, adopted technologies and standards, overview on personalization research works,

categorization of related works into a 3-circle Venn Diagrams to expose the lack of research in certain aspects and thus highlight the need for this research.

Chapter 3 presents a conceptual framework of personalization modeling that encompasses the processes of incorporating personalization into standardization. Models are generated from a role-based 3-circle Venn Diagrams and represented by a Set Theory. A list of personalization features and options in each communication layer is defined with the associated weaknesses and drawbacks. A Personalization Index is then proposed to evaluate the models. Finally, the design of the proposed push-pull modeling of a personalizable B2B information interchange is then introduced.

Chapter 4 presents the conceptual prototype development and evaluation. The realization of personalizability in the standardized B2B information interchange is highlighted in this chapter through Personalization Index computation. Models with their Personalization Index can then be categorized in a three dimensional cube. This is followed by a section to present the performance of the prototype.

Chapter 5 concludes this research with a discussion on the issues aroused from this research. This is followed by the foreseeable immediate future work of this research.

CHAPTER 2

PERSONALIZABILITY IN STANDARDIZED B2B INFORMATION INTERCHANGE

2.1 Introduction

This chapter covers a comprehensive literature review of B2B information interchange R&D pertaining to the low degree of flexibility in the communication layers that might have inhibited the pervasiveness of standardization adoption among the SMEs. In order to resolve the challenges encountered in B2B information interchange, this thesis explores the possibility of integrating personalization with standardization by analysing the relationships and roles of personalization with B2B standards and the three communication layers. The research niche is identified from the categorization of previous works in a communication-layer intersection diagram. Personalization approaches and techniques are discussed for possible adoption in this research.

2.2 Evolution in B2B Information Interchange

The evolution of B2B information interchange can be summarized into three stages namely the Electronic Data Interchange (EDI), the basic e-commerce model, and the B2B e-commerce (Sawhney, 2000). In the first stage of B2B evolution, EDI was a one-to-one transaction with closed/private infrastructure/access. The second stage is the basic e-commerce model using the one-to-one approach which leveraged the open/public infrastructure/access. In the third stage, the e-market or B2B e-commerce like the eHub brings many trading partners together to form a community which executes the transactions in the nature of many-to-many interactions through a mediating agent (Sawhney, 2000). This third stage of B2B e-commerce can either be manipulated as a closed/private or open/public infrastructure/access, as in the case of ChemConnect.com (ChemConnect, 2007)

and ChemDex.com (ChemDex, 2007) which create a virtual market place for enterprises to sell and buy chemical products. These development catalysts conceive the next level of B2B evolution which now has become an open source infrastructure with many-to-many transactions equipped with an additional feature comprising personalizable services. This research attempts to propose such an open source infrastructure with many-to-many transactions which is not limited to closed/private access.

Since the mid 1980s, B2B standards have been playing a key role in increasing the performance of inter-organization interoperability (Söderström, 2001; Cargill, 1999; Filos & Banahan, 2001; Van der Aalst, 1999). Hypertext Markup Language (HTML) and eXtensible Markup Language (XML) have been used by enterprises to exchange information on the web and integrate their business processes. Subsequently, standards consortium emerged to define and develop the specific standards of the industry. This is shown in RosettaNet Partner Interface Processes (PIPs) creating standards for high-tech industries (RosettaNet, 2007a), OASIS Electronic Business Extensible Markup Language (ebXML) for general business (OASIS, 2006b), Postsecondary Electronic Standards Council (PESC, 2006b) for educational community, Chemical Industry Data Exchange for chemical industry, Petroleum Industry Data Exchange, and the World Wide Web consortium (W3C, 2007).

The goal and objective of the abovementioned standards consortiums are to define a common set of rules to be adopted by the same industry partners to ease and enhance the interoperability. This standardized B2B information interchange has proven to reduce the time and cost consumption of transaction processes. For example, Dell has reported the overall performance of its RosettaNet PIPs implementation which has resulted in a total of three hundred and twenty five thousands USD cost savings in December 2005 (RosettaNet, 2007b), and TMNet reported reducing total processing time from twenty-one days to three days (RosettaNet, 2007b). Unfortunately, some standards require proprietary framework such as the RosettaNet Implementation Framework (RNIF) and EDI in PESC which are time

consuming and involve high deployment cost. Therefore, the next generation standard should have the element of web compatibility to reduce the deployment of time and cost. This would then encourage the SMEs to more widely adopt the B2B standards. This feature is referred to as dynamic e-business standards (Chung, 2002; Allison & Terry, 2008). Web services standards will then emerge to meet this need. The neutral features of this platform will enable and enhance scalable integration and business agility.

Section 2.2.1 and 2.2.2 will describe RosettaNet and ebXML standards respectively. RosettaNet has been identified to be predominantly used in Asia IT and electronics industries while ebXML is the potential XML-based process standardization (Kilgarriff et al., 2004). PESC is a standard similar to RosettaNet but it is more widely used in the educational sector. It will be utilized in this research as a prototype of RosettaNet and will be introduced in Section 2.2.3.

2.2.1 RosettaNet Partner Interface Processes Standards

RosettaNet which was launched in June 1998 is an independent, self-funded, non-profit consortium of over four hundred major Information Technology players, Electronic Components and Semiconductor Manufacturing vendors worldwide (OASIS, 2009). One of the popular standards of RosettaNet is PIPs. PIPs specify a sequence of business processes or business process choreography as well as its business document format and content. This contributes to the efficiency in controlling the information flowing through the network between trading partners. RosettaNet has been widely implemented in some 670 companies in Malaysia (RosettaNet, 2008). Hence, as the background of standardized B2B information interchange, this study has reviewed RosettaNet PIPs 3A4 Request Purchase Order (PO) which is one of the most famous PIPs implemented in Malaysia.

The PIP Request PO enables a buyer to issue a PO and a provider to acknowledge whether the order is accepted, rejected or pending. The acknowledgement may also include

related information about delivery expectations. The process of issuing PO occurs after steps 1. Request Price and Availability (PIP 3A2), 2. Request Quote (PIP 3A1), and 3. Request Shopping Cart Transfer (PIP 3A3). This process may be followed by 1. Request PO Change (3A8), 2. Request PO Cancellation (PIP 3A9), 3. Query Order Status (PIP 3A5), and 4. Distribute Order Status (3A6). If the status of an order is pending, Notification of PO Acknowledgement Order (PIP 3A7) is used to notify the buyer (RosettaNet, 2007b). Figure 2.1 shows the business process model for PIP 3A4.

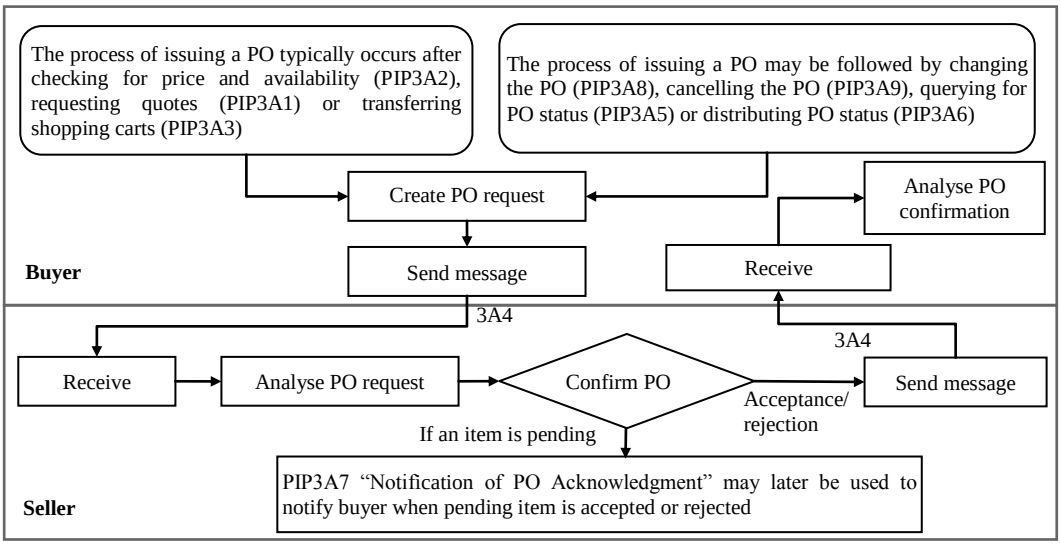

Figure 2.1: PIP 3A4 business process model (RosettaNet, 2007a).

PIP Request PO also provides a data dictionary (standard schema) to describe the content (data fields and its constraints) of the PIP. The content of the document is organized in a tree format. Each data field can be either a data by itself or a root to a group of data. The complete PIP 3A4 document file contents 551 data fields (RosettaNet, 2007a). Figure 2.2 shows the ontology of a data element (AccountDescription) in PIP 3A4.

```
<xs:element name="AccountDescription" type="tns:AccountDescriptionType">  </xs:element>
<xs:complexType name="AccountDescriptionType">
  <xs:annotation>    <xs:appinfo>
    <urss:Definition>The collection of business properties that describe a customer or supplier account.</urss:Definition>
    <urss:CreationDate>2005-03-24</urss:CreationDate>
    <urss:LastUpdatedDate>2006-06-07</urss:LastUpdatedDate>
    <urss:TypeVersion>01.04</urss:TypeVersion>
  </xs:appinfo>  </xs:annotation>
  <xs:sequence>
   <xs:element ref="dacc:AccountClassification" minOccurs="0"> </xs:element>
   <xs:element name="AccountName" type="xs:string" minOccurs="0"> <xs:annotation>
     <xs:appinfo> <urss:Definition>The name of a bank account.</urss:Definition>  </xs:appinfo>
   </xs:annotation>
  </xs:element>
  … …
  </xs:sequence>
  <xs:attribute name="schemaVersion" type="xs:token">
  </xs:attribute>
</xs:complexType>
```

Figure 2.2: AccountDescription data element in PIP 3A4.

2.2.2 Electronic Business XML (ebXML)

In the development of XML, ebXML is one of the most famous evolutions done by the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) and Organization for the Advancement of Structured Information Standards (OASIS, 2006b), for common electronic business document interchange. Its value lies in having a horizontal standard which is not specifically for any industry. ebXML is a set of specifications for business processes, business document content and format, registry to register and find business processes, company profiles and Trading Partner Agreement (TPA) (Mongiello, 2006; OASIS, 2006a; OASIS, 2006b). RosettaNet appears to be more comprehensive compared to ebXML (Table 2.1). Therefore, this research has chosen RosettaNet for further study. There are researchers who try to merge RosettaNet with ebXML (Dogac et al., 2002). In fact, efforts have been taken to merge other standards with RosettaNet, including ISO standards and World Wide Web Consortium (W3C) XML. This shows that RosettaNet standards are still being actively developed and improved, and its flexibilities are being incorporated into other standards.

Table 2.1: XML standards comparison (RosettaNet, 2001a).

XML Standards Components	RosettaNet	ebXML
Universal Specification Schema & Architecture	√	√
Supply Chain Business Processes	√	
Supply Chain Technical Dictionary Content	√	
Business Model Business Processes	√	
Universal Business Processes	√	√
Universal Technical Dictionary Structure	√	
Universal Business Dictionary Structure & Content	√	√
Universal Registry & Repository Structure		√
Universal Messaging Service	√	√

2.2.3 Postsecondary Electronic Standard Council (PESC)

PESC is a non-profit community-based umbrella association of the members from the higher education industry. PESC's primarily activity is to focus on the establishment of Web service-based data standard definition and adoption, which will help to ease the information interchange between higher education entities (PESC, 2006b). Among the standards which have been approved by the consortium are High School Transcripts, College Transcripts, and Academic Records. The standards use XML schema which suits Web services standards similar to RosettaNet standards. The available main standard of PESC is the CoreMain_v1.2.0 schema. An example of a group of data from CoreMain version 1.2.0 which shows its ontology similarity with RosettaNet is presented in Figure 2.3.

```
<xs:complexType name="AcademicProgramType">
    <xs:sequence>
        <xs:group ref="core:AcademicProgramCodeGroup" minOccurs="0"/>
        <xs:element name="AcademicProgramType" type="core:AcademicProgramTypeType" minOccurs="0"/>
        <xs:element name="AcademicProgramName" type="core:AcademicProgramNameType" minOccurs="0"/>
        <xs:element name="NoteMessage" type="core:NoteMessageType" minOccurs="0" maxOccurs="unbounded"/>
    </xs:sequence>
</xs:complexType>
… …
<xs:simpleType name="SchoolLevelType">
    <xs:annotation><xs:documentation>Indicates, Secondary, Postsecondary, etc.</xs:documentation></xs:annotation>
    <xs:restriction base="xs:string">
        <xs:enumeration value="Elementary"/>
        <xs:enumeration value="Secondary"/>
        <xs:enumeration value="Postsecondary"/>
    </xs:restriction>
</xs:simpleType>
```

Figure 2.3: PESC core main standards snippet (PESC, 2006b).

There are other higher education and research industry standards consortiums such as IMS Global Learning Consortium which specializes in on-line learning standards and integration (IMS, 2001). Another consortium is the Euro Current Research Information Systems with Common European Research Information Format standard in research information sharing. On the other hand, Shibboleth which provides Web Single Sign On middleware is used to simplify user authentication among educational entities. Both RosettaNet and PESC are vertical standards (Omelayenko & Fensel, 2001) which focus on the electronic and educational industries respectively. They use ontology-based XML schema too. With these similarities, PESC standards are adopted as a simplified version of RosettaNet PIP for implementation.

2.3 Related Works of Standardization in Communication Layers

2.3.1 Three Communication Layers in Standardization

Riemer and Totz (2003) define three layers of communication in Business-to-Consumer (B2C) Customer Relationship Management. These three layers namely Communication Subject, Communication Channel, and Communication Attributes are personalizable in order to substain customer loyalty. In this research, these communication layers are found to be able to describe personalization appropriately in B2B information interchange and are thus adopted to be the focus of this thesis.

In B2B information interchange communications, this thesis defines a Communication Subject as the message of interaction between two parties. In the case of B2B standards, ontology-based standard schema with XML-based messages is the Communication Subject in B2B information interchange. Today, XML has become the de facto data representation format in B2B information interchange (Mongiello, 2006) for integrating different Web-based systems. The medium used to interchange the messages is called a Communication

Channel, infrastructure or implementation framework. As B2B information interchange technology evolves, the EDI gateway is an indispensable Communication Channel in an electronic data interchange environment. However, in a standardized B2B information interchange environment, the Communication Channel has yet to be researched and further re-defined. The Communication Attributes define the configuration of an interaction process. Some of the parameters include communication intensity, frequency, time, and protocol (Riemer & Totz, 2003). In Section 2.3.2, three communication layers are facilitated to categorize related works in standardization and its inherent challenges.

2.3.2 Communication Channel

The major challenge in standardized Communication Channel is its inflexibility to be reused based on the user preferences. This issue has incurred a higher initial setup time and cost (RosettaNet, 2007a; RosettaNet, 2007b; Fuks & Wieczerzycki, 2006; Damodaran, 2004; Söderström, 2003; Burrows, 1999). Therefore, researchers are still working towards producing personalizable, reusable, and reliable solutions. Web services have emerged as the potential enabling technology (Lau, 2007; Elvis, 2007; Boncella, 2004; Chung, 2002). Some related works of Web service-based infrastructure are compared in the first place to highlight their strengths and weaknesses.

In Table 2.2, the *B2B standards applied* shows the B2B standards adopted in the publication. From the eight publications being compared, only two of them do not apply RosettaNet. Next, *Back-end integration* clarifies the involvement of private processes. Two of the compared works do not provide back-end integration. *Implementation described* is crucial to verify a proposed solution by a proof-of-concept prototype. *Evaluation* then measures the performance of a solution. Five types of work do not implement the solution and thus do not have any evaluation. *Technology used* highlights the utilization of Web services with other technologies, if any. All eight types of work utilize Web services. Since this research focuses on the SMEs, a criterion on *Suitability for SMEs* is included in the

comparison. This criterion will detect whether the work is providing the SMEs other alternative solutions. The first five criteria are adopted from Kontinurmi (2007), while the last criterion is required in this research. From the comparison, none of the researches has successfully implemented an end-to-end Web service-based solution in standardized B2B information interchange dedicated to the SMEs.

Table 2.2: Related works of Web service-based B2B standards.

Criteria / Publication	B2B standards applied	Beck-end integration	Implementa-tion described	Evaluation	Technology used	Suitability for SMEs
Cixing & Zhu (2009)	RosettaNet & other standards	Yes	No	No	Web services	No details
Yan et al. (2008)	RosettaNet	No	B2Bi gateway	No	Web services	Yes
Kotinurmi (2007)	RosettaNet	Yes	Yes	Qualitative comparison	RNIF, Web services, XSLT	No
Wang & Song (2006)	RosettaNet	Yes	No	No	RNIF, Web services	No details
Simmons (2004)	None	Yes	No	No	IBM Web services gateway	No details
Tikkala (2004)	RosettaNet	No	Yes	Quantita-tive experiment	Web services, XSLT	No details
Masud (2003)	RosettaNet	Yes	No	No	RosettNet-based Web services	No details
Willaert (2001)	ebXML, SOAP	Yes	No	No	ebXML, SOAP	No details

In order to further understand standardized B2B information interchange infrastructure, RosettaNet implementation models namely the Direct Model and Application Service Provider (ASP) Model are studied. The Direct Model requires the company to fully adopt the standardization and develop the implementation infrastructure, called RosettaNet gateway or RNIF which costs more than five thousand USD. The ASP Model involves a third party service/solution provider or an intermediary server which develops and hosts the implementation infrastructure or gateway. Each company is charged on the basis of the

messages interchanged. In this case, the company does not need to develop RNIF. Solution providers such as B-Global, B2B Commerce, CrimsonLogic, Dagang Net Technologies, Kompakar EBiz, (RosettaNet, 2008) offer services to the ASP hub sharable between several trading partners (RosettaNet, 2007b). For example, a company may be charged thirty USD per transaction. Besides the hardware and gateway cost, the implementation cost of defining standard schema for both models depends on the number and complexity of the messages transmitted (Stanley, 2008).

Other challenges in this layer include the capability of solution to integrate with legacy system (Kotinurmi, 2007; DINET, 2010), interconnection with other networks (DINET, 2010), enhance performance (Juriz et al., 2007), platform independence (Sanders et al., 2008; Juric et al., 2007; Isaacson, 2007; Gialelis et al., 2006), distribution (Sanders et al., 2008), and heterogeneity (Sanders et al., 2008; Gialelis et al., 2006). Some of these features are not necessarily to be embraced by the SMEs. For example, SMEs who do not have complicated legacy system would prefer to adopt a light-weight B2B information interchange gateway (Yan et al., 2008) and thus, the feature of integration with legacy system is not crucial. However, this research still takes into account the abovementioned features into consideration because the coverage will enable extension of this research to future works which might have different focus or fall under varied industrial applications.

2.3.3 Communication Subject

Since the early 1990s, research and development in standardization have been widely spread among researchers and industries after the importance and benefits of B2B standards have been recognized. The emerging organizations and research centres are carrying out R&D mostly in terms of standardizing business documents contents and its format (RosettaNet, 2001a; RosettaNet, 2007a; OASIS, 1993; OASIS, 2006b; Janner et al., 2006; OAGI, 1994; PESC, 2006a, 2006b). However, with the emergence of different standards in the markets, there exists a new problem in integrating the standards. As such, R&D in

semantics came in to resolve the problems of B2B standards integration (Kotinurmi et al., 2009; JBoss Redhat, 2009; Kotinurmi, 2007; Seng & Lin, 2007; Kilgarriff et al., 2004; Omelayenko & Fensel, 2001). These standards conformed to the same XML schema format and structure yet are different semantically. It is a challenge to integrate these standards. The heart to this problem is the desire of standard organizations to dominate the market and thus, withholding the information sharing and integration among the organizations. This issue will not be studied further in this research since it involves many standardization parties, which have different business strategies.

Besides the above mentioned issue, Damodaran (2004) reports that a large message with redundant content is another challenge in integrating the standards. This is due to the rigid standard schema which cannot be changed during deployment time. The issue can be further broken down into two sub problems which include inflexibility in simplifying lengthy standard schema (JBoss RedHat, 2009; RosettaNet, 2007b), and inflexibility to reduce redundant content. The first problem implies that the large standard schema is able to cater to an industry's need in sending the desired content but unable to increase the performance. In the second problem, a receiver cannot customize the standard schema once it has been deployed and results in a lack of flexibility. In both cases, if the user requests to change or simplify a standard schema, all trading partners that use the same standard must be informed and carry out the appropriate modification together accordingly. This will not only require modification to the Communication Subject but also the Communication Channel. This is a tedious and complicated process which not only involves one trading partner but also many others.

The exchange of transactions through PIPs has now greatly reduced the efficiency and data quality of the transactions interchanged (Damodaran, 2004; Surfcontrol, 2007). A study has been conducted for the comparison of two RosettaNet PIPs in a procurement transaction comprising the Request Purchase Order and Query Order Status. It was found that among

the first hundred data fields in both the PIPs, there were 90% of redundant contents while the entire PIPs were computed to contain 56% of redundant fields (Ting & Khoo, 2007b). Seng and Lin (2007) also reports a repeat rate of 92% between PIP Request Purchase Order (PIP 3A4) and PIP Request Purchase Order Change (3A8).

Another issue aroused in standardized B2B information interchange in this layer is the document transformation process from different proprietary format into XML-based document. Existing B2B standards are still imperfect in integrating disparate data sources in an organization (Xu & Zhao, 2010; JBoss Redhat, 2009). Most of the B2B standards focus on public process automation. Intermediary layer to bridge public and private process becomes crucial to fill the gap and automate the process (Seng & Lin, 2007; Kilgarriff et al., 2004). This issue also leads to the data accuracy problem. It is a challenge to validate the interchanged documents' data structure, attributes, and its types. Another level of validation which enables private process automation is the database table dependency that is lacking in B2B standard schema. Besides, Silver (2009) proposes XML-based dynamic web content delivery in which the content is scalable at real time. This means that the receiver controls the content and its currency. Although this is not a new technology in web content delivery, nevertheless, it is still a challenge in today's standardized B2B information interchange due to its rigid standard schema. All these challenges have yet to be resolved in personalizable standardized B2B information interchange.

2.3.4 Communication Attributes

Communication Attributes in standardized B2B information interchange can be researched from different perspectives. One of the R&D efforts in this field is business processes modelling. Besides the standardized business processes defined by standards organizations (OASIS, 2006a; RosettaNet, 2002; RosettaNet, 2007a), other researchers such as Schönberger (2006) and Shi and Chen (2007) utilize existing standards (RosettaNet PIPs) to model the collaboration of business processes. Re-modelling the business processes is an

important task in order to streamline and automate B2B information interchange process and thus, create more pervasive solutions in standardization.

In the case of the RosettaNet standards, the trading partners have to connect to the Internet 24 hours a day and 7 days a week (24x7), in order to ensure a smooth transmission of the transactions (RosettaNet, 2007a; Fuks & Wieczerzycki, 2006). In other words, the receiver does not have the flexibility to decide *WHEN* to receive the document from the senders (Ting & Khoo, 2007a; Ting & Khoo, 2007b). The 24x7 Internet connection requirement increases the implementation cost for the SMEs. In this context, the model of document interchange is a Conventional Push Model whereby the sender has the control to send a message at anytime (Ting & Khoo, 2007b; Bhide et al., 2002). It is described in the B2C e-commerce Conventional Push Model whereby the manufacturers 'push' the ready products to the customers. The B2C customers have no choice but to customize the products. In contrast to the Conventional Push Model, the Pull-only Model however allows the customer to customize the product in order to personalize it to fulfil the desired needs (Aspen, 2006).

In terms of information interchanged, sending an advertisement to a user's mobile phone is an example of a B2C Push Model. From a different perspective, a user subscribing to an online newspaper like the popular Really Simple Syndication (RSS) is an example of a B2C Pull-only Model. The push-pull mechanism denoted here is not complicated and applicable in any entity. On the contrary, the push-pull mechanism in the B2B information interchange environment for different market segments is diverse and complicated. This is due to the complex business processes (public and private processes), non-repudiation security, and the confidential information interchanged. The push-pull models in standardized B2B information interchange do not seem to have been fully researched so far, as evidenced in (Yan et al., 2008; Aspen, 2006; Aksoy & Leung, 2004; Acharya et al., 1997; Vrieze et al., 2003; Willaert, 2001; Bonifati, 2002; Mohan, 2001; Bhide, 2002). Although

the mail server can be an alternative to allow the receiver to 'pull' messages whenever they are connected to the Internet, the processes for downloading the message and transforming it to a database format is still being done manually. The push-pull models will be discussed in Section 3.2.1.

Table 2.3 compares the existing works on push-pull models in terms of five criteria. The first criterion is on push-pull models covered in the work called *Push-pull models used*. All the works utilize both push and pull models except that of Bonifati et al. (2002) which only utilizes the push model. Three of them also utilize the hybrid model (Aksoy & Leung, 2004; Vrieze et al., 2003; Bhide, 2002). The *Application area* denotes the utilization of the push-pull models in different applications such as B2B information interchange (Yan et al., 2008), B2C (Aspen, 2006), data broadcasting (Aksoy & Leung, 2004; Acharya et al., 1997), and information request/respond system (Vrieze et al., 2003; Mohan, 2001). The term *Technology/methodology used* further describes how the models are conceptualized or realized. Both the terms *Implementation described* and *Evaluation* justify the works through a proof-of-concept prototype and its evaluated performance.

Table 2.3: Comparing previous works.

Criteria / **Publication**	***Push-pull models***	***Application area***	***Implementation***	***Evaluation***	***Technology/ methodology***
Yan et al. (2008)	Push, Pull	B2B information interchange	Yes	No	Web services, RNIF
Aspen (2006)	Push, Pull	B2C purchasing process	Yes	Descriptive	N/A
Aksoy and Leung (2004)	Push, Pull, Hybrid	Data broadcasting	Yes	Quantitative experiment	Algorithms
Vrieze et al. (2003)	Push, Pull, Hybrid	Personalization of a system – user interface and functions	No	No	Rule-based, function-based
Willaert (2001)	Push, Pull	Document transformation process	No	No	ebXML, SOAP, XSLT
Bonifati et al. (2002)	Push	Web services registration to XML repositories	No	No	Web services, XQuery, SOAP, Rule-based
Mohan (2001)	Push, Pull	Database caching for dynamic web site	Yes	Descriptive	Caching technologies
Bhide (2002)	Push, Pull, Hybrid	Dynamic Web data dissemination	Yes	Quantitative	Algorithm

Based on the comparison in Table 2.3, push-pull modelling is not a new approach in many fields such as dynamic data broadcasting (Aksoy & Leung, 2004; Acharya et al., 1997; Bhide, 2002), B2C purchasing process (Aspen, 2006), personalization of a system's interface and functions (Vrieze et al., 2003), document transformation process (Willaert, 2001), and database caching for dynamic web site (Mohan, 2001). However, push-pull modelling is new in standardized B2B information interchange (Yan et al., 2008) and not much research work have been done on the modelling especially the hybrid model, the actual implementation and evaluation. This research has taken the initiative to study different models in standardized B2B information interchange with a proof-of-concept prototype and evaluation. Other challenges in this layer are the support of Internet connectivity types (Yan et al., 2008), communication protocol and connection method (DINET, 2010), and communication nature, either synchronous or asynchronous (RosettaNet, 2001b). Some of these characteristics are not supported in standardized B2B information interchange due to its restrictive standard procedures and thus become obstacles among the SMEs.

2.3.5 Standardization versus Personalizability

Eva Söderström in her survey found that striking a balance between standardization and personalization is one of the challenges in B2B standards (Söderström, 2003). Here, information interchange is controlled and limited by TPA and standards guidelines. On the other hand, personalizable standards allow modification of standards during implementation to suit an industry's specific need. For example, optional fields in standards can be decided by the receiver anytime. However, extremely flexible standards will cause inter-operability problems between the trading partners who adopt the same set of standards. On the other hand, online alteration is thereby not allowed or prohibited to avoid integration problems. Consequently, personalized B2B information interchange contradicts with standardized B2B information interchange. In short, standardization is the inverse of the level of

personalizability (Söderström, 2003). All the inflexibilities in all the three communication layers as discussed before have created low personalizability in the standardized B2B information interchange. In other words, standardization inhibits personalizability and thus, personalizability is the main challenge of all. Nevertheless, having personalized the communication layers, the outcomes are not determined. Therefore, it is required to define the parameters of personalization in order to determine the personalizability and ensure its interoperability.

2.4 Related Works of Personalization in Three Communication Layers

Many researchers and practitioners are interested in personalization due to the fact that personalization meets customer needs and sustains loyalty (Riemer & Totz, 2001; Riemer & Totz, 2003). Riemer and Totz (2001; 2003) studied the economic motivation through personalization in B2C e-commerce and found that customer loyalty had improved the relationship with the customers due to the increase in switching costs, which refer to the direct/indirect costs incurred in order to switch from one product/service to another, such as time and cost spent to search for a new supplier, and a negotiation process. A personalized product/service makes the product/service unique in the market. Customers find it difficult to search and compare new product/service and thus, are reluctant to switch to other product/service. Furthermore, personalized services are customized based on the customer's service request history and profile. For customers switching to a new company, they may run the risk of losing their profiles. All these factors have contributed to a high switching cost and thus increase customer retention rate. Likewise, inflexible B2B information interchange communication layers bring about issues discussed before which inhibit smaller businesses from adopting standards. The research challenge here is to find out whether higher personalizability is able to conceptually motivate the higher pervasiveness/adoption due to the economic benefits provided by personalizability. This would counter the myth