E-government Data Interoperability Framework in Hong Kong

Thomas Y. Lee and Patrick K. Yee and David W. Cheung
Center for E-Commerce Infrastructure Development
Department of Computer Science
The University of Hong Kong
{ytlee,kcyee,dcheung}@cecid.hku.hk

Since its first release in 1998, the Digital 21 Strategy has been serving the blueprint for Hong Kong to develop the information and communications technology infrastructure. With government leadership as an important component, the Strategy has put forward the development of an e-government that can realize one-stop delivery of public electronic services. In 2003, Hong Kong Government has launched the Interoperability Framework (IF) as an e-government initiative to facilitate implementation of cross-department joined-up government services. IF comprises two parts: (1) definition of a set of recommended technical specifications as a single point of reference for departments and contractors to implement joined-up projects, (2) a framework for formulating and managing XML message standards for G2G and G2B data exchange. This paper discusses about the second part of the IF, mainly on the XML Schema and Design Guide, and its real-life applications.

Keywords: Hong Kong e-Government Interoperability Framework, XML Schema Design and Management Guide

I. INTRODUCTION

A. HKSARG Interoperability Framework for E-Government

First launched by the Hong Kong Special Administrative Region Government (HKSARG) in 1998, the Digital 21 Strategy [1] is a strategy paper setting out the blueprint for development of information and communications technology (ICT) in Hong Kong. In 2001 Digital 21 Strategy, one focus area was “to ensure that HKSARG leads by example in driving e-government, providing e-option for more services, and actively pursuing e-procurement and outsourcing.” An important e-government initiative is to provide client-centric joined-up government services to the public. In HKSARG, public services are offered by different HKSARG Bureaux and Departments (B/Ds) in a distributed fashion. To provide one-stop comprehensive services, seamless flow of information across individual B/Ds is a prerequisite. In 2003, the HKSARG Information Technology Services Department, now known as Office of Government Chief Information Officer (OGCIO) [3], established the Interoperability Framework (IF) for E-Government [2] to meet this objective.

B. Three Dimensions of Interoperability

IF addresses three dimensions of interoperability, namely technical interoperability, data interoperability, and process interoperability. The XML Schema Design and Management Guide [4] describes these terms as follows:

- Technical interoperability: agreement on what communication protocol and message format to be used when one party sends information to another; e.g. the purchase order shall be encoded in XML, as defined by a specific XML schema, and XML Encryption and XML Signature shall be applied on certain content components, and the XML message shall be sent via HTTP.
- Data interoperability: agreement on what information has to be transmitted from one party to another, and the definition and representation of such information; e.g., the “delivery date” has to be specified on a purchase order, and the definition of “delivery date” is the date on which goods shall be received by the buyer, and the representation of “delivery date” adopts the ISO 8601 standard [16].
- Process interoperability: agreement on how the business activities of the concerned parties affect each other; e.g., when the seller receives a purchase order from the buyer, the seller should accept or reject this order within a specified time period.

To address these dimensions of interoperability, IF maintains two sets of documents. The first set of documents, HKSARG Interoperability Framework [5], recommends a collection of technical specifications for implementation for application in different interoperability areas. For example, in the Application Integration Domain, SOAP v1.1, WSDL v1.1, and UDDI v2 are recommended for simple functional integration in an open environment while ebXML Message Service v2 is recommended for reliable message exchange between application systems in an open environment for business document oriented collaboration. This means when a new e-government project is recognized to fit a particular interoperability area, the B/D responsible for the project and their IT contractors are required to adopt the recommended technical standards in system implementation.

The other set of documents are collectively called XML Schema Design and Management Guide [4] (Schema Guide). They are developed to address data and process interoperability. Section II outlines the contents of this Schema Guide. We describe the Design Guide part and the Management Guide part respectively in Section III and Section IV. Section V lists some e-government projects that have applied this framework in production environment. Finally, Section VI concludes this paper.
II. XML SCHEMA DESIGN AND MANAGEMENT GUIDE

The Center for E-Commerce Infrastructure Development [8] of The University of Hong Kong was commissioned to develop the XML Schema Design and Management Guide to help B/Ds and IT contractors define and adopt e-government data standards. This Guide provides the following to address the data and business interoperability of government-to-government (G2G) and government-to-business (G2B) joined-up services:

- a methodology for business analysts to specify the definitions and representations of information in a consistent and structured way as reusable information models,
- an approach for programmers to convert the information models of the data elements into W3C XML Schema Definition (XSD) [7] code,
- the guidelines for concertedly aligning and

![Diagram of Schema design process flow]

Figure 1. Schema design process flow.
standardizing the definitions and representations of
data elements that have potential for reuse, and
- the guidelines for project teams to adopt suitable
  Common Schemas and also to contribute reusable
data elements for creating new data standards.

The Guide consists for four parts:
- PART I: Overview states the objectives of the Guide
  and outlines its contents, and describes the data
  interoperability problems and strategy in e-
  government development. This part also covers the
  data interoperability measures and the guiding
  principles.
- PART II: XML Schema Design Guide provides a
  systematic schema design methodology for business
  analysts to model business process and information
  requirements. It also provides the rules for
  programmers to convert the information models into
  XSD code.
- PART III: XML Schema Management Guide serves
  as a handbook for various parties to align data
  elements from different projects to create data
  standards among B/Ds. It facilitates the development
  and management of reusable XML Schemas. It also
  helps project teams understand the process of data
  alignment, and their role in contributing reusable
  data elements for concerted alignment.
- PART IV: Appendices provide supplementary
  information to help the readers understand this
  Guide. Most importantly, there is a case study on
  using the Guide to define data standards for
  application for import and export licences for
  pharmaceutical products.

III. XML SCHEMA DESIGN GUIDE

The XML Schema Design Guide (Design Guide)
dокументs a comprehensive XML schema design
methodology. There are four core sections: (1) Schema
Design Process, (2) Business Process Modelling
Methodology, (3) Business Information Modelling
Methodology, and (4) XML Schema Definition
Development. The Design Guide provides analysts and
developers with an end-to-end and seamlessly integrated
schema design methodology. It covers modelling techniques
from business process modelling, data modelling to XSD
coding.

A. Schema Design Process

This schema design process guides a project team to
develop the XML data interface, i.e., XSD, for a government
system to exchange data with other government or business
systems. The process involves two types of schemas: Project
Schemas and Common Schemas. A Project Schemas only
needs to satisfy the specific requirements of a particular
project while Common Schemas are reusable across projects.
Common Schemas are centrally standardized in the
government and are shared publicly in the Central Registry
[11]. Another difference between Project Schemas and
Common Schemas is as follows. A Project Schema is usually
document-oriented, like purchase order, because it is ready
for use for data exchange. A Common Schema is data-
element-oriented, such as address, person name, while it is
used as a building block to construct Project Schemas.

The process flow is shown in Fig. 1 and is summarized as
follows. First of all, the project team collect and analyze the
user requirements to see if there is a suitable industry
standard. The team starts to design a custom schema only
when no industry standard is suitable. Then, the business
analyst models the business process in which a system
exchanges data with other systems, and identifies the
business documents involved in the exchange. For each
identified document, the business analyst builds its
information model, i.e., the document structure as well as
data constraints. The data models are converted into XSD
code of the Project Schema by programmers or by software.
If the project team anticipates some data elements in the
Project Schema are generic enough for reuse by other
projects and B/Ds in future, they can propose these data
elements for central standardization as Common Schemas.

B. Business Process Modelling Methodology

The business process modelling methodology is derived
A business process is modelled as a collection of business
collaborations. A business collaboration is a choreography
of business transactions. A business transaction is an
abstraction of a single document exchange, which can be a
one-way or two-way document flow between two parties
involved in the collaboration. When the flows of documents
between different parties are identified, these documents will
be modelled in the next stage.

The Design Guide provides a set of modelling
worksheets for business analysts to fill in the requirements of
business collaborations and business transactions. For
example, the worksheet in Fig. 2 documents an order entry
business collaboration, which involves a buyer as the
requesting role and a seller and the responding role. Two
business transactions, namely Request Quote and Create
Order, are identified in the collaboration. The Request Quote
transaction involves an exchange of two documents Quote
Request and Price Quote while the Create Order transaction
is an exchange of Purchase Order and Order Confirmation
documents. These documents will be used for business
information modelling. The process model can be specified
as a UML activity diagram.
C. Business Information Modelling Methodology

The business information modelling (BIM) process gathers the requirements for the business documents identified in the business process modelling. These requirements are usually in form of paper and electronic copies of those documents that are currently exchanged manually or in old electronic fashion. The modelling methodology simplifies and extends the Core Components Technical Specification (CCTS) [10] and allows business analysts to specify documents and data elements as the following information models:

- **Business Document** models an electronic document as a unit for business information exchange; a root Aggregate Business Information Entity is used to provide the representation of the document,
- **Aggregate Business Information Entity** (ABIE) models an object class and aggregates Basic and Association Business Information Entities as the properties,
- **Association Business Information Entity** (ASBIE) models a complex property in an object class,
- **Basic Business Information Entity** (BBIE) models a singular property in an object class,
- **Core Component Type** models a basic business data type as a building block for the above models.

BIM provides a set of Core Component Types (CCTs) as the basic business data types for building other the information models, e.g., Amount, Date Time, Code, Quantity, etc.

The relationships of these information models can be represented by a UML class diagram like Fig. 3. BIM also provides a modelling spreadsheet as shown in Fig. 4 for business analysts to enter the specifications of information models. The spreadsheet contains user-friendly macros that guide users to enter various attributes required by different models on specific dialog boxes. It also allows users to import and select existing models for reuse in defining new models. Moreover, the spreadsheet models can be converted into XSD through macro programs without coding.

D. XML Schema Definition Development

The Design Guide provides a comprehensive specification to standardize the conversion from an information model into XSD code. For example, each BBIE or ABIE is converted into an XSD complex type. A standard XSD library for all CCTs is provided for import into developed XSDs. Since the conversion mechanisms from information models to XSDs have been standardized in the Guide, the XSD generation can be automated by software.
IV. XML SCHEMA MANAGEMENT GUIDE

The XML Schema Management Guide (Management Guide) defines the framework (e.g., policy, organization structure, etc.) to manage the Project Schemas and Common Schemas defined using the Design Guide. The Management Guide describes:

- the considerations for the managing Project Schemas as well as their controlled vocabularies, e.g., code lists, namespace, etc.,
- the management process and organization structure for concerted alignment of reusable data elements proposed by B/Ds for standardization of Common Schemas, and
- the guidelines for building a Project Registry and the Central Registry for sharing Project Schemas and Common Schemas respectively.

Each Common Schema is associated with one of the following three maturity levels.

- **Level 0 – Agreed in principle.** Among the B/Ds that believe the Common Schema might be applicable to their business, the majority anticipate that they require further investigation and analysis before adopting the Common Schema.
- **Level 1 – Recommended for reuse.** Among the B/Ds that believe the Common Schema might be applicable to their business, the majority anticipate that they are ready to adopt the Common Schema in most of their future projects. Common Schemas are reviewed by all B/Ds regularly. In a review, when more B/Ds vote on a Common Schema for Level 1 than for Level 0, this Schema will be promoted to Level 1.
- **Level 2 – Matured for reuse.** Among the B/Ds that believe the data element might be applicable to their business, the majority anticipate that they are ready to adopt the Common Schemas in future projects. In addition, the Common Schema has already been used in some projects and has remained stable for a certain period. When at least 5 B/Ds are using a Level-1 Common Schema which has not been changed for 6 months, this Schema will be automatically promoted to Level 2.

V. APPLICATION IN E-GOVERNMENT PROJECTS

Many e-government projects have applied the Guide to create their Project Schemas. The following are some examples:

- **Automation of dangerous goods manifests submission** [14]. The project facilitates shipping companies to automate submission of dangerous goods manifests to the Marine Department through the Internet using ebXML Message Service.
- **XML weather information publishing** [12]. The Hong Kong Observatory publishes weather information in XML for businesses, such as media and transport companies, and government departments to automate processing of weather information.
- **Works Project Information Standard** [13]. The Works Branch Development Bureau defines a large-scale XML data standard for various building and construction domains to realize electronic data exchange among various stakeholders in construction projects.
The Macao Government Interoperability Framework [15]. The Guide is being translated into Chinese and Common Schemas for the Macao environment are under construction.

VI. CONCLUSIONS

The HKSARG XML data interoperability framework provides a comprehensive XML schema design methodology and the necessary schema management infrastructure to facilitate e-government data standardization. A library of reusable Common Schemas has been developed, and shared in a data standards registry for developing Project Schemas for new government services. While these schemas are designed for government projects, they are also published publicly to facilitate adoption by the business sector. Moreover, the XML Schema Design Guide is a generic integrated methodology, which covers all necessary steps in e-business modelling, from requirement analysis, to business process and information modelling, to XSD coding. Business enterprises and other governments can easily adapt it to establish their own data interoperability infrastructures.

REFERENCES