Regis University

ePublications at Regis University

Fall 2007

Xml Interfaces: a Growing Need for Standardization

Elizabeth A. Jackson
Regis University

Follow this and additional works at: https://epublications.regis.edu/theses

Part of the Computer Sciences Commons

Recommended Citation

Regis University
College for Professional Studies Graduate Programs
Final Project/Thesis

Disclaimer

XML Interfaces: A Growing Need for Standardization

Elizabeth A. Jackson

Regis University

School for Professional Studies

Master of Science in Computer Information Technology

Regis University

School for Professional Studies Graduate Programs

MSCIT Program

Graduate Programs Final Project/Thesis

<u>Certification of Authorship of Professional Project Work</u>

Print Student's Name   **Elizabeth A. Jackson**

Telephone   **970-635-9539**   Email   **beth_jackson@lpbroadband.net**

Date of Submission **June 15, 2007**   Degree Program   **MSCIT**

Title of Submission   **XML Interfaces: A Growing Need for Standardization**

Advisor/Faculty Name   **Douglas Hart**

Certification of Authorship:

I hereby certify that I am the author of this document and that any assistance I received in its preparation is fully acknowledged and disclosed in the document. I have also cited all sources from which I obtained data, ideas or words that are copied directly or paraphrased in the document. Sources are properly credited according to accepted standards for professional publications. I also certify that this paper was prepared by me for the purpose of partial fulfillment of requirements for the Master of Science in Computer Information Technology Degree Program.

_____   _____
            Student Signature   Date

6/15/07

Regis University

School for Professional Studies Graduate Programs

MSCIT Program

Graduate Programs Final Project/Thesis

<u>Authorization to Publish Student Work</u>

I, _____Elizabeth A. Jackson_____, the undersigned student, in the
Master of Science in Computer Information Technology Degree Program hereby authorize Regis
University to publish through a Regis University owned and maintained web server, the
document described below ("Work"). I acknowledge and understand that the Work will be freely
available to all users of the World Wide Web under the condition that it can only be used for
legitimate, non-commercial academic research and study. I understand that this restriction on use
will be contained in a header note on the Regis University web site but will not be otherwise
policed or enforced. I understand and acknowledge that under the Family Educational Rights and
Privacy Act I have no obligation to release the Work to any party for any purpose. I am
authorizing the release of the Work as a voluntary act without any coercion or restraint. On
behalf of myself, my heirs, personal representatives and beneficiaries, I do hereby release Regis
University, its officers, employees and agents from any claims, causes, causes of action, law
suits, claims for injury, defamation, or other damage to me or my family arising out of or
resulting from good faith compliance with the provisions of this authorization. This authorization
shall be valid and in force until rescinded in writing.

Print Title of Document(s) to be published: ____XML Interfaces: A Growing Need for____

Standardization _____

_____   _____6/15/07_____
                    Student Signature   *Date*

- -

Check if applicable:

_____ The work does <u>not</u> contain private or proprietary information of the associated parties
and their attached permission is <u>not</u> required.

_____ The Work contains private or proprietary information of the associated parties and their
permission to e-publish this Work is not forthcoming.

__XX__ The Work contains private or proprietary information of the following parties and their
attached permission is required as well: ___ADP Screening and Selection Services_____
                                        Name of Organization and/or Authorized Personnel
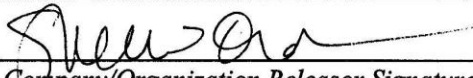
Regis University

School for Professional Studies Graduate Programs

MSCIT Program

Graduate Programs Final Project/Thesis

<u>Releasor Authorization to Publish Student Work WWW</u>

I, _____Shelli Oldham_____the undersigned, _____Product Manager_____

Print Name of Company/Organization Representative   *Representative's Title*

on behalf of ___ADP Screening and Selection Services_____ ("Releasor") do hereby authorize

             Company/Organization Name

Regis University to publish through a Regis University owned and maintained web server, the

document described below ("Work") and acknowledges that the Work contains personal or

proprietary information of the Releasor. Releasor further acknowledges and understands that the

Work will be freely available to all users of the World Wide Web under the condition that it can

only be used for legitimate, non-commercial academic research and study but that this restriction

on use will be contained in a header note on the Regis University web site but will not otherwise

be policed or enforced. This authorization shall be valid and in force until rescinded in writing.

Print Student Name: ___Elizabeth A. Jackson_____

Title(s) of document(s) to be published: ___XML Interfaces: A Growing Need for_____

_Standardization_____

BY: _____ DATE: ___6/13/07_____

       Company/Organization Releasor Signature

Note: It is the student's responsibility to obtain the necessary release(s) prior to submitting the Final Project for publication. Please print your name and list all applicable documents.

Regis University

School for Professional Studies Graduate Programs

MSCIT Program

Graduate Programs Final Project/Thesis

<u>Advisor/Professional Project Faculty Approval Form</u>

Student's Name: __Elizabeth A. Jackson_____ Program _____MSCIT_____
                              PLEASE PRINT

Professional Project Title: __XML Interfaces: A Growing Need for Standardization__
                              PLEASE PRINT

Advisor Name _____Douglas Hart_____
                PLEASE PRINT

Project Faculty Name _____Dee Bilo_____
                    PLEASE PRINT

Advisor/Faculty Declaration:

I have advised this student through the Professional Project Process and approve of the final document as acceptable to be submitted as fulfillment of partial completion of requirements for the MSCIT Degree Program.

Project Advisor Approval:

_____   _____06/20/2007_____
Original Signature   *Date*

Degree Chair Approval if:

The student has received project approval from Faculty and has followed due process in the completion of the project and subsequent documentation.

_____   _____
Original Degree Chair/Designee Signature   *Date*

Regis University

School for Professional Studies Graduate Programs

MSCIT Program

Graduate Programs Final Project/Thesis

<u>Disclaimer</u>

Use of the materials available in the Regis University Thesis Collection ("Collection") is limited and restricted to those users who agree to comply with the following terms of use. Regis University reserves the right to deny access to the Collection to any person who violates these terms of use or who seeks to or does alter, avoid or supersede the functional conditions, restrictions and limitations of the Collection.

The site may be used only for lawful purposes. The user is solely responsible for knowing and adhering to any and all applicable laws, rules, and regulations relating or pertaining to use of the Collection.

All content in this Collection is owned by and subject to the exclusive control of Regis University and the authors of the materials. It is available only for research purposes and may not be used in violation of copyright laws or for unlawful purposes. The materials may not be downloaded in whole or in part without permission of the copyright holder or as otherwise authorized in the "fair use" standards of the U.S. copyright laws and regulations.

Note: The following Disclaimer is for your information. You do not need to submit it with your project. It will be scanned into all student works that are available on the WWW.

Abstract

The incentive for this research came from the need to solve a fundamental problem many companies face when transferring data across multi-platform applications: how to deploy XML interfaces that integrate easily with other systems. The goal of this project was to identify a solution for Company ABC when integrating XML interfaces with customers, partners, and vendors. The project researches the need and development of data standards that can solve one of the key problems with XML – interoperability. Through the use of a common language, companies can utilize data formats and schemas developed by data standards organizations operating in their vertical industry. Implementing a standard XML interface will help Company ABC reduce the expense of customizing all of their XML interfaces.

Acknowledgement

I would like to dedicate a page in this publication to those who inspired and encouraged me to reach my dream of successfully completing my degree. Thank you to my husband who supported me throughout the process and encouraged me to hang in there and complete my degree. Many thanks as well for his help in reading through all my papers however technical they were. Thank you to my parents for their encouragement in all my pursuits. Thanks also to my advisor, Dee Bilo, for her help in getting me on the right track.

Table of Contents

List of Figures

Chapter One: Introduction

1.1 Introduction

In the competitive Business-to-Business (B2B) environment, it is critical that new business opportunities and strategies are integrated into the business application. The challenge is to achieve integration in the diverse computing environment that is universal in business today (Ennser, 2000).   Extensible Markup Language (XML) has become recognized as the clear technology choice for enabling inter-company data exchanges (HR-XML Consortium, 2003). The support for creating, exchanging, and processing of XML within programming languages, databases, enterprise software, and application servers has become universal (HR-XML Consortium, 2003).

XML provides the format for identifying information that applications need and XML schema languages describe a shared vocabulary of names used to markup XML documents (Ennser, 2000). The challenge for businesses is to create a XML schema language that can be used to define, describe, and catalogue XML vocabularies for classes of XML documents (Malhotra, 1999)

1.2 Problem Statement

The company documented in this project uses XML in its customer, partner, and vendor integrations and faces the challenge of creating a standard approach to developing, deploying, and maintaining its XML interfaces. To preserve the confidentiality of the organization, the name of the company was withheld and referred to as "Company ABC," and the specific software programs used to perform the XML processes were not identified. Company ABC's current use of XML integration is very costly, time consuming, and resource intensive. A common problem that occurs when IT systems communicate is data integration (Fohn, 2007).

1.2.1 Existing Integration Processes

Company ABC utilizes XML interfaces because it improves accuracy of data, reduces turnaround times, and reduces the cost to fulfill reports. The company currently uses XML interfaces in the following processes: 1) customer/partner interfaces, 2) vendor business processing, 3) USIS vendor interface, and 4) instant product interfaces. These interfaces account for approximately 40% of products fulfilled by automated means.

1.2.1.1 Customer/Partner Interfaces

This interface utilizes an automated order process. The products to be ordered have already been pre-determined by the customer/partner. These orders generally tend to have large amounts of data, specifically several applicants or employees that need to have background screens ordered. In this scenario, the customer sends in an order via XML using an XML schema. The system tries to validate the order based upon the data submitted. If any part of the order has missing data, then that piece of the order is skipped. All information is sent back and returned in pure XML.

In this process, the XML interface must be customized for each customer and partner requiring IT development time to create and maintain these integrations. The integration between customer and partner systems can be quite complex, requiring more development time to get the customer and partner integrated into Company ABC's system. This development time puts severe strain on IT and Operations resources because of the custom interface that must be developed between the customer/partner and Company ABC.

1.2.1.2 Vendor Business Processing

In this process, the interface uses a servlet to feed the XML to and from vendors. Requests for products are made directly from Company ABC to the vendor via XML and

responses are received back from the vendor via XML. The XML is then parsed and sent to

Company ABC's database.

This process is highly dependent on support from IT and Operations due to the significant

number of errors that occur during the processing of background reports. These errors occur

during the transmission of XML within the business process. Every time the process is changed

or a new business process is set up, the base code and internal application must be modified.

1.2.1.3 USIS Vendor Interface

This XML interface is between the USIS vendor and Company ABC's database. USIS

supplies court record reports to Company ABC and is its largest supplier. The USIS vendor

interface is based upon HR-XML and is a real-time interface used to send candidate data out for

search and return results directly back to Company ABC. In this process, the XML interface

works well for Company ABC and requires little maintenance from IT.

1.2.1.4 Instant Product Interfaces

In this interface, requests are manually requested by the customer. The customer has

specified the products to be ordered via Company ABC's application interface. After all

products have been ordered by the customer, the first step in the process requires that a request

be made to the vendor for an instant product. The request is handled in one of two scenarios. In

the first scenario, the request is submitted via a template that is specific to the vendor and

product. In this request, all XML has been written for the specific template based upon the XML

standards of the vendor. In the second scenario, the data from the order is placed into a data

class object which goes out to a web services library and builds all of the XML behind the

scenes.

The next step in the process is a response from the vendor. The response comes back from the vendor in XML format and then is sent to a Java parser. The parser turns the XML into Java objects, which in turn goes to the database returning readable PDF reports to the customer.

The instant product interface has some products using the HR-XML standard but not all. For those instant products that do not use a standard interface, the XML must be customized to fit the vendor's needs.

In all XML interfaces not utilizing the standard, the result is a very labor intensive request to add or modify products that customers, partners or vendors may wish to add within Company ABC's application. This is due to the limited number of resources available to support the XML interfaces. At this time, only one resource in Company ABC is able to manipulate the data in XML based upon vendors needs. This creates a domino effect because when XML is changed, changes must be tested and released into production, which may impact other vendors and cause their interface to work improperly.

1.2.2 Integration Considerations

XML is a common methodology used by background screening companies to easily interface with its customers, partners, and vendors. Company ABC should consider implementing a standard XML interface to reduce the risk and expense of integrating data between its system and the systems of its customers, partners, and vendors.

1.2.3 The Need for Standard XML Interfaces

The need for implementing a standard XML interface between Company ABC and its customers, partners, and venders is important to the structure of Company ABC's integration processes. Standardization allows Company ABC to free itself from the need to build custom interfaces for new products, customers, and partners.

Using a standard approach to all of Company ABC's XML interfaces would allow

Company ABC to improve upon its current customer, partner, and vendor relationships by

simplifying how integration is accomplished. The time to market is shorter, which benefits

Company ABC and its customers, partners, and vendors.

1.2.4 Research Goal

The goal of this project was to demonstrate the need for Company ABC to standardize

XML interfaces with its customers, partners, and vendors. Research shows that when XML

interfaces are standardized, companies benefit by reducing IT development time and

maintenance costs, reducing processing errors, and improving data quality.

1.3 Business Case

The need for providing a data interchange standard format for businesses exchanging

background check data is critical. Research shows that discovery, analysis, custom interface

development, and ongoing maintenance account for 60 to 70 percent of many HR information

system solutions (HR-XML Consortium, 2003).

1.4 Business Goals

The business goals for this project were lowering IT development time, developing easier

code to maintain, speeding up the time to market, easing the complexity of the integration

process, and reducing the expense of customizing all XML interfaces.

1.5 Scope of the Project

The scope of this project was to research and analyze data collected on XML interfaces to

determine the best solution for integrating XML interfaces in Company ABC's implementation

process. The proposed solution will address the multiple and unique XML interfaces Company

ABC has with their customers, partners, and venders in an effort to reduce costs and resource

development time. By creating one common XML standard interface, Company ABC can create

a uniform XML format that can be used by all customer, partner, and vendor systems.

1.6 Outline

Chapter 2 will review the research gathered on the topic of XML. In the first part of the

chapter the technical definition, evolution, and current status of XML will be presented. The

second part provides an in-depth report of the research conducted that is relevant to the problem

of XML interface integration. The third part summarizes all the research that is known and

unknown about XML integration. The fourth part defines the contribution this project will make

to the field.

Chapter 3 will address the research methodology used in the project. An analysis of the

research methodology and procedures used to conduct the project is addressed in the first part.

The second part reviews the presentation of results and deliverables. The third part discusses the

resource requirements necessary to conduct the research. The fourth part analyzes the outcomes

of the project. The chapter concludes with a summary of the research methodology.

Chapter 4 discusses the history of the project. The first part reviews how the project

began and management of the project. The second part provides insight into the milestones of

the project. The third part evaluates if the project successfully met the project goals. The fourth

part discusses what went right and wrong during the project and if any project variables impacted

the project. The fifth part provides a detailed analysis of the findings. The chapter concludes

with a summary of the results.

Chapter 5 focuses on the lessons learned during the project. This chapter is devoted to

discussing what was learned from the project experience, things that could have been done

differently, if the project met its initial expectations, and the next stage of evolution. The chapter

finishes up by providing conclusions and recommendations to Company ABC about possible solutions to the problem and a summary of the project.

1.7 Definition of Terms

B2B – Business-to-business: The marketing strategy used by businesses who sell services or goods to other businesses.

cXML – Commerce XML: A standard developed for companies that utilize business-to-business electronic commerce.

DTD – Document Type Definition: Defines the elements of an XML document.

ebXML – Electronic Business XML: Standards sponsored by OASIS and UN/CEFACT to provided formal XML concepts and methodologies for better e-business solutions.

E-Commerce – Electronic Commerce: The distribution, buying, selling, servicing, and marketing of services or products over the Internet or other computer networks.

HR-XML – Human Resources XML: Standard XML developed by the HR-XML Consortium.

HTML – Hypertext Markup Language: Markup language used for the creation of Web pages.

OAGI – Open Applications Group, Inc: An open standards group that builds process-based XML standards for B2B integration.

OASIS – Organization for the Advancement of Structured Information Standards: A not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards.

PMI – Project Management Institute: Organization focused on providing knowledge and skills to project management professionals, as well as set the standards for project management.

ROI – Return on Investment: The amount of profit or cost savings an organization realizes from an investment.

SGML – Standard Generalized Markup Language: A meta-language that defines markup languages for documents.

Unicode Standard: Provides a unique number for every character, no matter what the platform, program or language.

UN/CEFACT – United National Centre for Trade Facilitation and Electronic Business: An inclusive organization who encourages close collaboration between governments and private business to secure the interoperability for the exchange of information among the public and private sector.

USIS – Leading supplier of comprehensive employee drug and background screening services.

W3C – World Wide Web Consortium: A consortium who develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential (W3C).

WWW – World Wide Web: System of hypertext documents that are interlinked and accessed via the Internet.

XML – Extensible Markup Language: A markup language for documents containing structured information.

XML DOM – XML Document Object Model: Defines a standard way to access and manipulate XML documents.

1.8 Chapter Summary

The purpose of this project was to provide Company ABC with a solution to the problem of integrating its XML interfaces with customers, partners, and vendors. The research conducted for this project collected and analyzed data to support the need for implementing a common data standard across all of its product offerings. The need for a common data standard is a result from one of the key problem areas on interoperability, the data integration problem (Fohn, 2007).

Chapter Two: Review of Research

2.1 Introduction

XML has provided businesses with a tool that opens up data to be shared among customers, partners, and vendors without sharing or integrating critical business systems. When exchanging business data, XML acts as the uniform standard and as semantic middleware, through which heterogeneous applications can communicate with one another over uniform interfaces and in a language which everyone involved can understand (Herm, n.d.).

Determining the best mechanism for data interchange can be risky and expensive when companies must negotiate and agree upon a common interchange mechanism. This chapter analyzes the research conducted to determine the best approach for integrating XML interfaces and provides Company ABC with a possible solution for interacting with customers, partners, and venders without having to establish, engineer, and implement several different interchange mechanisms.

2.1 Overview of Literature and Research

2.1.1 Definition of XML

XML is a universal descriptive markup language for use in the delivery of information across the Internet whose primary purpose is to add structure and intelligence to data exchange (Ennser, 2000). It is an open, fee-free standard that provides the framework (rules and tools) for creating individual markup languages. Using these rules, programmers can design, name, and organize XML elements, attributes, and entities. In defining XML, programmers apply self-describing markup (or tags) around each element of data, making XML data easy for computers to access, index, search, transfer, and secure.

XML documents are comprised of markup (elements, attributes, and angled brackets) and content. Figure 1 shows an example of an XML document (Palermo, 2002).
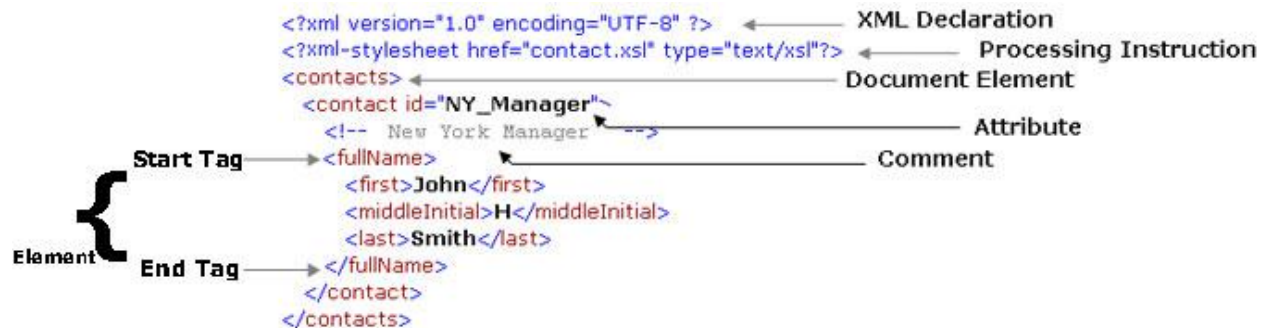
Figure 1: XML Document Sample

The XML Declaration identifies the document as XML and the version of XML to which it was authored. The declaration is not required but is generally supplied in order for the document to be understood clearly. The next line defines the processing instructions. Processing instructions provide information to an application that are not textually part of the XML document but required by the XML processor to pass them to an application. Elements identify the nature of the content they surround. Attributes are name-value pairs that occur inside start-tags after the element name. Comments are not part of the textual content of an XML document and the XML processor is not required to pass them along to an application. Comments may be placed anywhere between markup in the document (Walsh, 1998).

The names, allowable hierarchy, and meanings of the elements and attributes have been left open to allow developers to define customizable schema, which provide a syntactic foundation for the creation of a custom, XML-based markup language. The general syntax of markup languages is rigid. Documents must adhere to the general rules of XML, ensuring that all XML-aware software can at least read (parse) and understand the relative arrangement of information within them.

The schema purely supplements the syntax rules with a set of constraints. The purpose of the schema is to allow validation of the document structure and every specific element of the document, which does not violate any of the constraints defined by the model and considered valid according to the schema (Walsh, 1999). The schema is built with the following most central top-level constructs:

- a (global) *element declaration* – associates an element name with a type

- a *complex type definition* – defines the requirements for attributes, sub-elements, and character data in elements of that type

  - *attribute declarations* – describes which attributes may or must appear

  - *element references* – describes which sub-elements may or must appear, how many, and in which order

- a *simple type definition* – defines a set of strings used as attribute values or character data

Two types or two elements cannot be defined with the same name, but an element declaration and a type definition may use the same name (Moller, 2003).

There are two types of schemas: 1) document type definition (DTD) and 2) XML Schema. DTD is the oldest form of schemas and inherited from SGML. Employing DTD's, designers of the XML document may select or create a DTD that specifies the meaning of the tags. The DTD defines all of the elements used in a set of documents, as well as the relation of the tags to each other in a tree format. Every element is given an element type name (tag name) and a set of attributes. Every attribute is given a name and value. These elements and attribute names are strings or a restricted set of characters.

Since elements are not predefined, programmers must create vocabularies or use a collection of XML elements. Some open source XML vocabularies used by B2B's are Human

Resource XML (HR-XML), Electronic Business XML (ebXML), and Commerce XML (cXML).

These local and global vocabularies are defined using the W3C's XML Schema Definition

Language.

XML schemas are the successors of DTD's. XML Schemas also define the data elements

of particular transactions and specify the options and constraints governing the use of those

elements but go a step farther. They are designed to validate the results of collected information

based on specific data types. This makes XML schemas much more powerful than DTD's.

XML Schema is the first separate schema language for XML and became a W3C

Recommendation in May 2001 (Sperberg-McQueen, 2000).

To manipulate an XML document, you need an XML parser. An XML parser takes an

XML document as input and makes its content and structure available to other applications

(Karre, nd). The parser is a piece of software used to validate the XML document against a

DTD/schema to verify that the document is valid. The parser loads the document into a

computer's memory. Once the document is loaded, its data can be manipulated using a parser.

One type of parser is the XML Document Object Model (XML DOM) that defines the standard

way for accessing and manipulating XML documents. Figure 2 shows how DOM represents an

XML document as a tree-structure (a node tree), with the elements, attributes, and text defined as

nodes (W3School, nd).
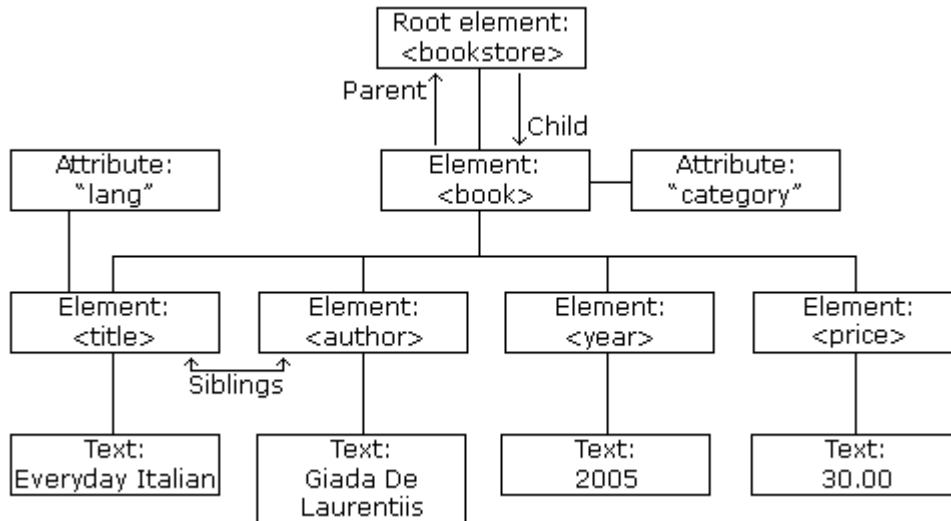
Figure 2: DOM Tree Node

According to the specification given in the World Wide Web Consortium

recommendation, every XML document needs to be well-formed and valid (Karre, nd). This

means that the document must contain a proper schema and the document obeys the constraints

of those declarations. To be considered valid, the document must meet the following

requirements as defined by Palermo (2002):

- All tags must be closed

- Tags are case sensitive

- The XML document must have a single root element

- Elements must be nested properly without overlap

- No element may have two attributes with the same name

- Attribute values must be enclosed in quotes (using either ' or ")

XML can be used in several ways: 1) store data and pass information over the Internet;

2) data syndication; 3) podcasting; and 4) configuring instructions for the behavior of

applications. It is a universal language that allows B2B organizations to transfer data across diverse platforms and continues to grow with the changing environment of the Internet.

2.1.2 History and Goals of XML

XML started its development in 1996 due to the success of HTML and the need for a simplified SGML standard that would be easier to implement on the Web. This need for a new document standard that could handle the problem of different document formats of various systems prompted the W3C to host the standards process and form the XML Working Group to develop XML. The group was chaired by Jon Bosak of Sun Microsystems with the active participation of the XML Special Interest Group. XML 1.0 became a W3C Standard in February 1998 (W3C Recommendation, 2004).

The XML language was developed similar to SGML and HTML but has three significant differences: 1) user is not locked into predetermined elements or tags; 2) it is easier to create document types; and 3) use of tags to describe both documents and data, making it easy to parse and exchange data between different applications on various platforms (Ennser, 2000). These differences create a language that can control the structure and content of the data being passed between applications.

The W3C defines the ten key design goals for XML as:

1. XML shall be usable in a straightforward manner over the Internet

2. XML shall support a wide variety of applications

3. XML shall be compatible with SGML

4. It shall be easy to write programs which process XML documents

5. The number of optional features in XML is to be kept to the absolute minimum, ideally zero

6. XML documents should be legible and reasonably clear

7. The XML design should be prepared quickly

8. The design of XML shall be formal and concise

9. XML documents shall be easy to create

10. Terseness in XML markup is of minimal importance (W3C Recommendation, 2004)

2.1.3 Evolution of XML

The core XML specifications have changed very little since their inception in 1998 because they are a collection of tried and true ideas. The XML 1.0 Technical Recommendation remains stable and unchanging with only one corrective release in 2000. Due to XML 1.0's use of strict rules regarding the construction of XML names and the frequently changing Unicode standard, W3C drafted version XML 1.1 which was released in October 2006.

XML 1.1 loosened the rules around name creation to keep up with the new versions of Unicode, which provides better support for legacy line-end codes and arbitrary Unicode characters. The popularity of XML and related initiatives has surpassed the expectations of the original XML Working Group, making it the most popular and successful development of the W3C (SYS-CON Media, Inc. (2007).

2.1.4 Current State of XML

XML plays a very strong role in its current form by providing an organization with a safe choice for data exchange, application integration, documentation, business messaging, and other areas. It is a standard that offers tremendous power to organizations by giving document creators the ability to represent all kinds of data with tags that associate values to each element making it

easy to work, group related information together, and leverage the Unicode standard by supporting all computing platforms and human languages.

2.1.5 Future of XML

From a programmer's point of view, the future of XML is moving toward schemas that are more compliant due to the validation issues and less reliance on DTD's. Creating more standard XML attributes will make integration much easier.

2.2 Literature and Research Specific to Project

2.2.1 Practical Uses of XML in Business to Business

In the B2B environment, there is a need for the exchange of data quickly and securely between applications and systems. The use of the Internet or extranet provides an invaluable means of improving B2B partnerships and inter-organizational relationships that can cut the cost of operations and production, improve business processes, and enhance e-business partnerships. The ability to integrate between businesses provides significant value to companies in the areas of delivering products to market faster, ability to serve customers better, better understanding of buying patterns, reducing supply chain costs, and quicker response to business change.

Within the B2B framework lies a comprehensive set of tools that allows for easy configuration, plug-in, customization, and execution of business transaction between business partners (Ennser, 2000). Using standard XML technology, businesses can help streamline the data exchange format due to its simplicity, extensibility, openness, and lower implementation cost.

Consider a business that would like to transfer data across the network and several applications. The first difficulty would be how the data is represented. Business A may define a purchase order differently from Business B. Secondly, each business may be using different

applications to run their business. Business A is using Oracle while business B is using SAP.

Thirdly, the businesses may be using different protocols to run their networks. Each business

may be standardized on different platforms where Business A is using NT and Business B is

using Sun Solaris. Data formatting is a large problem for B2B integration that XML can help

solve as shown in Figure 3.
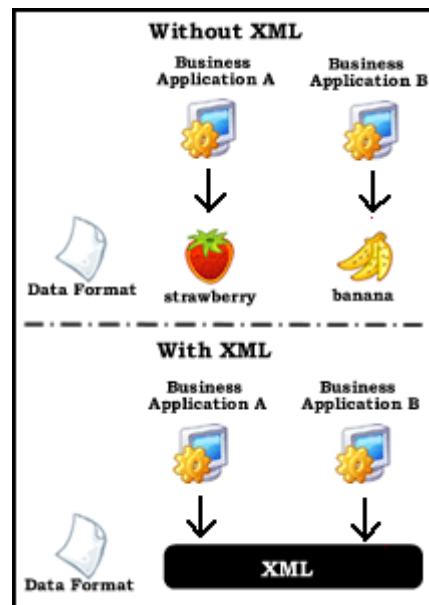
Figure 3: Data Formatting

    XML provides an efficient way to transfer data across various applications and platforms

in B2B situations. Each business in the transfer must agree upon what the data means and how it

is presented via the XML markup. Once the businesses have agreed upon the meaning of the

information and the XML markup, a parser is created that can easily move the data. Business

rules may even be applied to manipulate the data before transfer or after receiving.

    2.2.2 Data Integration Problems

    XML technology is an efficient way to transfer data but there is one defining problem:

interoperability. XML interoperability is the ability of an application to make effective use of an

XML document that it receives. XML gives programmers the freedom to create vocabularies to meet their needs, but this freedom also brings about a lack of a unified vocabulary making data integration a problem area for interoperability.

Fohn (2007) notes "the data integration problem results from the lack of a common language through which IT systems can communicate" (para 2). Common languages are necessary to offer both consistent and uniform syntax and semantics. By creating a common language, data standards can be built to solve the needs of organizations.

2.2.3 Data Standards Organizations

The need for data standards has lead to the creation of data standards organizations, which focus on a particular vertical industry. Data standards organizations help to establish best practices that streamline industry-wide integration. These organizations develop and publish open, freely available data exchange standards based on XML.

The research for this project was focused upon Company ABC whose business is to provide background checking products to its customers and partners, as well as work directly with vendors who provide the products. One such organization, HR-XML Consortium, has been formed and is well represented by many of Company ABC's competitors. The HR-XML Consortium is a global, vendor-neutral, non-profit consortium dedicated to enabling e-commerce and inter-company exchange of human resources data (HR-XML Consortium, 2003).

The HR-XML Consortium is one of the largest and best supported XML industry standards groups with over 100 member companies represented in 22 different companies. The Consortium's development of the HR-XML Background Screening specification defines a data format for requests to background check service providers and a format for search results back from the providers. The specification includes a complete set of flexible XML schemas defining

the data format of the background check data exchanges, as well as supporting documentation that explains and illustrates the schema components (HR-XML Consortium, 2003).

2.2.2 Standardized XML Schemas

The common thread throughout the integration process is the ability for organizations to utilize a standard XML technology, specifically the XML Schema. HR-XML's schema was developed for human resources and its data definitions were developed using the W3C's XML Schema Definition Language.

For the purposes of Company ABC, using HR-XML Background Requests and Background Reports schemas would provide the needed support of requests to providers of employment screening services and the return of search results from those providers. Company ABC would benefit tremendously from schemas developed specifically for background screening companies by lowering development time, code maintenance, and administrative costs.

2.3 Summary of What is Known and Unknown about the Project

Jim Green, CTO of webMethods states "Exchanging business information is a critical requirement for automating business processes within and between companies." (W3C, 2001) The capacity of Company ABC to use XML interfaces in its customer, partner, and vendor integrations is a significant piece of background-related business applications. The need for a standard approach to Company ABC's XML interfaces is necessary to reduce the dependency on its IT and Operations resources. Integrating a standard, such as HR-XML's Background Screening specification, will help control the unbounded growth of the many interfaces Company ABC is currently developing and maintaining.

The Open Applications Group, Inc, (OAGI) has developed a mathematical model for sizing the growth and costs of unbounded interfaces $n(n-1)$ where "n" represents the number of

components to be integrated (HRcertify-org, nd). For example, 20 components equal 380

interfaces. By developing one standard interface, the interface grows at a flat rate of n x 2

resulting in 20 components equal 40 interfaces. The HR-XML standards now change the many-

to-many integration equation into a many-to-one.

Adopters of the HR-XML Consortium's Background Checking Specification have

reported cost savings attributable to error reduction and improved data quality (HR-XML

Consortium, 2003). This would clearly be something Company ABC could realize if they

decided to implement the HR-XML standards in all of its XML interfaces.

2.4 Contributions Project will make to the Field

The research from this project shows that a standard XML interface is critical to

Company ABC's integration with its customers, partners, and vendors. Integrations are a key

aspect of the background screening industry and Company ABC needs to be able to interact with

its customers, partners, and vendors in a way that removes the risk and expense of having to

develop unique XML interfaces.

Chapter Three: Methodology

3.1 Research Methodology

The research methodology used for this project was formed from the following two methods: Waterfall Methodology and the Big6^{TM}. By combining these two methodologies, the author could benefit from a defined set of activities and deliverables with the flexibility to move between the different stages. The four consecutive stages of activity included *Analysis, Research, Outcome, and Reflectio*n and the percentage of work necessary to complete each stage is outlined below.
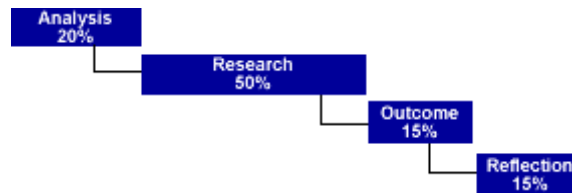
Figure 4: Waterfall Stages

Analysis: This stage defined the IT problem and identified the information needed to solve the problem. It also addressed the purpose of the project and the goals to be achieved, as well as the top strategies for seeking the best possible sources.

Research: Qualitative research was conducted in this stage in the form of interviews, observation of company data exchange processes, business document reviews, articles, and journal reviews. This stage looked at locating sources and finding relevant information to the problem.

Outcomes: In this stage, it was necessary to synthesize the information gathered to determine the outcome for the project. The information was organized based upon its relevance to the topic and then documented for presentation to management.

Reflection: This stage provided an evaluation and reflection of the project in terms of lessons learned. It provided the reader with knowledge about the subject matter and how companies could use this information to improve upon their XML interface processes.

3.2 Life-Cycle Model Used

The project resulted only in recommendations based upon the research conducted, and no life-cycle methodology was proposed for Company ABC to use. This project was a business analysis of how the XML interfaces currently in place could be better implemented with one common standard data exchange.

3.3 Specific Procedures

By incorporating the use of qualitative research methods (Myers, 2007), the author was able to look at the project from the user' point of view. As stated by Myers (2007), "qualitative research methods are designed to help researchers understand people and the social and cultural contexts within which they live" (para. 9). Understanding how the implementation process worked from the users' perspective enabled the author to draw conclusions about how to improve upon XML integration processes.

The author utilized the following three qualitative research methods for this project:

Action Research: Conduct research that would improve upon the XML implementation process by speeding up time to market, lowing costs, and easing maintenance issues.

Case Study: Investigate the business process of Company ABC, in this case the integration of XML interfaces.

Grounded Theory: Develop a theoretical account of the XML integration process Company ABC currently uses in an effort to solve a problem.

3.4 Review of Results/Deliverables

The deliverables from this project were a business analysis report documenting the process review of the XML product interfaces and presentation of findings to IT and managers. The documentation recommended Company ABC utilize the HR-XML standards across all XML interfaces, which would improve the company's ability to work with other customers, partners, and vendors, simplify the integration of systems, and decrease the development time required to create and maintain the integrations.

3.5 Resources

The project had limited resources even though the outcomes would be beneficial to Company ABC. The project's primary resource was the author who served as the researcher. Additional resources from Company ABC were used in the form of interviews with the product manager, IT director, and programmers.

To preserve the confidentiality of the organization, the name of the company was withheld and referred to as "Company ABC," and the specific software programs used to perform the XML processes were not identified.

3.6 Outcomes

The research gathered from this project determined that the best solution for Company ABC to improve upon its XML interfaces was to implement the HR-XML standard across all product offerings. This would be accomplished by using the Background Checking specifications developed by the HR-XML Consortium. Company ABC has decided not to implement this solution across all of its XML interfaces. This is due to limited resources, time requirements, and financial benefits. Company ABC feels it cannot write the new XML code necessary across all products and remove the old code while still maintaining its current

activities. The implementation of such a major project would severely impact the IT department, which is already under strain to complete current projects.

Since Company ABC has decided not to implement the changes recommended from this project, the problem will continue to severely hamper the implementation process it currently uses forcing IT to devote time to building and maintaining custom XML interfaces. This also puts a strain on some of the other divisions throughout the organization who are unable to support the complicated implementation process.

Persuading Company ABC to change its implementation process poses a challenge to the author. Getting more managers to buy-in to the concept of a standard implementation process would be helpful but it would require considerable persuasion from the IT Director to the General Manager that this change would substantially benefit the company financially.

3.7 Chapter Summary

This chapter outlined the research methodologies used to define the activities and deliverables of the project. The project was guided under the auspices of qualitative research methods since its intent was to research how Company ABC could improve upon a business process and provide recommendations for a solution.

The research resulted in determining a possible solution for Company ABC to standardize all of its XML interfaces by using HR-XML. Due to limitations on resources, Company ABC has determined it will not implement the recommendation.

Chapter Four: Project History

4.1 Brief History

The basis for this project emanated from discussions with the IT Director in December 2006. Numerous topic ideas were uncovered during these conversations, but XML stood out as a viable opportunity for Company ABC to uncover areas where the company could improve upon its integration processes.

Further discussions were conducted with the Product Manager to determine a problem Company ABC was having with its XML integration. This laid the groundwork for starting a project that would analyze the XML integration problem and the author would serve in the researcher role.

4.2 Project Management

This project was managed under the guidelines established by the Project Management Institute (PMI). The specific guidelines used included Project Integration Management, Project Scope Management, and Project Time Management. The use of these guidelines helped to facilitate the goal of conducting research necessary to ensure success of the project.

4.3 Milestones

The first major milestone was to determine the topic for the project. XML is a major player in the ordering process of Company ABC and would provide a good business analysis of how to improve upon the XML interface integrations.

The second major milestone was to define the problem statement and purpose for the project. This enabled the author to establish the goals and scope of the project.

The third major milestone was to clearly understand XML (technical definition, evolution, and current and future state), conduct interviews to see the XML interface process in action, and read journal and article reviews to analyze the topic.

The fourth major milestone required synthesizing all the knowledge gained from the research and putting it into a written document. This required parsing through all of the research and outlining only the information necessary for the paper.

4.4 Changes to the Project Plan

The project plan was changed slightly to establish more recommendations. As the project progressed into the outcome stage, it was necessary to conduct more interviews with associates at Company ABC to discuss how XML interfaces could be better implemented. This pushed the reflection stage farther out to devote more time into looking at ways Company ABC could improve upon its XML integration processes.

4.5 Did the Project Meet the Goals

Overall the project was a success based upon the goal outlined previously. The research provided Company ABC with a solution to its XML interface integration process problem. Implementing the standard HR-XML across all XML interfaces would provide many benefits to the development, deployment, and maintenance of customer, partner, and vendor integrations.

4.6 What Went Right or Wrong

In general, the project went fairly smoothly. There was enough data available to research the topic quite thoroughly. Discussions with Company ABC resources provided a great deal of insight into the XML interfaces and generated ideas on how the company could improve upon its XML integration processes. Analyzing the information gained from the interviews proved to be tricky. As the author talked to associates at Company ABC, it was sometimes difficult to assess

their thoughts correctly and get caught up in taking notes instead of listening to what was being said. The author should have tape recorded the conversations.

As far as what went wrong, it was easy to go down the wrong path and find other ideas to discuss due to the amount of information available on the topic of XML. This project was focused on improving the XML interfaces Company ABC uses with its customers, partners, and vendors. Much of the information available was technical. It was easy to get caught up in how XML works rather than how businesses should apply it to their needs.

4.7 Project Variables and Impact on the Project

The variables of this project included time and resources. The author had limited time to conduct research to establish recommendations for an analysis of the XML interface processes. It was necessary to conduct interviews of Company ABC associates and view the current processes during normal business hours, which impacted the time necessary to fully understand the business process in-depth. The IT department was under pressure to complete projects so interviews were conducted as appropriate to respect the associates' workload.

4.8 Findings/Analysis of Results

In reviewing the results of the project, the major point proved that utilizing a standard XML across all products would be beneficial to Company ABC. These benefits include lowering costs, reducing maintenance of XML interfaces, faster time to market, and attracting more customers, partners, and vendors because Company ABC uses a defined industry standard XML interface.

Utilizing a standard XML interface will improve relationships with its customers, partners, and vendors, simplify the integration of systems, decrease the need for building and

maintaining unique interfaces, and reduce costs affiliated with Company ABC's associates' time required to fulfill reports.

Consortiums have become an effective resource for businesses that use XML. The HR-XML Consortium is specific to the background check industry and has developed XML vocabularies for human resources. The Consortium can provide Company ABC with XML schemas, documentation, and examples based on use cases for all products that Company ABC is currently using.

Many of Company ABC's competitors have published their use of HR-XML standards on the HR-XML Consortium Web site. This gives Company ABC's competitors a significant advantage.

4.9 Summary of Results

Results are summarized below:

- The project was a success in determining a viable solution for Company ABC to implement a standard XML interface with all of its customers, partners and vendors using HR-XML developed by the HR-XML Consortium.

- The author developed a business analysis report documenting the research gathered to establish recommendations.

- Due to limited resources, the standard HR-XML interface will not be implemented across all product offerings Company ABC currently supports.

Chapter Five: Lessons Learned and Project Evaluations

5.1 Lessons Learned From the Project Experience

Company ABC does not see the results of this project as a priority for the organization. Prioritization is given to projects that impact ROI and Company ABC feels this project would not be a good utilization of resources compared to other projects currently in process. The research gathered from this project has highlighted a need for standardization and may be used when future discussions of IT projects are considered in yearly reviews.

Limited IT resources have impacted the outcome of this project. The standard HR-XML will be used only in cases where utilizing the standard HR-XML is easy to implement and does not affect IT resources. All other XML interfaces will continue to use the current process.

It is difficult to get buy-in from management even when the benefits outweigh the costs required to maintain the XML interfaces in their current state. This can be a frustrating aspect of research-related projects. Money and resources definitely play a pivotal role in how organizations manage their development processes.

The research project provided a look at how different people within organizations perceive an application. Some managers felt XML was an awful application, having too many problems and always needing to be fixed. This may well be due to the fact that the only time they interacted with XML was when there was a problem. As a result, the attitude about XML was negative. Other managers found XML to be a wonderful tool that worked effortlessly within the organizations systems. Of course there were minor problems, but nothing that could not be fixed.

It is also interesting to note that the actual programmers who work with XML all day told the author they felt XML is a very powerful tool, which is easy to use and maintain. Even with

some of the problems programmers experience such as invalid characters, UNIX to Windows conversion, illegal characters, and space handling, XML is the backbone to the success of Company ABC's applications.

5.2 What I Would Have Done Differently

With such a tight project timeline, it did not allow for the opportunity to consult with other organizations to examine how they have implemented XML interfaces and see their successes and failures. The information gathered from this opportunity would have provided the author with some understanding of how XML is used in different industries and possibly generate some ideas for other ways to improve upon implementation processes and reducing risks.

5.3 Initial Project Expectations

The project was expected to provide Company ABC with a solution to its XML interface integration. This expectation has been met by providing useful information to the organization on how to improve its process of integrating XML interfaces with customers, partners, and vendors through the use of a standard XML interface.

5.4 Where Would the Project Progress From Here

The project would progress on to implementing the standard HR-XML across all product offering XML interfaces. The IT Department would develop a new project, which would involve removing all the old XML code and implementing the standard HR-XML across all interfaces. At this time, limited resources and monetary benefits are not substantial enough to warrant the continuation of this project.

5.5 Conclusions/Recommendations

From the research collected, the conclusions/recommendations for Company ABC are as follows:

- Implement a standard data interchange across all XML interfaces.

- Develop a business model that defines the data needs for all XML interfaces.

- Develop a partnership with the HR-XML Consortium for access to cross-industry standards including the schemas, documents, and use cases necessary for background check companies.

- Utilize the HR-XML standards approved and developed by the HR-XML Consortium.

- Utilize the HR-XML standards so Company ABC may publish their name on the Consortium Web site, as many of their competitors are currently listed.

5.6 Summary

To summarize, this project has demonstrated that using a standard XML interface will improve Company ABC's process of integrating with its customers, partners, and vendors. By integrating a standard data exchange mechanism developed by an industry specific consortium across all product offering XML interfaces, Company ABC would 1) be more attractive to customers, partners, and vendors if it uses the HR-XML standard, 2) lower the cost of implementing XML interfaces, 3) develop XML code that is easier and less costly to maintain, and 4) speed up time to market by simplifying the integration of their system with other customer, partner, and vendor systems.

References

Eisenberg, M. and Berkowitz, B. (2006) *What is the Big6™?* Retrieved February 2, 2007 from http://www.big6.com/showarticle.php?id-415.

Ennser, L., Leo, P., Meszaros, T. and Valade, E. (2000) *The XML Files: Using XML for Business-to-Business and Business-to-Consumer Applications.* IBM Redbooks.

Fohn, S. (2007) *Whitepaper: HR-XML 3.0.* Retrieved March 23, 2007 from http://www.hr-xml.org/blog/?page_id=119

Herm, M. (n.d.) *XML - An opportunity for small and medium-sized enterprises*. Retrieved April 12, 2007 from http://www.softwareag.com/xml/library/herm.htm.

HR-XML Consortium (2003) *Background Checking FAQs*. Retrieved March 19, 2007 from http://66.179.192.180/resources/BackgroundCheckFAQs2003-April-23.pdf.

HRcertify-org (nd) *Frequently Asked Questions*. Retrieved March 19, 2007 from http://www.hrcertify.org/subchannels/faqs_01.php

Karre, S. and Elbaum S. (n.d.) *An Empirical Assessment of XML Parsers*. Retrieved April 9, 2007 from http://www.cse.unl.edu/~elbaum/papers/workshops/webe_final02.pdf

Malhotra, A. and Maloney, M. (1999) XML Schema Requirements. Retrieved May 23, 2007 from http://www.w3.org/TR/NOTE-xml-schema-req

Moller, A. and Schwartzbach, M. I. (2003) *Technologies for the future Web*. Retrieved April 9, 2007 from http://www.brics.dk/~amoeller/XML/index.html.

Myers, Michael D. (2007) *Qualitative Research in Information Systems*. Retrieved January 10, 2007 from http://www.qual.auckland.ac.nz/

Palermo, M., Singh, D., Mohr, S., Siegers, P., and Knowles, C. (2002) *Professional ASP.NET 1.0 XML with C#.* Peer Information Inc. publishing. Retrieved March 29, 2007 from http://www.perfectxml.com/XMLIntro.asp

Project Management Institute (2000) *A Guide to the Project Management Body of Knowledge*. Newton Square, PA: Project Management Institute, Inc.

Sperberg-McQueen, C.M. and Thompson, H. (2000) *XML Schema*. Retrieved March 29, 2007 from http://www.w3.org/XML/Schema

SYS-CON Media, Inc. (2007) *Tracking the Evolution of XML*. Retrieved February 9, 2007 from the XML-Journal Web site at http://xml.sys-con.com/read/40505_p.html

W3C (2001) *Testimonials for XML Schema Recommendation*. Retrieved April 23, 2007 from
http://www.w3.org/2001/05/xml-schema-testimonial.html

W3C Recommendation (2004) *Extensible Markup Language (XML) 1.1*. Retrieved February 5,
2007 from http://www.w3.org/TR/2004/REC-xml11-20040204/

W3Schools (nd) *XML Dom Tutorial*, Retrieved March 29, 2007 from
http://www.w3schools.com/dom/default.asp

Walsh, N. (1999) Understanding XML Schemas. Published by XML.com. Retrieved February
5, 2007 from http://www.xml.com/pub/a/1999/07/schemas/index.html?page=1.

Walsh, N. (1998) *A Technical Introduction to XML*. Published by XML.com. Retrieved
February 5, 2007 from http://www.xml.com/lpt/a/316.