Reference Model for Open Architecture for XML Authoring and Localization Version 1.0

Committee Draft 02

26 October 2009

Specification URIs:
This Version:
http://docs.oasis-open.org/oaxal/V1.0/cd02/oaxal-v1.0-cd02.html (Authoritative)
http://docs.oasis-open.org/oaxal/V1.0/cd02/oaxal-v1.0-cd02.pdf
Previous Version:
http://docs.oasis-open.org/oaxal/V1.0/cd01/oaxal-v1.0-cd01.html (Authoritative)
http://docs.oasis-open.org/oaxal/V1.0/cd01/oaxal-v1.0-cd01.pdf
Latest Version:
http://docs.oasis-open.org/oaxal/V1.0/oaxal-v1.0.html
http://docs.oasis-open.org/oaxal/V1.0/oaxal-v1.0.pdf
Technical Committee:
Open Architecture for XML Authoring and Localization (OAXAL) TC
Chair(s):
Andrzej Zydroń
Editor(s):
Andrzej Zydroń
Derek Saldana
Related Work:
This specification replaces or supercedes:
N/A
Declared XML Namespace(s):
N/A
Abstract:
The Open Architecture for XML Authoring and Localization (OAXAL) provides a comprehensive, efficient, and cost-effective model for building an XML lifecycle production framework based completely on Open Standards from OASIS, LISA OSCAR and W3C.

Status:
This document was last revised or approved by the OAXAL TC on the above date. The level of approval is also listed above. Check the "Latest Version" or "Latest Approved Version" location noted above for possible later revisions of this document.

Technical Committee members should send comments on this specification to the Technical Committee's email list. Others should send comments to the Technical Committee by using the "Send A Comment" button on the Technical Committee's web page at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=oaxal.

For information on whether any patents have been disclosed that may be essential to implementing this specification, and any offers of patent licensing terms, please refer to the Intellectual Property Rights section of the Technical Committee web page (http://www.oasis-open.org/committees/oaxal/ipr.php).

The non-normative errata page for this specification is located at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=oaxal.

Notices

Table of Contents

1 Introduction

The Open Architecture for XML Authoring and Localization (OAXAL) represents a comprehensive, efficient, and cost-effective model regarding the authoring and translation aspects of XML publishing. OAXAL encompasses the following key Open Standards:

- XML - Extensible Markup Language (XML) is a simple, flexible text format originally designed to meet the challenges of large-scale electronic publishing. XML also plays an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.
- Unicode - A character encoding scheme that encompasses all character sets.
- W3C ITS - An XML vocabulary that defines translatability rules for a given XML document type.
- SRX - Segmentation Rules eXchange, a LISA OSCAR standard defining text-subdivision rules for each

language.

- xml:tm - XML-based text memory, a LISA OSCAR standard for author memory (a history of segments and revisions) and translation memory (a history of translated segments).
- GMX - Global Information Management Metrics Exchange, a LISA OSCAR standard for word and character count and metrics (for volume, complexity, and quality) exchange.
- TMX - Translation Memory eXchange, a LISA OSCAR standard for exchanging translation memories.
- Unicode TR29 - The primary Unicode standard defining word and sentence boundaries.
- Open Standard XML Vocabularies, including DITA, Docbook, XHTML, SVG, ODF, and others that may emerge as standards.
- XLIFF - XML Localization Interchange File Format, an OASIS standard for exchanging Localization data.

The architectural model for OAXAL is as follows:

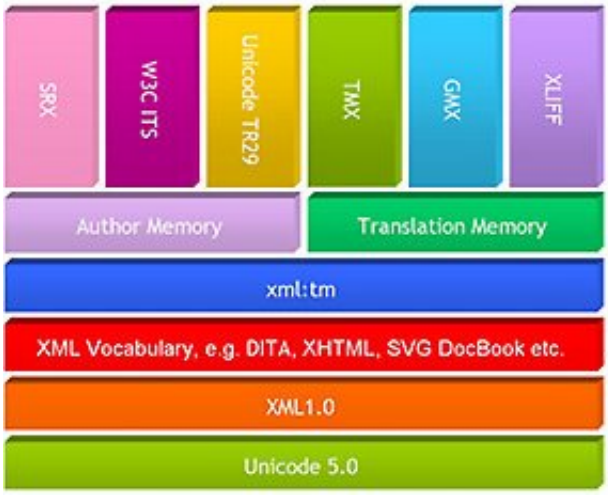

Figure 1: OAXAL Standards Component Stack

The following diagram represents the interaction of the key OAXAL standards:

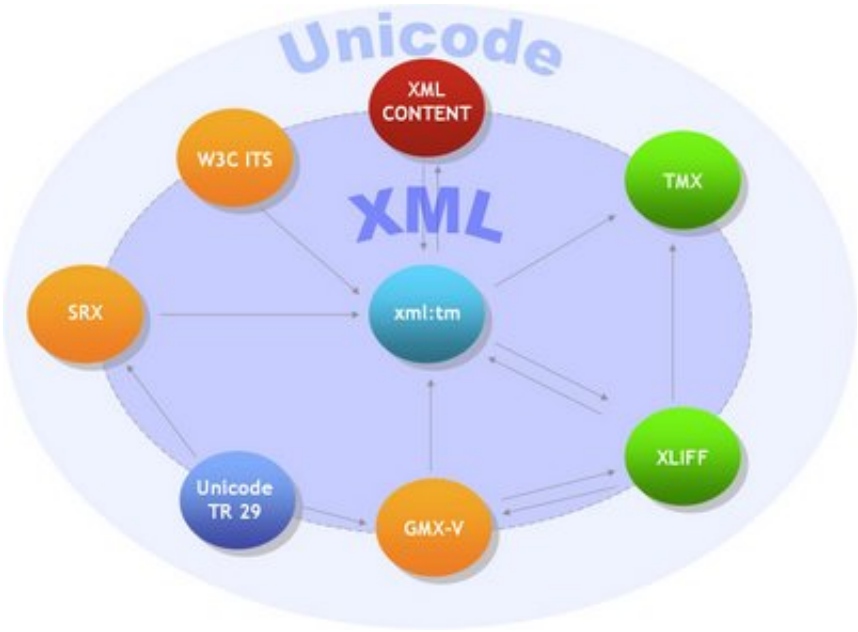

Figure 2: OAXAL Interaction of Standards

This diagram is annotated and described in detail in subsequent parts of this document. OAXAL is designed to cope with the common requirements for XML authoring and Localization. The authoring plus Localization aspects of OAXAL are most effective within a Content Management System (**CMS**) environment. For a translation-only **workflow**, OAXAL can be implemented without a CMS system.

OAXAL is designed to integrate tightly and transparently within the document-life-cycle workflow model which includes:

- document creation
- the authoring cycle
- Localization
- subsequent document updates and the Localization thereof

For the translation-only environment, OAXAL provides an elegant and open architecture for processing XML documents for translation.

1.1 What is a reference model

A *reference model* is an abstract framework for understanding significant relationships among the entities of some environment. It enables the development of specific reference or concrete architectures using consistent standards or specifications supporting that environment. A reference model consists of a minimal set of unifying concepts, axioms, and relationships within a particular problem domain and is independent of specific standards, technologies, implementations, or other concrete details.

As an illustration of the relationship between a reference model and the architectures that can derive from such a model, consider what might be involved in modeling important aspects of residential housing. In the context of a reference model, we know that concepts such as eating areas, hygiene areas, and sleeping areas are all important in understanding what goes into a house. There are relationships among these concepts and constraints on their implementation. For example, there may be a physical separation between eating areas and hygiene areas.

The role of a reference architecture for housing would be to identify abstract solutions to the problems of providing housing. A general pattern for housing, one that addresses the needs of its occupants in the sense of, say, noting that there are bedrooms, kitchens, hallways, and so on is a good basis for an abstract reference architecture. The concept of "eating area" is a reference model concept; a kitchen is a realization of "eating area" in the context of the reference architecture.

There may be more than one reference architecture that addresses how to design housing; for example, there may be a reference architecture to address the requirements for developing housing solutions in large apartment complexes, another to address suburban single family houses, and another for space stations. In the context of high-density housing, there may not be a separate kitchen but rather a shared cooking space or even a communal kitchen used by many families.

An actual – or *concrete* – architecture would introduce additional elements. It would incorporate particular architectural styles, particular arrangements of windows, construction materials to be used, and so on. A blueprint of a particular house represents a specific architecture as it applies to a proposed or an actual constructed dwelling.

The reference model for housing is, therefore, at least three levels of abstraction away from a physical entity that can be lived in. The purpose of a reference model is to provide a common conceptual framework that can be used consistently across different implementations and is of particular use in modeling specific solutions.

1.2 A Reference Model for Open Architecture for XML Authoring and Localization

The goal of this reference model is to define the component parts of XML publishing with respect to the authoring and Localization aspects of the process. It provides a normative reference that remains relevant for OAXAL as a comprehensive model.

The OAXAL standards components stack shows how the reference model for OAXAL is constructed from its constituent Open Standards. The concepts and relationships defined by the reference model are the basis for describing the reference architecture.

Architecture must account for the goals, motivation, and requirements that define the actual problems being addressed. While reference architectures can form the basis of classes of solutions, concrete architectures will define specific solution approaches.

Architecture is often developed in the context of a pre-defined environment, such as the protocols, profiles, specifications, and standards that are pertinent.

OAXAL implementations combine all of these elements, from the more generic architectural principles and infrastructure to the specifics that define the current needs, and represent specific implementations that will be built and used in an operational environment.

1.3 Audience

The intended audiences of this document include (non-exhaustively):

- Architects and developers designing, identifying, or developing a system based on OAXAL
- Standards architects and analysts developing specifications that rely on OAXAL
- Decision makers seeking a "consistent and common" understanding of OAXAL
- Users who need a better understanding of the concepts and benefits of OAXAL

1.4 Guide to using the reference model

New readers are encouraged to read this reference model in its entirety. Concepts are presented in an order that the authors hope promote rapid understanding.

This section introduces the conventions, defines the audience, and sets the stage for the rest of the document. Non-technical readers are encouraged to read this information because it provides background material necessary to understand the nature and use of reference models.

- The Open Architecture for XML Authoring and Localization section introduces the concept of OAXAL and identifies some of the ways that it differs from previous paradigms for authoring and translation systems. This section offers guidance on the basic principles of OAXAL. This section can be used by non-technical readers to gain an explicit understanding of the core principles of OAXAL and by architects as guidance for developing OAXAL-based architectures.
- The Reference Model section introduces the Reference Model for OAXAL.
- The Conformance Guidelines section addresses compliance with this reference model.

The glossary provides definitions of terms within the reference-model specification but does not necessarily form part of the specification itself. Terms that are defined in the glossary are marked in bold at their first occurrence in this document.

Note that while the concepts and relationships described in this reference model may apply to other "service" environments, the definitions and descriptions contained herein focus on the field of software architecture and make no attempt to completely account for use outside of the software domain. Examples included in this document that are taken from other domains are used strictly for illustrative purposes.

1.5 Notational Conventions

The key words MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL in this document are to be interpreted as described in RFC2119.

1.6 Normative References

[W3C ITS] Internationalization Tag Set (ITS) Version 1.0 http://www.w3.org/TR/2007/REC-its-20070403/ 03 April 2007

[SRX] SRX 2.0 Specification http://www.lisa.org/Segmentation-Rules-e.40.0.html OSCAR Recommendation, 7 April 2008

[xml:tm] XML Text Memory (xml:tm) 1.0 Specification http://www.lisa.org/XML-Text-Memory-xml.107.0.html 26 February 2007

[GMX] Global Information Management Metrics Volume (GMX-V) 1.0 Specification http://www.lisa.org/Global-information-m.104.0.html Version 1.0, 26 February 2007

[TMX] Translation Memory eXchange format (TMX) Specification http://www.lisa.org/fileadmin/standards/tmx1.4/tmx.htm Version 1.4b, 26 April 2005

[Unicode TR29] Unicode Standard Annex #29 http://unicode.org/reports/tr29/ Unicode 5.1.0 2008-03-25

[DITA] OASIS Committee Specification 01, Darwin Information Typing Architecture Version 1.1, May 2007. http://docs.oasis-open.org/dita/v1.1/CS01/overview/overview.html

[DocBook] OASIS Committee Specification, DocBook 4.2, July 2002. http://www.docbook.org/specs/cs-docbook-docbook-4.2.pdf

[XHTML] XHTML™ 1.1 - Module-based XHTML - Second Edition http://www.w3.org/TR/xhtml11/ W3C Working Draft 16 February 2007

[SVG] Scalable Vector Graphics http://www.w3.org/TR/SVG12/ Working Draft 1.2, April 2005

[ODF] OASIS Standard, Open Document Format for Office Applications (OpenDocument) 1.1, February 2007. http://docs.oasis-open.org/office/v1.1/OS/OpenDocument-v1.1-html/OpenDocument-v1.1.html

[XLIFF] OASIS Standard, XML Localization Interchange File Format (XLIFF) 1.2, February 2008. http://docs.oasis-open.org/xliff/v1.2/os/xliff-core.html

1.7 Non-Normative References

[LISA] Localization Industry Standards Association

[LISA OSCAR] LISA Open Standards for Container/content Allowing Reuse

[OASIS] Organization for the Advancement of Structured Information Standards

[W3C] The World Wide Web Consortium

2 Open Architecture for XML Authoring and Localization

Open Architecture for XML Authoring and Localization (OAXAL) is a reference model of how to construct an effective and efficient system for XML authoring and Localization based on Open Standards. OAXAL comprises the following standards:

- XML -- Extensible Markup Language (XML) is a simple, flexible text format originally designed to meet the challenges of large-scale electronic publishing. XML also plays an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.
- Unicode - A character encoding scheme that encompasses all character sets.
- W3C ITS W3C ITS - An XML vocabulary that defines translatability rules for a given XMLdocument type.
- Unicode TR29 - The primary Unicode standard defining word and sentence boundaries.
- SRX - Segmentation Rules eXchange, an XML vocabulary defining segmentation rules for each language.
- xml:tm - XML-based text memory, a LISA OSCAR standard for author and translation memory.
- GMX - Global Information Management Metrics Exchange, a LISA OSCAR standard for word and character count and metrics exchange.