



OASIS Open Office XML Document Format

Michael Brauer

Technical Architect

Sun Microsystems GmbH



About the Speaker

- Technical Architect in OpenOffice.org/StarOffice development
- OOo/StarOffice developer since 1995
- Main focus: XML Technologies
- OpenOffice.org XML Project Owner
- OASIS Open Office Technical Committee Chair

Agenda

- Office Applications and XML - Overview
- OASIS and OASIS Open Office Technical Committee
- Open Document Format – The Benefits
- OpenOffice.org XML and OASIS Open Office XML – The Changes

Office Applications and XML

An Overview

Office Applications and XML

An Overview

- Office versus Custom Schema
 - Office Schema: Schema for Office Documents
 - Tailored to Office Documents
 - Feature Complete
 - Custom Schema: Any Other Schema
 - Standard Schemas: UBL, XHTML, DocBook, etc.
 - User Defined Schemas: Defined by Customer
 - Bills, Orders, Meeting Minutes, etc.

Office Applications and XML

An Overview

- Office Schema
 - "Own your data" for Office Documents
 - Interoperability
 - Long term preservation
 - Loading and Saving of Office Documents
 - OpenOffice.org:
 - OpenOffice.org XML schema since OOo 1.0
 - OASIS Open Office schema since OOo 2.0

Office Applications and XML

An Overview

- Custom Schema I: XSLT Support
 - Import of Custom Schema Documents for
 - Editing, Printing, Publishing as PDF, etc.
 - Example: DocBook
 - Export of Custom Schema Documents for
 - Publishing, Viewing, etc.
 - Example: DocBook, XHTML
 - Rendering of Custom Schema Documents for
 - Printing, Publishing as PDF, etc.
 - Example: UBL, user defined order schema
 - OpenOffice.org: XSLT support since OOo 1.1

Office Applications and XML

An Overview

- Custom Schema II: XForms Support
 - Form-like Entry of Structured/Custom Schema Data For
 - Submission, Saving, etc.
 - Example: UBL, user defined order schemas
 - Form-like Editing of Structured/Custom Schema Data
 - Example: UBL, user defined order schemas
 - XForms Supports
 - Data Types, Calculations, Validations
 - OpenOffice.org: XForms support since OOo 2.0

OASIS and Open Office TC

OASIS

Organization for the Advancement of
Structured Information Standards

- **Not-For-Profit, Global Consortium**
 - Drives development, convergence and adoption of e-business standards.
 - Horizontal and e-business frameworks, web services, security, public sector, vertical industry applications
- **Technical Work Is Driven By Members**
- **OASIS provides guidance, process, and infrastructure**

OASIS

Organization for the Advancement of
Structured Information Standards

- Founded In 1993 As “SGML Open”
 - Renamed to OASIS in 1998
- More than 600 members
 - Corporate and individual members
 - Members from 100 countries
- OASIS Work/Standards: ebXML, UDDI, DocBook, UBL, etc.
- More info: <http://www.oasis-open.org>

OASIS Open Office TC

- Founded 2002
- Charter: Creation of an open, XML based file format for office applications
 - Called “OASIS Open Office XML Format”
- Members: include OpenOffice.org
- Status:
 - Committee Draft available
 - Standard approval planned for Jan. 2005
 - ISO standardization planned for 2005

OASIS Open Office TC

- Basis: OpenOffice.org format
- *"The OASIS Open Office TC is the place where office vendors and users that are interested in an open file format for office applications meet to agree on that format. For OpenOffice.org and StarOffice, it is the place where their file formats are developed."*
- More info: <http://www.oasis-open.org>

Open Office Document Format

Open Office Document Format – The Benefits

Open Office Document Format

Open Office Document Format – The Benefits

Completeness

Open Office Document Format – The Benefits

- Completeness - Why is this important?
 - Long term preservation for *all* office documents
 - Key requirement for interoperability
 - Precondition for being default file format
- Open Office Schema
 - Covers all OpenOffice.org applications
 - Covers all their features

Default File Format

Open Office Document Format – The Benefits

- Default file format - Why is this important?
 - Ensures *all* data created can be processed
 - Avoids use of historic binary formats
 - Documents can be reused
 - By other office applications
 - For tasks not yet known
- Open Office Schema
 - Default since OpenOffice 1.0

File Size

Open Office Document Format – The Benefits

- File Size - Why is this important?
 - Transfer of documents
 - Bandwidth for E-mail transfer
 - Storage requirements for archiving
- Open Office Schema
 - Zip files as container
 - Document and images in one physical file
 - Small file size through compression

Open Office Document Format

Open Office Document Format – The Benefits

Vendor Independence

Open Office Document Format – The Benefits

- Vendor Independence - Why is this important?
 - Independence of vendor's goodwill
 - Dependence can hinder standardization
 - Dependence can lead to monopolistic markets
- Open Office Schema
 - Controlled by standardization body (OASIS)
 - Development follows OASIS process

Open Licence

Open Office Document Format – The Benefits

- Open Licence - Why is this important?
 - Usage Restriction
 - May prevent open source/competing application from implementing/using a format
 - License changes
 - May impact business models
 - May have unforeseeable costs

Open Licence

Open Office Document Format – The Benefits

- Open Office Schema
 - Perpetual, royalty free, non discriminatory license
 - OASIS Open Office TC charter has openness standing rule

Extensibility

Open Office Document Format – The Benefits

- Extensibility - Why is this important?
 - Fosters format development and improvement
 - Enables building formats on top of a standard
 - Improves interoperability/backward compatibility
- Open Office Schema
 - Influenceable by OpenOffice.org community
 - Influenceable by becoming OASIS Open Office TC member
 - Extensibility part of specification

Interoperability

Open Office Document Format – The Benefits

- Interoperability - Why is this important?
 - Fosters innovation and competition
 - Supersedes need for proprietary format implementations
 - Ensures information is accessible
 - Cheaper development of tools
 - Cheaper support by ISVs

Interoperability

Open Office Document Format – The Benefits

- Open Office Schema
 - Interoperability explicit goal
 - Designed independently of OpenOffice.org application
 - Proof of concept by KOffice adoption

Reuse of Standards

Open Office Document Format – The Benefits

- Reuse of standards - Why is this important?
 - Tools already support established standards
 - Cheaper support by ISVs
 - Cheaper training costs
- OASIS Open Office Schema
 - Makes use of XHTML, SVG, XSL, MathML, XLink and Dublin Core Meta Initiative
 - Reuses its own schema fragments

Conclusion

Open Office Document Format – The Benefits

- Requirements depend on each other
- Open Office Format Meets all requirements
- No borderline between important/unimportant requirements
- Openness is 1st class goal - from the beginning

Technical Details

Technical Details

Changes To the OpenOffice.org XML Format

- New XML Namespaces
- Schema Language: Relax-NG
 - XSD can be generated from Relax-NG
 - More powerful than DTD, Simpler than XSD
- Schema/Validation improvements
 - Elements for Document Classes, etc.
 - Attribute Value Data Types (XSD)
- Adoptions to new Standard Versions
 - SVG 1.1, CSS3, etc.

Technical Details

Changes To the OpenOffice.org XML Format

- Adoption to new/other Office Applications
- Removal of Inconsistencies
- Error Corrections



OASIS Open Office XML
Document Format

michael.brauer@sun.com

