



# OASIS Open Office XML File Format

Dieter Loeschky  
Sr. Manager Software  
Engineering  
Sun Microsystems



# What's Wrong With Your Office

- Voice of the Customer
  - office productivity applications
  - need additional processing, integration
    - archiving and indexing
    - content checking
    - database, work-flow integration
    - operation on many files
  - unstructured work-flow, many document types
- Requirement

# XML Office File Format

- Solution: XML Office File Format
  - XML, for easy integration, processing
  - define office vocabulary
  - requirements for office file format
    - full featured, cover full office productivity space
    - easy to process, easy to generate
    - not vendor or application specific
- OpenOffice.org XML File Format
  - tailored to these requirements

# Standardizing The Format

- A File Format is an Investment
  - must be open, documented
  - must be stable, controlled evolution
- Win-Win with Widespread Adoption
  - more support, more tools
  - need user and industry acceptance
- Develop Standard Format at OASIS
  - based on OpenOffice.org XML Format
    - format has proven useful in real life

# OASIS

- OASIS = Organization for the Advancement of Structured Information Standards
- Non-Profit, global consortium for e-business standards
  - Web services, business transactions, electronic publishing, interoperability within and between marketplaces, etc.
- Lightweight and open processes
  - expressly designed to promote industry consensus and unite disparate efforts
- More than 600 corporate and individual

# OASIS

## Open Office XML File Format

- Technical Committee (TC)
  - Office Application Vendors
    - Corel, KOffice, OpenOffice.org, Sun Microsystems (chair)
  - XML Tool Vendors
    - Arbortext, Corel, Stellent
  - Industry and User Groups
    - Boeing, National Archive of Australia, 1dok, SpeedLegal, Sony, etc.

# Designing a File Format

- Consistent Design
  - common format across all applications
  - one concept, one representation
- Reviewed Design Process
  - examine existing formats
    - use MathML, XLink, Dublin Core
    - reuse from XHTML, SVG, XSL-FO, CSS
  - specify, review, finally implement

# Packages

- XML File Format Concerns

- file size
- embedded images, objects

- XML Package

- ZIP format, XML-based manifest
- XML streams + binary streams
  - images, OLE objects, printer setup data
- used by StarOffice, OpenOffice.org, Gnome, KOffice



# OASIS Roadmap

- TC founded in Dec. 2002
- Phase 1 – Standardizing the Format
  - using the OpenOffice.org XML Format as base
  - TC specification in late summer 2003
  - OASIS specification in the fall 2003
  - implement specification in the next release of StarOffice / OpenOffice.org
- Phase 2 – Enhancing the Format

# XML in StarOffice 6.0

- StarOffice 6.0 relies on the OpenOffice.org XML File Format
  - XML is default file format
    - binary file format is deprecated
  - full featured
  - open sourced and documented at OpenOffice.org
  - OpenOffice.org XML File Format is base for the standardization at OASIS

# XML in StarOffice 6.1

- New XML Filter Dialog
  - easy to integrate all XML formats by using XSLT transformations
  - easy to validate custom XSLT transformations
  - easy to register custom XSLT transformations as an import / export filter
  - sample XSLT transformations for
    - upcoming MS Word 2003 XML format
    - DocBook
    - XHTML

# XML in StarOffice 6.1 (cont.)

- XML related projects
  - FIGARO
    - 'XML-based Approaches for ETD's within OpenOffice'  
Volker John, University of Hamburg
  - XDiML
    - 'Save as XDiML (Dissertation Markup Language), Writing and Converting digital Theses and Dissertations using OpenOffice'  
Matthias Schulz, Humboldt University
  - StarOffice / TAMINO integration
    - integration with Software AG's TAMINO XML server

# XML in StarOffice Q

- XML related issues
  - incorporate the standardized OASIS Open Office XML File Format
  - more XSLT transformations
    - upcoming MS Office 2003 Excel format
    - enhanced DocBook transformation
  - Digital signature and encryption

# Conclusion

- Office XML File Format
  - fully supports office users & applications
  - enables
    - processing of office documents
    - integration into custom infrastructure
- Standardization at OASIS
  - open standard, open development
  - secures your investment



Dieter Loeschky

Dieter.Loeschky@sun.com



We make the net work.