

Interoperability **by Design**

Connecting People, Data and Diverse Systems

Document Interoperability Initiative

Vijay Rajagopalan

Principal Architect, Interoperability Strategy

Microsoft Corporation

Agenda

- Interoperability Principles
- DII Goals
- Industry Participation & Previous DII Results
- Scenarios for Today's Discussion
- Open discussion

Microsoft Interoperability Principles

Announced on February 21, 2008

Ensuring Open Connections

Data Portability

Enhanced Support for Standards

Open Engagement

 **Windows Vista™**
including the .NET Framework

 **Microsoft® Office 2007**

 **Windows Server® 2008**

Microsoft® Exchange Server 2007

Microsoft® SQL Server™

 **Microsoft® Office SharePoint® Server 2007**

“The interoperability principles and actions announced today by Microsoft present significant opportunities for the vast majority of software developers and will benefit the broader IT community.”

- Thomas Vogel, Information Management, Novartis Pharma

Document Interoperability Initiative

Lab Events

- Test interoperability between existing implementations
- Roundtable to increase dialogue on key issues

Conformance Test Suites

- Identify developer & customer needs
- Facilitate community development for most popular formats

Templates

- Facilitate creation of a library of templates optimized for translation between popular formats (OXML, ODF, UOF, etc.)

Address issues of interoperability and effective data exchange between widely deployed document format implementations

Document Interop Roundtable

Dialogue and collaborate on these key issues

- What document interoperability issues exist today?
- Where should we collaborate as a community to solve issues?
- What tools, test suites and templates will enable developers to better implement formats?
- Where are customers demanding interoperability? (applications, formats, platforms)
- 6 successful DII roundtables so-far (Boston, Korea, China, Germany, Redmond)
- Planning Europe & Asia in Q3

Document Interop Roundtable – Boston, MA, USA

Dialogue and collaborate on key issues

Participants included: **Novell, Nuance, DataViz, Quickoffice and Mark Logic**

- Customers want high fidelity with no data loss
- Consider both presentation, authoring and exchange scenarios AND enterprise server scenarios (back office, unstructured data, schema across supply chains)
- Test suites are critical to ensure quality, reduce development time
 - Think creatively to solve this – could be web –based or web service validation tool
- Speed of document rendering is important to customers (esp. in mobile & roaming scenarios) and view and edit modes may need to be treated differently to accommodate
- Access to an accepted library of quality test documents is a gap

Document Interop Roundtable – Beijing, China

Participants included: CESI, Kingsoft, Evermore, Red Flag Chinese 2000, China Standard Software Co(CS2C), UOF Translator working group

- How to conduct conformance testing? In what level of OOXML implementation can be claimed that this SW comply with OOXML.eg:- iWork can't open all documents created by MS Office.
- Translator can't solve all interop issues, as mapping between different formats will lose something.
- Multiple translations between different formats will result in more interop issues.
- Consistency in Visual Fidelity among local vendor implementations.
- Managing Chinese Special Features

Document Interop Roundtable – Seoul, Korea

Dialogue and collaborate on key issues

- Create developer 'friendly' authoring tools for new document formats (ala 'XAML pad')
 - Provide real-time “WYSIWYG” feedback
- We need a "trusted" source for files
 - A library of pre-validated documents that are 100% compliant with spec
- Need better authoring tools for custom defined schema in Office Open XML
- Event Improvements
 - Showcase the output of past events and what other ISVs are doing

Document Interop Round Table – Munich, Germany

Participants included: German Library of the Blind, Daisy Consortium, Stunnware, BusinessBase BV, Dialogica GMBH, COC

- Emphasis on presentation, accessibility, exchange and enterprise integration scenarios (back office, unstructured data, schema across supply chains)
- High Fidelity Roaming Scenarios (Mobile, Browsers)
- Need seamless back office Document Conversions
 - Automation & Authoring(eg:- Digital Certificates)
 - Printer Friendly document conversions.
 - Extracting Custom Data & Integrations
- Daisy
 - Preserving Structure & ensuring no data loss
 - Support for Browsers applications & content
- Developer Productivity & Toolset

A large, stylized wireframe sphere is positioned on the left side of the slide. It is composed of a grid of thin, light-colored lines that form a series of triangles across its surface. The sphere is partially cut off by the left edge of the frame. The background of the entire slide is a solid, medium-orange color.

DII RESULTS TILL DATE

Translation options

Translation

1. Built-in features of the product/application (authoring tools):

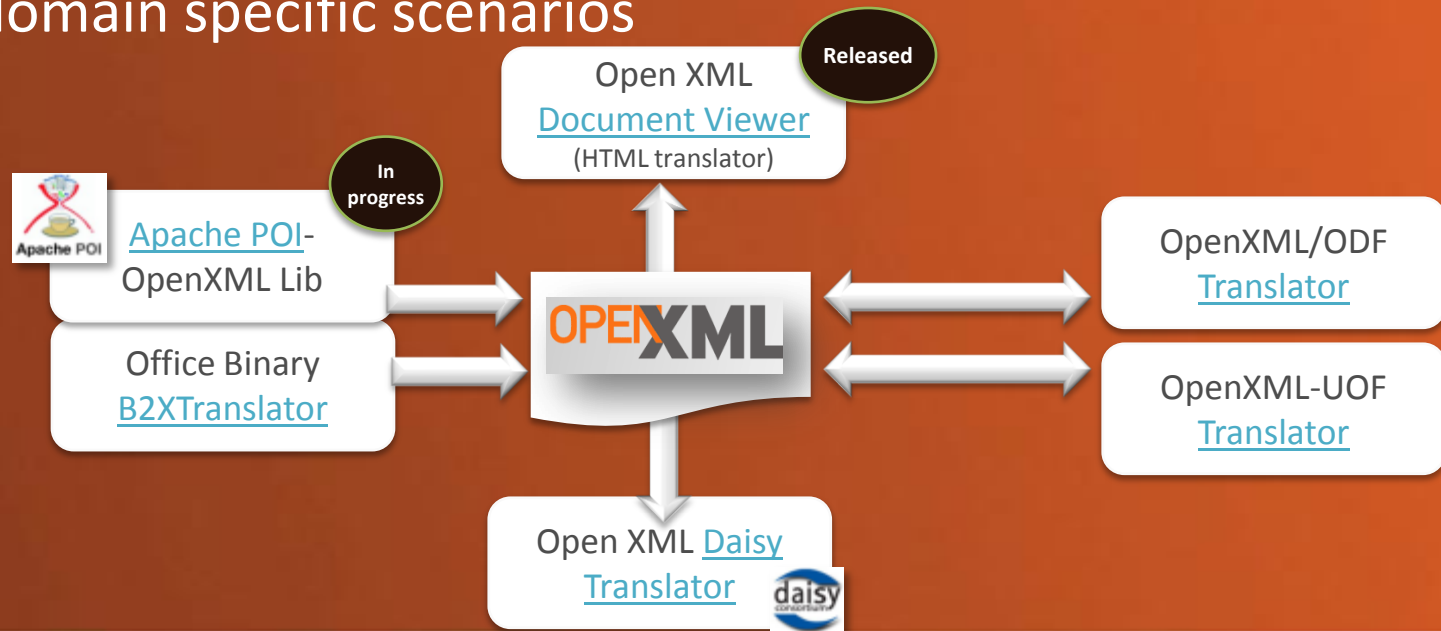
Product
Features &
Functionality

MS Office

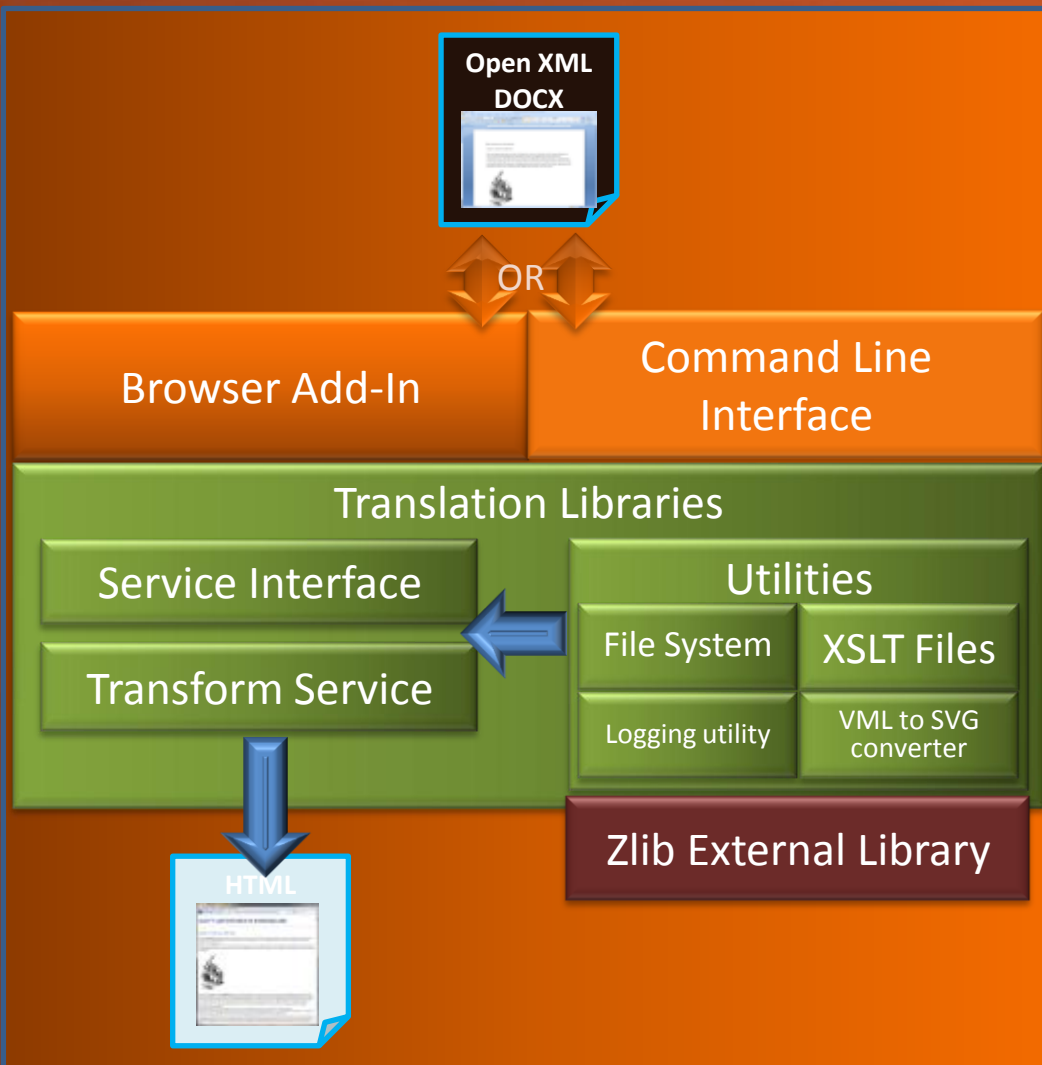
Save as OpenXML, Office Bin, PDF, HTML...
Save as ODF

Office
2007 SP2

2. Open Source Translation Engines projects sponsored by Microsoft : add-ins to Products/Applications & Pluggable on domain specific scenarios

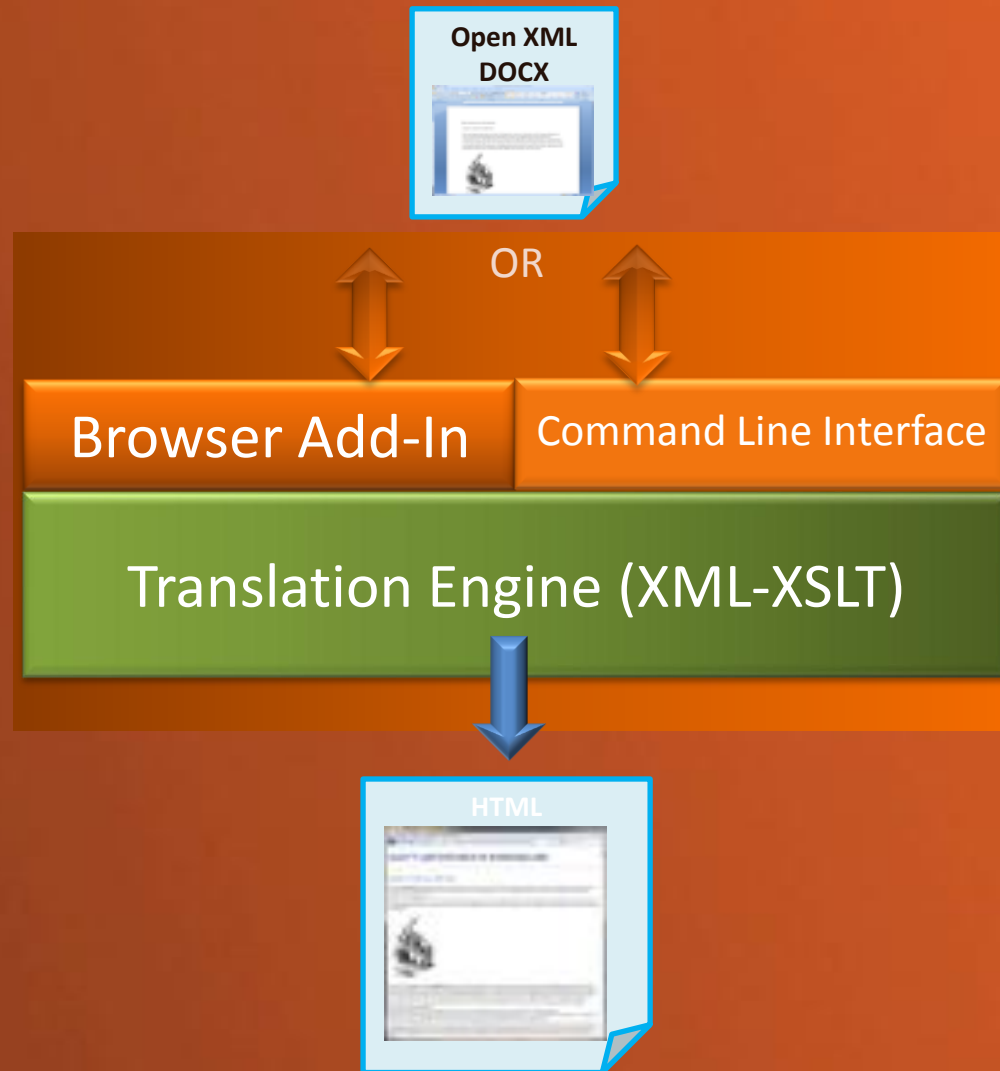


Announcing Open XML Document viewer/HTML Translator



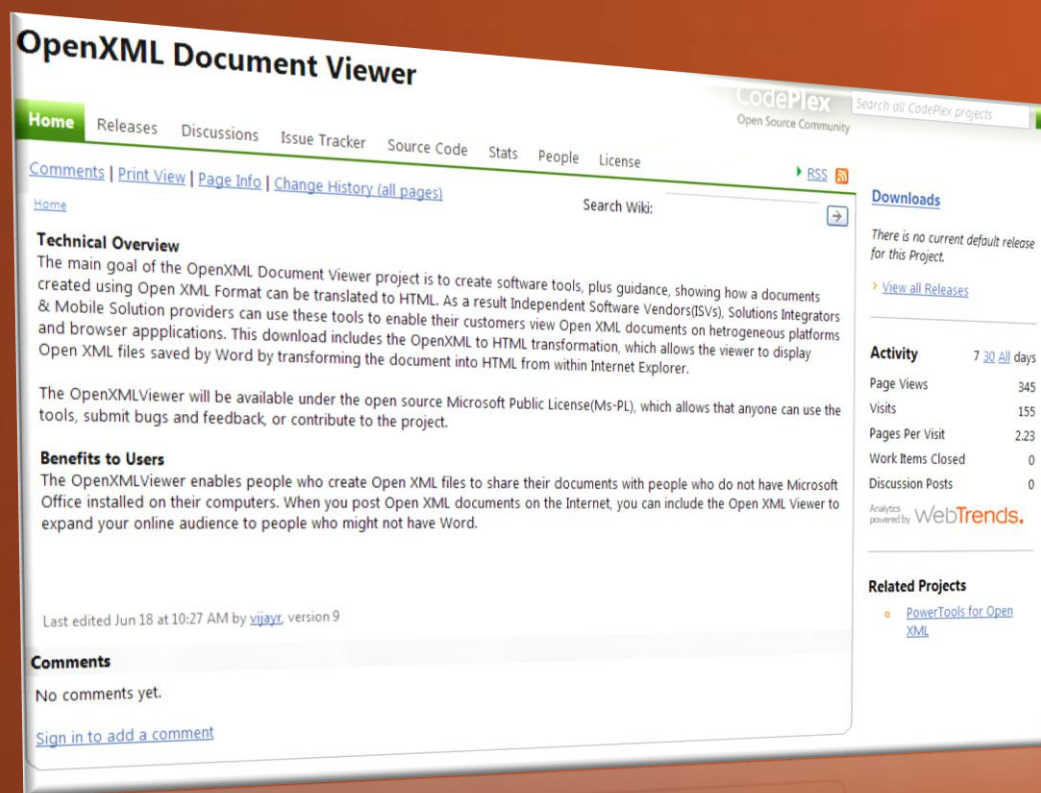
- Open XML viewing in “Roaming Scenarios” – aka web friendly formats, form factor independence
- Architecture
 - Use Componentized/Pluggable architecture to support heterogeneous domain specific scenarios
 - Implementations:
 - Browser Add-In
 - Command Line Interface
- Features
 - Converts .DOCX files into HTML data
 - Converts VML graphics within .DOCX files to SVG graphics
 - Leverages XSL Transformations

Open XML Document viewer/HTML Translator



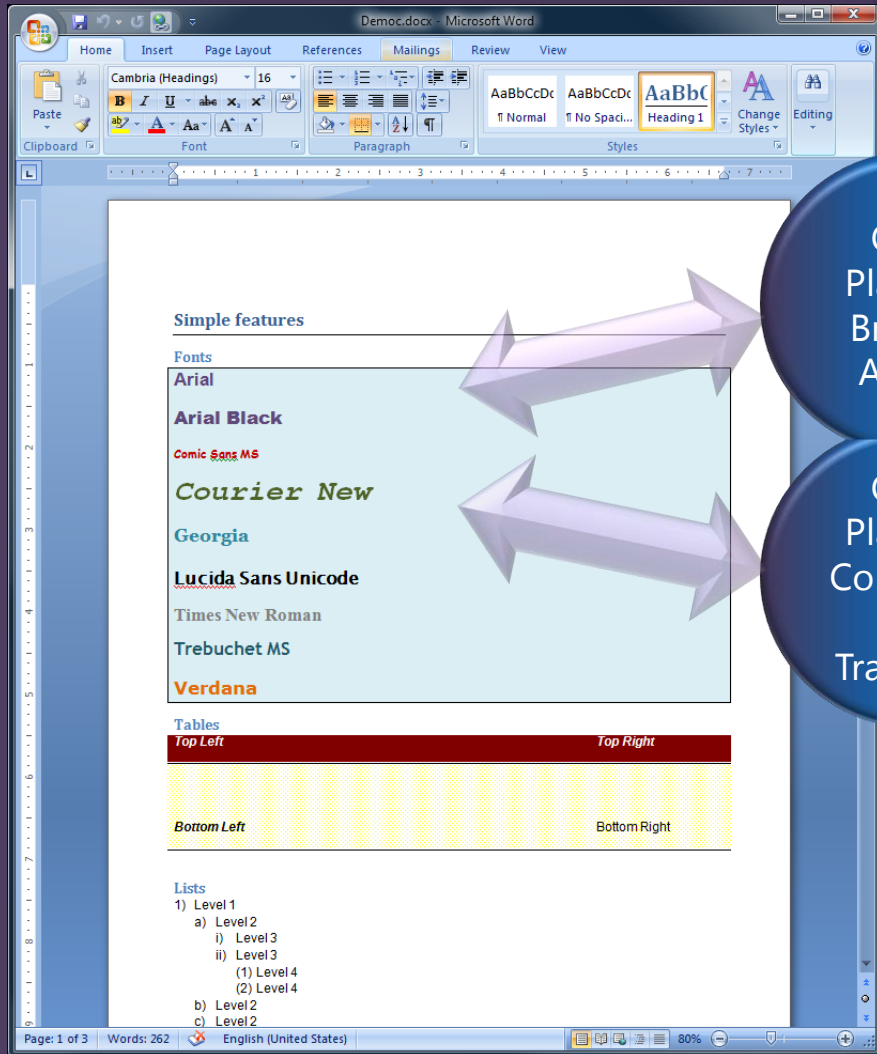
Open XML Document viewer/HTML Translator

- Project source available on Codeplex:
<http://www.codeplex.com/OpenXMLViewer>



Open XML Document viewer/HTML Translator

OPEN XML DOCX



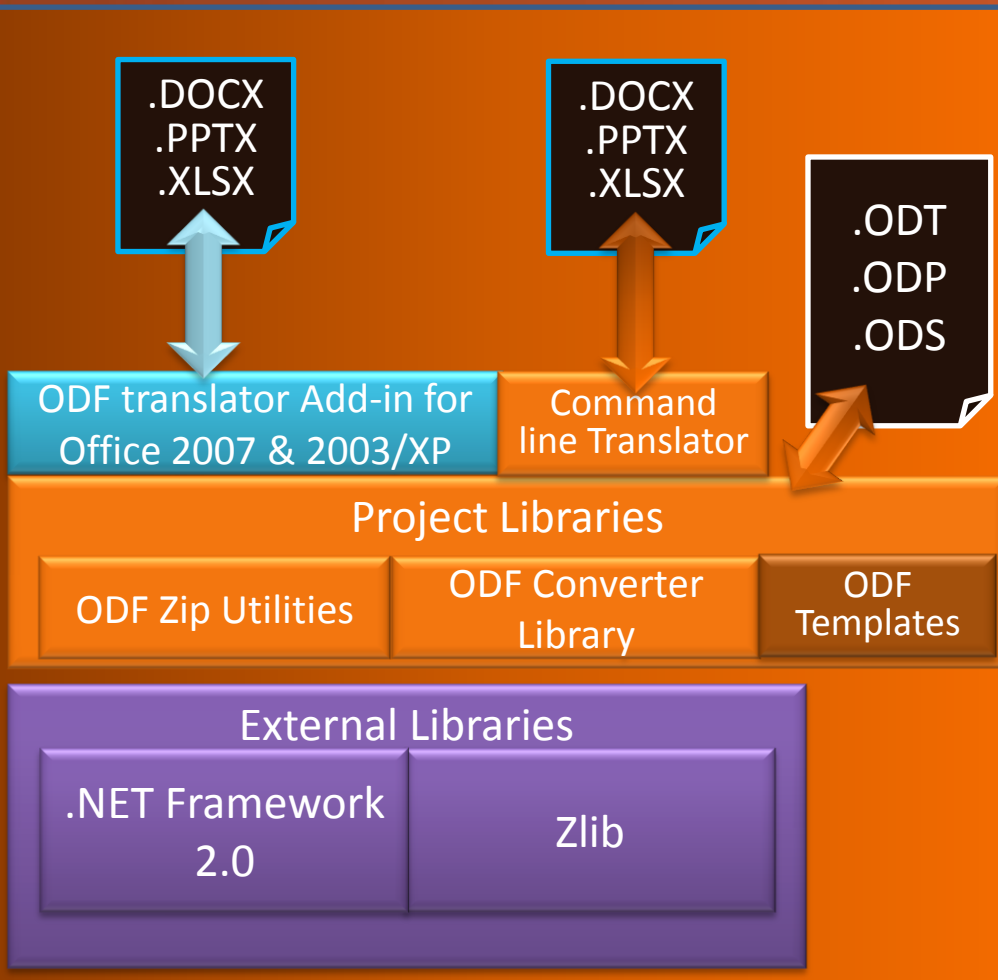
Cross
Platform
Browser
Add-In

Cross
Platform
Command
Line
Translator

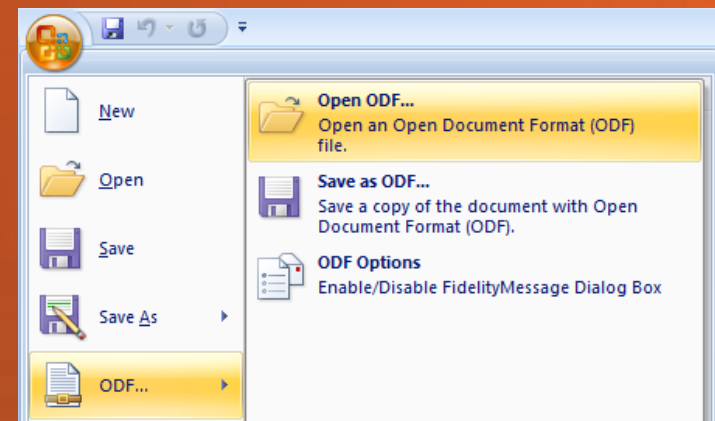
HTML



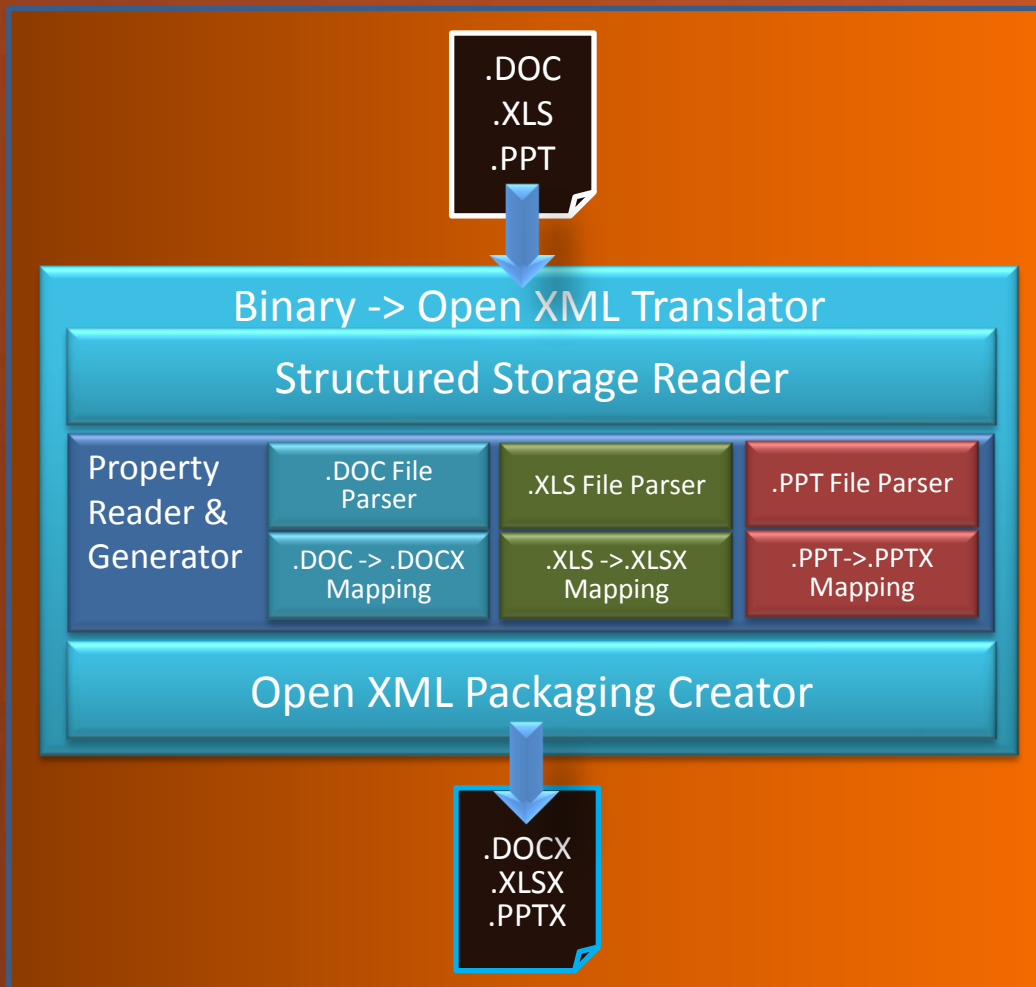
OpenXML/ODF Translator overview



- XML/XSL Transformation techniques
- Localized in six languages
- Transformation Engine can be decoupled:
 - Deployed in a number of domain specific scenarios: Back office, Hosted Service, Batch processing...
- Enables “Open/Save as ODF” in Word, Excel & Powerpoint menus:
 - Supports Office 2007 & 2003/XP
- Novell redistributes for SuSe Linux



Binary (.doc, .xls, .ppt) -> Open XML Translator Project



- Goal:
 - provide Pragmatic software tools showing how a document written using the Binary Formats (.doc, .xls, .ppt) can be translated into the Office Open XML format
 - Prescriptive Guidance & Feature Mapping
- Openly available to community under BSD license.
- Initial release to support .doc
- support for .xls and .ppt to follow in later revisions

Technical specifications for the binary formats available under Open Specification Promise

(<http://www.microsoft.com/interop/docs/officebinaryformats.mspx>)

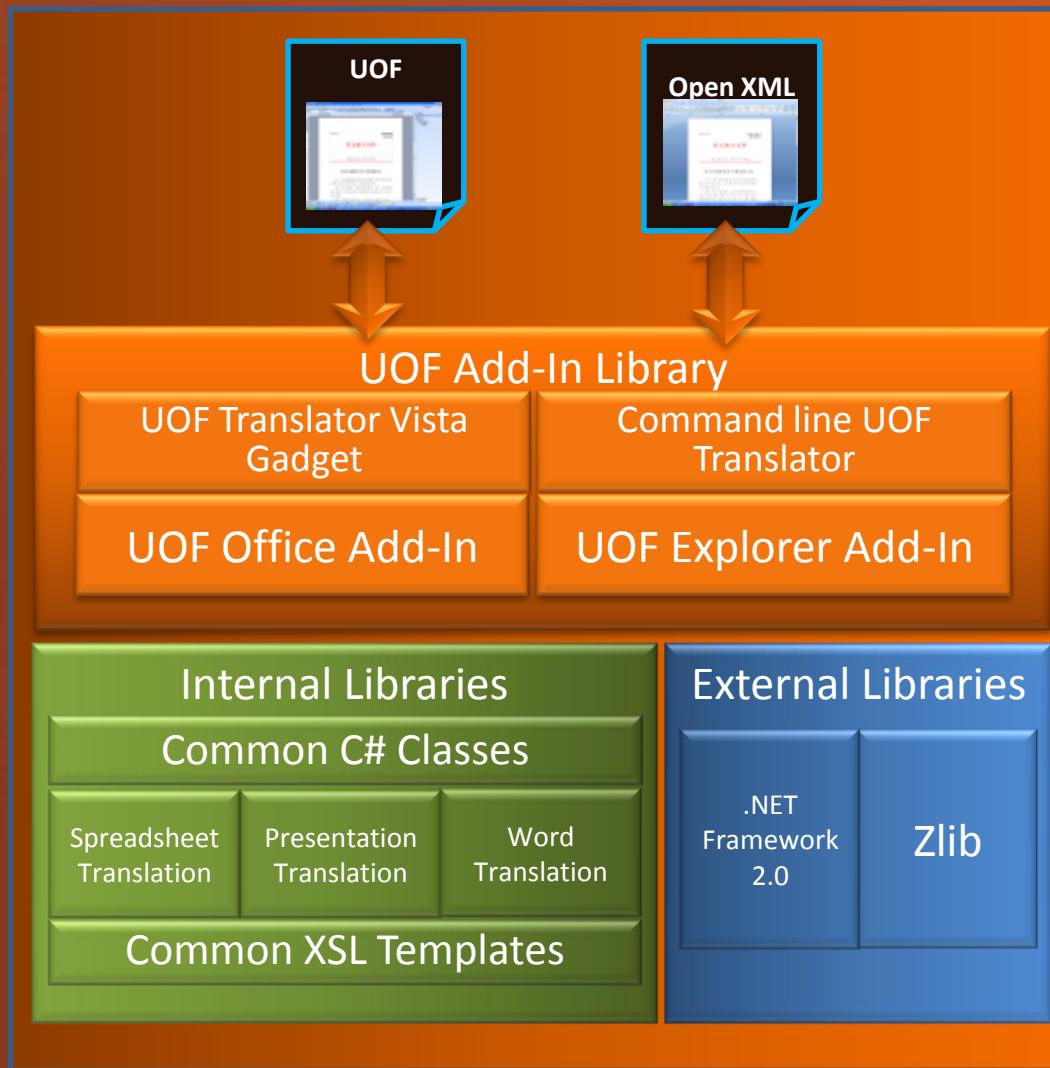
Download: <http://msdn.microsoft.com/en-us/library/cc216514.aspx>

Project on Sourceforge:

<http://b2xtranslator.sourceforge.net/>

Microsoft

OpenXML/UOF Translator overview



- Allows translation of the Chinese specific UOF format to the Open XML standard
- Bi-directional translation architecture
- Command line interface for scripting/batch file translations
- Enables “Open/Save as UOF” in Word, Excel & Powerpoint menus:
 - Supports Office 2007 & 2003
- Windows Vista Translation Gadget
- Windows Explorer Add-In
- Open Source project available on Sourceforge.net

UOF FORMAT

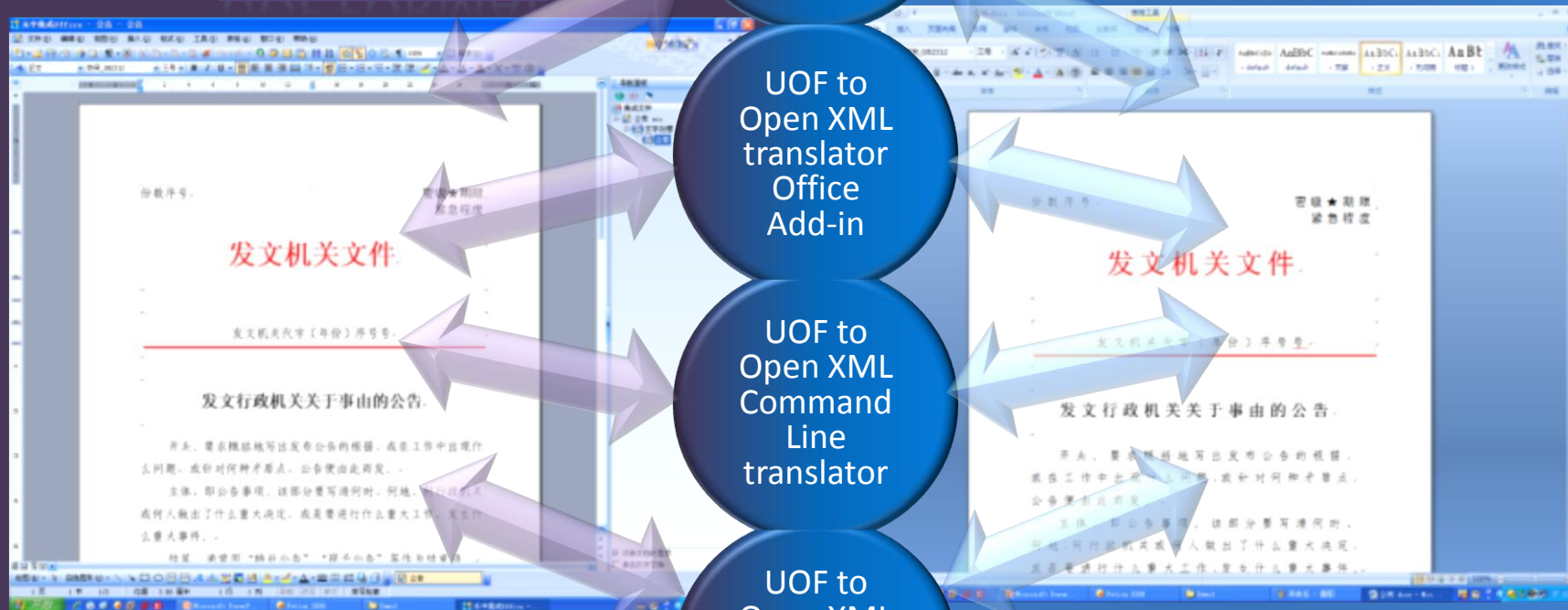
UOF to
Open XML
translator
Explorer
Add-In

OPEN XML FORMAT

UOF to
Open XML
translator
Office
Add-in

UOF to
Open XML
Command
Line
translator

UOF to
Open XML
translator
Vista
Gadget



Partner Solutions & Presentations

Scenario	Solution	Partner
DII Follow up & Translator Updates	OpenXML Document Viewer ODF Translator 2.5 Release	Microsoft
Document Archival, Preservation & Translation as a Service	Case study on Archival & Preservation efforts co-funded by EU, Microsoft & Others	Dialogika
Apache POI Case Study	Software Development Kit for Java Developers	Sourcesense
Document Interoperability from and Open Source perspective: PHPExcel and OPENXML4J	Cross Language Support for Document Interoprability	Maarten Balliauw & Julien Chable

A large, stylized wireframe sphere is positioned on the left side of the slide. It is composed of a grid of thin, light-colored lines that form a series of triangles across its surface. The sphere is partially cut off by the left edge of the frame. The background of the entire slide is a solid, medium-orange color.

FOR MORE INFORMATION

<http://www.microsoft.com/interop>

Microsoft®