



ODF API - ODFDOM

Svante Schubert
Software Engineer
Sun Microsystems, Hamburg



Do you know ODF?

- The OASIS / ISO standard for office documents (2005/06)
- The document format of many office applications
- A zipped package of XML and related files (image, sound, user files)
- Origin from OpenOffice.org's default format

OASIS 

Open Document Format for Office
Applications (OpenDocument) v1.0

OASIS Standard, 1 May 2005



International
Organization for
Standardization

ISO/IEC 26300

What to do with ODF documents?

- Usually store your Texts / Spreadsheets / Presentation
- Edited and viewing by an Office application



What to do with ODF documents?

- But how do I process 1.000.000 ODF documents?



We need an ODF API!

- API to automate ODF processing
 - > Creating, manipulating ODF documents
 - > Lightweight API
 - > API close related to ODF
 - > Opensource



Where to find an ODF API?

- Use StarOffice / OpenOffice.org API?
 - > Office API not ODF centric
StarOffice funded 1984, ODF standardized 2005 / 06
 - > Office viewing & editing functionality undesired,
Lightweight editing functionality not easy extractable



The Idea of a new ODF API!

- We need a new lightweight ODF API!!
 - > New API will focus on ODF
 - > Taking over ideas from previous ODF APIs (OOo API, AODL, ODF4J, etc.)



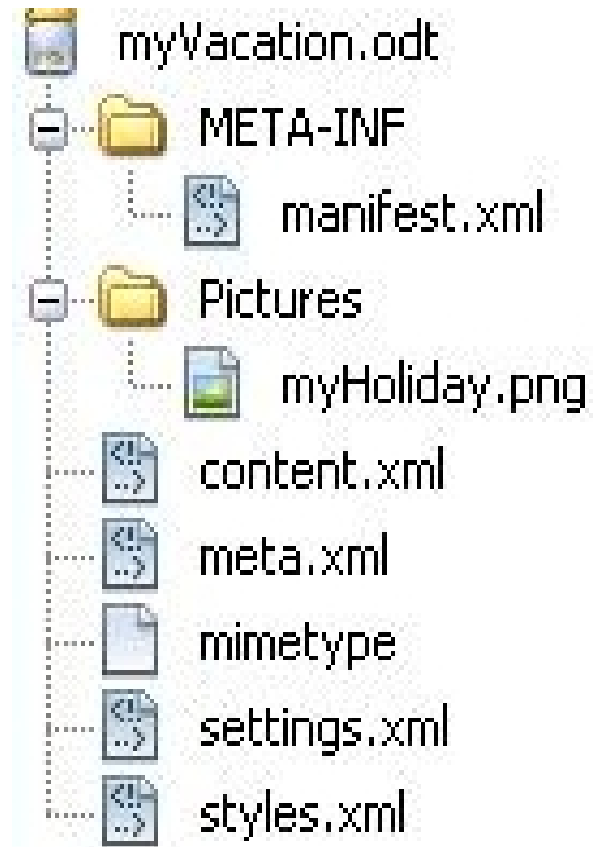
New ODF API - ODFDOM

- Sun open sourced ODFDOM
 - > Lightweight API
 - > OpenDocument centric
 - > Opensource (Apache 2)
 - > Multi-layered
 - > Java 5 reference implementation



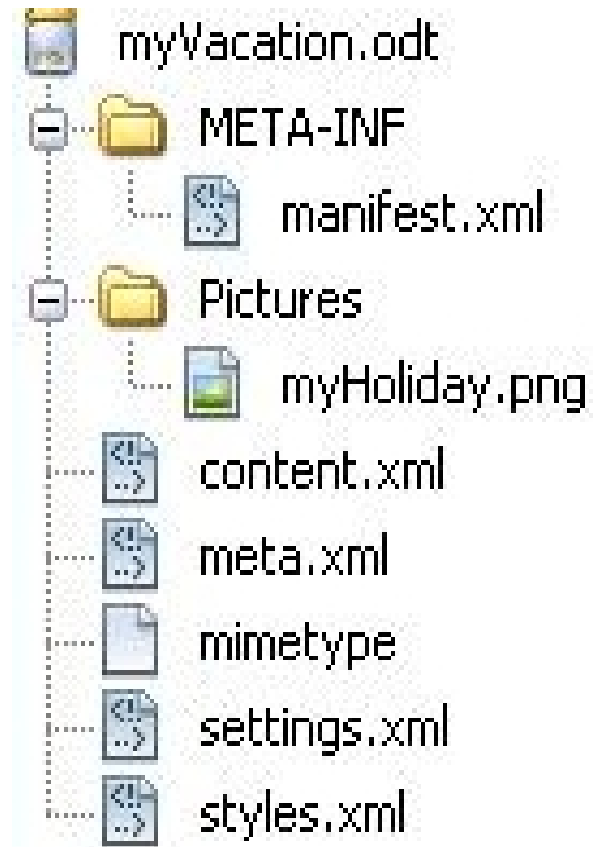
ODF Basics - Package & Files

ISO standardized default content
(as shown, but NOT Picture folder)



ODF Basics - Package & Files

Manifest as an Inventory /
'table of content'
Any user content..



ODFDOM - Layered Model

- ODFDOM featuring:
 - > Adding / removing file streams from the ODF package (ZIP)

ODF Package / Physical Layer
(managing package file streams)

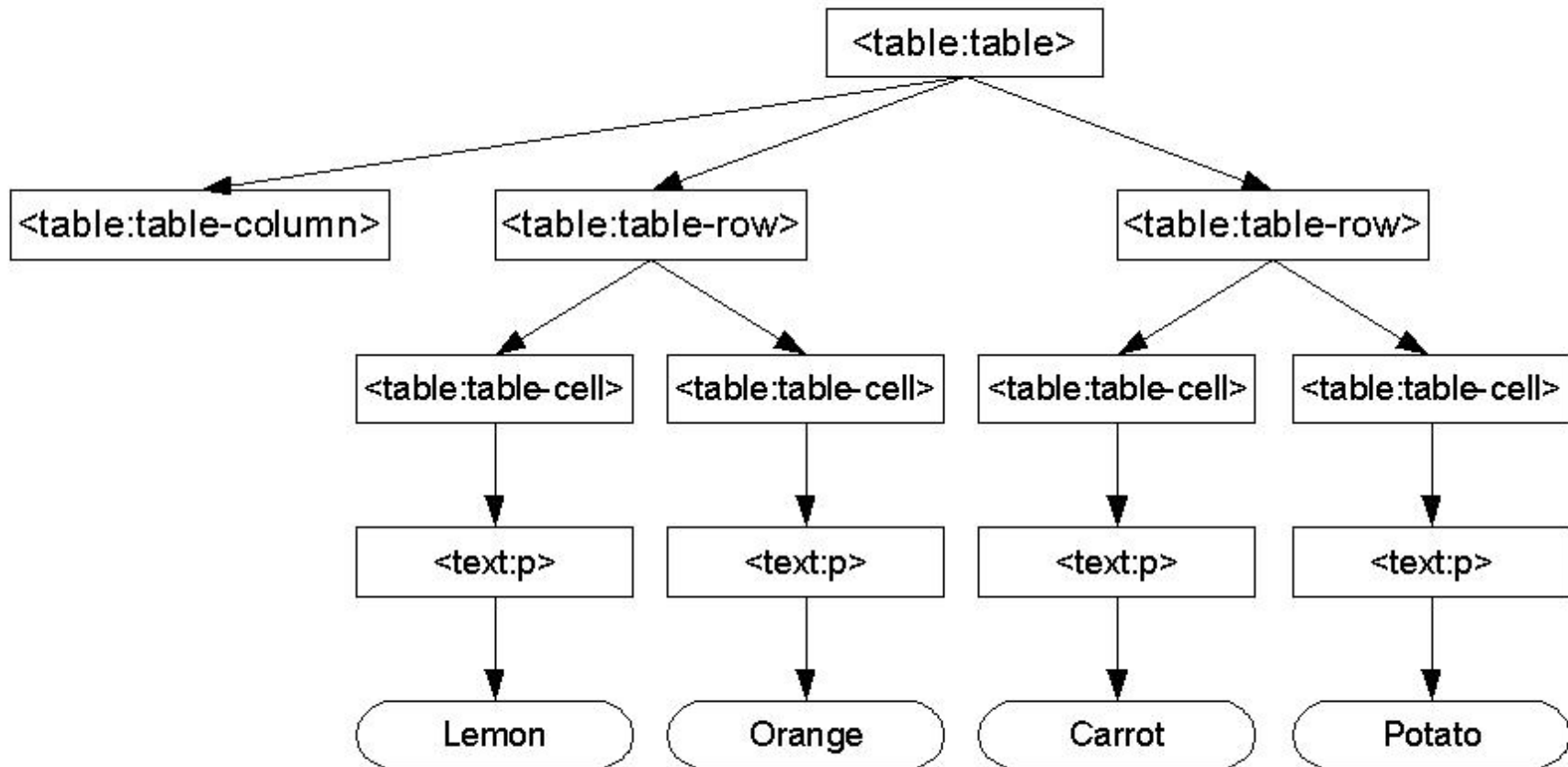
ODF Basics – XML Table Example

```

<table:table table:name="Table - fruits vs. vegies">.
  <table:table-column table:number-columns-repeated="2" />.
  <table:table-row>.
    <table:table-cell>.
      <text:p>Lemon</text:p>.
    </table:table-cell>.
    <table:table-cell>.
      <text:p>Orange</text:p>.
    </table:table-cell>.
  </table:table-row>.
  <table:table-row>.
    <table:table-cell>.
      <text:p>Carrot</text:p>.
    </table:table-cell>.
    <table:table-cell>.
      <text:p>Potato</text:p>.
    </table:table-cell>.
  </table:table-row>.
</table:table>.

```

Design Idea - DOM API for ODF



ODFDOM - Layered Model

- ODFDOM featuring:
 - > Processing ODF documents on ODF XML element level

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

ODFDOM - Layered Model

- ODFDOM featuring:
 - > Common high-level convenience functionality (e.g. add table, add table row, etc.)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

ODFDOM - Layered Model

Customized ODF Document / Extendable Layer
(optional layer not part of ODFDOM)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

ODFDOM - Layered Model

Customized ODF Document / Extendable Layer
(optional layer not part of ODFDOM)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

ODFDOM - Resources

- Quick Look
 - > Project of ODF Toolkit - <http://odftoolkit.org>
 - > Wiki - <http://odftoolkit.org/projects/odftoolkit/pages/ODFDOM>
- Deep Look (Packages)
 - > The zipped JavaDoc API
 - > The JAR of the reference Java 5 implementation
 - > The zipped NetBeans package containing the sources of ODFDOM

Why an ODF TOOLKIT Project?

- Round up ODF developer story
- Provide a site to share programs & ideas on ODF
 - > ODF SDK
 - > ODF Conformance Tools
 - > ODF Applications and Tools
- ODF SDK - ODFDOM
Provide a layered model for ODF reusable for different programming languages (ODFDOM)
- To discuss common ODF application problems...

Shared Problems of ODF Application

- How to support multiple ODF versions?
 - > ODFDOM supports by default latest ODF version
 - > Provides downgrade functionality to older ODF versions
 - > Problem: Not yet specified when a feature was added to ODF
- How test ODF application for conformance?
 - > Why not sharing ODF conformance tests and documents with other ODF applications?

Shared Problems of ODF Application

- ODF RelaxNG schema not sufficient for Code Generation
 - > Misses relation between stylable element and style:family (e.g. text:p is related to the value 'Paragraph')
 - > Creation of attribute default values from ODF 1.2 schema possible, missing example for default of mandatory attributes (As in OOo - create an image just by an image source)



Questions & Answers

Svante Schubert
Svante.Schubert@sun.com



ODF API - ODFDOM

Svante Schubert
Software Engineer
Sun Microsystems, Hamburg



Do you know ODF?

- The OASIS / ISO standard for office documents (2005/06)
- The document format of many office applications
- A zipped package of XML and related files (image, sound, user files)
- Origin from OpenOffice.org's default format

OASIS

Open Document Format for Office
Applications (OpenDocument) v1.0

OASIS Standard, 1 May 2005



International
Organization for
Standardization

ISO/IEC 26300

What to do with ODF documents?

- Usually store your Texts / Spreadsheets / Presentation
- Edited and viewing by an Office application



What to do with ODF documents?

- But how do I process 1.000.000 ODF documents?



We need an ODF API!

- API to automate ODF processing
 - > Creating, manipulating ODF documents
 - > Lightweight API
 - > API close related to ODF
 - > Opensource



Where to find an ODF API?

- Use StarOffice / OpenOffice.org API?
 - > Office API not ODF centric
StarOffice funded 1984, ODF standardized 2005 / 06
 - > Office viewing & editing functionality undesired,
Lightweight editing functionality not easy extractable



The Idea of a new ODF API!

- We need a new lightweight ODF API!!
 - > New API will focus on ODF
 - > Taking over ideas from previous ODF APIs (OOo API, AODL, ODF4J, etc.)



The Idea is to create a new library and the merging of both after some time..

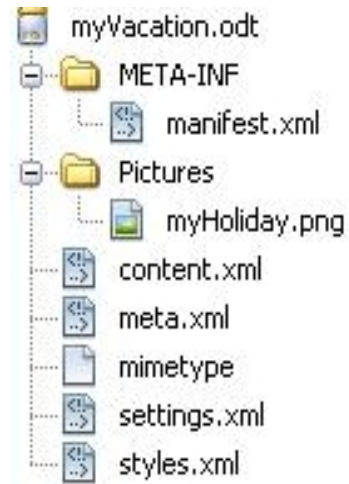
New ODF API - ODFDOM

- Sun opensourced ODFDOM
 - > Lightweight API
 - > OpenDocument centric
 - > Opensource (Apache 2)
 - > Multi-layered
 - > Java 5 reference implementation



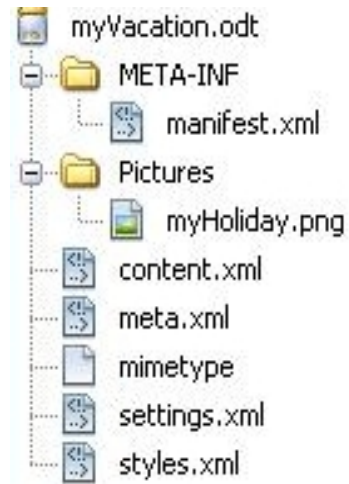
ODF Basics - Package & Files

ISO standardized default content
(as shown, but NOT Picture folder)



ODF Basics - Package & Files

Manifest as an Inventory /
'table of content'
Any user content..



ODFDOM - Layered Model

- ODFDOM featuring:
 - > Adding / removing file streams from the ODF package (ZIP)

ODF Package / Physical Layer
(managing package file streams)

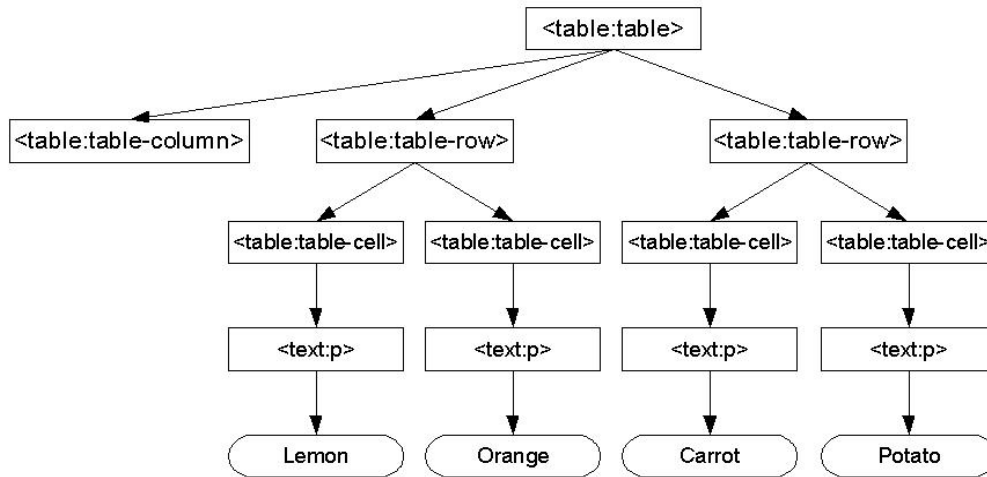
11

Generated from the ODF RelaxNG Schema (typed W3C XML DOM, similar to HTML DOM part of Apache Xerces parser)

ODF Basics – XML Table Example

```
<table:table table:name="Table - fruits vs. vegies">.
  <table:table-column table:number-columns-repeated="2"/>.
  <table:table-row>.
    <table:table-cell>.
      <text:p>Lemon</text:p>.
    </table:table-cell>.
    <table:table-cell>.
      <text:p>Orange</text:p>.
    </table:table-cell>.
  </table:table-row>.
  <table:table-row>.
    <table:table-cell>.
      <text:p>Carrot</text:p>.
    </table:table-cell>.
    <table:table-cell>.
      <text:p>Potato</text:p>.
    </table:table-cell>.
  </table:table-row>.
</table:table>.
```

Design Idea - DOM API for ODF



ODFDOM - Layered Model

- ODFDOM featuring:
 - > Processing ODF documents on ODF XML element level

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

14

Page 14

Generated from the ODF RelaxNG Schema (typed W3C XML DOM, similar to HTML DOM part of Apache Xerces parser)

ODFDOM - Layered Model

- ODFDOM featuring:
 - > Common high-level convenience functionality (e.g. add table, add table row, etc.)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

15

Page 15

Generated (at least skeleton at the first time) from the ODF RelaxNG Schema (typed W3C XML DOM, similar to HTML DOM part of Apache Xerces parser)

ODFDOM - Layered Model

Customized ODF Document / Extendable Layer
(optional layer not part of ODFDOM)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

16

ODFDOM - Layered Model

Customized ODF Document / Extendable Layer
(optional layer not part of ODFDOM)

ODF Document / Convenient Functionality Layer
(frequently used functionality)

ODF Typed DOM / XML Layer
(DOM classes generated from ODF RelaxNG)

ODF Package / Physical Layer
(managing package file streams)

17

ODFDOM - Resources

- Quick Look
 - > Project of ODF Toolkit - <http://odftoolkit.org>
 - > Wiki - <http://odftoolkit.org/projects/odftoolkit/pages/ODFDOM>
- Deep Look (Packages)
 - > The zipped JavaDoc API
 - > The JAR of the reference Java 5 implementation
 - > The zipped NetBeans package containing the sources of ODFDOM

Why an ODF TOOLKIT Project?

- Round up ODF developer story
- Provide a site to share programs & ideas on ODF
 - > ODF SDK
 - > ODF Conformance Tools
 - > ODF Applications and Tools
- ODF SDK - ODFDOM
Provide a layered model for ODF reusable for different programming languages (ODFDOM)
- To discuss common ODF application problems...

Shared Problems of ODF Application

- How to support multiple ODF versions?
 - > ODFDOM supports by default latest ODF version
 - > Provides downgrade functionality to older ODF versions
 - > Problem: Not yet specified when a feature was added to ODF
- How test ODF application for conformance?
 - > Why not sharing ODF conformance tests and documents with other ODF applications?

• Possible ways to support multiple ODF versions:

- Only support the latest ODF version?
- Every ODF version gets its own ODFDOM version?
- Every ODF version can be accessed by latest ODFDOM?
- Possible solution: annotate element and attribute accessors with an annotation of ODF version of their introduction

Shared Problems of ODF Application

- ODF RelaxNG schema not sufficient for Code Generation
 - > Misses relation between stylable element and style:family (e.g. text:p is related to the value 'Paragraph')
 - > Creation of attribute default values from ODF 1.2 schema possible, missing example for default of mandatory attributes (As in OOo - create an image just by an image source)



Questions & Answers

Svante Schubert
Svante.Schubert@sun.com