OASIS

Publishing DITA with DocBook XSL

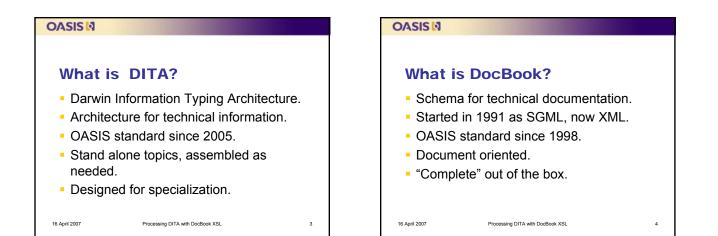Bob Stayton
Sagehill Enterprises
OASIS Symposium 2007
16 April 2007

OASIS

Who am I?
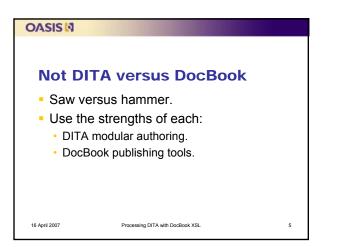
- DocBook Technical Committee member
- DITA TC observer
- Author of DocBook XSL: The Complete Guide
- DocBook consulting
- DITA consulting

OASIS

What is DITA?

- Darwin Information Typing Architecture.
- Architecture for technical information.
- OASIS standard since 2005.
- Stand alone topics, assembled as needed.
- Designed for specialization.

OASIS

What is DocBook?

- Schema for technical documentation.
- Started in 1991 as SGML, now XML.
- OASIS standard since 1998.
- Document oriented.
- "Complete" out of the box.

OASIS

Not DITA versus DocBook

- Saw versus hammer.
- Use the strengths of each:
  - DITA modular authoring.
  - DocBook publishing tools.

OASIS

DITA Open Toolkit

- Project on SourceForge since 2005.
- Reference implementation of DITA.
- Generates multiple outputs.
- bookmap XSL is "proof-of-concept"

OASIS

Basic idea

- Author content in DITA.
- Temporarily convert DITA to DocBook.
- Process the temporary DocBook XML.
- Discard the DocBook XML.

OASIS

Not …

- Not a permanent conversion.
- Not round trip conversion.
- Not lossless.

OASIS

Not lossless

- Goal is publishing
- Ignore or collapse as needed
- Let entities expand
- Nothing that can't be undone or redone

OASIS

Why

- DocBook is oriented to publishing books.
- DocBook XSL is better.
- DocBook schema is more stable.
- More complete documentation.

OASIS

Book oriented

- DITA: amorphous cloud of topics.
- DocBook: ordered hierarchical instance.
- Separate the tasks.
- Book features are automatic.

OASIS

DocBook XSL is better

- Under development since 1999.
- More mature, more bug fixes.
- More complete coverage of its schema.
- More easily customized, up to a point.

2

OASIS

Stylesheet customization

- Extensively parameterized.
- Empty placeholder templates.
- Title pages from XML specifications.

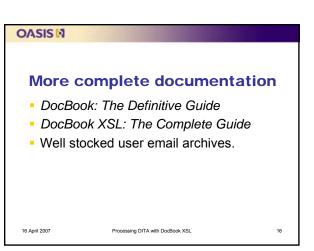**OASIS**

DocBook XSL book features

- TOC, index, List of Figures, etc.
- Extensions for XSL-FO processors:
  - PDF bookmarks
  - Page crop marks
  - Index page ranges
- Control of HTML chunking levels
- Syntax highlighting

OASIS

DocBook schema more stable

- Committee Spec: 4.5 versus 1.0
- Not undergoing major development.
- New features are incremental, not architectural.
- Strong commitment to backwards compatibilty.

OASIS

More complete documentation

- *DocBook: The Definitive Guide*
- *DocBook XSL: The Complete Guide*
- Well stocked user email archives.

OASIS

Use cases

- Publishing as book, not topics.
- When you don't have the resources to develop DITA bookmap XSL.
- When you have an existing DocBook XSL customization.
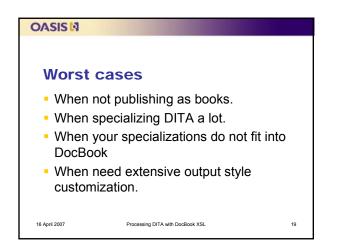- When transitioning from DocBook to DITA.

OASIS

More use cases

- When you need a DocBook feature.
- When department tools are DocBook and you want to author in DITA.
- When mixing both source types.

OASIS

Worst cases

- When not publishing as books.
- When specializing DITA a lot.
- When your specializations do not fit into DocBook
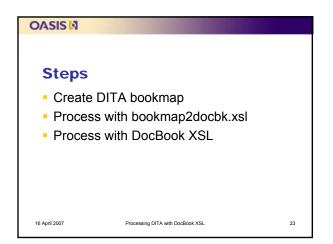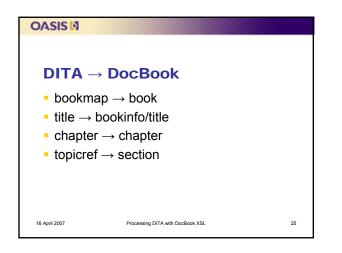- When need extensive output style customization.

OASIS

Downside

- More complicated.
- Must learn DocBook.
- May need to customize both:
  - Conversion
  - DocBook XSL
- Compare to custom DITA bookmap development.

OASIS

What you need

- DITA 1.1 DTDs
- DITA Open Toolkit from SourceForge
- DocBook XSL

OASIS

Other tools

- XSLT processor such as Saxon
  - http://saxon.sourceforge.net/
- XSL-FO processor
  - FOP
  - XEP
  - Antenna House

OASIS

Steps

- Create DITA bookmap
- Process with bookmap2docbk.xsl
- Process with DocBook XSL

OASIS

DocBook is hierarchical

- DITA is modular topics
- Map and bookmap create hierarchy
- Bookmap in DITA 1.1

OASIS

DITA → DocBook

- bookmap → book
- title → bookinfo/title
- chapter → chapter
- topicref → section

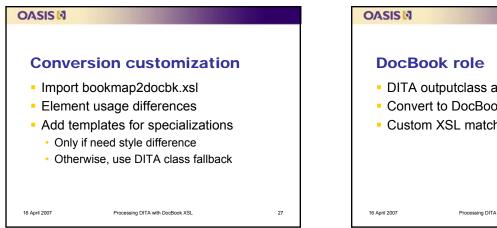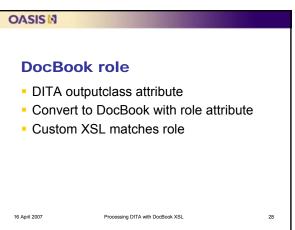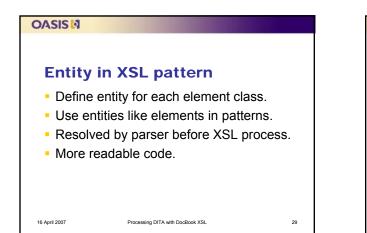**OASIS**

DITA special features

- Conrefs
- Cross references with DITA syntax
- DITA map inheritance
- Resolve these during conversion.

OASIS

Conversion customization

- Import bookmap2docbk.xsl
- Element usage differences
- Add templates for specializations
  - Only if need style difference
  - Otherwise, use DITA class fallback

OASIS

DocBook role

- DITA outputclass attribute
- Convert to DocBook with role attribute
- Custom XSL matches role

OASIS

Entity in XSL pattern

- Define entity for each element class.
- Use entities like elements in patterns.
- Resolved by parser before XSL process.
- More readable code.

OASIS

Entity example

```
<!ENTITY booktitle "*[contains(@class,
 ' bookmap/booktitle ')]">

Replace:
<xsl:template match="*[contains(@class,
 ' bookmap/booktitle ')]/*[contains(@class,
 ' bookmap/mainbooktitle ')]">

With:
<xsl:template match="&booktitle;/&mainbooktitle;">
```

OASIS

DocBook validation

- DocBook stylesheets expect valid DocBook
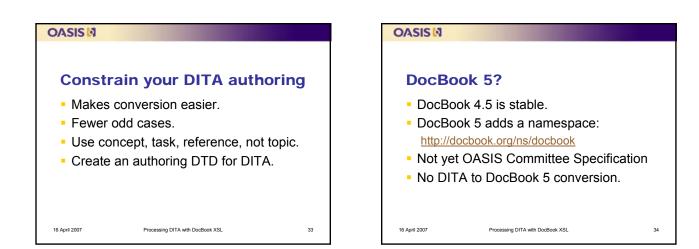- Test conversion for validity.
- Test with your content.

OASIS

DITA content models

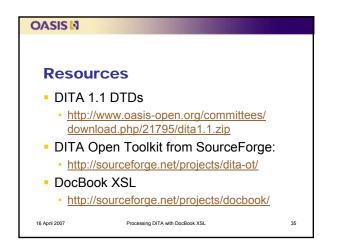- Loose models enable specialization
- May not validate in DocBook
- Conversion challenges:
  - Mixed content models
  - Mixes block and text
  - Permits odd orders (e.g., title last)
- Similar problems with DITA publishing stylesheet

OASIS

Constrain your DITA authoring

- Makes conversion easier.
- Fewer odd cases.
- Use concept, task, reference, not topic.
- Create an authoring DTD for DITA.

OASIS

DocBook 5?

- DocBook 4.5 is stable.
- DocBook 5 adds a namespace:
  http://docbook.org/ns/docbook
- Not yet OASIS Committee Specification
- No DITA to DocBook 5 conversion.

OASIS

Resources

- DITA 1.1 DTDs
  - http://www.oasis-open.org/committees/download.php/21795/dita1.1.zip
- DITA Open Toolkit from SourceForge:
  - http://sourceforge.net/projects/dita-ot/
- DocBook XSL
  - http://sourceforge.net/projects/docbook/

OASIS

More resources

- DocBook: The Definitive Guide
  http://docbook.org/tdg/en/html/docbook.html
- DocBook XSL: The Complete Guide
  http://www.sagehill.net/docbookxsl/index.html
- sagehill.net DITA
  http://www.sagehill.net/dita