Experiences with an XML topic architecture (DITA)

Don R. Day, IBM
James H. (Jamie) Roberts, IBM

WinWriters Conference
February 12, 2002
Overview

- Brief history
- Lessons learned
- Top 10 benefits
History: Darwin
Information Typing
Architecture

- Synopsis: SGML a mainstay; HTML making headway; XML a better way
- Analysis: Our information is moving towards being more topical and discoverable
- Result: A small DTD and an extensible methodology:
  - Simple: About 60 content elements; only four are required (topic, title, body, paragraph-like)
  - Extensible: A classing mechanism defines new "info-types"
  - Reliable: XSL and CSS based processing supports new vocabularies
  - Applicable: Currently being used for product support portals and application user assistance
Prototypes 2001

- Lotus: Early demo involving the migration of entire, word-processor based Notes user assistance
  www.notes.net/notesua.nsf/find/notes503xml

- IBM product support portals: major libraries (up to 1200 topics) underway

- Feedback has improved the DTD size, usefulness of the metadata, and element naming
Lessons learned & techniques applied

- **Verified Technologies:**
  - XML, XSL, CSS and other World Wide Web Consortium processing technologies
- **Specialization:**
  - class-based mapping of new vocabularies to previous vocabularies
  - common XSLT transform to provide reliable fallback processing for new DTDs based on the architecture
  - override XSLT transforms to provide specific behaviors for new vocabularies
  - specialized vocabularies to improve the contextual relevance of searches

- **Content reuse by reference**
Lessons learned...

- Verified methodologies:
  - Topics are appropriate for future information
  - Metadata and semantics are important for discoverability
  - Maps can support alternative uses of topics
  - Topic relationships can be maintained internally or externally
  - Migration of previously untyped content is possible.
Lessons learned...

- **Verified Design Process:**
  - Having a tested information architecture in place at the start was enormously valuable.
  - Having an information architecture behind authoring and production tools helps writers and their deliverables (and ultimately the user).
  - Having both prototypes and community feedback (both public and internal) was enormously useful for confirming the goals!
Top Ten benefits of authoring with DITA

10. DITA is no longer a prototype
   • Published and updated
   • Not platform-dependent (being standards based)
   • Has many potential uses (markup being separated from presentation)
   • Supports different rendering models
Top Ten benefits of authoring with DITA

9. Clarity of Markup:
   - Simple (small base set)
   - Recognizable
   - Flexible (build complex structures with simple set of tags)
Top Ten benefits of authoring with DITA

8. Supports Multiple Workflows
   - authoring environments, what you have to target for delivery
   - Adobe workflow
   - XML editors
Top Ten benefits of authoring with DITA

7. Discoverable
   • Specialized vocabularies can have meaningful element names
   • Metadata (Dublin Core, and more)
   • Accessible to search engines, topic mapping, etc.
Top Ten benefits of authoring with DITA

6. Re-usable
   - Context-free topics may be referenced as content at any level of a map
   - Topics may be nested (which usually is contrary to reuse), but...
   - Nested topics can be used independently
Top Ten benefits of authoring with DITA

5. Reliable

- Processing is based on "import" model of XSLT, CSS
- New specializations have fall-back support on prior tools
- Processors for new vocabulary can modify or rewrite the "base classes"
Top Ten benefits of authoring with DITA

4. Editable
   - Tested with most major XML editors
   - CSS-based editors support highly consistent views
   - Use metadata or not
Top Ten benefits of authoring with DITA

3. Multiple formats
   - Web pages, PDF, RTF, man pages, etc..

2. Available delivery infrastructure:
   - UA systems: HTML Help, Java Help, Web Help (by transformation to HTML or as directly viewed XML)
   - Transcoding servers: WebSphere, Apache (on-the-fly transforms)

And...
Top Ten benefits of authoring with DITA

1. You can use it with Eclipse
   - New open source platform for development
   - Includes an information system for navigating and displaying user assistance
   - Documentation takes advantage of the Eclipse plug-in framework
Sources & Additional Information

<metadata>
  • Don R. Day: dond@us.ibm.com
  • James H. (Jamie) Roberts: robertsj@ca.ibm.com
  • Introduction to the Darwin Information Typing Architecture:
    (tools and articles)
  • Eclipse: www.eclipse.org
</metadata>
Summary

- DITA is available
- Useful in a number of ways
- Supports information architecture for user assistance
- Can be exploited in the Eclipse infrastructure