

ODF Metadata and Interoperability

Rob Weir
2010-04-09

Basic question

- 1 What can we say about behavior?
- 2 Editing semantics and what do we say about metadata we don't recognize?
- 3 Best practices for interoperable metadata

Skin, but no bones

- By itself, ODF 1.2 metadata does not do anything
- It is a framework that can represent anything, but it means nothing.
 - Unless we agree on what it means
- Are there a handful of commonly used vocabularies that we should try to implement broadly?

Some hypothetical candidates

- FOAF: a person and their social network
- vCard: electronic business card
- iCalendar/RDF: events, alarms, to-dos
- ccRel: Creative Commons Rights Expression Language
- SCOVO: Statistical Core Vocabulary
- BasicGeo: geotagging

How to make these interop?

- 1 Agree on a subset of vocabularies to support
- 2 Agree how the can be used in ODF.
- 3 Document this agreement as a “profile standard” in OASIS, e.g., “A Profile of ODF 1.2 + ccRel”
- 4 Test interop at Plugfests

Foreign Metadata

- Your editor loads a document that contains metadata that you do not understand. What do you do?
 - Do you preserve it?
 - Do you allow it to be edited?
 - Do you allow it to be cut & paste?
 - Across documents?

The Hidden Constraint Problem

- Metadata can have implicit constraints that are easily broken by editors
 - “<p>When you have completed the required paper work, please submit the documentation to Roberto González for further processing</p>
 - What happens when I copy that text to another document? To a document that already has an approver? What happens if I edit the document and change the name (but not the email address, which is hidden)?

Examples of Hidden Constraints

- Cardinality. The given metadata attribute can only occur 1x, 2x, or Nx in a document.
- Referential integrity. The given metadata attribute must be consistent with another metadata attribute or with some other document content.
- Volatile metadata based on an editing timestamp, transaction ID or a random number.
- Should I be digitally signing a document that contains metadata that I cannot see?

“Safe” Metadata

- Behaves like a visual attribute: `<p>This is an example of bold text.</p>`
 - Can be cut and pasted, in total or in part
 - Can be inserted into
 - No dependency between attribute and content
 - Can be copied into another document with metadata preserved.
 - More thinking needed to define what “Safe” means.
- Go beyond this and metadata is not interoperable.
- Net effect is that interoperable metadata is quite weak. So we need to agree on some specific vocabularies.
- Semantic web doesn't have this problem, since the web, for the most part, is not editable.

Best Practices for Interoperable Metadata

- If there is a widely adopted vocabulary that already exists, adopt that rather than reinventing.
 - Work with the “network effect”, not against it.
- See if we can agree on a set of generally useful and mature vocabularies and define them via a Profile Standard
- If you must invent your own vocabularies (and if you don't, your customers surely will) then avoid hidden constraints, use “safe” metadata.