

The Ladder, Initial Thoughts

Bill and I are leaving Boston today after a good set of meetings with the DSpace folks. I'll be hitting the road shortly, but I wanted to get some thoughts out on what we're calling "The Ladder" for lack of a better term.

The concept of the ladder is this: Assuming we have Akubra, what are the next "rungs" of agreement in a shared persistence model? We played with the idea of different levels, starting with the purely structural, and moving carefully up into semantics.

At Level 0, you have Blobs with Ids. This is Akubra.

At Level 1, Relationships

At Level 2, Aggregation

At Level 3, Metadata/Semantics

I should point out that this ordering was a result of brainstorming, and we did not attempt to flesh out the details for each level. But it gave us a sort of strawman to frame the discussion.

Here's my first attempt at refinement after sleeping on it. Two key goals:

1. It is possible to persist everything in Level N + 1 into N. If this is done, rebuilding everything from level 0 is possible.
2. Though RDF might not be presumed, it is possible to represent each level in RDF (minus blob content)

Level 0

Blobs with Ids and readonly properties (size, minimal others).

Level 1

Entities with readonly + writable relationships and properties.

Entities optionally have content (represent bitstreams aka Blobs).

Level 2

Entities with everything in level 1, plus key inferred (read-only) relationships. A big requirement here is to support two-way links for whole-part relations. Might also include transitive relationships.

It's possible to infer such relationships from what's expressed in level 1 and a set of inference rules.

Level 3

Higher-order repository domain objects. The concept of a metadata entity relating to content entity would be represented here.