

BIBFRAME “House Rules”

No cake until your chores are done.

Steven Folsom, Cornell University

BIBFRAME Workshop in Europe - September 23, 2020

Steven Folsom - Coordinator, Metadata Design and Operations at Cornell University Library, heavily involved in the work of LD4L, LD4P, LD4P2, and now LD4P3, with a particular focus on an implementable strategy for BIBFRAME.

This short slide deck was adapted from a talk I gave at the beginning of LD4P2 <https://docs.google.com/presentation/d/1ly8QOMZKy-N9bGavQyKkl6XyJiaQaNZRpNIBHsFUoKA/edit?usp=sharing>, where I was projecting application profiles as one of the core activities of the grant and BIBFRAME implementations in general. Also hinted at the need for a governance model for community approved profiles.

Why is this conversation is so important...

- It's a lot of hard work to interpret an ontology or set of ontologies (including BF extensions), and decide how exactly we want to use them.
 - Multiply this by every content type we want to describe across every workflow.
- Interoperability is an obvious goal. With profiles in place tools will be in

a better position to collaboratively create and consume the data.

- That said, we still need to agree on what we should agree on.
- Profiles provision for field labels when property labels aren't enough; this impacts how we train and discuss practice.
 - If every form displays the same aspects of the model slightly differently, how to do train or communicate across platforms/institutions?

Semantic Web Layer Cake

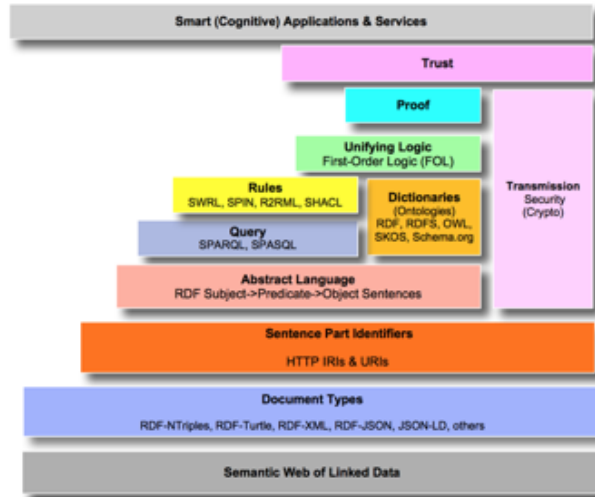


Image credit: <https://medium.com/openlink-software-blog/semantic-web-layer-cake-tweak-explained-6ba5c6ac3fab>

I tend to think of the rules/profile discussion through the lense of the larger semantic web community and related practices beyond what libraries are currently working on.

The rules we're talking about today fall under this yellow section, where there is a long history and current activity to provide constraints in applications and RDF data flows.

Profile Layer Cake

Application	The syntax the profile is captured in	Likely to change depending on the app.
	Decisions around blank nodes	Escalates into a community discussion quickly
	Which look up service are we using?	Usually up to developers, but the decision needs to be recorded
	Labels and order of fields	May need greater consideration at the Community level
Workflow	Default values for the institution	True across the board for institution
	Default values for specific workflows	True for specific workflows and/or content types
	Levels of description	Think MARC encoding levels
	Administrative metadata	Deciding how to track data flows
Institution	Preferred vocabularies	LC Authorities only? FAST, Wikidata, ISNI, local
	How we say something	E.g. bf:originDate vs. Contributions with dates, both?
	Caps on certain types of assertions	Upper limits on the quantity assertions using a particular property
	Decisions on community optionals	Shoring up the floor
Data	What are the things we're describing?	Community decided and maintained, including extensions
	What can/should we say about things?	Often considered "the floor"
	Cardinality, repeatability, datatypes	Usually pretty loose to account for practicalities
	How tied to AAPs and RDA are we?	A consideration for BF-> MARC conversions

If you zoom in on the yellow portion of the cake there are even more choices. Again, we need to decide what we'll agree to agree on.

I propose we first need to decide cross-institution floor level data targets.

These will be the foundation on which to consider institutional decisions. Within an institution we might have to support different workflows based on skill and training levels.

And lastly our platforms will shape some of the more technical decisions about how we capture and maintain workflows.

More focus should be on shareable machine actionable rules. (That green “frosting” on the last slide.)

- They become shareable configurations (assuming systems support the format and semantics)
- Increase chances of data interoperability (assuming a certain amount of agreement across implementations, little room for interpretation)
- If we start with machine actionable, human readable can* be generated.
- Save time [and frustration] when data standards are developed, but then aren't implemented in the tooling.
- [Some] control over application UI and integration with content standards
- *NOTE: the PCC is currently organizing effort around this, and hope it evolves into a transparent community-involved standard that helps with the practical implementations.*

*Results will vary depending on how the machine actionable profiles are created.

With gratitude

Steven Folsom, Cornell University Library

Email: sf433@cornell.edu

Twitter: @sf433