Encoding Patron Information in RDF

Jakob Voß (VZG)

Semantic Web in Libraries (SWIB12), November 28th, 2012
Library data in the Semantic Web

Patron information

Essential patron information

Summary

Sources
Sorry, no final ontology yet, but work in progress.
Feedback and contributions are very welcome!
Section 1

Library data in the Semantic Web
It’s been around for some time...

- Talis!
- LoC BIBFRAME (successor of MARC21) will be LOD: http://www.loc.gov/marc/transition/.
...data is being published.

- BNB (bibliography): http://bnb.data.bl.uk/
- LIBRIS (bibliography, authority): http://data.libris.kb.se/
- DNB (bibliography, authority): http://data.nature.com/
- VIAF (authority): http://viaf.org/viaf/data/
- LoC (authorities): http://id.loc.gov/
- Lobid (organizations): http://lobid.org/
- Europeana (authorities): http://data.europeana.eu/
- ...your library next (?) http://datahub.io/group/lld
What kind of library data?

- bibliographic data (title, author, date...)
- authority data (thesauri, classification, subjects...)
- organizations (to a limited degree...)

Is this really the core stuff that libraries deal with day by day?
Questions you should be nervous about

- How does LOD actually increase efficiency (to save money)?
- Does LOD model how data actually is (instead of how it should be)? In fact practical library data is quite dirty.
- What about the data that makes libraries unique:
  - Not bibliographic data
  - Data about holdings, access, buildings, opening hours...
- What about the patrons?
Section 2

Patron information
Eventually it is not that much

- user profiles
- loans and reservations
Why is patron information so neglected?

- Privacy: it is not Open Data
- Difficulties to get data out of legacy systems
- Lack of motivation (is it just boring?)
Motivation at GBV

- Access to library patron information for mobile apps and discovery interfaces
- Primarily required as API
- Alignment with RDF only as by-product to facilitate reuse and to enforce quality
- Same procedure as DAIA (API, data model & ontology)
Potential ontologies to build on

- BIBO, FRBR, RDA... (bibliographic data)
- FOAF (people)
- SIOC (online communities, access, services)
- DAIA (availability and library services)
- Organization ontology (organizations and places)
- OWL-S (discontinued Service Ontology)
- ...

Section 3

Essential patron information
Data modeling rules of thumb

- RDFS and OWL are **not** conceptual modeling languages but schema languages, such as XSD, SQL Schema etc.
- Better don’t begin with RDF at all.
- Begin with:
  - **Requirements**: what information do we need?
  - **Possibilities**: what information do we have?
- Strip down to the least common denominator
Which patron information do we care about most?

1. Personal data (name, email . . . ) : FOAF
2. Account data (state, type, expiration, fees . . . ) : ?
3. Loans and reservations : ?
2. Account data

Instances of foaf:OnlineAccount or sioc:UserAccount with:

- date of expiration (*no ontology found yet*)
- fees (*not ontology found yet*)
- account states and types
Account states and types

The general state of a patron’s account in a library.

0. active (may use most services)
1. inactive (may not use most services)
2. inactive because account expired
3. inactive because of outstanding fees

n. inactive because of …

This does not involve types of accounts (e.g. student, professor, external user etc. each as sioc:Role) because it’s difficult to find a consensus about account types among all libraries.
Many possible ontologies exist:

a) One class for each account state

   _:pa lib:hasPatronState [ a lib:PatronState ] .

b) Open world assumption with inactive as default

   lib:ActivePatron rdfs:subClassOf lib:Patron .

   _:pa a lib:Patron . # could be active
   _:pa a lib:ActivePatron . # active for sure

c) Open world assumption with active as default

   _:pa lib:isInactiveBecause ?reason .
3. Loans and reservations: What information?

Each loan or reservation combines information about

I) a library patron
II) a document held by a library
III) a current state of the loan or reservation
IV) additional properties such as:
   ▶ date issued
   ▶ number or renewals
   ▶ where to pick up the document
   ▶ . . .
II) A document held by a library

- Patron might be interested in a specific work or edition
- Most loans are about a specific copy
II) A document held by a library

- Patron might be interested in a specific work or edition
- Most loans are about a specific copy
- Problem already addressed in DAIA ontology

\[
\text{[ a bibo:Document ] daia:exemplar [ a frbr:Item ] .}
\]

- At least two URIs for each request:
  - URI of the patron originally requested
  - URI of the document the patron finally gets
III) Current document status for loan or reservation

Relation between a particular document and a particular patron:

0. no relation
1. reserved (the document is not accessible for the patron yet, but it will be)
2. ordered (the document is being made accessible for the patron)
3. held (the document is on loan by the patron)
4. provided (the document is ready to be used by the patron)
5. rejected (the document is not accessible at all)
Section 4

Summary
First result: we got an acronym!

**Patron Account Information API (PAIA)**

*Figure: Paia, Hawaii*
Second result: conceptual model with basic definitions

- Patron account states: active, inactive, inactive++
- At least two URIs for each document that is requested/loaned
- Document status:
  none, reserved, ordered, held, provided, rejected
What’s next

- Implement PAIA as API to get real world data instead of toy examples.
- Express this conceptual model in terms of RDF with existing ontologies and a new PAIA ontology, yet to be created.
Section 5

Sources
Current specification of Patron Account Information API: https://gbv.github.com/paia/

Source code of this presentation (CC-BY-SA): https://github.com/jakobib/swib2012

Images:

- Paia_beach_looking_east.jpg CC-BY-SA by Wikimedia Commons user Skier Dude.