

Schema.org, SchemaBibExtend

Schema.org - Charles

- vocabulary derived from and used by search engines
- not designed for librarians/info professionals, but for ordinary folks using the web
- “organically evolving,” not “one true vocabulary”
- public-vocabs@w3.org
- github.com/rvguha/schemaorg
- Yandex, Yahoo!, Bing, and Google are the consumers
- internationalisation - translate, multilingual examples, etc.
- JSON-LD, RDFa 1.1

SchemaBibExtend - Richard

- summer 2012 - WorldCat published schema.org for 300M bib records
- is schema.org for bibliographic data?
- Richard formed W3C community group to discuss bib extensions for schema.org
- Challenges
 - most who joined were from a rich metadata background (both a plus and minus)
 - focused on what else is needed to describe a resource (for the world, not librarian/info community)
 - FRBR - resisted it, but the language still worked its way in
 - assuming we are different/better
 - what else is going on in schema.org
- Achievements
 - Property Name Proposals
 - Citation
 - Offer
 - Work Relationships (workExample, exampleOfWork)
 - Periodicals, Articles, & Multi-volume Works (Periodical, PublicationIssue, PublicationVolume, issn, pageStart, pageEnd, pagination, hasPart, isPartof)
 - Recipes and Guidelines
 - good starting point for newbies
 - Recognition / Discussion

Schema.org in the Wild: Libraries ++

- GoodRelations e-commerce vocabulary
 - agent, object, promise, location
- Library holdings-as-schema
 - library, product, offer, availableAtOrFrom
- AbeBooks.com uses Product / Offer

- Karen Coyle has gallery of various holdings displays
- rdfa.info/tools

Extending Schema.org - Richard

- WorldCat structured data objectives
 - linking with hubs of authority on web
 - viar.org - persons
 - LOC - subjects
 - Dewey.info - classifications
 - Dbpedia - most things
 - Widely distributed and understood
 - web standard data access patterns
 - common vocabs on web
 - visibility in search engines
 - Conclusions
 - linked data - RDF, Canonical URIs
 - Schema.org - backed/recognized by major search engines, widely adopted and understood
- Phase 1
 - mine the data from MARC records
 - ID entities in data (person, organization, creativework, etc.)
 - match strings into things (viaf.org, loc subjects)
 - use of Hadoop
- Phase 2
 - model what is of interest to the Web - what will draw people to resources?
 - share the way the web does (linked data, schema.org)
- Phase 3
 - see what it looks like, note gaps
 - lobby for updates/extensions via public-vocabs list, group pages/wikis
- not everything is appropriate for schema.org
 - I'll create my own vocab, as an extension, like frosting on the schema.org cake
 - bibliograph.net - not a standalone, but an extension to add domain specific gaps; currently used by VIAF and WorldCat linked data
 - extension vocab=adds your domain specifics, but mostly schema.org
- namespace proliferation is the enemy of adoption
 - too many in data model causes others who may want to copy/use model to wonder what they can/can't use

BIBFRAME

Kevin Ford, LOC

Core bibframe classes:

- Work
- Instance
- Authority
- Annotation

RDA

- Work
- Expression
- Manifestation
- Instance

bibframe + RDA

- can be mapped pretty easily (see slide)

bibframe profile spec

- info model and reference serialization to support a means for identifying and describing ***structural constraints***
- primarily a means for an application (e.g. cataloging tool) to guide a cataloger in creation of a bibframe record
- defined by set of resources templates
- resource template further defined by array of property templates
- value constraints can be applied to properties to define or constrain permissible values

bibframe editor demo

- kitchen sink profiles yield forms with every single property possible
- simple monograph profiles
- 3windmills.com/wemi-in-bfe example

questions

- repo of bibframe profiles? kevin wonders if that will be effective - uniqueness of profiles inevitably renders them less re-usable - can technically share
- bibframe profile editor project in progress (to help keep local profiles up to date)
- karen: what are the chances libraries can get a downloadable version w/ a back end? they don't have developers!
 - kevin: connecting to storage is easy; but that leads to request to see the stored data, which requires a search interface, which leads to request to display it in a certain way, and edit if needed, which is 80% of OPAC.

Jackie Shieh at GW, their bibframe odyssey:

- environmental scan (data and staff)
- data strategies
- staff training
- enlist fellow sojourners nearby
- tech support
- current status

ILS voyager w/ 1.8 million bib records

Staff: GW Libraries - resource description group, global resource center, digital archivist; GW Law Library - cataloging and metadata services

data strategies: building block approach - single volume book record > multi-volume > multi-format

training: 2009-2014, staff required to complete training

- FRBR WEMI
- RDA TEST classroom, group exercises, self directed, weekly mtg
- BIBFRAME - biweekly meetings and study hall
- Linked Data RDF - webinars, self directed online
- BIBFRAME TestBed -
- huge leap for staff to understand triples, URIs - took ~1 year training to cover this
 - actionable URIs example of vocabulary challenge (you mean <http://?>)

enlist sojourners

- gathered colleagues at other DC based libraries to invest in Zepheira training

tech support

- only installed bibframe editor locally - they copy/paste output to shared drive for comparison

current status

- giving feedback as official implementer to LOC
- experimenting with blacklight interface

Peter Green - First 6 Mos of BIBFRAME testing at Princeton - April 2014-current

- anticipated shift from “cataloger” to “metadata librarians”
- message that bibframe is inevitable
- core group of 4 staff (head of C&MS, ER librarian, Metadata librarian, tech staff)

approach

- start w/ notification for prof catalogers
- linked data > bibframe model > bibframe game (cut records up into pieces to reassemble)
- formed bibframe testers group (specialties in rare books, music, etc.)
 - 11 testers, weekly schedule
 - initial questions
 - test what?
 - how?
 - guidelines?
 - what are others doing?
 - decide tasks from week to week
 - test marc2bibframe and submit issues
 - analyze model
 - test bibframe editor at lc
 - read / discuss bibframe forum listserv
 - read / discuss av study on bibframe.org site
 - create profiles
 - check MARC records enriched with URIs
 - create use cases
- tools for testing
 - simple group website
 - issue tracking spreadsheet
 - local web service that helps folks visualize data structure, get used to xml
 - partial marc to bibframe mapping (from marc2bibframe module)
- early experiments
 - python script to enrich MARC records with URIs
 - fuseki sparql server + snorql + pubby to html
 - LodLive free visualization tool online
 - visualRDF
 - refinder
- outcomes so far
 - reported marc2bibframe issues
 - bibframe editor feedback
 - contributions to listserv discussions
 - interesting internal discussions
 - experimentation toward pilot projects
 - initiated collaborations with library systems, digital humanities center
- challenges
 - time
 - steep learning curves
 - silence in meetings
 - core group as instructors
 - technology (esp. editor output and input)

- sense of testing in isolation
- next steps
 - pilot projects
 - latin american periodicals
 - large scale MARC enrichment
 - creating og desc data for small projects
 - visualizations
 - training
 - collaborations (anyone?)

Phil Schreur, Stanford University Libs

From Conversion to Creation, From Reaction to Reality

Linked Data for Libraries Mellon Grant for Stanford, Harvard, and Cornell, with aim to:

- create ontology
- provide collection of linked data resources
- develop/share open source software suite

Started initially looking at MARC to bibframe; now looking at other data formats to convert

Review of bibframe converter itself

Institutional isolationism - each doing different things with MARC data, so standard conversions tricky

Processes to Reroute for production to happen

- tool for creation (bibframe editor)
- mechanism to share new properties & authorities created for local needs with wider community
- authorities
- registries
- database of record (ILS + Digital collections + triplestore...ack!)
- external commitments

Using momentum/cataloger interest/experience with vague timelines to set 2015 timeframe as production goal