LibrisXL, MARC & BIBFRAME

Notes on Building a Vocabulary Driven Catalogue

niklas.lindstrom@kb.se

niklas.lindstrom@kb.se

National Library
of Sweden

New Cataloguing System

In-House, Open Source

https://github.com/libris/
librisxl
whelk-core
lxlviewer
definitions

Technical Motivation

Why?

To effectivize library data **usage**, by means of integration, sharing and reusing, through modernized foundations and tooling.

We handle a lot of searches, imports and exports (through user interfaces and APIs).

Our Process

From MARC21 to DC, BIBO, SKOS, MADSRDF, FRBR, RDA,
Schema.org, BF1, BF2 …

Common foundation: **RDF** (because it's formally grounded in
simple semantics).

In order to work with wider contexts, do we need:

A) Simpler data?
B) More details?

Digesting MARC

And a legacy of disparate details

"**bib**" turns into *data about* instances and works… ;

"**hold**" into items (and "compounds"…) ;

"**auth**" is both extended and dissolved into reusable terms, enumerations and "all other" entities…
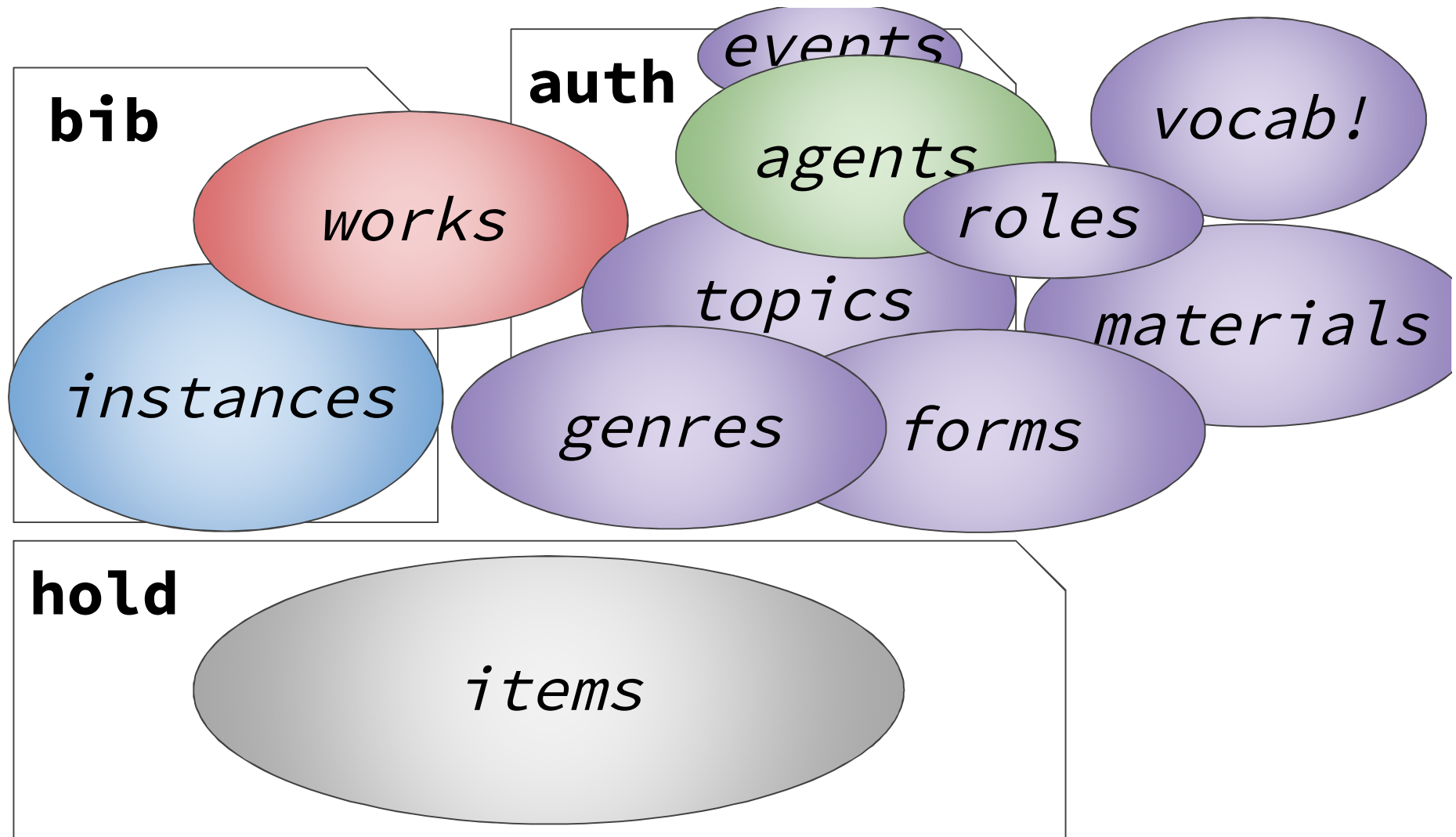
… including works.

Disseminating MARC

For now…

Our implementation ("marcframe") is driven by a **declarative** spec of **mappings**.

It works in both directions.

It's undergoing vetting as we speak.

bib
- works
- instances

auth
- events
- agents
- roles
- vocab!
- topics
- materials
- genres
- forms

hold
- items

FILTRERA

Typ

Instans (6699400)
Elektronisk (2274532)
Text (31217)
Mikroform (1630)
Ljudinspelning (211)
Taktil resurs (179)
Videoinspelning (89)
Karta (74)
Musiknoter (43)
Stillbild (25)

Verkstyp

Text (9007418)

Verksinnehållstyp

marc:Periodical (358798)
marc:Bibliography (346704)
marc:Thesis (255998)
marc:MonographicSeries (91274)
marc:Catalog (54262)
marc:Review (29819)
marc:Statistics (13651)
marc:Dictionary (9973)
marc:Encyclopedia (8429)
marc:Handbook (7018)

Verksspråk

Engelska (3273802)
Svenska (2726257)
Odefinierat språk (845226)
Tyska (767031)
Franska (399105)

SÖK

[   ]

AKTIV FILTRERING

Text ✕

Sökning på "*" gav 9007418 träffar.   Visar 200

FÖRSTA   FÖREGÅENDE   **1**   2   3   4   5   6   7   …   NÄSTA   SISTA

Instans, Monografisk resurs
Statens ungdomsråd nuläge samt överväganden och förslag
[Title saknas] Svenska swe

1987 Arbetsgruppen BUS, Jordbruksdepartementet Stockholm, 55 s., ([1], 55 s.), 21 cm

Instans, Monografisk resurs
Essays in imperial government presented to Margery Perham
[Title saknas] Engelska eng
by Kenneth Robinson and Frederick Madden [editors]

1963 Blackwell Oxford, 293 s.

Instans, Monografisk resurs
Mätmetodik för syrgasdiffusion i jord med särskild hänsyn ti...
[Title saknas] Svenska swe
Carin Tyler

1973 Lund, [2], 33 bl.

Instans, Seriell resurs
Raija, Raija
[Title saknas] Svenska swe
Bente Pedersen

1987-1994 Boknöje Helsingborg

Instans, Monografisk resurs
Die Protestantisierung der Landschaft Saanen 1555/56
[Title saknas] Tyska ger

1961 Freiburg i. S.

Instans, Monografisk resurs
Ho poietēs Giōrgos Sepherēs

id.kb.se

Our "world view", defined

Part of the catalogue (just nice IRIs and a curated start page).

Shared definitions.

Our vocabulary.

A hub for mapping notions.

In English: https://id.kb.se/doc/model

Putting Semantics To Work

Our Umbrella Vocabulary

Incorporating select linked data patterns, now mainly aligned
with BIBFRAME 2 (and DC, and Schema.org, etc.).

Application "constraints" using idioms within the vocabulary
(no formal profiles or validation).

A Grab-Bag of Vocabulary Mappings

Useful notions are reinvented. We need **explicit mappings**, and tools for **using** them. (RDFS/OWL + …)

Multiple alignments makes it *possible* to "peel off" unused specifics and shift perspectives according to shifting needs.

(Lots of technical details to discuss here!)

BIBFRAME 2?

Useful

A Pragmatic Move

Work/Instance works.

Real things, not "authorities".

provisionActivity makes sense.

Linked "enumerations" seem good overall.

Provisional?

Variously thick layers of varnish over MARC structures?

Codes instead of IRIs.

Notes instead of specific types (bf:noteType "…").

Qualification obscuring simple relations…

Administrative Metadata? Named Graph…

Principles for Differentiating?

subclasses / subproperties
"type-like"?
structured values
qualified relations
quotations?

bibo:isbn vs.
bf:identifiedBy + bf:Isbn

dc:title vs. bf:title + …

bf:Instance vs.
        bf:Media
                bf:Carrier
bf:Work vs. bf:Content

bf:GenreForm
bf:Issuance

bf:contribution
lcbf:relation

Realism – Admittedly Naive and Skewed

A "thing" for us* is defined by **our usage** of it. Not by physics. Nor by ideals.

Start with the **concrete**. Compare **variants**. Abstract only to factor out reusable, generalized **patterns**.

Prefer simple and direct. Details get mixed up, at odds with intended disambiguation.

*) humans

Really …

Everything is an **ongoing change**. Everything's **connected**.
Nothing is really completed, neither perfect, nor
everlasting.

We need to reduce data obscurity, like:

  - repetition,

  - partial overlap,

  - divergence and fragmentation.

Questions?