

National Library Platform Based On BIBFRAME

National Library of Sweden



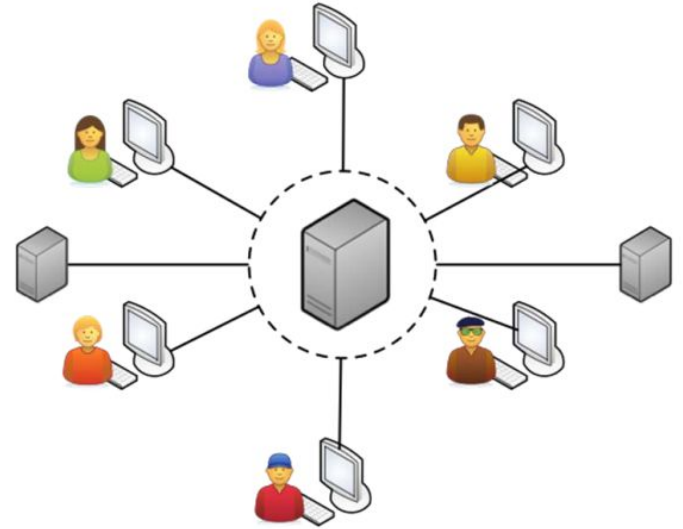
Libris: Services/Cooperation/Catalogue

- **A centralized** National Union Catalogue.
- Data is **collaboratively edited**.
- A publication is catalogued **only once**.
- Other libraries only need to register a **holding**.
- Records exported to each local library system.

Built on principles of openness and cooperation.

Includes the Swedish national bibliography and authority file.

Libris Public OPAC, ILL, etc.



XL: the core of a new generation of Libris systems

In June 2018, XL went into production, replacing the old MARC21 system with one based on Linked Data and building upon BIBFRAME.

From Ridgid/Fragile to Robust/Flexible

Currently

We only accept data conforming to our internal shape. (“Just JSON”. Still conforming JSON-LD, i.e. it is RDF. We just don’t use all the flexibility yet.)

Goal

Use our RDF vocabulary mappings to enable a richer I/O system. On the vocabulary, data granularity and identity matching levels alike.

Operational Mappings

“**KBV**” is *our application* vocabulary. It’s a bit like “BFLC”, but is an umbrella of equivalences and sub-terms-of to BIBFRAME and more (including “profile-like” choices).

It’s our goal to have it stable and put everything “odd” in a separate “marc” namespace. But “*stable*” is *highly contextual* and grounded in practise.

Our **MarcFrame** mechanism converts our data from and to (the “Libris” flavour of) MARC21.

MarcFrame and KBV Continued

To take us out of the bounds of MARC.

More intelligent handling of various shapes of RDF. (Differences in vocabulary, data granularity and identity; types vs. “type-like”.)

Normalizing and reducing the MARC exports by:

1. Phasing out redundant or costly details.
2. Extracting and disambiguating Works!
3. Aligning (e.g. types with GenreForm, Content, Media, Carrier).

What's in a Data Platform?

Data Platform

- Model (table, tree, **graph**, ...)
- **Vocabulary (-ies)** (terms/properties/classes for a (set of) domain(s))
- (Syntax)
- **Units** of Administration & Transport
- Data flows: Queries & **Federation**

We're Not There Yet (No One Is)

LDP? LD-API? Hydra? LDF/TPF?

Triples or Quads => Named Graphs

Provenance

Vocabulary interoperability

Federation (ResourceSync, LDN, WebSub...)

Services/API:s? Old, new, search, ...

(MARC I/O, OAI-PMH)

Cherry-picked, with extensions

JSON-LD documents

Improving...

In progress...

Evaluating...

LD-API (SPARQL working, but not live)

(Of course; but for legacy reasons!)

Past & Future Ecosystems

How do we avoid walking in circles?

- Z39.50 vs SPARQL?
- *MARC(21) vs. RDF (BF) is mainly about **repositioning**.*
- OAI-PMH or “just linking”? Other kinds of federation?

Data Flows in practise: dealing with deviations, misconceptions, practises stuck in a rut.

MARC vs RDF?

*The decision we are facing in terms of bibliographic data is often couched in terms of "MARC vs. RDF", however, **that is not the actual question** that underlies that decision.*

*Instead, the question should be couched as: **entities and relations, or not?** if you want to*

share *entities like works and persons, and if you want to create actual relationships between bibliographic entities, something other than MARC21 is required.*

– Karen Coyle <<http://kcoyle.blogspot.se/2017/05/two-frbrs-many-relationships.html>>

What Links Enable

Open World Assumption =

There's More To Know

No Unique Identifier = Resilience

Libris + LoC, DNB, BnF, *Wikidata*...

Requires:

- Federation & Provenance
- Shapes & Mappings

| | |
|--------------------------------------|--|
| FÖREDRAGEN BENÄMNING | Hedeby (övergiven stad) |
| BREDARE | › Geografiskt ämnesord: Övergivna städer |
| VARIANT | › Geografiskt ämnesord: Haithabu (Extinct city) › Geografiskt ämnesord: Hedeby |
| MARC:HASADDEDENTRYGE OGRAPHICNAME | › Marc:AddedEntryGeographicName: {Namnlös} |
| SAMMA SAK SOM | resource/auth/345651 |
| EXAKT MATCH [WIKIDATA] | https://www.wikidata.org/entity/Q165414 |
| ALTERNATIVNAMN [WIKIDATA] | Га́ддеби (<i>ru</i>) Слисторп (<i>ru</i>) Хэдебю (<i>ru</i>) Haithabu (<i>es</i>) |
| BESKRIVNING [WIKIDATA] | <ul style="list-style-type: none">○ ciudad de Dinamarca (<i>es</i>)○ bedeutende Siedlung dänischer Wikinger bzw. schwedischer Waräger (<i>de</i>)○ city (<i>en</i>)○ stad in Duitsland (<i>nl</i>)○ פיירד בגרמניה (<i>he</i>)○ ドイツの都市 (<i>ja</i>)○ qytet në Gjermani (<i>sq</i>) |
| KOORDINATER [WIKIDATA] | Point(9.5652777777778 54.4911111111111) |
| BILD [WIKIDATA] |  |

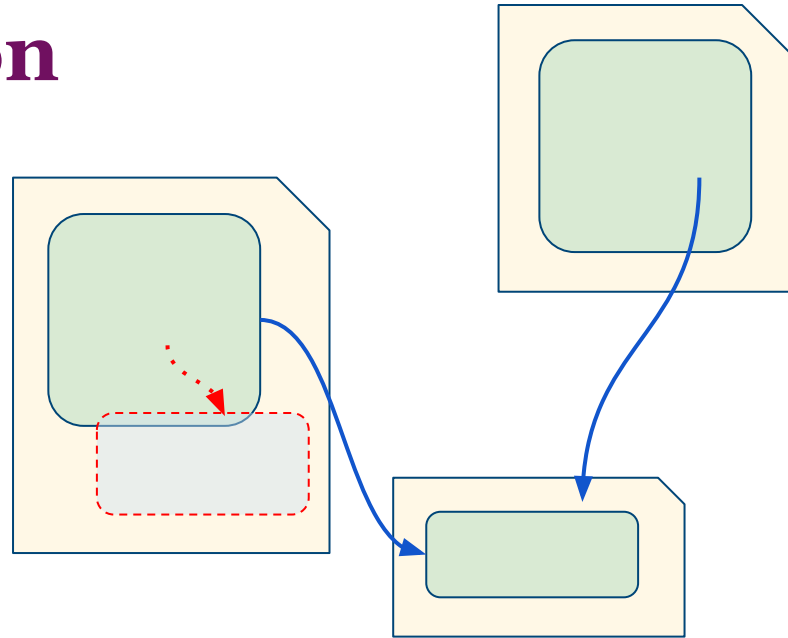
Libris

Wikidata

Data Normalization

Linking:

- ❑ Enables Reuse
- ❑ Eliminates Redundancy
- ❑ Ensures Consistency



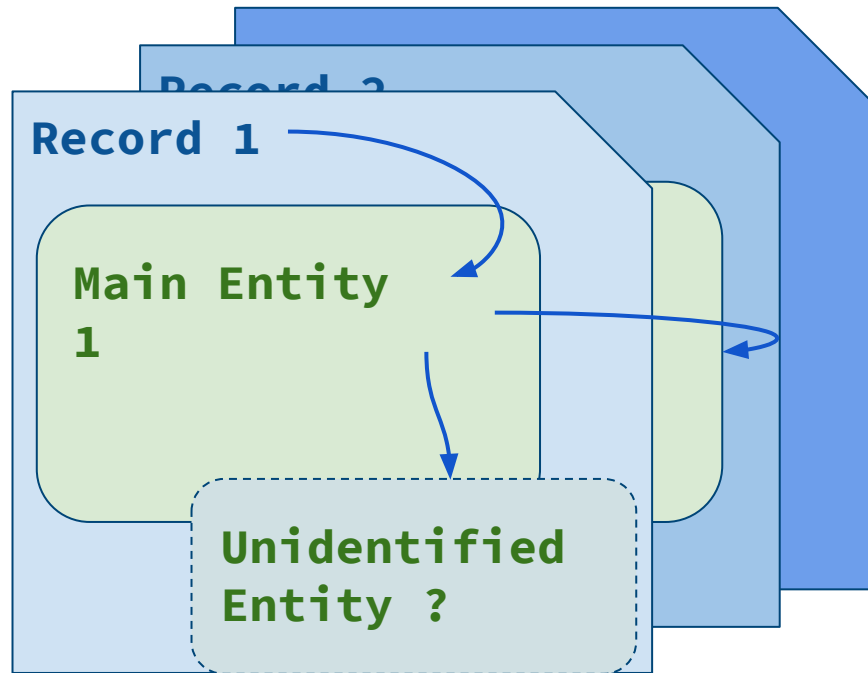
Units Of Description

Named Graphs*

One **Main Entity**

— as normalized as *possible*. (BNodes do abound here.)

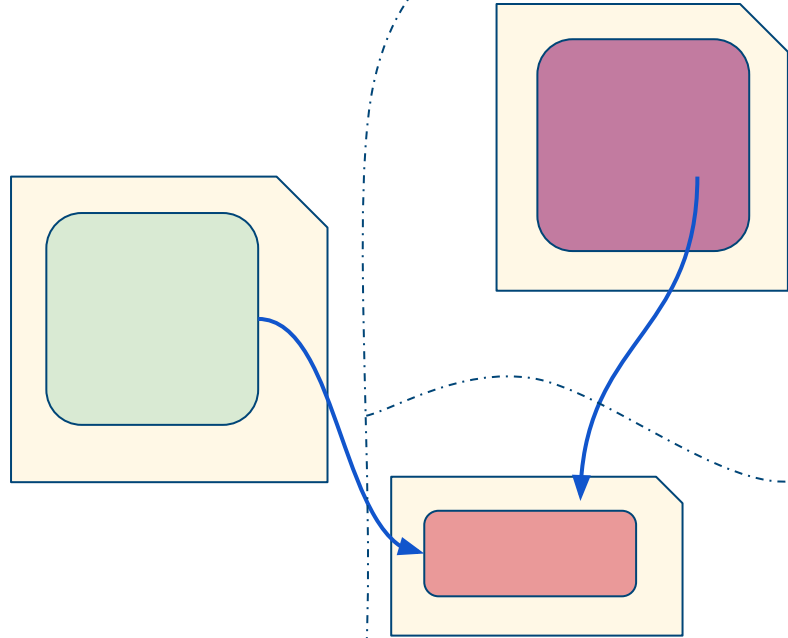
* = (We call them records, since they describe 2+n entities:
1) themselves; 2) the main entity; 3) all composite bnodes;
4) sometimes denormalized other things (to be linked).)



Federating Data

Linking outside the box:

- ❑ Requires linkable data.



Federation? Just The Web of Data?

Think about how you paste links into common online services (e.g. chats, forums, etc).

This often results in concise, summaric boxes/**cards**.

They work through structured data in the **source**, mapped into **target data shapes**, stored with provenance (source page, timestamp) and **cached**.



The image shows a screenshot of a Facebook post from 'The New York Times'. The post is titled 'When Great Minds Don't Think Alike' and is by Stephen Heyman. The main content is an illustration of two brains, one orange and one black, connected by a line. To the right of the black brain are four yellow lightning bolts. The post text reads: 'How an identical brainstorming exercise produced very different results in two distinct cultural contexts.' Below the illustration, the title 'When Great Minds Don't Think Alike' is followed by the subtitle 'How much does culture influence creative thinking?' and the author 'NYT.LMS | BY STEPHEN HEYMAN'. At the bottom of the post, there are interaction options: 'Like · Comment · Share · Send' and engagement counts: '153 6 54'.

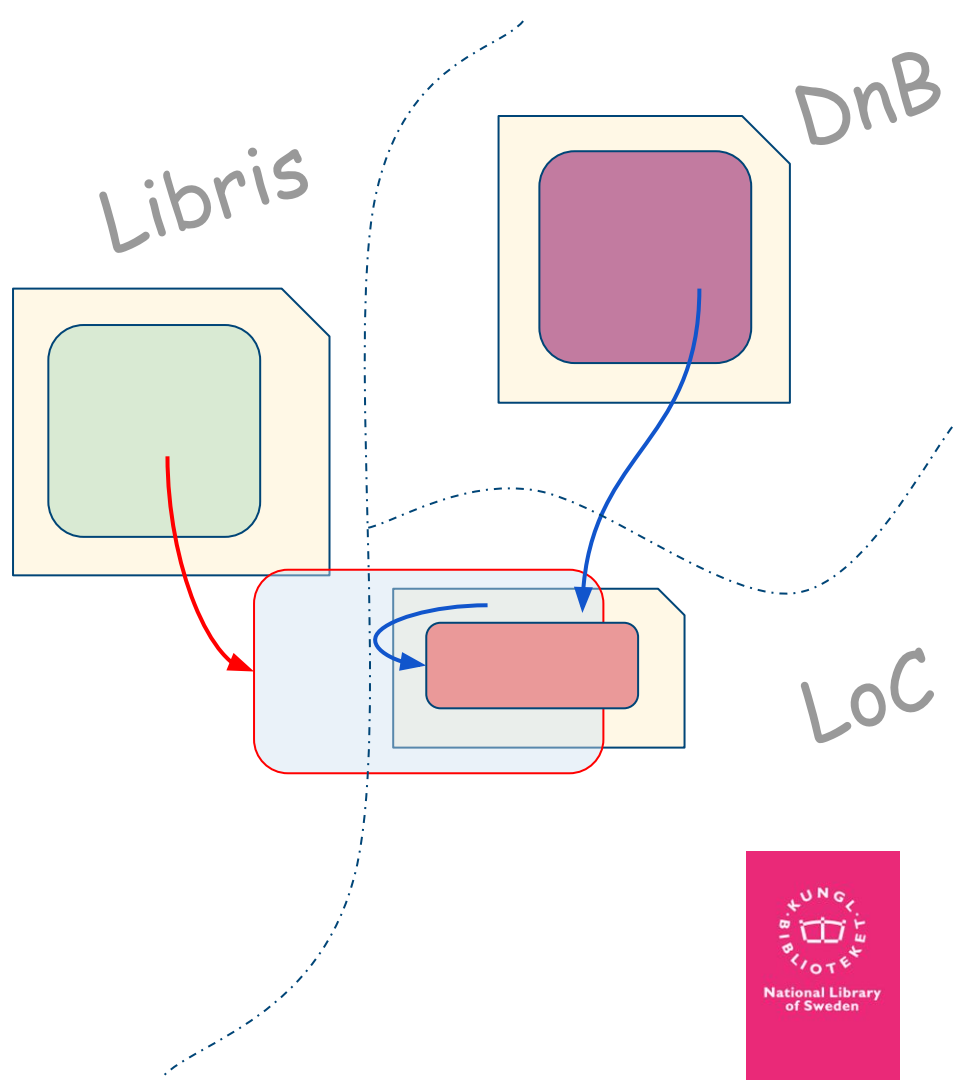
<https://developers.facebook.com/docs/sharing/webmasters/>

Data Proxies

Boxed links:

- ❑ **Cached** snapshots
 - ❑ Skewed to our views
- ❑ Indexed in our services
- ❑ Reloaded “when needed”
- ❑ Links to source data for the entity, avoiding crude aggregation*

* = a conflation of cache and workspace



Matching Identities

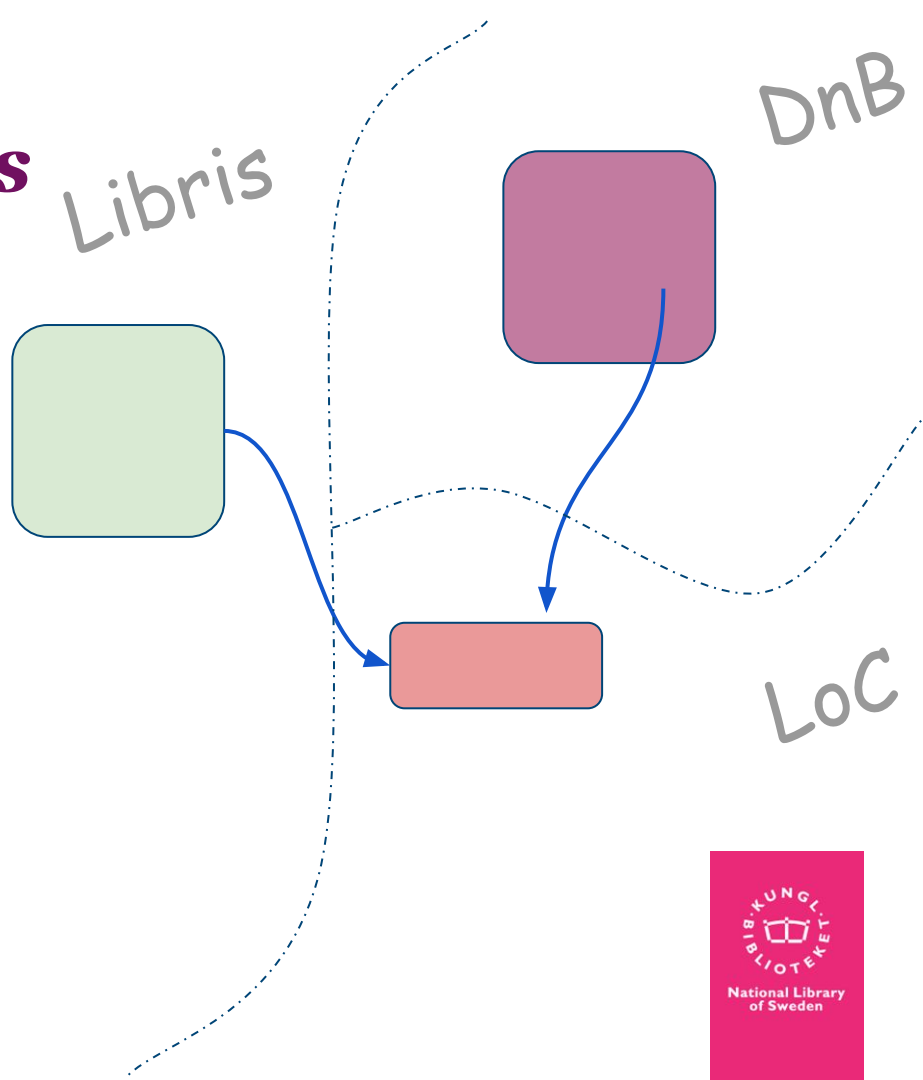
Ideally multiple matches.

Improves durability of the notion.

“Holographic lenses”.

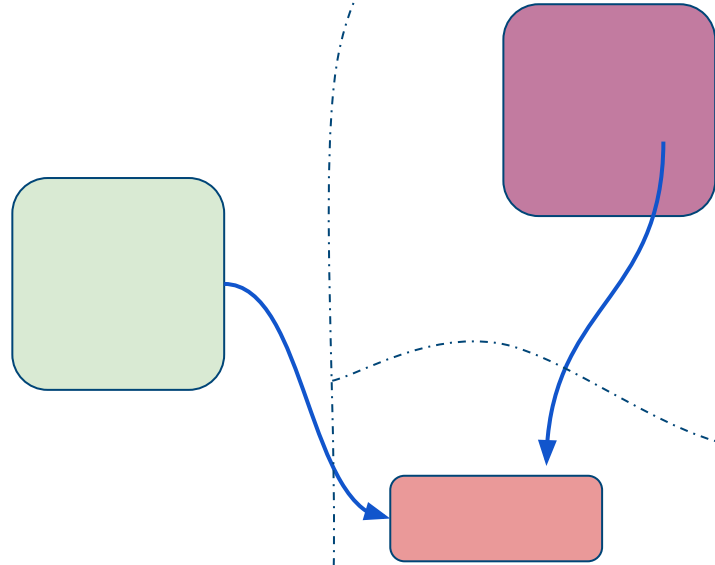
Everything rarely 404:s at once...

A poignant suggestion: no sameAs to the outside, only exactMatch...



Using Linked Data

Data *usage* requires ***comprehension***
and ***compatibility***...



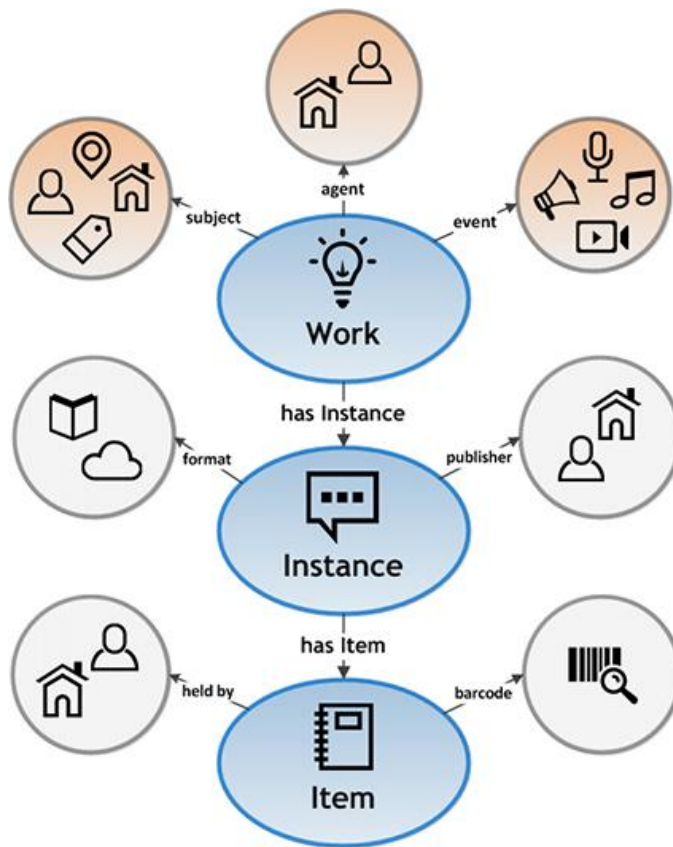
BIBFRAME 2.0

The model created to replace MARC21.

We formally decided to

align our model with

BF2 in 2017.



Not Just* BIBFRAME 2

KBV: Our local application ontology, for our specific needs.

It has a core of **BIBFRAME 2** *equivalencies* (+ some of RDA, SKOS/MADS, Schema.org where needed).

* = Not *actually*? This depends on how someone reads/views/*interprets* our data...





Descriptions are Maps

Different Data Shapes? ETL or ...?

Class & Property & Concept Alignments.

Ad hoc: SPARQL Constructs (or XSLT, or GraphQL, or ...)

Selection & Chunking: JSON-LD Framing? We use a custom implementation, plus FRESNEL...

Which Vocabulary Terms? Granularity?

Vocabulary Mappings (Crosswalks, Hubs, ...): RDFS, OWL, SKOS, (...).

Inference mightn't be the One Solution™!

Vocabulary Maps (BF → DCT / Schema.org)

```
prefix : <http://id.loc.gov/ontologies/bibframe/>

</work/a> a :Print ;
  :carrier lcarrier:nc ;
  :instanceOf </abstract/a> ;
  :identifiedBy [ a :Isbn ;
    rdf:value "12-3456-789-0" ] ;
  :provisionActivity [ a :Publication ;
    :agent </org/a> ;
    :date "2017" ] .

</abstract/a> a :Text ;
  :content :Text ;
  :title [ :mainTitle "A" ] ;
  :contribution [ :agent </person/a> ;
    :role lcrel:ill ] .
```

```
prefix : <http://purl.org/dc/terms/>

</work/a> a :BibliographicResource ;
  :format bf:Print ;
  :isFormatOf </abstract/a> ;
  :identifier "12-3456-789-0" ;
  :issued "2017" ;
  :publisher </org/a> .

</abstract/a> a :BibliographicResource ;
  :contributor </person/a> ;
  :title "A" .
```

```
prefix : <http://schema.org/>

</work/a> a :Book,
  :Product ;
  :exampleOfWork </abstract/a> ;
  :isbn "12-3456-789-0" ;
  :datePublished "2017" ;
  :publisher </org/a> .

</abstract/a> a :Book ;
  :illustrator </person/a> ;
  :name "A" .
```


Eschewing Legacy

The dependency on MARC21 is so pervasive that **everything from systems integration to cataloguing productivity has been contingent on its** abundance of varying and overlapping details.

Normalizing all this data and **adjusting dependent behaviour** by using the features of linked data is an ongoing challenge, which we've only begun to address.