



BIBFRAME in / for Hungary

Conclusions of a PhD Thesis

Miklos Hubay
Library IT Group Leader
Petőfi Literary Museum



A Little History

First informations (November, 2014)

Master's Thesis (June, 2016)

Doctoral School (2016-2020)

National Scientific Students' Associations Conference (April, 2017 - Special Award)

Presentation with **Ádám Horváth** (April, 2017 - Networkshop Conference)

National Széchényi Library's BIBFRAME Test Conversion Project (2018-2019)

University Courses (2019-)

PhD Thesis and **Petőfi Literary Museum's** Test Conversion Project (2021-2022)

Thesis Defence (late 2022?)



Table of Contents

- Introduction
- Age of Markup Languages From SGML to XML
- Entities in LIS
 - *Historical Approach and FRBRization Techniques*
- Appearance of RDA
- Basics of Linked Data
 - *RDF and RDFS, OWL, Vocabularies, VESs, Data Reconciliation, FRESNEL, SHACL, ShEx, etc.*
- Linked Data in Libraries
 - *BIBFLOW, Share-VDE, BLUECloud, Library.Link, ALIADA, Innovative Inspire, Stanford Blacklight*
- SEO with Linked Data
- *BIBFRAME*

Table of Contents - BIBFRAME

- Model 1.0 and Model 2.0
- Extensions (**bibfra.me, ARM, bibliotek-o, PMO, bflc**)
- Interoperability (**RDA Element Sets, Europeana, schema.org**)
- Conversion Tools
- Cataloging Platforms
 - **BFE+BFPE (now MARVA), Sinopia, LIBRIS XL, RERO, etc.**
- **NSZL Project Results and Consequences**
- **PLM Data Conversion Project**
 - **Analysis, Critical Approach and Consequences**
- Summary and Evaluation

*there are much more:
Scribe, CEDAR, Metadata
Maker, etc.*

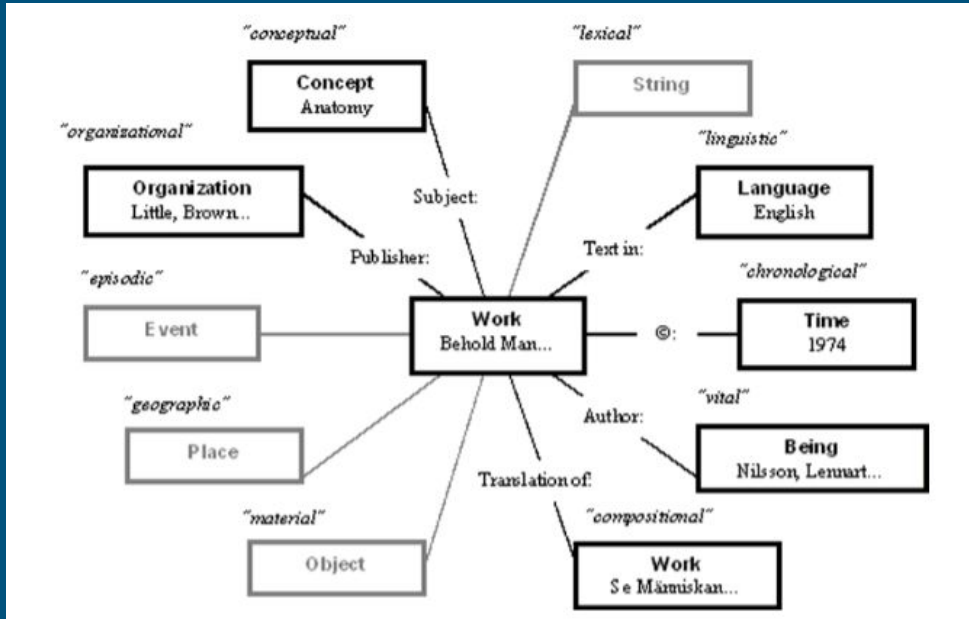
A Snippet From the Past: XOBIS



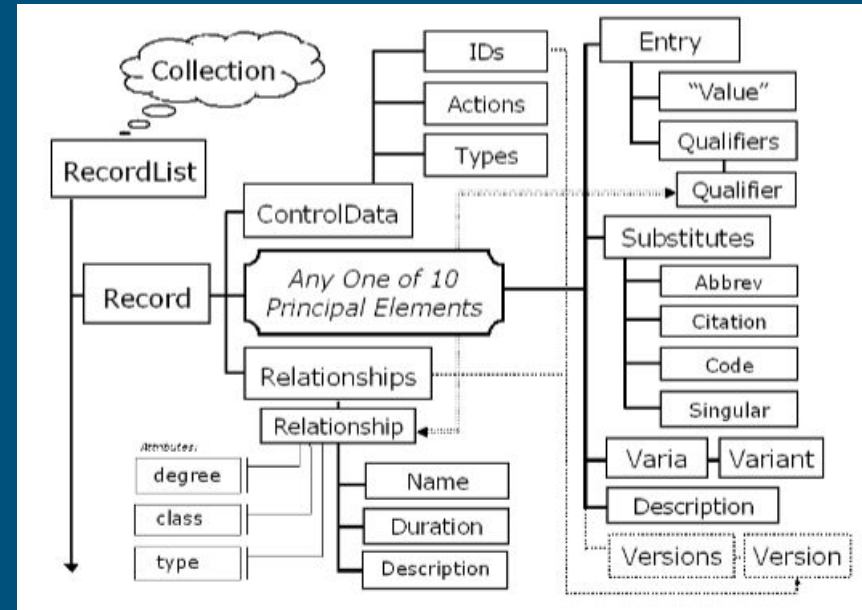
A Forgotten Model

- one of the most interesting and most forward-looking implementations of XML in LIS
- created in 2002(!) - emphasizes **connections** four years before Linked Data
- *“XOBIS is an XML schema which reorganizes bibliographic and authority data elements into a single, integrated structure. It explores balancing valuable traditions with new technologies to create a potential foundation for future access to information in a distributed digital environment.*
- *„a middle path between the complexity of MARC and the oversimplification of the Dublin Core.”*

Principal Elements and Connections



10 Principal Element
100 Connection Type



Basic Metadata
Structure



Key-title The Serials librarian



Resource information



Title proper: The Serials librarian.

Other variant title: SL

Country: United States

Medium: Print

BIBFRAME Prefix on ISSN Portal

```
<j.0:status rdf:resource="http://issn.org/vocabularies/RecordStatus#Register"/>
<j.1:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">20220319163000.0</j.1:modified>
▼<j.2:wasAttributedTo>
  <j.3:Organization rdf:about="http://issn.org/organization/ISSNCenter#_1"/>
  </j.2:wasAttributedTo>
▼<j.3:mainEntity>
  ▼<j.0:Instance rdf:about="http://issn.org/resource/ISSN/0361-526X">
    <j.3:url>http://firstsearch.oclc.org</j.3:url>
    <dc:identifier>0361-526X</dc:identifier>
    <j.3:url>http://www.tandfonline.com/loi/wser20</j.3:url>
    <dc:format rdf:resource="http://issn.org/vocabularies/medium#Print"/>
    <j.1:isFormatOf rdf:resource="http://issn.org/resource/ISSN/1541-1095"/>
    <j.1:type rdf:resource="http://marc21rdf.info/terms/formofmaterial#a"/>
  ▼<j.0:identifiedBy>
    ▼<j.0:IssnL rdf:about="http://issn.org/resource/ISSN/0361-526X#ISSN-L">
      <j.0:status rdf:resource="http://issn.org/vocabularies/IdentifierStatus#Valid"/>
      <rdf:value>0361-526X</rdf:value>
    </j.0:IssnL>
  </j.0:identifiedBy>
  <j.3:issn>0361-526X</j.3:issn>
  ▼<j.0:identifiedBy>
    ▼<j.0:Identifier rdf:about="http://issn.org/resource/ISSN/0361-526X#KeyTitle">
      <rdf:value>The Serials librarian</rdf:value>
      <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/KeyTitle"/>
    </j.0:Identifier>
  </j.0:identifiedBy>
```

```
▼<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:j.0="http://id.loc.gov/ontologies/bibframe/" xmlns:j.1="http://purl.org/dc/terms/"
  xmlns:j.2="http://www.w3.org/ns/prov#" xmlns:j.3="http://schema.org/" xmlns:j.4="http://purl.org/ontology/bibo/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
```

HUNMARC and BIBFRAME

Some Previous Hungarian LD-projects

National Széchényi Library - LD publication project (2010)

http://nektar.oszk.hu/wiki/Semantic_web

- RDFDC for bibliographic data
- FOAF for name authority entries
- SKOS for subject terms and geographical names

NSZL-Hungarian Electronic Library

NSZL-Digital Archive of Pictures

<https://mek.oszk.hu/html/lod.html>

<https://dka.oszk.hu/html/export.html>

- DCTerms, BIBO, FOAF, ORE

MOKKA-UP

- experimental BIBFRAME data set from the Hungarian Common Catalogue with Share-VDE front-end

- <http://www.w3.org/2000/01/rdf-schema>
- <http://www.w3.org/1999/02/22-rdf-syntax-ns>
- <http://purl.org/dc/terms/>
- <http://purl.org/ontology/bibo/>
- <http://www.w3.org/2002/07/owl>
- <http://www.openarchives.org/ore/terms/>
- <http://xmlns.com/foaf/0.1/>
- <http://www.sparontologies.net/mediatype/>
- <http://viaf.org/viaf/>
- <https://creativecommons.org/licenses/>
- <http://id.loc.gov/authorities/subjects/>
- <http://id.loc.gov/vocabulary/iso639-2/>
- <http://nektar.oszk.hu/data/auth/>

Some Previous Hungarian LD-projects

Hungarian Museum of Fine Arts (2015) - with ALIADA

http://archiv.magyarmuzeumok.hu/muhely/2337_potty_lesz_a_szepmuveszetibol (only in Hungarian!)

- FRBRoo, WGS84, SKOS
- MARC, DC and LIDO as input

Hungarian National Museum (2017-2018)

<https://mnm.hu/hu/cikk/potty-lett-magyar-nemzeti-muzeum> (only in Hungarian!)

- BIBFRAME → enrichment with BNF, ISNI, LCNAF, VIAF, Wikidata
- Virtuoso and Monguz OPAC3 integration

Joining Forces



Magyar Nemzeti Múzeum Központi Könyvtár

< Vissza a találati oldalhoz

< 1./13 >

Teljes

Részletes

MARC

MARCXML

Cédula

RIS

 Letöltés

Cím: Sirály

Személynév: [Márai Sándor \(1900-1989\)](#) 

Kiadás helye: [\[Budapest\]](#)

Kiadó neve: [Helikon](#)

Megj. éve: 2003

Nyomda/Gyártó: Szombathely : Sylvester

Terjedelem, fizikai jellemzők: 272, [2] p ; 19 cm

Megjegyzés: Kisregények

data.bnf.fr

id.loc.gov

viaf.org

www.wikidata.org

isni.org

Problem

*There were no publications (and still there aren't)...
for librarianship...
with critical approach...
about the details of metadata or conversion processes...
and data quality!*


Conversion Project in Petőfi Literary Museum

- **CHO data in HUNMARC(-based) format** with relevant data fields and subfields
 - can also be expressed as LIDO
 - unique content in standardized fields
 - new data fields for museums, not just in 9XX

*~ 22.000 bibliographic records from the library of the museum →
Relevant consequences for the whole Hungarian library landscape*

God Bless the MARCEdit!


bf: BibFrame TestBed




The BibFrame2 Testbed is a simple to use utility designed to allow metadata specialists the ability to model their data using BibFrame Concepts.

Process

Close

Source File: 

Save File: 

File Type:

BaseURI:

Profile:

Translation Rules: [Edit](#)

Data Serialization:

Id Field:

[Convert BIBFrame 2 MARC](#)

What is in the
RDF/XML?

What is **NOT** in
the RDF/XML?

Translations

```
<bf:noteType>Physical details</bf:noteType>
<bf:noteType>Accompanying materials</bf:noteType>
<bf:noteType>Numbering</bf:noteType>
<bf:noteType>description source</bf:noteType>
<bf:noteType>with</bf:noteType>
<bf:noteType>report type</bf:noteType>
<bf:noteType>issuance information</bf:noteType>
<bf:noteType>additional physical form</bf:noteType>
<bf:noteType>funding information</bf:noteType>
<bf:noteType>issuing body</bf:noteType>
<bf:noteType>additional codes</bf:noteType>
<bf:noteType>Computer file
characteristics</bf:noteType>
```

```
<rdfs:label>invalid</rdfs:label>
<rdfs:label>incorrect</rdfs:label>
<rdfs:label>canceled</rdfs:label>
<rdfs:label>linear horizontal</rdfs:label>
<rdfs:label>linear vertical</rdfs:label>
<rdfs:label>Includes translation</rdfs:label>
<rdfs:label>former frequency</rdfs:label>
```

```
<bf:part>summary</bf:part>
<bf:part>sung or spoken text</bf:part>
<bf:part>libretto</bf:part>
<bf:part>table of contents</bf:part>
<bf:part>original</bf:part>
```

The URI Problem in Entry Fields

No subfields for URIs (\$1 and \$0)

100/110/111/130

440

600/610/611/630

700/710/711/730

`<bf:Agent rdf:about="[...]PIM2255316#Agent100-8">`



`<bf:Agent rdf:about="https://resolver.pim.hu/auth/PIM94986">`

Person Name Subfields

```
<marc:datafield tag="100" ind1="1" ind2="0">
<marc:subfield code="a">Mitrovics</marc:subfield>
<marc:subfield code="j">Gyula</marc:subfield>
<marc:subfield code="m">ifj</marc:subfield>
<marc:subfield code="d">1871-1965</marc:subfield>
</marc:datafield>
```

RDF/XML:

```
<bflc:name00MatchKey>Mitrovics Gyula 1871-1965</bflc:name00MatchKey>
<bflc:primaryContributorName00MatchKey>Mitrovics Gyula 1871-
1965</bflc:primaryContributorName00MatchKey>
<bflc:name00MarcKey>10010$aMitrovics$jGyula$mifj$d1871-1965</bflc:name00MarcKey>
<rdfs:label>Mitrovics Gyula 1871-1965</rdfs:label>
```

Do we need bflc relations at all?

MARC/XML 2:

```
<marc:datafield tag="100" ind1="1" ind2="0">
<marc:subfield code="a">Krúdy</marc:subfield>
<marc:subfield code="j">Gyula</marc:subfield>
<marc:subfield code="d">1878-1933</marc:subfield>
<marc:subfield code="c">nemes</marc:subfield>
<marc:subfield code="g">széchenykovácsi</marc:subfield>
</marc:datafield>
```

RDF/XML 2:

```
<bflc:name00MatchKey>Krúdy Gyula 1878-1933 nemes</bflc:name00MatchKey>
<bflc:primaryContributorName00MatchKey>Krúdy Gyula 1878-1933
nemes</bflc:primaryContributorName00MatchKey>
<bflc:name00MarcKey>10010$aKrúdy$jGyula$d1878-
1933$cnemes$gszéchenykovácsi</bflc:name00MarcKey>
<rdfs:label>Krúdy Gyula 1878-1933 nemes</rdfs:label>
```

Place Conversion Proposal in X10, X11

710 21 \$a Metropolitan Museum of Art
\$c New York, N.Y
\$4 rend., közread.

```
<bf:contribution>
  <bf:Contribution>
    <bf:agent>
      <bf:Agent rdf:about="http://hubaymiklos.hu/test/MNMKVT67445#Agent710-11">
        <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Organization" />
        <bflc:name10MatchKey>Metropolitan Museum of Art New York, N.Y</bflc:name10MatchKey>
        <bflc:name10MarcKey>71021$aMetropolitan Museum of Art$cNew York, N.Y$4rend.,
közread</bflc:name10MarcKey>
        <rdfs:label>Metropolitan Museum of Art New York, N.Y</rdfs:label>
      </bf:Agent>
    </bf:agent>
    <bf:role>
      <bf:Role>
        <bf:code>rend., közread</bf:code>
      </bf:Role>
    </bf:role>
  </bf:Contribution>
</bf:contribution>
```

Series

440 - Data on Work and Instance level - **not defined in MARC 21**

490 - Data on Instance level

761 - Series of / Part of series linking field - **not defined in MARC 21**

800-810-811-830 - **not defined in HUNMARC**

```
<bf:hasSeries>
  <bf:Work rdf:about="[...]PIM2058639#Work440-12">
    <rdfs:label>Európa zsebkönyvek</rdfs:label>
    <bf:title>
      <bf:Title>
        <bflc:title40MarcKey>440 0$aEurópa
zsebkönyvek$x0324-2722$v286</bflc:title40MarcKey>
        <bf:mainTitle>Európa
zsebkönyvek</bf:mainTitle>
      </bf:Title>
    </bf:title>
    <bf:identifiedBy>
      <bf:Issn>
        <rdf:value>0324-2722</rdf:value>
      </bf:Issn>
    </bf:identifiedBy>
  </bf:Work>
</bf:hasSeries>
```

```
<bf:hasSeries>
  <bf:Work rdf:about="[...]PIM2173011#Work440-12">
    <rdfs:label>Európa zsebkönyvek</rdfs:label>
    <bf:title>
      <bf:Title>
        <bflc:title40MarcKey>440 0$aEurópa
zsebkönyvek$x0324-2722$v265</bflc:title40MarcKey>
        <bf:mainTitle>Európa
zsebkönyvek</bf:mainTitle>
      </bf:Title>
    </bf:title>
    <bf:identifiedBy>
      <bf:Issn>
        <rdf:value>0324-2722</rdf:value>
      </bf:Issn>
    </bf:identifiedBy>
  </bf:Work>
</bf:hasSeries>
```

Series Data as Authority Record

- Sorozat (130) -

Sorozatcím (130)	Series Statement	
Alcím	Subtitle	<input type="button" value="+"/> <input type="button" value="-"/>
Ágazati megjelölés	Number of part/section of a work	<input type="button" value="+"/> <input type="button" value="-"/>
Ágazati cím	Name of part/section of a work	<input type="button" value="+"/> <input type="button" value="-"/>
Szerzőségi közlés	Statement of responsibility	
ISSN		
Kiadó	Publisher	
Keletkezés éve	Year	
Megjegyzés	Notes	

- Sorozat - lásd utaló (430) - Series Statement - See also -

Sorozatcím	Series Statement	
Ágazati megjelölés	Number of part/section of a work	<input type="button" value="+"/> <input type="button" value="-"/>
Ágazati cím	Name of part/section of a work	<input type="button" value="+"/> <input type="button" value="-"/>
Utalás jellege	---- Relation Type	<input type="button" value="+"/> <input type="button" value="-"/>

Title Types

HUNMARC: defined by distinct **data fields** (741-749)

MARC 21: defined by **ind2** in field 246

Second Indicator

Type of title

- No type specified

0 - Portion of title

1 - Parallel title

2 - Distinctive title

3 - Other title

4 - Cover title

5 - Added title page title

6 - Caption title

7 - Running title

8 - Spine title

245\$c - 245\$e

245 10 a Apám meséi
c Rostás-Farkas György, Karsai Ervin
e [előszót írta Czövek Judit]
e [graf. F. Tóth Zsuzsa]

```
<rdfs:label>Apám meséi</rdfs:label>
<bf:title>
  <bf:Title>
    <rdfs:label>Apám meséi</rdfs:label>
    <bflc:titleSortKey>Apám meséi</bflc:titleSortKey>
    <bf:mainTitle>Apám meséi</bf:mainTitle>
  </bf:Title>
</bf:title>
<bf:responsibilityStatement>Rostás-Farkas György, Karsai Ervin</bf:responsibilityStatement>
```

UDC

080 - UNIVERSAL DECIMAL CLASSIFICATION NUMBER (R)	nac
First - Type of edition	
# - No information provided	
0 - Full	
1 - Abridged	
Subfield Codes	
\$a - Universal Decimal Classification number (NR)	
\$b - Item number (NR)	
\$x - Common auxiliary subdivision (R)	
\$0 - Control or standard number	
\$1 - Real World Object	
\$2 - Edition identifier (NR)	
\$6 - Linkage (NR)	

Class:	ClassificationUdc
URI:	http://id.loc.gov/ontologies/bibframe/ClassificationUdc
Label:	UDC Classification
Definition:	Universal Decimal Classification number used for subject access.
SubClass Of:	Classification
Change Notes:	2016-04-21: New

Note Fields - 5XX

505 - HUNMARC: Note on Content (rather includes general notes on content, not TOC)

MARC 21: Formatted Contents Note

BIBFRAME: <bf:tableOfContents>

588 - HUNMARC: Note on Extent and Physical Details

MARC 21: Source of Description, Etc. Note

BIBFRAME: <bf:noteType>description source</bf:noteType>

No relations for HUNMARC fields:

- 503 - Note on (Bibliographical) History
- 535 - Note on Location of Item
- 562 - Note on Uniqueness of Version/Item
- 565 - Note on File Details
- 567 - Note on Methods (Detailed in Work)
- 584 - Note on Usage (growth/year, ref. questions, etc.)
- 590 - Note on Details of Bibliographic Work
- 591 - Note on Copy Number
- 592 - Note on Offprint
- 593 - Note on Abstract in Foreign Language
- 594 - Note on Language of Table on Contents
- 596 - Note on Edition
- 597 - Note on Publication
- 598 - Note on Serial
- 599 - Local / Internal Note

Four years later...

Is the librarianship strong enough to make
BIBFRAME
well-known?



Do we have to use parallel conversion
methods?

Will history repeat itself?

Are these data really **LOUD**?

© Rob Sanderson, 2018

What is the critical data
quantity in BIBFRAME?

Thank You!

Miklos **HUBAY**

Metadata Librarian, Library IT Group Leader

Petőfi Literary Museum

hubay.miklos@pim.hu