



BIBFRAME Interoperability Group (BIG): Update

Ian Bigelow (University of Alberta)
Nancy Lorimer (Stanford University)

19 September 2023
BIBFRAME Workshop in Europe



Outline



- Background
- Terms of Reference
- Initial accomplishments
- 2023 Work Plan
 - Current Work
 - Validation: SHACL and DCTap
 - Defining BIBFRAME Interlingua
 - Levels of adherence
 - First pass for Work (level 1 - 2)
 - Other issues under discussion
- Next steps:
 - Getting to early outputs
 - Testing within BIG and through consultation
- Q&A

Origin of the BIBFRAME Interoperability Group



Result of the BIBFRAME Data Exchange Meeting, a virtual meeting organized by the PCC, September 9-10, 2021, to discuss exchange of BIBFRAME data between systems and implementations

Attendees represented national libraries, PCC committees, LD4 community, vendor community, European BIBFRAME Group, and other interested parties

Major challenge identified:

Interchange of BIBFRAME data caused by different choices in expressing the BIBFRAME ontology in original data creation and different results from data conversion from MARC

International Group focused on interoperable BIBFRAME data approved by PCC Policy Committee (PoCo): January 2022

More information available on the BIG Wiki:

<https://wiki.lyrasis.org/pages/viewpage.action?pageId=249135298>

Membership



Membership is institution-based

One member and one alternate per institution of the following:

- Standards bodies
- Libraries that have implemented BIBFRAME (or are actively working towards implementation)
- BIBFRAME data hosting organizations

Current Chairs:

- Current - Ian Bigelow (University of Alberta Library)
- Outgoing - Melanie Wacker (PCC)
- Incoming - Xiaoli Li (UC Davis)

Current membership



National Library of Sweden



Consultants



As BIG reaches a point where there is more to share, we will be involving consultant organizations to review work and hopefully assist with testing. While not a comprehensive list, so far this includes:

- EBSCO
- Ex Libris
- MODS Editorial Committee
- National Library of Medicine
- NISO
- University of Illinois at Chicago Library

Terms of Reference and Charge



Work collaboratively on the development and maintenance of interoperable BIBFRAME data guidelines

- to support production level implementation
- to address issues restricting interoperability, and
- to inform development of associated toolings and infrastructure.

BIG is not responsible for further development of the BIBFRAME ontology itself. While members may use open and/or proprietary tools to support BIBFRAME data creation and exchange locally, this group is primarily focused on interoperability for unrestricted metadata reuse.

Accomplishments



From the start of work in July 2022 through mid-2023:

- Incorporated work done by several other working groups, such as the Strawperson Working Group, Communication Working Group, and the Use Case Working Group
- Reviewed several BIG members BIBFRAME implementations and discussed their requirements for interoperability and issues encountered
- Surveyed BIG members on the cataloging standards they are currently using
- Conducted a BIBFRAME Implementation survey
- With LD4, Share-VDE, OCLC held a discovery session at the Library of Congress hosted by LC Policy, Training & Cooperative Programs Division (PTCP)
- Incorporated feedback and actions from the 2022 Linked Data Summit held at LC
- Developed a work plan
- Work plan established February 2023

Implementation Survey Questions and Analysis



- 11 implementations
 - Significantly different model from bf:2.0?
 - *50% said yes*
 - *svde:Opus & svde:Work; bflc; local extension vocabularies; BIBFRAME lite*
 - BIBFRAME version base?
 - *Mostly 2.0; one 2.1; one in the process of moving to 2.1*
 - MARC to BIBFRAME processing and version?
 - *RDFizer tool (SVDE); local conversion logic (Sweden, Finland); LC MARC2BIBFRAME convertor*
 - BIBFRAME to MARC processing and version?
 - *logic based on LC's conversion, local conversion logic (Sweden)*
- Documentation & Sample data

Speaks to the challenges associated with the charge!

2023 Work Plan



1. Define standard BIBFRAME “shape” necessary for data exchange
 - a. Utilize PCC data and standards as a test case and starting point
 - b. Start with Monographs, but include others as possible or at a later date
 - c. Review needs based on native BIBFRAME description versus from conversion (from MARC)
2. Ensure recommendations are readable by technical staff and librarians, but with updates preferably made in only one place
 - a. Investigate how to produce tabular format, possibly using DC TAP, and generate SHACL
 - b. Codify interoperability scope (formats/extensions/legacy or new, etc)
 - c. Document best practices for technical aspects of BIBFRAME interchange as identified through the work of the group
3. Share with consultants for testing and validation of assumptions

Current work: Guiding Principles



Properties/classes are evaluated against guiding principles for interchange interoperability:

1. **Identification**

Guiding principles: Distinguishing principle; also disambiguation and deduplication of resources.

2. **Discovery**

Guiding principle: The shape of the data needs to be predictable for discovery (data consistency)

3. **Data quality evaluation**

Guiding principles: Does the existing data meet cataloging requirements of the institution.
Records data provenance in a shared data environment.

4. **Data Re-Use**

Guiding principles: Catalogers should not have to re-enter the facts about the original work.
The work has to have a stable identity, so that distinguishing characteristics are preserved.
Linking entities for resource discovery.

BIG Subgroup Efforts



- SHACL/DCTap Subgroup:
 - Objectives:
 - Part 1: Setup structure and practices for the spreadsheet (influenced by DCTap) so that decisions can be captured consistently and support corresponding SHACL
 - Level of validation
 - pass/fail from spreadsheet level vs more detailed conditional SHACL based on some of the discussions at BIG
 - Part 2: Testing the process and iterating (tbd)
- BIBFRAME Interlingua Subgroup:
 - Objectives:
 - Define BIBFRAME Interlingua for Monographs - more information later in the presentation

Investigating SHACL



SINOPIA - DEV

Dashboard **Editor** Resource Templates Actions Search Sinopia

SHACL Node Property SHACL NodeShape

SHACL NodeShape

SHACL Shape Label

SHACL Node Shape Target Class

SHACL Node Shape Property

SHACL Node Shape Node Kind

Class: <http://www.w3.org/ns/shacl#NodeShape>

SHACL Shape Label

Property: <http://www.w3.org/2000/01/rdf-schema#label>

Enter a literal

SHACL Shape Label

SHACL Node Shape Target Class

Property: <http://www.w3.org/ns/shacl#targetClass>

URI SHACL Node Shape Target Class

Label Label for SHACL Node Shape Target Class

+ Add another SHACL Node Shape Target Class

SHACL Node Shape Property

SHACL Node Property

Property: <http://www.w3.org/ns/shacl#property>

Class: <http://www.w3.org/ns/shacl#PropertyShape>

- + Add SHACL Property Path
- + Add SHACL Property Class
- + Add SHACL Property DataType
- + Add SHACL Property Node
- + Add SHACL Property Minimum Count
- + Add SHACL Property Maximum Count
- + Add SHACL Property Regex Pattern

SINOPIA - DEV

Dashboard **Editor** Resource Templates Actions Search Sinopia

SHACL Node Property

SHACL Property Path

SHACL Property Class

SHACL Property DataType

SHACL Property Node

SHACL Property Minimum Count

SHACL Property Maximum Count

SHACL Property Regex Pattern

Class: <http://www.w3.org/ns/shacl#PropertyShape>

SHACL Property Path

Property: <http://www.w3.org/ns/shacl#path>

URI SHACL Property Path

Label Label for SHACL Property Path

SHACL Property Class

Property: <http://www.w3.org/ns/shacl#class>

URI SHACL Property Class

Label Label for SHACL Property Class



Current Work: Investigating DCTAP



- DC Tabular Application Profiles
- DCMI documentation: <https://github.com/dcmi/dctap>
- Elements (column headers):

shapeID	shapeLabel	propertyID	propertyLabel	mandatory	repeatable	valueNodeType	valueDataType	valueConstraint	valueConstraintType	valueShape	note

Minimal Data Exchange Requirements

C	D	E	F	G	H	I	J
Top Class	SubClasses	Property	Property Label in Sinopia	Language Support	Required? (by BIG)	Validation Severity	General Note
bf:Work	bf:Text	bf:hasInstance	has Manifestation (BIBFRAME Instance)		FALSE		required, applicable
		bf:title	Work Title		TRUE	Violation if not present	
		bf:title	Variant Work Title		FALSE		
		bf:contribution	Primary Contribution		FALSE	Warning if not present	required
		bf:contribution	Contribution		FALSE	Warning if not present	required
		bf:genreForm	Form/Genre of Work		FALSE	Warning if not present	From
			Government Publication Type		FALSE		
		bf:originDate	Date of Work		FALSE	Warning if not present	
		bf:legalDate			FALSE		
		bf:originPlace	Place of Origin of the Work		FALSE	Warning if not present	
		bf:language	Language		TRUE	Violation if not present	
		bf:notation	Script		FALSE		
		bf:geographicCoverage	(Geographic) Coverage of the Content		FALSE		

Selection of work property decisions for validation of minimal data exchange requirements

Defining BIBFRAME Interlingua



BIBFRAME Interlingua Adherence Levels

Institutions exchanging BIBFRAME data should consider the following adherence levels:

- **Level 1 (Violation if not present):**

The data elements included in level 1 are those considered essential for the functional exchange of BIBFRAME data. Any institution using BIBFRAME data in production should adhere to this level and be able to publish and receive data preserving this information. Similarly, those working on the development of BIBFRAME should exercise extreme care with any changes to these elements, working and communicating with BIG to ensure local systems are ready for any updates.
- **Level 2 (Warning if not present):** -severity =~ necessary to reduce ambiguity*

These data elements are less critical for identification and interchange, but considered key for data re-use and discovery for many communities of practice across BIG. To adhere to Level 2 these elements are not required, but where they exist they should adhere to the indicated approach to structuring and presenting the data for re-use.

Evaluating bf:Work Properties

A	B	C	D	E	F	G	H	I
Resource template name	Resource template ID	Top Class	SubClasses	Property	Property Label in Sinopia	Label in Sinopia	Required for BIG?	Validation Severity
_PCC BF2	pcc:bf2:Monograph:Work	bf:Work	bf:Text ; bf:Monograph				TRUE	Violation if not present
				bf:title	Work Title		TRUE	Violation if not present
				bf:contribution	Contribution		FALSE	Warning if not present
				bf:genreForm	Form/Genre of Work		FALSE	Warning if not present
				bf:originDate	Date of Work		FALSE	Warning if not present
				bf:originPlace	Place of Origin of the Work		FALSE	Warning if not present
				bf:language	Language		TRUE	Violation if not present
				bf:subject	Subject of the Work		FALSE	Warning if not present
				bf:classification	Classification numbers		FALSE	Warning if not present
				bf:content or bf:Text (subclass of)	Content Type		TRUE	Violation if not present
				bf:adminMetadata	Administrative metadata		TRUE	Violation if not present

Example: Evaluating bf:title

Required for:

- Identification
- Data Re-use
- Discovery

Next step:

Compare exact title shape across institutions

Example:

```
:l1 a bf:Instance>
    bf:title <http://sinopia.io/resource/title/Fugitive_Telemetry>.
    <http://sinopia.io/resource/title/Fugitive_Telemetry>
        rdfs:label "Fugitive telemetry"@en.
```

```
:l1 a bf:Instance>
    bf:title [ a bf>Title ;
              bf:mainTitle "Fugitive telemetry"@en
            ] .
```

```
:l1 a bf:Instance>
    bf:title "Fugitive telemetry"@en
```

Example: Evaluating bf:AdminMetadata Properties

Required for:

- Data evaluation
- Data Re-Use

Further review required to determine which properties of bf:AdminMetadata are required for interchange

bf:adminMetadata	Administrative metadata	TRUE	Needs more discussion with group. Which specific parts are required?
------------------	-------------------------	------	--



BIBFRAME Work AdminMetadata (Draft)*

Top Class	Property	Property Label in Sinopia	Required for BIG Work
bf:AdminMetadata			
	bf:catalogerId	Cataloger ID	FALSE
	bf:creationDate	Date Cataloged or Updated/Changed	TRUE
	bf:changeDate		FALSE
	bf:assigner	Cataloging institution	TRUE
	bf:descriptionModifier	Modifying institution	FALSE
	bf:descriptionAuthentication	Description authentication	FALSE
	bf:encodingLevel	Encoding level	FALSE
	bf:descriptionConventions	Description conventions	FALSE
	bf:descriptionLanguage	Description language	FALSE
	bf:generationProcess	NA	FALSE
	bf:generationDate	NA	FALSE
	bf:identifiedBy	Identifiers--Local	FALSE
	bf:status	Status	FALSE

**Proposal made by subgroup; needs confirmation from BIG*

Other Issues Under Discussion



MARC Conversion

- How much should MARC to BF conversion limitations affect our requirements?
- How much should the need to convert BF to MARC affect our requirements (*properties, vocabularies*)
- Are there properties that we think are intended only for conversion purposes or do they have other uses as well? (*citations, corresponding author in open source articles*)

Content Standards

- Are we working only in BF here or BF + RDA?
- If something is core to RDA, should it be required for interchange of BF?
 - if yes, what do we do with AACR2 records converted to BF which lack many RDA fields?
 - if no, how do we interact with content standards?

Other Issues Under Discussion



Value Vocabularies

- A challenge when not everyone uses the same vocabularies
- Use of range often optional when present
 - good when using a non-library vocabulary (e.g., Art & Architecture Thesaurus) where it is inappropriate to claim that a term is an instance of a BF class
 - not so good if your only existing vocabulary is a literal

Language & Script Tags

- Need for language tags on literals?
- Script tags aren't used much, but are useful

Other Issues Under Discussion



BF Extensions

- What BF extensions should we take into consideration?
 - **BF**LC (*LC BIBFRAME extension*)
 - **SV**DE (*Share Virtual Discovery Environment Ontology*)
 - **AR**M (*Art & Rare Materials Ontology*)
 - **PM**O (*Performed Music Ontology*)
- How do we evaluate them for general use?
- Do we require (i.e., request) mappings to general BF?

Next Steps



- Determine shapes (data models) for bf:Work properties based on sample data
in process
- Determine minimal interchange requirements and property shapes for bf:Instance
in process
- Create tabular data and generate SHACL
in process
- Codify interoperability scope (formats/extensions/legacy or new, etc)
- Document best practices for technical aspects of BIBFRAME interchange
- Share with consultants for testing and validation of assumptions

References--BIG Background



- BIBFRAME Interoperability Group wiki
<https://wiki.lyrasis.org/pages/viewpage.action?pageId=249135298>
- Use Case Working Group Final Report
https://docs.google.com/document/d/1n-Cmm8vfGnWp2mig2bpmvqFcKlbEIOW3ud4jiSR_c5Y/edit
- Communication Working Group Final Report
<https://docs.google.com/document/d/1CZNCSAszm4z bzUb itoQiDQ8C4gJ6zchALG2UrvnrPOw/edit>
- BIBFRAME Data Exchange Meeting 2021 (Summary)
<https://www.loc.gov/aba/pcc/bibframe/PCC-BIBFRAME-Data-Exchange-Summary.pdf>
- Linked Data Summit - Interoperability of Library Data (November 2022)
<https://www.loc.gov/aba/pcc/bibframe/Linked-Data-Summit-2022-Summary.pdf>
- Discovery Session (2 February 2023)
Zoom recording is linked [here](#). Passcode: %7UNP21g
- Wacker, M. (12 July 2023). The International BIBFRAME Interoperability Group (BIG)--Background and Current Work. LD4 Conference 2023.
<https://www.youtube.com/watch?v=Pv2hg21d1FQ>

References--Standards



- BIBFRAME Model
<https://www.loc.gov/bibframe/docs/bibframe2-model.html>
- BIBFRAME Ontology
<https://id.loc.gov/ontologies/bibframe.html>
- BIBFRAME LC Extension Ontology
<https://id.loc.gov/ontologies/bflc.html>
- Shapes Constraint Language (SHACL)
<https://www.w3.org/TR/shacl/>
- DCTAP (Dublin Core Tabular Application Profiles)
<https://www.dublincore.org/specifications/dctap/>

Questions?

Thanks to all BIG members who contributed to these slides

nlorimer@stanford.edu

bigelow@ualberta.ca