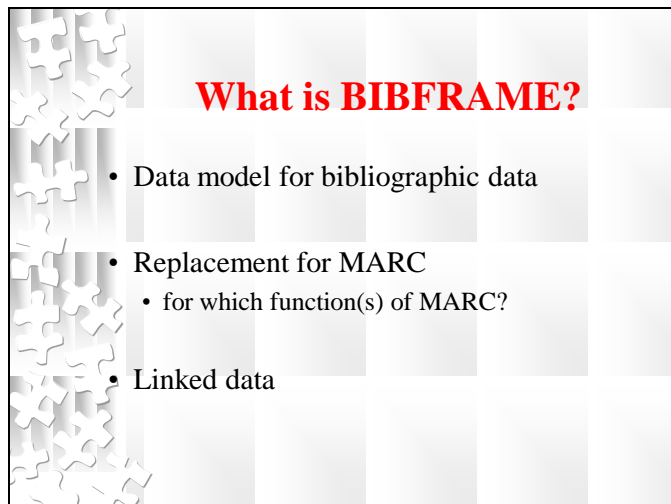
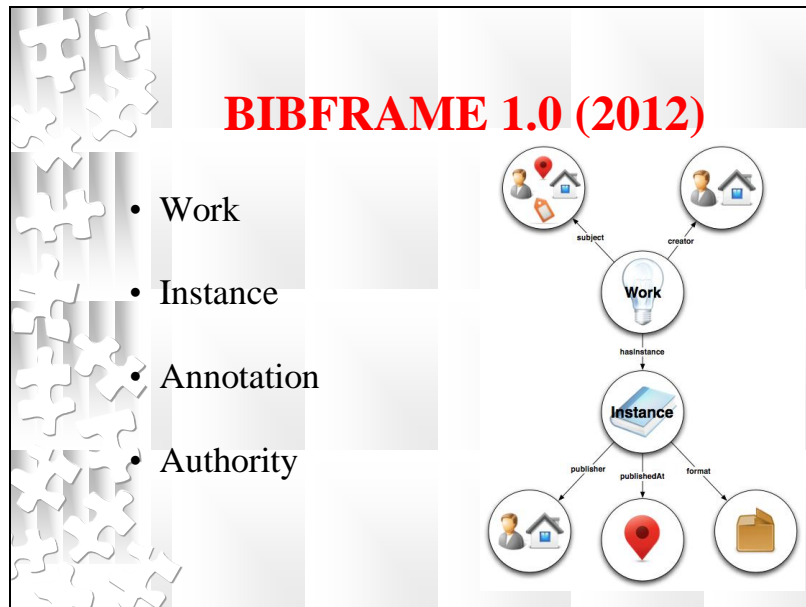


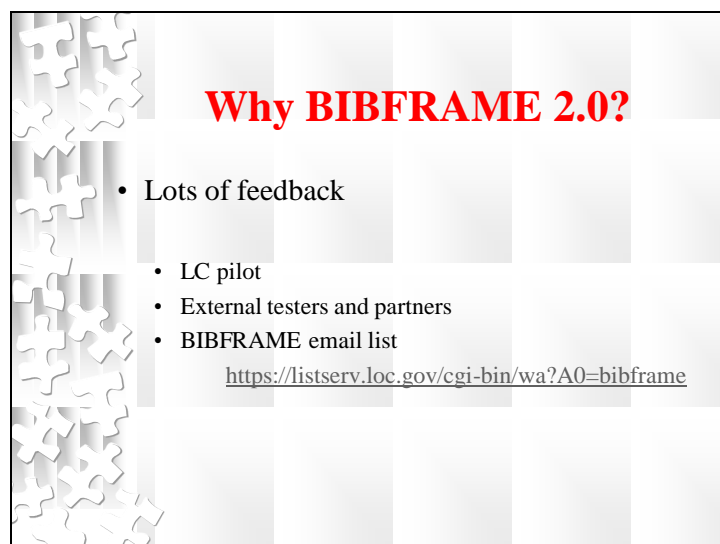
I agreed to give this talk many months ago and, as with many things, it seemed like a much better idea when it was far off in the future. So I'd like to start off with the caveat that I have not been deeply engaged with either BF 1.0 or BF 2.0 nor am I especially knowledgeable about linked data. However, I hope that some things that I've been able to glean from reading about BF and some related thoughts will be useful to you. I think this information is largely correct, at least in broad strokes, but please let me know if you notice anything I've misunderstood or misinterpreted.



Most of us have been hearing about BF for many years now so I'm just going to touch on a few brief highlights. It's a data model for bibliographic data and it's billed as a replacement for MARC. However, MARC does a lot of things and it's not clear to me exactly which function(s) of MARC BF is intended to replace. BF uses linked data, which is a standard for publishing information on the web in a structured form. Linked data promises many benefits, such as making library data more visible on the web, making it possible to share library data more broadly and effectively and being a more accessible format for users outside the library world to manipulate.



LC released BF 1.0 back in 2012. BF 1.0 has four main classes. You may not be used to the word class in this context, but it just means a type of thing or an entity, like the FRBR entities. Work and Instance are shown in this diagram. Annotation and Authority are a little more complicated, which is probably why they aren't incorporated into this simple picture.



About a year ago, LC published BF 2.0. Why did they do this? Their approach to BF 1.0 seemed to be similar to taking a whole bunch of pasta, throwing it a wall and seeing what sticks. They came up with something and put it out for people to look at and test. This got them a lot of feedback from many places, including their own internal pilot, formal and informal testers and partners and other interested parties. LC took this feedback and came up with BF 2.0.

BF 2.0: <http://www.loc.gov/bibframe/docs/>

BF list: <https://listserv.loc.gov/cgi-bin/wa?A0=bibframe>

Why BIBFRAME 2.0?

- Better modeling of audiovisual materials (AVPreserve)
- BIBFRAME AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources
<https://www.loc.gov/bibframe/docs/pdf/bibframe-avmodelingstudy-may15-2014.pdf>
- BIBFRAME AV Assessment: Technical, Structural, and Preservation Metadata
<http://www.loc.gov/bibframe/docs/pdf/bf-avtechstudy-01-04-2016.pdf>

I'd like to highlight a couple of important sources of feedback that influenced the development of BF 2.0. LC has a large off-site campus in Virginia focused on the conservation of audiovisual materials, primarily recorded sound and moving images. Catalogers there were concerned about how well BF would support description of the materials they work with. To address these concerns, LC commissioned two reports from a company called AVPreserve that does consulting related to preservation of and access to audiovisual materials.

<https://www.loc.gov/bibframe/docs/pdf/bibframe-avmodelingstudy-may15-2014.pdf>

<http://www.loc.gov/bibframe/docs/pdf/bf-avtechstudy-01-04-2016.pdf>

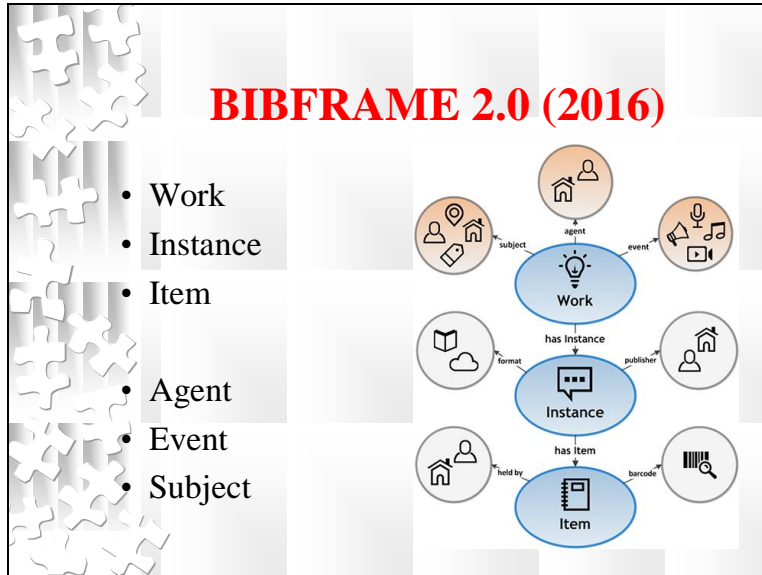
Why BIBFRAME 2.0?

- Better adherence to linked data best practices (Rob Sanderson)
- Analysis of the BIBFRAME Ontology for Linked Data Best Practices
https://docs.google.com/document/d/1dly-gQsH67Ay0T0O0ulhyRiKjpf_I0AVQ9v8FLmPNo/edit
- Linked Data Best Practices and BIBFRAME
<https://www.slideshare.net/azaroth42/linked-data-best-practices-and-bibframe>

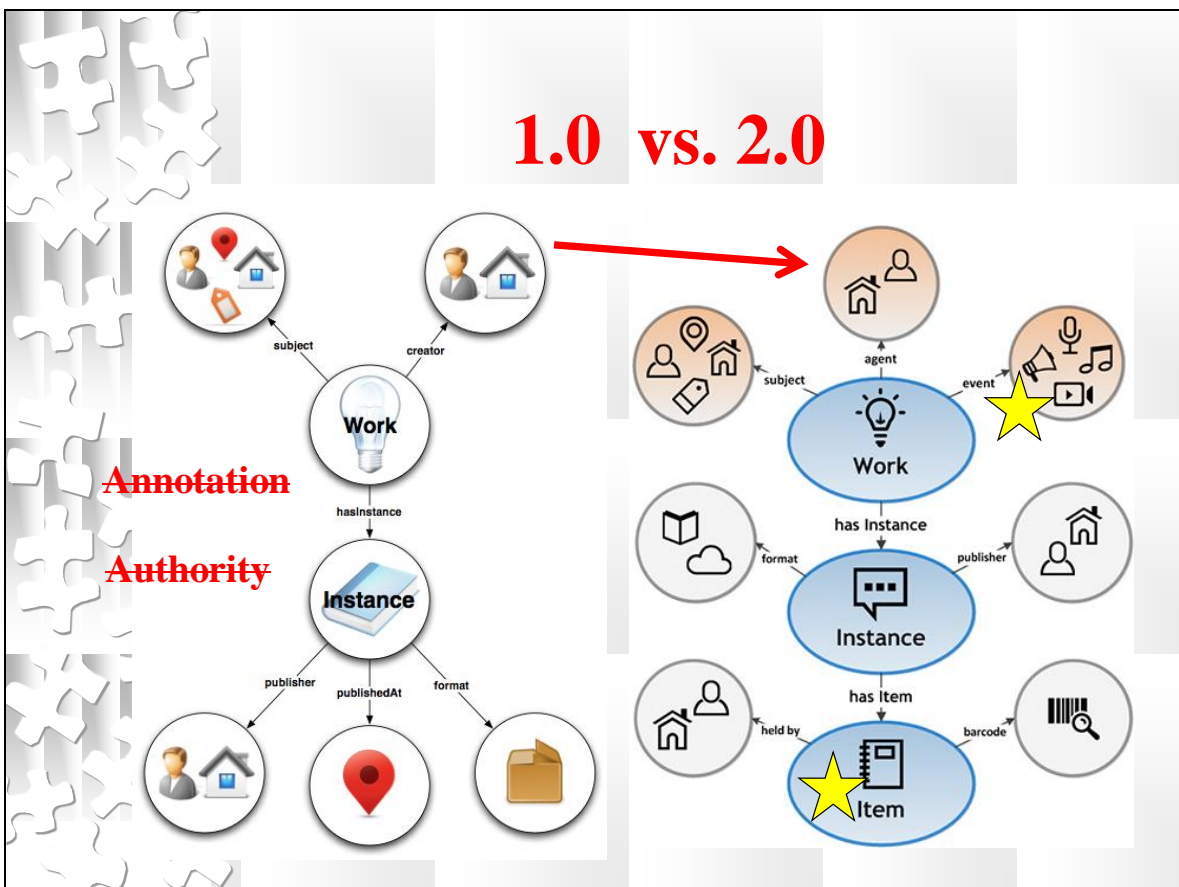
Many linked data experts have been critical of the ways in which BF fails to conform to generally-accepted best practices for linked data. In 2015, Rob Sanderson, one of these experts, prepared a detailed report on this topic at LC's request. The other link is for an interesting presentation he gave comparing BF 1.0 and 2.0 in terms of their compliance with linked data best practices.

https://docs.google.com/document/d/1dly-gQsH67Ay0T0O0ulhyRiKjpf_I0AVQ9v8FLmPNo/edit

<https://www.slideshare.net/azaroth42/linked-data-best-practices-and-bibframe>



Here's the new Bibframe 2.0 diagram. BF 2.0 features six core classes or entities.



Here are the big-picture views of BF 1.0 and 2.0 side-by-side. Annotation and Authority have gone away and Event and Item have been added to BF 2.0. In the diagram, Creator has turned into Agent.

Events

BIBFRAME AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources (2014)

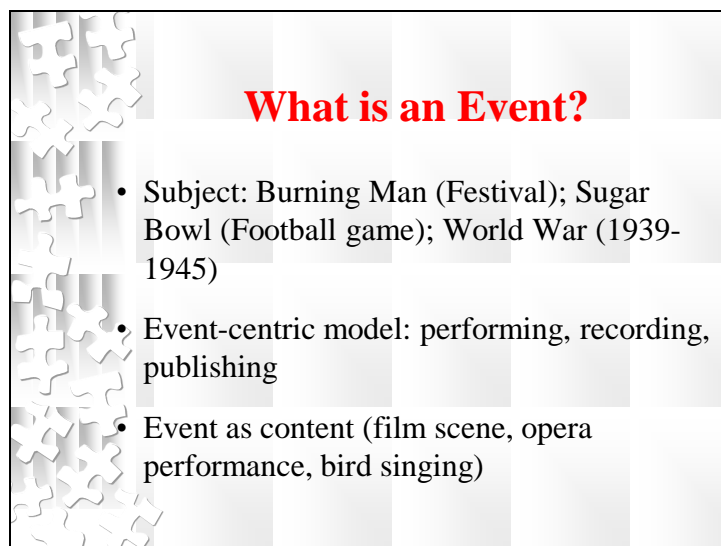
Content is

- Distinct intellectual creation
- Documentation of event that occurred in time and space

<http://www.loc.gov/BIBFRAME/docs/pdf/BIBFRAME-avmodelingstudy-may15-2014.pdf>

One of the big conclusions of the AV modeling report is that some resources do not contain works, but consist of mere documentation of events. I think there are ways of interpreting a FRBR work so that these types of resources can be said to have works, but the term work does come with a lot of baggage in the library world. Most of the videos collected by ordinary libraries are features, documentaries or other published works where the creators of the content clearly had artistic and/or intellectual intentions that they were trying to communicate. A moving image archive, on the other hand, is likely to have a lot of unpublished footage that was made for other purposes. However, libraries do collect some things that might fall into this category, such as filmed lectures or concerts or the Zapruder film.

The report recommends that instance be related to a super-class or super-category called content, which contains both works and events. A given instance could be related to a work, an event or both.



What is an Event?

- Subject: Burning Man (Festival); Sugar Bowl (Football game); World War (1939-1945)
- Event-centric model: performing, recording, publishing
- Event as content (film scene, opera performance, bird singing)

There are various ways that event could be interpreted for data modeling. The one we are probably most familiar with is event as subject. These examples are from FAST.

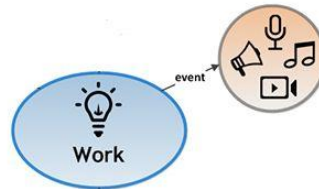
Events, not surprisingly, play a prominent role in event-centric modeling. In event-centric modeling, actions or processes like recording or publishing that happen to entities are just as important to describe as the objects themselves. FRBRoo, the object-oriented interpretation of FRBR, is an example of this approach.

The final possibility is the one suggested by the a/v modeling report, in which an event is the content of the resource, as in the examples given here, which I took from the report. [“That event may have taken place specifically for the purpose of realizing the work (a film scene), as a type of realization of a work (an opera performance), or it may have been independent of any intention to document that event (a bird singing).”]

This usage reminds me of the MARC relator term for “depicted [dpc],” except that “depicted” has a broader meaning that also includes still depictions of objects as in a painting (<http://www.loc.gov/marc/marbi/2005/2005-06.html>). On the other hand, depicted doesn’t intuitively work so well for sound.

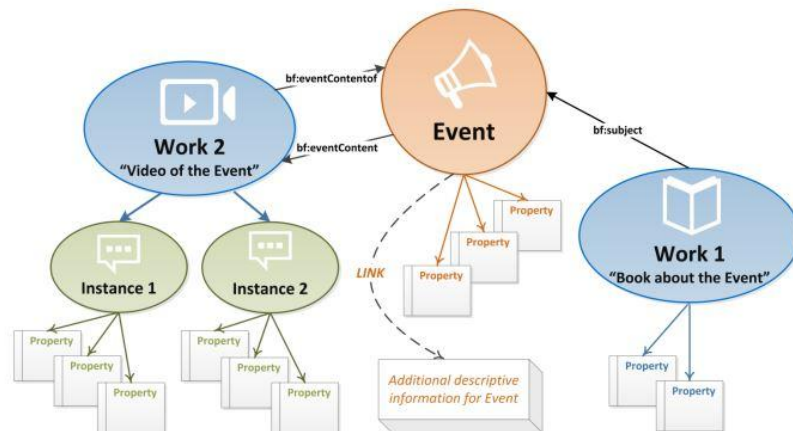
Event 2.0

Something that happens at a certain time and location, such as a performance, speech, or athletic event, that is documented by a resource.



Although the super-class or super-category of Content did not make it into BF 2.0, there is now a core class of Event attached to the Work.

Event 2.0



Here is an example from LC showing two ways in which Event might be used. On the left is a Work that consists of a recording of the Event where the event is the content of the resource. On the right is a book about the event where the Event is the Subject of the Work.

Event 2.0

```
##Description of a work with the event as its subject
<http://bibframe.example.org/work/text/bravesPirates19590526>
  a bf:Work , bf:Text ;
  bf:title [rdfs:label "Hard-luck Harvey Haddix and the greatest game
  ever lost" ] ;
  bf:subject
    <http://bibframe.example.org/event/bravesPirates19590526> .

##Description of a Work with event as its content
<http://bibframe.example.org/work/movingImage/bravesPirates19590526>
  a bf:Work , bf:MovingImage ;
  bf:title [rdfs:label "May 26, 1959, Braves vs. Pirates" ] ;
  bf:eventContentOf
    <http://bibframe.example.org/events/bravesPirates19590526> .
```

This is an RDF view of the same situation. At the top is a book with the event as Subject and below that is a film of the event itself.

Annotation 1.0

Linked Data Best Practice:
Define Your Domain


Define only terms from your domain model

- Reviews and Summaries (Open Annotation?)
- Cover art and Tables of Contents
- Items and Holdings

Better to directly link to Works, Instances...

One of the linked data best practices listed by Rob Sanderson says that you should figure out the scope of what you want to describe and then only make your own terms for things that are core to your domain. Annotations in BF 1.0 were a way to describe “third-party provided links between resources.” There were two main types of annotations. The first provided additional information about a resource. This includes the sort of content that libraries often get from external providers, such as reviews and summaries. Sanderson doesn’t think these are core to the library domain. He recommends that BF reuse existing terms defined by others where available, such as those from Open Annotation. He also believes that this information should be linked directly to the bibliographic entity rather than indirectly via an Annotation. It appears that BF 2.0 has kept classes for reviews, summaries and so on, although it does now link them directly to the resource.

The other function of Annotations was to allow libraries to make assertions about their holdings and specific items, which is a central part of the data the libraries have to manage.




Item 1.0

Item is an Annotation of an Instance

- bf:Annotation
 - bf:HeldMaterial
 - bf:HeldItem

in BF 1.0 the process of describing items was convoluted and indirect. An Instance was linked to an Annotation. The Annotation could be linked to HeldMaterial, which then could be linked to HeldItem.



Item 2.0

Core Class

An item is an actual copy (physical or electronic) of an Instance. It reflects information such as its location (physical or virtual), shelf mark, and barcode.

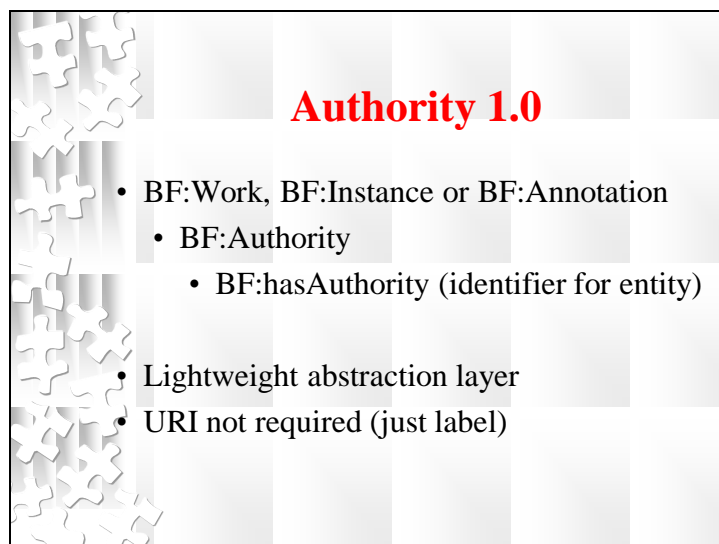
- simple
- compound (summary description of group of items)

BF 2.0 recognizes the central importance of items to library data and makes them into their own core category. BF 2.0 items can either be simple, that is a single thing, or compound. Compound items are descriptions of groups of simple or compound items so you can have multiple layers.

Item 2.0

```
<http://BIBFRAME.example.org/item/itemZ >  
  a          bf:Item ;  
  bf:heldBy  [rdfs:label "Library of Congress" ] ;  
  bf:itemOf  <http://BIBFRAME.example.org/instance/instanceY> ;  
  bf:shelfMarkLcc [rdf:value "LB2395.C65 1991" ] ;  
  bf:usageAndAccessPolicy [  
    a          bf:AccessPolicy ;  
    rdf:value  "unrestricted" ] ;
```

Here is a brief example of some of the properties associated with the Item.




Authority, like annotation, was an indirect process. You might assume that the URI for a subject or creator should be linked directly to the URI for the Work, but this wasn't what happened in BF 1.0. Instead, BF 1.0 used what LC called a lightweight abstraction layer, where a Work was linked to an Authority entity and the Authority was then linked to labels and URIs for the subject or creator.

In this scenario, a URI is not required and you could just have a label for the agent or topic that you're describing. That is essentially equivalent to what we do today when we type a string that isn't backed by an authority record into a controlled field in a MARC record. One common prediction about the move to linked data is that we'll have to do a lot more authority work. I don't think this is absolutely true from a technical perspective. Since the label strings don't have to be unique, you could just mint a new URI for every name that you come across that isn't already in the national authority file. Later, if you found out that two occurrences of the same label described the same person, you could link them with sameas. This probably isn't the most effective approach, but it is also unlikely that we will be able to rigorously identify every entity that appears in bibliographic data.

Agent 2.0

Entity having a role in a resource, such as a person or organization.

Range not specified;
not just Work



```
graph BT; Work((Work)) -- agent --> Agent((Agent));
```

in BF 2.0 the Agent class has become more prominent. The picture is a little misleading as the type of thing that an Agent can be linked to is not specified in BF 2.0 and therefore isn't limited just to Works.

Creator 1.0 & Contribution 2.0

- Linked Data Best Practice:
Define only one pattern for each feature
- BF:creator & BF:contributor → BF:contribution

```
<http://bibframe.example.org/resource> bf:contribution [
  a bf:Contribution ;
  bf:role <http://id.loc.gov/vocabulary/relators/ill> ;
  bf:role [rdfs:label "illustrator"] ;
  bf:agent <http://id.loc.gov/rwo/agents/n94064763> ] .
```

As I mentioned, BF 2.0 got rid of the Creator that appeared in the BF 1.0 picture. Another linked data best practice is to define only one pattern for each feature. In BF 1.0, there were several ways to express the role that a person played in relation to a resource. This can be confusing for the cataloger and makes it more complex to use the data because all the different possible patterns have to be accounted for.


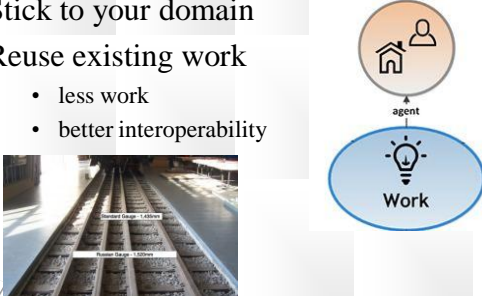
In BF 2.0, they appear to be using only one method of relating agents to resources and that is using the BF:contribution property as shown here. This is still an indirect method, but it takes a different approach than the BF 1.0 authority used.

“A BIBFRAME Agent may be associated with a BIBFRAME resource (e.g. Work) through some role, like author, illustrator, or editor. Role Association Expressed as a Contribution Property bf:contribution and Class bf:Contribution The property bf:contribution has expected value a bf:Contribution, which pairs an agent with a specific role. Example: Role is illustrator, and the association is expressed as a Contribution.”

Person 2.0

Linked Data Best Practices:

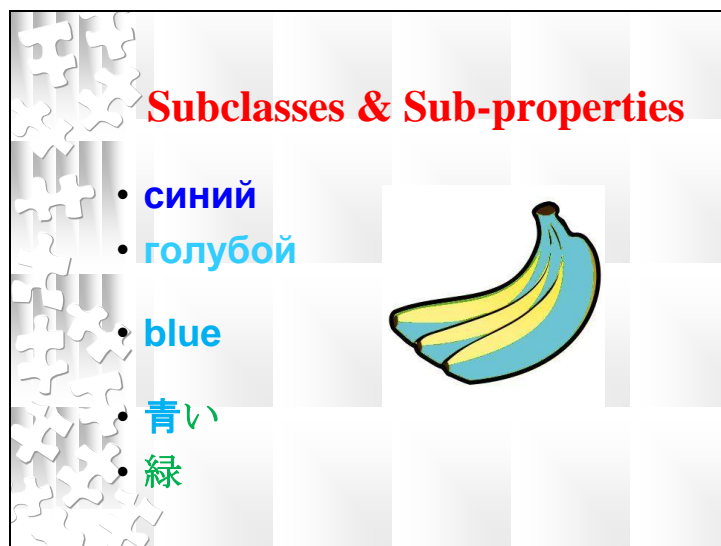
- Stick to your domain
- Reuse existing work
 - less work
 - better interoperability



Going back to this idea of linked data best practices, in addition to sticking to your domain, it is recommended to reuse existing work where possible. This is less work for vocabulary developers and supports better interoperability.

Many things work together better if everyone does them the same way. In January, the first freight train ran all the way from China to the UK. However, due to the different gauges of track used by different countries, it wasn't actually the same exact train all the way. They had to unload and reload everything on to a different physical train 5 or 6 times when the gauge changed.

In linked data, the ideal seem to be just to use someone else's term as-is. The next best alternative is to define your own term and then use sameas statements to link your term to others' terms.



Sometimes, you don't want to say the exact same thing as an existing term does. Another way to relate your term to someone else's is to say that your term is a subcategory of theirs. In English, we have the basic color term blue. In Russian, you can't call something just blue; it's either dark blue or light blue. If a Russian speaker wants to use basic color names in English, they have to be less specific. To go the other way, in ancient Japan, there was no separate word for green. The things that we call green were subsumed under blue. A word for green later entered the Japanese language, but it apparently didn't come into common use until after WWII. In daily conversation, Japanese still often refer to some types of green objects as blue. My husband is Japanese and he periodically temporarily disorients me by referring to blue bananas or traffic lights. However, as long as we are aware of the differences in specificity between our terms, we can still communicate reasonably well. A Russian and ancient Japanese person could still converse at the lowest-common-denominator level of blue-green even though some precision is lost.

So linked data does hierarchy reasonably well with this dumbing down approach, just like all sort of titles can go under the big tent of DC title.

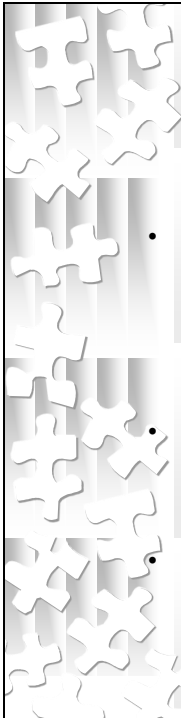
[https://en.wikipedia.org/wiki/Ao_\(color\)](https://en.wikipedia.org/wiki/Ao_(color))

Differing Worldviews


Middle School Model	Junior High School Model
12 th grade	12 th grade
11 th grade	11 th grade
10 th grade	10 th grade
9 th grade	9 th grade
8 th grade	8 th grade
7 th grade	7 th grade
6 th grade	6 th grade
5 th grade	5 th grade
4 th grade	4 th grade

However, linked data doesn't have a neat solution to the gnarlier problem of what to do when people have different ways of splitting up the universe. This is especially a challenge for multilingual thesauri, but comes up in all sorts of situations. Suppose that you know that Jack is a middle school student and you know that he is moving to a new town that uses the junior high system. That doesn't give you enough information to know whether or not Jack will be going to junior high. The likely next-highest level of specificity is K-12 student, which is not terribly informative.

[There are better examples of this, but I was not very inspired, unfortunately]



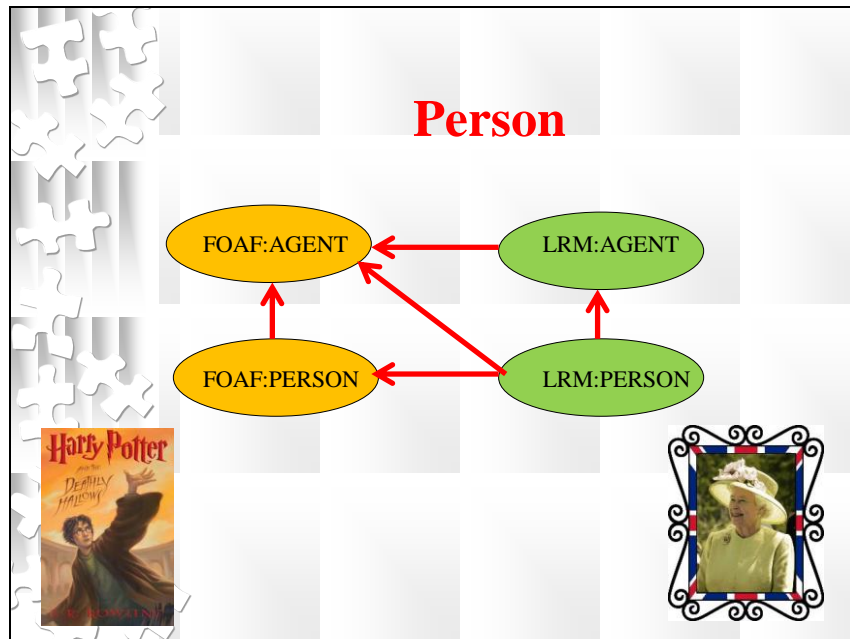
Person 2.0



- **FOAF:PERSON**
The Person class represents people. Something is a Person if it is a person. We don't nitpic about whether they're alive, dead, real, or imaginary. The Person class is a sub-class of the Agent class, since all people are considered 'agents' in FOAF
- **BF:Person**
Individual or identity established by an individual (either alone or in collaboration with one or more other individuals).
- **LRM:Person**
An individual human being
The entity person is restricted to real persons who live or are assumed to have lived.

Why might you want to define your own terms and what might be some drawbacks? Let's take a more bibliographically relevant example. The FOAF person is defined very loosely. As they say, they don't nitpic. Although some people have advocated that BF just use FOAF:Person, BF has made its own term. The BF 2.0 definition is similar to what we currently put in MARC 100 and 700 and explicitly incorporates the concept of bibliographic identity.

As you may know, IFLA has been working on reconciling the three FRBR models, FRBR, FRAD (authority data) and FRSD (subjects), for many years. These were all developed independently and each takes a somewhat different approach to data modeling. Or in the case of FRSD, a wildly different approach. IFLA's FRBR replacement is now called the Library Reference Model (LRM) and the final version is in the process of being signed off on by the IFLA bureaucracy. The final version of LRM is apparently pretty close to the last public draft, which defined a person as a real human being. This is a much narrower definition than the others. Apparently, other cultural heritage communities, such as museums and archives, define person as a real human being and the creators of LRM think it is more important for LRM to cleanly interoperate with these other cultural heritage communities than with groups that have a broader definition of person.



What happens when you mix these different definitions? This seems to be a case where it would be appropriate to say that the LRM person is a subclass or sub-category of the FOAF person, which is a subclass of FOAF agent. The two terms are linked, but there are some limitations on what you can do with them. Suppose I am making a linked data description and want to describe an entity that needs to be a FOAF:Person. If someone else has stated that Queen Elizabeth is an LRM:Person, I don't have to confirm myself that she is also a FOAF:Person. Because an LRM:Person is a FOAF:Person by definition, I can take advantage of someone else's assertion.

However, the converse is not true. Suppose I am creating an RDA description. Because RDA is committed to following LRM, if I need to describe an entity as an LRM:Person, I can't just accept a statement from someone else that Queen Elizabeth is a FOAF:Person. I have to determine for myself that she also meets the definition of an LRM:Person. Since Harry Potter can be a FOAF:Person, but not an LRM:Person, entities that others have defined as a FOAF:Person cannot be incorporated into an LRM-based description without review.



RDA & BIBFRAME 2.0

Much more “RDA-friendly”

University of Washington Mapping of RDA
Core to BIBFRAME 2.0

<http://faculty.washington.edu/kiegel/ld/rda-core-to-bibframe.pdf>

Although MARC was developed to print AACR cards and that heritage shows in its design, it aspired to be hospitable to other cataloging rules. Similarly, the proximate cause of LC finally beginning to develop a successor to MARC was feedback from the U.S. national libraries RDA test that testers didn’t believe the full potential of RDA could be reached in MARC and a more suitable carrier was needed. However, LC also wants BF to be usable with a broad range of cataloging rules not just RDA.

It seems that LC may have gone too far with this neutral approach in the development of BF 1.0 and not paid enough attention to RDA. Catalogers in the LC pilot and others identified many RDA elements that couldn’t be mapped to BF. LC has attempted to rectify these omissions in BF 2.0. The University of Washington has done some useful work trying to map the RDA core elements to BF 2.0.

<http://faculty.washington.edu/kiegel/ld/rda-core-to-bibframe.pdf>

MARC to BF 2.0 Comparison Tool

<http://id.loc.gov/tools/bibframe/compare>

Only records from LC's catalog

BIBFRAME Comparison Tool Compare MARCXML to BIBFRAME2

Identifier Bib ID LCCN Serialization Turtle RDF XML

No Yes

MARC

```

#2145cgm a22003733a 4500
001 1095322
005 20120313145642.0
007 V|||||
008 050505s2002 xx 124 mlang
010 $a 2005643144
017 $aPA1:192:344$bu:5. Copyright Office
040 $aOLCS:DLCS$aamims:DLIC
050 00 $aDUA 1310 (viewing copy)
245 00 $aMetropolis /$an UFA production ; director, Fritz Lang ; screenplay by Thea Von Harbou
.
250 $a[Restored authorized edition].
260 $c2002.
300 $3viewing copy.$a1 videodisc of 1 (ca. 124 min.) (DVD) :$bkd., $bu ;$c4 3/4 in.
500 $ccopyright: Kino International Corporation. Prev. pub.: Silent motion picture preexisti
ng. $m: English translation of titles. DCR 2002; PUB 03oct02; REG 10May03; PA1:192-344.
508 $aCopyright notice: Friedrich-Wilhelm Murnau Stiftung ; 1927 and (Renewed) Kino Internat
ional ; 2002.
509 $aRestored version has 43 minute documentary on the making of Metropolis, featurette: Th
e Digital Restoration, photo galleries featuring production stills; missing scenes; architectur
al sketches; poster artwork, 13 cast and crew biographies, audio commentary, titles in English,
French, and Spanish.
509 $aResources used: copyright data base; copyright data sheet; videodisc container.
  
```

BIBFRAME (Turbo)

```

[ a bf:noteType "Physical details" ],
[ a bf:note :
  rdfs:label
  ) Kino International ; 2002." ],
[ a bf:note :
  rdfs:label "Restored version has 43 minute documentary on the making of Metropolis,
featurette: The Digital Restoration, photo galleries featuring production stills; missing scene
s; architectural sketches; poster artwork, 13 cast and crew biographies, audio commentary, titl
es in English, French, and Spanish." ];
bf:provisionActivity [ a bf:provisionActivity,
  bf:Publication ;
  bf:date "2002" ],
[ a bf:provisionActivity,
  bf:Publication ;
  bf:date "2002"^^<http://id.loc.gov/datatypes/edtf/ ;
  bf:place <http://id.loc.gov/vocabularies/countries/xx > ] ;
bf:provisionActivityStatement "2002." ;
bf:responsibilityStatement "an UFA production ; director, Fritz Lang ; screenplay by Thea V
on Harbou" ;
bf:title [ a bf:title ;
  rdfs:label "Metropolis /" ;
  bf:titleSortKey "Metropolis /" ;
  ]
  
```

LC recently released a MARC to BF 2.0 comparison tool. It can only be used with records that are in LC's catalog. Inputting an LCCN will generate a side-by-side view of the MARC record and the BF 2.0 description.

<http://id.loc.gov/tools/bibframe/compare>

Thank you!

Kelley McGrath
kelley@uoregon.edu