BIBFRAME Must Die

Jeff Edmunds[1]

« Nous perdons un temps précieux sur une piste absurde, et nous passons sans le soupçonner à côté du vrai. » --
Marcel Proust

Today marks the 21st anniversary of Roy Tennant's seminal article "MARC Must Die" (*Library Journal*, October 15, 2002). The article begins with the observation that "When MARC was created, the Beatles were a hot new group and those of us alive at the time wore really embarrassing clothes and hairstyles,"[2] which, while making the point that MARC is "old" (a relative term), neatly overlooks the fact that the Beatles' music lives on today and that clothes and hairstyles in 2002—or even 2023—were/are at least as embarrassing as they were in the 1960s.

It's 2023, and MARC lives on too. Proponents of BIBFRAME, once touted as the "replacement" for MARC, and then, when replacement became obviously impossible, shifted their rhetoric to refer to BIBFRAME as MARC's "successor." But even "successor" is misleading, for a successor comes after, and so far, there is no "after" for MARC. It's still here.

The BIBFRAME movement originated in ideas about the modeling of bibliographic data that arose in the mid-1990s. Unfortunately, much of the resulting conceptualizations about how people look for and find library materials were developed before Google, before Facebook, before smartphones, before 5G networks, before Alexa and the Internet of Things, before ChatGPT and generative AI, and before a host of other technologies that have utterly changed the way people search for, find, and think about information.

Tennant argued that MARC must die because it represented an old, and therefore outdated, approach to the capture and exchange of bibliographic metadata.

I will argue that it is BIBFRAME, not MARC, that is fundamentally flawed, already outdated, and a danger to the future of cataloging. Pretending to be modern, BIBFRAME—along with the RDA/LRM that underpins it—is in fact premised on antiquated modes of thought about information discovery and retrieval and about the optimal methods for describing resources in a networked world.

Amanda Cossham succinctly pointed this out in her 2017 PhD dissertation:

> "The explicit focus on user needs in the FRBR model, the International Cataloguing Principles, and RDA: Resource Description and Access does not align well with the ways that users use, understand, and experience library catalogues nor with the ways that they understand and experience the wider information environment. *User tasks, as constituted in the FRBR model and RDA, are insufficient to meet users' needs*."[3] (p. 11, emphasis in the original)

---

[1] The views expressed here are the author's and not (necessarily) those of my colleagues or employer.
[2] "MARC Must Die," Roy Tennant, *Library Journal*, October 15, 2002, pp. 26-28.
[3] *Models of the Bibliographic Universe*, 2017,

Proponents of BIBFRAME and the "new" (i.e., Official) RDA ignore this simple fact: the conceptual underpinnings of BIBFRAME and RDA are wrong.

They are wrong because the conceit that users or catalogers do, can, or should, think of library materials in terms of Works, Expressions, Manifestations, and Items is absurd now, and probably always was. They are wrong because the bibliographic universe on which they were supposedly based has morphed into a multiverse whose features and workings would have been nearly unimaginable thirty years ago.

Seemingly enamored of the oracular pronouncements of Tim Berners-Lee, proponents of BIBFRAME have advanced an agenda divorced from reality. A sizable influx of grant money from the Mellon Foundation and the Institute of Museum and Library Services ($500,000 for the BIBFLOW project, several millions more for the various phases of the LD4L project, to say nothing of the resources and thousands of person-hours the Library of Congress has sunk into its own BF efforts) has fueled the illusion that BIBFRAME is the future of resource description.

It isn't, and it shouldn't be.

The future lies elsewhere, with full-text indexing, big data, and increasingly impressive AI systems that can parse data and draw accurate inferences in the absence of clean, consistently packaged data, rather than on tenuous and ever-fluid relationships between entities that BIBFRAME and RDA try to "bake in," at enormous cost to catalogers, vendors, and libraries.

I am reminded of the statement made by David J. Fiander in 2001, in an article referenced by Tennant in his call for MARC's death:

> "Computing power is now cheap enough that it's no longer necessary to make things easy for computers at the expense of people."[4]

What was true in 2001 remains true in 2023. "Making things easier for computers at the expense of people" perfectly conveys the guiding principle of the BIBFRAME movement.

Having painted themselves, and us, as catalogers, into a conceptual corner from which there is no exit, the Library of Congress and the coterie of LD4L adherents find themselves compelled to push for various "simplifications" and tweaks to MARC, none of which will solve much but will allow them to more easily or less lossily convert rich MARC data into the machine-friendly RDF triples that will supposedly flow from the adoption of BIBFRAME.[5]

Of the few libraries who have "fully" transitioned to BIBFRAME, only the National Library of Sweden, to my knowledge, has published an assessment of the results:

In their 2019 article, "Leaving Comfort Behind: A National Union Catalogue Transition to Linked Data," about the Swedish National Library's transition, Bodil Wennerlund and Anna Berggren conclude "The

---

https://figshare.com/articles/Models_of_the_bibliographic_universe/5216347

[4] "Applying XML to the Bibliographic Description," David J. Fiander, *Cataloging & Classification Quarterly*, Vol. 33(2) 2001, pp. 17-28.

[5] Examples of these changes are the proposed abandonment of Romanized data in MARC records (which was rejected by the library community) and the proposal for the 264$s, an un-subfielded string to record publication data (also rejected as of this writing for being incompatible with the spirit of MARC *and* with the spirit of BIBFRAME).

advantages of linked data have not been immediately obvious."[6] In a similar vein, during the 2021 BIBFRAME European Workshop, Alan Dunskin from the British Library said, "if the benefits of linked data are not to be exploited, then the justification to invest in it becomes much more difficult […] and so the incentive to move away from MARC will be weakened."[7]

I agree.

Here are five reasons BIBFRAME must die.

BIBFRAME WOULD COST LIBRARIES DEARLY

As Marshall Breeding pointed out during a presentation in November 2022

> "Transition to BIBFRAME will involve substantial investment: technology development, staff training, documentation, etc. Libraries can expect cataloging costs to be *dramatically higher* for BIBFRAME compared to MARC, *for many years beyond the initial transition*. Efficiencies may be gained over time."[8] [emphasis mine]

Note that "dramatically higher costs" are framed as a given, and "efficiencies" as a hypothetical. Note too that these dramatically higher costs will continue for many years beyond the transition.

The LD4L libraries most enthusiastic about BIBFRAME, among them the Library of Congress, Stanford, and the University of Alberta, have been the beneficiaries of millions of dollars in grant money from the Mellon Foundation and the IMLS. Those grants, to my knowledge, have ceased.

Financially, libraries are now on their own.

The large academic research library where I have worked for 34+ years is commonly described as being one of the top ten in North America. The university it serves is facing, as of this writing, a $140 million budget deficit. In this climate, making a case for the adoption of BIBFRAME—a hugely costly endeavor whose benefits to anyone are dubious and eventual efficiencies hypothetical—would be like pitching a $10,000 voice-activated pneumatic nail gun equipped with Bluetooth® to a carpenter building a bookcase who already has a $20 hammer hanging from her tool belt.

BIBFRAME IS INFEASIBLE

For BIBFRAME to work in any nominally meaningful way, thousands of dominoes would have to fall into place. To cite but a handful:

- Libraries, vendors, and publishers would have to abandon MARC for BIBFRAME.
- Inconsistencies in various local implementations of the BIBFRAME ontology would have to be reconciled.

---

[6] https://library.ifla.org/id/eprint/2745/1/s15-2019-wennerlund-en.pdf

[7] For recordings of the 2021 Workshop, see https://www.casalini.it/bfwe2021/

[8] "BIBFRAME and Linked Data: System Readiness," Marshall Breeding, presentation for the Linked Data for Libraries Users Group of the Midwest Collaborative for Library Services on November 3, 2022. https://librarytechnology.org/document/27923

- A standard interchange flavor of BIBFRAME would have to be agreed upon and adopted.
- Editors that allow cataloging natively using BIBFRAME would have to be reconciled so that the data output from one is indistinguishable from, or at least compatible with, the data output from another.
- Entity management systems would have to be designed, deployed, tested, and adopted that allow shared management of linked data, obviating the need for thousands of separate silos.
- All major and minor library vendors (Clarivate, EBSCO, SirsiDynix, etc.) would have to convert their ILS and discovery systems to be based on BIBFRAME and linked data rather than on MARC.
- Vendors that specialize in the creation of MARC data, such as SkyRiver, Cassidy Cataloguing, and Bibliographic Data Services (BDS), would have to be cajoled into completely retooling their operations to generate BIBFRAME data instead of MARC records.
- OCLC would have to transform WorldCat, or a significant portion thereof, from a MARC-based database of bibliographic and authorities records into a BIBFRAME-based triplestore of entities and relationships between those entities.
- OCLC would have to devise and deploy editors and APIs to allow the native creation of BIBFRAME data in WorldCat, for all formats of materials, including serials for CONSER, and the exchange of this BIBFRAME data with its members and customers.
- Official RDA would have to be mapped to BIBFRAME and the mapping agreed upon and adopted by all interested stakeholders.
- Official RDA would have to be universally trained for, adopted, and implemented.
- Catalogers and technical staff at many or most of the 117,000+ libraries in the United States, and at the hundreds of thousands of libraries around the world, would have to be trained to catalog using BIBFRAME, or at least to understand its principles and the data in which it is expressed.
- All (or most) library catalogs and discovery systems would have to be redesigned to operate on the principles of BIBFRAME linked data rather than document-centric MARC records.
- A central triplestore would have to be designed, created, and maintained, since one of the fundamental tenets of the linked data ethos is not to copy data locally but to store it centrally.

The scenario above defies credulity.

Even if, through some unimaginable confluence of miracles large and small, a few of these massive dominoes were to fall, would users, catalogers, or libraries be better off?

There is zero evidence that they would.

BIBFRAME ADDS NO VALUE TO THE DISCOVERY ECOSYSTEM

Proponents of BIBFRAME have repeatedly claimed in vague terms about its superiority over MARC in a networked world. The University of Alberta Library's Linked Data Implementation Plan (2023)[9], for example, refers to "enhanced user experience" and "better discovery" a total of seven times in the course of the thirty-page document. Nowhere are these terms defined.

[9] https://docs.google.com/document/d/1heSEuYO_9fry764ywAZak2Zc5J5vpvMB-zqEg4cAPY8/edit#heading=h.kqpepx3zcs6d

But what, if any, problem, are they trying to solve? As Kyle Banerjee wrote, "[Linked data for libraries] doesn't solve human problems."[10] Expanding on this idea, he continues: "Linked Data is a powerful tool, but only for problems that have technical origins. As the term implies, Linked Data depends on data. Metadata needs consistent and complete access points. Ontologies and vocabularies need to be comprehensive and well-maintained. Systems need to know what to do with the data they retrieve. None of those requirements is met for general library use, nor is there reason to expect they will be."

Banerjee is correct. Our time is best spent improving our metadata and cataloging massive quantities of un- and under-cataloged materials, not wringing our hands over data models devised decades ago by an elite few librarians and theorists who are now attempting to completely retool the art and practice of cataloging to fit ideological aims.

During a presentation in 2022, Ruth Kitchin Tillman said: "I'm not yet convinced that a shift to BIBFRAME will help us achieve the best outcomes for the 99% of users who are not going to do deep relational dives."[11] If this assessment is correct, and I believe it is, what added value is brought by BIBFRAME to the table?

None that has yet been demonstrated.[12]

BIBFRAME IS USER UNFRIENDLY, WHOEVER THOSE USERS MIGHT BE

As Marshall Breeding pointed out in the presentation cited above, a transition to BIBFRAME would involve, at a minimum, "technology development, staff training, [and] documentation." The Library of Congress has stated that it will not implement the Official RDA Toolkit until all its catalogers are "brought into" BIBFRAME. As of now this has yet to occur. The Program for Cooperative Cataloging has repeatedly kicked the can of Official Toolkit implementation down the road as increasing numbers of

---

[10] *"The Linked Data Myth,"* Kyle Banerjee, *Library Journal*, August 13, 2020.
https://www.libraryjournal.com/story/the-linked-data-myth
[11] "Between Exports and Infrastructure: Linked Data Systems in 2022," Ruth Kitchin Tillman, presentation on October 14, 2022 at the Potomac Technical Processing Librarians 98th Meeting. https://ruthtillman.com/talk/ptpl-exports-infrastructure-bibframe/. (This observation should remind us of the fundamental flaw of the BIBFRAME/linked data approach to resource description: that the primary value of bibliographic and authorities data is *as data*, and not as pointers to the things described. The "99%" figure should remind us that BIBFRAME is an elitist enterprise conceived by an actual 1% to serve a hypothetical 1% of library users.)
[12] The SHARE-VDE project, to date the most ambitious multi-institution implementation of BIBFRAME data in the context of a database and discovery interface, has demonstrated no functionality to substantively differentiate it from commercial discovery layers built on MARC data, full-text indexing, and algorithms. The only user testing of SHARE-VDE of which I am aware, presented at the 2022 BIBFRAME Workshop ("User Experience Testing in the ShareVDE 2.0 Catalog," see https://www.bfwe.eu/budapest_2022), was performed on a tiny group of test subjects at the University of Pennsylvania, who were presented with user tasks that no users, in my experience, would normally undertake on their own. During the day's wrap-up session, Beth Picknally Camden (UPenn) observed that "Users also have a record-based thinking." In other words, they think more in a MARC-like way, associating a record with a single resource, than a BIBFRAME/WEMI way, picturing a cloud of entities and relationships between those entities. This should come as no surprise.

voices are raised in resistance to its rocambolesque jargon and user-unfriendly design.[13] The current plan—if it can be called a plan—is to roll out the Official Toolkit over a 4-year period, from 2023 to 2027.

The unpleasantness of swallowing an extremely bitter pill is not mitigated by swallowing it slowly.

When, in early 2023, the PCC Sinopia Cataloging Affinity Group presented on the PCC's test of the Toolkit using Sinopia (the BIBFRAME editor created as part of the LD4L grant-funded projects), it noted the "Circularity and 'rabbit hole' effects of new Toolkit structure."[14]

Complexification is not the solution to whatever ails cataloging practice, if indeed anything ails it.

Simplicity is.

BIBFRAME IS BORNE OF, AND STEEPED IN, INEQUITY

BIBFRAME implementation and the Official RDA Toolkit are costly, out of the reach of most libraries, and therefore elitist, serving the very few at the expense of the many. Libraries are facing, and have faced for many years, an age of austerity, and austerity always punches down. The handful of libraries, especially in North America, who have shown the most interest in BIBFRAME, are some of the most well-resourced, serving Predominantly White Institutions (PWIs) with multi-billion-dollar endowments.[15]

No less an authority than Karen Coyle has pointed out how a small group drove the development of FRBR, the mindset that underpins the LRM/RDA/BIBFRAME movement:

> "In the end, FRBR was developed and reviewed by a very small constituency that was not representative, by any measure, of the library community that prompted its development. […] [O]ne must acknowledge the fact that the end result of such a process may not actually serve the larger library community."[16]

In other words, a tiny minority of library thinkers set the cataloging world on a path toward expensive irrelevance.

---

[13] Disenchantment with RDA, especially the debacle of the Official RDA Toolkit, is so widespread and so keen that an alternate set of rules is being crafted by a community of career catalogers: the Open Rules for Cataloging (ORC), see https://openrulesforcatalogingproject.github.io/

[14] Recording of presentation at https://drive.google.com/file/d/1ejLPLmlZa16fzr-oOpvXYh4rSITXtJEF/view

[15] Stanford's endowment as of August 2022 was $36.3 billion (https://facts.stanford.edu/administration/finances/#:~:text=4%25-,Endowment,aid%20during%20the%20fiscal%20year); the University of Alberta's endowment a few year ago stood at CA$1.3 billion (https://web.archive.org/web/20181222030353/https://www.ualberta.ca/about/facts/giving); the University of Pennsylvania (a private Ivy League university not to be confused with the public Pennsylvania State University) had an endowment of $21 billion as of June 2023 (https://investments.upenn.edu/about-us) ; Cornell's endowment was, in 2022, $9.8 billion (https://cornellsun.com/2022/10/03/cornells-endowment-reports-1-3-percent-loss-amid-global-market-turmoil/); the Library of Congress's 2021 budget topped 802 million tax payer dollars.

[16] Karen Coyle. *FRBR Before and After: A Look at Our Bibliographic Models*. Chicago, IL: ALA, 2016. p. 74.

In Conclusion (a Personal Note)

When I began studying BIBFRAME many years ago, I approached the topic with an open mind and was operating on the assumption that BIBFRAME would replace MARC within a few years. The more I learned, however, the more my ignorance graded into disbelief, thence into skepticism.[17]

Because I have come to realize how much time and how many resources are being squandered on a solution to a nonexistent problem, my skepticism of BIBFRAME has tipped toward disdain.

Roy Tennant famously proclaimed that MARC must die, for theoretical reasons.

I call for BIBFRAME's demise for practical ones.

The path being laid out and championed by the Library of Congress and others—abandonment of MARC, adoption of the unusable Official RDA Toolkit, and of BIBFRAME—is expensive, infeasible, and inequitable, and offers no significant improvements to catalogers, libraries, discovery vendors, or the public we collectively serve.

BIBFRAME must die.

---

[17] For examples of this skepticism, see my three 2017 discussion papers: "BIBFRAME as Empty Vessel" (https://dx.doi.org/10.26207/zs35-2546), "Roadmap to Nowhere: BIBFLOW, BIBFRAME, and Linked Data for Libraries," and "Zombrary Apocalypse!: RDA, LRM, and the Death of Cataloging" (available from the author, jhe2@psu.edu, as PDF files).