

8 September 2026

BIBFRAME Workshop in Europe 2026

Using BIBFRAME data at scale

Jeff Mixter

Director of Global Metadata and Innovation

Chris Madden

Principal Architect, Data Services

Quick updates

Dewey: Meeting people where they are

Knowledge organization only works in a language you actually read.

- Dewey already exists in six languages; Spanish and Arabic now reach the linked data too
- Readers meet subject headings they recognize, in their own script
- The URIs never change, so every language points at the same shared concept

Who this reaches

LIVE

Spanish-language access to Dewey data

SOON

Arabic data follows in the coming months

NEW

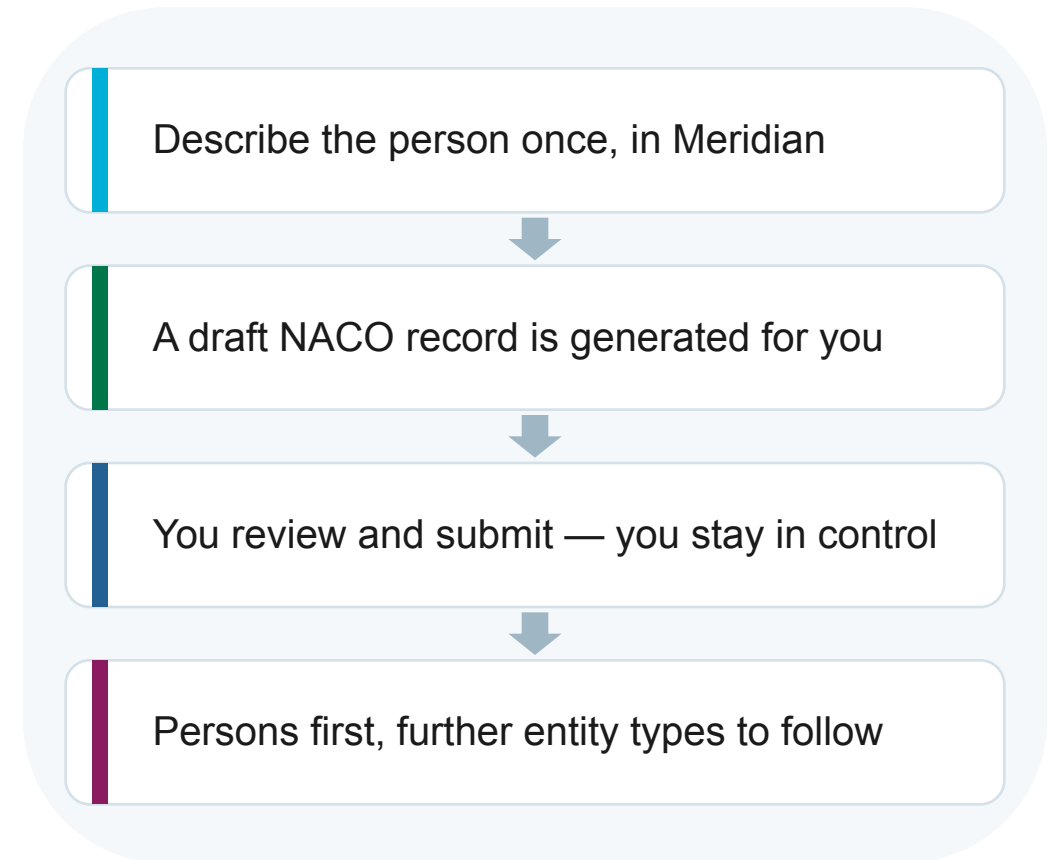
Arabic interface, read right to left

One vocabulary. Many more people served.

Taking the double work out of authority work

Contributing to NACO should not mean describing the same person twice.

- Libraries doing NACO through OCLC can generate a draft NACO record from a WorldCat Entity
- One descriptive act instead of two, with no rekeying into a second system



Evolution of WorldCat

WorldCat: 50+ years of cooperative metadata

50+ years

of continuous cooperative cataloging

Tens of thousands

of libraries across the world

Most trusted

actively managed data quality

Platform agnostic

interoperable with virtually every system

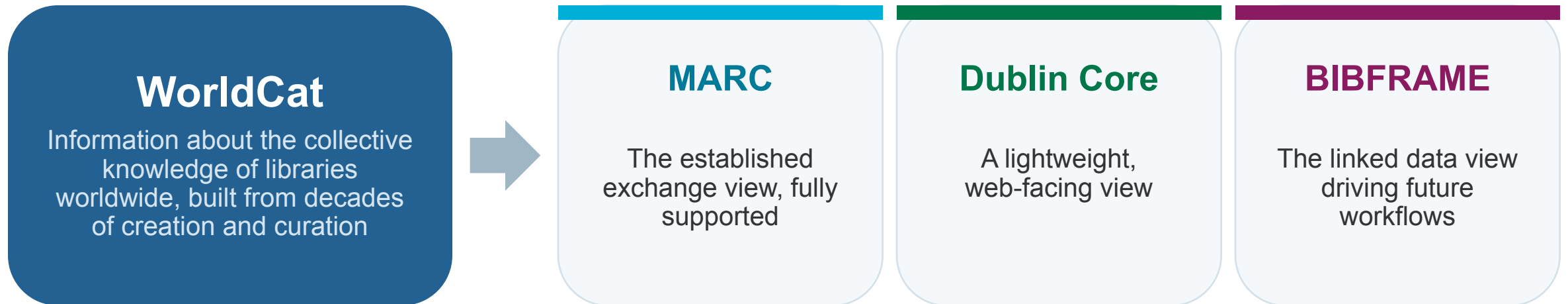
Metadata at scale is not an aspiration—it's 50+ years of practice.

- Built with the library community and continuously shaped by it
- Serving academic, public, national, school and special libraries alike
- The most comprehensive, highest-quality bibliographic dataset in the world
- One shared description, enriched by OCLC staff and expert librarians around the world

WorldCat today: One description, many lenses

WorldCat is not simply 630M+ MARC records. It is a description of the world's collective knowledge, built and curated over decades.

- Created with the library community and still adapting as their data needs evolve
- The same underlying information, expressed through whichever lens a library needs



WorldCat tomorrow: A connected knowledge graph

As AI and evolving technologies further impact library services, the importance of trusted data increases.

- The work underway is turning WorldCat into a connected knowledge graph
- That graph gives library AI services something verifiable to stand on: trusted entities, trusted relationships, real provenance
- Discovery, research, and reference tools that can show their sources, not guess at them



All of WorldCat, modeled as BIBFRAME

Not a subset. Not a sandbox. All of it.

- Years of work by OCLC's BIBFRAME specialists to build a BIBFRAME expression of WorldCat
- Quality and ingest pipelines rebuilt to treat MARC and BIBFRAME as equals
- Foundations in place, with no disruption to existing MARC workflows

What the work covers	
Scope	The full WorldCat corpus, not a sample
Model	Entity modelling reviewed by our specialists
Data	Ingest and quality pipelines rebuilt
Parity	MARC and BIBFRAME views held in step

“BIBFRAME-ifying” WorldCat

WorldCat as BIBFRAME: By the numbers

Works

450 Million

in the BIBFRAME view of WorldCat

Instances

650 Million

in the BIBFRAME view of WorldCat

RDF Triples

250 Billion

in the BIBFRAME view of WorldCat

This is BIBFRAME at production scale, not a demonstration set.

- The complete WorldCat corpus, expressed as BIBFRAME and continually refreshed as libraries catalog
- Hundreds of billions of assertions are normalized, compared, and reconciled through a multi-stage evidence pipeline
- **Work formation also reveals relationships among distinct Works, creating a foundation for graph-native discovery**

From MARC to BIBFRAME at WorldCat scale (1)

1 | INTERPRET

Create BIBFRAME Assertions

Transform MARC descriptions into BIBFRAME entities, properties, and relationships while preserving identifiers and provenance.

☐ BIBFRAME Instances

2 | NORMALIZE

Make Evidence Comparable

Normalize titles, dates, languages, identifiers, contributors, formats, and relationships into stable comparison signals; informed by decades of experience

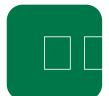
☐ Comparable Evidence

3 | GATE

Build Candidate Neighborhoods

Apply strict initial gates to cluster records that could describe the same Work, while excluding unrelated material from deeper evaluation.

☐ Bounded Candidate Sets



Normalization and candidate selection are well-understood prerequisites.

At scale, they create the foundation for evidence-based identity and relationship discovery.

From MARC to BIBFRAME at WorldCat scale (2)

4 | EVALUATE

Accumulate Evidence

Evaluate dozens of exact, fuzzy, and weighted signals across bibliographic dimensions, preserving agreement, conflict, and insufficient evidence as distinct outcomes.

- ☐ Evidence Profiles

5 | RESOLVE

Determine Intellectual Identity

Group records when the evidence supports a shared Work, preserve hard boundaries when it does not, and retain ambiguous candidates as unresolved.

- ☐ Work Assignments

6 | MATERIALIZE

Publish a Trusted Graph

Create stable BIBFRAME Works, connect their Instances, and retain confidence, evidence summaries, versions, languages of cataloguing, and provenance.

- ☐ Works + Graph Assertions



BIBFRAME Works are not created through a single match. They emerge from the progressive accumulation and resolution of evidence across multiple bibliographic dimensions.

Building identity, discovering adjacency

WORKS

Stable intellectual identity



Identity is the immediate product,
the expected outcome

IDENTITY FORMATION

Compare • Corroborate • Distinguish



Adjacency is the hidden product,
revealed through evidence

INSTANCES

Coherent bibliographic resources

ENTITY FORMATION

Normalize • Reconcile • Enrich

250 BILLION TRIPLES

**Identity organizes the corpus,
adjacency unlocks it**

Work formation establishes trusted
intellectual identity across the corpus.

Near-match relationships connect distinct
Works, enabling discovery, exploration, and
future graph-native services.

A graph-native discovery question

Combining computed adjacency with explicit BIBFRAME assertions

Starting with one Work, which distinct Works belong nearby, and what evidence connects them?

Ask the Global Graph

```
SELECT ?relatedWork ?sharedSubject

WHERE {

    <work-uri>
      skos:relatedMatch ?relatedWork ;
      bf:subject ?sharedSubject .

    ?relatedWork
      bf:subject ?sharedSubject .

}
```

Illustrative query

1 | Begin with discovered adjacency

`skos:relatedMatch` limits the query to distinct Works connected through evidence discovered during Work formation.

2 | Add explicit bibliographic context

Shared `bf:subject` assertions help explain why two distinct Works belong in the same neighborhood.

3 | Corroborate and prioritize

Creator, classification, expression, and related-title evidence can corroborate, rank, and contextualize the connection.

What becomes visible at global scale

Global context | No single record sees the whole picture

Relationships emerge across the complete cooperative corpus, allowing distinct Works to be evaluated in broader context and meaningful proximity to become visible beyond individual collections.

Corroborated evidence | Multiple dimensions strengthen relationships

Creator • Subject • Classification • Expression • Related titles

Independent evidence can reinforce, challenge, rank, and contextualize a relationship before it is surfaced.

Explainable relationships | The graph preserves more than identity

The graph records **what belongs together, what remains distinct, and why distinct Works remain connected**. That context can be reused across discovery, research, and future graph-native services.

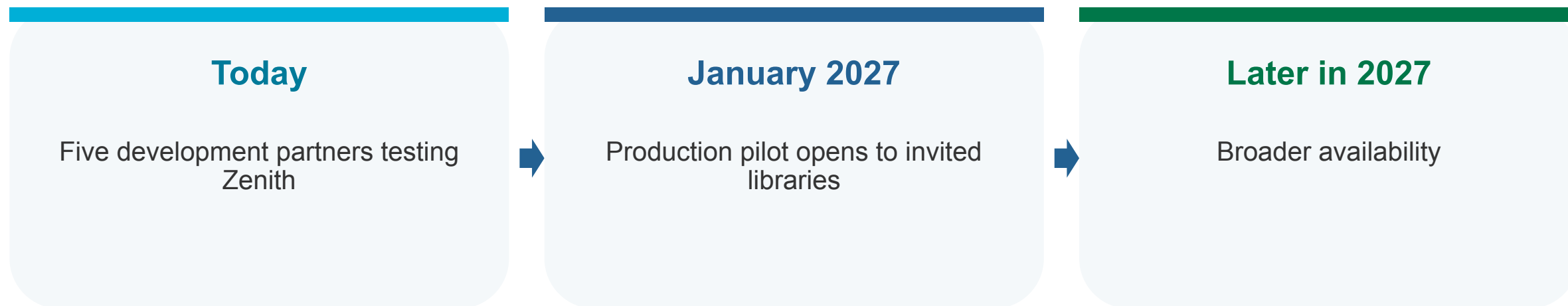
**Exact matches organize WorldCat.
Near matches make WorldCat navigable.**

What's next

Zenith: from prototype to production

Zenith, our BIBFRAME editor, becomes a production pilot.

- Shaped from the outset by five development partner libraries
- Native BIBFRAME cataloging backwards compatible with MARC



Technical BIBFRAME documentation

The BIBFRAME counterpart to Bibliographic Formats and Standards (BFAS).

- A full technical specification of how BIBFRAME is expressed in WorldCat
- Entity by entity, property by property, with worked examples
- The day-to-day reference catalogers have had for MARC for decades
- Published alongside the Zenith pilot and maintained as the model evolves

What will be published

Entity and class specifications

Property-by-property guidance

Worked examples and usage notes

Versioned as the model changes

Learning our way into BIBFRAME, together

Engagement programming to assist the whole community through the transition.

- Open to the library community—on-demand and freely available
- From fundamentals through to hands-on cataloging in BIBFRAME

BIBFRAME fundamentals

Concepts, entities and why the model matters

Transitioning to BIBFRAME

What managing data as BIBFRAME means in practice

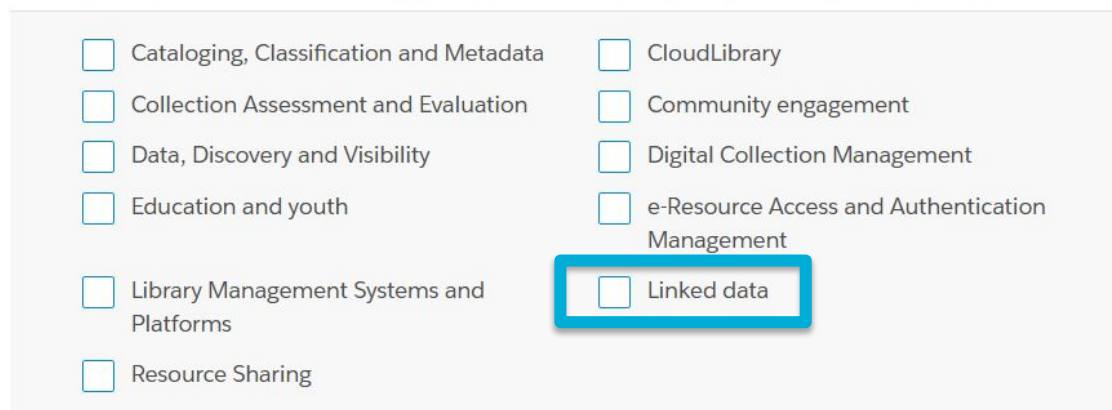
Cataloging in BIBFRAME

Working natively in BIBFRAME, hands on

Where we go from here

BIBFRAME does not need a new pool of data. It needs the world's collective knowledge already curated by librarians over decades.

- Talk to us about joining the Zenith production pilot
- Tell us what your BIBFRAME workflows need next
- Stay tuned for more information about our engagement programming
 - Visit **oc.lc/email**
 - Enter your info
 - Select **Linked Data**



A screenshot of a survey form with two columns of checkboxes. The 'Linked data' checkbox is highlighted with a red border.

☐ Cataloging, Classification and Metadata	☐ CloudLibrary
☐ Collection Assessment and Evaluation	☐ Community engagement
☐ Data, Discovery and Visibility	☐ Digital Collection Management
☐ Education and youth	☐ e-Resource Access and Authentication Management
☐ Library Management Systems and Platforms	☒ Linked data
☐ Resource Sharing	

**Because
what is
known must
be shared.[®]**

Jeff Mixter

Director of Global Metadata and
Innovation
mixterj@oclc.org

Chris Madden

Principal Architect
maddenc@oclc.org