

M25 Conference 11<sup>th</sup> May 2022

# Linked Data as a Cooperative Effort

**Fiona Carr**

Sales Account Manager



# Fiona Carr

Sales Account Manager, UKI

---



# Annette Dortmund

Senior Product Manager

---

---

# OCLC OVERVIEW

---

# OCLC – all about the community

- Non-for-profit library member cooperative
- Fully functional suite of products and services
- Engagement as a differentiator
  - OCLC Research
  - Partnership
  - Community driven

---

# LINKED DATA AS A COOPERATIVE EFFORT

---

# A decade with Linked Data

[oclc.org/linkeddataresearch](https://oclc.org/linkeddataresearch)



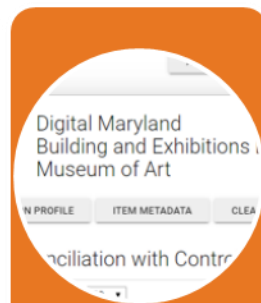
Publish linked data - FAST, VIAF, WorldCat (2009 - )



EntityJS Research Project (2013)



Person Entity Lookup Pilot (2014)



CONTENTdm Metadata Refinery (2015-16)



Project Passage (2017-18)

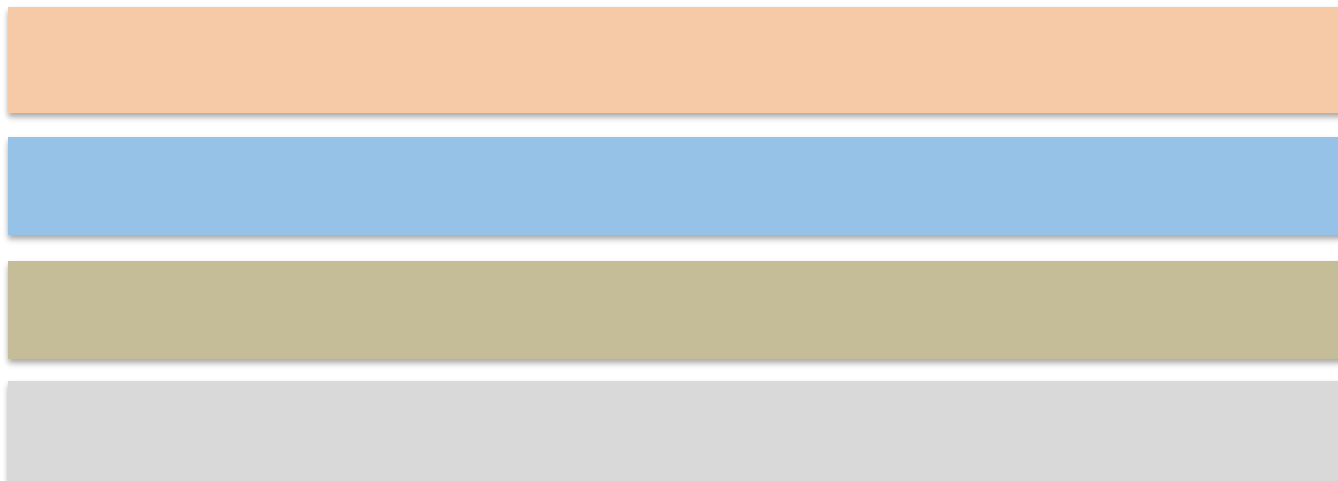
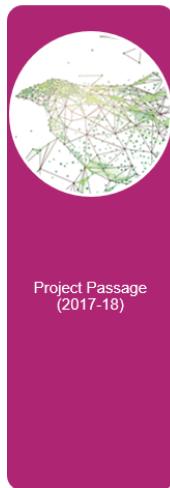


CONTENTdm Linked Data Pilot (2019-20)



Shared Entity Management Infrastructure (2020-21)

# A semantic continuum



# A semantic continuum



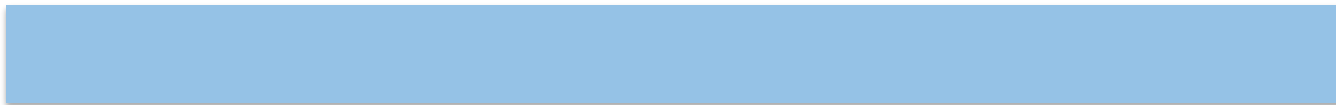
Project Passage  
(2017-18)



CONTENTdm Linked  
Data Pilot (2019-20)

**Shared, homogeneous,  
and centralized entities...**

**...accounting for the reality of localized,  
heterogeneous, de-centralized collections**





# A semantic continuum



Project Passage  
(2017-18)



CONTENTdm Linked  
Data Pilot (2019-20)

**Shared, homogeneous,  
and centralized entities...**

**...accounting for the reality of localized,  
heterogeneous, de-centralized collections**

**Machine-matching, highly  
automated reconciliation...**

**...with tools for hand-matching, semi-  
automated reconciliation**

# A semantic continuum



Project Passage  
(2017-18)



CONTENTdm Linked  
Data Pilot (2019-20)

**Shared, homogeneous,  
and centralized entities...**

**...accounting for the reality of localized,  
heterogeneous, de-centralized collections**

**Machine-matching, highly  
automated reconciliation...**

**...with tools for hand-matching, semi-  
automated reconciliation**

**Well-accepted context:  
Persons & Works**

**Granular context:  
About, Depicts, Annotations, Notes**

# A semantic continuum



Project Passage  
(2017-18)



CONTENTdm Linked  
Data Pilot (2019-20)

**Shared, homogeneous, and centralized entities...**

**...accounting for the reality of localized, heterogeneous, de-centralized collections**

**Machine-matching, highly automated reconciliation...**

**...with tools for hand-matching, semi-automated reconciliation**

**Well-accepted context:  
Persons & Works**

**Granular context:  
About, Depicts, Annotations, Notes**

**Blurs the line between  
bibliographic & authority work**

**Blurs the line between object description & context description**

# A semantic continuum



**Shared, homogeneous, and centralized entities...**

**...accounting for the reality of localized, heterogeneous, de-centralized collections**

**Machine-matching, highly automated reconciliation...**

**...with tools for hand-matching, semi-automated reconciliation**

**Well-accepted context:  
Persons & Works**

**Granular context:  
About, Depicts, Annotations, Notes**

**Blurs the line between bibliographic & authority work**

**Blurs the line between object description & context description**

**A large centralized infrastructure is needed.  
Custom applications and interfaces are needed.**

**This will facilitate new Knowledge Work in libraries.**

---

# WHY A LIBRARY INFRASTRUCTURE?

---

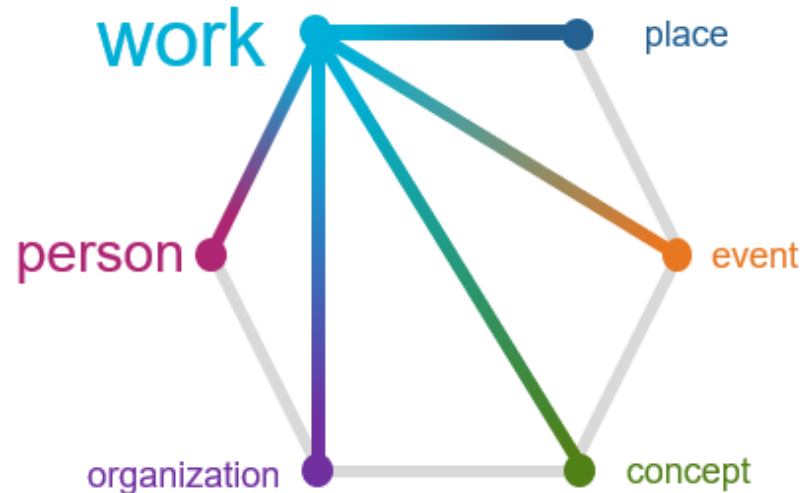
# Vision: Linked data for any library



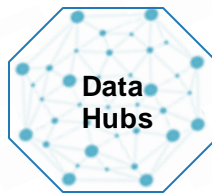
# So what are we building?



- An ecosystem for entities
  - Sustainable production infrastructure for libraries, at scale
- Metadata management tools for legacy and next-gen data
  - Workflow and integration support
  - Ability to ingest, transform, publish in (almost) any schema
- In partnership with libraries and cultural heritage institutions



**Metadata management & Workflow integration**

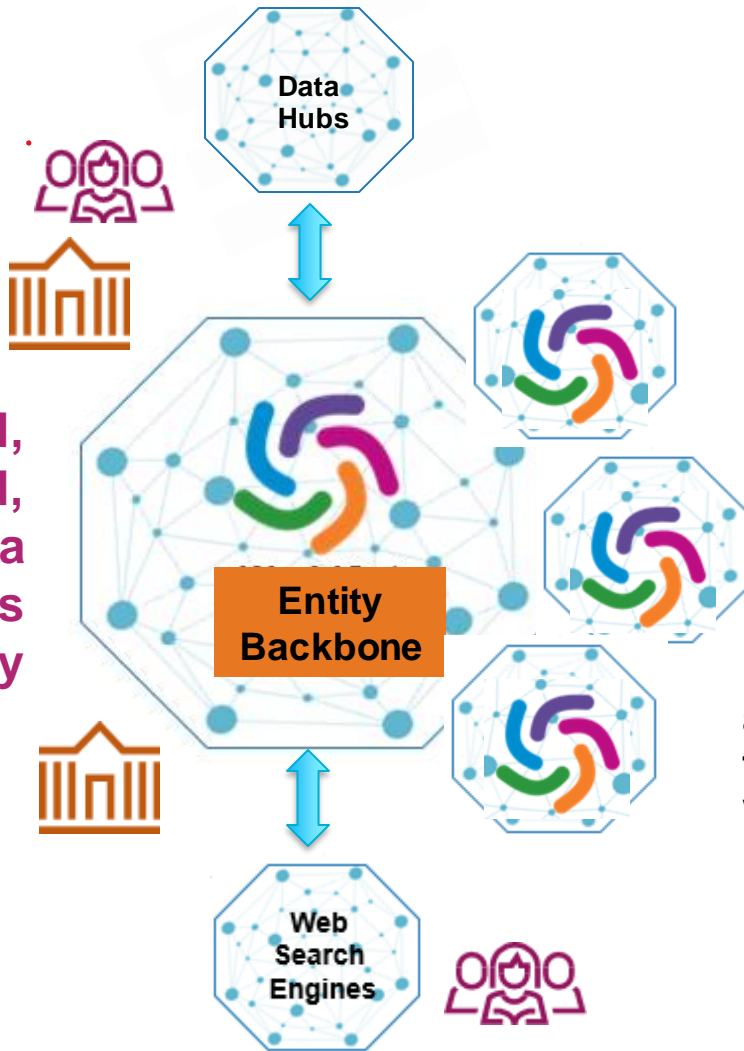


**Multi-lingual,  
interconnected,  
context-rich metadata  
for all material types  
and any library**



**Metadata management & Workflow integration**

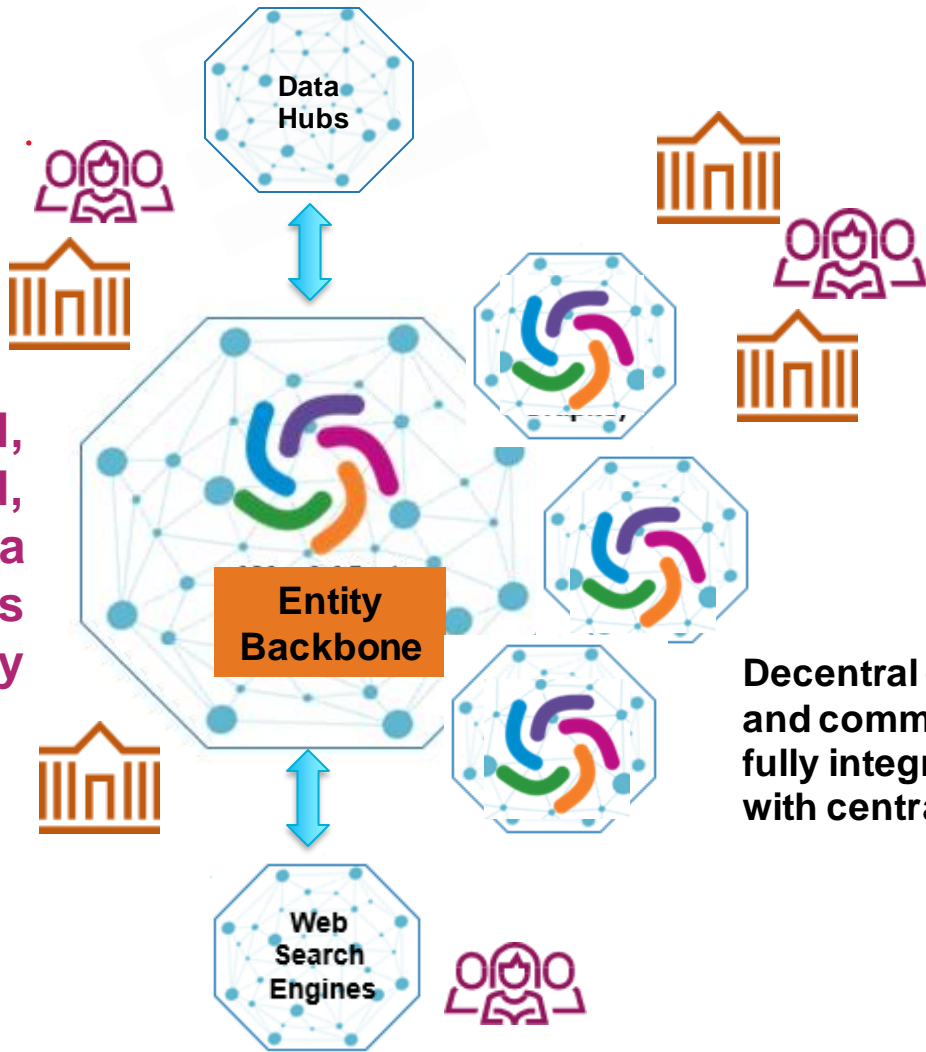
**Multi-lingual, interconnected, context-rich metadata for all material types and any library**



**Decentral graphs and communities, fully integrated with central graph**

**Metadata management & Workflow integration**

**Multi-lingual, interconnected, context-rich metadata for all material types and any library**



**Metadata management & Workflow integration**

**Decentral graphs and communities, fully integrated with central graph**

# Long-term Guiding Principles

- Persistent URIs (OCLC commitment)
- Optimized for storage, use and exchange of unique identifiers of ALL types
- Infrastructure tolerant of the heterogeneity of standards, practices, and metadata structures libraries need
- All libraries can benefit
- Continued support for and integration with MARC

# Project outputs...

## Entity infrastructure

- 150+ million entities w/ persistent URIs
- Production
- Foundational infrastructure



## Public website (free)

- Widely available
- No sign-in and no fee
- Search and view WorldCat Entities



## Entity APIs

- WorldCat Entity Search API (*search and view*)
- WorldCat Entity Management API (*search, view, create, and edit*)



## Basic user interface

- AKA SEMI UI
- Search, view, create, or edit WorldCat Entities



# ... and continued development

## Entity infrastructure

- 150+ million entities w/ persistent URIs
- Production
- Foundational infrastructure



## Public website (free)

- Widely available
- No sign-in and no fee
- Search and view WorldCat Entities



## Entity APIs

- WorldCat Entity Search API
- WorldCat Entity Management API



## Basic user interface

- Search, view, create, or edit WorldCat Entities



## Moving forward

- Development partner phase
- General release
- Additional functionality (ongoing)



# Stay tuned

- Stay tuned slide, mention future press release

# Relevance in the UK

- Governance and mission of OCLC extends globally
- Plan M and Jisc efforts – National Metadata Agreement
  - What's in it for libraries?
- Advocacy and Engagement – essential
- Community driven efforts will ensure success

Our work forms part of Jisc’s ‘Plan M’ - a vehicle being used to address the following issues identified by UK academic and specialist libraries:

Practical	Strategic
<p>Misdirected effort</p> <p>Having to look in multiple places to find suitable records or creating records where a suitable one already exists</p>	<p>Fragmented Infrastructure</p> <p>Having multiple delivery mechanisms and different working practices is inefficient and incoherent</p>
<p>Duplicated effort</p> <p>Different organisations and suppliers making the same or similar changes to a record</p>	<p>Uneconomic practices</p> <p>Precious resources and investment are being wasted</p>
<p>Unexploited value</p> <p>The inability to share enhanced records with other organisations who may be able to benefit from them</p>	<p>Sectoral transformation inertia</p> <p>The focus of cataloguing practice needs to shift towards automation of routine cataloguing and investment in special collections</p>



# Key Activities: Supply Chain Engagement



Establish a Strategic Metadata Taskforce to lead on engagement, advocacy and action: consisting of strategic representatives from across the supply chain



Work with library metadata practitioners to define a 'fit for purpose' metadata standard which clearly specifies consistent requirements valid throughout the supply chain



Engage with primary publishers through their library advisory groups and by convening a Round Table event to open a dialogue about collaboration



Develop pilot projects with key publishers to share their metadata openly



Work through Jisc Collections and regional procurement consortia to leverage dialogue with secondary publishers



Link up with existing standards bodies helping to connect experts and stakeholders

# Questions?

## **Fiona Carr**

Sales Account Manager, UKI - South

[fiona.carr@oclc.org](mailto:fiona.carr@oclc.org)

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