



Achieving Better Content: Leading Metadata to Higher Standards

Tamar Ganor | Content Product Manager
Christine Stohn | Discovery Product Management

Agenda



The Content Mountain - a story of growth



Growth challenges



Keeping up with the content



Data improvements in CDI to enhance discoverability and user experience



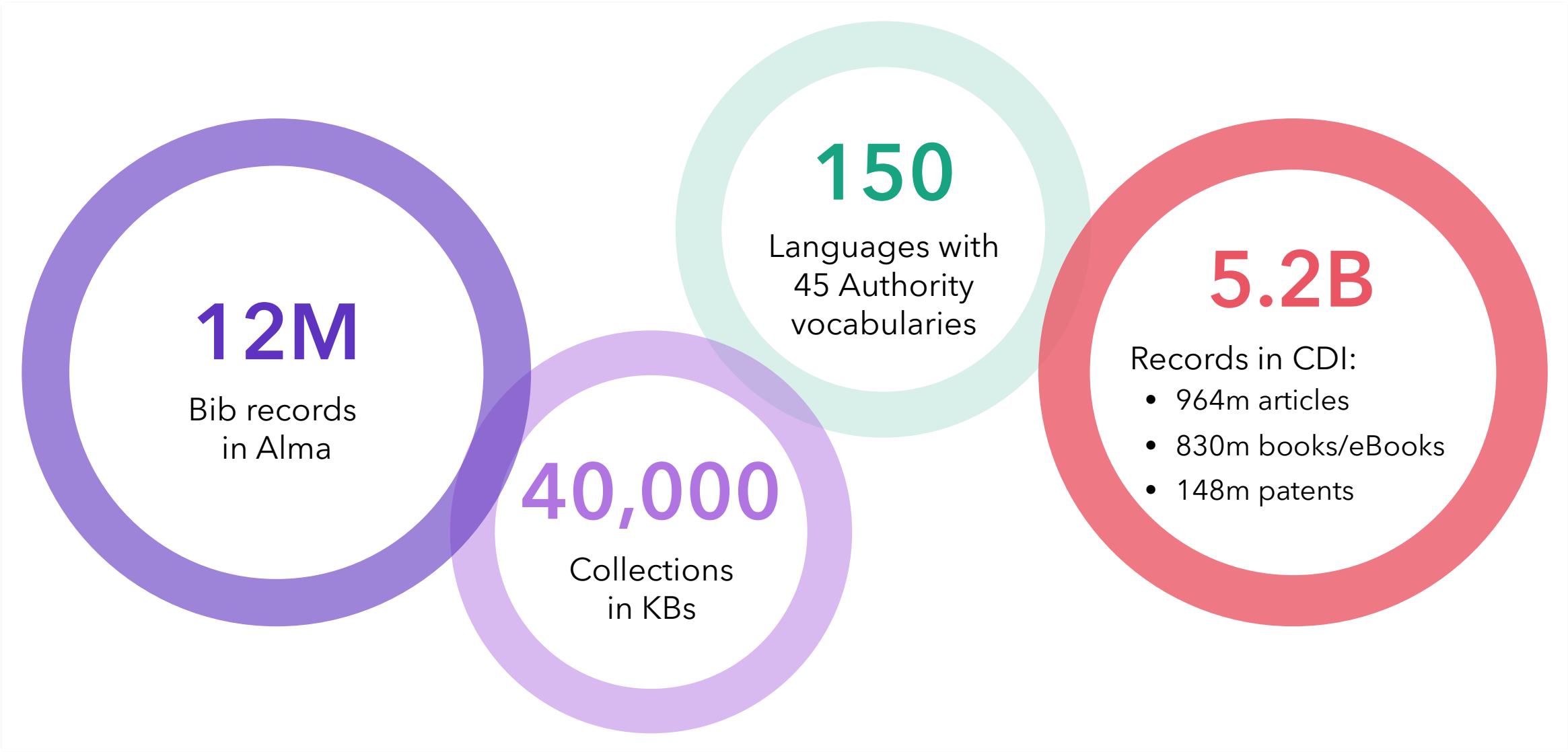
Data flows



The Content Mountain

Depth, Breadth and Perspectives

The Content Mountain



Let's talk metadata



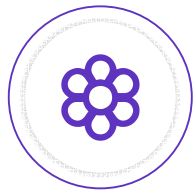
Highly Activated

44 collections with over 1,000 activations



Local Representations

1,667 local collections in CDI



Great Variety

8,408 collections with fewer than 10 activations



Content Providers

Over 5,000 in CDI



Open Access

333M OA items in CDI



Authority Vocabularies

57.1M Authority records in Alma CZ

Growth Originated Challenges

- Staying current
- Metadata quality





The Way to Data Excellence

Data Excellence



**Capacity, Efficiency,
Speed and Quality**



Quicker Updates

Increase update frequency of Alma



Better Quality

Improve and expand MARC metadata in Alma CZ



Improved Discoverability

Enhance linking and rights for CDI users

Transparent communication and collaboration with community and providers

Data Excellence: Updates Challenge



Daily Ingest

- Reduce turnaround time to daily updates
- 20 million portfolios affected
- ~165,000 activations
- By end of 2023:
70% of automated CZ portfolios will be updated daily

Create automation, improve ingest process,
clean up collections

Data Excellence: Quality Challenge

Better metadata through various sources

- 1 Selecting highly activated eBook collections:
 - So far over 153K Bib records enriched
 - Planning to enrich ~200K Bib records by end of 2023
 - Over 20,500 activations

Data Excellence: Quality Challenge

2 Cooperating with providers

- Reviewing available metadata
- Increase capacity of Provider Relations team

3 Alternative sources enrichment

- LoC print records
- Project Gutenberg
- Future prospect - AI



Data Improvements in CDI to Enhance Discoverability and User Experience

Document Attributes (Secondary Resource Types)

Definition



- **Attributes** associated with a record to provide more information
- Does not replace but **adds value to** the existing 'primary' resource type

Purpose



- **Display additional information** about the record/item (display flag)
- Used in search and as filters (facets)

Priority 2023



- *Review article*
- *Primary source*
- *Preprint*
- *Retracted publication*
- *Retraction notice*
- *Withdrawn publication*
- *Withdrawal notice/ item withdrawal*

1



ARTICLE

Music Therapy for Children With Autistic Spectrum Disorder and/or Other Neurodevelopmental Disorders: A Systematic Review

Mayer-Benarous, Hanna ; Benarous, Xavier ; Vonthron, François ; Cohen, David

Frontiers in psychiatry, 2021, Vol.12, p.643234-643234



PEER REVIEWED



OPEN ACCESS

REVIEW ARTICLE



Get PDF >



Available Online >



1



ARTICLE

Structural Evolution and Onset of the Density Wave Transition in the CDW Superconductor LaPt_2Si_2 Clarified with Synchrotron XRD

Nocerino, Elisabetta ; Lorenzo, Irene San ; Papadopulos, Konstantinos ; Medarde, Marisa ; Lyu, Jike ; Klein, Yannick Maximilian ; Minelli, Arianna ; Hossain, Zakir ; Thamizhavel, Arumugam ; Lefmann, Kim ; Ivashko, Oleh ; von Zimmermann, Martin ; Sassa, Yasmine ; Månsson, Martin

2022; 2022



PEER REVIEWED



OPEN ACCESS

PREPRINT



Available Online >





Tweak your results

Expand My Results

Search in Full Text

Sort by Relevance

Availability

Resource Type

Articles (8)

Newsletter Articles (1)

Retracted Publication (1)

Retraction Notice (1)

Newspapers search

Subject

0 selected 1-9 of 9 Results Personalize

1



ARTICLE

A 25-year experience with Carpentier-Edwards Perimount in the mitral position

Ayegnon, Kouakou Gregoire ; Aupart, Michel ; Bourguignon, Thierry ; Mirza, Alain ; May, Marc-Antoine ; Marchand, Michel

Asian cardiovascular & thoracic annals, 2011, Vol.19 (1), p.14-19

“ Data of 401 patients who underwent mitral valve replacement with the Carpentier-Edwards Perimount bioprosthesis between 1984 and 2009 were evaluated... ”

RETRACTED PUBLICATION

Available Online

2



ARTICLE

Retraction. "A 25-year experience with Carpentier-Edwards Perimount in the mitral position"

Asian cardiovascular & thoracic annals, 2011, Vol.19 (3-4), p.292-292

RETRACTION NOTICE

95 results

Sorted by **relevance** ▾

 This publication has been retracted/withdrawn
Notice of retraction/withdrawal









1  **Correction to 'Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*...**
by Abramson, J. Z.; Hernández-Lloreda, M. V.; García, L. ; More...


Peer-Reviewed 
Open Access 



 Quick Look Proceedings of the Royal Society. B, Biological sciences, 02/2018, Volume 285, Issue 1873



Journal Article  PDF  Full Text Online

2  **Comments on "Killer whale (*Orcinus orca*) behavioral audiograms" [J. Acoust. Soc...**
by Heffner, Henry E.; Heffner, Rickye S.

Peer-Reviewed 
Primary Source 
Preprint 
Review Article 
Open Access 

 Quick Look Branstetter and his colleagues present the audiograms of eight killer whales and provide a comprehensive review of previous killer whale audiograms. In their...

Journal Article  PDF  Full Text Online

Citations  Cited by 



Documentation is available on the Knowledge Center

[https://knowledge.exlibrisgroup.com/Primo/Content_Corner/Central_Discovery_Index/Documentation_and_Training/Documentation_and_Training_\(English\)/CDI_-_The_Central_Discovery_Index/CDI_Document_Attributes](https://knowledge.exlibrisgroup.com/Primo/Content_Corner/Central_Discovery_Index/Documentation_and_Training/Documentation_and_Training_(English)/CDI_-_The_Central_Discovery_Index/CDI_Document_Attributes)

OA Handling for Hybrid (and other) OA Articles



Handling of hybrid Open Access articles in CDI

- DBID level collections can be activated for some publishers (where we have reliable flags on article level), including SAGE, Emerald, OUP, IEEE, Taylor & Francis, Springer (more will be added)



NEW: we are indexing Unpaywall

- Can be activated as DBID level database
- Can be used as Link in Record resource or/and in combination with the API
- **Available in November**

Other CDI Data Enhancements in 2023

- **New resource types** (for Primo): magazines, magazine articles, questionnaire
<add release date and affected collections>
- Improve consistency of **language** and **publication date** fields between Index and display
- Better identification of **book reviews** by a set of rules



Subject Heading Normalization Project



Three Fields

- **Subject:** populated from controlled vocabulary, deduplicated and normalized
- **Keyword:** all subjects/keywords that cannot be mapped to the subject field
- **LATER Discipline:** all discipline level subjects, according to a controlled list



Iteration 1

- Subject and keyword field
- LCSH, MeSH, and some additions from PQ thesauri
- Optional: opt-in by activating it - **November** (Summon), **February** (Primo)
- Test option (for the facets) available for Primo customers **&searchCDI=normalizedSubjects**

Subject ^ Old

Facet

- Education (129)
- Music (121)
- Philosophy Of Mind (83)
- Compilation Of References (8)
- Functionalism (80)
- Globalization And Language Teaching
- Publish-Or-Perish Syndrome
- American Philosophy In The Twentieth Century
- Cognitive Neuroscience And The Theory And Practice Of Social And Political Inquiry
- Computational Functionalism
- Consciousness (74)
- Developmental Psychology: Philosophy, Concepts, Methods

Subject ^ New

- Education (908)
- Families (612)
- Gender Identity (574)
- Rites And Ceremonies (568)
- Sex Role (563)
- Social Problems (536)
- Holidays (513)
- Interpersonal Relations (513)
- Fiction (455)
- Employment (436)
- Folklore (432)
- Manners And Customs (417)
- Recreation (415)
- Music (379)

&searchCDI=normalizedSubjects

- Test facets and subject field in full record
- No key word field is displayed with the test option

Full view

- DJing secondary school > Old
- Edexcel Music Resources >
- Education, Secondary >
- enabling students music >
- GCSE Music >
- Great Britain >
- Instruction and study >
- KS3 Music >
- master level music >
- MCing secondary school >
- Music >
- Music & Education >
- OCR Music Resources >
- PGCE Music >
- PGCE Music books >
- Secondary Education >
- teaching music >
- Turntablism >

- Great Britain > New
- Music >
- Instruction and study >
- Turntablism >
- Education, Secondary >

All other subjects are going into the keyword field (visible in the PNX)

Another Important Data Project



Author identification

- Enrich CDI with **ORCIDs** and connect to author names where available, to identify authors better



In future

- Use additional sources for author identification

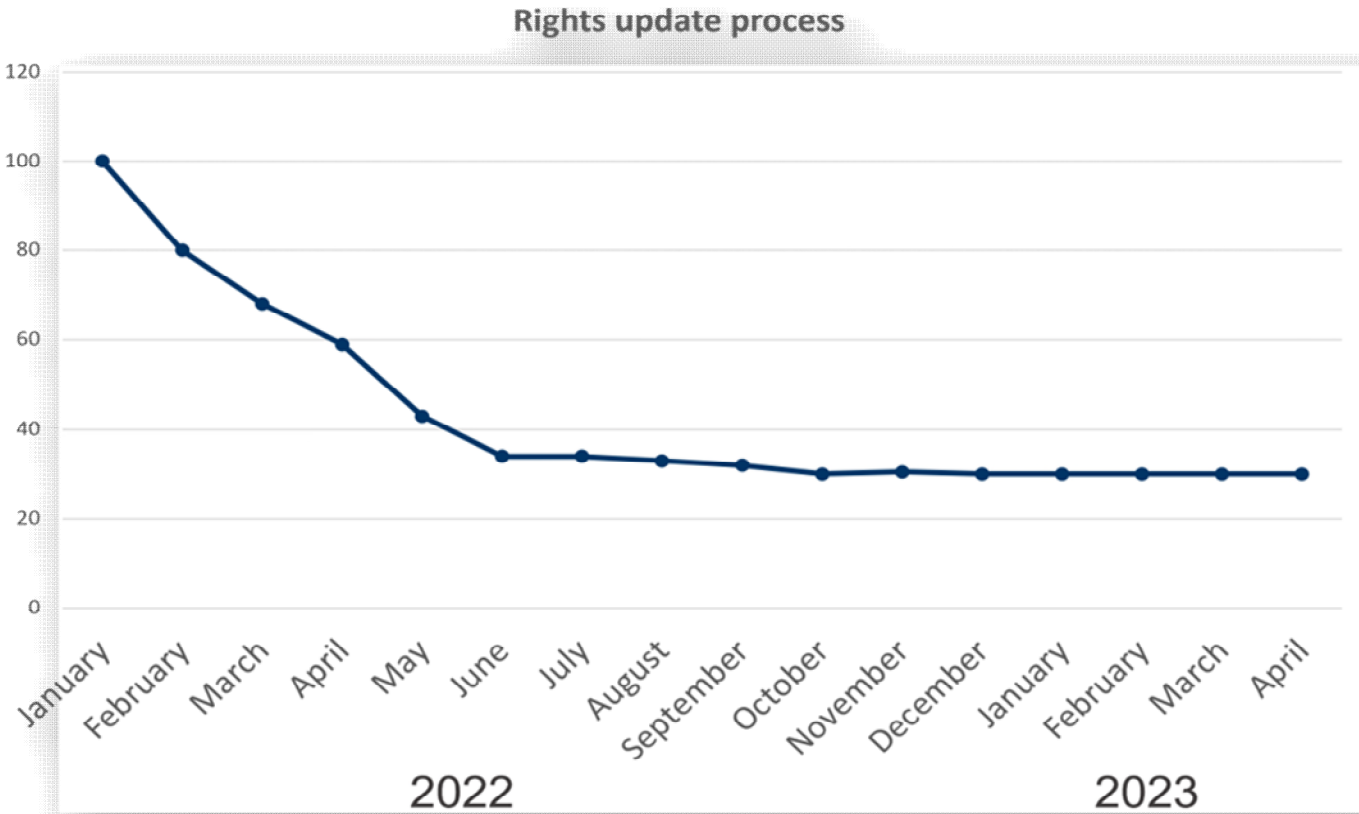
Both the subject and the author project help us to move towards a Linked Data Network in CDI



Data Excellence for Rights and Access

Quicker Availability with Faster Rights Processing

From 100 hours to an average of 30 hours to process activation to discovery



Fulltext Access Enhanced: Linking Improvements

○ Increased use of document ID for linking accuracy

○ Focus on **Quicklinks**

- Now with 37 supported providers
- Over **16,304** collections with **1.56 billion** records
- Linking success rates 99%
- Over 400 institutions currently active or testing
- More than 1.3m clicks in March 2023

More providers coming in 2023

◀ Quicklinks in Primo and Primo VE | Search Configurations ▶ [Table of contents](#)
[No headers](#)

Quicklinks enables users to go directly to the full text with one click. For more information about Quicklinks, see the following articles:

- Quicklinks in Summon
- Quicklinks in Primo and Primo VE

The following table lists which providers currently support Quicklinks and is updated as providers are added or changed.

Provider	Content	Type of Quicklink (PDF/HTML)	Last Updated
ACM Digital Library	All	PDF	
Elsevier	All	HTML	
ProQuest	All Content excluding eBooks on the ProQuest Platform: https://www.proquest.com	HTML (in most cases a link to download the PDF full text is included with the article)	
SAGE Publications	Journals under https://journals.sagepub.com	PDF, HTML	
Springer Nature	Journals and Books under https://link.springer.com	PDF, HTML	
Wiley-Blackwell	Journals and Books under https://onlinelibrary.wiley.com	PDF, HTML	
BioOne	Journals Content under https://bioone.org	PDF	
IOP Publishing	All	PDF	
National Library of Medicine	Content which is part of PubMed Central collection under https://www.ncbi.nlm.nih.gov	PDF, HTML	
New England Journal of Medicine	Content under https://www.nejm.org	PDF, HTML	
Emerald	Journals and Books under https://www.emerald.com/	PDF, HTML	
Credo Reference	Content under https://search.credoreference.com/	PDF, HTML	March 27, 2023 - Added PDF
Thieme	All journal content besides conference proceedings under https://www.thieme-connect.de/products/journals	PDF, HTML	
IEEE	Content under https://ieeexplore.ieee.org	HTML	
Mary Ann Liebert, Inc.	Journals under https://www.liebertpub.com/	PDF, HTML	
SPiE	Content under https://www.spiedigitallibrary.org/	PDF, HTML	
University of Chicago Press		PDF, HTML	
Royal Society Publications	Content under https://royalsocietypublishing.org	PDF, HTML	



The Data Flow

Where is the Data Coming from... and Where Does it Go?



Providers



Through the
Products



End-User
Functionality



In Summary

- **Data Excellence** (DX) is the main initiative for Ex Libris to achieve better content

- We are leading metadata to higher standards using **internal and external solutions**

- Pushing selective efforts to improve **quality**, get quicker **updates** and enhance **availability**

- **Community + Providers + Ex Libris**

Please join the Q3 content webinar, for a discussion with Sage and Oxford representatives.

Metadata at its Origin -

Publishers Talk about Metadata

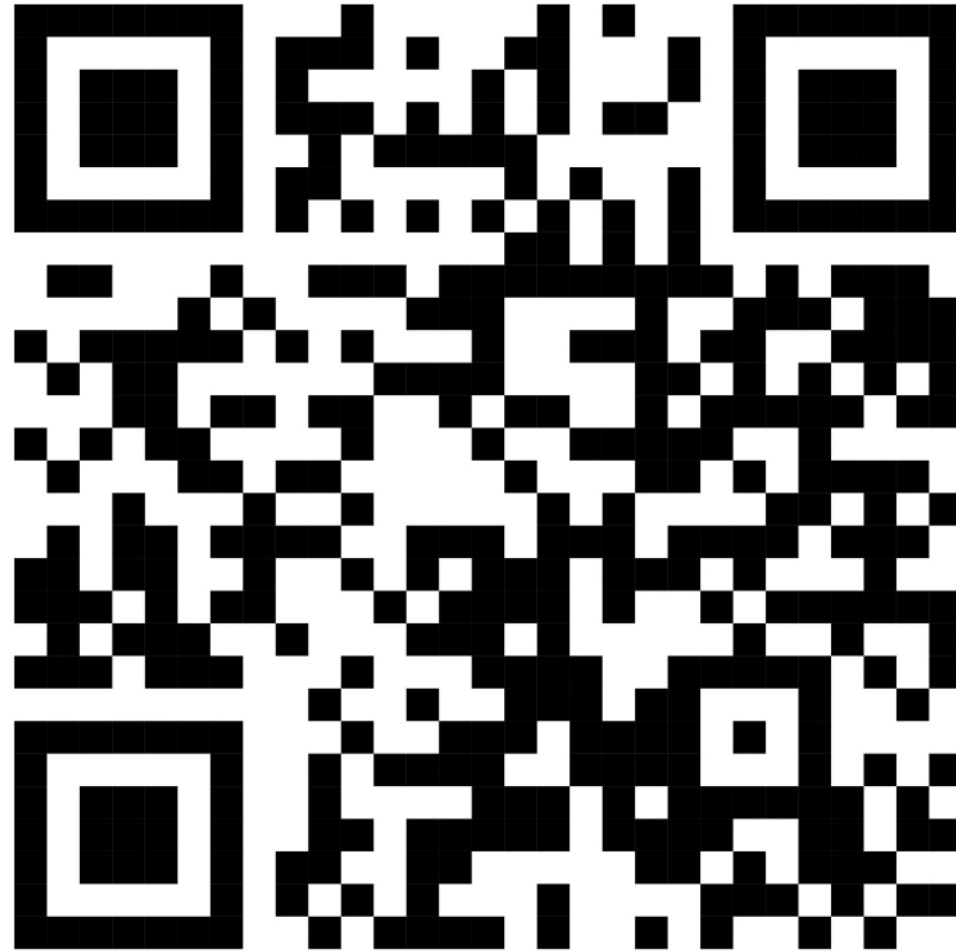
Thursday, September 28th

15:00 GMT

*Feel free to send your questions by September 22nd



Not yet registered for the content mailing list??



Don't miss out!!



Thank You!

tamar.ganor@clarivate.com

christine.stohn@clarivate.com

© 2023 Clarivate

Clarivate and its logo, as well as all other trademarks used herein are trademarks of their respective owners and used under license.