

The background features a dark blue gradient with abstract, wavy lines in shades of purple and blue. Small, glowing dots in blue and purple are scattered across the scene, some connected to the wavy lines by thin vertical lines, creating a digital or network-like aesthetic.

Monitoring Transitional Agreements: the challenges and successes of implementing article level metadata collection

Beth Harris

Licensing Intelligence Team

Jisc agreements in 2021



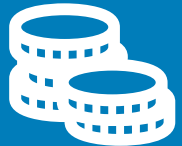
31,000

UK articles published
immediately OA under Jisc
agreements



287 million

articles read by UK
member organisations



£156 million

is spent on subscriptions and publishing of digital
content by Jisc subscribing institutions

What is this session about?

As of 2022 Jisc offers 40+ open access agreements (including TAs, S2O, Native OA and Compliant green)

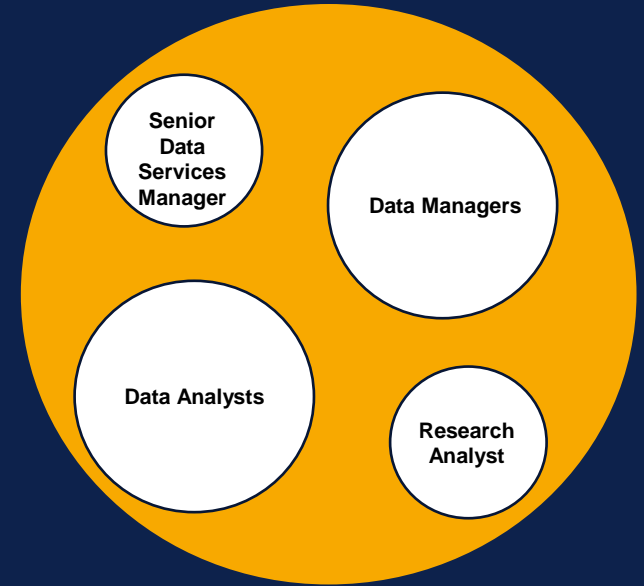
- We need to be able to monitor, assess and evaluate them:
 - Are they meeting institutional needs for read and publishing access?
 - Are they enabling compliance with funder mandates?
 - Are they cost-effective for library budgets?
 - Are they facilitating the transition to open?
- This session shares Jisc's work to implement a data collection schedule on publishing output and to combine datasets to produce analysis to evaluate TAs.

Who are we? What do we do?

Licensing intelligence team at Jisc

New team (est. Jan 2021) who:

- Collect data from content providers
- Standardise and verify datasets
- Combine this with additional datasets
- Produce data in a usable format
- Demonstrate value and impact

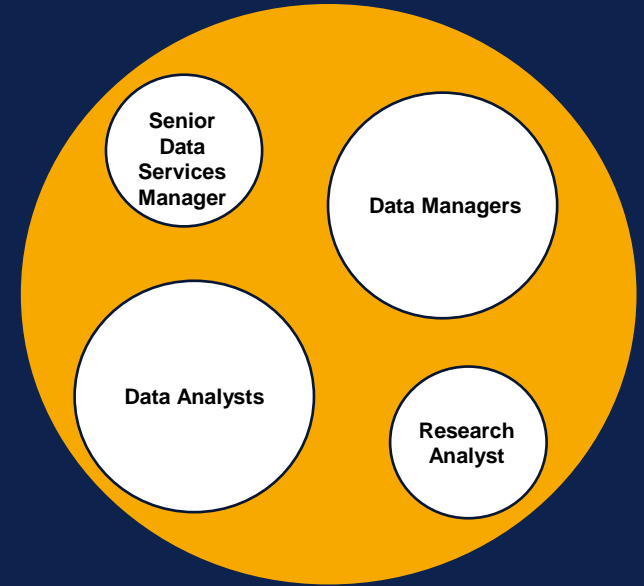


Who are we? What do we do?

Licensing intelligence team at Jisc

Some of our team aims include:

- Developing comprehensive quality assured data.
- Provide flexible products and tools to support data access and reuse.
- Develop strategic partnerships and engaged communities to support effective data access and reuse.
- Establish common frameworks and a standards-based infrastructure to support interoperability.



History of ALM – KE project

- A 2019 Jisc collaboration with the Knowledge Exchange demonstrated the difficulties in monitoring the publishing element of TAs:
 1. Lack of consistency in metadata elements provided across publishers
 2. Inefficient workflows between publishers and consortia
 3. Inability to compare agreements across publishers
 4. Reports from publishers received at different frequencies.
 5. Burden on institutions to evaluate a TA's cost and value.
- The outcome of the project was the creation of the Knowledge Exchange OA group checklist which was repurposed into a template (available on Zenodo) that was designed to collect all output under TAs as a standard.
 - <https://zenodo.org/record/3407214#.YWP4A9rMKUI>

Data collection and Jisc

Data collection

- Contact publisher with our article level metadata template (32 fields, 20 are mandatory)
- Discussion with the publisher to establish:
 - Details about the publisher's data workflows
 - Answers to metadata queries
 - Reporting schedule (monthly/quarterly)
 - Receiving test data
- Include data requirements, as agreed, into the license
- Data received from 34 publishers with TAs from 2019-2021.

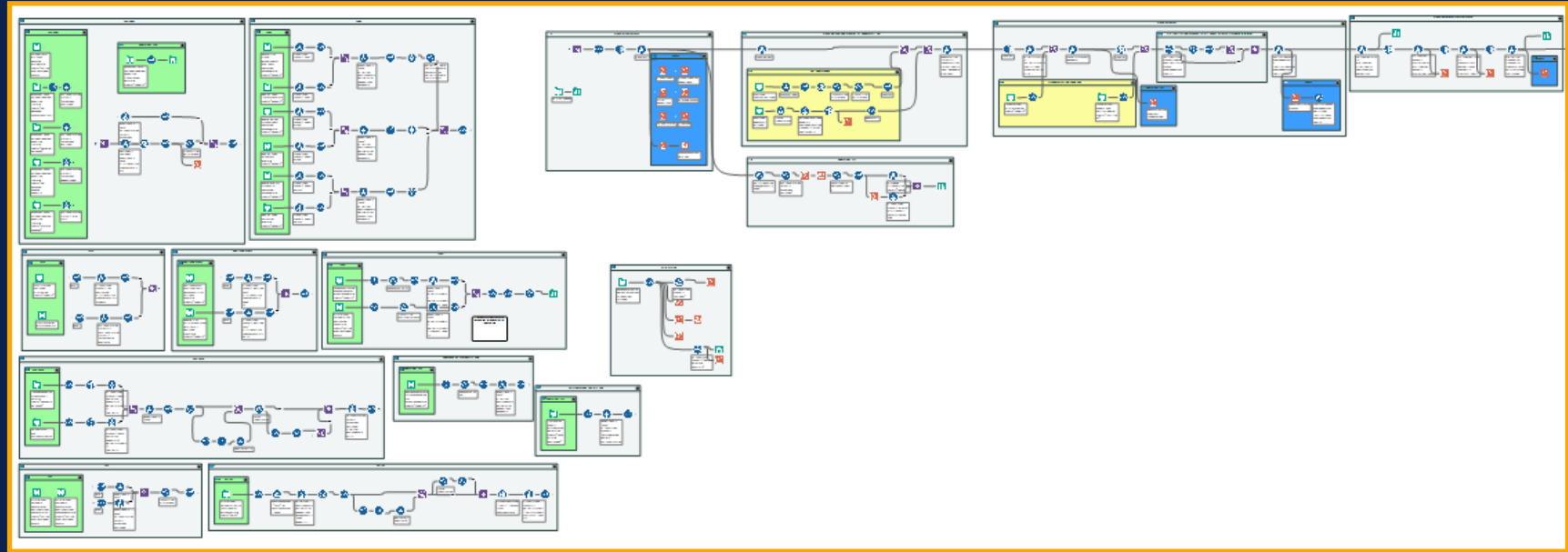
OA article-level metadata template for publishers

Note:
Publishers are requested to fill this template and send it via email to the consortia as well as to all the academic institutions that are part of the consortium. Publishers are meant to deliver the OA article-level metadata report on a monthly basis (e.g. on the 15th day of every month) and the OA report should be submitted to the consortia.

Year of agreement (start and finish date):
Name of agreement:

#1: DOI	#2: Is the article Open Access?	#3: Institution Name	#4: Article Title	#5: Article Type	#6: Journal ID (publisher ID)	#7: Journal Title	#8: Journal Subject / Discipline	#9: Journal ISSN	#10: Journal E-ISSN	#11: Article License (CC license)
http://doi.org/xxxxx	Yes/No	University of xxxxx	xxxxx	e.g.		xxxxx	xxxxx	xxxxx	xxxxx	e.g. CC BY
http://doi.org/xxxxx	Yes/No	University of xxxxx	xxxxx	e.g.		xxxxx	xxxxx	xxxxx	xxxxx	e.g. CC BY
http://doi.org/xxxxx	Yes/No	University of xxxxx	xxxxx	e.g.		xxxxx	xxxxx	xxxxx	xxxxx	e.g. CC BY
http://doi.org/xxxxx	Yes/No	University of xxxxx	xxxxx	e.g.		xxxxx	xxxxx	xxxxx	xxxxx	e.g. CC BY

Data Cleaning, Standardisation and Verification



Data Cleaning, Standardisation and Verification

1. Cleaning the DOI

- Remove hyperlinks, whitespace, unnecessary punctuation, check format.
- Why – to run the DOI past external APIs such as Crossref and Unpaywall to obtain additional information such as funder information and OA status.

2. De-duplicating DOIs

- Identify any duplicate records
- Work with publisher to understand duplication(2 institutions co-authored the paper or a genuine error) and agree method for dealing with such duplications internally.
- Why – to ensure we're not double counting articles.

3. Institution Name

- Institutions can have many names or can be known by an acronym, so we standardise these to the legal name used by Jisc's internal systems. We also add PIDs (Ringgold and Jisc ID).
- Why – this helps us check the publishing institution is subscribed to the agreement and allows for accurate institutional level analysis.

Data Cleaning, Standardisation and Verification

4. Currency

- Ensure where the APC list price is included the currency is also included. We then standardise this to GBP.
- Why – this allows for us to total APC list prices to show overall value/ cost in one standard currency.

5. Article Type

- Article types are mapped across publishers to the COAR 3.0 standard
- Why – to ensure we have a coherent article type across all publishers

6. License Type

- License types are mapped across publishers to the CC BY standard
- Why – enables us to monitor compliance across publishers.

Data Cleaning, Standardisation and Verification

7. Verify data

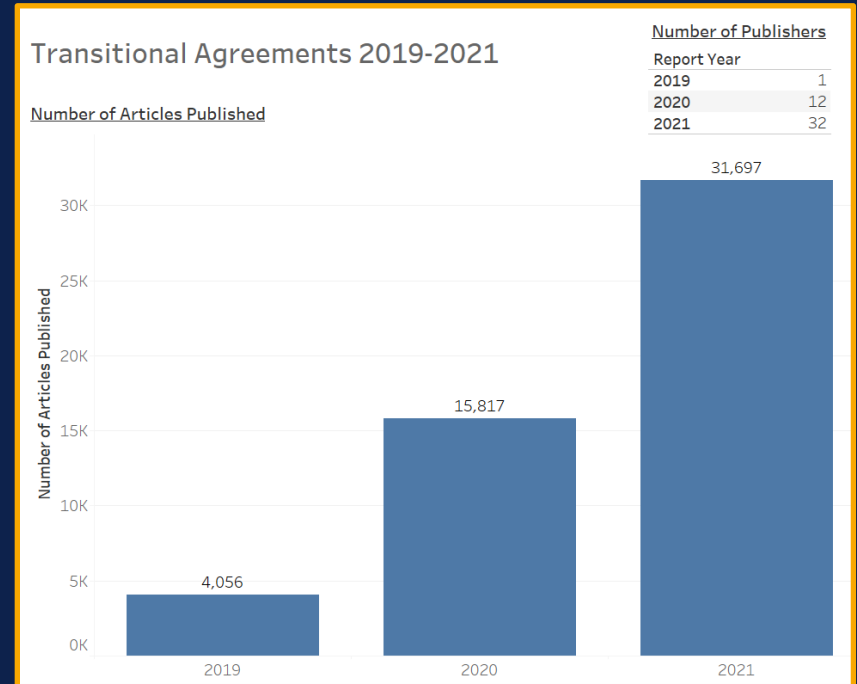
- Check that all institutions in a publisher's report are subscribed to the agreement.
- Why – ensures we exclude any records that have been sent in error.

8. Enhance the data

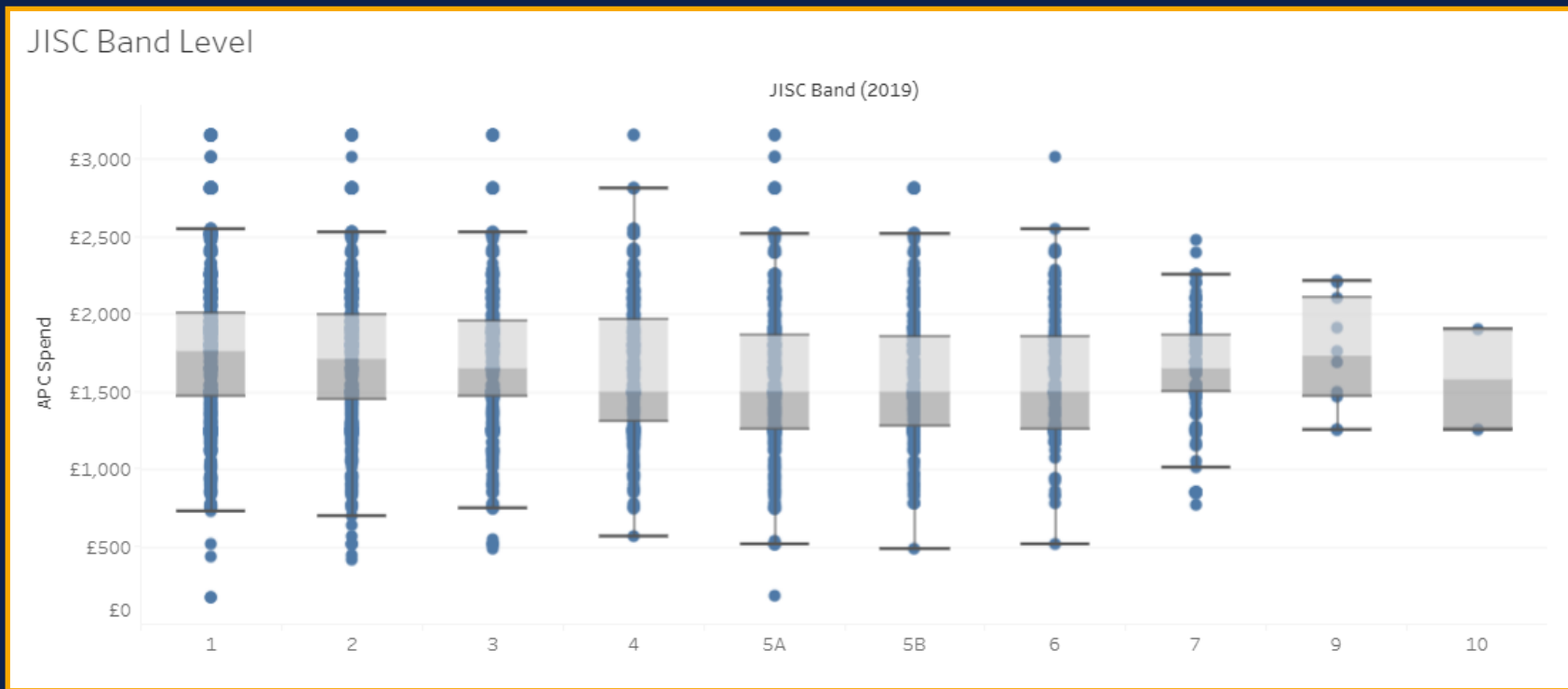
- Add Crossref funder fields and any missing metadata
- Add Unpaywall OA status and any missing metadata
- Add in Jisc band, mission groups, identify UKRI and ex-COAF funding.
- Why – this allows us to fill in any gaps such as article title, license type or article type and have more complete data and by adding other information such as mission group we can complete more relevant analysis.

Outcome

A cleaned and standardised dataset on all articles accepted and published under Jisc negotiated Transitional Agreements



Outcome



Outcome

Investment per article = Total spend (reading + publishing) ÷ total number of OA published

Publisher	Investment per Article
Publisher 1	£2,293
Publisher 2	£3,346
Publisher 3	£2,876
Publisher 4	£3,258
Publisher 5	£2,967

Successes from the project

Use of PIDs

Standardisation

Good working
relationships

Understanding
of external
stakeholder's
workflows

Challenges



Limits of smaller publishers



Articles without a unique identifier (DOI)



Organization names



Article type and license type



Publisher workflows

Looking ahead



Collaboration between all stakeholders

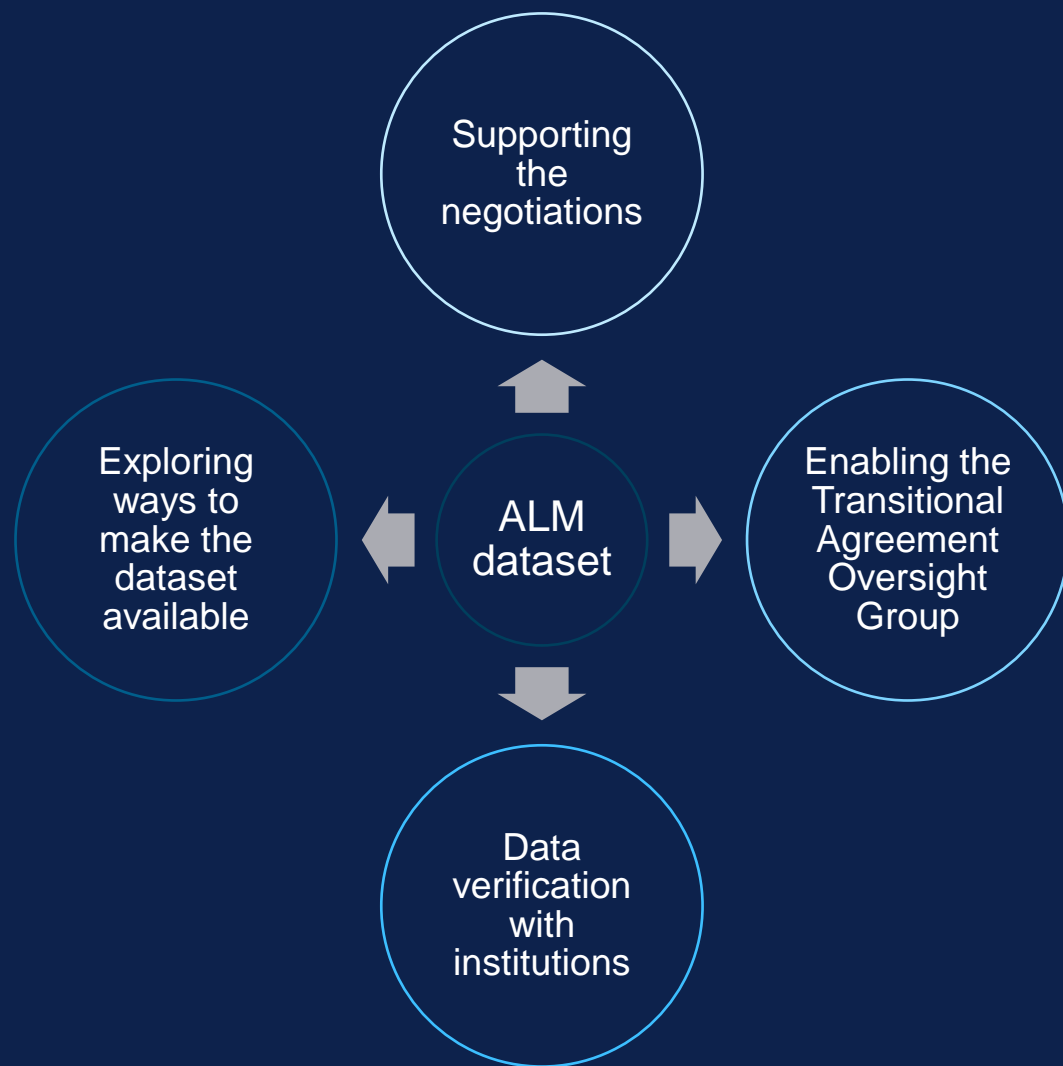


Adoption and engagement of suppliers



Encourage the adoption of PIDs

Next steps



Thank you for listening.

bethany.harris@jisc.ac.uk

