

# RECOMMENDATIONS FOR REMODELLING THE UK LIBRARY DATA MARKETPLACE

November 2020

# Executive Summary

- The development and delivery of an efficient bibliographic metadata supply model is essential for the 21<sup>st</sup> academic library.
- UK academic libraries spend considerable amounts of money on acquiring this metadata, but too frequently it is not fit for purpose.
- We can safely say that the landscape is messy, metadata flows from player to player, often being stripped of prior intellectual input
- This situation requires leadership and engagement with the entire library data marketplace.
- Jisc has a proven track record in achieving significant marketplace change and it is well placed to lead it by playing an outward looking and international role.



# 1. Introduction

- Jisc commissioned us to engage with stakeholders and to make recommendations about steps that could help the sector to realise efficiencies.
- Our task was to present one or more transformative but realistic recommendations that could introduce efficiencies and streamline the marketplace, and that would inform a practical pilot or other concrete next step.
- We also gratefully acknowledge members of the steering group at Jisc and the assistance of many individuals who contributed their expertise via interviews and focus groups or submitted financial data to the project team.



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>

# 2. Methodology

- We conducted 10 interviews with a mixture of higher education libraries, infrastructure providers, and specialist libraries.
- We developed a template for gathering detailed input from libraries - this exercise was focussed on the procurement and amendment of records, capturing both the cash and in-kind costs relevant to this work.
- Further interviews were undertaken with those that operate in the supply chain. These interviews explored current practice and business models, and appetite for new workflows and business models.
- We held an online focus group with members of the library community using the [Mentimeter Tool](#) to validate our findings, test assumption and proposals.



# 3. Defining the metadata marketplace

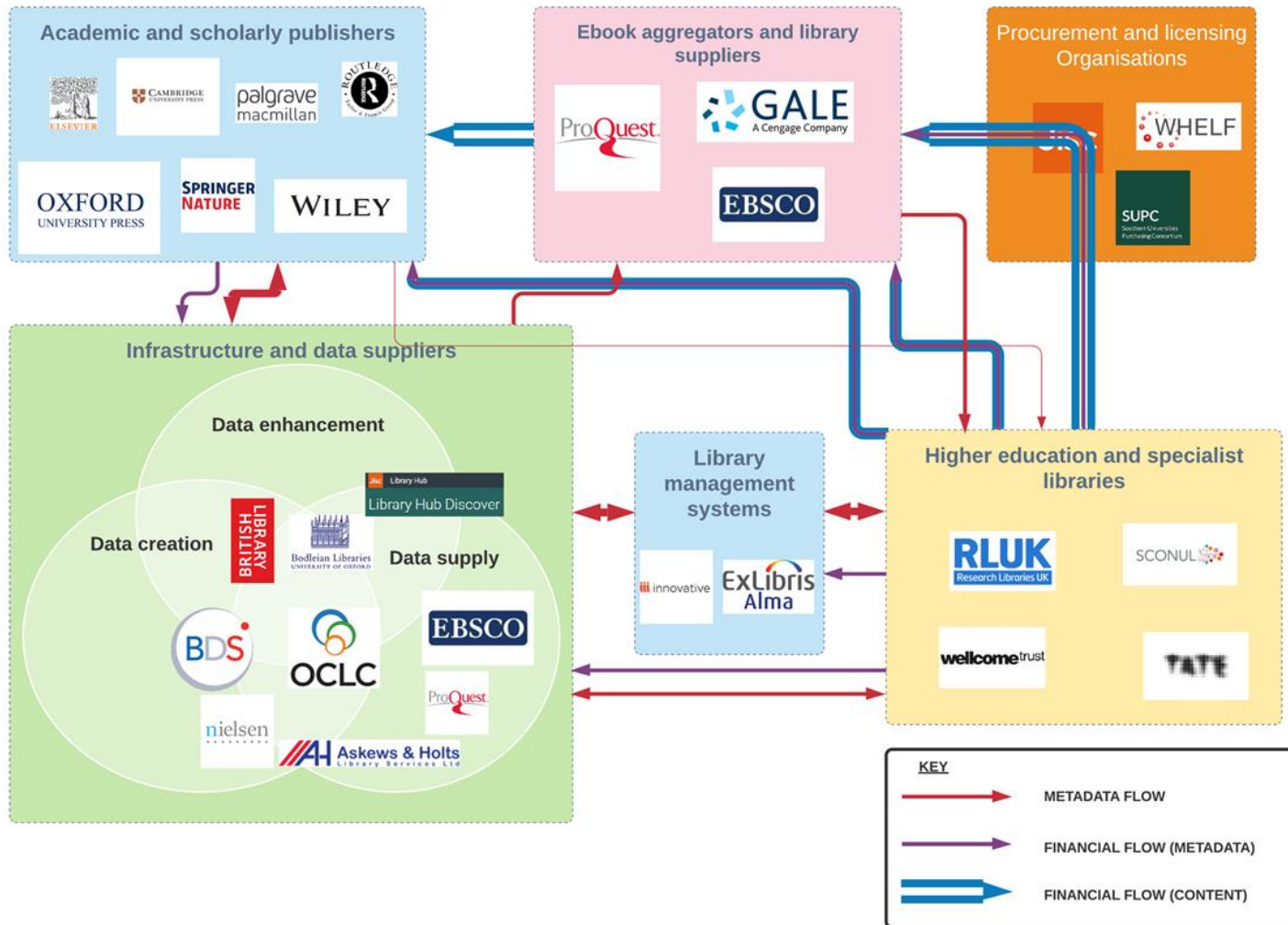
There is a wide spectrum of metadata of relevance to UK academic and specialist libraries, and we focussed on book bibliographic metadata for print and e-books.



# Overlapping marketplaces

- There are overlapping marketplaces for book bibliographic metadata because academic, educational, and trade books are included in library collections.
- The supply chain carries books from publisher to library, and serves not only academic and specialist libraries but also public libraries, national libraries, and more.
- In the library world the MARC standard has evolved over time, but in the retail supply chain the ONIX standard is deployed internationally, and national bodies exist to convene and coordinate supply chain partners.





# The supply chain is complex

- There are multiple intermediaries.
- Metadata and financial flows are not always linked.
- Many organisations fulfil multiple roles, particularly with regard to the creation of metadata.
- Library management systems are a critical source of metadata for libraries.
- The majority of financial transactions pass through E-book aggregators.
- Procurement and licensing organisations to date had limited involvement in metadata transactions.



# 3. Stakeholder overview and issues

We began our research with stakeholder interviews to understand the role for metadata in their organisations, and their aspirations and concerns with respect to book bibliographic metadata.



*We value better quality metadata. There was a time we thought it did not matter, but high-quality metadata gets the reader to what they want.*



# UK Academic and Specialist Libraries Pain points (1)

- Librarians spend a great deal of effort correcting book bibliographic metadata that they have purchased.
- They often cannot acquire the e-books they need because they cannot find them when they search e-book aggregator platforms.
- There is unease about 'double-dipping' - paying twice for the same metadata and/or different libraries paying twice for the same metadata.
- There was no clear consensus about what a reasonable first copy cost would be for bibliographic metadata record nor acceptable ways to recover these costs.



# UK Academic and Specialist Libraries Pain points (2)

- There are many channels through which metadata is made available, and none is viewed as entirely fit for purpose.
- Teasing out what the issues might be is made more challenging as primary publishers, secondary publishers, aggregators, metadata suppliers, library system providers, and other intermediaries were lumped together under the umbrella term “vendors”.
- There did not seem to be a clear understanding of how metadata is created, how it moves through the supply chain, or how it is licensed and is sold.



We also interviewed a wide range of stake holders across the supply chain

- LMS Providers
- OCLC
- British Library
- BDS
- Jisc
- Purchasing consortia
- Ebook aggregators
- Primary publishers





# 3. Forcefield analysis

We used this helpful input from stakeholders to inform a forcefield analysis. This is an approach to understanding drivers and obstacles to change, and the forces we believe are at play in the book metadata marketplace



## Driving Forces (Positive forces for change)

- Quality and timeliness concerns from librarians
- Jisc searching for ways to articulate needs, reduce friction and improve efficiency
- Range of other stakeholders also looking for cost effective ways to provide quality services in a more efficient and sustainable way
- Largest primary publishers are engaged and innovative about metadata creation and management
- Willingness to open content and infrastructure where there are sustainable business models
- Improving technology to automate the creation and exchange of dynamic metadata
- Changes in library management systems are driving changes in metadata
- Covid-19 crisis has accelerated the shift to e-books

## State of the current metadata landscape

## Restraining Forces (Obstacles to change)

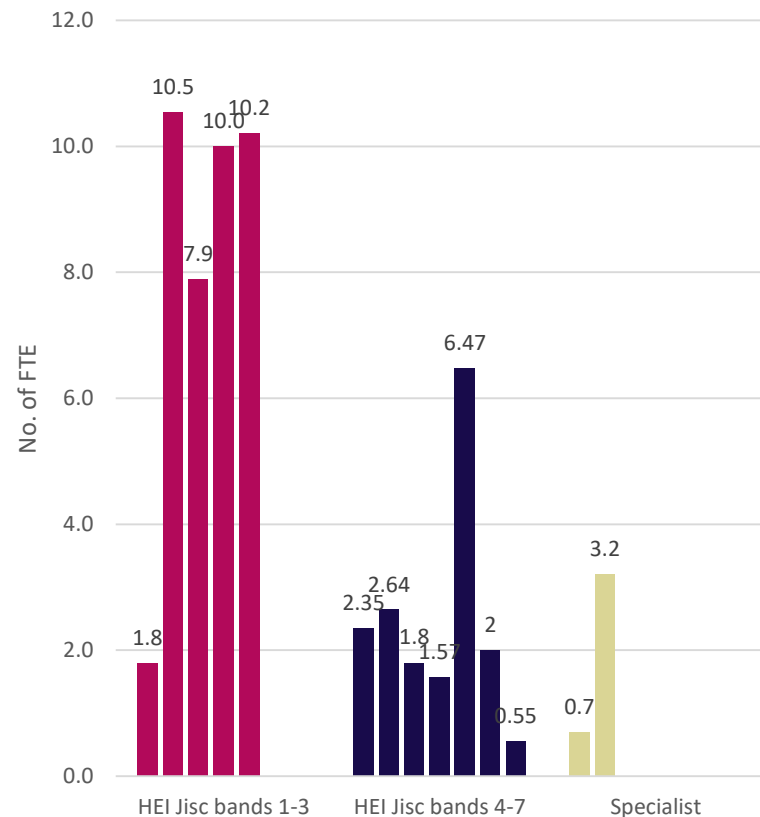
- No clarity on costs to create and curate dynamic metadata
- Limited understanding of all stakeholder roles and contributions
- Insufficient consensus about what librarians need vs. want
- Expectation that libraries should have metadata their way without paying a premium price
- Poor rights management and unclear licensing for metadata
- Defensive posture from many intermediaries whose core business models are under threat
- Need to invest in innovative technology and skills
- Complex, diverse stakeholder community
- Fragmented library communication to and influence on the *academic* library supply chain re metadata
- High quality metadata for open access content is often unavailable or difficult to source

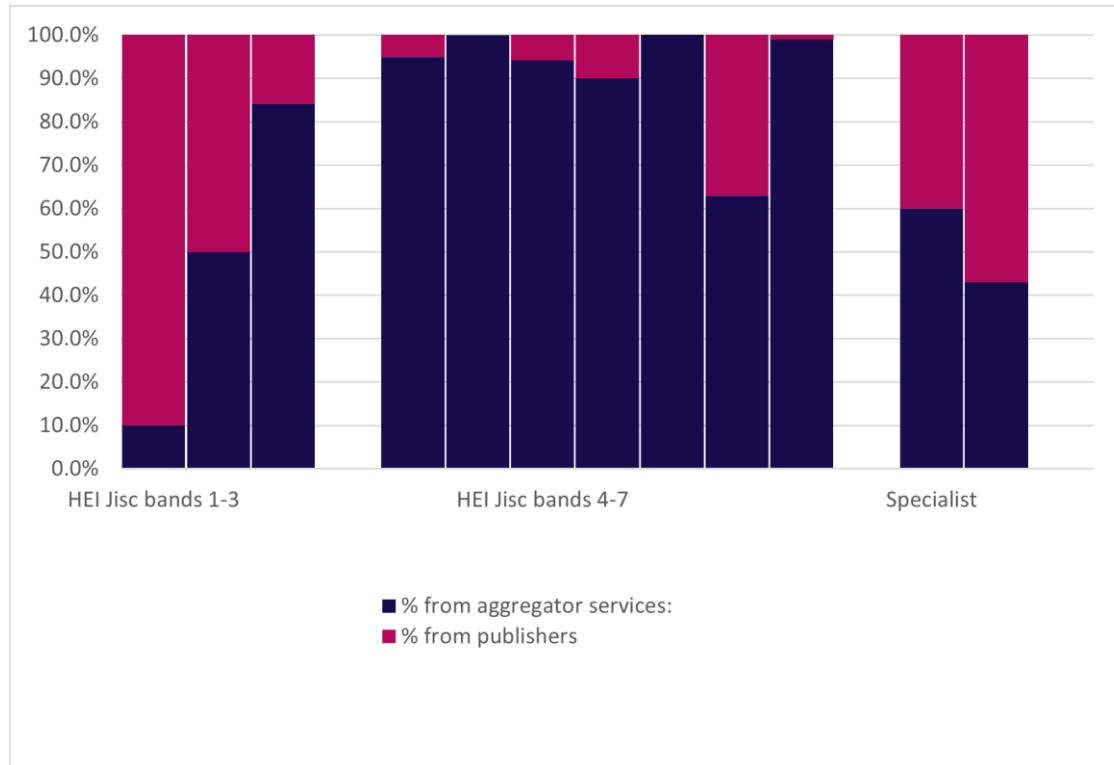
# 3. Financial Modelling

We sought information on costs and pricing models from other actors within the metadata marketplace, however we were unable to obtain sufficient comparable data from these actors.



- The sampled HE institutions suggests 4-6 million records are ingested per year across HE.
- There was not sufficient evidence to provide a sector wide view for price per item of metadata.
- Scaled up across the UK HE sector we estimate that metadata management involves ~400 FTE at a cost of £15-£20 million per annum.
- The headcount across the sector is estimated at over 1,000, as for many library employees metadata management is just part of their role.





For HE libraries in Jisc bands 4-7, there is a strong preference for acquiring e-books from aggregator services.

# 4. Library Focus Groups

Next we convened two library focus groups, one each for representatives of institutions in Jisc bands 1-3 and 4-7 to gain feedback on everything presented in these slides so far.



- There was a lot of support for the view that it is about opportunity cost and reducing the amount of time spent doing basic repair work.
- Articulating it in terms of adding value is a good way to do it. Not cost saving but it will free their time up to do things their Directors want them to do such as spending more time on identifiers and free up time start to add or pull in the value.
- Another example given was work on legacy data. This what they would like to do instead of making enhancements and firefighting all the time. They would like to be doing something for the community but currently don't have time.



*It's opportunity cost not financial saving. Less time on basic 'repair' type metadata work = more time working on using good data to drive usage, access and availability of what we buy to support Institution mission.*



- Participants ranked options.
- The combined results and average scores (out of 10) are presented right.

1	Influence the whole library supply chain – 8.4
2	Engage with primary and secondary publishers – 8.1
3	Engage with procurement organisations – 7.6
4	Engage libraries in additional ways –7.4
5	Engage more deeply with OCLC – 6.5
6	Engage more deeply with the British Library -6.2
7	Procure records from BDS – 4

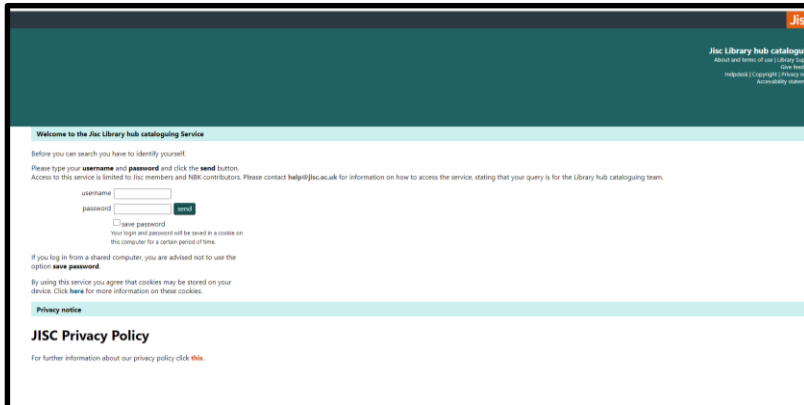




# 5. Recommendations

Our task was to present transformative but realistic recommendations that could introduce efficiencies and streamline the marketplace, and that would inform a practical pilot or other concrete next step.

# 1. Clear branding of NBK and its services in preparation for international engagement



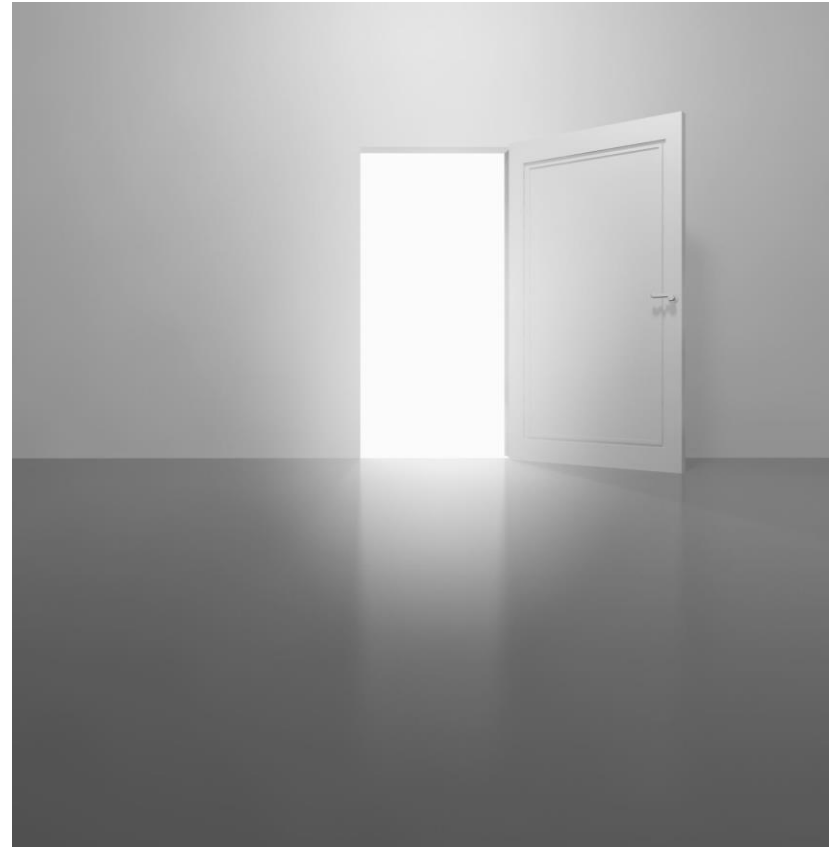
The screenshot shows the login page for the Jisc Library Hub cataloguing Service. The page has a dark green header with the Jisc logo in the top right corner. Below the header, there is a navigation menu with links for 'Jisc Library Hub catalogue', 'About and terms of use', 'Library Support', 'Data feeds', 'Helpdesk', 'Copyright', 'Privacy notice', and 'Accessibility statement'. The main content area is white and contains a 'Welcome to the Jisc Library Hub cataloguing Service' message. It instructs users to enter their username and password and click the 'send' button. There is a checkbox for 'I have password' and a note that login information is stored in a cookie. A 'Privacy notice' link is also present. At the bottom, there is a 'JISC Privacy Policy' link.

- This is a global game with national pieces, crossing into the domain of national and public libraries and the commercial book supply chain.
- In our interviews with international players there was no awareness of NBK or Jisc's important role in providing this national infrastructure.
- Raising awareness of Jisc's role in providing these services is important, because it demonstrates Jisc's expertise in the area.
- Careful branding and communication of these services will help Jisc to position itself for an international leadership role .



## 2. Leadership and the supply chain

- Jisc could ultimately evolve the NBK to provide a centrally managed dynamic metadata store for open metadata that was integrated with the community's LMS and discovery layer systems.
- To achieve this the supply chain challenges need to be addressed and supply chain opportunities realised.



## 2.1 Establish Strategic Partnering Advisory Board :

An engaged group convened and facilitated by Jisc could be a great vehicle for kick-starting a new conversation by articulating 'customer demand' in a clear way which the supply chain could respond to.

- We recommend that Jisc establishes a Strategic Partnering Advisory Board to focus on the collaboration activities required to ensure that all content items are delivered with metadata suitable for a 21<sup>st</sup> century library.
- An important part of its terms of reference would be leading engagement and advocacy with the supply chain.
- There is likely to be the need for sub-groups to deal with some of the specific projects and engagement programmes.
- There is a danger that a library only group, just reiterates the problem, but there is huge potential to convene a different type of group with senior librarians, librarians involved in the governance the regional higher education purchasing consortia, primary publishers and others in the supply chain.



## 2.2 Engagement of metadata practitioners in UK academic libraries

The players in the supply chain need to be provided with a clear and consistent message about what is required.

- Librarians agree that current poor metadata is a significant issue for efficiency and resource discovery.
- The players in the supply chain need to be provided with a clear and consistent message about what is required.
- While individual libraries may always have some specific requirements agreeing a shared minimum cataloguing standard or standards in the UK will be essential.
- The National Acquisitions Group, CILIP, and the NBK Community Data Group have already done some good work in this area, and may provide a starting point for a sub-group to engage in the development of this standard, which would inform the tender and licensing requirements for procurements from primary, secondary publishers and other vendors.



## 2.3 Engagement with the players in the metadata supply chain (1)

Primary publishers: all book metadata begins life here, and primary publishers are essential stakeholders.

The largest publishers create very rich metadata records in XML for their books and build these up throughout the book's lifecycle starting from project inception and continuing right through to post-publication.

- **Library advisory groups:** Jisc has a high profile and is well positioned to engage at a high level with these publishers. This engagement could be done in several ways. Many of them have library advisory groups, and we are sure that an offer of involvement from members of Jisc's NBK team would be welcome. We would suggest an approach to Springer in the first instance.
- **Round table:** Inviting metadata experts from the primary publishers to a round table event may engage primary publishers in the issues and kick starting a dialogue about how Jisc can collaborate them. Organising this event (or events) in conjunction with professional bodies such as STM and the Book Industry Communication (BIC) might be a good way to begin.



## 2.3 Engagement with the players in the metadata supply chain (2)

Primary publishers: all book metadata begins life here, and primary publishers are essential stakeholders.

The largest publishers create very rich metadata records in XML for their books and build these up throughout the book's lifecycle starting from project inception and continuing right through to post-publication.

- **Round table:** Inviting metadata experts from the primary publishers to a round table event may engage primary publishers in the issues and kick start a dialogue about how Jisc can collaborate them. Organising this event (or events) in conjunction with professional bodies such as ALPSP, CrossRef, and Book Industry Communication (BIC) might be a good way to begin in the organisation of such round tables.



## 2.3 Engagement with the players in the metadata supply chain and pilots (3)

If there is appetite from primary publishers, representatives could be invited to join the Strategic Partnering Advisory Board, or a sub-group with primary publishers and libraries could be set up.

- **Pilots:** Many publishers also generate/procure and deliver MARC records direct to library customers, or to intermediaries in the library retail supply chain. Engaging them to make all the academic book metadata they produce openly available throughout the supply chain with a CC0 licence could be a huge win as it should drive discoverability, access, and sales across myriad platforms.
- Pilot projects to test the concept and build engagement could be highly influential.
- We would suggest approaching Sage, as a first step to explore a project along these lines, perhaps focusing on all or part of the SAGE Knowledge eBook and Reference is a social sciences digital library of more than 5,000 titles, which is available via Jisc Collections. Springer Nature might make another ideal partner. Although not a primary publisher, NBK may want to work with Knowledge Unlatched to ingest the metadata it already makes available under a CC0 licence.





## 2.3 Engagement with the players in the metadata supply chain and pilots (4)

The National Acquisitions Group has done some helpful work in unpacking the specific problems libraries have with MARC records procured from secondary publishers but the proposed solution (i.e. ‘tell’ aggregators what they need to do) is perhaps over-simplistic.

- **Jisc Collections and regional procurement consortia:** strong relations exist between Jisc Collections and primary publishers and between the regional procurement consortia such as SUPC and secondary publishers. These relationships can be brought into play to influence the book bibliographic metadata supply chain, particularly if good feedback loops can be created between the bibliographic data community and these organisations.
- **Secondary publishers (i.e. aggregators):** a really key message in this part of the supply chain is to please add to metadata where needed, but not to remove any metadata provided by primary publishers, for example on accessibility features or the OA status of books.



## 2.3 Engagement with the players in the metadata supply chain and pilots (5)

Great cataloguers, influential with many stakeholders, and technologically savvy the British Library could be a particularly powerful partner to the Jisc in terms of influencing the supply chain, and Jisc could be particularly powerful partner to the British Library in terms of understanding the academic library community.

- **Metadata providers:** organisations such as **BDS**, **the British Library**, and **OCLC** could also be approached to provide open metadata adapted to UK academic library requirements. Here the key challenge will be to pay the ‘first copy’ costs. For each metadata record if there will only ever be one customer and the record will then be openly available for all, then the full economic cost of creating and managing that metadata records must be paid. This needs to be sustainable both for the metadata provider and also for the customer.



## 2.3 Engagement with the players in the metadata supply chain and pilots (6)

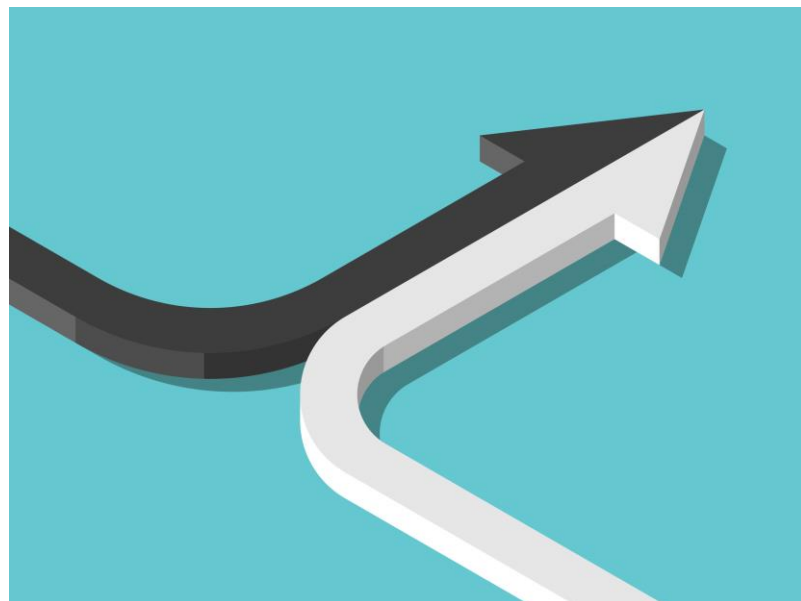
Jisc's role in providing national infrastructure in this space firmly qualifies it for the expert group.

- **Standards bodies:** Bodies such as **BIC/CrossRef/DOAB already** engage with and influence both primary publishers (i.e. educational and trade for BIC, academic for CrossRef, and fully open for DOAB) and other links in the UK academic library supply chain. They are key allies on this journey. It is essential that Jisc gets involved with these group. Jisc could apply to sit on the BIC Operational Board, for example. A priority of this group is to connect experts and stakeholders.



### 3. Consider other strategic partnerships

- NBK may want to consider if strategic partnerships may be advantageous.
- BDS and OCLC may be potential strategic partners in enriching NBK.
- They might, at a reasonable cost, be excited by partnering with Jisc in making long lasting and sustainable change.



## Conclusion

...what starts life as a robust metadata record can be stripped of intellectual input, so that it become a rather useless nub. At the same time libraries are investing millions of pounds in providing the books that are essential for research and teaching.

- We are not recommending quick fixes, they do not really exist, and would merely be sticking plaster.
- Rather we are recommending leadership for long-term change, which will ensure metadata suitable for 21<sup>st</sup> century academic libraries.
- Jisc is well placed to deliver the leadership needed
- This will require outward looking active engagement with the players in the supply chain, ideally with governance from a Strategic Partnering Advisory Board, and supported with some groundbreaking pilots.



# Streamlining the Metadata Marketplace

Next Steps

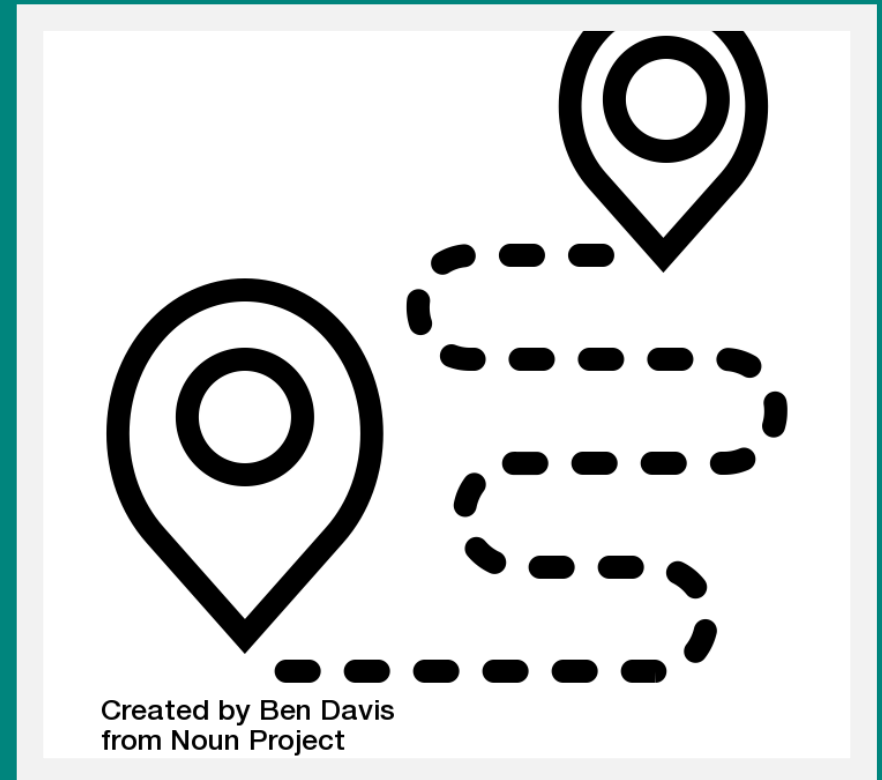
# Reflections from Jisc

## Our response to the report

- Confirmation of the complexities of the metadata marketplace, with valuable insights into the driving and restraining forces for change.
- Engagement with a wide range of stakeholders throughout the supply chain, highlights the need for more dialogue between key players.
- Financial modelling, while imperfect, provides useful indicative figures which contextualise issues.
- Captures an appetite for change amongst some (but not all) stakeholders.
- Jisc cannot solve the complexities alone, but can offer leadership and a vision for change.
- Recommendations suggest practical steps forward, which we will develop into a roadmap for the next five years.

# Creating a Roadmap

- Disseminate the report recommendations
- Validate the findings with stakeholders
- Consult on plans for action
- Develop a programme of activity

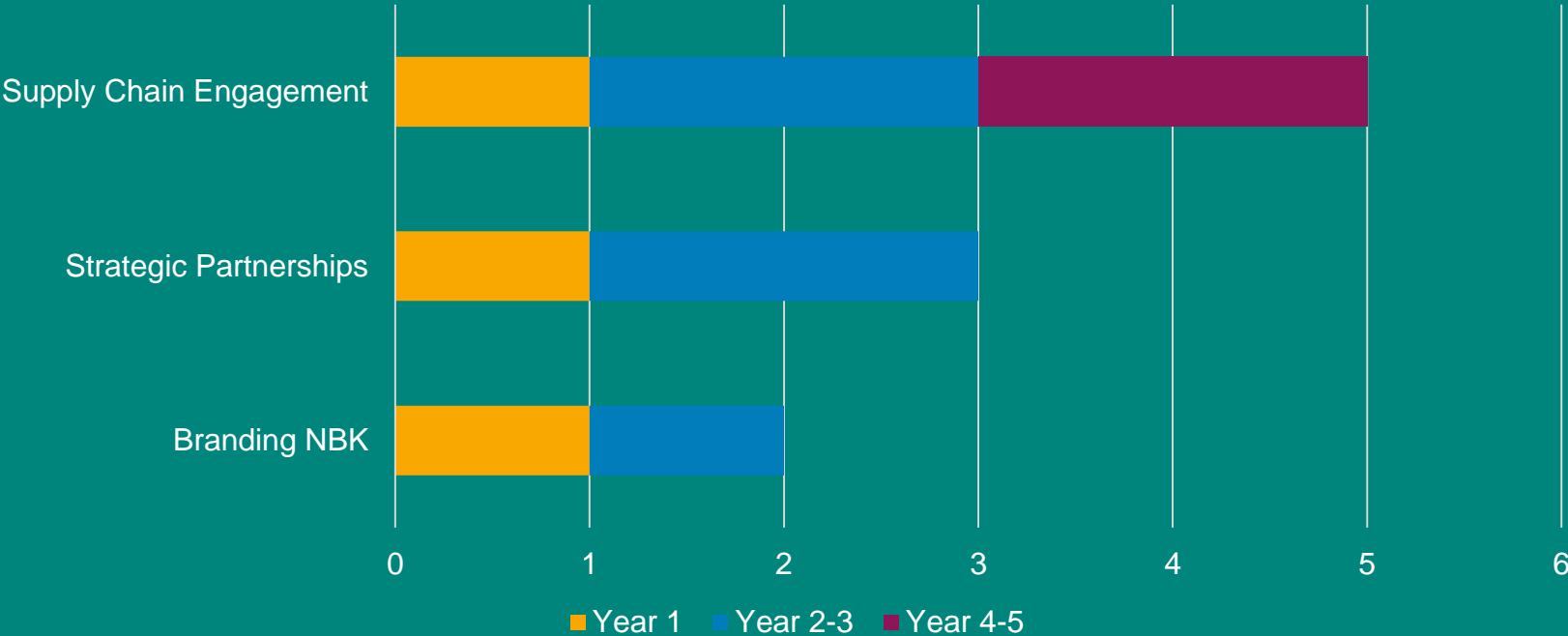




# A Phased Approach

Aiming to deliver change over time

Action Plan



# Key activities: Branding and Partnerships




Branding the NBK: position the NBK as a vital piece of *National Infrastructure* and assert Jisc's expertise, developing the concept of thought leadership


**Throughout the metadata supply chain  
Internationally**





Consider strategic partnerships with suppliers to enrich the NBK with openly licensed metadata


# 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