THE BRITISH LIBRARY

Foundations for the Future

The British Library's Collection Metadata Strategy

2019-2023

Summary

"Our vision is that by 2023 the Library's collection metadata assets will be unified on a single, sustainable, standards-based infrastructure offering improved options for access, collaboration and open reuse"

Collection metadata is an umbrella term for the structured information required to efficiently manage, access and preserve the collection. It is a key organisational asset, representing centuries of resource investment. Efficient exploitation and stewardship of collection metadata underpins delivery of The British Library's core purposes, strategies and portfolios.

"Foundations for the Future" is The British Library's second Collection Metadata Strategy and covers the period 2019-2023. Its predecessor, *"[Unlocking the Value](#)"*, described the foundational principles, priorities and governance required for centralised management of metadata assets. However, the ability to exploit these assets to deliver service transformation has often been restricted by complex legacy systems infrastructure and related technical standards.

The British Library's Technology Transformation and Library Systems Transformation programmes are addressing the challenge of outdated infrastructure and coincide with the period covered by "*Foundations for the Future*". Collectively, these initiatives combined with the Collection Metadata Strategy offer an opportunity to finally integrate and present the Library's metadata assets in a way unachievable since its foundation.

Strategic activities will be guided by the following priorities:

- **Enhance the Library's ability to exploit its collection metadata assets in order to deliver its strategic priorities, portfolios and programmes**

- **Ensure that the value of the Library's Collection Metadata investment is maintained through effective stewardship**

- **Enable open access to collection metadata in order to improve resource discovery and promote wider community reuse**

"*Foundations for the Future*" records the measures used to manage implementation of the strategic priorities together with high-level objectives. While the rapidly evolving operational environment makes it difficult to predict all aspects of implementation with precision, high-level deliverables are noted in the *Collection Metadata Strategy Roadmap 2019-23*. Annual implementation plans will also be created and coordinated with current Library priorities.

Purpose

"*Foundations for the Future*" is the Library's second Collection Metadata Strategy and covers the period 2019-2023. Its predecessor, "*Unlocking the Value*", was the first document to describe the foundational principles, priorities and governance needed to manage the Library's collection metadata assets in support of wider strategic objectives. Like its predecessor, the new strategy will be supplemented by annual implementation plans covering the actions necessary to address legacy issues, support current business needs and achieve future goals. A strategic roadmap summarising activities planned for 2019-23 has also been produced to accompany this document.

Context

The Library's 'Collection metadata' has evolved from a basic inventory listing to encompass all structured information required to efficiently manage, access and preserve the collection. It is therefore a key organisational asset, representing centuries of resource investment. Delivery of the Library's core purposes, strategies and portfolios is underpinned by the efficient exploitation and stewardship of collection metadata. However, despite its strategic importance, collection metadata's potential to support service transformation has been restricted by over-complex systems infrastructure and competing technical standards.

The first collection metadata strategy was primarily concerned with establishing centralised management and support of metadata to spread best practice and develop new capabilities. The new strategy will build on the achievements of the first and seek to implement a unified systems and standards infrastructure for collection metadata. This unification process will inevitably require some compromises arising from short and longer-term goals. However, only through convergence on a standard, unified approach to collection metadata management will it be possible to support the Library's strategic ambitions in an affordable and sustainable manner.

Challenges

The period covered will see the Library facing several key metadata management challenges.

Unifying Collection Metadata Infrastructure

Collection metadata assets are often the result of decades of uncoordinated management and therefore often inconsistent and confusing to users and staff. The lack of centralised control of metadata policy and best practice was addressed by the previous strategy. However, the Library's ability to unify its metadata has remained constrained by legacy infrastructure and standards. This situation has led to over-complex services incapable of supporting: comprehensive inventory control, coherent resource discovery, efficient digital preservation or innovative collaborations.

Systems- The Library Systems Transformation (LST) and Technology Transformation (TT) programmes are addressing the challenge of outdated infrastructure and coincide with the new Strategy period. These programmes, coupled with activities outlined in "*Foundations for the Future*", offer an opportunity to finally integrate and present collection metadata in a manner unachievable since the Library's foundation.

Standards- Standards underpin the creation, management and dissemination of collection metadata and are increasingly essential as dependence on automated workflows for efficiencies grows. A proliferation of standards increases operational costs and complexity while making it hard to offer a consistent user experience. Procurement of a new generation of systems offers an opportunity to a more integrated approach based around a uniform, system-neutral, metadata model for all collections. This model would also address important new properties of digital acquisitions including: increased granularity, complex relationships between local and remote resources, persistent identification, dynamic rights management and digital preservation.

Access-The LST and TT Programmes also offer an opportunity to unify options for sharing and access to metadata and collection content via the pooling of previously fragmented resources. Creation of a single central repository holding all collection metadata combined with a unified metadata model would allow the comprehensive presentation of the Library's collection as a cohesive whole. Such a repository could also offer more efficient and sophisticated access options for wider communities via common APIs to collaborative platforms, shared national infrastructure and public sector open data initiatives.

Maintaining and Developing the Value of Collection Metadata Assets

Making Older Metadata Suitable for Current and Future Needs - The bulk of the Library's metadata is the legacy of centuries of effort by multiple institutions in compliance with differing cataloguing rules. While some enhancement and harmonisation has been feasible, we will seek new and innovative

ways of migrating and enhancing legacy metadata to support current and future business needs.

Improving Visibility of Hidden Collection Resources - If collection metadata is unavailable to users, the content described is effectively invisible and its value and utility greatly diminished. Categories include:

- Printed collection catalogues available only in reading rooms
- Local collection databases inaccessible via our resource discovery infrastructure
- Metadata encoded in proprietary formats or subject to licensing restrictions for reuse

A programme of metadata creation and migration activities is required to improve the visibility of any 'hidden' content and support its use for digitisation, resource discovery and community reuse.

Delivering Efficient and Sustainable Collection Metadata Processes

Increasing Volume and Complexity -Traditional methods of metadata generation, management and dissemination are not scalable or appropriate to an era of rapid digital change, rising audience expectations and diminishing resources. The volume and complexity of content requiring metadata is rising as new technologies enable bulk digitisation and publishers offer entire backlists via multiple licensing channels. However, the extension of automated metadata processing with manual exception handling offers the ability to maintain quality while dealing with increased volume.

New Options for Metadata Creation and Enhancement -The changing operational environment offers opportunities for capability improvement and service enhancement. New open web resources, automated generation of metadata from full text content and crowdsourcing initiatives offer many interesting new possibilities for exploration. We must ensure staff have opportunities to acquire the skills required to exploit such possibilities in order to develop innovative solutions to increasingly complex and time-critical challenges.

Collection Metadata for Preservation, Rights and Management Information

Preservation Metadata – The long-term future of the rapidly evolving digital collection will be underpinned by the use of accurate and timely preservation metadata. Many content formats are currently unstable as publishers experiment with new

technologies and delivery models. Only by accurately and consistently recording key content properties to recognised metadata standards will it be possible to preserve and present digital content to future generations of Library users.

Rights Management – In order to manage increasingly diverse and dynamic options for hybrid print/digital and local/remote access to collection content, we must support an equally sophisticated rights metadata infrastructure. The Library's digitisation projects also require complex rights management together with reliable metadata to enable efficient selection of collection items.

Collection Management Information –Suitably designed and maintained Collection Metadata offers further benefits in supporting collection management and audit queries. Metadata enhancements required for an improved user experience can also offer rich management information and collection comparison options when coupled with appropriate systems.

Spreading and Maintaining Best Practice

Communication – The success of the Strategy will rely on adoption and maintenance of best practice by Library staff. To encourage this process, case studies will be developed and shared showing how agreed standards, policies and best practice help deliver service and workflow efficiencies. Support documentation also will be available to staff via the Library Wiki.

Staff Development – In order to support continuous improvement of metadata assets, it will be essential to enable staff to undertake skills development and participate in knowledge sharing with internal and external colleagues. We must identify development needs and offer both formal training and informal opportunities for professional development.

Strategic Fit

The British Library Act, *(1972, Ch. 54)* records the organisation's corporate responsibility to be "*a national centre* for *reference, study and bibliographical and other information services.*" This core requirement resulted in the Library offering a range of metadata services from its foundation. In addition to recording UK publishing output via the *British National Bibliography*, these services also served to increase international visibility of the collection. The Library's ability to comply with key Government initiatives relating to public bodies (*e.g. Open Public Data, Freedom of Information (FOI) & Public Sector*

Information (PSI)) is also underpinned by its metadata assets.

Effective management of collection metadata supports the Library's mission and vision statements, delivery of key strategies and service portfolios, notably Legal Deposit, Everything Available and Digital Preservation. Delivery of the Library's Living Knowledge Vision also depends upon the possession of appropriate collection metadata, specifically:

Custodianship – Curation and preservation of the national collection requires well-maintained collection metadata for secure inventory control and access

Research – Research is supported and stimulated via the use of collection metadata for resource discovery, document delivery and data mining.

Business – Innovation can be stimulated via the release of open metadata for commercial exploitation in line with Government public sector guidance

Culture – Engagement with the public via exhibitions and creative communities via release of digital content is supported through metadata use in identifying suitable resources and creating related web pages.

Learning – Schools and other academic institutions are helped via our release of free metadata to support efficiencies in the creation of their library catalogues

International – We work with partners around the world to exchange metadata, advance international standards and host agencies for standard identifiers.

Strategic improvement programmes for Save Our Sounds, British Library on Demand and Boston Spa Renewed together with the Everything Available Programme objective of single search navigation are predicated on the availability and capability of suitable collection metadata.

Our Vision

Our vision is that by 2023 the Library's collection metadata assets will be unified on a single, sustainable, standards-based infrastructure offering improved options for access, collaboration and open reuse.

Strategic Principles

The following core principles support the vision and will guide implementation.

- The Library must maintain a management strategy for collection metadata, as it does for other key assets.

- The level of detail and availability of collection metadata should be determined by operational service and product requirements *(e.g. Universal Viewer),* collection security, preservation and audit.

- All collection metadata must be hosted on a recognised Strategic Metadata Repository.

- All collection metadata should be created and managed in compliance with recognised international standards and legal obligations.

- The Library must be responsive to the expectations of its users and stakeholders in the management and exploitation of collection metadata.

Strategic Priorities - 2018-2023

The Strategy will focus on the opportunity offered by replacement of ageing systems to build a fully unified collection metadata infrastructure. Specific activities will be determined by the following priorities.

Enhance the Library's ability to exploit its collection metadata assets in order to deliver its strategic priorities, portfolios and programmes by:

- **Managing migration to a new unified infrastructure for collection metadata management** *to reduce operational costs and complexity*

- **Developing and maintaining innovative collection metadata creation and enhancement processes** *to support efficient exploitation of third party data sources e.g. automated entity extraction from full text, crowdsourcing and bulk record enhancement.*

- **Designing and supporting accurate and flexible rights management metadata for both content and descriptions** *to enable the Library to take advantage of new licensing options and increase reuse of the collections*

- **Providing colleagues with self-service tools, expert assistance and best practice guidance** *to improve consistency and efficiency in the*

exploitation of collection metadata and integrate with established training initiatives

- **Ensuring requirements for new management information tools utilising collection metadata are included in LST procurement specifications** *to improve accurate assessment of collection strengths and weaknesses.*

Ensure that the value of the Library's Collection Metadata investment is maintained through effective stewardship by:

- **Collaborating with international, cross sectoral partners** *to ensure appropriate collection metadata standards are developed and implemented to support operational efficiency and preserve long-term value, e.g. creating opportunities to consume high-quality metadata from external sources*

- **Ensuring collection metadata accurately and persistently incorporates all necessary information** *to deliver the Library's portfolios and wider strategic objectives*

- **Developing processes to prioritise collection metadata enhancements** *that support delivery of improved services to internal and external users*

- **Managing long-term risk to the integrity of collection metadata** *against short-term efficiencies*

- **Undertaking collection metadata maintenance, enhancement and QA processes** *to support increased operational efficiencies and smooth transition to new systems*

- **Ensuring licensing of third party collection metadata is negotiated with a long-term view** *i.e. accurately documented and securely stored*

Enable open access to collection metadata in order to improve resource discovery and promote wider community reuse by:

- **Ensuring suitable collection metadata is available** *to enable and support new shared service initiatives, open research and commercial collaborations*

- **Increasing opportunities for interaction with our collection metadata** *to show the richness of the collection and its relevance to all*

- **Ensuring our collection metadata is accurately represented and accessible** *via all relevant national or global discovery channels and platforms*

- **Exploring opportunities for community engagement in metadata enhancement** *e.g. via development and promotion of crowdsourcing challenges*

- **Taking steps to expose any collection metadata assets currently unavailable** *for resource discovery, partnership or collection management*

- **Ensuring comprehensive, accurate and timely rights** metadata **is available** *to support the sharing or purchase of any digital content.*

- **Making sure our collection metadata is openly available** *in full compliance with current public sector best practice and related legislation.*

Delivering the Priorities

Governance & Support

The governance structure responsible for ensuring that an effective Collection Metadata strategy is delivered comprises three core groups:

- **The Collection Metadata Strategy Group (CMSG)** *– a strategic coordination group chaired by the Head of Collection Management.*

- **The Collection Metadata Advisory Group (CMAG)** *- an expert group with representation from collection areas across the Library responsible for advising the CMSG on technical issues and chaired by the Collection Metadata Standards Manager (Collection Metadata)*

- **The Collection Metadata Processing Group (CMPG)** *- a group responsible for resolving operational metadata issues in line with the Metadata Strategy and chaired by the Head of Content and Metadata Processing (Collection Management)*

The Collection Metadata Team *(in Collection Management) is* responsible for supporting activities required to deliver of the Strategy.

Implementation

Annual strategy implementation plans will be created covering the following areas:

- Asset Management
- Process Improvement
- Policy & Standards
- Communications
- Systems Infrastructure
- Enhancement & Preservation
- Licensing & Rights
- Open Metadata & Reuse
- Discovery & Access
- Staff Development

While the changing operational environment makes it problematic to predict all aspects of implementation, high-level delivery objectives are documented in the *Collection Metadata Strategy Roadmap 2019-23* and include:

Throughout

- **Collection Metadata systems, standards and business process harmonization and convergence** – *will be supported to enable delivery of the Library Systems Transformation Programme and Everything Available Discovery Services Roadmap*
- **Specific metadata development requirements for projects, portfolios and operational efficiencies** *will be delivered (e.g. Save our Sounds, Heritage Made Digital, non-print legal deposit, etc.)*
- **Government data initiatives will be implemented** *in line with public sector policies and best practice (e.g. open public data, data privacy, etc.)*
- **Key strategic metadata collaborations will be supported** *(e.g. JISC National Bibliographic Knowledgebase, UK Research Reserve, Single Digital Presence, etc.)*
- **Staff awareness of best practice and efficiency for the creation and exploitation of metadata** *will be improved*
- **Inadequate collection metadata will be made fit for purpose** *and migrated to new systems in line with corporate priorities via targeted automated or manual interventions*

- **The Library's open metadata service portfolio** *will be further enhanced and extended*

2019/20

- **Metadata assessment, migration and configuration activities for the Digital Asset Management and Preservation System (DAMPS) and other LST Project strands** *will be supported in line with agreed Project deadlines*
- **Activities where corporate policy and best practice for collection metadata are required** *will be identified for further work (e.g. converged standards, crowdsourcing metadata projects, entity extraction from full text, etc.)*
- **A single target metadata model for use on next generation systems** *will be developed in consultation with stakeholders.*
- **New options for metadata skills sharing by Library staff** *will be implemented (e.g. Knowledge Exchange Forum, mentoring, etc.)*
- **Library metadata systems usage** *will be benchmarked against that of peer organisations and the results documented to inform new systems requirements.*
- **Activities supporting publisher circulation of good quality metadata** *via trade supply chains will be investigated and piloted (e.g. use of the ISNI)*
- **Collection Metadata for Library holdings** *will be available via the new JISC National Bibliographic Knowledgebase*
- **An options appraisal for migration of the Sound & Moving Images (SAMI) dataset** *will be completed and recommendations documented*

2020/21

- **Technical options for the persistent identification, hosting and management of collection metadata assets and related services** *will be identified and prioritised*
- **New metadata tools and templates for data capture and basic record creation** *will be developed for digitisation and crowdsourcing projects*
- **Implementation of a replacement for the current Integrated Archives & Manuscripts System (IAMS)** *will be completed and all metadata migrated*
- **A programme of work to digitise and migrate the last printed catalogues** *will be undertaken*

2021/22

- **All remaining catalogue metadata outside of core repositories** will be migrated
- **New metadata transformation options implemented**

2022/23

- **Implementation of a replacement for the current Aleph system** will be completed and all metadata migrated
- **Central metadata repository implemented**
- **Convergence process completed for systems and standards**
- **Development of new service options in consultation with stakeholders**

Metrics

Suitable metrics will enable the effectiveness of the strategy to be measured will be supplemented by an annual audit reporting the current state of the Library's metadata assets.

Communications

For transformation to succeed, cultural and technical and business change will be required. This change will require a comprehensive range of communications to support standards convergence and adoption of best practice. Priority will be given to clear 'outcome-based' communications avoiding complex jargon. External stakeholders also require communication of the Library's collection metadata activities, standards and systems initiatives. Both sets of requirements will be addressed by the following measures.

- Continuing development of Collection Metadata Wiki as a central staff resource for dissemination of best practice
- Maintenance of a set of public Collection Metadata web pages
- A centralised British Library email support function for collection metadata queries.
- Use of social media to provide updates on progress
- Talks at external events to disseminate information about the Strategy

Expected Outcomes

By 2023:

- **The complexity of the Library's collection metadata infrastructure** *will be reduced by convergence on an agreed set of supported standards and systems*
- **The unified collection metadata infrastructure** *will offer new access and processing options enabling a greatly improved user experience of Library services*
- **Efficient, sustainable collection metadata workflows** *will match the increasing scale and complexity of collection content via implementation of new techniques for record creation and exploitation of external data source*

Glossary

Term	Definition and Scope
Collection Metadata	'Collection metadata' is an umbrella term for the metadata used to: manage, access, preserve and describe the collection
Hidden Metadata	Collection metadata that is unavailable to someone to whom it would be of value. Hidden metadata includes: printed catalogues; local databases that cannot be accessed through the discovery layer; metadata encoded in a specialised schema or format; metadata that is subject to restrictions on reuse or distribution.
Legacy Metadata	Inherited metadata that is not capable of satisfying use cases. Legacy metadata includes: metadata that is being used for purposes for which it was never intended; metadata that has been converted from a printed format to a machine readable format; metadata that is not compatible with current standards. These categories are not mutually exclusive.
Metadata	Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource
Open Data	Data that is made available by organisations, businesses and individuals for anyone to access, use and share
System-neutral metadata	Metadata that is independent of proprietary or system specific constraints that restrict opportunities for development or data migration.