Enrichments Policy

for the common European data space for cultural heritage

Revision	Version 1
Last updated	August 2023
Author(s)	Hugo Manguinhas, Ariadna Matas, Maša Škrinjar
Contributor(s)	Europeana Foundation staff Europeana Data governance stakeholder working group
Dissemination Level	Public

Table of contents

Introduction and definitions

The Europeana Initiative has been conducting activities around enrichment for several years, often developed in partnership with other organisations through past projects, and will continue to do so in projects and activities supporting the common European data space for cultural heritage (hereinafter the 'data space').

Enrichments are data about a cultural heritage object (i.e. metadata) or of the object itself (i.e. its content) that augment, contextualise or rectify the authoritative data made available by cultural heritage institutions. Enrichments of metadata aim to improve it by adding new, or refining or rectifying information about the object, while content enrichments aim to produce alternative representations in various media types of the object.

Examples of common types of enrichments in the data space are: connecting subject terms in the metadata of the object with vocabularies; suggesting corrections to the name of an author in the object metadata; creating an audio description for an object; annotating a manuscript with linked data resources about historical names of persons or events in history; annotating an artwork with scientific literature where it is analysed; applying Object Character Recognition (OCR) to a digitised book or Handwritten Text Recognition (HTR) to a manuscript; generating audio to text transcription of a recorded speech or generating a subtitle for a video; etc.

In this document, **enrichment effort** is used to refer to an activity or workflow run by an organisation and/or project that either leads to new enrichments or the development of tools for producing or reviewing enrichments. It may reflect manual, crowdsourced, user-assisted/semi-automated or fully automated processes that contribute to the production and/or review of enrichments.

Enrichments contribute to improving data quality and can help make the cultural heritage object more meaningful, substantial, or rewarding to a wider audience. They improve discoverability and access to the object, contextualise it or make it easier to understand for an individual, define relations with other objects and/or resources, enable multilingual access, etc.

The purpose of this policy is to bring direction and consistency across enrichment efforts in the data space. It does so by providing a clear vision and a number of principles. As enrichments become more established and relevant in cultural heritage, this helps avoid inefficiencies and supports the positive contribution that enrichments can make.

The terminology used in this document is defined in the [Europeana Glossary](.).

Scope

This policy supports the conception and development of purposeful enrichment efforts that contribute to the objectives of the data space. To achieve this, it establishes a vision and a set of principles that guide enrichment efforts in a consistent way across projects, organisations and activities in the ecosystem in and around the data space.

The principles apply to:

- End-to-end enrichment efforts in the context of the data space, including those taking place in cultural heritage institutions or aggregators with the intention of publishing data in the data space.
- Everyone contributing to and participating in the creation of enrichments in the context of the data space. This includes individuals and organisations acting as data partners, technical partners, users, etc. up to the operator of the data space.

This document does not include considerations on specific types of enrichment, and instead provides overarching principles to which each type of enrichment must adhere. It does not include implementation measures either. These will follow the adoption of the policy. This policy is a starting point for updates to the Europeana Licensing Framework (ELF), Europeana Publishing Framework (EPF), Europeana Data Model (EDM), as well related processes, standards and future policy tools that the data space might rely on, and will ensure that this is done consistently.

Process

This policy has been developed collaboratively by consulting and bringing in the views, priorities and concerns of stakeholders, namely:

- **Data partners**, which include two different groups: cultural heritage institutions and aggregators. They provide the data that is enriched and may provide enrichments. They may include data partners in Generic Services and data space projects and/or data partners from across the data space.
- **Technical partners**, which include research institutes & infrastructures, universities, tool developers/vendors, including aggregation infrastructures not designed by aggregators themselves, Generic Services and data space projects and/or technical partners from across the data space. They are not data providers, but provide tools and/or maintain infrastructures for producing enrichments.
- **The Europeana Foundation**, when acting as responsible for updating frameworks in relation to this policy, as facilitator, mediator or supporter of enrichment initiatives, and as a producer/provider of enrichments as part of its aggregation activities.
- **Users of the data space**, including users of the Europeana website and APIs, and users of tools that produce enrichments.

- **The European Commission**, in particular DG Connect as the department responsible for the data space.
- **Member States**, in particular the Ministries and cultural heritage institutions represented in the Expert Group on a common European Data Space for Cultural Heritage (CEDCHE) and its relevant sub-groups.

These stakeholders have participated in the definition of this policy through a [general stakeholder consultation](#), followed by targeted consultation through the [Data Governance Stakeholder Working Group](#).

Vision and principles for enrichment efforts

Enrichments efforts are conducted with the aim of increasing the overall quality of data and maximising its value for further use, whilst engaging and harnessing the knowledge of experts and different user communities. Experiences and skills are shared within and outside the data space, leading to new partnerships and collaborations. Reuse of enrichments by individuals and organisations is supported through removing legal barriers, tracking the provenance of enrichment efforts, and the application of interoperable technologies and standards. While planning and executing enrichment efforts, diverse perspectives and environmental concerns are taken into account.

Each section below extends this overarching vision with respect to a specific topic and identifies the main principles to be respected.

Quality

Enrichments should enhance the quality and value of the data by, for example, adding more context, creating connections with other resources and making data more useful to various groups of users with different abilities and preferences.

Principles

1. The outcomes of enrichment efforts lead to improved quality and increase the value of data for users.

2. The quality improvements resulting from enrichment efforts are measured on the basis of transparent and standardised criteria established via relevant frameworks (such as the Europeana Publishing Framework's tier system).

3. A process for validating enrichments is designed prior to the execution of enrichment efforts as an instance of a validation methodology developed by the Europeana Initiative and is based on (pre)agreed evaluation criteria and quality targets for the respective enrichments.

Transparency of provenance

We aim to be transparent about the provenance of enrichments, including the approaches behind them, so that users can make informed decisions about if and for which purposes they can use them.

Principles

4. Authoritative data provided by cultural heritage institutions is differentiated from data added through enrichment efforts made by actors other than the holding institution.

5. Detailed information about the purpose, actors, processes and workflows for producing and validating the quality of enrichments is made available to users.

6. The efforts made by the producer(s) of enrichments are acknowledged where possible and/or relevant to the audience.

Interoperability, reusability and copyright

We strive to facilitate the reuse of enrichments and to rely on open standards that facilitate interoperability. Barriers to (re)using data and software resulting from enrichment efforts are removed to the extent possible, balancing legal and ethical considerations.

Principles

7. Enrichments are made interoperable and reusable through the use of openly accessible, usable and machine-readable standards (such as EDM and associated extensions and profiles). The same applies to the contextual data (such as vocabularies) that is part of enrichments or is used in enrichment efforts, and to the data about the provenance of enrichments.

8. Reuse conditions are clearly communicated using interoperable standards in a way that is accurate and clear (in line with the Europeana Licensing Framework).

9. No new or additional copyright or related rights exist in most enrichments, regardless of whether they are machine or human-generated. If they do exist, they are waived for the purpose of facilitating the functioning and reuse of data in the data space. This does not affect pre-existing rights in data being used (e.g. a poem that is transcribed).

Collaboration and capacity building

Through enrichment efforts, we strive to foster new collaborations within and between sectors and disciplines. We also aim to build capacity among all stakeholders from the same or different domains to contribute, develop and share their knowledge and experience.

Principles

10. Enrichment efforts foster partnerships based on both technical and non-technical expertise.

11. Enrichment efforts enable better integration and connections with other data spaces and sectors by enhancing the data in a way that is relevant for and can be used by other data spaces and their focus industries.

12. Sharing experiences and findings obtained from enrichment efforts builds new or extends existing digital capacity for cultural heritage professionals by bringing new understanding and skills around (new) technologies and standards, and demonstrating their application to address existing challenges through the creation of business cases, workflows, methodologies, policies, case studies, good practice, and sharing outputs of enrichment activities.

Participation, diversity and inclusion

We strive to encourage diverse user participation and community engagement in enrichment efforts as well as in the data available in the data space. Enrichment efforts should contribute to diverse and inclusive data, create opportunities for under-represented voices to shape the narrative about cultural heritage objects, and care for the needs of different users.

Principles

13. Organisers of enrichment efforts define their scope, conditions and guidelines for participation, and desired outcomes. They are mindful of users' accessibility needs, skills and knowledge and proactively involve relevant users in a way that ensures diverse participation. They implement mechanisms of control and moderation for outputted information.

14. Alongside organisers' efforts outlined in the previous principle, producers of enrichments also take responsibility for their own actions, making sure enrichments are not inappropriate, offensive and harmful, and that they comply with pre-defined criteria and/or guidelines. They make an effort to think critically about their contributions and overcome their own cultural biases.

15. Where possible, underrepresented groups, minorities and minoritised communities are encouraged to get involved in enrichment efforts, especially when the data that is being enriched relates to, was created by, or is about these communities, by reducing the barriers to their participation and by approaching them in a mindful and targeted way.

Environmental sustainability

While enrichment efforts are meant to have a positive impact, the production of new data, its storage and reuse often have a negative impact on the environment. We believe that all data management activities, including enrichment efforts, should be designed and implemented with awareness of their environmental cost, and by considering whether their value is worth the negative impact.

Principles

16. All resources used are identified and optimised when selecting, developing and executing the methods used in enrichment efforts. This includes not only human and software/hardware resources but also other resources, such as energy consumption and its corresponding digital carbon footprint.

17. Any enrichment effort includes the consideration of whether the quality and value brought by enrichment efforts sufficiently justify the investment and environmental cost.

18. Replication of enrichments or enrichment efforts is avoided unless they result in a foreseeable increase in quality and value.