Methodology for validating enrichments

Revision	Version 1.1
Last updated	29/2/2024
Editor(s)	Antoine Isaac, Europeana Foundation
Contributor(s)	Monica Marrero, Nuno Freire, José Eduardo Grano de Oro, Eleftheria Tsoupra, Hugo Manguinhas, Europeana Foundation
Dissemination Level	Public

Version history

Revision no.	Date	Author	Organisation	Description
0.1	10/2022	Antoine Isaac	Europeana Foundation	Initial draft
0.2	11/2022	Antoine Isaac, Mónica Marrero, Nuno Freire, José Eduardo Grano de Oro, Paolo Scalia, Eleftheria Tsoupra	Europeana Foundation	Review and update
0.3	02/2023	Hugo Manguinhas, R&D team members above	Europeana Foundation	Review and update
1.0	08/2023	Antoine Isaac, Nuno Freire, Mónica Marrero	Europeana Foundation	More details on sample datasets, cleaning up appendices
1.1	02/2024	Antoine Isaac	Europeana Foundation	Formalisation of workflow, update past & future work, enhanced presentation.

Table of contents

Document status

The present version of this document is largely based on experience gathered with data space partners and published in earlier project deliverables, especially the Europeana DSI Generic Services and Data space - supporting projects. However, it contains material that has not been yet reviewed by all relevant stakeholders. Do not hesitate to contact us if you have any feedback!

1. Introduction

Many partners in the common European data space for cultural heritage (hereafter "data space"), including projects, aim to produce *enrichments*[1] of metadata and/or content for the cultural heritage objects they contribute. This includes (semantic) tagging using the Linked Open Data vocabularies supported by Europeana[2] (or not), translation, transcriptions, etc. These *enrichment efforts* use various processes to produce these enrichments, from fully manual, e.g., using crowdsourcing) to fully automated.

As part of the work on the common European data space for cultural heritage, an Enrichments Policy[3] has been created, which lays down principles to bring direction and consistency across enrichment efforts and ensure that enrichments bring value. Some of these principles focus on the quality of enrichments, and call for a "methodology for validating enrichments designed based on predefined guidelines for evaluation and agreed quality targets, shared prior to the execution of enrichment efforts". This document is set to articulate such a methodology.

This document is also a complement to the Europeana Publishing Framework (EPF), which encourages data producers to contribute richer metadata and content. While ongoing work on the EPF seeks to measure the impact of enrichments in quantitative terms - number of statements, presence of entities with multilingual data, etc. - the methodology presented here focuses on assessing the quality of the enrichments produced.

Recommendation: the process exposed here should be included in project proposals so that no enrichments are ingested in the data space without prior validation.

---

[1] See definitions of *enrichment* and *enrichment effort* in the Enrichments Policy referenced below.
[2] See https://pro.europeana.eu/page/europeana-semantic-enrichment#enrich-your-own-metadata
[3] https://pro.europeana.eu/post/enrichments-policy-for-the-common-european-data-space-for-cultural-heritage

Before these enrichments - or the tool that produces them - are shared in the data space, their quality must be validated. If acceptance criteria are not met, the enrichments are either rejected or pushed back for improvement[4].

This document sets up a process and guidelines for the validation of enrichment processes or their outputs when these are to be integrated in the data space's core services (Europeana). It is aimed at any party that will be involved in (a part of) such a validation process.

This process should be applied:
- for projects delivering enrichments, upon each ingestion milestone
- for external services contributing enrichments to Europeana and software to be integrated in Europeana, the first time it is integrated and then on a regular basis

Ideally, aggregators which perform enrichment should seek to regularly validate their enrichments, especially when these are applied to datasets different from those who were previously subjected to enrichment (different types of object, different languages, depending on which part of the data or metadata is used as base for the enrichment process).
The guidelines could also be used in a wider context, i.e., when enrichments are not integrated in the core data space services.

The methodology consists of two parts: 1) analysis and documentation of enrichment cases and their validation, and 2) actual validation of enrichments. Note that these parts are complementary and should be ideally carried out in parallel: establishing relevant documentation requires that most of the process for validating the enrichment effort has been fully specified and already started, if not completed!

The validation efforts may be carried out by partners or the Europeana Foundation (EF) - ideally both being involved in some way. Each part or sub-part specifies the actors expected to be involved and their role. Many projects are prepared so as to include an evaluation effort to assess the quality of the enrichments they produce. But the operator responsible for the quality of data in the data space should always be involved in the final vetting.

---

[4] Note that validation should aim to assess the overall quality of the enrichments produced by a project or a tool, but it does not need to result in a global accept/reject decision: the methodology can for example help to select a subset of the produced enrichments that is deemed trustable enough. For example see the description of Task 3.4 ("Enrichments integration into Europeana CSP") for the Jewish History Tours project:"T*he task aims to design a validation methodology for assessing quality of the automated geographical enrichments delivered to EF within the Action and define the threshold above which automatic enrichments will be considered reliable.*"

2. Part 1: Analysis and documentation of enrichment cases

In this part, the partner in charge of creating and/or running the enrichment effort provides information about the effort, using a structured questionnaire. The Europeana Foundation advises and verifies that the information presented in the questionnaire is informative enough.

Understanding the context and goal of an enrichment effort should be the first step in any validation. To support this goal, we have created a questionnaire to be filled by each project or individual enrichment effort (see Appendix). This questionnaire aims to identify:

- the main principles of the enrichment performed, especially, its type of outcomes, and how confidence scores are produced
- the datasets that the enrichment is performed against
- the amount of enrichments (to be) produced by a project

This analysis and documentation effort includes the description of any quality evaluation actions planned for the enrichment process and results, which may of course include validation actions as described below. This could include the quality requirements that the enrichment must meet (otherwise they would have to be determined in the next phase).

Another important part of analysis and documentation regards the identification of enrichment issues[5]. Be aware of the possible issues found for your type of enrichment. Make a list and give everyone in your project a chance to contribute to it. It is likely that new issues are identified during the project, therefore this list may eventually be updated at any moment of the enrichment process lifecycle.

The output of this phase is not necessarily a self-contained, independent document. It can refer to more complete documentation, such as project deliverables. What matters is that one can access all relevant information. Do not fear giving complete documentation, as it helps build trust[6].

In line with the effort of documenting as much as possible the process, as the project progresses, it should make evaluation data openly accessible whenever possible, be it training data used before validation or evaluation data used for validation. This contributes to building trust in the overall process and it benefits the common

---

[5] Enrichment issues include for example faulty automatic translation of a metadata field, linking an object to a place with wrong coordinates, etc.
[6] Some projects report for example on "tailoring of enrichment tools for the cultural domain", but they do not document how they have done it.

European data space for cultural heritage at large, by allowing others (including Europeana, if it needs it) to re-use the data in the future[7].

To give enough time to validate the enrichment efforts, we expect the questionnaire to be filled way before - at least 3 months - the deadline for acceptance of the enrichment efforts.

3. Part 2: Enrichment acceptance validation workflow

Acceptance validation is based on the following workflow: a sample of enrichments is extracted and validated against a quality criterion to guarantee that the quality meets the requirements. The requirements and quality criteria should be agreed with EF and if the threshold is not met, the enrichments are either rejected or sent back for improvement.
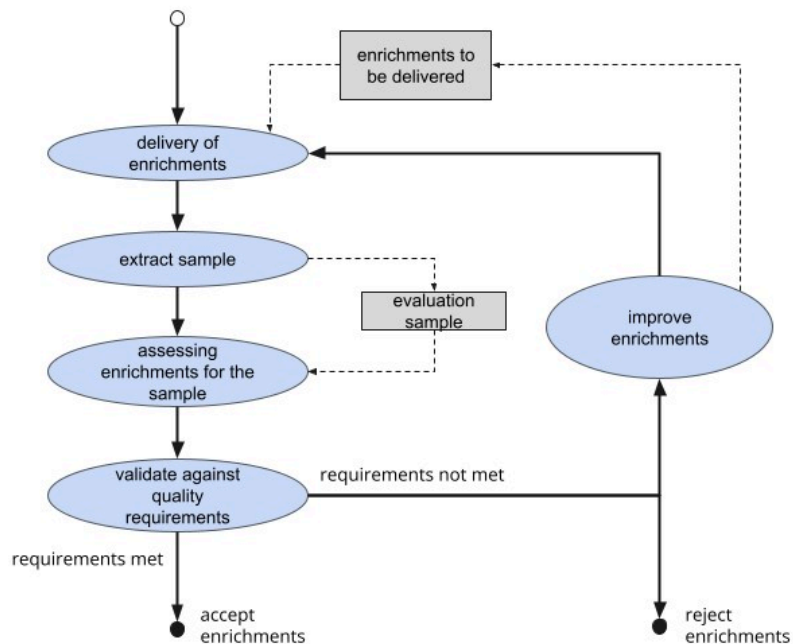
Of course this validation is not needed when all enrichments are manually reviewed in a trusted way during their production process, unless the manual review is not 100% trusted. In this case it should be confirmed the same way a confidence score should be confirmed, cf. below.

The following sections detail the actions to be undertaken in the various steps of the workflow.

---

[7] The datasets could also be referred to on the AI4Culture platform, which will seek to support capacity building actions and a wider sharing of data and tools for AI across the data space.

3.1. Step 1: Sampling enrichments for validation [mandatory]

In this step, a data provider or a partner in charge of designing and/or running the enrichment effort selects a sample of data for evaluation. The Europeana Foundation may advise, and verifies that the sample is representative enough.

The goal here is to select a sample of enrichments that allows to validate the general set. This sample shall be relevant for the goal which guided the production of enrichments. Especially, the sample must be representative of the enrichments and the original datasets that these enrichments apply to, quantitatively and qualitatively.

In addition, the sample should have an adequate size that can provide statistical confidence of the results. The minimum size of the sample will depend on the exact assessment to be done by the evaluators. For example, in the simple case where the evaluators assess whether an enrichment is correct/incorrect, and we specify[8] a confidence level of 95% and a 5% margin of error (i.e., the statistics calculated from that sample will be within 5 percentage points of the real population value 95% of the time), then the sample should include over 385 enrichments.

NB: in some cases, a pre-existing gold standard[9] can be used to validate the enrichments. In such a case the 'sampling' can be quite straightforward. Still it is important to verify that the gold standard is representative.

The extraction of the sample is expected to be automated or semi-automated. A stratified sampling approach can be followed to ensure better representativeness, where the following criteria can be used - possibly in combination - to define sampling strata:
- representativeness of enriched dataset:
  - general characteristics: domain, origin of collections (data provider and/or aggregator), type of objects, topical coverage, temporal coverage...
  - for metadata enrichment, type (property) of source statement, characteristics of source value
  - for content-based enrichment, format and size of media, media quality, language, type of font (for text recognition)...
- representativeness of enrichment target[10]
  - type of target (e.g., class of entity)
  - topical coverage (entity belonging to a vocabulary or sub-vocabulary)...

---

[8] Such criteria should of course be agreed among partners prior to assessment, depending on the requirements of the specific enrichment effort(s).
[9] An annotated dataset that includes all the enrichments required and only them. For example, for subjects, the gold standard would have records enriched with all the subjects that should appear. An enrichment process then should find all those subjects, and only those, to be 100% accurate.
[10] These criteria are probably relevant to a lesser degree, unless the application case specifically requires a focused (or balanced) enrichment output.

- representativeness for a specific application, i.e., having an "extrinsic evaluation" in mind. For example by selecting a set of search queries and evaluating the effectiveness of enrichments (e.g. translations) to bring relevant search results for these queries.

NB: the last two categories are probably less relevant in a "basic" Europeana case where enriching source data is an objective in and of itself. Some application cases may however specifically require a focused (or balanced) enrichment output.

For examples, see the decisions made in the Europeana Subtitled project[11] (for representativeness of collection and types of videos), Europeana Foundation's AI experiments (for representativity of type of content and collections represented), the discussion in Europeana Translate[12] for representativeness of collections and metadata fields, and the evaluation of the Spanish pilot for multilingual search at Europeana[13] (for representativeness of queries). We can make available relevant documentation upon request.

3.2 Step 2: Assessing enrichments for the sample dataset [mandatory]

In this step, a data provider, a partner in charge of designing and/or running the enrichment effort, or a third party evaluates the enrichment produced for the selected sample. The Europeana Foundation advises on the methodology and may help with the evaluation itself (depending on experience required).

The enrichments for the selected sample dataset must be assessed for their quality, ideally in terms of fitness for a application case that the enrichment is meant to support (display on Europeana, improved search, assistance to data curation, etc).

Typically this step aims at evaluating if the results of the enrichment are correct (*precision*) and complete (*recall*), where the former is usually regarded as more important than the latter. But quality is not only about precision or recall. Other aspects of information value may be considered; for example when enrichments are very general and repeated across many objects in a dataset, or when the same enrichments are repeated for one object. Sometimes the aim of the assessment could be to compare a new enrichment process with an existing one that is judged trustful enough.

The assessment is often done manually, where human evaluators examine the enrichments produced for the sample dataset and assess whether they are appropriate. In some cases the assessment can be partly or fully automatized, for example when:
- there is a pre-existing gold standard (cf. previous section) that provides an appropriate reference of enrichments for the sample set.

---

[11] https://pro.europeana.eu/project/europeana-subtitled
[12] https://pro.europeana.eu/project/europeana-translate
[13] https://doi.org/10.5281/zenodo.7123543

- there is a *trusted* confidence score[14], and the score for the enrichments produced for the sample set is above an agreed confidence threshold. This requires that a prior validation of the truthfulness of the score has been conducted[15].

Further, the manual assessment of enrichments, as an exercise quite akin to creating descriptive metadata for an item, raises important problems of inter-annotator variability. Some tasks, for instance determining the subject of an item or identifying relevant visual elements in an image, are notably hard to get an agreement on, because of their high subjectivity: what makes a result relevant is often in the eye of the beholder and depends on their view on the application. This issue should be considered in the assessment process, for example by having enrichments evaluated by several annotators, as was done for computer vision-generated enrichments in the Saint George on a Bike project[16] (relevant documentation available upon request).

Sometimes the assessment task will require to refine the assessment methodology during the project, especially when it's a new kind of enrichment, or when the quality requirements of the project are not precisely defined at the start.

The following recommendations should be considered:
- If you plan an evaluation campaign, test it sooner rather than later, and as many times as possible. Insights gathered while having would-be evaluators exposed to successive (fast) iterations of mock-ups, even on paper, are more useful than dozens of calls discussing an evaluation interface in the abstract. Make sure however that these discussions involve some real enrichments to be evaluated, even if it is for a handful of objects only.
- Try to re-use existing tools (e.g., LabelStudio, which has been used by the Europeana XX project[17] and Jewish History Tours project[18]) as much as possible. Do not be afraid of using simple solutions like Google Sheet. Anything works, especially when it allows rapid prototyping of the evaluation.
- Dedicate a lot of effort to writing good guidelines for human evaluators. Good documentation speeds up the evaluation process and alleviates some typical evaluation issues like variability. Do not hesitate to test documentation with would-be evaluators. It is generally very useful to mention typical examples of enrichment issues (see Appendix) that an evaluator may encounter. Do it carefully though, as it could bias an evaluation by having evaluators focus on the same problems again and again.

---

[14] I.e., enrichments are provided with a confidence score, which properly reflects the probability of an enrichment to be correct.
[15] This validation of the confidence score can be carried out using a process similar to the one presented here. In fact it may be included as a side objective of the acceptance validation workflow, see following section.
[16] https://saintgeorgeonabike.eu/
[17] https://pro.europeana.eu/project/europeana-xx
[18] https://pro.europeana.eu/project/jewish-history-tours

3.3. Step 3: Validate enrichments against quality requirements [mandatory]

In this step, a partner in charge of designing and/or running the enrichment effort computes relevant metrics based on the evaluated sample. The Europeana Foundation advises and verifies that the level of quality of enrichments (or a part thereof) meets the agreed target.

The enrichments submitted for validation shall be tested against the quality requirements and approved or not for ingestion in Europeana.
This is intended to be an automatic process, exploiting the results of the previous steps.

In this step, evaluation metrics such as precision, recall and F-measure are needed. Some specific decisions need to be made, and we foresee that the following points can be relevant:
- Deciding a specific quality metric and KPI before the project starts is a good idea to force the project to embed quality checking work in its process. But it seems counterproductive to close the door to using other metrics and KPIs in the course of the project, when these would appear to be more relevant in the light of discussions between the partners involved, and/or the progress of the state-of-the-art.
- Evaluators should strive to have evaluation measures that try to better reflect the impact of quality issues. For example, it is likely to be preferable to have precision/error rates based on a "micro-averaged" approach (i.e., counting every instance of enrichment) than a "macro-averaged" one (i.e., counting whether a general 'enrichment rule' is correct or not). Imagine a situation where 1 wrong enrichment rule negatively impacts 1000 records and 99 correct enrichment rules impact 1 record each: the general impression for users would be very negative despite the fact that 99% of the rules are correct. In a way, the macro-averaged figure tells us about the conceptual quality of enrichment but less about its practical impact.

We envision two main options with respect to exploiting these indicators to perform validation:
- Global acceptance/rejection. The evaluation in the earlier step meets the quality expectation agreed for the enrichment case. For example, when the precision of evaluated enrichments is 95% while a threshold of 90% would be deemed acceptable.
- Partial acceptance/rejection: only some enrichments are accepted because they correspond to a stratum for which the assessed enrichments meet the quality expectation (while enrichment on other strata may fail). For example: (1) select all enrichments with a confidence score above 85%; (2) select translations that are computed for metadata values longer than a certain number of words. As for the sampling, the filtering could combine several criteria; for instance, accept only translations of long-enough descriptions in French and German.

The crucial point is to identify the quality requirements, namely the criterion to use (overall precision, confidence score…) and the desired level of quality (threshold for the score). At this stage we do not have general guidelines yet. Consensus should emerge in the context where the enrichment occurs, e.g., between the partners of the project that applies the enrichment. This could for example be agreed on during the preparation of the project.

In the future, we hope to further identify general quality requirements that can be considered across projects, as part of the Enrichments Policy for the common European data space for cultural heritage. For this, ideally all enrichment tools should output a *meaningful* confidence score, which would allow the deployment of reliable filtering. This endeavour however faces two critical challenges:
- It is going to be hard to obtain confidence levels that behave in a comparable way across tools that employ very different techniques to produce very different output.
- Even for one tool, some partners may have quite different quality expectations, i.e., they would use different thresholds for the same score. This seems undesirable but some providers may require it.

3.4. Step 4: Using assessment to identify reliable confidence score and thresholds [optional]

In this step, a partner in charge of designing the enrichment process defines a confidence score that can be used to filter good enough enrichments. The Europeana Foundation may advise.

If the enrichment tools output a confidence score for the enrichments, try to measure whether manual evaluation confirms with automatically (or generically) generated confidence scores. Some projects compute confidence scores and do manual evaluation but there is no data on whether the confidence level is correlated with manual acceptance.

Establishing reliable confidence scores and associated thresholds can be based on assessment efforts like the ones presented here, where evaluation is used to:
- confirm an existing score and threshold: when one filters out all enrichments that are below the threshold, the evaluation confirms that the remaining enrichments meet the acceptability requirement.
- identify a threshold for an existing score: sometimes an enrichment tool comes with a generic confidence score but it has not been tested in the specific conditions of Europeana. Evaluations can help to define a threshold above which enrichments meet the acceptability requirement. In the project Europeana Translate, a sophisticated formula was applied to the original translation engine's confidence score, so that a threshold can be later applied to retain translations that correspond to a certain level of human quality assessment.

Evaluation can sometimes help to create a new score and threshold if there are no pre-existing ones. For example, a translation process can assign higher confidence levels to translation of longer metadata values, when they are in languages for which previous evaluations have confirmed better performance of the tool. Or an enrichment process may be assigned a 100% confidence level when it is never proven wrong by evaluation. As a matter of fact it is not always possible to adapt confidence scores, for example, simply because it is too late in a project's life cycle, or because an enrichment process is based on output by tools (or AI models) that are used out-of-the-box, as was the case of the geo-enrichment in the Jewish History Tours project. In such cases, as a clearly less optimal solution, one could envision to apply to all enrichments the accuracy (precision) that was measured in the sample evaluation.

3.5. Step 5: Improve enrichments [optional]

In this step, a partner in charge of designing the enrichment process tries to improve the enrichment results. The Europeana Foundation may advise.

This step includes any kind of improvement effort (automated, manual, or both) that can make the enrichments meet acceptance thresholds, such as
- crowdsourcing campaigns for validation and correction of enrichments
- improving enrichment software

For some examples of action that could enhance an enrichment process, see Appendix "Recommendations on enrichments".

While trying to alleviate the enrichment issues observed during evaluation, keep an eye on potential evaluation biases. For example, it could be that a project reports that enrichments are very correct but users do not find them so useful, while this feeling is caused by issues that apply only to a small part of the evaluated dataset.

5. Conclusion and future work

This document sets up guidelines for the validation of enrichments in the data space context, so that no enrichments are ingested in the data space without prior validation. It is aimed at any party that will be involved in (a part of) such a validation process. Especially, we believe it should be considered in project proposals that include enrichment efforts

This version is not final. We expect other versions to be produced in the future, integrating more relevant work from projects, and continuing to solicit feedback from relevant stakeholders in the Europeana(Tech) community and beyond[19]. The 2024 action

---

[19] The EuropeanaTech Data Quality Committee has already been consulted and the methodology and its application in the projects JHT and EnrichEuropeana+ have been presented at two technical conferences: Semantic Web in Libraries (SWIB23, https://swib.org/swib23/) and the

plan for the data space Enrichment Policy also includes actions that can have an impact on the methodology presented here, especially with respect to the information requested for enrichment efforts:

- review and extend EDM provenance profile
- run a questionnaire to identify resources consumed for enrichment efforts
- develop an enrichment policy compliance checklist

We will also keep updating this document by gathering feedback received while applying this methodology in coming data space projects.

The main areas of consolidation that we have identified are:

- consolidating the list of recommendations, including further identifying general quality requirements that can be considered across projects, and proposing recommendations for the publication of evaluation data and reports
- consolidating the list of example issues.

International Conference on Dublin Core and Metadata Applications (DCMI 2023, https://www.dublincore.org/conferences/2023/) under the title "Towards a methodology for validation of metadata enrichments in Europeana".

Appendix - Enrichment questionnaire

Please make a copy of this for your project

Enrichment evaluations carried out by [name of the project]

Objective

This document is intended to collect information on how evaluations were carried out by the project for enrichment tasks (e.g. named entity recognition, translation). The objective of this is to have information on delivered enrichments and methodology applied to assess them (for example how quality is measured to produce confidence scores) that could help Europeana, after the end of the project, to get recommendations on good practises and thresholds to be applied during the exploitation of enrichments. An initial set of recommendations for enrichments can be found in Section 6.2 of the report on the Comparative evaluation of semantic enrichments[20].

Enrichment evaluations

Please make a copy of the table below for each different source or type of enrichments submitted to the data space during the duration of the project e.g enrichments/annotations/subtitled/crowdsourced annotations. Also replace the text in italics with the answers specific to your work.

Enrichment performed and outcomes	
Goal of the enrichment	*Explain what is the expected benefit for the enrichment and how it will bring value to the collection.* *E.g improve discoverability of the item and help navigate across collections* *E.g. improve readability and accessibility of text*
Source of enrichments	*E.g., crowdsourcing, expert annotation, specific (automatic) enrichment tool - please provide links or references where possible.*
Type and target of enrichments	*E.g. semantic enrichment (with persons, locations, concepts), translations, etc.* *Also, when applicable, indicate the target of the enrichment such as the data source ie. Wikidata or another linked data vocabulary in case of semantic enrichment, and the type of link (metadata field) used to express the semantic enrichment*
Dataset(s) selection	*Criteria used to make the selection of the dataset(s), size of resulting dataset*

Overall enrichments produced by the project and delivered to Europeana Foundation	*Total enrichments done and enrichments delivered to Europeana from this source. Please indicate the number of enrichments and the number of records enriched.*
Overall enrichments produced by the project and accepted by Europeana Foundation *(to be filled in in collaboration with Europeana)*	*Total enrichments ingested in Europeana from this source. Please indicate the number of enrichments and the number of records enriched.*
Enrichments delivered and not accepted and motive (to be filled in in collaboration with Europeana)	*Total enrichments delivered to Europeana but not accepted (e.g., the vocabularies were not supported, the enrichments already existed). Please indicate the number of enrichments and the number of records enriched.*
Other issues	
Quality Assessment	
Framework	*For evaluation, please indicate details on tool(s) and methodology employed (for example the revision process during crowdsourcing). NB: some details are expected to be indicated in the following rows.*
Coverage	*Indicate whether evaluation was carried on the full set of enrichments or a part of it.*
Reviewers / Annotators	*Indicate who were the reviewers/annotators, e.g. providing institutions, aggregators, partner experts from the project, external experts, end-users from a specific audience such as teachers.*
Results	*Details on results should include statistics on accepted and rejected enrichments.*
Source and/or rationale for the confidence scores (or any other quality metric)	*Please give details about the assessment methodology followed (e.g. the level of agreement between reviewers of crowdsourcing) and/or the tool employed that calculates the confidence score for (final) assessment of individual enrichments.*
Recommended threshold for confidence scores	*Indicate here if you have a recommendation regarding the exploitation of enrichments (e.g., for search and display purposes).*
Main issues observed during evaluation	

Appendix - Enrichment efforts for which this methodology has been applied

The table below presents some efforts (mostly in projects) where the methodology has been applied - especially, the enrichment questionnaire has been filled for these projects and can be made available on request.

Project/effort	Type of enrichment
Europeana Subtitled	closed captions, subtitles
PAGODE	metadata semantic enrichment (places, persons, organisations, custom concepts, PhotoVocabulary terms, Wikidata terms)
Europeana XX	metadata semantic enrichment, translation, subtitles
WEAVE	metadata semantic enrichment (places, persons, organisations, custom concepts)
EnrichEuropeana+	metadata semantic enrichment (places, persons, dates), translations,language detection
Europeana Translate	translation
Jewish History Tours	metadata semantic enrichment (places)
AI4Culture	*ongoing*
De-Bias	*ongoing*
5Dculture	*ongoing*

Appendix - Recommendations on enrichments

This appendix gathers recommendations that were gathered from various validation efforts - notably in the Europeana XX and PAGODE[21] projects (relevant documentation available upon request):
- Invest more time making sure that the selection of the data you will be enriching matches the objective of the project. If possible, involve the respective data providers in that process.

- Review the selection to make sure that the data selected do not have quality issues. If that's the case, fix the data quality issues or remove such resources from the selection.
- Analyse and take into consideration the original metadata that is existing translations/enrichments of the object's fields (e.g. main or/and alternative title, description, subject etc.) in order to avoid providing duplicate translations/enrichments.
- Produce enrichment that exploits at most the possibilities for fine-grained semantic data offered by the Europeana Data Model
- Follow existing guidelines for interoperability and quality of data, such as using or aligning to the vocabularies already supported by Europeana, try to have your vocabularies supported by Europeana, and produce metadata values that can be exploited by the existing data normalisation processes at Europeana.
- Consider investing more time testing the automatic tools to discover and fix bugs as early as possible and before submitting the data to Europeana. For example, make sure that the translation tools handle special characters appropriately. Check lists of common errors that could apply for similar enrichments (see other appendix here for more examples)

Appendix - Example issues with validity and relevance of metadata enrichments

The following list has been gathered during the iterative development and testing of the geo-enrichment service developed by the Jewish History Tours project (documentation available upon request). Some of these issues have been fixed in the last version of the service but they could be nonetheless informative for other enrichment efforts.

- Enrichment is correct, but too general or incomplete. I.e. out of several possible locations only one was found or found city only, while street or place were in the original text. For example, "Heroic act of the Hero of the Soviet Union Jr. Lieutenant M. P. Zhukov" ca 1943" enriched with "Soviet Union" or "Receipt for the train ticket from Moscow to unknown destination, 1940s" enriched with "Moscow". These enrichments are still useful for the users due to their relation to the geo-entity. For example, if a user is looking for "RSFSR actors" or "Moscow train", these enrichments will bring relevant results during search. However, showing a train ticket from Moscow to another city on a map of Moscow is going to be quite less relevant than a photo of a train in a Moscow station.
- Something in an enrichment is correct, though main locations are not found. For example, the same street name in different cities. Or enrichments with concepts or places mentioned in the metadata but not directly relevant to the main topic of the item. "Solomon Mikhoels and Evgenii Kiselev, USSR Consul General in New York, Philadelphia, 1943" is enriched with the Consulate-General of Russia in New York City. But the item is a photograph taken in Philadelphia, not New York.

- Enrichments that are technically correct but pick on a metadata aspect that is not core to the target application. For example "[Osip Liubomirskii's biography of Solomon Mikhoels, Moscow, 1938](#)" enriched with "[Moscow](#)", while Moscow is "only" the publication place of the book. The book content itself is not necessarily centered on Moscow.
- Enrichments based on metadata values with question marks in the metadata, as in ("[Ester Karchmer, GOSET actress, Moscow (?), 1920s](#)"). The data provider found this is probable enough to be written in the title so it should be included among the enrichments.
- Enrichments that are for known places but without a link to an established vocabulary like Geonames or Wikidata, while such vocabulary was a target for the enrichment.