



MS3. User requirements and use cases definition

Version 0.8

Documentation Information

Contract Number	2018-EU-IA-0104
Project Website	www.saintgeorgeonabike.eu
Contractual Deadline	01/03/2020
Nature	Report
Author	Albin Larsson (EF), Antoine Isaac (EF)
Contributors	Mónica Marrero (EF)
Reviewer	Maria-Cristina Marinescu (BSC), Ariadna Lobo (BSC), Maria Cristina Marinescu (BSC), Joaquim More (BSC), Sergio Mendoza (BSC),
Keywords	Use cases, requirements, Europeana, artificial intelligence



Co-financed by the Connecting Europe Facility of the European Union

Change Log

Version	Author	Description Change
V0.1	Antoine Isaac (EF), Mónica Marrero (EF)	Initial Draft
V0.2	Maria Cristina Marinescu (BSC)	Reviewed
V0.3	Ariadna Lobo (BSC)	Reviewed
V0.4	Joaquim More (BSC)	Reviewed
V0.5	Maria Cristina Marinescu (BSC)	First iteration
V0.6	Ariadna Lobo (BSC)	Reviewed
V0.7	Albin Larsson (EF) Antoine Isaac (EF)	Second iteration
V0.8	Maria-Cristina Marinescu (BSC), Sergio Mendoza (BSC), Ariadna Lobo (BSC)	Reviewed

Table of Contents

1. Introduction -----	3
2. Target audiences -----	3
3. Use cases -----	5
3.1 General service for enriching collections -----	5
3.2 Ingesting results from general enrichment service into Europeana -----	6
3.3 Search based on enrichment-----	6
3.4 Populate a crowdsourcing tool with candidate enrichments -----	7
3.5 Upload in data sharing platforms -----	8
3.6 White-label, custom trained enrichment service [TBC]-----	9
3.7 Support enhanced accessibility of Europeana content for visually impaired persons. [TBC]	9
3.8 Rejected use cases-----	9
4. Requirements-----	9
4.1 Media type of data input and semantic level of output -----	10
4.2 Scope and data coverage/constraints-----	11
4.3 "Iconographic level" requirements -----	12
4.4 Data vocabulary requirements -----	13
4.5 Rights requirements -----	13
4.6 Language requirements -----	13
4.7 Captions for accessibility [TBC]-----	14
Appendix A. Technical options for supporting end-user semantic search-----	15
Appendix B. An example of computer-aided tagging campaign on Wikimedia Commons -----	17
Appendix C. List of types of image objects-----	18

1. Introduction

In our experience the sources of data for cultural heritage images lack good metadata that relates to the content, although they may have descriptions that give information on the authors (of the art piece or of the picture of the art piece), style, and period. The ultimate goal of the Saint George on a Bike project is providing rich information about what is depicted, when the metadata available is insufficient.

To better focus its effort, the project's Activity 2 set ups to better identify end users, use cases and system requirements, focusing on Europeana's main interests but also considering the possibilities given by the existence of the data and their suitability for the proposed type of AI-supported analysis.

This milestone represents a second iteration of this work, listing in Section 2 the first use cases we have identified, breaking them down into significant user actions (steps). We also begin to list possible requirements and constraints in Section 3, which are not necessarily use case specific at this stage, but rather aim at capturing the various options that can apply to several cases, for example, narrowing down the scope of the datasets the use cases could be first applied to, identifying the various "shapes" of enrichments and their iconographic level, etc.

At this stage of the project, we have decided to not make final decisions on some of these points, as we felt we needed more insight on the feasibility of applying AI technology on the data we can gather in order to produce the desired output with some reasonable chance of success. The challenges raised by data availability and the complexity of the problem vary indeed greatly from one option to another. Some of the options and variants available for the use cases (such as the potential crowdsourcing environments) are still open at this stage (see "TBC" items), and the best venue may highly depend on the choices made for the types of enrichment and coverage of training for the enrichment service.

We will therefore produce at least another iteration of this document, which will further define the use cases and - especially - the context in which we seek to apply them. Note that this version may remove some of the use cases we have identified, if they prove impossible to realize, or if they would deliver too little value considering the human and data resources at hand.

Our plan is to release the next version together with the results of the activity's second tasks, focused on identifying the technical and functional requirements on the project's enrichment service in order to meet the specifications of the use cases. With this plan in mind, we have begun to write down some functional and technical considerations, which may not be directly related to the definition of the use cases themselves but may become highly relevant in the next stages.

Note that the third task of Activity 2, the definition of the HPC resources needed and of the scalability and portability of the services is out of scope of any further iteration of this document. We expect this to be rather discussed in MS5 and MS6.

2. Target audiences

The project proposal has identified several main target audiences:

- Research institutions, particularly those interested in the European cultural heritage, history, and anthropology
- Cultural and creative industries (advertising, architecture, arts, crafts, design, fashion, film, performing arts, publishing, etc)
- Cultural tourism
- Cultural institutions such as museums
- Public and private educational institutions
- Public administration bodies interested in preserving the European cultural heritage

- Citizens
- Pan-european policy makers

Among them, some correspond to Europeana's core "markets"¹:

- Cultural Heritage Institutions willing to share digitized cultural heritage objects with a wide audience
- Academic researchers, creative industries and education

The use cases we present in the following section have been guided by these audiences, with some cases focusing on the needs of certain audiences:

	CHIs or related cultural projects	Researchers and research institutions	Creative industries, incl. tourism	Education	Citizens	Pan-european policy makers	CHIs or related cultural projects
3.1 General service for enriching collections	++	+	+			+	3.1 General service for enriching collections
3.2 Ingesting results from general enrichment service into Europeana	++ (for Europeana partners)						3.2 Ingesting results from general enrichment service into Europeana
3.3 Search based on semantic enrichment	+	+	+	+	++		3.3 Search based on semantic enrichment
3.4 Populate a crowdsourcing tool with candidate enrichments	++	+		+	+		3.4 Populate a crowdsourcing tool with candidate enrichments
3.5 Upload in data sharing platforms	+	++	+	+		+	3.5 Upload in data sharing platforms
3.6 White-label, custom trained enrichment service	++	++					3.6 White-label, custom trained enrichment service
3.7 Support enhanced accessibility for	+				++	+	3.7 Support enhanced accessibility for visually

¹ <https://pro.europeana.eu/what-we-do>

visually impaired persons							impaired persons
3.1 General service for enriching collections	++	+	+			+	3.1 General service for enriching collections

3. Use cases

This section presents the outcome of a first effort of identifying relevant users and use cases for the work of the project. Note that in the contexts of a large data aggregator like Europeana, or of scientific infrastructures like EUDAT, the actors involved in a use case may not be necessarily the "end user" chiefly targeted by the part of the services that is more visible to the general public. For example, Europeana intends to deliver value to a range of Cultural Heritage actors that work for Cultural Heritage Institutions and related projects. Needless to say, the goals of these actors will be eventually to serve an audience of "end users" in the more intuitive sense. However, the best way to serve these goals may not be easy to identify at the level of a back-end service like the one developed by Saint George on a Bike, if just because one "back-end use case" may serve several "end user use cases" or even other back-end use case, as when a search service is provided as part of an API that will be used by applications not yet designed.

In this section two use cases stand out. The first is enriching Europeana collections, as a generic back-end use case where metadata on Europeana images is enriched in order to serve a diversity of applications built on top of metadata, including some other use cases in this section. The second is searching based on enrichments, as it corresponds to a functionality that Saint George on a Bike has been chartered to deliver as part of Activity 8 "End user general service definition".

In this section we also include a subsection on rejected use cases, which lists the scenarios that we have envisioned in the earlier stages of our effort but that we decided to drop when realizing they would be too difficult and/or not interesting enough. We include them should they prove interesting ideas for further projects.

3.1 General service for enriching collections

This use case provides a general service/technical interface enabling enrichment of cultural heritage datasets. This interface would allow Europeana to integrate this service in its ingestion and enrichment process (use case 3.2). This service could also potentially be used by aggregators and data partners as well as other third parties, independently from the results being ingested in Europeana.

Note that this use case is different from the use case outlined in 3.6 as this is a generic service and not a (customizable) service trained on custom datasets given by users.

This technical interface (API) should take the input defined in 4.1 "Media type of data input and semantic level of output" (such as an image) and return the data defined as output in the same section together with a confidence score for each enrichment. The level of output could potentially be configurable by the client implementing the interface.

It has to be taken into account that this is not a real time scenario and therefore the progress of an enrichment task needs to be represented and communicated through the service.

3.2 Ingesting results from general enrichment service into Europeana

In this use case Europeana ingests the results or a subset of the results of the general enrichment service covered in use case 3.1, so that this data can be exposed through Europeana's APIs and its clients. This includes quality assurance/ingestion processes, and storage of the data.

The enrichment process could be triggered by Europeana data officers and go through the following generic steps:

1. Select a collection or a part of it, usually based on a metadata-based query.
2. Run our enrichment tool to generate metadata for each of the images in the input dataset.
3. Assess and validate the results manually, possibly on samples.
4. Load (good quality) resulting metadata in Europeana's data environment.

We contemplate various scenarios for results validation. The basic methodology is validation by Europeana staff, including data officers. We will explore other approaches such as crowd-based validation as detailed in use case 3.4.

3.3 Search based on enrichment

The project is set up to implement a search service based on enrichments produced with AI technology. This would make Europeana objects easier to discover based on richer criteria related to image content and potentially higher level concepts recognized in image content (e.g. icons and symbols). This use case corresponds to the project's Activity 8.

This text[metadata]-based image search tool will allow users to understand how common patterns depend on the historical moment, the symbolic, and other cultural particularities of cultural significance. It may be used for many research, education, cultural, or social projects. The main challenge here is to understand the context at the time that an image was created, context outside of which the symbols, traditions, and rules that it reflects lose meaning. The hope is that by using both images and text, and linking them via Linked Open Data² (LOD), we can find relevant, hidden, and related information, from which the context can be created and learned.

In this use case, the end user of Europeana's portal or a similar service:

1. enters a search query that corresponds to her information need
2. gets a list of Europeana objects as results

This use case has two (independent) variants of the basic case. The first variant, on the input side, allows the user to select a "semantic" search term which corresponds to the URI of a concept, person, or place that represents her search. This URI could come from a controlled vocabulary or a similar resource, such as Europeana's own "contextual entities" gathered into the Europeana Entity Collection³ from various third-party sources, and used for enrichment, search and browsing⁴ purposes. The selection of the semantic resource corresponding to the query can be made using an assist function such as auto-suggestion of candidate matches for the user query. Such interactive query building can be especially valuable in the education context, where keywords are not always mastered by users. "Semantic" search also enables better handling of multilingual issues, where the data for contextual entities can come in many languages.

² <https://www.w3.org/DesignIssues/LinkedData.html>

³ <https://pro.europeana.eu/page/entity#entity-collection>

⁴ A specific application case is the detection of images in the Europeana Collection, which depict a specific entity, and can be used as illustration on the web page Europeana maintains for this entity.

The second, on the output side, lets the system give some rationale to the user, on why some results are shown to her, for example by presenting her some extracts from the database (semantic enrichments) that underlies the search function.

There are two aspects that determine which features we will provide:

- whether the service will be visible on the Europeana portal or will be developed and hosted as a project-specific effort, at BSC.
- whether the search functionality will be supported by Europeana or not (in terms of search and storage), regardless of visibility.

The answer to these questions will not only be based on decisions regarding end user priorities. The technical feasibility of re-using existing components will play a key role. Appendix A provides an initial discussion of this matter.

Note that in any case, the search service needs to be multilingual, which may require (automatic) query or metadata translation.

3.4 Populate a crowdsourcing tool with candidate enrichments

In this use case, we intend to set up or incentivize crowdsourcing campaigns in which the end users may validate metadata enrichments, either as a purpose in itself or as a means to access another service. The manager of such a campaign could be a Europeana data officer, a Europeana data provider or aggregator, or the owner of a crowdsourcing tool like WithCrowd.

This use case includes the following steps, from the manager perspective:

1. Select a collection
2. Run enrichment tool
3. Load results (annotations) in an annotation tool that allows validation of annotations.

From the end user (crowd) perspective, the use case involves:

1. Accessing an object and inspecting the enrichment to validate
2. Validate the enrichment, according to the possibilities offered by the crowdsourcing tool (ticking suggestions, commenting...)

The possibilities will be much influenced by the context of the campaign: type of enrichments, annotation tooling, engagement strategy, etc

To begin with, we will have to determine which communities are a good target for an annotation campaign in our context. It could be the general public or a more specific, "niche" audience. Since the Europeana V2.0 project⁵, Europeana and other projects have run annotation pilots and campaigns around specific domains ("nichesourcing"), some of them featuring the validation of automatically produced tags. The domains involved includes sound archives (Europeana Sounds), historical manuscripts (Enrich Europeana), fashion (CrowdHeritage⁶), art and natural history (SEALINCmedia), etc.

We could also target communities that have general interest in crowdsourced data creation, especially the Wikimedia community and its Wikidata sub-community. Our enrichment service could for example be implemented as an backend to the MediaWiki extension MachineVision⁷. This extension powers the Wikimedia Commons "Suggested tags" tool, whereby community members

⁵ <https://pro.europeana.eu/post/social-semantic-web-interview-with-europeanatech>

⁶ <https://pro.europeana.eu/post/crowdheritage-a-crowdsourcing-platform-for-enriching-europeana-metadata>

⁷ <https://www.mediawiki.org/wiki/Extension:MachineVision>

validate tags resulting from the application of Google Cloud Vision, cf Appendix B. In the Wikidata context there is a strong community of contributors interested in cultural heritage material.⁸

Typically, general campaigns have focused on specific objects ("road", "tree", "asphalt") or rather low-level descriptors ("red"). These are insufficient in more than one way. First, they focus on one type of concepts, mostly those that appear in the foreground, and obliterate that an object for example probably shows a town, that there are rolling hills in the background. Second, the suggestions are isolated labels rather than scene descriptions.

In some specific campaigns the focus was on named places and persons and concepts, especially when relying on annotations created based on automatic text or metadata analysis. But there still seems to be an opportunity for campaigns with higher-level enrichment based on image analysis such as the ones our project seeks to produce.

An alternative idea is associations, friends and relatives of people with disabilities, particularly visually impaired people, who may be able to help towards a better web accessibility experience when art pictures are present.⁹

Another option may be to launch an effort for annotating works that have to do with certain lifestyles, minorities, or with a gender slant, with the hope of using these to understand how these themes were treated in the past relative to the present, and thus improve the context of interpretation for these objects.

3.5 Upload in data sharing platforms

The project has a commitment to upload the results of the enrichment process to open data sharing platforms.

Of course a first scenario is to upload enrichments to Europeana, which has data sharing APIs. But the project is also committed to contribute data to the European Commission's European Data Portal¹⁰. Public administrations could use the rich heritage metadata to generate quality material for the websites, leaflets that promote their regions, improve the advertising and quality of cultural tourism and other cultural activities and fairs. This information can also serve as education material.

Additionally, we will attempt to contribute data to established research platforms like EUDAT (via B2FIND/B2SHARE) and EOSC. This scenario fits the idea of a research-focused use case, whereby SGoaB contributes a rich set of metadata of image descriptions that researchers may use for many types of statistical, symbolic, and linguistic analyses. This metadata may also be used by researchers in heritage and culture-related domains (including digital humanities) for more specific tasks of interpretation, such as mapping historical, social events reflected in cultural artifacts.

Finally, we believe there exists an opportunity to add enrichments to tools available for Wikimedia communities, so that these enrichments might be contributed to Wikidata or Wikimedia Commons. This could be combined with some of the crowd validation options envisioned for the use case in section 3.4.

As part of this use case, SGoaB will contribute the semantic enrichments produced by the project for Europeana objects, into these platforms. For non-Europeana services, we envision uploading

⁸ <https://blog.wikimedia.org/2016/08/23/wikidata-glam/>

⁹ <https://mw17.mwconf.org/proposal/project-describe-an-experiment-in-crowdsourcing-and-accessibility/>

¹⁰ As an alternative, we could contact national (and possibly local) portals to assess the interest of integrating our result on their sites. We expect this to be a much more difficult task than for the European Data Portal, though, given the diversity in the data standards they use, the access services they offer, and national (and local) legislation and publication methodologies. It is also a less good fit with the essentially European nature of the dataset we are going to produce.

the metadata enrichments produced by the project, as per the project's commitments. We may include Europeana's original metadata, as enrichments without the original data would make little sense to the users of these platforms, who will not necessarily be familiar with Europeana, or just be ready to accept fetching data from several places.

In all scenarios, we will also consider the data sharing platforms as potential sources of data to enrich. The SGoaB service will seek to generate enrichment for that subset of research data catalogues that refers to heritage data – arts, fashion, pictures of historical value or of archeological artifacts, etc.

Note that there is an impact from the data re-users on the platform: if the platform has a specific usage by say, researchers, this could lead to specific choices on the shape of the data to upload, the training to perform, etc. These aspects, which overlap with some of the requirements discussed in Section 3, will be studied in a next iteration of the use cases and requirement analysis.

3.6 White-label, custom trained enrichment service [TBC]

A user, for example a researcher or a data operator from a cultural heritage institution, performs her own training (or transfer learning) using her own datasets, to produce a "tuned" version of the service that will perform better for her own needs.

This can allow for fine-grained learning, and to lift some rights issues (as the tuned enrichment service could be confined within a user's context).

This would require more work in terms of setting the service and instances thereof, but would bring a clear value in terms of HPC. Users interested in applying SGoaB technology to visual data from another domain may indeed do so efficiently, since our service will be implemented using high-performance technologies (though not offering real-time enrichment).

3.7 Support enhanced accessibility of Europeana content for visually impaired persons. [TBC]

Metadata enrichments generated by the project could be used to improve the web accessibility of the associated images for people with disabilities. Concretely, web portals like Europeana could use the descriptions in the alternate tags of images to make their web pages accessible.

This would be of immediate interest to public administrations and associations. As a matter of fact the extent and success of this use case depends on concretizing some form of collaboration with the Illunion facility service, associated with the ONCE foundation for the blind. We will otherwise not implement this use case during the project, but rather offer the enhanced metadata such that other actors could use it for web accessibility in the future.

Explicit practices for image descriptions striving for accessibility exists among CHIs (Cooper Hewitt). Existing practices should be taken into account.

3.8 Rejected use cases

- A data partner can contribute their training data to the "general" enrichment service
- Similarity-based search for images, based on some fingerprints
- Enrichment by end-user: User uploads one cultural heritage object (either an image or an image with metadata) and gets a description of its content.
- Employ cultural heritage images and enrichments in captcha for crowdsourcing.

4. Requirements

Below we list the current "requirement dimensions" that can define the expected input and output for applying AI technology for the various cases. We expect each use case to focus on one or several of the possible options listed for these dimensions.

NB: we do not aim to do this selection of options now, as much depends on further analysis of the case and their still evolving context (stakeholders' needs, technical infrastructure, datasets available). We do not claim the current dimensions are exhaustive, either; they may be revised in the next document iteration.

Note that we are not reporting here on the data requirements on datasets for training models for enrichment, given that this is an issue related to the implementation rather than the (enriched) results. These datasets are subject to challenges that have to do with an insufficient number of paintings, as well as of the low availability of metadata or descriptions associated with images. However, if the project does not manage to source training sets that are extensive enough, then either additional techniques will be needed to obtain good results, or restrictions will have to be enforced on the scope of the use cases.

4.1 Media type of data input and semantic level of output

This dimension focuses on the kind of data that the use case is expected to handle, both for input and output.

Possible types of data input:

- image
- aligned image and metadata, when available

The output may consist of complex semantic representations (graphs), but simpler levels of representation should also be present in the final output:

The following table gives possible preference in terms levels of semantic output, per use case¹¹:

¹¹ The use case 3.6 for a custom trained enrichment service is not featured here. Users of such service would probably expect to focus the tuning on the level of output of their choice, making it hard to make a preference explicit at this stage.

Semantic Level	Use Case	Examples
Semantic resources (tags) <i>From vocabularies preferably with web URIs.</i>	3.1, 3.2, 3.3, 3.4, 3.5	<u>Adoration of the Magi</u> ¹² -> Jesus Christ, Virgin Mary, Wise Man. (as subjects coming from a vocabulary) http://iconclass.org/rkd/73B57/ (which actually corresponds to a concise description: "adoration of the kings: the Wise Men present their gifts to the Christ-child (gold, frankincense and myrrh)")
Textual tag	3.1, 3.2, 3.3, 3.5	<u>Arrivals on the Leviathan</u> -> "atlantic liner" or/and "boat". <u>The Leviathan in dry dock</u> -> "boat" <u>Neptuno</u> -> "man", "trident"
Textual captions	3.1, 3.2, 3.3 ¹³ , 3.5, 3.7	"Man reading a book in a dark room." "Woman plays a guitar outdoor during sunny weather." ¹⁴
Semantic/knowledge graph <i>As opposed to the level of semantic tags, here we have graphs with edges (standing for relationships, where the action/link also has a URI).</i>	3.1, 3.2, 3.3, 3.5, 3.7	<u>St. George, kill, dragon</u> "X, seated" <u>Arrivals on the Leviathan</u> -> graph representing that two women are standing on the deck of the boat/liner/Leviathan <u>Adoration of the Magi</u> -> graphs representing that a Wise Man is adoring Jesus Christ, that Virgin Mary is holding Jesus Christ.

4.2 Scope and data coverage/constraints

The project proposal notes that context, symbols, styles and traditions are relevant.

Our current guidelines for selecting relevant collections for enrichment are the following:

1. Type of object: Paintings, pictures of paintings.
2. Period: 14th to 18th century. Paintings before 1300 are few and not well conserved, while those after 1900 involve more modern concepts, lots of which are present nowadays, concrete rather than symbolic, and could be captured in a picture. Our focus for now is the

¹² As it is a glass slide, there is no description of the painting.

¹³ Although being intended for the specific requirements given by 3.7 these captions could potentially be indexed for search usage.

¹⁴ See also 4.7

pre-industrial era, where symbols were more prevalent and paintings were the only way to visually capture a real or imaginary scene.

3. Theme: Religious art, Mythology. These are the themes with the most number of paintings in the 14th-18th century, but also representing most of the symbols we are interested in detecting.

We could use the same criteria to guide the selection of relevant targets for enrichments.

These guidelines will have to be updated to reflect new requirements emerging from the analysis of use cases.

Types of objects

Types of objects could be extended beyond paintings, for example photographs, including photographs of other objects (sculptures, glass slides with art work, furniture, clothes, etc.). Appendix C gives a list of types of objects contained in the Europeana Entity Collection that we could potentially include (these are expected to be applicable to the objects categorized as IMAGE in the edm:type by data providers). In terms of topics, we could focus on concepts that are not time-specific (that could avoid detecting cell phones in old paintings!)

Amount of metadata available

We could consider to focus on well known objects (e.g. this [photograph of La Gioconda](#)¹⁵) as opposed to unknown art works or items for which we do not have (good) metadata (e.g. this [glass-plate slide of Kremlin](#)).

We will have to discuss whether requirements are the same for training and for applying the model once created. For example, once the model is created, can we use it for items with no metadata, just the image?

For example we have portraits with no subject or no description:
https://www.europeana.eu/portal/en/record/2048414/item_AN2SIQHWH72K2FN4AB4K3JMX4BB_IQTQ5.html

https://www.europeana.eu/en/item/92034/GVNRC_NESA02_B0258

4.3 "Iconographic level" requirements

Echoing the traditional categories of iconography¹⁶, there is a hierarchy of possible levels of objects present on an image, which we can target for enrichment:

1. Basic level classes ("person", "bird", etc.)
2. Basic actions with preposition-style connectors that can be viewed as positions (e.g: "riding a horse" denoting "on a horse")
3. Higher level concepts ("knight"), possibly derived from combinations of lower concepts (e.g: "Man" + "Spear" + "Horse" => "Knight")

¹⁵ For the object examples we provide URLs to web pages, but the metadata of any object can be obtained just by replacing "html" by "json" in the URL. For example https://www.europeana.eu/portal/en/record/2024903/photography_ProvidedCHO_KU_Leuven_998814993_0101488.html by https://www.europeana.eu/portal/en/record/2024903/photography_ProvidedCHO_KU_Leuven_998814993_0101488.json

¹⁶ Cf. the categories in the work Erwin Panofsky and pointers on the project space at https://3.basecamp.com/4181566/buckets/12787583/google_documents/2219897092

4. Named entities¹⁷, also possibly via combinations (e.g: "Man" + "Trident" => "Neptunus")
5. Higher level actions ("Virgin Mary holding Christ") or scenes that can also be viewed as positions (e.g: "Man" + "on" + "Woman's lap" => "Piéta")
6. Concepts, possibly reifying some of the above objects ("Piéta", "seascape", "crucifixion")
7. Imaginary places ("Hell") and persons. This category applies to some of the previous levels.

Note that the capacity of AI image analysis should not necessarily limit us here. Crowdsourcing could be interesting for identifying basic level classes and actions, but tentatively also for choosing from candidate (high-level) objects. Natural Language Processing applied on titles may provide the named entity and image subject, etc.

4.4 Data vocabulary requirements

This dimension refers to identifying which controlled vocabularies we could pick concepts and semantic tags from, and which ontologies we could use in the case of generating semantic graphs.

Options include Europeana Entity Collection and related (source) datasets: DBpedia, Geonames, Wikidata, as well as vocabularies used by Europeana providers¹⁸.

4.5 Rights requirements

Images in different datasets have very different rights. Out of Europeana, this applies to metadata as well. This opens the question of which proportion of these can be legally used. Conditions on original content used for training (e.g. CC-BY) may propagate on the results of the trained enrichment service.

We do not yet have an answer to this, but different use cases may come with different possibilities (and constraints) regarding the rights of the training and enriched data. Some cases, assuming wide publication of enrichment, could require fully open collections, both for training and publication. The type of output and storage of information could also matter. For example, producing and ingesting enrichments in Europeana as annotations (see Appendix A) may enable compliance with more re-use conditions than if enrichments are only represented as "regular" metadata, because annotations can keep the provenance and attribution of training data more easily than if we try express the provenance of training data in an object's enriched record. When an organization trains their own version of the enrichment service and applies it on their collections, such limitations would probably not have to be considered. Finally, we can imagine a process where we begin working with fully open training data and then we move to constrained datasets that are interesting for training or applying the service, asking their providers if they can lift some of the constraints.

This dimension is related to the work of task 8.3, which we intend to orient towards an assessment of the legal/copyright constraints that apply around the services SGoaB wants to develop.

4.6 Language requirements

We foresee our basic enrichment functionality will work with (meta)data in English, both for training and generation of enrichments. Several use cases are probably going to require translation of (i) training data, for better applicability of the enrichment service (ii) (meta)data being subject to the enrichment and/or (iii) produced enrichment, to allow pan-European usage.

First we will experiment with applying automatic translation (using the eTranslation DSI) in the training phase, especially as an attempt to enhance the impact of NLP in that phase.

¹⁷ Cf. https://api.europeana.eu/api/v2/search.json?rows=0&wskey=api2demo&query=title:*portrait*&profile=facets&facet=title at

¹⁸ <https://pro.europeana.eu/page/europeana-semantic-enrichment#enrich-your-own-metadata>

We will also use automatic translation DSI to generate metadata in other languages than English, when needed. These translations will however probably need to be verified if they have to be employed in a service (or uploaded in a repository) that expects a certain level of quality.

Finally, we also plan to try and translate user queries to English for the use cases where such queries need to be handled and matched against SGoaB enrichments.

4.7 Captions for accessibility [TBC]

The use case 3.7 addresses alternative texts for use by assistive technology. The purpose of these alternative texts is to provide a short visual description of the motive and the context of the image's content. An alternative text should if possible not duplicate information found in other structured text such as dc:description and dc:format.

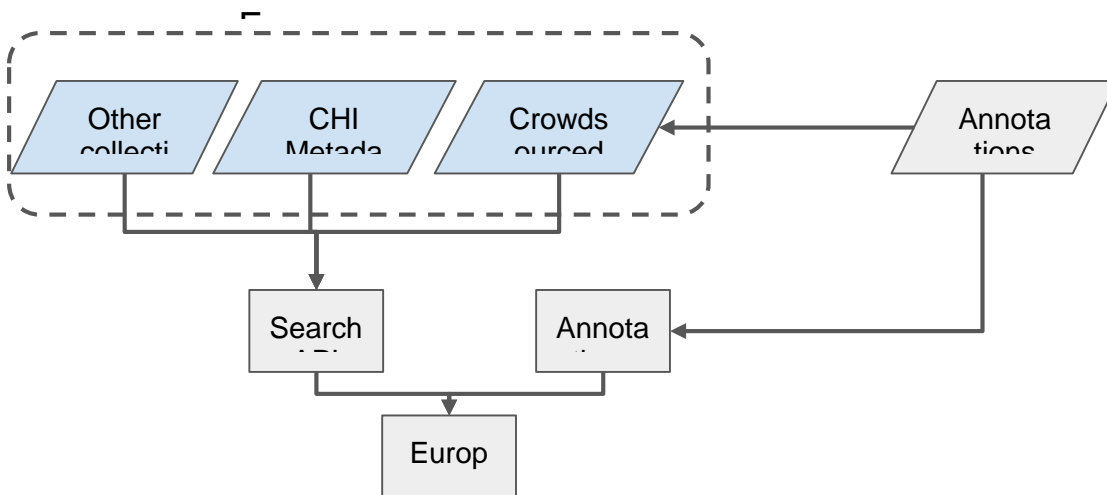
It's valuable to capture not only the main subject of the depiction but also to provide some additional context, "Man reading a book in a dark room." is therefore better than "Man reading a book.".

Appendix A. Technical options for supporting end-user semantic search

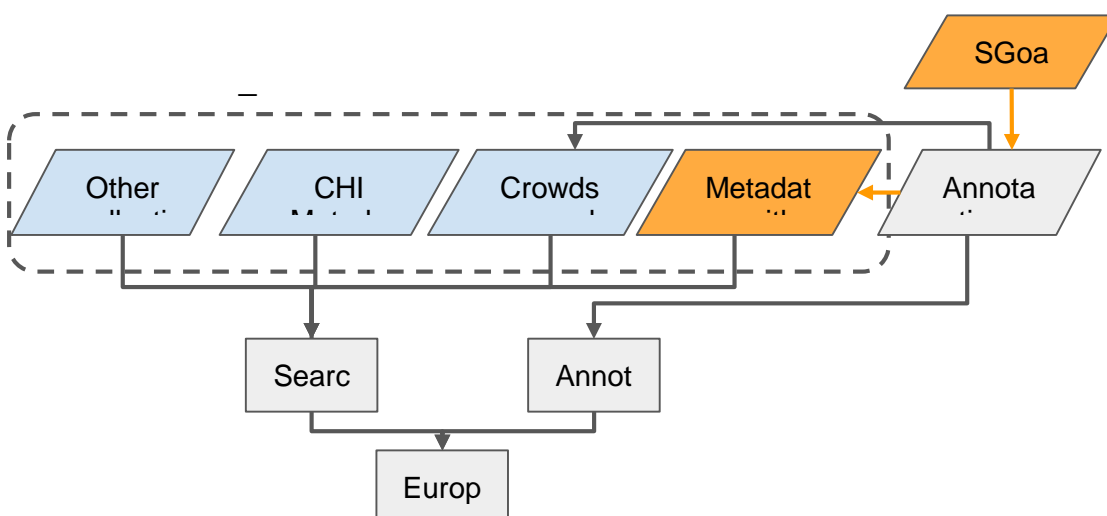
NB: this section has been incorporated in MS5, we may remove it from the next iteration of this document.

We will seek to develop the project's semantic search functionality on top of Europeana's APIs as much as possible. Europeana already hosts two search APIs. Europeana's "basic" Search API only accesses "regular" metadata provided by Europeana's institutional data providers and only a small portion of possible annotation data (i.e. transcriptions of textual documents), while the Annotation API accesses all other annotations, possibly including some more structured enrichments that could be produced by the project's service.

The following figure illustrates the current Europeana configuration for its metadata and annotation repositories and APIs.

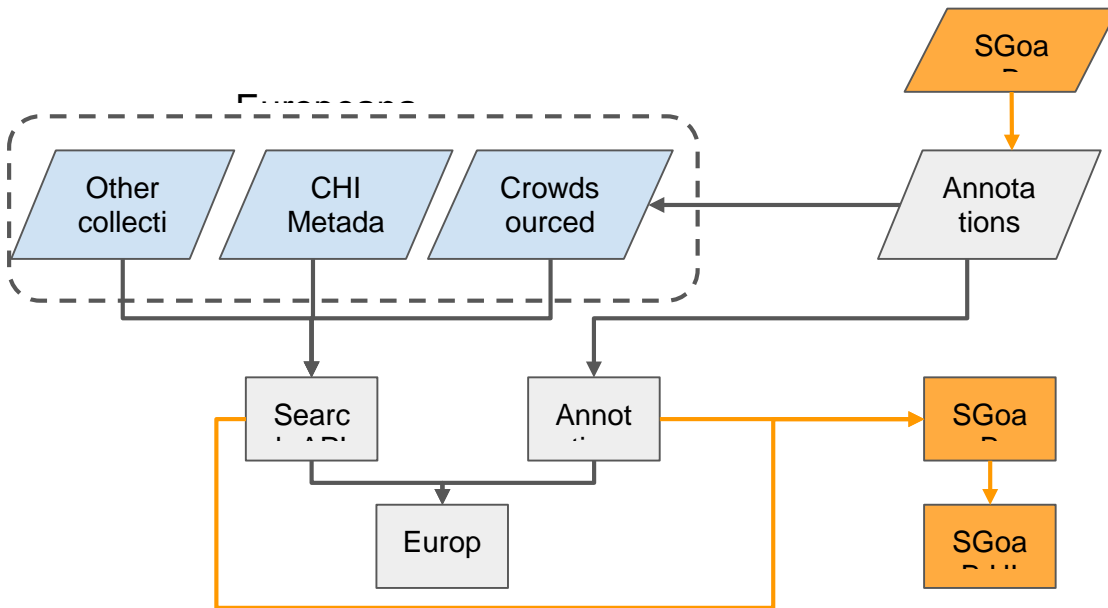


The following one illustrates how the project could build upon Europeana's main APIs and portal, after loading the Europeana annotations database with the results of the SGoaB enrichment tool:

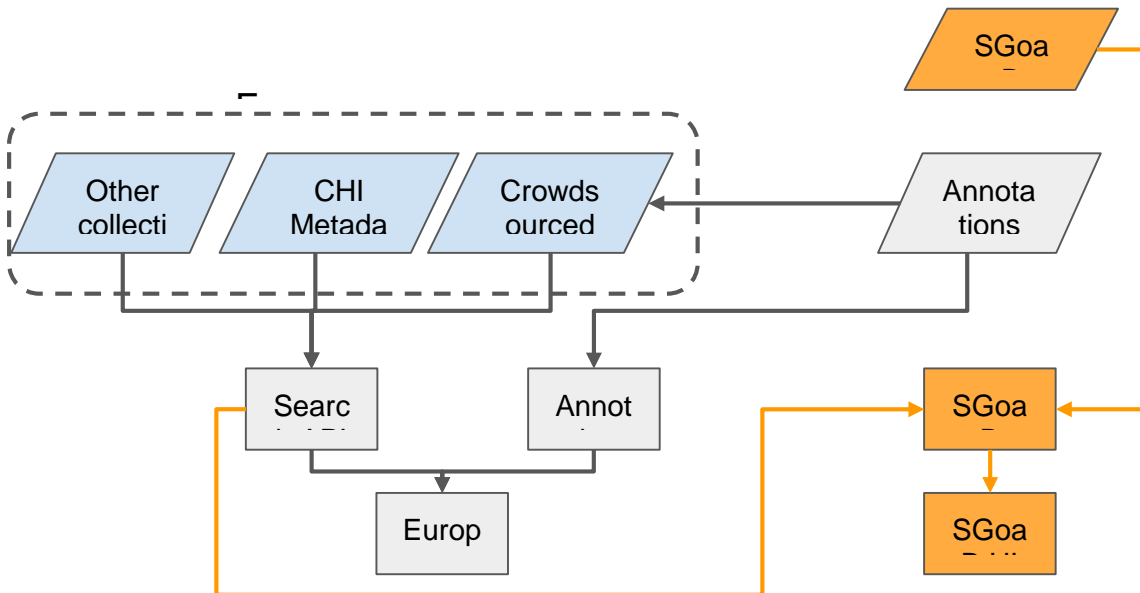


Using Europeana's APIs and portal would save the need of creating a project search service as a separate stack, but it would require them to be adapted to exploit annotations that represent our

project's enrichments (currently the Search APIs only exploits transcriptions, which do not include the enrichments our project would create). Should this prove impossible, the project would have to build its own Search API and Search UI, as shown below:



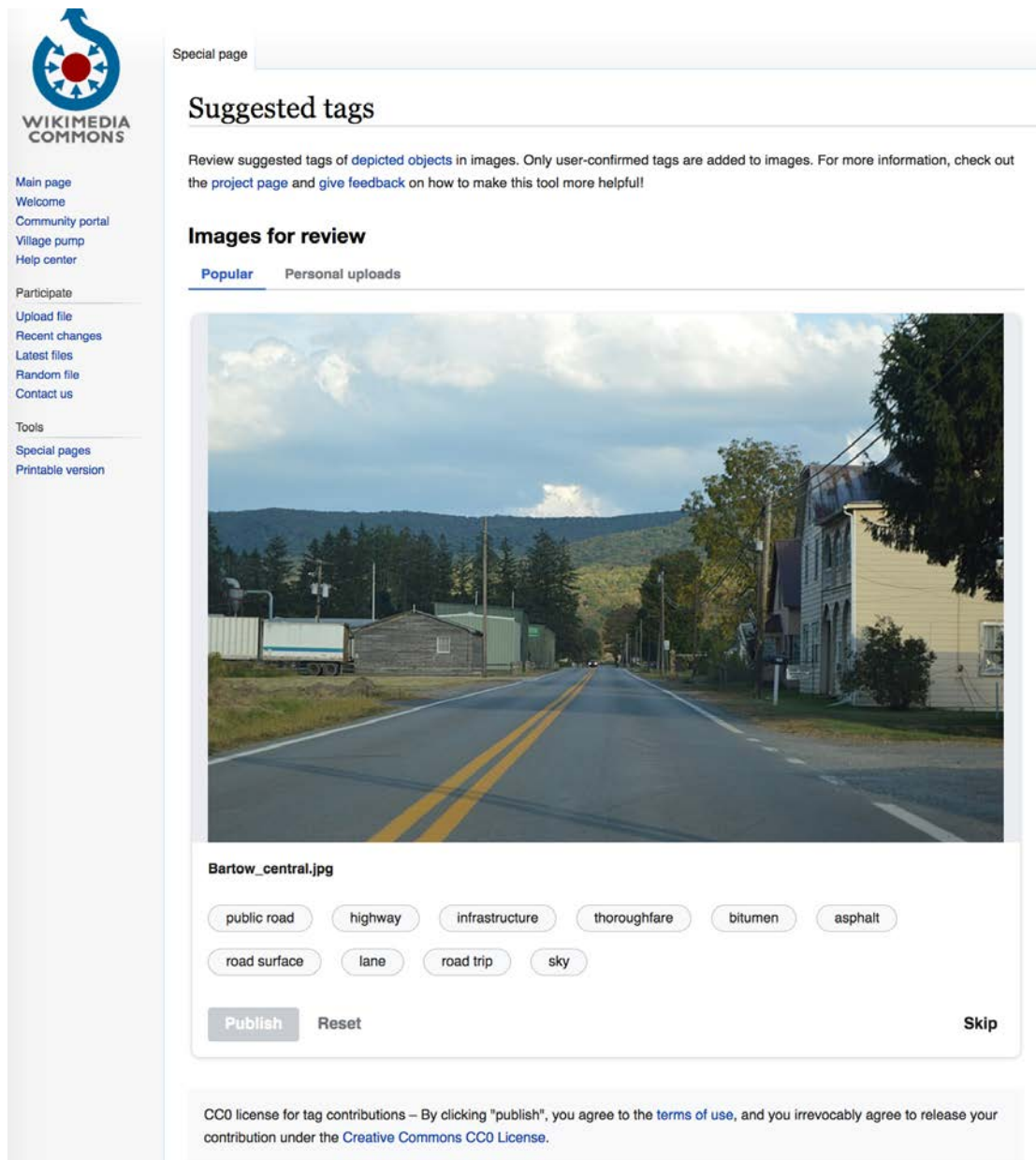
The next uncertainty concerns the use of Europeana's Annotation API. It may happen that enrichments of specific shape or quality cannot be loaded in the Annotations API, or not easily exploitable from there, in which case, the project-specific Search API would have to rely on its own enrichment storage and access layer:



Appendix B. An example of computer-aided tagging campaign on Wikimedia Commons

A description of the Wikimedia Structured Data computer-aided tagging is available at https://commons.wikimedia.org/wiki/Commons:Structured_data/Computer-aided_tagging

Tag validation can be done at <https://commons.wikimedia.org/wiki/Special:SuggestedTags> (requires login)



The screenshot shows the 'Suggested tags' interface on Wikimedia Commons. On the left is a navigation sidebar with the Wikimedia Commons logo and various links. The main content area is titled 'Suggested tags' and includes a 'Special page' tab. Below the title is a brief instruction: 'Review suggested tags of depicted objects in images. Only user-confirmed tags are added to images. For more information, check out the project page and give feedback on how to make this tool more helpful!'. The 'Images for review' section is active, showing a preview of the image 'Bartow_central.jpg'. Below the image, there are two tabs: 'Popular' (selected) and 'Personal uploads'. A list of suggested tags is displayed in rounded buttons: 'public road', 'highway', 'infrastructure', 'thoroughfare', 'bitumen', 'asphalt', 'road surface', 'lane', 'road trip', and 'sky'. At the bottom of the tag selection area are 'Publish', 'Reset', and 'Skip' buttons. A footer note states: 'CC0 license for tag contributions – By clicking "publish", you agree to the terms of use, and you irrevocably agree to release your contribution under the Creative Commons CC0 License.'

Appendix C. List of types of image objects

The table shows a list of types of objects which are expected to be applicable to the objects categorized as images by data providers (using IMAGE in the edm:type metadata field)¹⁹. These concepts are included in the Europeana Entity Collection, which is used for automatic enrichment²⁰.

Type of object	Corresponding DBpedia concept
Archaeological Site	http://dbpedia.org/page/Archaeological_site
Artefacts	http://dbpedia.org/page/Artifact_(archaeology)
Book Bindings	http://dbpedia.org/page/Bookbinding
Book Plates	http://dbpedia.org/page/Bookplate
Buildings	http://dbpedia.org/page/Building
Cartoons	http://dbpedia.org/page/Cartoon
Ceramics	http://dbpedia.org/resource/Ceramic
Collotype	http://dbpedia.org/page/Collotype
Costume	http://dbpedia.org/resource/Costume
Cultural Property	http://dbpedia.org/page/Cultural_property
Daguerreotype	http://dbpedia.org/resource/Daguerreotype
Digital Image	http://dbpedia.org/page/Digital_image
Drawings	http://dbpedia.org/resource/Drawing
Engravings	http://dbpedia.org/resource/Engraving
Equipment	http://dbpedia.org/page/Equipment
Furnishings	http://dbpedia.org/page/Furniture
Heraldry	http://dbpedia.org/page/Heraldry
Illuminations	http://dbpedia.org/resource/Illuminated_manuscript
Illustrations	http://dbpedia.org/page/Illustration
Jewellery	http://dbpedia.org/resource/Jewellery

¹⁹ The original list for all types of objects is at <https://docs.google.com/spreadsheets/d/1qjyyneq6aMoPC2v5hwC8YinmHKNYJtvTJp1HJdnnPc8/edit#gid=0>. Types in red were expected to be a less good match with the Europeana datasets.

²⁰ <https://pro.europeana.eu/page/europeana-semantic-enrichment#automatic-semantic-enrichment>

Lithography	http://dbpedia.org/resource/Lithography
Maps	http://dbpedia.org/resource/Map
Metalwork	http://dbpedia.org/page/Metalworking
Monuments	http://dbpedia.org/page/Monument
Mosaic	http://dbpedia.org/resource/Mosaic
Musical Instruments	http://dbpedia.org/resource/Musical instrument
Paintings	http://dbpedia.org/resource/Painting
Photographic Plate	http://dbpedia.org/page/Photographic_plate
Photographs	http://dbpedia.org/resource/Photograph
Posters	http://dbpedia.org/resource/Poster
Postcard	http://dbpedia.org/resource/Postcard
Prints	
Sculpture	http://dbpedia.org/resource/Sculpture
Sheet Music	http://dbpedia.org/resource/Sheet_music
Sketches	http://dbpedia.org/resource/Sketch_(drawing)
Specimen	http://dbpedia.org/page/Biological_specimen
Stained Glass	http://dbpedia.org/resource/Stained_glass
Studies	http://dbpedia.org/page/Study_(art)
Tapestry	http://dbpedia.org/resource/Tapestry
Technical Drawing	http://dbpedia.org/page/Technical_drawing
Textiles	http://dbpedia.org/page/Textile
Toy	http://dbpedia.org/resource/Toy
Watermarks	http://dbpedia.org/resource/Watermark
Woodcut	http://dbpedia.org/resource/Woodcut
Woodwork	http://dbpedia.org/resource/Woodworking