How can we facilitate the use of Europeana's Cultural Heritage Collections in History Education? Recommendations for Europeana

Research Framework on Improving Discoverability, Opening Up Historiana Project, part of Europeana Digital Service Infrastructure 4

Fani Partsafyllidou, EuroClio, European Association of History Educators, October 1st 2019, the Hague, the Netherlands

In this paper technical suggestions that will facilitate searching for historical sources in Europeana platform will be presented.

Contents

What problem do these recommendations solve?

In order to provide the [needs assessment](#) EuroClio conducted a survey and examined teachers' patterns when choosing an item as a source. The results of the survey showed that the more contextualized an item is, the more useful it is for history education, provided that the item is of historical interest in the first place. In fact, Items with adequate, comprehensible descriptions are 58% more likely to be included in a lesson.

Then, we took into account the difficulties and challenges in our project. Most objects do not have a description, which makes searching difficult.

Meanwhile, the prospect of having all European items (48 million, in May 2020) curated by professionals is not probable in the foreseeable future. Currently, stakeholders acknowledge the importance of curation, but find an obstacle in the perception that curation is not scalable, that millions of items cannot possibly be curated.

This is exactly the issue we want to contribute solving in EuroClio. In this research we will share a set of practical ideas for new functionalities in Europeana's search engine that will result in automated curation. These new functionalities use the data we already have on each item, then process them and combine them, creating more information. The questions we have to answer are: Is the item of historical interest? Is it curated?

Search Engine Optimization for History

Each sector approaches our cultural heritage in a different way; an artist needs to find items with high resolution, or with a specific colour -a history teacher needs to find items that are dated and curated. This means that search engine optimizations has to act differently according to the user's profession or to provide different filtering options specialized for each sector.

Filter 1. Historical dating precision

Before we get to the suggestion of a new functionality let's first see how the concept of time is saved in metadata. When we are dealing with cultural heritage, and thus the past, the perception of time changes. Items can belong to a decade, a century, or even more, depending on the time period. However, here's where things get ineffective. An item whose dating span is two centuries will be visible more times in a search by year, whereas an item with a precise dating will show up less times. Meanwhile, the item with the exact dating is more likely to have educational value because it is related to a more specific historical context.

Users can now search by year, which is admittedly more than enough for the vast majority of users. However, this can be improved in a fascinating way. If users can filter by historical dating precision, Europeana will become the most important platform for finding new sources in History, both in schools and in academia.

It is not only valuable, but also easy to add this filter. First, we define what historical dating precision is. Since the later something was created the easier it is to classify it in a time period, the concept of historical dating precision is relative: Specifically, there are also exceptions to this rule, with the most well-known being the Dark Ages, in which we have less historical sources than Classical Antiquity from which we have more historical records. In this case it is not necessary that items from the Dark Ages can be easier to date than the ancient ones. Although this contemplation is interesting, it goes beyond the limits of the current research. This is an exemplary taxonomy of historical dating precision.

Option A – Relative Taxonomy			
If the items dates to	The dating is considered exact if its span is less than	The dating is considered approximate if its span is less than	
40.000 BC – 20.001 BC	-	-	
20.000 BC – 1201 BC	-	1000 years	
1200 BC – 401 BC	50 years	100 years	

400 BC – 399 AD	20 years	50 years	
400 – 1199	50 years	100 years	
1200 – 1599	50 years	100 years	
1600 -- 1799	20 years	50 years	
1800 – 1939	5 years	20 years	
1940 – now	Day	5 years	

Option B – Absolute Taxonomy			
If the items dates to	The dating is considered exact if its span is less than	The dating is considered approximate if its span is less than	
anytime	25 years	150 years	

Technical requirements.

Date entries in Europeana's metadata should be streamlined into the same date formats. Subtracting the value of the earliest possible dating (terminus post quem) from the value of the latest possible dating (terminus ante quern) gives us the estimated range. This number should be compared to the second column of the table. If it is smaller, then the item is classified as having an exact date. If it is larger, then it is classified as approximate.
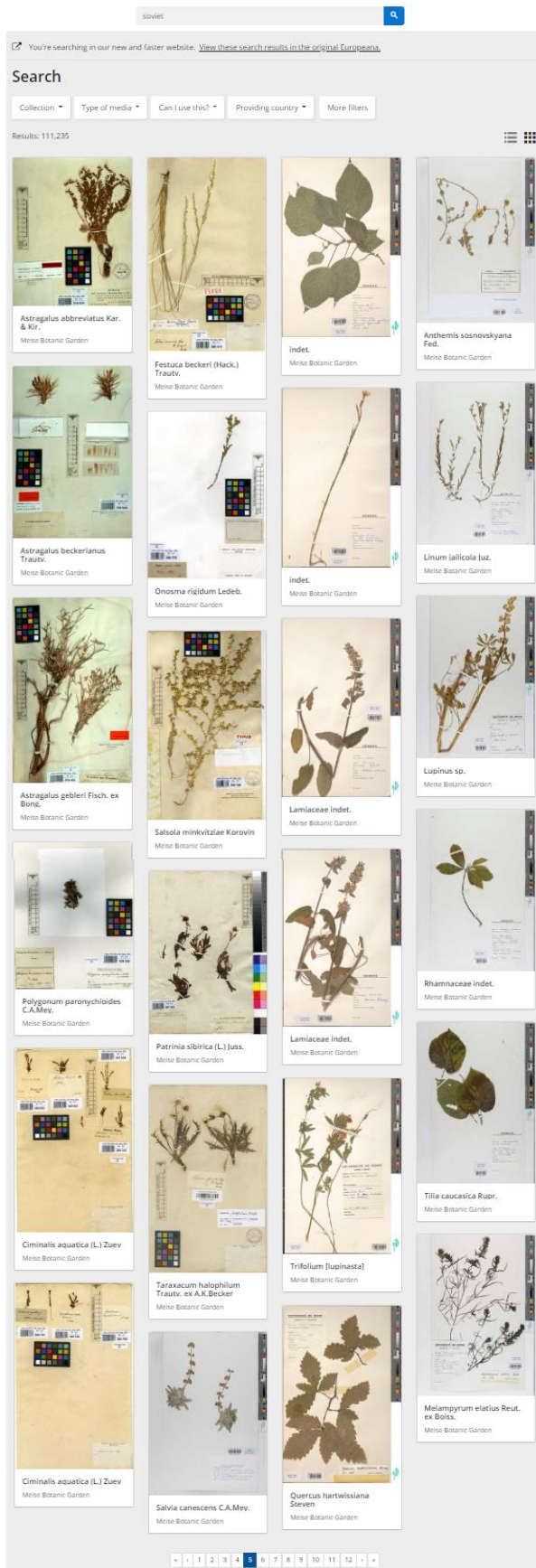
Filter 2. Face Recognition for Images

What problem does it solve?
How can it be applied in History Education?
Technical specifications
Problems that might arise
Examples before and after

Within minutes of browsing in Europeana's digital collections, one can understand the vastness of the term cultural heritage. Practically anything that has ever been created with a religious or aesthetic motive in the history of human kind is part of our cultural heritage. This becomes overwhelming for the average history teacher, who needs 2-3 sources for her lesson.

This is a filter that concerns Images in Europeana, which is currently the predominant type of media (26 million images, 20 million texts, May 2020). These are by far the two most popular media types, combined they amount for 95,83% of all Europeans items (46 out of 48 million). Searching texts is more straightforward, once they are transcribed --the technology behind this procedure is currently under construction. The words inside the text can be searched as well. We need a different approach when searching images.

In our survey, in which we asked history educators from Europe which sources they are most likely to use in a lesson, we found that teachers would prefer the pictures that included a human figure. Two out of ten images did not include a person and were not a map, and these images were 95% least likely to be used in a lesson.

Assuming three categories that encompass everything: People, manmade objects, and nature. History is the study of human societies and as such, we are interested in people and manmade objects. Face recognition can speed up the process of finding a historical source by omitting irrelevant results. In the example on Screenshot (i) we search 'Soviet' using no filters. Botanical results start in the third

Screenshot i Fifth page of search query 'soviet' is full of plant specimens.
https://www.europeana.eu/en/search?page=5&query=soviet&view=grid

page, and fully cover the fifth page. If we could use a filter expressing that we want results that contain people, face recognition would remove an impressive amount of the irrelevant results.

Same goes for manmade objects. Needs assessment shows that images with people are preferred in education. This means that out of two results, one with a vase, and one with a person holding a vase, the latter would be preferred. This is attributed to the pedagogical value, the teacher wants to show the relationship between the person and the vase, not the vase itself – unless it is an Art History lesson. Another reason is that images with people are simply more appealing, which is reflected in school textbooks too.

On the other hand, a problem that may arise when using face recognition for historical purposes is that portraits might be overrepresented in the results. History is the study of human societies, not facial characteristics. It is not of a historical significance how do people look, yet vanity is a timeless human flaw so hundreds of thousands of close-up portraits have remained as cultural heritage. These portraits will bypass the face recognition filter. An advanced implementation of the filter could include the option to search for items with 2+ people in them, which depict the relationships between people.

Option: Filtering altogether

The above filters can be used independently, or they can be merged into one filter, called 'Recommended for History Education'. It should be noted that an item which does not show up under the filter 'Recommended for History Education' is not necessarily irrelevant for History Education. On the contrary, most items can be used to teach history if someone devotes time to find more information and contemplate on its historical value.

The question we should be asking is: Can the historical value of this item be understood by someone who has no ties with its origin country? Since Europeana has reached dozens of millions of items, history teachers need to be able to explore the collections effectively. 'Recommended for History Education' means that the item is more likely to be clearly useful to History Education regarding its content and the pieces of information provided.

Taxonomy: Maps need to be tagged as text instead of image

Maps play a vital role in studying and teaching History. In this section we will explore how these items can be accessed in a more effective way.

Item i Map of Crimea https://www.europeana.eu/en/item/9200579/agmdy33a
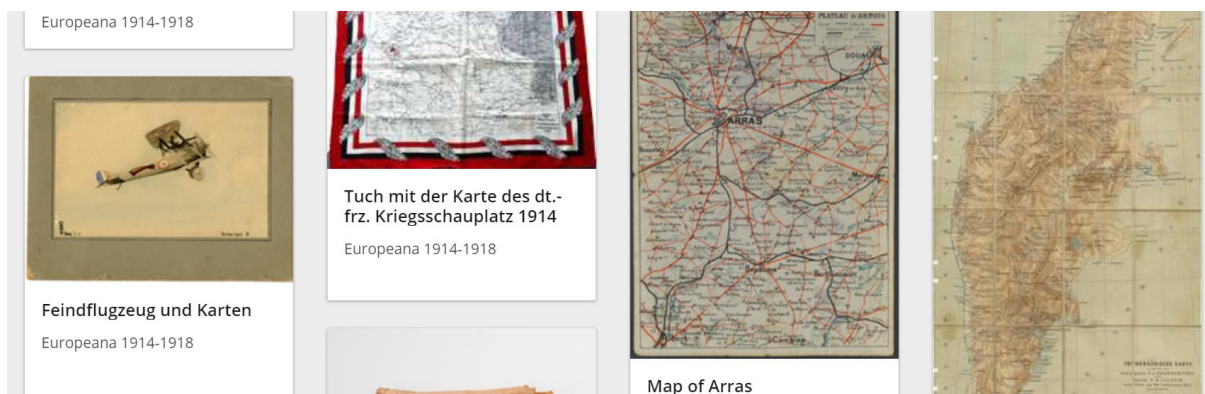
Currently, maps are characterised either as images or as texts interchangeably. I have failed to understand a distinction between the two given types. Please compare Screenshots I and ii below.


Screenshot ii First page results in the query 'map', free to use, Item type: Image

Screenshot iii First page results in search query 'map', free to use, Item type: Text

Providing a consistent way to present maps can speed up the user's experience. On a second note, maps are most commonly characterised as images. It would be helpful if photographs of maps are characterised as text instead of image. First, it would be more accurate. The core pieces of information in a map are the geographical names, which are written inside the map in the form of a text. The colours and the design are secondary elements that support the text.

Second, choosing text as source type for maps is the most useful and effective way to go. It will enable the text-processing services to be used on maps. Handwritten maps will be transcribed in the same process the manuscripts are, providing a list of geographical names for every map. All those names then become searchable. This can boost Europeana's traffic; given that two-thirds of traffic to Europeana Collections comes from Google[1], anyone who searches an unknown geographical name will be lead to Europeana, if there is a map mentioning this place and the map is processed as text. For example, the location 'Karasuhasar' which is seen in Item I above, returns zero results in Google using the query 'Karasuhasar map. All in all, if a map has an x number of geographical names inside it, if transcribed it will appear as a result in x more search queries than now. Currently, only the title is a keyword. In the aforementioned example, this map will show up only to the users who type Crimea.
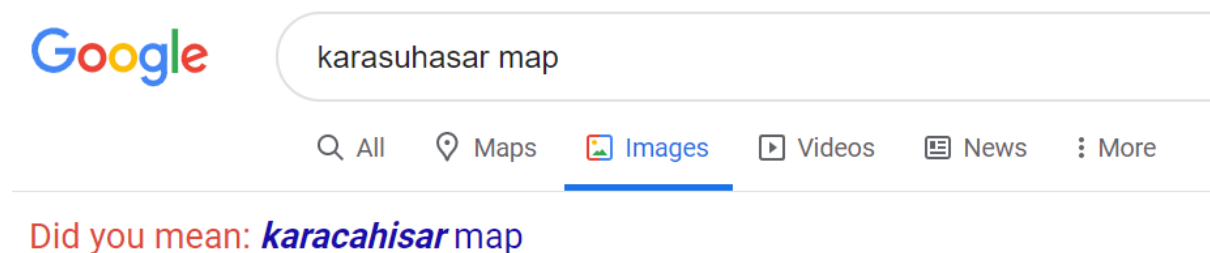
Screenshot iv Geographical name depicted in Europeana item i returns zero results in Google

Subsequently, processing maps as texts enables translation, which further improves the search.

[1] Beth Daley, Making it easier to search and browse collections online
https://pro.europeana.eu/post/making-it-easier-to-search-and-browse-collections-online