Deployment and Maintenance of Europeana DSI core services - SMART 2016/1019

CONTRACT NUMBER - 30 - CE - 0885387/00-80

DELIVERABLE

C.3 Data access pattern report 1 (M4)

Revision	1.0
Date of submission	22 January 2018
Author(s)	Dasha Moskalenko, Europeana Foundation, Antoine Isaac, Europeana Foundation, Tim Hill, Europeana Foundation, Henning Scholz, Europeana Foundation, Hugo Manguinhas, Europeana Foundation
Dissemination Level	Public

Co-financed by the European Union
Connecting Europe Facility

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
1.0	19-1-2018	Dasha Moskalenko, Antoine Isaac, Ash Marriott, Tim Hill, Henning Scholz, Hugo Manguinhas	Europeana Foundation	Final version
1.0	22-1-2018	Victor-Jan Vos, Jill Cousins, Julia Schellenberg	Europeana Foundation	Final review

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

The sole responsibility of this publication lies with the author. The European Union is not responsible for any use that may be made of the information contained therein.

'Europeana DSI is co-financed by the European Union's Connecting Europe Facility'.

TABLE OF CONTENTS

Executive summary 4

1. Introduction 5

2. Data access questions 6

3. Data access methods 6

4. Data access measurements 8

4.1. Quantitative measurements based on user searches and visited content 8

4.1.1. Most commonly queried topics on Europeana Collections 8

4.1.2. Type of content that is most often queried: images, video, audio, text, 3D 9

4.1.3. Most visited blogs, galleries and exhibitions 11

4.2. Measuring performance for Europeana search 13

4.3. Quantitative usage of the API 16

4.3.1. Content accessed via SPARQL endpoint 16

4.3.2. Content accessed via OAI-PMH harvesting 16

4.4. Link maintenance activity 17

5. Analysis of the data access measurements 18

5.1. Assumptions based on present understanding 18

6. Conclusion 19

Annex 1: Browse entry points for topics on Europeana Collections 20

Annex 2: List of data sets with broken links 21

Executive summary

This report describes the data we gather and the methods and tools we use with the aim to assist our users in finding the objects relevant to their needs. It will introduce the measurements that we collected so far and which measurements we aspire to collect in the near future to give us a more complete overview of how users access our data.

Comparing the total number of queries executed with the total number of query strings shows that query strings are quite diverse and generally not used often. This is also because users use terms in many different languages or spell terms differently. Content most accessed was content promoted by our browse entry points. Most viewed are recent editorials (blogs, galleries, and exhibitions) although older editorials remain popular.

Search patterns show that users exhibit a clear preference for images. Users also prefer to search for cultural heritage content with open licenses. Additionally, users show a preference for large and extra large imagery in their searches.

At the time of writing we are working on refining the framework for evaluating search and discovery functions. EF will also work on defining targets for improvement. This will be reported on in the next version of this deliverable (M8, April 2018). We also performed activities to enhance performance of the search and discovery algorithms (i.e. autosuggest for search accuracy) to help harmonize user queries and provide a more satisfying and reliable search experience.

1. Introduction

The purpose of this report is to show how users interact with content and metadata on Europeana Collections, thus their most common data access patterns. Further, to analyse content and metadata accessed by users with the aim to anticipate and serve our users the objects they need.

Main question explored in this report is: “How can we assist our users in finding the objects relevant to their need at a moment in time?”

This question is interesting to explore because users are becoming more demanding and less patient when interacting with online search engines. This report explores methods for anticipating and fulfilling our users needs, quickly and accurately, based on their current data access patterns. Methods used to collect data currently are highlighted and consideration is given to methods we have at our disposal but which are not yet fully operational.

The report concludes with an analysis of the data we have at hand including a number of assumptions about the most common data patterns. We also propose methods to validate these assumptions which will be reported on in the next C.3 report which is due in April 2018.

2. Data access questions

To answer the question: “How can we assist our users in finding the content relevant to their need at a moment in time?”, we decided to answer the following subquestions:

- What is the most queried content in Europeana? What are the most popular blogs, galleries and exhibitions?
- What type of media is most commonly requested?
- What content is most commonly downloaded?
- How many active API key users do we have and what type of content are they accessing?
- How many broken links do we have and how do we deal with them?
- Which datasets are most often requested to be harvested?

We chose to explore these questions to see if we can detect common data access patterns in our content. This will help us create relevant browse entry points, configure best bets to offer users suggestions on content they might be interested in, create top 10 lists for most accessed content, create themed galleries to help users explore related content, create more accurate autocomplete suggestions in the search bar, and rank the most commonly requested content higher in our search results. This will serve our users the content they need quickly and in general give us new insights to be able improve our overall browse user flow.

All the methods mentioned above will facilitate a better experience for our users since they help our users find the content they are looking for and explore new relevant content.

3. Data access methods

We use google analytics to see what content is most searched in Europeana Collections, and what blogs, galleries and exhibitions have the most clicks and shares. This gives us a good overview of the most popular content we hold. We can use such information to promote content and other related content on Europeana Collections, make it easier for our users to find.

Another interesting question we explored was the type of content most commonly queried, images, audio, video or text. This gives us a good picture regarding the type of content we should aim to source from data providers and the type of new collections that would be worth introducing.

Content that is most commonly downloaded is a good source of information on what type of records lead to conversion. Records with a high download ratio can be seen as the ones that users find most useful and have practical implementations for.

Another source of information for funneling out the most popular content on Europeana Collections is the use of our API keys[1]. We looked at how many active API users[2] we have i.e. users that were either active for more than five days or made more than five accesses in a month, and what content they are accessing. To measure the content that API users are accessing we rely on our logs.

Besides offering access for API users to our search engine in exactly the same way as is done for Europeana Collections (through the Search API[3]), Europeana also offers a SPARQL[4] endpoint as a more advanced way for researchers to search through Europeana content that facilitates a more precise interaction with our data. It enables one to search and correlate data based on specific metadata fields and data patterns[5]. Logging and analyzing SPARQL usage offers a granular view on how users, in this case researchers interact with our data and which topics they find most relevant for their research.

Not all users of the API interact with it in a synchronous way. Some require that data is collected beforehand and indexed on their side so that they can provide a better service to their users. For this purpose, Europeana offers a OAI-PMH service[6] which allows API users to selectively gather a large amount of metadata records based on datasets and when the metadata records have been last updated (or created). This also makes it possible for API users to keep data up-to-date with the data that Europeana holds. Logging and analysing which datasets are being accessed and the frequency of their update gives us a general idea of what content is seen to be more relevant and the level of demand.

In this report we also look at our activities to improve our search and discovery functionality. Specifically exploring the effectiveness and completeness of the search algorithm. In future reports this activity will be highlighted in report C.2 Usage pattern report as it is more related to how users use the features of the Collections portal than to which content they search on them.

---

[1] An API user may reflect one or more actual users given that an API key may be associated to a single platform/portal such as Europeana Collections (which has its own API key).
[2] The Europeana REST API receives an average of 45 million requests from about 200 different users every month. From these we consider as active users the ones that demonstrate either a relatively frequent or high use of the API as they reflect services or applications that contribute to the reuse of the Europeana content in several of its markets.
[3] Europeana Search API. View at https://pro.europeana.eu/resources/apis/search
[4] Europeana SPARQL API. View at https://pro.europeana.eu/resources/apis/sparql
[5] E.g. what are the French 18th-century painters with at least five artworks available through Europeana
[6] OAI-PMH service. View at https://pro.europeana.eu/resources/apis/oai-pmh-service

We also analyse how many broken links there are on Europeana Collections that may be preventing users from accessing the content they need and damaging their experience. This report also explains what activities we undertook to fix broken links and our plans to automate this time consuming activity.

4. Data access measurements

4.1. Quantitative measurements based on user searches and visited content

4.1.1. Most commonly queried topics on Europeana Collections

The table below describes the top 10 query strings and number of times they were queried on Europeana Collections, during the period of 1 September - 31 December 2017.

Topic query strings	**Theme**	**Nr. of times searched**
Women in World War 1	1914-1918	2289
Letters from World War 1	1914-1918	1762
Postcards from World War 1	1914-1918	1645
World War 1 videos	1914-1918	1410
The Eastern Front	1914-1918	1369
Man Ray	Photography	1216
Carte de visite	Photography	1408
Stereoskop	Photography	960
Karl Heinrich Lammel	Photography	954
Mappa Mundi	Maps	886

Europeana Collections had about 592,000 total queries, executed via the search box, in the period from 1 September – 31 December 2017. This is an average of about 5,000 queries per day.

Topics requested are highly diverse, with about 344,000 unique query strings on Europeana Collections, in the period from 1 September – 31 December 2017. Comparing

the total number of queries executed (c. 592,000) with the total number of query strings (c. 344,000) shows that query strings are quite diverse and generally not used often. This is also because users use terms in many different languages or spell terms differently. Compared to the high diversity of query strings the amount of times our top ten query strings were queried seems to be noticeable.

Note that all numbers here are extracted from Google Analytics. In future iterations of this deliverable we plan to complement them with numbers from our search logging framework, as it will (amongst other things) make it easier to distinguish 'canned' queries from blogs and social media, and queries where users actually type in their query strings. At the time of this report, we still need to consolidate these numbers.

4.1.2. Type of content that is most often queried: images, video, audio, text, 3D

Through analysis of search filters, we can understand the patterns of how users are using search to find the content they require. From September 2017 to December 2017, users employed the following search facets to filter by content type. These figures show the number of pageviews for searches employing these facets.

Type	**Nr. of times queried**
Image	358430
Text	62015
Video	28590
Audio	18697
3D	5332

Within an image search, users used the following filters to refine their search by file size.

Filter	**Nr. of times used**
Small	1146
Medium	2620
Large	7751
Extra Large	7581

Across all content types, users filtered their searches, as follows, to determine licenses of cultural heritage content.

License	Nr. of times used
Openly licensed	82362
Restricted license	8931
License only with permission	2835

These search/data access patterns show that users exhibit a clear preference for images. Users clearly prefer to search for cultural heritage content with open licenses, with more than 7 times of users deploying this filter than looking for material with restricted access. Additionally, users show a preference for being able to find large and extra large imagery in their searches.

These search filter patterns show that the Europeana Content Strategy to encourage Cultural Heritage Institutions to submit high-quality openly-licensed content is in line with users' expectations. As a result of this, we've created a new promotional tile for high quality openly licensed content, which should aid users in finding this material without the need for complex filtering.

Tile as browse entry point on Europeana Collections to openly licensed images. View at https://www.europeana.eu/portal/en

4.1.3. Most visited blogs, galleries and exhibitions

In the tables below you can find blogs, galleries and exhibitions visited most by our users in the period between 1 September 2017 - 31 December 2017. Most viewed are recent editorials, however older editorials remain popular.

Blog title	First published	Thematic Collection	Number of times accessed	% of total blog pageviews
Marco Polo: the man who brought China to Europe	January 2014	n/a	3660	5%
Winners of GIF IT UP 2017	November 2017	Europeana Collections	2654	3.6%
The Mauritshuis arrives in Europeana	September 2017	Art	2354	3.2%
GIF IT UP 2017	September 2017	Europeana Collections	1958	2.7%
Top 20 Searches on Europeana	December 2017	Europeana Collections	1636	2.2%
Francis Bacon, Shakespeare and Secret Societies	January 2013	n/a	1610	2.2%
The Harry Orvomaa collection of Jewish recordings	October 2017	Music	1187	1.6%
Did Columbus really see mermaids?	January 2012	n/a	1147	1.6%
Winners of Picture this competition	September 2017	n/a	1145	1.6%
Make a GIF from a vintage video	October 2017	Europeana Collections	1054	1.5%

NB: The pageview figures in this report represent when the user accesses that blog post in isolation. These figures do not account for viewing the contents of the blog through the overall blog feed, which accounts for 15% of the overall traffic to Europeana Blog.

Gallery title	**First published**	**Thematic Collection**	**Number of times accessed**
Calligraphy across cultures	October 2017	n/a	1634
Winter Wonderland	October 2017	n/a	1409
Treasures of the Mauritshuis	September 2017	Art	1192
Clothes for the Ballroom	May 2017	Fashion	1178
Artists' self-portraits	April 2017	Art	968

Exhibition title	**First published**	**Thematic Collection**	**Number of times accessed**
Power to the People	August 2017	Photography	8233
Art Nouveau: A Universal Style	March 2017	Art	8052
Leaving Europe	n/a	Migration	7739
Music and Mechanics	October 2017	Music	5684
An Ecstasy of Beauty	December 2017	Art	4492

4.2. Measuring performance for Europeana search[7]

Major efforts go into the further development of the search functionality. We have seen major improvements in Europeana DSI-1 and DSI-2, and under this project, Europeana continues to invest in improving the search functionality. Currently, there is an emphasis on how to measure the search functionality effectiveness and completeness, so that we can more objectively assess the results of our work.

Since 2014, Europeana has embarked on a long-haul effort of defining new, finer indicators for appraising search and discovery on Europeana (in Europeana DSI-2 there was only one KPI 'Percentage of all user searches matched to an entity from the Entity Database'. Previous search improvement reports developed during the Europeana DSI-1 and DSI-2 projects[8] guide this ongoing Europeana DSI-3 work[9]. The emphasis is on metrics immediately recoverable from our log data. That is to say, on data currently being harvested, rather than information that would require further software-engineering work to be done (other than processing of logs implemented earlier).

The metrics are also connected to specific search and discovery components, making it clearer where improvements can be made to enhance user experience on Europeana.

The current set of defined metrics of the search functionality is:

Component	**Evaluation Criteria**	**Description**	**Value for the period of September to December 2017**
Search box	nDCG	Normalised Discounted Cumulative Gain: a measure of how high in the result list clicked items are found. See https://en.wikipedia.org/wiki/Discounted_cumulative_gain#Normalized_DCG. Used as a reporting metric in previous DSIs' Search	0.562

[7] NB: in future iterations this section will be moved to C.2 Usage pattern report, as it relates rather to how users use the features of the Collections portal than to which content they search on them. A part based on the search logging framework will still appear in this report (C.3). However, in the report on the queries made by users (see section on content most commonly searched in Europeana).

[8] The most recent Search Improvement Report is DSI-2 D6.3 available at https://pro.europeana.eu/files/Europeana_Professional/Projects/Project_list/Europeana_DSI-2/Deliverables/d6.3-search-improvement-report.pdf

[9] We have gathered a Europeana Information Retrieval Component Inventory (https://docs.google.com/spreadsheets/d/1RAv1oZ3rVyfKC7bnwW29RYUWOsy-rnDZE6j0d2Yup5k) and are defining new KPIs with the University of Sheffield, (https://docs.google.com/document/d/16TKUfpZVM7m3SXjgfPD1_9Z2QvScxrJ8MIpdGHbCgb4/).

		Improvement Plan Progress Reports.	
	Percentage of queries with clicked results	The proportion of queries that are followed by a click on their results, as opposed to queries without clicks (which are then assumed not to have brought relevant results). Note that this figure may be strongly depressed by internal operations checking datasets, when many searches are launched but few results viewed.	3.4%
	Entity Collection coverage	Percentage of searches matchable by entities within the Entity Collection. Used as a reporting metric in previous DSIs' Search Improvement Plan Progress Reports (with a KPI of 30%)	34.2%
	nDCG upon completion of autosuggest term by user	nDCG as applied to the entity list supplied by the autocomplete. Available only for the testing phase of the Entity Collection API (by Europeana staff) so we expect it to be artificially higher than (coming) production results.	0.887
Filters	Frequency of use	The number of times filters of search results are activated by users	50.1%
	Frequency of item access after filter applied	The number of times the use of a filter is followed by a click on one of the items in the new (refined) search results.	31.4%
Similar Items	Frequency of use	The number of times users have clicked on one of the items suggested as part of the 'similar items' listed for an object being browsed.	0.15%
	nDCG	Normalised Discounted Cumulative Gain: a measure of how high the similar item clicked was in the list of all similar items suggested.	0.60

At the time of writing we are working on refining the framework for evaluating search and discovery functions, which means that the set of indicators is likely to be different in the next iterations of this deliverable. EF will also work on defining targets for improvement. It is the first time that many of these metrics are generated so it's not possible to conclude and interpret the metrics yet. This will be done once more data is collected. This will be reported on in the next version of this Deliverable (M8, April 2018).

The activities to enhance performance of the search and discovery algorithms are continuing along the lines mentioned in the previous search improvement reports (i.e. changing weights of fields used for the ranking, considering completeness of metadata records and presence of thumbnails, etc.). EF also redesigns some interface components that will impact some metrics (such as the design of the search result page).

One example is the contribution of the autosuggest for search accuracy. In September 2017, we introduced an auto suggest query strings functionality for our users when entering a search term in the search box to help harmonize user queries and provide a more satisfying and reliable search experience. In this first release, the auto suggest terms available are persons and concepts (e.g. Leonardo da Vinci[10], Art Nouveau[11]). The functionality appears on the main search bar on Europeana Collections, on Thematic Collections and on a smaller search bar on object pages.

Our expectation is that having a query completed by the autosuggest will help users arrive faster at the items that are relevant for them, and we can measure this using the nCDG. The assumption is that clicking an item higher up the list of results signifies that this list was more useful.

There are two stages of development of the autosuggest feature:

- Stage 1: Autosuggest by keywords - This offers the completion of search queries using terms available from the Entities API. - This option is already developed.
- Stage 2: Autosuggest using Entities URI - This completes the search query, but searches using the Entity URI which covers items dereferenced to an entity, as well as the equivalent searches in different languages. - This option is scheduled to be developed.

Once the second stage of development is completed, we'll be able to do a comparison of two metrics:

- nDCG of results obtained when search uses a keyword
- nDCG of results obtained when search uses an Entity URI

We expect that results obtained using an Entity URI will be more relevant and users will find content quicker through these result sets.

[10] Entity page on Leonardo da Vinci. View at https://www.europeana.eu/portal/en/explore/people/146741-leonardo-da-vinci.html
[11] Entity page on Art Nouveau. View at https://www.europeana.eu/portal/en/explore/topics/96-art-nouveau.html

4.3. Quantitative usage of the API

Users also access data via the Europeana REST API. Since January 2017, EF has collected monthly usage statistics for the Europeana REST API. From these statistics we derived two metrics to determine the number of active users: 1) number of API keys that exceeded the average of 5 calls a day; 2) number of API keys that were active for more than 5 days in each month. In the table below you find information of the months September - December 2017.

Period	**Number of API keys that…**	
	exceeded the average of 5 calls a day	**were active for more than 5 days in each month**
September	75	83
October	76	74
November	Not available (due to technical problems resulting from the migration to the new hosting provider)	
December		

4.3.1. Content accessed via SPARQL endpoint

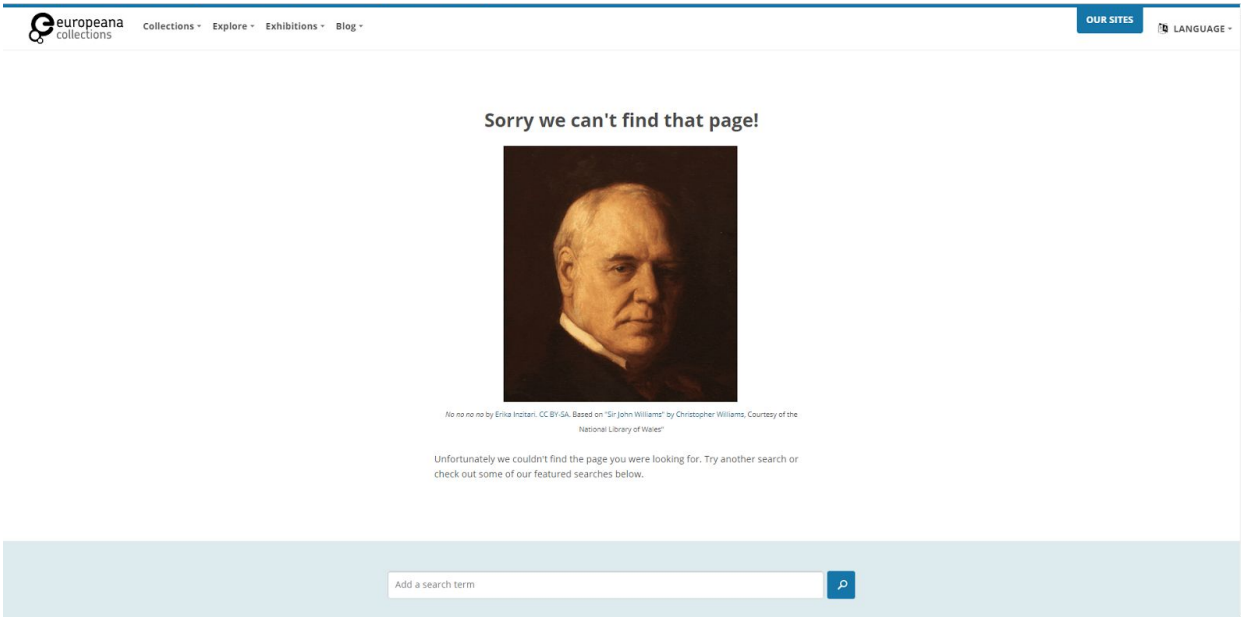
The SPARQL endpoint was developed as a pilot service and therefore does not yet track usage. As part of our efforts to bring this service into a full production state, we will start logging all user requests and collect the relevant content statistics for this report.

4.3.2. Content accessed via OAI-PMH harvesting

The OAI-PMH service is still in an alpha state and therefore does not yet track usage. This will be addressed once this service is moved into beta state by adding logging of all user requests and collect the relevant content statistics for this report.

4.4. Link maintenance activity

We worked on broken links as part of a bigger analysis of Europeana datasets to improve data according to the Europeana Publishing Framework[12] (EPF). We made an inventory of the tier compliance of records across the entire database in September 2017. This inventory also included the datasets that are not compliant to any of the EPF tiers. Reasons for not being compliant with the EPF include insufficient thumbnail size, lack of direct links to digital media and broken links.



Page a user views when he clicks on a broken link.View at https://www.europeana.eu/portal/en/record/08559/en_root551__show_image__id_353.html

EF prioritised the list of datasets that are not compliant to the EPF based on size of the dataset and reason for being not EPF compliant. About sixty datasets (about 1.2 million records) with broken links have been identified to date. The majority of these datasets have been de-published since September 2017 (after consultation with data partners), including contributions from the former Europhoto and Assets projects. For the remaining ten datasets with broken links, data partners have been approached to discuss how to improve

[12] The EPF clarifies the relationship between the quality and openness of data and the impact it can have. Implementing this framework means that EF and all its data partners work towards high content quality so it is compliant with the higher tiers of the framework. In the EPF four tiers of participation were developed, to allow data partners to decide what they want to achieve with their digital collections, based on their own agendas and capabilities. The higher the quality of data provided by data partners, the more benefits they will create for their audiences. View at https://pro.europeana.eu/post/publishing-framework

the datasets or otherwise depublish them within the next three months. Following this approach, all datasets larger than 5,000 records will be free of broken links by April 2018. A list of main datasets with broken links can be found in Annex 2.

Process for detecting and fixing broken links
Datasets with broken links are mainly detected through the gallery[13] broken link detection. When a displayed item in a gallery encountered an error in retrievability, the data partner services team gets notified and can take appropriate action. The broken links can then be rectified by either contacting the data provider or removal of the datasets should the source no longer exist.

5. Analysis of the data access measurements

This work has highlighted the need for more data to analyse. We will collect this over the course of Europeana DSI-3 and report on new findings to be able to draw more solid conclusions on the most common data access patterns. We can see some trends but need to collect more data using the methods described above to validate our current assumptions. See below our assumptions based on the data we have collected and our plans to validate these for the next C.3 report which is due in April 2018.

5.1. Assumptions based on present understanding

Our first assumption is based on the most popular search terms queried. Seeing that historical content specifically related to the First World War received a lot of traction on our website in the past few months we might assume that this is either related to the centenary of the First World War or our users are First World War enthusiasts. To validate such assumptions we need to know if the queries are coming from unique users. If that is the case then we need to validate if there was any marketing or promotional activity around this content. This will confirm if this is a temporary or a long lasting pattern. Alternatively, we survey our users using a tool called Hotjar asking which content they want to see most on Europeana.

Our second assumption is regarding the popularity of the First World War Collection and the Photography Collection. These are searched more often than any of the other Collections. We assume that this is because they are newcomers and are both highly visual photographic Collections. This content is easier to consume because it is specific and clear in what it represents and gives a feeling of nostalgia. To validate this we will survey the users visiting the Europeana 1914-1918 Collection and Europeana Photography using a tool called Hotjar on their reasons for visitation.

[13] Galleries are one of our tools to highlight content by displaying records bundled by theme. View at https://www.europeana.eu/portal/en/explore/galleries

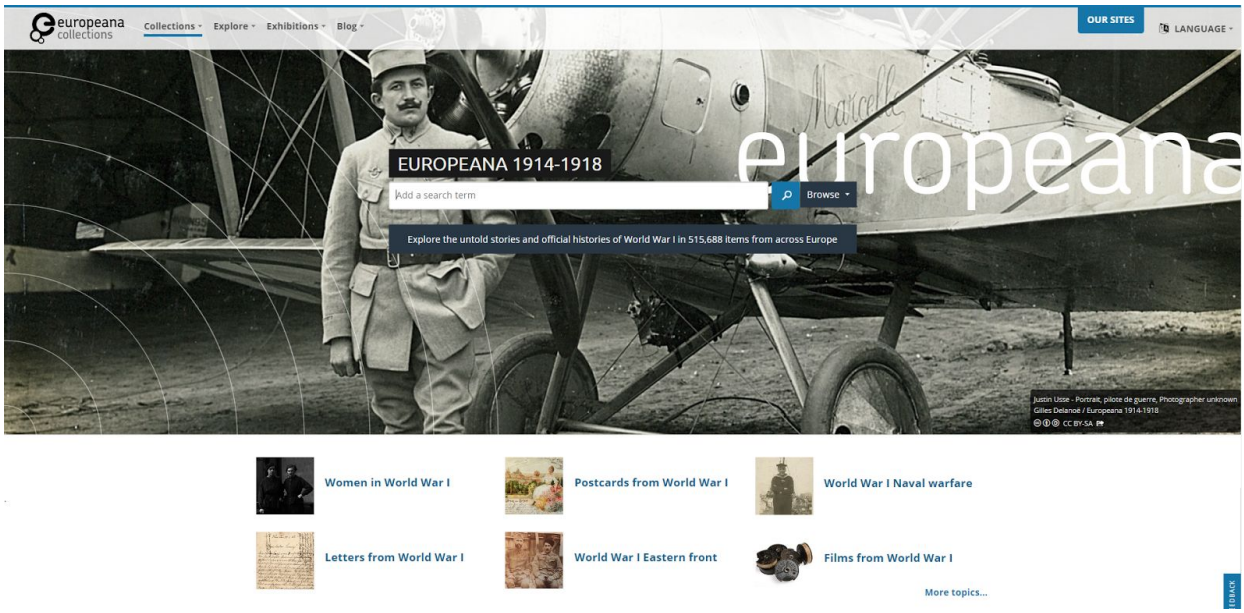
Our third assumption is that returning users prefer browse entry points over search. The reasons being that 28% of the users are not using the search query and 60-70% go directly to an item page leaving only 10% of users to use the search engine. To validate this we will either count the unique users that use our search compared to our overall usage. Or we will create a minimal viable product (MVP) with a hollitics browse page so a niche page that focuses on browse and see the visitation on that page versus the unique search queries.

6. Conclusion

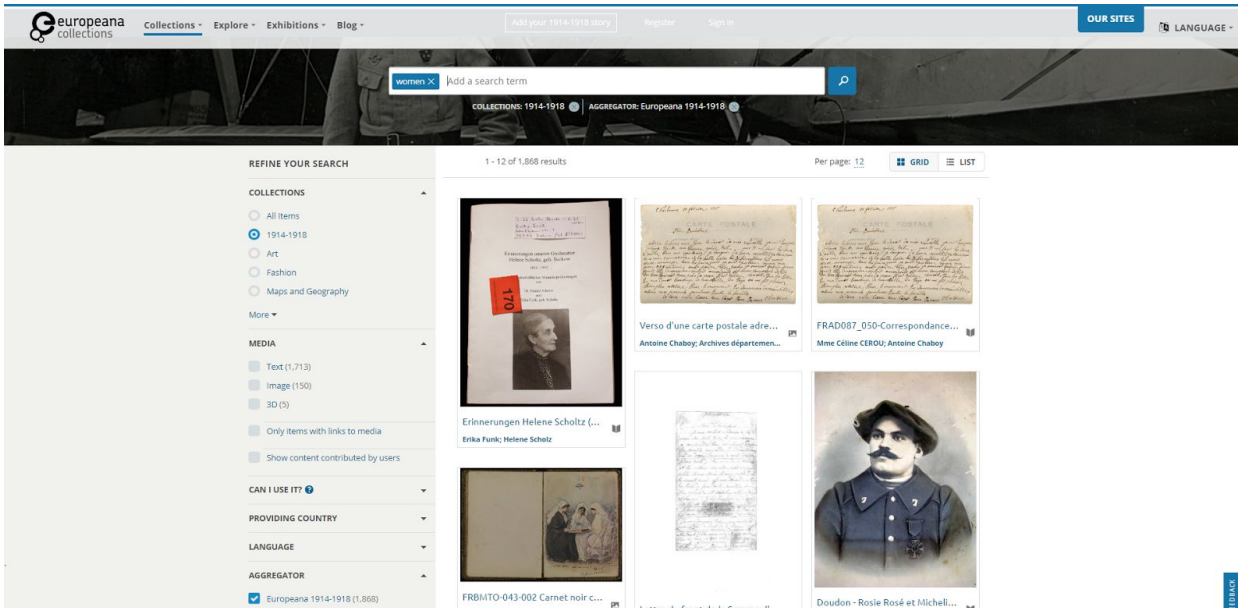
We need to collect more data to draw usable conclusions but expect that the results of the ensuing analysis could have a significant impact on our users' experience. To make this data truly usable we need to mine further information about our users by presenting them with alternative designs to measure engagement and using surveys to validate the assumptions we have so far drawn from the data itself.

We will continue our activities to improve the user's search experience by improving autosuggest, and will expand the coverage of the Entity Collection and improve the search algorithm.

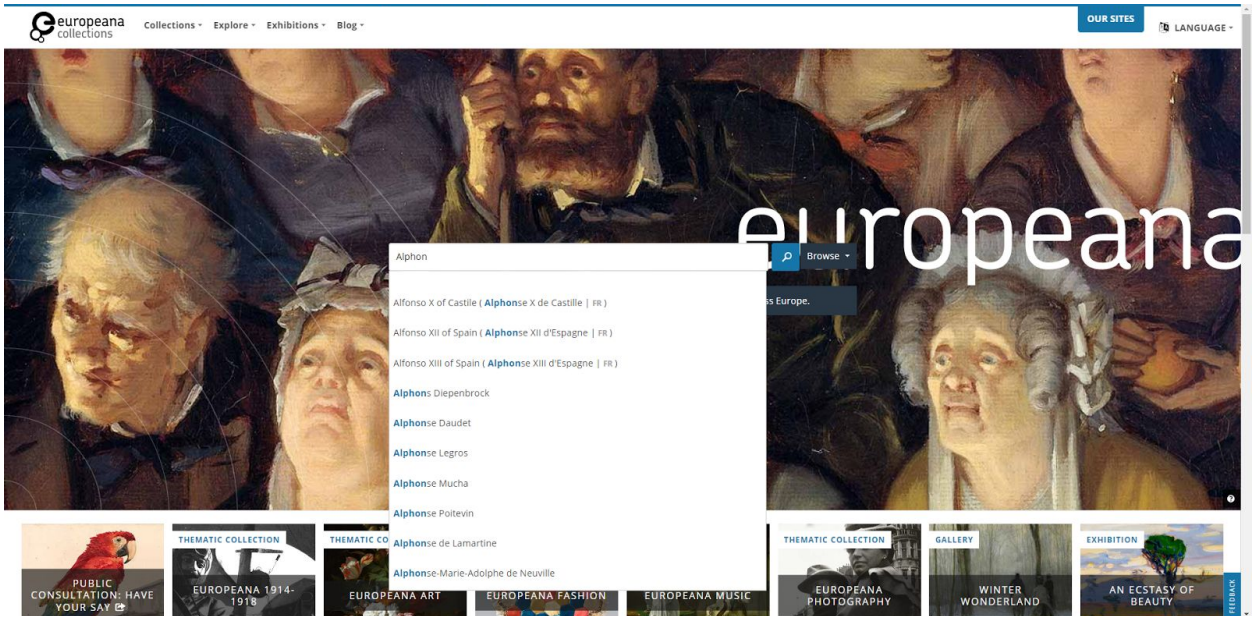
Annex 1: Browse entry points for topics on Europeana Collections



Browse entry points on topics on Europeana Collections (at the bottom of the screen). View at https://www.europeana.eu/portal/en/collections/world-war-I



Browse entry point Women in World War 1. View at https://www.europeana.eu/portal/en/collections/world-war-I?f%5BPROVIDER%5D%5B%5D=Europeana+1914-1918&q=women&view=grid



Pop-up menu with suggested terms when starting to type the term 'Alphon' in the search box. Search box can be viewed at www.europeana.eu.

Annex 2: List of data sets with broken links

Data provider	Country	No of datasets	No of records
EuroPhoto ANP provider	Netherlands	1	155448
EuroPhoto DPA provider	German	1	132175
EuroPhoto MTI provider	Hungaria	1	120870
EuroPhoto PAP provider	Poland	1	119548
EuroPhoto EFE provider	Spain	1	118855
EuroPhoto ANSA	Italy	1	61192
EuroPhoto BELGA provider	Belgium	1	37616
EuroPhoto LUSA provider	Portugal	1	31162

EuroPhoto EPA provider	Germany	1	20356
Fundacion Albeniz	Spain	1	15684
Accademia Nazionale di Santa Cecilia	Italy	3	26051
Centre Virtuel de la Connaissance sur líEurope	Luxembourg	16	23063
Hellenic Archive of Scientific Culture (HASC)	Greece	1	500
DW-WORLD.DE | Deutsche Welle	Germany	8	349
Slovenská národná knižnica - Slovak National Library in Martin	Slovakia	1	5548
Heritage Malta	Malta	1	7218
Stedelijke Musea Mechelen	Belgium	1	7412
National Library of Wales	UK	1	9099
Bowes-OAI	UK	1	15314
ERT SA	Greece	1	21001
Rijckheyt, centrum voor regionale geschiedenis	The Netherlands	1	33886
DK-National Aggregation Service	Denmark	1	83432
Digitale Collectie	The Netherlands	1	85049
Athena	Italy	1	92311
Slovenské národné múzeum	Slovakia	2	24000