europeana

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

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

DELIVERABLE

C.2 Usage pattern report 1 (M4)

Revision	1.0
Date of submission	15 January 2018
Author(s)	Dasha Moskalenko, Europeana Foundation Ash Marriott, 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	8-1-2018	Dasha Moskalenko, Ash Marriott	Europeana Foundation	Final version
1.0	15-1-2018	Jill Cousins, Victor-Jan Vos, 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This deliverable contains original unpublished work except where clearly indicated otherwise.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.

TABLE OF CONTENTS

Executive summary

Research and analysis of usage patterns led to two areas of work to improve access to Europeana:

a. Finding ways to encourage General Public users into the site (thus reducing the bounce rate on item pages), and improving Google indexing of pages;
b. Analysis of professional users led to the redesigning of Europeana Pro.

On Europeana Collections, to reduce the bounce rate on item pages, Europeana has redesigned the item page. A redesigned item page will offer more possibilities to further explore Europeana content. This entices users to further engage with the website. Additionally, Europeana works on enhancing the number of indexed pages by Google. Google is the main entry point of users to Europeana. To achieve this, Entity pages have been introduced. These pages link information about concepts or persons to Europeana content and we hope Entity pages will improve the crawl rate of Europeana Collections by Google. However, we need more data to assess if this is really the case. We are therefore investigating how best to test this hypothesis.

Following up on user research for the Europeana Pro website that existed before September 2017, Europeana introduced a redesigned Europeana Pro website in September 2017. We are also investigating the findability of content on the new website and how satisfied users are with the new design.

Introduction

This document describes how Europeana determines where and how users experience difficulties with the use of the Platform and the actions we take to overcome such difficulties with systematic improvements over time. This report refers to the user-facing parts of the Platform, Europeana Pro (for all professional markets) and Europeana Collections and the thematic collections (for end-users).

For both products, Europeana has defined research goals in the reporting period M1 - M4, as derived from a key performance indicator of Europeana DSI-3, or an already known user issue, as indicated by user research or Commission Services. In general, development of a new feature or a design change is based on either or both of these.

Europeana's most common methodologies to collect user feedback are: surveys, heatmaps, nps scores, and feedback buttons. A more detailed overview of the user recommendations Europeana receives via the feedback button and how these are incorporated into our development cycle is explored in deliverable C.5, which will be delivered at the end of month 5 of Europeana DSI-3 (January 2018). Yearly, Europeana also commissions external agencies to perform usability and accessibility tests.

Europeana will further develop its methodologies on how to best to evaluate where users are experiencing difficulties with the user-facing products and thus be able to improve the service.

This deliverable should be read in combination with Deliverable C.3 Data Access Pattern Report, which describes how users access and engage with Europeana data and the forthcoming Deliverable C.5, which describes how users have provided feedback to Europeana and how we have addressed this feedback.

Europeana Collections usage patterns

Europeana Collections is Europeana's main service for end-users to access Europeana data. As a product, Europeana Collections provides access to the full Europeana dataset, as well as the thematic collections on the themes that the Tender has identified (e.g. Art, 1914-1918, Fashion, Music). That means that any improvement or design change will also be seen across all thematic collections as well as on Europeana Collections itself.

The aim of the first four months of development of Europeana Collections under Europeana DSI-3 was to improve the traction of Europeana Collections for first time visitors and hence reduce the bounce rate of our item pages. Europeana faced a bounce rate of about 60% (Google Analytics) in mid-2017. Therefore, we needed to find motivations for users to explore Europeana content further.

Europeana also made efforts to increase the indexing of Europeana pages by Google, so that we could get visitors coming in the first place. Visits (measured using Google Analytics, please see screenshot below) to the website coming from Google decreased throughout 2017.

The solution that has been put in place to both improve on site user experience and indexing by Google was Entity pages. These link high quality content with concepts or persons. Such pages have a higher value for indexing and therefore rank higher in Google searches. We also changed our strategy and only submitted records that have high-quality metadata to Google for indexing. Previously, we submitted all Europeana pages for indexing (over 53 million pages) to Google. As a result, this should give Europeana pages, which we assume are estimated by Google based on the quality of the total amount submitted, higher value and consequently rank them higher in Google searches. This gives users better possibilities to find Europeana content on Google. This strategy appears to be working as the biggest increase in indexed pages occurred following the release of the Entity Pages in September 2017. Entity pages are now searchable in Google and their contents are indexed, unlike our search pages. We made it easier for Google's crawlers to access our highest quality pages, which we think, is why the number of indexed pages increased.

Traffic on Europeana end-user products

Europeana Collections, Europeana 1914-1918, Europeana end-user blog, Transcribathon
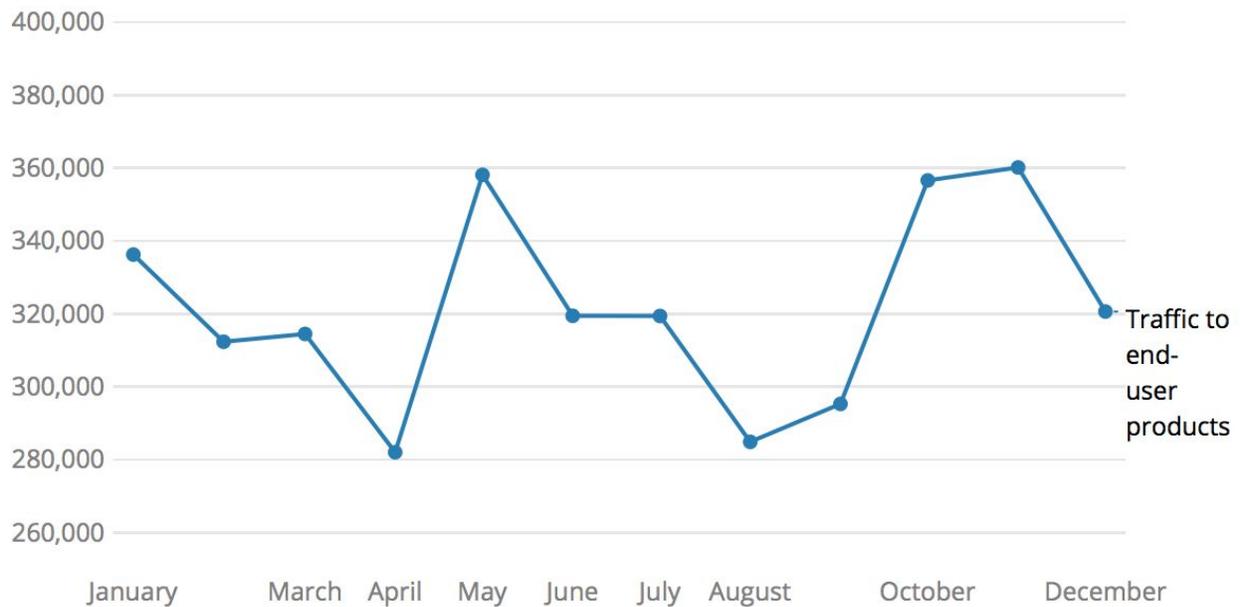
Traffic on Europeana to end-user products (including Europeana Collections, exhibitions, blogs). Note: Due to technical problems resulting from the migration to a new hosting provider, Google Analytics data from November/December is incomplete.

Entity page of Alphonse Mucha. View at
https://www.europeana.eu/portal/en/explore/people/61016-alphonse-mucha.html

Reducing Europeana's bounce rate on Europeana Collections

The bounce rate is of particular importance as the majority of Europeana's incoming traffic arrives via Google. Europeana is looking to reduce the bounce rate through redevelopment of the item page and cross-promotion of other content available on Europeana Collections.

Hypothesis
With the introduction of the new item page, the bounce rate should significantly be reduced as the redesigned item page will hold better promotion of content to the user.

Bounce rate
Due to a failure in capturing analytics on the item page, data from 27 October until 4 December is missing. The following data is therefore split into two ranges. In average the bounce rate was 60% .

Average bounce rate - September 1 to October 26	60.64%
Average bounce rate - December 5 to December 31	62.39%

Conclusion
The release of the new item page in the first quarter of 2018 is expected to contribute to an improvement in these figures, but until that point no conclusion can be drawn on whether the improvements have had the desired effects. Below you can find an example of the new record page design. The following changes are anticipated to improve the bounce rate:

- Automatic promotion of related galleries that may interest a user (and in a future iteration, related exhibitions and blogs) through 'cards' on the right hand column. There is currently no promotion of other Europeana content on the item page giving a user, arriving from organic search, a reason to continue to search on Europeana Collections. The promotion of related content aims to help users to find similar content. Entity pages additionally improve our crawlability for Search Engine Optimization (SEO) because of the related items listed on these pages for crawl.
- Customisable promotional cards will allow Europeana to place content of its choosing on the right hand side of the item page based on content type. This will allow promotion of other areas of Europeana that were not previous possible (i.e. campaigns).

We will measure the effectiveness of these changes and report on them in upcoming reports.

Example of new record page design

Improving Google indexing of Europeana Collections

With the high number of records available on Europeana Collections, it is difficult to get all indexed by Google. Google appraises the quality of the total amount of pages submitted for indexing. If most of the pages have poor metadata (that can be indexed to facilitate retrieval) the whole dataset is ranked poorly. Europeana undertook two activities which aimed to improve the number of indexed pages in Google:

- Introduction of Entity pages which connect records of a similar topic (i.e. Johannes Vermeer[1]);
- All sitemap records submitted have a higher amount of metadata populated than other records.

---

[1] Entity page on Johannes Vermeer. View at
https://www.europeana.eu/portal/en/explore/people/60404-johannes-vermeer.html

Hypothesis
Entity pages: the introduction of Entity pages will provide a crawlable link between content of specific topics and people. The linking should improve the crawl rate[2] of Europeana Collections.

Completeness rank on sitemaps: submitting only pages with completed metadata fields will improve the searchability of Europeana content. More pages will be indexed due to the perception of higher quality to Google, as well as exposing more keywords to a users searching Google.

Conclusion
The biggest increase in indexed pages occurred following the release of entities in September 2017. Entities pages are now searchable in Google and their contents is indexed, unlike our search pages.

Overall indexing
- In September 2017 (pre-experiment/entities) Google had c. 7.4 million pages indexed.
- As of January 2018 Google has indexed c. 15 million pages.

Sitemap indexing based on completeness rankings
- In August 2017 (pre-experiment/entities pages) we had submitted c. 53.6 million pages to Google, of which c. 3 million were indexed.
- In January 2018 we had submitted c. 10.5 million pages to Google, of which c. 3.7 million were indexed.

Despite removing the lower quality pages from the sitemap thus cutting the number of indexed pages by around 80%, we have had a 20% increase in the number of indexed pages. We made it easier for Google's crawlers to access our highest quality pages, which is why the number of indexed pages increased.

The dramatic reversal in indexing since September points to this strategy working. However, we would like to confirm whether the introduction of entity pages or the condensed sitemap caused the increase in indexed pages, which we will assess by reintroducing the full sitemap. Ideally we would have the whole sitemap indexed, not just for Google's indexing, but also for other software engines to retrieve content. Therefore in the next experiment we're going to assess the number of indexed pages, following submission of the full sitemaps, which should determine if the positive change in indexed pages was based on the entity pages or the lower number of high quality pages submitted. We would like to emphasize that success in this area is subject to any changes Google might make in future.

---

[2] Crawl rate is the number of pages Google accesses on the website to find search results for users.

The graphs below show the number of indexed pages throughout 2017 and the beginning of 2018. The graphs show indexed pages across both sitemaps that are currently available on Europeana. HTTPS sitemaps were introduced in October following the introduction of SSL across the majority of Europeana content. Where unsecured external content is shown on a page (for example, content embedded by third parties), we are redirecting HTTP to HTTPS. The majority of pages on Europeana are now on HTTPS.
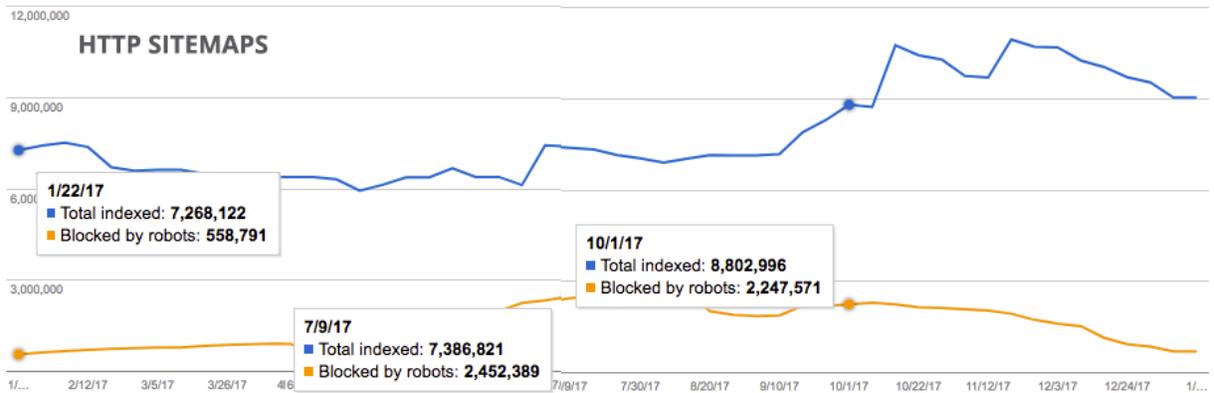
The numbers shown in the graph may be up to a week old, and may not reflect some filters that can prevent indexed URLs from appearing in our search results. Learn more.

Europeana Pro usage patterns

Europeana Pro is Europeana's main communication tool towards our professional target markets (Cultural Heritage Institutions, Academic Research, Education, and Creative Industries) as well as towards the wider Network of data partners and experts. The aim of the first four months of development of Europeana Pro was to improve the findability of information and to enhance user satisfaction of the site. User research (in Appendix 1) validated both issues and revealed that there is a need for a redesign of the website. Consequently, the new Europeana Pro website was beta launched in September 2017 and user research is currently being undertaken to see if the website made a positive change in findability of content and user satisfaction.

Conclusion

Based on the results from of the Europeana DSI-2 NPS survey we concluded that Europeana Pro should be redesigned, with particular attention to be paid to the information architecture elements of the website:

(1) sitemap, the way the information is clustered and structured on the website;
(2) navigation, the way users can reach the information;
(3) labelling, the words used to describe the information clusters.

Our propositions for the redesign of Pro were as follows:

- 27,8% of users rated the old Europeana Pro was easy to use with 9 and 10. We aim to increase this to 35%.
- 27,8% of users rated the ease of navigation of the old Europeana Pro with 9 and 10. We aim to increase this to 35%.
- 33.8% of users rated that old Europeana Pro met their expectations with 9 and 10. We aim to increase this to 35%.

In September 2017, we launched the new Europeana Pro website[3] (please see screenshot below), based on the knowledge we had gained from the surveys under Europeana DSI-2. Currently, we are running a survey asking the same questions as before on the new Europeana Pro website to compare the results. Based on these results we will see if the redesign had the expected impact. This will be reported on in the next version of this deliverable (M8, April 2018).

---

[3] New Europeana Pro website. View at https://pro.europeana.eu/

New Europeana Pro navigation

Annex 1: User research on Europeana Pro

Improving findability of information on Europeana Pro

We wanted to know if users can easily find the information on Europeana Pro that they are looking for. User feedback revealed that information could not easily be found on the previous (old) version of Europeana Pro.

We ran a survey on Hotjar asking our users:
1. How easy is it to use the website?
2. How easy is the website to navigate?

The survey ran for three months from May to August 2017. We received 166 responses.

How easy is it to use the website? (Users answered these questions using a 0-10 scale)

Responses are grouped as follows:
- Promoters (score 9-10) are enthusiastic users that find the website is easy to use. These are loyal users who will keep returning to the website and referring others, fueling growth.
- Passives (score 7-8) are satisfied but unenthusiastic users who are not too satisfied with how easy the website is to use. These users are vulnerable to competitive offerings.
- Detractors (score 0-6) are unhappy users who find the website difficult to use. These users can damage our brand through negative word-of-mouth.

Results can be summarized as follows:
- Promoters (score 9-10): 27,8% of users answered that they found the website easy to use.
- Passives (score 7-8): 43,4% of users answered that they are not too satisfied with how easy the website is to use.
- Detractors (score 0-6): 28,8% of users answered that they find the website difficult to use.

To get the Net Promoter Score (NPS) we subtracted the percentage of Detractors from the percentage of Promoters yielding a NPS score of -1. A NPS score of -1 signals that our users find the website difficult to use. A NPS that is positive (e.g. higher than zero) is felt to be good while an NPS of +50 is excellent[4]. A NPS score of -1 indicates that the website is not easy enough to use by users.

---

[4] Net Promoter. View at https://en.wikipedia.org/wiki/Net_Promoter

How easy is the website to navigate? (Users answered these questions using a 0-10 scale)

Responses are grouped as follows:
- Promoters (score 9-10) are enthusiastic users that find the website is easy to navigate. These are loyal users who will keep returning to the website and referring others, fueling growth.
- Passives (score 7-8) are satisfied but unenthusiastic users who are not too satisfied with how easy the website is to navigate. These users are vulnerable to competitive offerings.
- Detractors (score 0-6) are unhappy users who find the website difficult to navigate. These users can damage our brand through negative word-of-mouth.

Results can be summarized as follows:
- Promoters (score 9-10): 27,8% of users answered that they find the website easy to navigate.
- Passives (score 7-8): 44% of users answered that they are not too satisfied with how easy the website is to navigate.
- Detractors (score 0-6): 28.2% of users answered that they find the website difficult to navigate.

To get the Net Promoter Score (NPS) we subtracted the percentage of Detractors from the percentage of Promoters yielding a NPS score of -1. A NPS score of -1 signals that our users find the website difficult to navigate. A NPS that is positive (e.g. higher than zero) is felt to be good while an NPS of +50 is excellent. A NPS score of -1 indicates that the navigation on the website is not sufficient.

Improving Europeana Pro User satisfaction

When research showed that it was difficult for users to use Europeana Pro and to navigate on the website, Europeana further investigated possible reasons. We wanted to know if the website meets the expectations of users. The aim of this question was to understand if the information on the website is what people are looking for and find valuable. This question was asked additionally to the previous questions to see why people found it difficult to use and navigate the site. Was it because the content of the site is not what they expected?

We ran a survey on Hotjar asking our users:
1. To what extent does the website met your expectations?

The survey ran for three months from May to August 2017. We received 166 responses.

To what extent does the website met your expectations? (Users answered these questions using a 0-10 scale)

Responses are grouped as follows:
- Promoters (score 9-10) are enthusiastic users whose expectations are met by the website. These are loyal users who will keep returning to the website and referring others, fueling growth.
- Passives (score 7-8) are satisfied but unenthusiastic users whose expectations the website doesn't completely satisfy. These users are vulnerable to competitive offerings.
- Detractors (score 0-6) are unhappy users whose expectations the website does not meet. These users can damage our brand through negative word-of-mouth.

Results can be summarized as follows:
- Promoters (score 9-10): 33.8% of users answered that their expectations are met by the website.
- Passives (score 7-8): 35.6% of users answered that the website doesn't completely satisfy them.
- Detractors (score 0-6): 69.4% of users answered that the website does not meet their expectations.

To attain the Net Promoter Score (NPS) we subtracted the percentage of Detractors from the percentage of Promoters yielding a NPS score of 3. A NPS score of 3 signals that the website generally meets the expectations of our users. A NPS that is positive (e.g. higher than zero) is felt to be good while an NPS of +50 is excellent. A NPS of 3 is positive, however the number is very low.