



# **DELIVERABLE**

Project Acronym: Europeana Newspapers

**Grant Agreement number:** 297380

**Project Title:** A Gateway to European Newspapers Online

# D4.6 Newspaper Indexing and Interface Development D4.7 Usability Testing

Revision: 1.0

Authors: Alastair Dunning (The European Library / Europeana Foundation)

Alena Fedesenka (The European Library / Europeana Foundation) Anastasia Gasia (The European Library / Europeana Foundation)

Markus Muhr (The European Library / Europeana Foundation)

Project co-funded by the European Commission within the ICT Policy Support Programme				
Dis	Dissemination Level			
Р	Public	x		
С	Confidential, only for members of the consortium and the Commission Services			





## **Revision History**

Revision	Date	Author	Organisation	Description
0.1	16-03-2015	Alastair Dunning	TEL	Updated draft
		Alena Fedesenka	TEL	version
		Anastasia Gasia	TEL	
		Markus Muhr	TEL	
		Joris Pekel	EF	
0.2	22-03-2015	Neil Fitzgerald	BL	Internal review
		Adam Sofronijevic	UB	
0.3	30-03-2015	Clemens Neudecker	SBB	Internal review
		Hans-Jörg Lieder	SBB	
		Sandra Kobel	SBB	
0.4	13-04-2015	Alastair Dunning	TEL	Following feedback from reviewers
0.5	29-04-2015	Alastair Dunning	TEL	Following request from Commission to present D4.6 and D4.7 as separate from D4.5.
1.0	30-04-2015	Clemens Neudecker	SBB	Internal review and
		Sandra Kobel	SBB	final version

# 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This deliverable contains original unpublished work except where clearly indicated otherwise.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.





## **Table of contents**

1.	Introduction	4
2.	Creating an Interface to the Aggregated Data	4
3.	Usability Testing	6
4.	Surfacing the Data in Europeana	11
5.	Future Plans	12
6.	Conclusion	14
Apı	pendices	15
	I - Display Options	15
	II - End-User Statistics - The European Library	16
	III - End-User Statistics - Europeana	21





# Introduction

While deliverable *D4.5 Newspapers aggregated by TEL* explained the aggregation process for Europeana Newspapers, these co-joined deliverables (D4.6 and D4.7) look at the process of creating an interface from the aggregated corpus of metadata, full text and images.

This deliverable explores the process of planning and designing the interface, taking into account the needs of users but also the needs of the libraries contributing data. It provides oversight of the results from the usability testing and how the recommendations were incorporated into The European Library. The appendix includes quantitative evidence of the usage of TEL and Europeana. For those wanting more detail on the usability testing, the report produced in 2014 will be of interest.<sup>1</sup>

The report also shows how the newspapers were incorporated into the Europeana portal.

It concludes by looking at the future plans for the newspapers, which include plans for closer integration with Europeana. More technical details are available in a separate document.<sup>2</sup>

# Creating an Interface to the Aggregated Data

At the start of the project in 2012, the plans for searching and browsing the aggregated data had still to be formulated. There was a clear mandate to share metadata with Europeana but it was unsure as to how this would incorporate the images and full text that was part of the project.

As The European Library was undertaking the aggregation and had built expertise in text aggregation from a previous project, it was decided that a dedicated newspapers interface would be built by the team at The European Library.

Before beginning any design work each of the library partners was contacted to discuss how they would like their data represented in The European Library. This was necessary to build trust with the library partners and to ensure that the newspapers were represented in a way that did not conflict with the strategic priorities and business models. Libraries could choose from various options - these are outlined in Appendix 1

The following part of this process was to develop wireframes that offered various functions to the user - not just full text search but the option to browse through various highlighted newspapers and limit searches according to date.

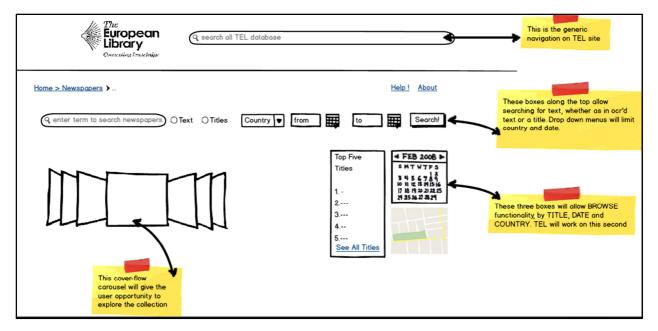
\_

<sup>&</sup>lt;sup>1</sup> Alan Blackwood, The European Library: Newspaper Archive – Usability Testing <a href="http://www.europeana-newspapers.eu/wp-content/uploads/2014/05/The-European-Library-Newspaper-Archive-Usability-testing-Report-April-2014.pdf">http://www.europeana-newspapers.eu/wp-content/uploads/2014/05/The-European-Library-Newspaper-Archive-Usability-testing-Report-April-2014.pdf</a>

<sup>&</sup>lt;sup>2</sup> Markus Muhr, Europeana Newspapers: Integration into the Europeana Infrastructure, https://docs.google.com/document/d/1VLULSdSMP4Dg2MWmtBbB5cRcuErVFaewX8y-06FulNs/pub

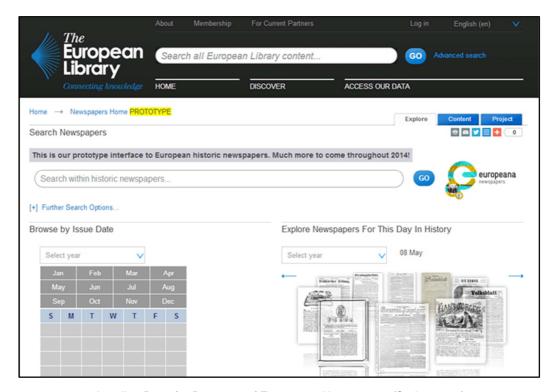






Early wireframe for Newspapers interface

Once an initial prototype was created and some of the first aggregated content added, it was tested informally by project members. Further adjustments were made as more content was added. Once the back end of the interface was also stable, the project could proceed with the first round of usability testing.



Landing Page for Prototype of Europeana Newspapers (Spring 2014)





# **Usability Testing**

An Edinburgh-based firm UserVision was appointed to undertake user testing, working with researchers that were patrons of four of the participating libraries. Working remotely, each of the volunteers was given a set of tasks by UserVision who then recorded their behaviour and interaction with the site.

The key findings of the usability testing were

- 1. A positive impression on the breadth and depth of content
- 2. Good understanding of the basic layout and functionality of the site
- 3. Appreciation of the different functionality although in many cases this functionality was difficult to find

The full report is available from the project website.<sup>3</sup>

Following this report, the interface was further developed. Notable changes included.

- 1. Allowing the basic search box to stand out more
- 2. Allowing a greater range of functionality via a drop down box (i.e. filtering by date, language and library)
- 3. Making a clear visual distinction between searching browsing functionality (e.g. including icons for exploring by country, date and title or showing all titles or issues
- 4. Making it easier to navigate between individual issues and search results
- 5. Adding in a full screen functionality

Other functionality was included as more content arrived from the partners in the project.

Where article level issues had been provided to The European Library, this could be reflected in the search results. Also, where a single search term appeared multiple times in the same issue that could also be shown in the newspapers, as seen in the screenshot below.

<sup>&</sup>lt;sup>3</sup> Alan Blackwood, The European Library: Newspaper Archive – Usability Testing <a href="http://www.europeana-newspapers.eu/wp-content/uploads/2014/05/The-European-Library-Newspaper-Archive-Usability-testing-Report-April-2014.pdf">http://www.europeana-newspapers.eu/wp-content/uploads/2014/05/The-European-Library-Newspaper-Archive-Usability-testing-Report-April-2014.pdf</a>.





You are looking at 10 items out of 432,309 for dollar Relevance Sort by: Items per page: Deutsches Nachrichtenbüro / 02 Sep 1936/ Page 1 (Article Title: amerikanische Staatshaushalt für 1936/37.) Der amerikanische Staatshaushalt für 1936/37. Washington, 2. September. In der vom Schatzamt am Dienstag ausgegebenen überprüften Haushaltsübersicht schätzt Präsident Roosevelt die Gesamteinnahmen des am 30. Juni 1937 endenden Haushaltsjahres auf 5 665 839 000 Dollar, einbegriffen 410 000 000 Deutsches Nachrichtenbüro / 03 Jun 1939/ Page 1 (Article Title: Amerika hat imperiale Pläne am Südpol.) Anfang Herbst vorgeschlagen und die Bewilligung der veranschlagten 340 000 Dollar befürwortet habe. Der amerikanische Admiräl erklärte, daß die Expedition zur "Sicherung der Rechte der Vereinigten Staaten" not wendig sei: Die USA. sollten rasch handeln, um sich das an Erdschätzen reiche Gebiet zu sichern [-] More occurrences in the same newspaper issue. Deutsches Nachrichtenbüro / 03 Jun 1939/ Page 1 (Article Title: USA-Ministerium hat es eilig.) USA-Ministerium hat es eilig. Washington, 3. Juni. Das amerikanische Marinemini sterium vergab am Freitag an Privat- und Regierungswerf ten die Bauaufträge für 24 Kriegsschiffe im Gesamtwert' von 359 Mill. Dollar, darunter Zwei 45 000 Tonnen- Schlachtschiffe, die von den Regierungswerften in New Deutsches Nachrichtenbüro / 03 Jun 1939/ Page 1 (Article Title: "Darlehen" des jüdischen "Rechts"-Anwalis.) Co. "Dar lehen" im Gesamtbetrage Von etwa 300 000 Dollar erhalten habe, als ein Prozeß dieser beiden Konzerne vor seinem Ge richt verhandelt wurde. Der Staatsanwalt stellte fest, daß Manton damals den die Tobacco vertretenden jüdischen Rechtsanwalt Louis Levy fragte, wer einer von Manton

Search Results - Article titles included and multiple occurrence of same search page in same issue. Try <a href="http://www.theeuropeanlibrary.org/tel4/newspapers/search?query=dollar.">http://www.theeuropeanlibrary.org/tel4/newspapers/search?query=dollar.</a>





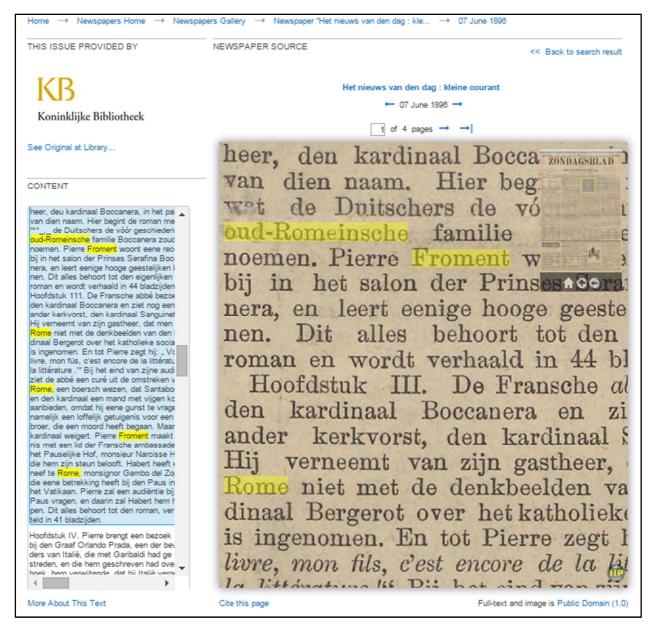


Individual item page featuring text highlighting, licensing mark, citation details, links to other pages and options for resizing the image.

 $\frac{\text{http://www.theeuropeanlibrary.org/tel4/newspapers/issue/3000100609254?hp=1\&count=10\&page=1\&query=rome\&deca}{\text{de}=1890\text{-}1899}$ 



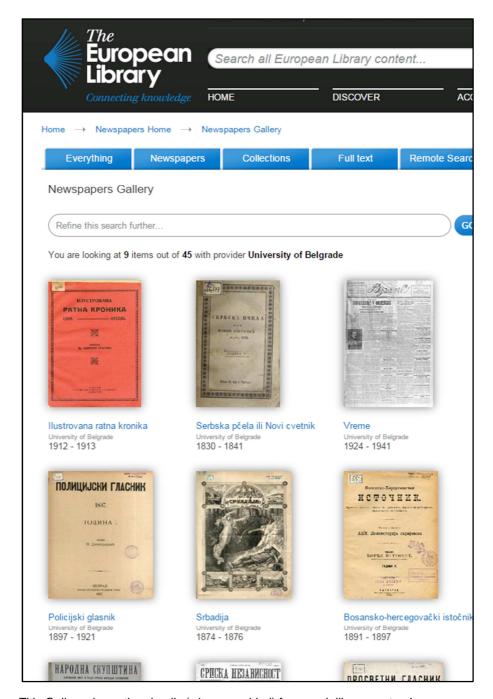




Clicking on the text in an individual issue page allows a paragraph to be highlighted and the image to be subsequently resized.







Title Gallery shows thumbnails (where provided) from each library partner's newspaper <a href="http://www.theeuropeanlibrary.org/tel4/newspapers/gallery?provider-id=P01360&count=9.">http://www.theeuropeanlibrary.org/tel4/newspapers/gallery?provider-id=P01360&count=9.</a>

A second round of usability testing was also planned. However, this has now been postponed and will be take place when the portal is migrated to Europeana - further details of this are in the section on Future Plans.

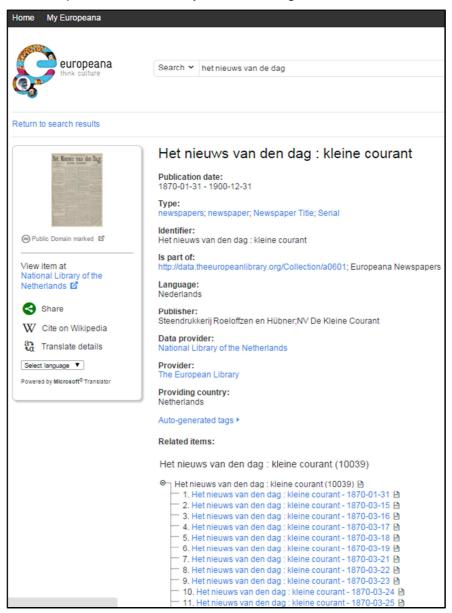




# 4. Surfacing the Data in Europeana

As part of the aggregation process, each metadata record for every issue and title was converted from its native format into a representation in the Europeana Data Model (EDM). This allows the data to be shown in the Europeana Portal.

Europeana could also explore some new ways of visualising their data:



Issue level record in Europeana featuring hierarchical metadata, showing relationship between different issues of the same title







Clicking on the thumbnail on Europeana opens up the full screen browser.

# 5. Future Plans

Originally, there had been plans for The European Library to become formally independent of Europeana and develop their own business model to sustain the content and newspapers interface.

However, the boards of The European Library and Europeana decided that there was too much financial and legal risk to separation, and it was decided to follow a process of closer reintegration.

This therefore gives the possibility of creating closer integration between TEL and Europeana and the newspapers.

Throughout the course of 2015 the following actions will take place:





- The TEL portal will continue to serve users in 2015. Minor adjustments will be made to the interface from May 2015 onwards. This work will be funded by the library membership fees paid to TEL by national libraries.
- Full-text content will be harvested from Associate Partners. At time of writing full text from Iceland and Belgium has already been aggregated.
- Various avenues to explore the corpus of newspapers data for the digital humanities are being explored. A workshop to explore the possibilities is organised for May 2015.<sup>4</sup>
- An experimental API is being released to select partners for those that want to test building application on top of the corpus.
- Simultaneously with the above, the data will be migrated from the TEL servers in London to Europeana Cloud, Europeana's data storage system. Much further details are provided in the Migration to Europeana document.<sup>5</sup>
- Design work will begin on a Europeana Channel for showing newspapers, incorporating previous user feedback. Channels are Europeana's concept for thematic portals, i.e. discovery services that focus on particular themes or subjects (e.g. Europeana Art, Europeana Music). This new interface will be ready in early 2016.
- A business model will be developed extending the TEL membership fee to allow for the aggregation of metadata and full text. An early version of this will be presented at the CENL Annual General Meeting in Bern in May. This will allow for the development of the interface and the integration of new content. Several libraries have indicated they can share more data (Iceland, Spain, Belgium, Wales) and discussions have already started on the technical aspects of this.
- Once the Europeana Newspapers channel is complete, the existing TEL interface will be deprecated.

<sup>&</sup>lt;sup>4</sup> See <a href="http://research.europeana.eu/blogpost/found-researchers-wanting-to-use-historic-newspapers">http://research.europeana.eu/blogpost/found-researchers-wanting-to-use-historic-newspapers</a> for more details.

<sup>&</sup>lt;sup>5</sup> https://docs.google.com/document/d/1VLULSdSMP4Dg2MWmtBbB5cRcuErVFaewX8y-06FulNs/pub.





# 6. Conclusion

#### **Interface Development**

In terms of interface development, the project decision to launch an early version of the interface was justified. This allowed the team to use the project funding to receive feedback, iterate the design and produce a more user-focused interface by the end of the project. It is difficult to see when this agile form of development should not be recommended for any cultural heritage project.

#### **User Testing**

User testing clearly indicates the preference of users to **be able search over multiple archives and to do so over full text**. This is demonstrated in qualitative terms, i.e. the responses from user testers in the usability testing report. It is also made apparent in quantitative terms, i.e. the user statistics. The preference for users to search through the detailed interface made available by the European Library, as opposed to the metadata only option that Europeana provides.





# **Appendices**

# I - Display Options

Each of the full partner libraries had six options as to how to deliver their images to The European Library

- 1. Full Text, Descriptive Metadata and Viewing Images images hosted by TEL
- 2. Full Text, Descriptive Metadata and Excerpts of Images images hosted by TEL
- 3. Full Text and Descriptive Metadata Only
- 4. Descriptive Metadata Only
- 5. Full Text, Descriptive Metadata and Viewing Images images hosted by library
- 6. Full Text, Descriptive Metadata and Excerpts of Images images hosted by library

Library	Option			
	Option 5) Full Text, Descriptive Metadata and Viewing Images - images hosted by library- up to 1876			
National Library of Austria	Option 4) Descriptive Metadata Only - 1876 onward			
National Library of France	Option 5) Full Text, Descriptive Metadata and Viewing Images - images hosted by library			
	Option 5) Full Text, Descriptive Metadata and Viewing Images - images hosted by library - up to 1900			
National Library of Netherlands	Option 4) Descriptive Metadata Only - from 1900			
Dr. Friedrich Teßmann State Library	Option 2) Full Text, Descriptive Metadata and Excerpts of Images - images hosted by TEL			
National Library of Estonia	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			
National Library of Latvia	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			
National Library of Finland	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			
State and University Library of Hamburg	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			
State Library of Berlin	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			
University Library of Belgrade	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL			





National Library of Poland	Option 1) Full Text, Descriptive Metadata and Viewing Images - images hosted by TEL
National Library of Turkey	Option 5) Full Text, Descriptive Metadata and Viewing Images - images hosted by library (yet to be fully implemented)

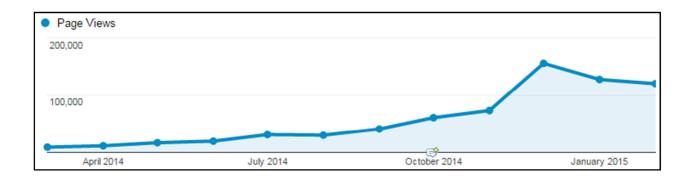
# II - End-User Statistics - The European Library

The following statistics are drawn from Google Analytics for <a href="https://www.theeuropeanlibrary.org/tel4/newspapers">www.theeuropeanlibrary.org/tel4/newspapers</a> for the period of 1 March 2014 to 28 February 2015.

## **Highlights**

- The quantity of visits is not particularly high (c.700,000 page views in a year) however, **the quality of visits is exceptionally high** (15 minutes per visit in Feb 2015)
- Wikipedia is an important source of traffic, providing nearly a quarter of all incoming links
- Ongoing marketing is also essential for driving traffic to the site as is shown by the spike in December
- Users come from across the world; the largest number of users comes from Germany (21%)

#### What is the quantity of page views?

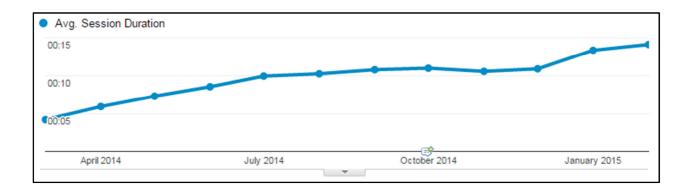


There were 694,845 page views in total, with a peak of 155,164 in December 2014. The increase in autumn 2014 can perhaps be assigned to the large increases in content added to the site. The statistics show the importance of marketing - The rise in December 2014 is due to the significant marketing push made at that time.





## How long do their sessions last?



The average session has lasted 11 minutes and 8 seconds and has risen to nearly 15 minutes by February 2015. **This is exceptionally high**; the average for the TEL site as a whole (i.e. all its content, not just newspapers) is 4 minutes and 28 seconds.

## What do you users search for?

Se	earch Term ②	Total Unique Searches ? ↓	Results Page Views/Search ?	% Search Exits ?	% Search Refinements ?	Time After Search	Average Search Depth ?
		23,052 % of Total: 100.00% (23,052)	4.30 Avg for View: 4.30 (0.00%)	18.71% Avg for View: 18.71% (0.00%)	20.48% Avg for View: 20.48% (0.00%)	00:04:44 Avg for View: 00:04:44 (0.00%)	0.61 Avg for View: 0.61 (0.00%)
1.	"alessandro volta"	79 (0.34%)	1.73	96.20%	1.46%	00:00:31	0.01
2.	japon	60 (0.26%)	12.23	85.00%	1.23%	00:32:19	0.85
3.	landshut	49 (0.21%)	2.47	28.57%	38.02%	00:02:20	0.00
4.	"karl storch"	40 (0.17%)	2.58	25.00%	29.13%	00:01:50	0.35
5.	"Léon Delafosse"	35 (0.15%)	3.00	42.86%	24.76%	00:04:53	0.03
6.	schach	32 (0.14%)	8.88	56.25%	8.45%	00:12:10	0.59
7.	ringkampfkonkurrenz	31 (0.13%)	1.00	90.32%	9.68%	00:00:03	0.00
8.	titanic	31 (0.13%)	2.32	93.55%	2.78%	00:00:41	0.03
9.	"Käthe Hildebrandt"	29 (0.13%)	1.52	3.45%	81.82%	00:03:22	0.00
10.	"max linder"	27 (0.12%)	12.74	33.33%	1.45%	00:13:06	4.93
11.	horbaczewski	26 (0.11%)	1.15	96.15%	0.00%	00:01:57	0.04
12.	magnussen	26 (0.11%)	1.15	11.54%	83.33%	00:01:35	0.00
13.	Groche	<b>25</b> (0.11%)	1.04	12.00%	88.46%	00:00:19	0.00
14.	Kohlhage	<b>25</b> (0.11%)	1.28	4.00%	84.38%	00:00:14	0.00
15.	papus	<b>25</b> (0.11%)	3.32	64.00%	21.69%	00:09:45	0.00

Unsurprisingly, the search terms entered by users show the long tail effect - a very wide range of queries and interest are shown by the users, with no term being repeated 100 times.

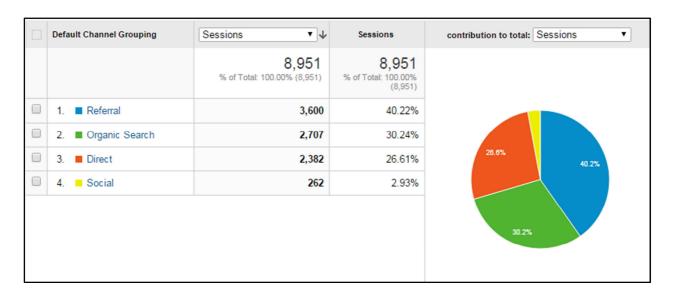




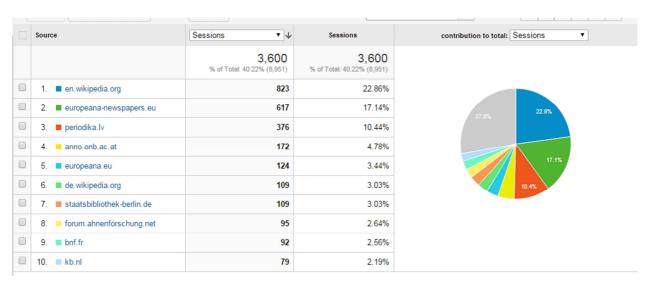
#### How do users locate the site?

Google Analytics groups four types of sources

- 1. Referral link from the world wide web
- 2. Organic Search from a search engine
- 3. Direct a user entering the URL into the browser address bar
- 4. Social from social media platform



#### Where are users referred to from the site?



Wikipedia is the most important source of users. As far as the authors are aware, there are only around 20 to 25 links from Wikipedia to the newspapers site but they provide nearly a quarter of incoming referrals.

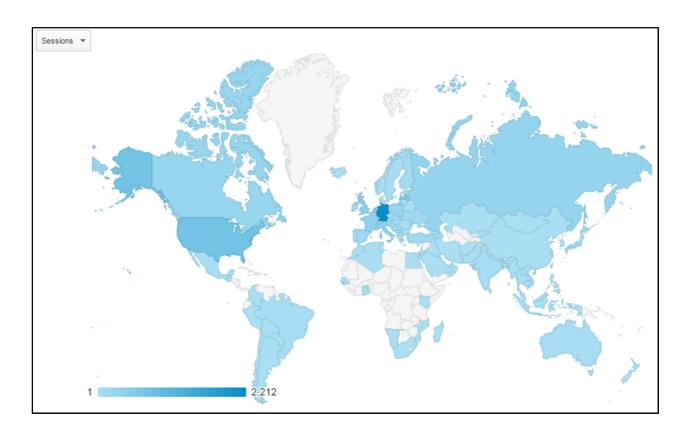




#### What countries do users come from?

	Country	Sessions	% Sessions
1.	Germany	10,679	21.18%
2.	United States	4,802	9.52%
3.	■ Netherlands	3,246	6.44%
4.	<b>■</b> Latvia	2,470	4.90%
5.	器 United Kingdom	2,461	4.88%
6.	<b>■</b> Italy	2,316	4.59%
7.	■ Poland	1,982	3.93%
8.	France	1,911	3.79%
9.	Austria	1,659	3.29%
10.	Russia	1,329	2.64%
			view full report

While Germany provides the largest number of users they are far from being the majority. Despite little English content there are also significant users from US and UK.

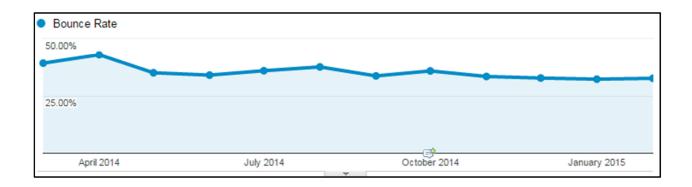


Users come from all across the world. This map shows number of sessions per country from 1 (various countries) to 2,212 (Germany). This may be because of the German newspaper content on the site that has not been published elsewhere.





## What is the Bounce Rate? I.e. How many users only look at one page then leave?



There is an average bounce rate of 34%. This was relatively stable, after the first few weeks (when there was very little content on the site and therefore less reason for users to stay). Again, this is a low result - the average for the TEL site is 52%.

#### How many pages do users look at per session on the site?



As content has been added to the site, the average number of pages per session has risen to 13.78 with a high of 15.60 in January 2015. Again, this compares favourably with TEL site as whole, where the figure is 5.60.





# III - End-User Statistics - Europeana

(Contributed by Joris Pekel, Europeana)

The following statistics are drawn from Google Analytics for the Europeana portal for the period 1 March 2014 to 28 February 2015. The same depth of analysis is not possible for Europeana as they do not have a dedicated newspapers interface. The statistics are therefore based on an items that have been marked as coming from this project, i.e. items included within this subset <a href="http://europeana.eu/portal/search.html?query=europeana\_collectionName%3A92\*ewspapers\*&rows=24">http://europeana.eu/portal/search.html?query=europeana\_collectionName%3A92\*ewspapers\*&rows=24</a>

What is the quantity of page views? 8,455 page views

What is the average time spent on site? 1 min 43 seconds

What is the Bounce Rate? I.e. How many users only look at one page then leave? 46,89%

21 / 21

How many pages do users look at per session on the site? 3.4 pages