



# Report of the Task Force Benchmarking Europeana

26 June 2017



## The Taskforce in a nutshell

### *Participants*

1. Stephan Bartholmei (Deutsche Digitale Bibliothek)
2. Laura Carletti (University of Nottingham)
3. Alexandros Charkiolakis (MIAM - Istanbul Technical University / Friends of Music Society in Athens)
4. Francesca Di Donato (Net7)
5. Monika Hagendorn-Saupe (Institute for Museum Research, SMB-SPK)
6. Adrian Murphy (on behalf of the Europeana Office)
7. Marco de Niet (DEN Foundation, chair)
8. Gerhard Jan Nauta (DEN Foundation)
9. Olivier Schulbaum (Platoniq)
10. Marcin Werla (Poznań Supercomputing and Networking Center)

### *Purpose*

To evaluate the methodology and content of the first study to benchmark Europeana as a cultural information service in order to propose a framework for future evaluation and benchmarking of Europeana services and a work plan for its implementation. The Study Report was written by Trilce Navarrete-Hernandez of the DEN Foundation as a contribution to the Europeana DSI-project. The report was published on 12 September 2016 and is available on Europeana Pro (<http://pro.europeana.eu/publication/benchmarking-europeana-against-other-web-services>). For the evaluation, we will also look at other reports in which the Europeana portal is evaluated or benchmarked.

### *Expected outcomes*

- An assessment of the methodology used in the Study Report "Europeana as online cultural information service" (September 2016).
- An assessment of the content of the Study Report "Europeana as online cultural information service" (September 2016) as input for the formal evaluation of Europeana in 2017.
- Recommendations about the continuation of benchmarking activities in the context of Europeana
- Proposal to set up a Europeana MC Working Group on benchmarking



# INDEX

<b>The Taskforce in a nutshell</b>	<b>2</b>
<i>Participants</i>	2
<i>Purpose</i>	2
<i>Expected outcomes</i>	2
<b>1. Introduction</b>	<b>4</b>
<b>2. The Benchmarking Europeana report</b>	<b>6</b>
<b>3. Why benchmark Europeana?</b>	<b>7</b>
<b>4. Evaluation of the methodology used</b>	<b>9</b>
<b>5. Evaluation of specific elements in the report</b>	<b>12</b>
<b>6. Evaluation of choice for benchmarked services</b>	<b>20</b>
<b>7. Conclusion and recommendations</b>	<b>21</b>
<b>Appendix A: List of references</b>	<b>22</b>
<b>Appendix B: Proposal for description of indicators</b>	<b>23</b>



# 1. Introduction

In May 2016 Europeana approached the DEN Foundation, partner in the Europeana DSI project for the ENUMERATE statistical framework, to perform a study on benchmarking Europeana as a cultural information service. The report from this study had to be delivered by the end of June 2016 (the end date of the Europeana DSI-1 project). DEN found Trilce Navarrete-Hernandez, formerly researcher at the Universities of Amsterdam and South-Denmark, available and willing to take on this study. The final version report was delivered in July 2016 and afterwards published on Europeana PRO.

Because of the time constraints during the study, there was no time for an extensive consultation among members of the Europeana Network. Both the development of the methodology for the benchmark and the selection of and comparison with other cultural information services was done on the authority of the team at DEN in close cooperation with the Europeana Office. It was agreed then that DEN would submit a proposal with the Europeana Members Council to set up a Taskforce to evaluate the benchmark study and its results. A proposal for such a Taskforce was submitted in November 2016 and approved by the Europeana Members Council in December accordingly. The Taskforce ran from January 2017 until July 2017.

The purpose of the Taskforce was "to evaluate the methodology and content of the first study to benchmark Europeana as a cultural information service in order to propose a framework for future evaluation and benchmarking of Europeana services and a work plan for its implementation. [...] For the evaluation, we will also look at other reports in which the Europeana portal is evaluated or benchmarked."

The following people volunteered to participate in this Taskforce and contributed to the evaluation of the benchmarking report.

- Stephan Bartholmei (Deutsche Digitale Bibliothek)
- Laura Carletti (University of Nottingham)
- Alexandros Charkiolakis (MIAM - Istanbul Technical University / Friends of Music Society in Athens)
- Francesca Di Donato (Net7)
- Monika Hagendorn-Saupe (Institute for Museum Research, SMB-SPK)
- Adrian Murphy (on behalf of the Europeana Office)
- Marco de Niet (DEN Foundation, chair)
- Gerhard Jan Nauta (DEN Foundation)
- Olivier Schulbaum (Platoniq)
- Marcin Werla (Poznań Supercomputing and Networking Center)

This document is the final report of the Taskforce to the Europeana Members Council. All members of the Taskforce contributed actively in the creation of this document. A physical meeting of the Taskforce took place on Wednesday 3 May 2017 in The Hague, in which 7 of the 10 Taskforce members were able to participate.

## *Final report on Benchmarking Europeana*



In this report we will first describe briefly the methodology and the outcomes of the benchmarking report as published in 2016 and address the question why it is useful to benchmark Europeana. Then we will zoom in on the methodology used in the 2016 study, assess the results from that study and discuss the other information services that were used to benchmark Europeana. We end this report with recommendations for future activities to benchmark Europeana.



## 2. The Benchmarking Europeana report

The study report *Europeana as online cultural information service, Benchmarking project* came about during the months of May-July 2016 and was published in a final version on 15 September 2016. The report was written in close collaboration between staff at the DEN Foundation and the Europeana office. The final report is available on Europeana Pro: <http://pro.europeana.eu/publication/benchmarking-europeana-against-other-web-services>.

During the study, the needs of Europeana for benchmarking were taken into account. Europeana had been actively developing its impact framework, which could use input from the benchmark; also the benchmark would be useful to bring data to Europeana Pro and its dashboard for data providers; finally the European Commission had announced that it would do a formal evaluation of Europeana in 2017, as a result of the Council Conclusions from 2016, established under the Dutch presidency of the European Union.

For these reasons, it was decided in a preparatory meeting that the three core values of Europeana ('usable', 'mutual' and 'reliable') as defined in the Europeana Business Plan were used as the starting point for the benchmark of Europeana.

The benchmark report consists of two main parts: the first part contains the results of the benchmark study, the second part (the annex) is an extensive explanation of the methodology used in the study.

The first part of the report is divided into four chapters:

- Chapter 1 presents the 19 other online cultural information services that were used as a benchmark for Europeana, describing the similarities and differences among these services and Europeana.
- Chapter 2 provides the results of a quantitative analysis and comparison of the online Europeana service and the 19 other cultural information services. It looks at topics like online popularity, size of digital collections and social media presence.
- Chapter 3 is based on an analysis of the 100 most popular items in Europeana in 2015 as supplied by the Europeana office. This chapter is meant to provide context to the quantitative comparisons from chapter 2 and support conclusions about the core values of Europeana (usable, mutual and reliable).
- Chapter 4 presents the final conclusions of the study in relation to the three values of Europeana

The annex explaining the methodology was added as an integral part of the report as this was the first time that Europeana was benchmarked on such an extensive scale and it was the intention of Europeana to repeat this benchmarking in the future. Also, both DEN and Europeana were aware of the limitations of the benchmark, caused by time restrictions and lack of predecessors. A careful explanation of the choices made during the study would be helpful for later evaluation of the study and future benchmarking activities.



### 3. Why benchmark Europeana?

During the preparations for the study report there was no opportunity to discuss the rationale underlying benchmarking at length, but it came up as an important question during the meeting of the Taskforce on 3 May 2017 in The Hague. The Taskforce looked at the importance of benchmarking Europeana from four perspectives, representing four major stakeholders:

- 1) Europeana itself
- 2) the European Commission
- 3) the Cultural Heritage Institutions in Europe
- 4) Europeana's users.

A comprehensive benchmark of Europeana should take all these perspectives into account, as the understanding of the benefits of benchmarking from these four perspectives will help to ask all the right questions in the next benchmarking study.

#### *1) Europeana's perspective*

There are several reasons why Europeana would want to be benchmarked. Firstly, it will help identify strong and weak points, both within Europeana and other cultural information services. Understanding your own strong points contribute to your advocacy, identify your unique selling points and finding support for your mission.

It is likely that a benchmark study will not come up with positive results only. The Taskforce feels strongly that negative outcomes for Europeana should not be brushed under the carpet but treated as learning points, even though they may be received with criticism in the outside world.

This relates to another reason why benchmarking is important: to be transparent in your accountability. Can you prove through comparison with others that you have used your resources wisely? As such, benchmarking relates to all the strategic goals of Europeana as expressed in the Business Plan 2016-2017.

Benchmarking also relates to the ambition of Europeana to develop a comprehensive Impact Framework, that is needed to better understand the value that Europeana creates for its various user groups ('markets'). Benchmarking results can provide useful input for such impact studies.

Finally, at a more generic level, it was stated that Europeana can also set an example by having itself benchmarked. Benchmarking digital services is not yet a common practice in the European Cultural Heritage sector and Europeana could prove itself as a frontrunner and trendsetter by doing so.

#### *2) The European Commission*

For the European Commission it is important to understand that Europeana does its job well and deserves to remain an EU flagship project. A benchmark study may provide factual and objective information of the position of Europeana in the digital information landscape in general and the digital heritage infrastructures in particular, and as such may inform policy



makers how to position Europeana in funding and tendering schemes. The Commission could also use benchmarking results to set priorities for Europeana.

### *3) European Cultural Heritage Institutes*

For the Cultural Heritage Institutions in Europe the benchmarking of Europeana will help to make informed decisions about how the mission of Europeana aligns with their own mission and what, in that respect, the added value of Europeana is for them in relation to other cultural online services, e.g for multiplying their digital content to user groups that they cannot reach themselves. Also, for those CHI's that already supply their digital content to Europeana, the benchmark results of Europeana may support their accountability to their own funders.

### *4) Europeana users*

Finally, for the users of Europeana, seeing how Europeana relates to other information services they may or may not use as well, contributes to an enhanced understanding of the values of Europeana for them, e.g. as a provider of trusted content, or regarding the quality, the reuse options or the coverage of the full content of Europeana. The communication of the benchmarking results to Europeana's end users may need a different approach than to the other stakeholders.

In the context of the users' perspective the Taskforce discussed the relationship between the benchmark and Europeana's Impact Framework more closely. It was agreed, that the benchmark study is about the connection between Europeana as a whole and the outer world by comparing characteristics, while the impact studies mostly relate to the changes brought about by Europeana and its content to Europeana's user groups (the five markets: European citizens, Research, Education, Creative industries and the Cultural heritage institutions). It will be worthwhile to explore whether impact on a specific user group (e.g. Education) can be benchmarked among various online information services. But this will have to wait until the Europeana Impact Framework has been fully developed and implemented.



## 4. Evaluation of the methodology used

The Taskforce addressed the evaluation of the report in two steps. First, the individual members provided answers to a set of 10 questions that were prepared by the chair of the Taskforce. During the physical meeting in The Hague on 3 May 2017, these answers were discussed collectively in order to find common ground in the opinions of the Taskforce members. This document reflects the common ground as found during the meeting and refined afterwards in reviewing draft versions of this report.

### *Europeana's mission as the starting point*

The first issue discussed about the methodology as used in the benchmarking report was the choice to use the three core values of Europeana ('mutual', 'reliable', 'usable') as the starting point for the benchmark. The Taskforce didn't think this was a good choice. The values are on the one hand too broad as a concept, lacking specificity to support comparison and measurement, and on the other hand, as a combination they are selected specifically for Europeana to address internal targets, making overall comparisons with other cultural information services more difficult. This is reflected in the conclusions of the report, which are very generic and not necessarily useful to understand the challenges that Europeana is facing and to implement improvements based on what we have learned from other services. Also, the values relate to the 'product' Europeana Collections and its underlying ecosystem, and but less so to processes, communication activities and costs involved, which also deserve to be included in a benchmark.

In the benchmark study, the three values were broken down into characteristics, and these characteristics led to the formulation of 40 indicators used in the benchmark. This three-step approach is in itself well done, and a good compensation for using the three generic values as a starting point. Discussing this approach led the Taskforce to the suggestion to not use the values of Europeana as starting point for the benchmark, but the mission (aspiration) of Europeana. The Europeana Foundation Board is currently in the process of rewording the mission. The current version, which has not yet received its final approval, reads: "*We transform the world with culture. We build on Europe's rich cultural heritage and make it easier for people to use for work, learning or fun. Our work contributes to an open, knowledgeable and creative society.*"). The mission is (or should be) a short, condensed verbal expression of everything Europeana should be about, and is (or should be) able to immediately highlight the similarities and differences with other online information services with a related mission, more so than the values. So the Taskforce suggests to use the mission of Europeana as the starting point, break that down into elements and for each element choose one or more indicators for the benchmark.

### *A modular approach*

By putting the mission of Europeana at the heart of the benchmarking, the Taskforce also addresses the question whether the benchmark should only address Europeana as a cultural information service (Europeana Collections) or have a wider view on Europeana. The Taskforce is, evidently, in favour of benchmarking wider aspects of Europeana as well (e.g. the governance, the budget and networking aspects). As such, it is clear that not all other entities (organisations, services, data collections) that will be selected for the benchmark of



Europeana, can be used in the full framework of the methodology. Therefore the Taskforce proposes a stepwise approach, in which not all entities are compared to each other in one run using a single set of indicators, but in which the benchmark is broken down into components or modules and for each component the relevant indicators and entities to benchmark Europeana with are selected. This modular approach also makes it possible to benchmark specific characteristics of Europeana against generic services like Google, Amazon etc without giving the impression that Europeana as a whole can be compared to such parties. Also, this diversification makes it possible to benchmark from different user perspectives: a scholar has different requirements and thus evaluation criteria of a service than a cultural entrepreneur, etc.

### *Indicators*

As for the indicators, the Taskforce strongly suggests that instead of inventing our own indicators for benchmarking Europeana, we first look at existing models that already provide principles and indicators for information services, such as the Fair Data Principles (FDP)<sup>1</sup>, the HEART framework (Happiness / Engagement / Adoption / Retention / Task success)<sup>2</sup> and the ENUMERATE indicators<sup>3</sup>. Applying existing indicators enlarges the chances of obtaining relevant data from other entities and it may even enhance the reliability of the outcomes as it may become possible to compare the outcomes of the Europeana benchmark with other benchmarking studies. Also, the benchmarking report doesn't always provide clear references to the indicators used. In Appendix B the Taskforce provides a model to formally document the indicators that can be used in future benchmarking activities.

### *Generic or specific?*

This leads to the question whether the methodology of benchmarking Europeana should be so generic and standardized that it can be used in other contexts as well. The Taskforce was ambiguous about this. We know that the Digital Public Library of America has shown an interest in the benchmark of Europeana, and it would be interesting to see whether the methodology framework could be re-used in other contexts. However, the Taskforce is cautious about turning this into a holistic approach to benchmarking, as we lack the proper skills and know-how to do this. The core of this activity is, ultimately, to better understand the role and position of Europeana, and this can be managed within the context of the Europeana Members Council and this Taskforce. A higher ambition needs a different setting with other resources.

### *Definitions*

A final issue discussed by the Taskforce concerns the terminology: the benchmark study contains a list of definitions, which is useful, but some definitions triggered discussions. The Taskforce recommends to update and finalize the definitions based on a wide consensus in the Europeana Network before a new benchmarking study takes place.

---

<sup>1</sup> <https://www.force11.org/group/fairgroup/fairprinciples>

<sup>2</sup> <https://research.google.com/pubs/pub36299.html>

<sup>3</sup> <http://pro.europeana.eu/enumerate/>



*Conclusion on the methodology*

The Taskforce recommends that the benchmarking of Europeana is multi-faceted and should reflect the mission of Europeana as a whole ('a body of products and services') and not just the website of Europeana Collections. Different benchmarking activities will be needed to reflect the different aspects, set against a varying landscape of competitors per aspect. The different components of the benchmark should be based on indicators that directly reflect the mission of Europeana and that are measurable. The indicators should ideally not be conceived by ourselves but adopted from existing frameworks in the field.



## 5. Evaluation of specific elements in the report

In this part of the report we will look at the graphs from the *Benchmarking Europeana* report. During the Taskforce meeting in The Hague, the individual graphs were discussed briefly and rated against four criteria:

- Relevance: Does the graph provide useful information in the context of benchmarking Europeana?
- Objectivity: Can you exclude a potential bias in the way the data is interpreted?
- Validity: Does the indicator really measure what it is supposed to measure?
- Reliability: Is there a formal measurement protocol or framework in place that documents the indicator?

For each of these four criteria, the graphs were rated on a scale from 1 to 3, where 1=poor, 2=average and 3=good.

The overall judgement of the graphs isn't meant as a final word on them, but as input for the party that will prepare the next benchmarking study for Europeana.

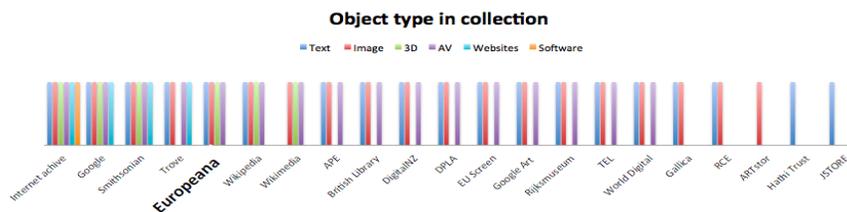


Fig.2: "the diversity of the type of objects made available"

This graph indicates the types of digital objects represented in all the services used in the benchmark. It is considered relevant to know this, but you would also need to know the amount of these objects as they occur: it makes a difference to know if there are 10 or 10.000 3D objects in a service. Also, the six objects types are very generic. It may be useful to use a more specific breakdown of types of collections, e.g. as provided in the ENUMERATE framework.

Taskforce Score:

- Relevance: 3
- Objectivity: 2
- Validity: 1
- Reliability: 1

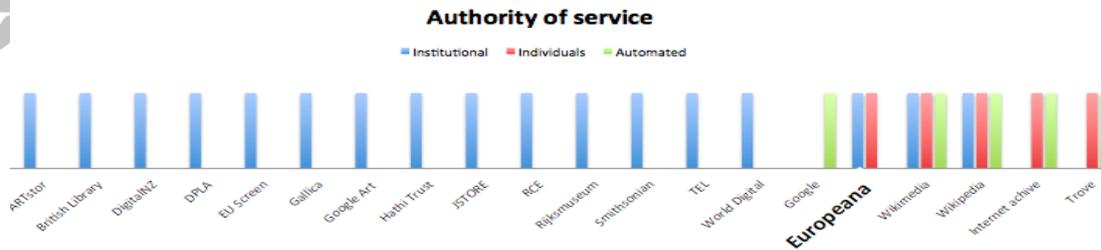


Fig.3: "authority of the content provided"

This graph explains the origin of the content in the information service. It is indeed relevant to know whether the content is curated or not, as this contributes to the trust that people have for a specific service. The Taskforce recommends not to only look at the origin of the data, but also at the trustworthiness of the source, how that trustworthiness is guaranteed and what the quality of the data is.

Taskforce Score:

- Relevance: 2
- Objectivity: 1
- Validity: 2
- Reliability: 1

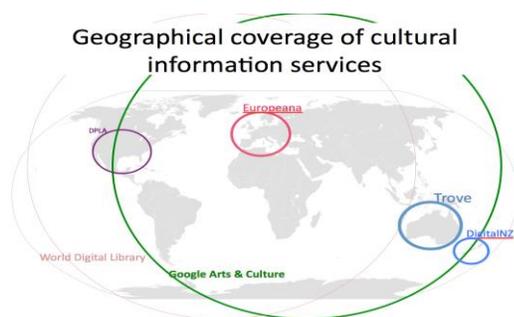


Fig.4: "covering all countries and several languages (...) where the thickness of the line represents the size of the collection."

This graph is not meant as a benchmark, but as a typology of the services compared in the benchmark. In the future, such typologies should be separated from those based on indicators meant to benchmark. Apart from the geographical coverage (in this case meaning: location of participating data providers) it would also be helpful to understand the linguistic coverage of the services compared (what languages are covered in the services?).

Taskforce Score:

- Relevance: 1
- Objectivity: -
- Validity: 1
- Reliability: -

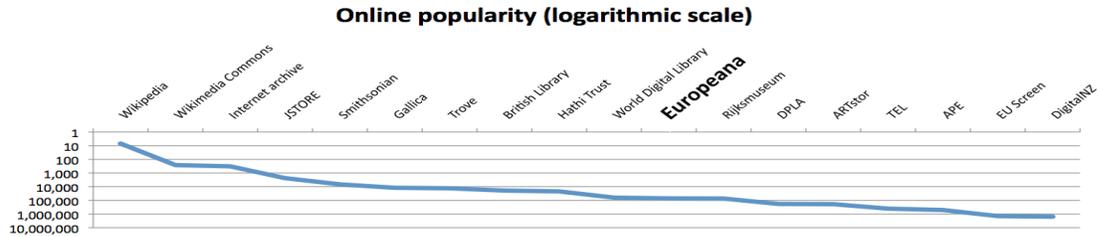


Fig.5: "online popularity, based on the Alexa ranking"

This graph expresses useful and relevant information, but the shape of the graph shouldn't be presented as a line. A table with numbers would be more appropriate than a logarithmic line graph, as it hides interesting differences. The Taskforce recommends to not only include Alexa rankings, but also use other tools like SimilarWeb.com. This is something to look into for the next benchmark.

Taskforce Score:

- Relevance: 3
- Objectivity: -
- Validity: 2
- Reliability: 3

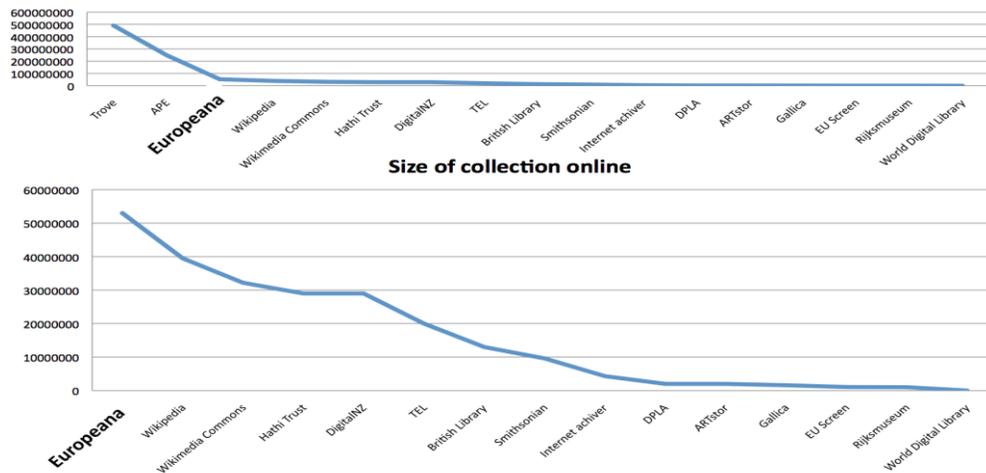


Fig.6: "a ranking was made by size of the collection made available"

This graph is useful in a benchmark, but, as was already mentioned in the report, there may be a bias as the way the size of the collection is measured (by the services themselves) is not normalized across the services.

Taskforce Score:

- Relevance: 3
- Objectivity: 2
- Validity: 2
- Reliability: 1

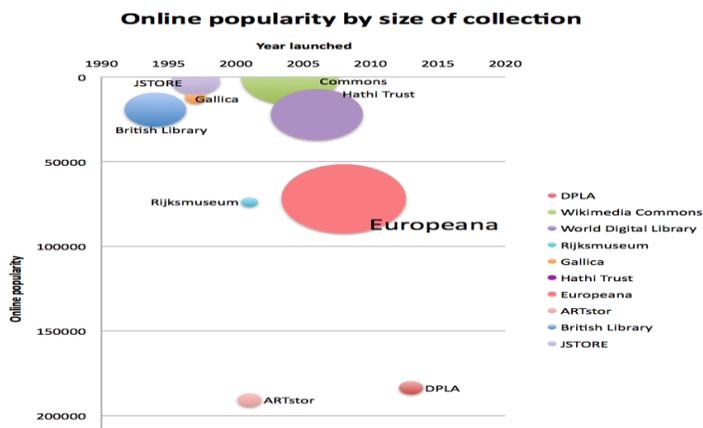


Fig.7: "the size of the collection, the online popularity and the year of the launch of the service [combined]"

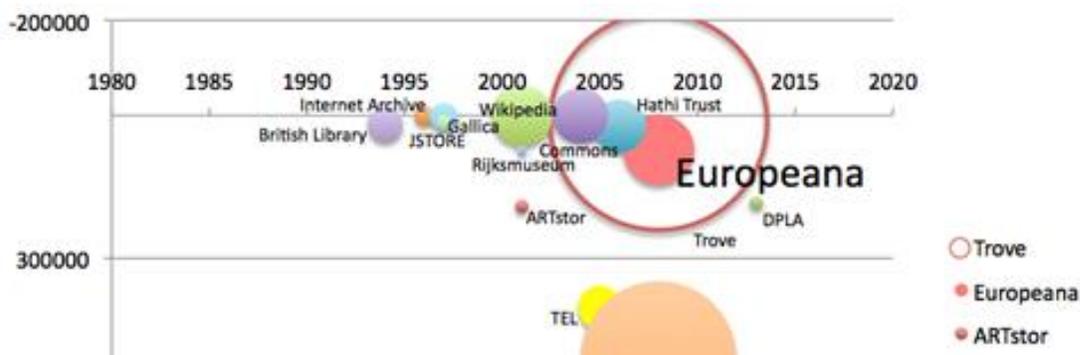


Fig.8: "the size of the collection, the online popularity and the year of the launch of the service [combined]" / "general view"

These two graphs are considered to be useful in a benchmark, but, as was already mentioned in the report, there may be a bias as the way the size of the collection is measured (by the service itself) is not normalized across the services. The combination of size of collection and popularity is interesting, as the graphs show that the services with the most content are not necessarily the most popular ones.

Taskforce Score:

- Relevance: 3
- Objectivity: ?
- Validity: ?
- Reliability: ?

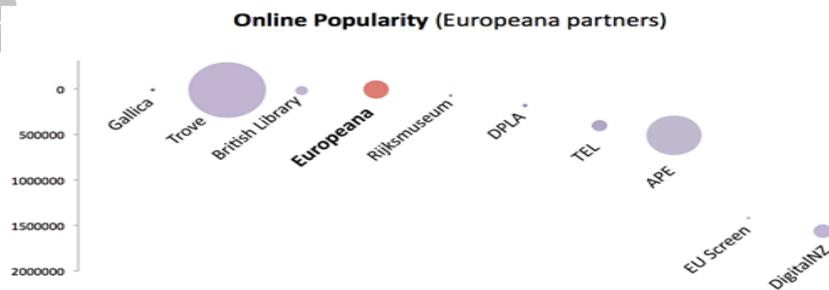


Fig.9: "popularity ranking by size of collection shows that Europeana ranks fourth among Europeana partners"

This graph would be more useful if it would be restricted to cultural information services that contribute data to Europeana. Now institutional websites (BL) are compared to Europeana as well, which is comparing apples and oranges.

Taskforce Score:

- Relevance: 3
- Objectivity: 3
- Validity: 2
- Reliability: ?

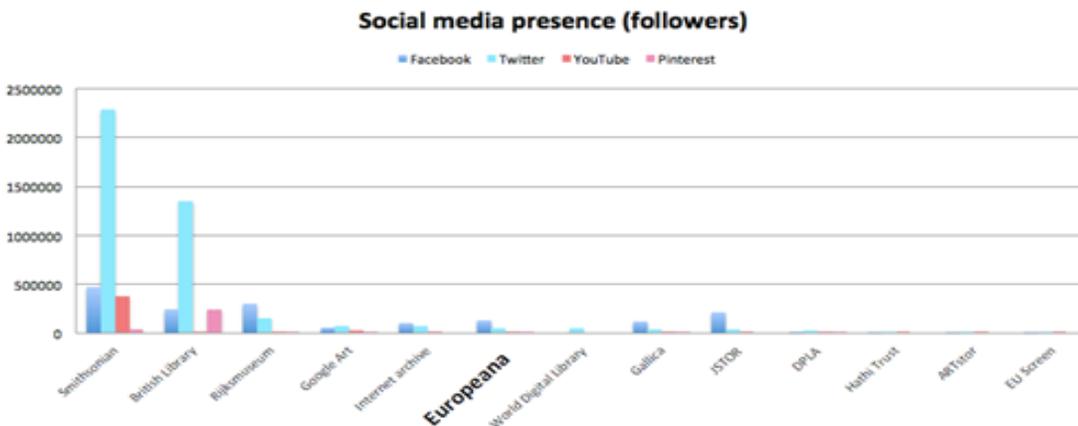


Fig.10: "number of social media followers"

The Taskforce likes the inclusion of data on outreach via social media in the benchmark, even though the use of the social media as a communication channel may occur without making use of digital collection items (but e.g. for marketing purposes instead). The Taskforce encourages to include other social media like Instagram, Pinterest and LinkedIn into the next benchmark study.

Taskforce Score:

- Relevance: 3
- Objectivity: 3
- Validity: 1
- Reliability: 3

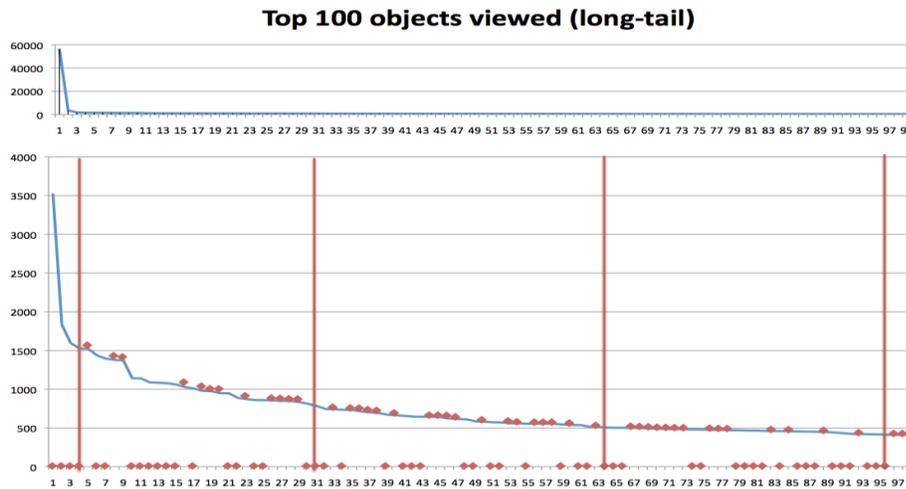


Fig.11: "100 most viewed objects in Europeana during 2015 (...) number of views per object" / "no image or preview (...) visible at the Europeana portal" [red diamonds] / "Europeana link to four objects is no longer available" [red vertical lines]

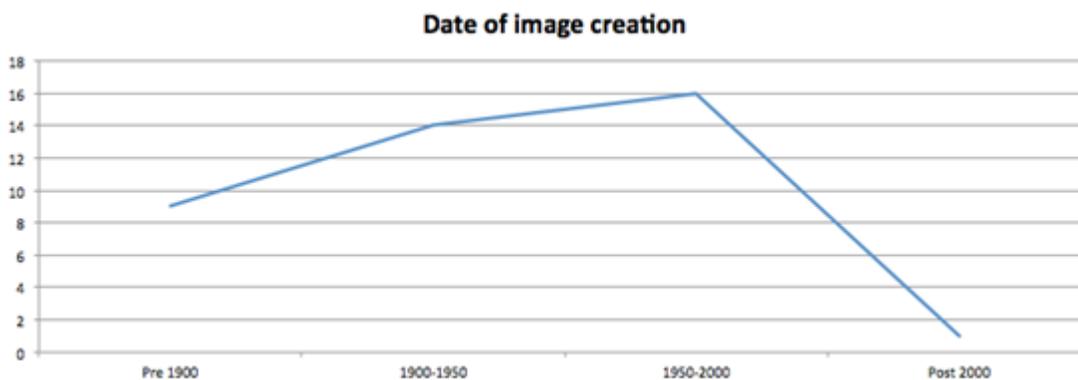


Fig.12: "average view of objects in Europeana (...) per date of creation [of the original object]"

Graphs 11 and 12 are interesting to know (especially as there are quite some obscure objects in the Top100), but this is statistics, not benchmarking. This kind of information should be available on Europeana Pro (e.g. via the dashboard). For the next benchmarking study, the Taskforce recommends to take the Top 10 (or 20) of the most viewed objects in Europeana and compare the rate of retrieval of these objects in Europeana with other services in which the same objects occur (e.g. Wikipedia, World Digital Library etc.). It would also be interesting to pay attention to the long tail: e.g. how much content in the service is not retrieved by end users during the course of a month or gets only 1 or 2 hits per year?

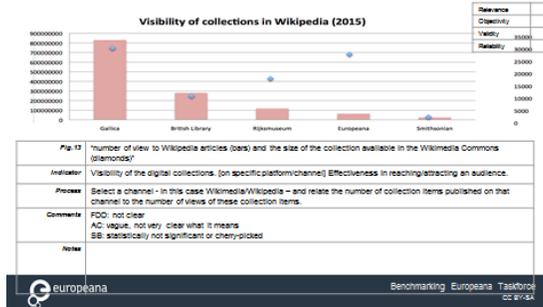


Fig.13: "number of views to Wikipedia articles (bars) and the size of the collection available in the Wikimedia Commons (diamonds)"

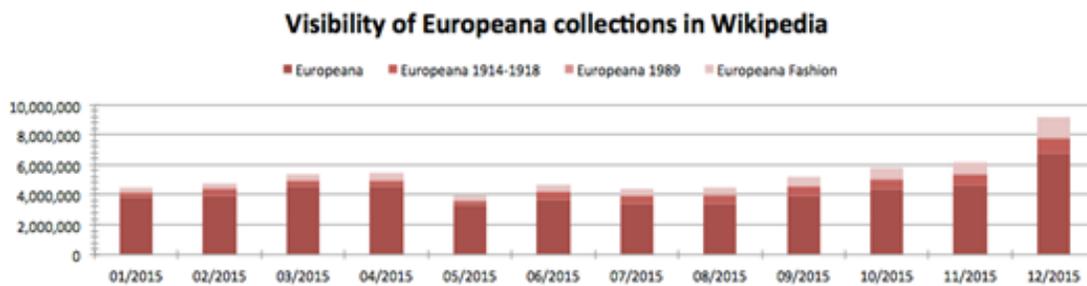


Fig.14: "use of collections from Europeana during a year (a growing use of objects in Wikipedia articles)"

Graphs 13 and 14 both provide useful information, e.g. in graph 13 that the size of the data collection in Wikimedia Commons does not correspond 1-on-1 with the volume of use of that data on Wikipedia. On the one hand, the Taskforce thinks it is useful to single out Wikipedia as a platform for dissemination of cultural heritage information, but, again, this is more statistics than benchmarking.

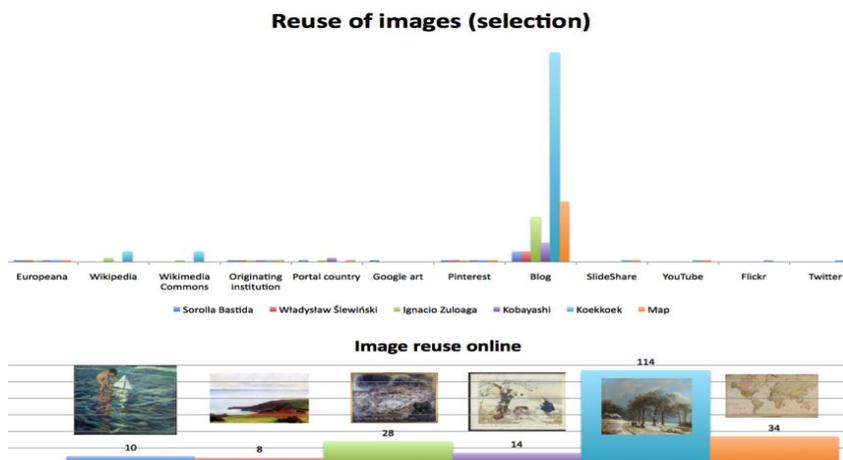


Fig.15: "top six most viewed paintings (with image) and a map reveals that blogs are the most important channels for further distribution"



The Taskforce thinks that it is clever to look at these kinds of re-use of popular objects from Europeana, but questions the reliability and the representativeness of such data. This needs further methodological refinement.

*Additional indicators*

In addition to these graphs from the report, the Taskforce brainstormed about topics that were not included in this benchmark study, but that may be included in the next one. The following suggestions were made:

- User interface comparison
- User satisfaction. This could be done with the following approach:
  - a. Run a survey asking users what are the most important characteristics of good cultural information service for them (e.g. size of collection, diversity, relevance of results, open licensing, quality of data etc.). This can be run very widely, not just on Europeana users, with some pre-defined characteristics and with possibility to add own ideas.
  - b. Analyse the results of the survey to extract the key service aspects that should be taken into account when evaluating/benchmarking a service.
  - c. Use these aspects in a second round of surveys to see how users of Europeana and other services see these services via the proposed set of aspects/characteristics, using a Likert scale.
- Spending & income: How much is spent and how much is earned per year?
- Data provider satisfaction: how easy/difficult is it for a cultural heritage institution as a data provider to contribute their data to a service?
- The volume (amount) of data that is available for free reuse.
- The number of APIs available and the use of these API's
- Record quality (both as standardized format and in richness of the content)



## 6. Evaluation of choice for benchmarked services

The study report used 19 institutions and cultural information services to benchmark Europeana with:

- APE
- ARTstor
- British Library
- DPLA
- DigitalNZ
- EUscreen
- Gallica
- Google
- Google Arts & Culture
- Hathi Trust
- Internet archive
- JSTOR
- Rijksmuseum
- TEL
- Trove
- Wikipedia
- Wikimedia Commons
- World Digital Library
- Smithsonian

In itself, this is a diverse and good list of institutions and services that have a relationship with Europeana. But as mentioned in the paragraph on methodology, the Taskforce suggests to make a longer list, including more information services (e.g. global services like IMDB, national services like ArtUK, commercial services like Amazon) and use that longlist to select per indicator a set of institutions, products or services that make the best fit for that indicator.

The Taskforce is aware of the fact that retrieving the relevant data from such other parties will remain a big challenge, but with more time given for a benchmark study, and by approaching individuals through the vast networks of memory institutions across Europe and Europeana itself, the Taskforce thinks that it is doable to obtain such data.



## 7. Conclusions and recommendations

The Taskforce appreciates the work done for the first benchmarking study and recognizes that a lot of important ground work was done in a very short time span.

The Taskforce also thinks that it is worthwhile to see the benchmarking of Europeana as a coordinated effort that should be done on a regular basis (e.g. every 18 or 24 months).

However, the Taskforce recommends to revise the methodology thoroughly, as the methodology used in the first benchmarking study was too much based on internal aspects of Europeana and could not come up with proper recommendations for improvement, based on lessons learned from comparing Europeana with other services.

The Taskforce invites the Members Council to reflect on the future options for benchmarking Europeana during their meeting of 6/7 July 2017 in Copenhagen. For the MC to do this, the Taskforce identified three scenarios:

1. to create a Working Group under the Europeana Members Council to continue the work of the Taskforce on a more regular basis. This Working Group would have the following responsibilities:
  - a. Decide on the new methodology and indicators
  - b. Document the indicators properly (see Appendix B)
  - c. Improve the definitions/terminology as used in the benchmarking study
  - d. Negotiate with Europeana office about funds for a new benchmark study as part of DSI-4
  - e. Act as a steering group for a new benchmark study
2. to incorporate the benchmarking of Europeana in a wider set of activities dealing with performance measurement, impact and statistics, including the Europeana Impact Framework, the Statistics Dashboard and the ENUMERATE Framework. The responsibilities as mentioned under 1 would then become part of a bigger set of activities. This would need consultation with the Europeana DSI-partners as well.
3. To discontinue the work on benchmarking Europeana.

We would appreciate it if the Members Council can spare some time to discuss this and decide on a preferred scenario.



## Appendix A: List of references

- Ayres, Marie-Louise, Trove and the world: cultural collections in a global environment. Paper for ALIA Online 2015. LINK (accessed 23 June 2017): [http://information-online.alia.org.au/sites/default/files/trove\\_and\\_the\\_world\\_-\\_cultural\\_collections\\_in\\_a\\_global\\_environment.doc](http://information-online.alia.org.au/sites/default/files/trove_and_the_world_-_cultural_collections_in_a_global_environment.doc)
- ENUMERATE, Survey Report on Digitisation in European Cultural Heritage Institutions 2015, June 2015 (Europeana/ENUMERATE). LINK (accessed 14/2/2017): [http://www.den.nl/art/uploads/files/Publicaties/ENUMERATE\\_Report\\_Core\\_Survey\\_3\\_2015.pdf](http://www.den.nl/art/uploads/files/Publicaties/ENUMERATE_Report_Core_Survey_3_2015.pdf)
- Europeana Dashboard (since April 2016). LINK (accessed 14/2/2017): <http://statistics.europeana.eu/>
- Europeana Publishing Framework (2015). LINK (accessed 14/2/2017): <http://pro.europeana.eu/publication/publishing-framework>
- Hughes, L.M. (ed.) Evaluating and Measuring the Value, Use and Impact of Digital Collections, London: facet publishing, London, facet publishing (2012). LINK (accessed 14/2/2017): [http://www.facetpublishing.co.uk/title.php?id=047203&category\\_code=](http://www.facetpublishing.co.uk/title.php?id=047203&category_code=)
- Navarrete, T. / DEN Foundation, *Europeana as online cultural information service*, Study Report (2016). LINK (accessed 14/2/2017): [http://pro.europeana.eu/files/Europeana\\_Professional/Publications/europeana-benchmark-report-sep-2016.pdf](http://pro.europeana.eu/files/Europeana_Professional/Publications/europeana-benchmark-report-sep-2016.pdf)
- Rodden, Kerry / Hilary Hutchinson / Xin Fu, *Measuring the User Experience on a Large Scale: User-Centered Metrics for Web Applications*. Proceedings of CHI 2010, ACM Press (2010) LINK (accessed 23 June 2017): <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/36299.pdf>
- Sherratt, Tim, From portals to platforms LINK (accessed 23 June 2017): <https://www.nla.gov.au/our-publications/staff-papers/from-portal-to-platform>
- Tanner, S., *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. King's College London, October 2012. LINK (accessed 24/2/2017): [https://www.kdl.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel\\_SimonTanner\\_October2012.pdf](https://www.kdl.kcl.ac.uk/fileadmin/documents/pubs/BalancedValueImpactModel_SimonTanner_October2012.pdf)
- Terras, M., "Opening Access to Collections: the Making and Using of Open Digitised Cultural Content, in: *Online Information Review* (2015). Special Issue on "Open Access: Redrawing the Landscape of Scholarly Communication" edited by G.E. Gorman and J. Rowley. LINK (accessed 14/2/2017): <http://www.emeraldinsight.com/doi/pdfplus/10.1108/OIR-06-2015-0193>



## Appendix B: Proposal for description of indicators

	<b>SPECIFICATION</b>
<b>Name of the indicator</b>	
<b>Theme (e.g. Costs, Use, Popularity)</b>	
<b>Ratio (why measure this?)</b>	
<b>For which stakeholder(s) is this indicator relevant?</b>	
<b>What is measured?</b>	
<b>How often should this be measured?</b>	
<b>References (e.g. if indicator is derived from existing framework)</b>	
<b>Remarks</b>	