

Involvement of OpenStreetMap in European H2020 Projects

Damien Graux^{1,*} and Thibaud Michel²

¹ Inria, Université Côte d'Azur, CNRS, I3S, Sophia Antipolis, France; damien.graux@inria.fr

² Wemap, Montpellier, France; thibaud@getwemap.com

* Author to whom correspondence should be addressed.

This abstract was accepted to the Academic Track of the State of the Map 2021 Conference after peer-review.

Since 1984, the European Commission has been supporting research through various successive programmes. Recently, from 2014 to 2020, the EU invested approximately 80 billion euros into its eighth programme, named Horizon 2020 [1]. Among various focuses such as the excellence of science or industrial secondments, H2020 emphasised supporting an open access policy for all research results [2]. Moreover, H2020 projects were strongly encouraged to use open source software and tools.

Practically, all research domains were eligible to be supported by the H2020 programme, and therefore, the scopes of the projects vary from e.g. computer science, to philology passing by agriculture. Technically, as these projects are almost always involving several partners located in several European member states joining forces from multiple institutions, there is often a need to deal with data coming from different places. And, more generally, geo-data are often involved to tag information which may be research data, meeting localisation, partner addresses, etc.

In such a context where open source tools and open access databases are recommended by the European Commission, we analyse the presence of OpenStreetMap (OSM) in H2020 projects. In addition, we also review the presence of other geographic services such as Google, Bing and Baidu maps, in order to better understand how researchers tend to choose one over the other. Furthermore, while OSM is a database, several services are created around it, including the basemap at openstreetmap.org, our study thereby reviews the use of OSM as an ecosystem of services.

Thanks to the open access policy, participants of H2020 projects had to make their results available. To do so, various types of materials were submitted to the European portal which then offers them publicly. As a consequence, for each project, one can access the articles (through DOIs), the blog posts, the slide decks, or also the deliverables. In particular, in our study, we decided to focus on the deliverables as they are accessible on the EC portal directly [3, 4] and as they are the common reports written by the partners to describe their approaches. Indeed, these deliverables (usually written on a regular basis during the course of the project) report on the findings and methodology set up to achieve the project's goals and the authors explain their architectural choices in depth such as describing the tools

Graux, D. & Michel, T. (2021). Involvement of OpenStreetMap in European H2020 Projects

In: Minghini, M., Ludwing, C., Anderson, J., Mooney, P., Grinberger, A.Y., (Eds.). Proceedings of the Academic Track at the State of the Map 2021. Online Conference, July 09-11 2021. Available at <https://zenodo.org/communities/sotm-2021>

DOI: [doi will appear here as a hyperlink]



© 2021 by the authors. Available under the terms of the Creative Commons Attribution ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)) license.

used. As a consequence, cartographic services, if involved at some stage in the project, are likely to be mentioned in these documents either as acronyms (e.g. OSM) or as website references (e.g. <https://www.openstreetmap.org/>).

In order to obtain the deliverables together with projects' information, we combined two European sources of information to gather all of the facets we wanted to cover: CORDIS [3] and Data.Europa [4]. In particular, we extracted from CORDIS various high-level information about the projects themselves: from their names and acronyms to their durations passing by the specific European call-for-fundings they answered and obtained their money from. This latter category can be useful in order to have a finer-grained understanding of the domains which are prone to involved cartographic services. Next in order, Data.Europa was used to download the deliverables themselves, which required several days of computing resources.

Overall, during the course of the H2020 programme, 33636 projects were funded by the European Commission. Depending on the type of action which was set by the projects, not all of them had some open deliverables written (and thereby available on the Europa platform). Actually, a large part of these projects did not have deliverables per se but rather articles or web posts. We indeed counted 25157 projects without deliverables which restricted our study to the remaining 8479 projects. Out of them, we listed a total of 92612 distinct deliverables to be analysed, representing more than 260GB.

Technically, once all of these deliverables were downloaded, we searched them for various terms to know if some cartographic services are involved in the text. We therefore set up several regex rules (e.g. 'open.?street.?map' or '[^a-z0-9]osm[^a-z0-9]') which were run over the 92000+ deliverables. This allowed us to systematically count all the occurrences of the considered cartographic solutions. In the end, we found that 1840 deliverables (from 651 projects) mention OpenStreetMap. More precisely (see Figure 1), through all the H2020 deliverables, there are approximately: 18600 mentions to OSM, 2800 to GoogleMaps, 226 BingMaps and 4 to BaiduMaps. Empirically, we notice that 1) one order of magnitude separates the occurrences of each cartographic service and 2) OpenStreetMap is, by far, the most represented solution and thereby the one on which public European researchers rely the most. Contextually, it is also interesting to note that not all the deliverables (1796 of them) mentioning "point of interest" refer to a cartographic service.

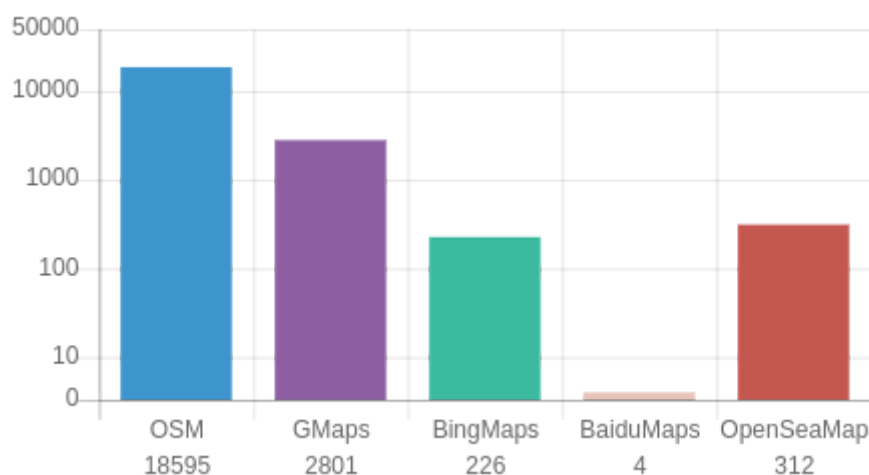


Figure 1. Number of occurrences over all H2020 open deliverables (logarithmic y-axis).

Moreover, we also analysed the co-occurrence cases, where different cartographic providers are jointly mentioned within a single deliverable. Notably, there are not that many. Indeed, only 59 deliverables mention both OSM and BingMaps, over the 226 occurrences of the latter; and only 291 deliverables mention both OSM and GoogleMaps, over the 2800 occurrences of GMaps. Besides, only 39 deliverables mention OSM, GoogleMaps and BingMaps. Such figures tend to suggest that once a group of researchers has chosen a cartographic solution, they tend to stick to it and do not try to compare them.

Furthermore, regarding OpenSeaMap, we counted 312 mentions from 27 deliverables, among which 20 ones mention both OSM and OpenSeaMap, showing how connected the two initiatives are.

In this study, we systematically analysed all the available H2020 deliverables, searching for cartographic service references, with a specific focus on OpenStreetMap. Our efforts show that OSM is the most used cartographic service in European H2020 projects in terms of mentions in the deliverable's texts, followed by GoogleMaps with one order of magnitude less mentions. It is worth noting that these projects involving OSM were backed by almost 4 billion euros of public money.

Based on these first interesting results, we plan to extend our scope of analysis following three axes. First, we think that it could be worth reviewing the other types of project's results such as the articles or the software source code bases. Second, we hope our approach paves the road to similar reviews of public funded initiatives, and based on this observation we plan to apply our scripts to other European funding programmes. Third, additional cartographic services could also be integrated into our pipelines such as ApplePlans or other OSM-related initiatives like OpenCycleMap in order to extend the covered scope.

Finally, for reproducibility purposes, we also share on a public github repository [5] all the necessary scripts to download the deliverables and generate the statistics. Furthermore, <https://dgraux.github.io/OSM-in-H2020> provides the reader with additional and detailed analyses together with visualisations, hoping these will help the community better understand the impact of OSM within the public European research landscape.

References

- [1] Horizon 2020. (n.d.). Horizon 2020 - European Commission. Retrieved April 25, 2021, from <https://ec.europa.eu/programmes/horizon2020/>
- [2] European Commission. (n.d.). Open access - H2020 Online Manual. Ec.Europa.Eu. Retrieved April 25, 2021, from https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm
- [3] The European Commission. (n.d.). CORDIS | European Commission. CORDIS, the primary source of results from EU-funded projects since 1990. Retrieved May 2, 2021, from <https://cordis.europa.eu/projects/en>
- [4] The European Commission. (n.d.-b). The official portal for European data. Publications Office of the EU. Retrieved May 2, 2021, from <https://data.europa.eu/en>
- [5] Graux, D., & Michel, T. (n.d.). OSM-in-H2020. GitHub. Retrieved May 2, 2021, from <https://github.com/dgraux/OSM-in-H2020>