

EDITORIAL

Open Access

Geospatial openness: from software to standards & data



Marco Minghini^{1*}, Amin Mobasher², Victoria Rautenbach³ and Maria Antonia Brovelli⁴

Abstract

This paper is the editorial of the Special Issue “Open Source Geospatial Software”, which features 10 published papers. The editorial introduces the concept of openness and, within the geospatial context, declines it into the three main components of software, data and standards. According to this classification, the papers published in the Special Issue are briefly summarized and a future research agenda in the open geospatial domain is finally outlined.

Introduction

The Open Definition qualifies knowledge as open “if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness” [1]. This definition, derived from the Open Source Definition [2] which is specific to software, puts forward the idea of a common good based on access to information that everyone can benefit from. This vision of a so-called ‘Open world’ is brilliantly invoked by Pollock, who calls for an open revolution – where everyone would have the freedom to use, enjoy and build on everything – in a digital age where most information is ‘closed’ and controlled [3]. There is no doubt that the geospatial domain holds the potential to play an important role in achieving this open digital revolution, given that an increasing number of georeferenced data is produced everyday – e.g. from sensors, including citizen sensors [4] – and geospatial technology has become mainstream in a plethora of domains. This Special Issue is exactly dedicated to geospatial openness declined in all its components, especially software, data and standards.

Open source software in the geospatial domain has a long and successful tradition. In the 1980s (almost the prehistory of GIS) the first versions of GRASS GIS (<https://grass.osgeo.org>, [5]) were released by the US

Army Construction Engineering Research Laboratories, while in the 1990s the history of Web Mapping began with MapServer (<https://mapserver.org>), originally developed by the University of Minnesota. The success of both projects as well as the communities they were able to attract explain the important role they still play today in the open source geospatial arena. The establishment in 2006 of the Open Source Geospatial Foundation (OSGeo, <https://www.osgeo.org>), dedicated to the development and support of open source geospatial technologies, was the key milestone to certify the maturity of open geospatial software and the related community. Today, open source geospatial technology is regularly used by governments, businesses, professionals and academics in a broad range of high-level applications [6–13]. However, open source software alone would have never succeeded without other parallel movements in the same direction of geospatial openness. Open geospatial standards are intrinsically linked to the activities of the Open Geospatial Consortium (OGC, <https://www.opengeospatial.org>), started at the end of last century to ensure geospatial interoperability and still leading standardization activities with more than 500 member organizations and tens of new standards released or updated every year. Similarly, open geospatial data represents the counterpart of open source geospatial software when moving from the domain of technology to the domain of data. Over the last few decades there has been an increasing awareness of the value of data and actions to ‘open the data’ (especially publicly funded data) have multiplied, resulting into open data policies included in political agendas at different levels, see e.g. [14–16]. In

* Correspondence: marco.minghini@ec.europa.eu

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

¹European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy
Full list of author information is available at the end of the article

addition, the amount of (geospatial) data produced by citizen-driven initiatives is usually distributed under open licenses and represent valuable input for several applications. The most popular of such initiatives in the geospatial domain, OpenStreetMap, is discussed in Section 4.

The Special Issue features 10 papers (all accessible at <https://www.springeropen.com/collections/osgs>), which are grouped in the following three sections according to the category they primarily belong to: open geospatial software, standards and data.

Open source geospatial software

Several papers published in this Special Issue describe specific open source technology in the geospatial domain. This section presents three of these papers. The software tools described in the first and second papers are general-purpose products, targeting users from (potentially) different domains and specifically developed to implement or make use of one or more OGC standards; in contrast, the third paper describes applications of a geospatial algorithm specific to the hydrology domain.

Geospatial datasets are typically made available to end users through standard web services, which are usually not easy to discover using traditional search engines. To address this issue, the Hypermap Registry developed by Corti et al. [17] and released under the MIT open source license, is a standards-based geospatial registry and search platform that harvests and manages a catalogue of OGC Web Map Services (WMS). The platform consists of various components, such as an OGC Catalogue Service for the Web (CSW), a database, a search engine (queryable through a REST API) and a map caching engine. An instance of the Hypermap Registry, known as HHypermap, is maintained by the Centre for Geographic Analysis of Harvard University and is used in conjunction with their WorldMap application. HHypermap has harvested about 120,000 layers from 15,000 services and regularly checks whether these services and layers are available. Hypermap Registry holds the potential to become an addition to a geoportal to accomplish the needs of Spatial Data Infrastructures (SDIs). Additional development is planned to extend its functionality to include, for example, support for OGC Web Feature Service (WFS) and to be included as a native application in the open source GeoNode platform.

The paper by Moreira de Sousa et al. [18] presents the latest developments of PyWPS, an open source implementation of the OGC Web Processing Service (WPS) standard, which allows the remote execution of geospatial processes. Developed since 2006, PyWPS is extensively used in research projects to integrate, publish and execute geospatial processes on the web. Written in Python 3 and available under the MIT license, since 2018

it has also become an official OSGeo project after completing the incubation process. The paper describes the latest PyWPS version (4.0), implementing WPS 1.0.0 and developed by over a dozen individual contributors during a period of almost 3 years. In addition to the new advanced features of PyWPS 4 (logging using Object-Relational Mapping, scalability using WSGI, containerisation), authors provide a comparison against other WPS server-side implementations and a set of real-world use cases where PyWPS is used. Finally, the paper offers a glimpse of the OSGeo way to ensure sustainability of its projects. This is achieved through the collaboration of different people and organizations with a common interest in the success of the incubation of the proposed product, being the positive incubation a proof of its reliability and maturity.

The improvements in sensor and satellite technology currently make it possible to collect high-resolution datasets in real-time. Along with the increased size and amount of data, the need for real-time processing has become crucial. One example in the hydrological field is the delineation of watersheds. Currently, many desktop GIS applications allow users to perform watershed analysis but there is a lack of libraries optimized for client-side and server-side web applications. Sit et al. [19] developed an optimized watershed delineation algorithm for web applications that was tested and benchmarked using various languages on both the server-side (Python, Go, C, Node.js) and the client-side (JavaScript, WebGL, WebAssembly). The software package is open source under the GPL. To optimize the delineation algorithm, WebGL was incorporated to allow using the Graphics Processing Unit (GPU). The benchmarking results show that the server-side implementation in C performs best, but on the client-side JavaScript becomes significantly more effective when WebAssembly is employed. In addition to independence from third-party software, the tool allows decision makers and stakeholders to make informed decisions through web-based analyses that exploit real-time information.

Open geospatial standards

Section 2 above has already demonstrated the strong link between open source geospatial software and open standards for interoperability of geospatial data and services, particularly those from the OGC. This Special Issue also includes two papers specifically describing advances in open geospatial standards. They both focus on the OGC CityGML standard, which is an open data model and XML-based format for the storage and exchange of semantic 3D city models.

Due to the generic nature of CityGML, which is intended to be a universal and application-independent geographic information model, there may be a need to

extend it to support certain applications. For this purpose, CityGML provides the Application Domain Extension (ADE), a built-in mechanism to augment its data model with additional information (new feature classes and attributes) required by particular use cases, such as energy studies, noise simulation, modelling utility networks, cultural heritage or property administration. The paper by Biljecki et al. [20] provides a technical overview of the ADE mechanism as well as a literature review of more than a hundred research papers on ADE developments since its introduction a decade ago. The review found out that there are currently 44 ADEs, either supporting specific applications or built for application-agnostic purposes such as CityGML adaptation to national geographic information standards. An overview of software support for CityGML ADE is also provided, so that overall the paper represents a useful reference for ADE developers.

In the second paper focused on the same domain of 3D GIS, Labetski et al. [21] propose a metadata ADE for CityGML. While the concept of metadata does not sound thrilling to most, it is a crucial sidekick to discover and access any type of data, including 3D city models. The CityGML standard does not offer a mechanism to store metadata in a structured way. Given the size and complexity of most CityGML files, having access to metadata would assist users in quickly understanding the nature of a dataset and its fitness-for-purpose. The study provides a metadata extension for CityGML, that includes categories compliant with ISO 19115 (the standard for geospatial metadata) as well as additional metadata elements that are specific to 3D city models (e.g. textures, materials and process of model acquisition). In the future this ADE might be incorporated as a core module of the CityGML standard. Finally, to further assist users, the authors offer open source software to automatically generate metadata for CityGML 3D city models.

Open geospatial data & OpenStreetMap

Geospatial software would not exist without geospatial data. And many experts in the geospatial field would agree that the data source which is most strongly associated with the concept of open data is OpenStreetMap (OSM, [22]). Initiated in 2004, OSM is an open-access, crowdsourced geospatial database having a global extent and including any real-world feature that is verifiable on the ground [23]. Given its open license, richness, level of detail and proven overall quality, OSM is currently used in a multitude of real-world and scientific applications as well as deeply studied by the academic community [24]. Since they both develop from the same idea of 'openness', the communities of open source geospatial software and open geospatial data (and in particular

OSM) feature a high level of intersection. Thus, it is not surprising that most of the software specifically developed to work with OSM data is also made available under open source licenses.

Half of the papers published in the Special Issue are relevant to OSM. Four of them describe software applications aimed to extract useful information from the OSM database through analysis, processing, visualization or a combination of them. The papers by Minghini and Frassinelli [25] and Raifer et al. [26] specifically address the analysis of OSM history, which represents a peculiar component of the OSM database itself.

Minghini and Frassinelli [25] offer two significant contributions. First, the authors provide an up-to-date, comprehensive review of available software applications focused on OSM history. These are classified according to the main way OSM data is used (visualization, statistics, analysis or conversion), the type of application (web page, web app, framework or tool), the source of OSM history data (OSM API, Overpass API or Full History Planet File) and the software license. The latter shows that only 6 out of the 34 software implementations analysed are not open source. The second contribution of the paper is the software "Is OSM up-to-date?", a web application accessible at [27] which offers ad hoc quality visualizations of OSM nodes and ways according to different history-based parameters. The main target users are OSM researchers, contributors and data users who can achieve a visual, intuitive idea of the intrinsic quality of the data. The application, mainly written in Python and licensed under the AGPL, can be also used through the command line and run inside a Docker container.

The paper by Raifer et al. [26] presents the OpenStreetMap History Database (OSHDB), a framework for spatio-temporal analysis of OSM history data. Unlike "Is OSM up-to-date?", which offers insights on the history of single OSM objects, OSHDB allows to perform efficient spatio-temporal analyses of OSM data on a global scale thanks to fast data access as well as flexible analysis methods. The tool is of special interest to researchers, since it provides a user-friendly and scalable way to create large scale visualizations, assess intrinsic data quality and investigate user contributions in specific areas and/or time frames. Available at [28] and licensed under the LGPL, the software offers a flexible Application Programming Interface (API) usable for a broad range of applications. It allows to extract, filter, aggregate and process OSM history data and extract outputs both on the status of the data at specified points in time and their evolution in specified time intervals.

The investigation of OSM data quality through the development of software infrastructure is also addressed in the paper by Mocnik et al. [29], who, however, do not only focus on the data itself but analyse also the OSM

community-driven processes behind it. The paper presents an open source infrastructure to collect and process data about different aspects of OSM. The merits and limitations of this infrastructure are discussed in details with exemplary use cases. In one of these, the authors analyse OSM data and generate insights about the OSM community and its mapping behaviour by a visual exploration of the datasets. Furthermore, the infrastructure focuses on producing datasets that can be easily stored and archived. While this allows to perform analyses in a simple and reproducible way, the authors argue that extensive adaptations of the data mining process to the particular use case are partially possible. They conclude that the proposed infrastructure is able to ensure two out of several factors involved in the concept of reproducibility: public availability of the data, the software, as well as workflows.

Patriarca et al. [30] explore the specific use of OSM as the source to generate Land Use/Land Cover (LULC) maps. These maps are fundamental for several applications (including planning, natural resources management, risk management and disaster response) but their production is typically a long and expensive process. Thus, the authors propose an open source tool, available under the GPL at [31] and in turn re-using other open source technology, to automatically convert OSM data into LULC maps with the nomenclature of official products (e.g. Corine Land Cover). Four versions of the algorithm, differing for the LULC data structure (vector or raster) and the underlying software used, were implemented and tested on different study areas characterized by heterogeneous size and volume of OSM data. Results are compared in terms of the outputs (LULC maps) generated and the performance, outlining the pros and cons of the different versions and documenting strengths and weaknesses of the open source software used. A jupyter notebook is also provided to ensure full reproducibility of results.

The last paper presents instead a case study on the humanitarian value of the OSM database and the need to educate people in contributing map data to address natural disasters and humanitarian crises. The case study described in [32] focuses on Nepal, which is amongst the 20 most disaster-prone countries in the world. In 2015, when Nepal was affected by a high magnitude earthquake, the power and usefulness of OSM data collected for disaster response was evident. However, while large urban centres such as Kathmandu and Pokhara were extensively mapped, smaller cities and rural Nepal remained unmapped. To address these data gaps, the Kathmandu Living Labs (KLL) started the Digital Internship and Leadership (DIAL) programme, which recruited young Nepalese people and trained them to not only collect OSM data, but also build capacity in skills

relevant to the digital age. The paper provides a detailed description of the DIAL activities and highlights its overall success, both from the mapping perspective – the interns performed about 2 million map edits, resulting in the creation of basemaps for large vulnerable areas in Nepal – and the educational perspective.

Conclusions

With the 10 papers described above, this Special Issue offers a glimpse into the current status of open geospatial solutions, differentiating them into software, standards and data but at the same time highlighting the intrinsic interconnections existing between them and their communities. As a matter of fact, the geospatial arena has evolved into a rich and complex ecosystem where more and more (types of) data become available, new standards are created to allow the use of such data in order to satisfy emerging user needs, and software improves to ensure that standards are implemented and data can be successfully manipulated. Openness of these solutions makes the whole process more transparent, agile, and overall efficient.

Future research in the geospatial domain should consider a number of new trends and related needs. First, the widely popular statement that spatial data is special [33] is currently more and more refuted. Among the reasons, the fact that geospatial data management has become mainstream and that spatial data often represents only a portion of the amount of digital information that is available and needs to be used to solve today's challenges. In addition, emerging technologies such as the Internet of Things (IoT) and APIs are shifting the focus from the nature of data to the way that data can be actually used, i.e. – once again – on interoperability. From the standardization perspective, the geospatial domain has already embraced this changing scenario, as demonstrated by the new OGC API family of standards (largely still to be published) that will modernize and revolutionize the way geospatial information is searched, accessed and retrieved on the web [34]. From the data and software perspectives, only in recent years a significant emphasis has been put on the need for efficient data management and stewardship through the promotion of so-called Findability, Accessibility, Interoperability and Reusability (FAIR) guiding principles. They apply to any digital object (not just 'data' in the traditional sense, but also to the algorithms, tools and workflows that lead to that data) and ultimately allow to maximise the benefits of data publication investments [35]. This brings to the closely-related issue of the reproducibility and replicability of scientific analyses and results, which is a cornerstone of science itself but has not yet been systematically addressed. Only recently a number of scientific journals have begun to include specific

requirements for the open publication, together with the research papers, of the data, software (mainly using containers) and code (mainly using web-based Git repositories) used therein. Other journals have been created which are specifically dedicated to research data and/or software, one of the best examples being the “Open Geospatial Data, Software and Standards” journal where this Special Issue is published.

Abbreviations

API: Application Programming Interface; CSW: Catalogue Service for the Web; DIAL: Digital Internship and Leadership; FAIR: Findability, Accessibility, Interoperability and Reusability; KLL: Kathmandu Living Labs; LULC: Land Use/Land Cover; OGC: Open Geospatial Consortium; OSGeo: Open Source Geospatial Foundation; OSHDB: OpenStreetMap History Database; OSM: OpenStreetMap; SDI: Spatial Data Infrastructure; WFS: Web Feature Service; WMS: Web Map Service; WPS: Web Processing Service

Acknowledgements

The authors would like to thank the Editors-in-Chief of the journal “Open Geospatial Data, Software and Standards” for supporting and promoting the publication of this Special Issue.

Authors' contributions

MM wrote Section 1; MM, VR and MB wrote Section 2; MM and AM wrote Section 3; MM, AM, VR and MB wrote Section 4; MM wrote Section 5. All authors read and approved the final manuscript.

Competing interests

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission. The authors declare that they have no competing interests.

Author details

¹European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy. ²Institute of Geography, Heidelberg University, Im Neuenheimer Feld 348, 69120 Heidelberg, Germany. ³Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria 0028, South Africa. ⁴Department of Civil and Environmental Engineering, Politecnico di Milano, 20133 Milan, Italy.

Published online: 30 January 2020

References

1. Open Definition 2.1. <http://opendefinition.org/od/2.1/en>. Accessed 9 Dec 2019.
2. Open Source Initiative. The Open Source Definition. 2007. <https://opensource.org/osd>. Accessed 9 Dec 2019.
3. Pollock R. The open revolution: A/E/T Press; 2018. <https://openrevolution.net>. Accessed 9 Dec 2019.
4. Foody G, Fritz S, Fonte CC, Bastin L, Olteanu-Raimond A-M, Mooney P, See L, Antoniou V, Liu H-Y, Minghini M, Vateva R. Mapping and the citizen sensor. In: Foody G, See L, Fritz S, Mooney P, Olteanu-Raimond A-M, Fonte CC, Antoniou V, editors. Mapping and the citizen sensor. London: Ubiquity Press; 2017. p. 1–12.
5. Neteler M, Mitasova H. Open source GIS: a GRASS GIS approach. Cham: Springer; 2013.
6. Brovelli MA, Minghini M, Moreno-Sanchez R, Oliveira R. Free and open source software for geospatial applications (FOSS4G) to support future earth. *Int J Digit Earth*. 2017;10(4):386–404.
7. Moreno-Sanchez R. Free and open source software for geospatial applications (FOSS4G): a mature alternative in the geospatial technologies arena. *Trans GIS*. 2012;16(2):81–8.
8. Minghini M, Kotsev A, Lutz M. Comparing INSPIRE and OpenStreetMap data: how to make the most out of the two worlds. *Int Arch Photogrammetry Remote Sensing Spatial Inform Sci*. 2019;42(4/W14):167–74.
9. Minghini M, Brovelli MA, Vandenbroucke D, Carbonaro M, Prüller S, Painho M, Martirano G, Frigne D. FOSS4G as a key building block for case-based learning in geographic information education. *Int Soc Photogrammetry Remote Sensing*. 2017;42(4/W2):129–35.

10. Bocher E, Neteler M. Geospatial free and open source software in the 21st century. Cham: Springer; 2010.
11. Steiniger S, Hunter AJ. The 2012 free and open source GIS software map—a guide to facilitate research, development, and adoption. *Comput Environ Urban Syst*. 2013;39:136–50.
12. European Commission. Open source strategy: history. 2019. https://ec.europa.eu/info/open-source-strategy-history_en. Accessed 9 Dec 2019.
13. Eom KS, Arias R, Brovelli MA, Criloux G, Kang HK, Li KJ. United Nations open GIS initiative: the first year of activities. *GEAM*. 2017;151(2):5–8.
14. European Parliament and European Council. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast). *Official J Eur Union*. 2019; 172:56–83 <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024>. Accessed 9 Dec 2019.
15. Australian Government. Australian government public data policy statement. 2015. https://www.pmc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf. Accessed 9 Dec 2019.
16. Indian Ministry of Science and Technology. National Data Sharing and Accessibility Policy (NDSAP): The Gazette of India; 2012. http://dst.gov.in/sites/default/files/nsdi_gazette_0.pdf. Accessed 9 Dec 2019.
17. Corti C, Lewis B, Kralidis AT. Hypermap registry: an open source, standards-based geospatial registry and search platform. *Open Geospatial Data Softw Stand*. 2018;3(8):1–12.
18. de Sousa LM, de Jesus JM, Čepický J, Kralidis AT, Huard D, Ehbrecht C, Barreto S, Eberle J. PyWPS: overview, new features in version 4 and existing implementations. *Open Geospatial Data Softw Stand*. 2019;4(13):1–10.
19. Sit M, Sermet Y, Demir I. Optimized watershed delineation library for server-side and client-side web applications. *Open Geospatial Data Softw Stand*. 2019;4(8):1–10.
20. Biljecki F, Kumar K, Nagel C. CityGML application domain extension (ADE): overview of developments. *Open Geospatial Data Softw Stand*. 2018;3(13):1–17.
21. Labetski A, Kumar K, Ledoux H, Stoter J. A metadata ADE for CityGML. *Open Geospatial Data Softw Stand*. 2018;3(16):1–16.
22. OpenStreetMap Wiki - Main Page. https://wiki.openstreetmap.org/wiki/Main_Page. Accessed 9 Dec 2019.
23. OpenStreetMap Wiki - Verifiability. <https://wiki.openstreetmap.org/wiki/Verifiability>. Accessed 9 Dec 2019.
24. Mooney P, Minghini M. A review of OpenStreetMap data. In: Foody G, See L, Fritz S, Mooney P, Olteanu-Raimond A-M, Fonte CC, Antoniou V, editors. Mapping and the citizen sensor. London: Ubiquity Press; 2017. p. 37–59.
25. Minghini M, Frassinelli F. OpenStreetMap history for intrinsic quality assessment: Is OSM up-to-date? *Open Geospatial Data Softw Stand*. 2019;4(9):1–17.
26. Raifer M, Troilo R, Kowatsch F, Auer M, Loos L, Marx S, Przybill K, Fendrich S, Mocnik F-B, Zipf A. A framework for spatio-temporal analysis of OpenStreetMap history data. *Open Geospatial Data Softw Stand*. 2019;4(3):1–12.
27. Is OSM up-to-date? web frontend. <https://is-osm-uptodate.frafra.eu>.
28. OSHDB source code. <https://github.com/giscience/oshdb>.
29. Mocnik F-B, Mobasher A, Zipf A. Open source data mining infrastructure for exploring and analysing OpenStreetMap. *Open Geospatial Data Softw Stand*. 2018;3(7):1–15.
30. Patriarca J, Fonte CC, Estima J, de Almeida J-P, Cardoso A. Automatic conversion of OSM data into LULC maps: comparing FOSS4G based approaches towards an enhanced performance. *Open Geospatial Data Softw Stand*. 2019;4(11):1–19.
31. OSM2LULC source code. <https://github.com/jasp382/gasp/tree/master/gasp/osm2lulc>. Accessed 10 Dec 2019.
32. Kanal K, Budhathoki NR, Erbstein N. Filling OpenStreetMap data gaps in rural Nepal: a digital youth internship and leadership Programme. *Open Geospatial Data Softw Stand*. 2019;4(12):1–10.
33. Anselin L. What is special about spatial data? Alternative perspectives on spatial data analysis. 1989. <https://escholarship.org/uc/item/3ph5k0d4>. Accessed 9 Dec 2019.
34. Open Geospatial Consortium. OGC APIs and the evolution of OGC standards. 2019. <https://www.opengeospatial.org/blog/2996>. Accessed 9 Dec 2019.
35. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data*. 2016;3(160018):1–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.