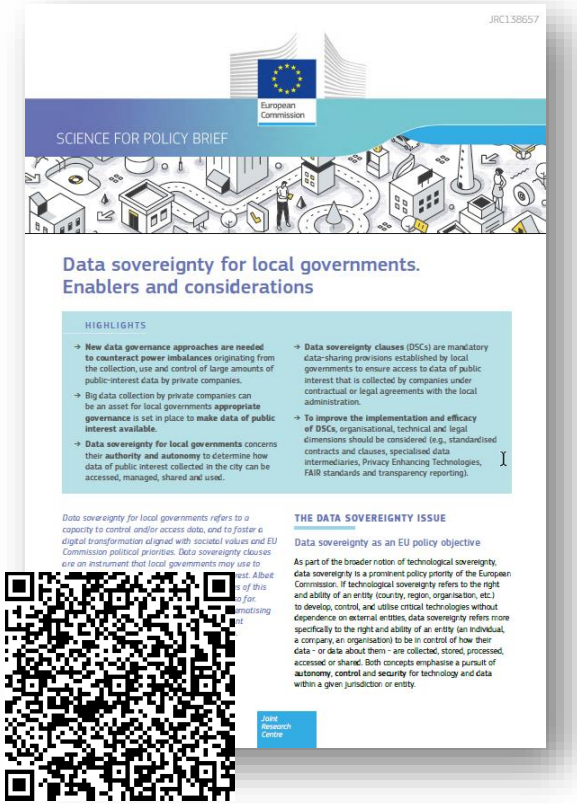
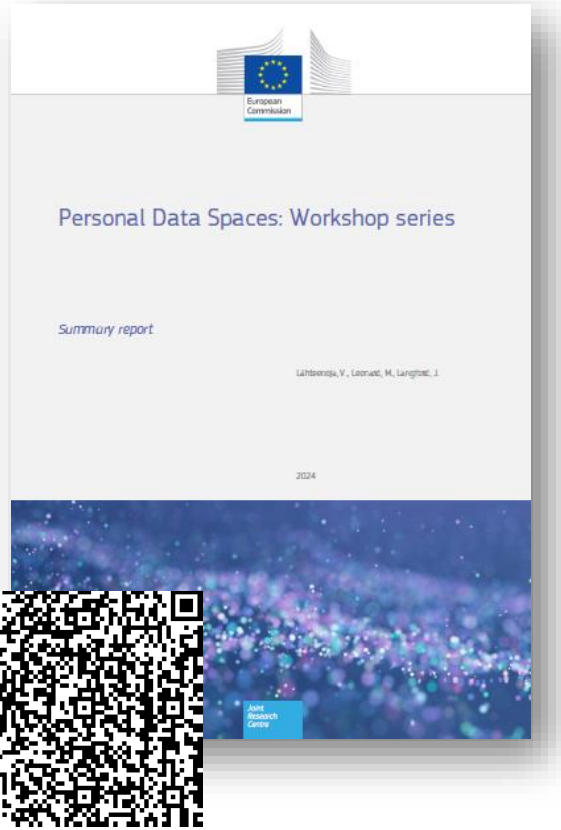
New JRC publications






The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-4/W12-2024
FOSS4G (Free and Open Source Software for Geospatial) Europe 2024 – Academic Track, 1–7 July 2024, Tartu, Estonia

Pan-European open building footprints: analysis and comparison in selected countries

Marco Minghini[1], Sara Thabit Gonzalez[1], Lorenzo Gabrielli[1]

[1] European Commission, Joint Research Centre (JRC), Ispra, Italy - (marco.minghini, sara.thabit-gonzalez, lorenzo.gabrielli)@ec.europa.eu

Keywords: Buildings, Open data, OpenStreetMap, Geoprocessing, GeoPython.

Abstract

This paper presents a comprehensive analysis of four non-governmental open building datasets available at the European Union (EU) level, namely OpenStreetMap (OSM), EUBUCCO, Digital Building Stock Model (DBSM) and Microsoft's Global ML Building Footprints (MS). The objective is to perform a geometrical comparison and identify similarities and differences between them, across five EU countries (Belgium, Denmark, Greece, Malta and Sweden) and various degrees of urbanisation from rural to urban. This is done in a two-step process: first, by comparing the total number and the total areas of building polygons for each dataset and country; second, by intersecting the building polygons and calculating the fraction of the area of each dataset represented by the intersection. Results highlight the influence of urbanisation on the dataset coverage (with increasing completeness when moving from rural to urban areas) and the varying degrees of overlap between the datasets based on a number of factors, including: the amount and up-to-dateness of the input sources used to produce the dataset; the presence of an active OSM community (for OSM and the datasets based on OSM); and the accuracy of Machine Learning algorithms for MS. Based on these findings, we provide insights into the strengths and limitations of each dataset and some recommendations on their use.

1. Introduction

The current data economy is characterised by a multitude of actors involved in the production of data. Compared to the past, when the public sector was the main societal player responsible for collecting, maintaining, updating and disseminating datasets, today's landscape is much more heterogeneous. Leveraging new technologies such as Artificial Intelligence (AI), Internet of Things (IoT) and cloud, private, research and citizen-led initiatives have become relevant producers of valuable geospatial data for several applications and use cases (Kotsev et al., 2021). In the European Union (EU), the European strategy for data (European Commission, 2020) emphasised the need to make sense of the huge amount of data produced by all societal actors through both technological and organisational measures, supported by legal interventions, with the ultimate goal to create a fair, trustworthy and interoperable data sharing environment known as the common European data space, which is currently in the making (Farrell et al., 2023).

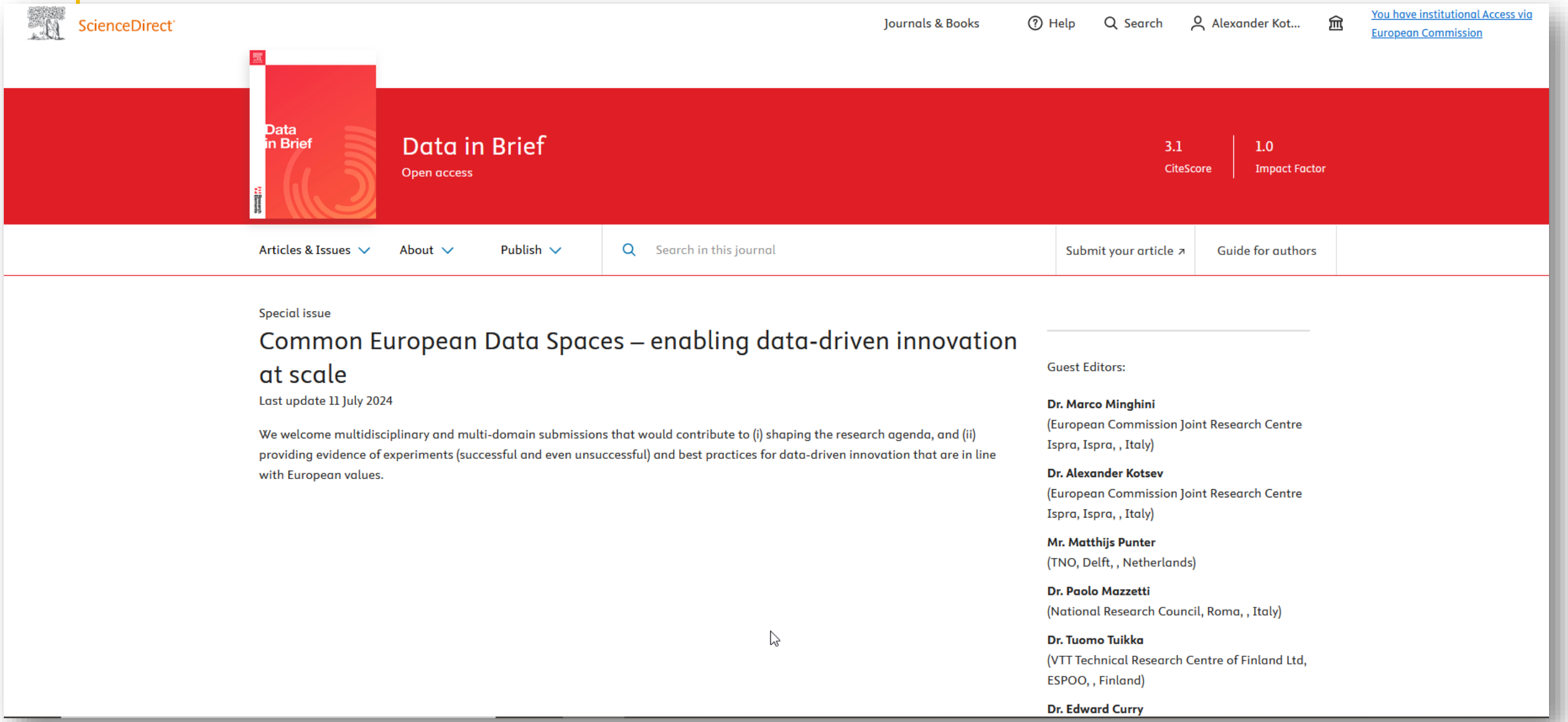
With this broad and complex context in mind, this work addresses the geospatial dimension of data sharing, which is horizontal across several societal domains, and zooms into specific geospatial datasets — building footprints (hereinafter simply referred to as buildings). These are fundamental datasets for several applications, including city planning, demographic analyses, modelling energy production and consumption, disaster preparedness and response, and digital twins. As key resources of Spatial Data Infrastructures (SDIs), buildings have been historically produced by governmental organisations — National Mapping or Cadastral Agencies — as part of their cartographic databases, with coverage ranging from local to national and licensing conditions being heterogeneous and not always open. This has typically made it challenging to derive open building datasets with a continental or global scale.

Over the last decade, however, the unparalleled developments in the resolution of satellite imagery, AI techniques and citizen engagement in geospatial data collection have enabled the production of several building datasets available at least at the continental scale under open licenses. A crucial role in this process was played by OpenStreetMap (OSM, https://www.openstreetmap.org), the most popular geospatial crowdsourcing project started in 2004 with the goal to create and maintain a database of the whole world licensed under the Open Data Commons Open Database License (ODbL, https://opendatacommons.org/licenses/odbl). With more than 2 million contributors active so far (OpenStreetMap Wiki, 2015), OSM has become the largest, most complete and most up-to-date geospatial database currently existing and its usage spans across multiple use cases and applications (Mooney and Minghini, 2017). OSM buildings are typically mapped through the digitisation of high-resolution satellite imagery, while in some cases they derive from the import of third-party datasets (e.g. released from governments) having an ODbL-compatible license. The quality of OSM buildings has been heavily studied in literature. While being heterogeneous across countries, regions and cities, OSM quality (mainly measured as positional accuracy and completeness) usually increases when moving from rural to urban areas, where it can equal or even outperform the quality of authoritative datasets (Hecht et al., 2013; Fan et al., 2014; Fram et al., 2015; Brovelli et al., 2016).

Nevertheless, concerns about OSM quality have stimulated the birth of multiple initiatives, led by either private or research actors, to create open building datasets using OSM as the foundation. Among the most promising initiatives led by the private sector is Overture Maps (https://overturemaps.org), founded in 2022 by four companies (Amazon, Meta, Microsoft and TomTom) with the mission to provide global, high-quality and interoperable open datasets from the combination of several input sources. Instead, open building datasets produced by research-led initiatives include EUBUCCO (Milojevic-Dupont et al., 2023; https://eubucco.com) and the Digital Building Stock Model (Florio et al., 2023; https://europa.eu/!W9YJqy), which combine OSM buildings with other sources to create

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-4-W12-2024-97-2024 | © Author(s) 2024. CC BY 4.0 License. 97

European Commission



https://www.sciencedirect.com/special-issue/10D1GF3RVVK

European Commission