Open-source tools and Open Science practices inspiring co-creation of geospatial data infrastructures – reflections from Finland and Tanzania

Niina Käyhkö[a*], Juha Oksanen[b], Antti Vasanen[c], Msilikale Msilanga[a], Masoud M. Hamad[a,d], Kylli Ek[e], Venla Aaltonen[a], Zakaria Ngereja[f]

[a] Department of Geography and Geology, University of Turku, Finland, 1st Author - niina.kayhko@utu.fi, 4th Author – msilikale.msilanga@utu.fi, 5th Author – masoud.hamad@utu.fi, 7th Author – venla.o.aaltonen@utu.fi
[b] Finnish Geospatial Research Institute, National Land Survey of Finland, Finland, 2nd Author - juha.oksanen@nls.fi
[c] The Regional Council of Southwest Finland, Finland, 3rd Author - antti.vasanen@varsinais-suomi.fi
[d] Department of Computer Science and IT, The State University of Zanzibar, Tanzania, 5th Author - massoud.hamad@suza.ac.tz
[e] CSC – IT Center for Science Ltd., Finland, 6th Author - kylli.ek@csc.fif
[f] Department of Geospatial Sciences and Technology, Ardhi University, Tanzania, 8th Author – ngereja@gmail.com

*Corresponding author

Keywords: geospatial data, open data, SDI, FOSS, citizen science

Abstract:

During the last ten years, we have been witnessing a rapid growth in both sources and quantities of digital data, open-source (FOSS) technologies and geospatial tools and services in the society. Through expansion of mobile technologies, new digital devices, Global Positioning Systems (GPS), and cloud infrastructures, societies and individuals have become more responsive to fluent access to up-to-date digital geospatial data. These rapid technological advancements and diverse data sources have provided a conducive environment for the co-creation of SDIs and different platforms and digital open spaces for data sharing, refinement and value-creation (Dangermond and Goodchild 2020).

Spatial data infrastructures (SDIs) of the 21st century have a substantial yet also a greater role in ensuring sustainable digital transformation. This is since our societies are increasingly depended on open-access, dynamic and replicable data ecosystems for improved decision-making. Today's geospatial data infrastructures operate in the web to meet the openness, access and community engagement demands and they also facilitate integration of diverse data sources effectively whilst ensuring data provenance for replicability (Kotsev et al. 2020).

In our presentation, we are comparing and presenting lessons learnt from the co-creation processes of four different digital geospatial data-driven research and learning infrastructures, which are currently operative in Finland and Tanzania. These spatial data infrastructures are: 1) Geoportti Research Infrastructure (Geoportti RI, https://www.geoportti.fi/), which is an open Geospatial Research and Education Hub for Finnish researchers and students, 2) Lounaistieto Regional SDI (https://www.lounaistieto.fi/), which is an open data service and community in Southwest Finland, 3) Climate Risk Database (CRD, https://geonode.resilienceacademy.ac.tz/), which is an open geospatial data platform and community-curated geospatial data service hosted by the Tanzania Resilience Academy, and 4) UTU Geospatial Data Service (https://geonode.utu.fi/), which is an open geospatial data sharing platform of the University of Turku, dedicated for the researchers and students implementing Open Science agenda.

Geoportti Research Infrastructure (Geoportti RI) provides Finnish scientists with an easy access to geospatial data and geocomputing resources at centralized high-performance computing (HPC) and cloud infrastructure at the IT Center for Science (CSC). Geoportti RI's objective is to encourage widespread utilization of geographical data and geoinformatics methods in both research and university education. Furthermore, it seeks to stimulate innovation in geospatial data and digital technology on a broader societal scale.

The Geoportti research infrastructure has been co-created through 20 years of cooperation between key geospatial research and academic institutions in Finland. During the last few years, more emphasis has been put on research competence, innovation, and excellence development, alongside a significant upgrade of a number of the existing services. Currently, Geoportti RI consists of seven categories of services (Figure 1): 1) Geodata and metadata services, 2) Research data and source code storing and sharing services, 3) Emerging technology laboratories and tools services, 4) High-performance computing (HPC) geoprocessing and geoanalytics services, 5) Virtualization services, 6) Skills

development and innovation services, and 7) Community services. The services are either fully established and maintained by Geoportti RI, established and maintained by Geoportti RI to bring significant added value to third-party services, such as national/global SDI services or CSC, and third-party services that are not produced, but promoted and instructed by Geoportti RI in the geographical information science context.

Figure 1. The current service portfolio of the Geoportti Research Infrastructure (2022-24).

Lounaistieto Regional SDI is a collaborative space of regional geospatial experts in South-west Finland with a co-creation history of over 20 years between key public open data actors in the region. Currently, Lounaistieto is a broad open data ecosystem with a map service, statistical analysis service, and wide range of community activities. Lounaistieto services are directed at authorities and decision-makers, researchers, residents, organizations, the media and business.

Lounaistieto SDI's transitioning from a data resource for local geographical information specialists to a comprehensive geospatial service for public information has evolved through five adaptable phases in response to evolving local demands and technological trends. As a result, the Lounaistieto-SDI has boosted the region's geospatial skills, extending advantages to academia, public entities, and businesses (Kalliola et al. 2019).

Tanzania Resilience Academy's Climate Risk Database (CRD) is a digital geospatial data repository that supports research, education, and disaster-risk management practices targeted for Tanzania cities. CRD contains around 200 open-access geospatial data sets, which have been shared by multiple users engaged in the Tanzania Urban Resilience Program. CRD is a platform maintain by the participants of the Tanzania Resilience Academy and the service is primarily for university students, staff and volunteers to curate geospatial data and learn data-related skills via volunteered participation to the community events, such as Data Visualisation Challenge. Over 1500 of students in Tanzania have benefited from the learning and innovation activities of CRD.

UTU Geospatial Data Service provides an open repository for geospatial research data sharing. The platform enables the University of Turku researchers and students discover, download, upload and share their own or research group's geospatial data sets via the service and create maps and get inspired by reading stories from the community. UTU Geospatial data service is part of the University of Turku data infrastructure, and it is also used in learning data curation skills via organising Data Party events and using the platform in undergraduate education.

All these infrastructures have been co-created by joint activities of several actors, who have engaged themselves in the planning, co-creation and operative activities of the infrastructure. They are operational through wide use of FOSS technologies. Unlike national policy-driven SDIs, the main motivation behind establishment of these services has been a desire to enhance open access and diversity of opportunities related to reusing and refining digital data for learning, research and innovation. Our aim is to reflect confronted challenges and successes of these open data services through the lenses of SDIs as cooperative digital data and knowledge spaces of the 21st century (see Kotsev et al. 2020), and provide lessons learnt.

Acknowledgements

References

Dangermond, J. and Goodchild, M.F., 2020. Building geospatial infrastructure. Geo-spatial Information Science 23, Vol. 1, pp. 1-9.

Kalliola, R., Käyhkö, N. and Jokela, S., 2019. Adaptive Development of a Regional Spatial Data Infrastructure Facing Local Prospects and Socio-Technological Trends. Bulletin of Geography, Socio-economic Series 44, pp. 73–80.

Kotsev, A., Minghini, M., Tomas, R., Cetl, V. and Lutz, M. 2020. From Spatial Data Infrastructures to Data Spaces—A Technological Perspective on the Evolution of European SDIs. ISPRS Int. J. Geo-Inf. 9, 176.