CARSTEN KEẞLER

Department of Planning
Aalborg University
Copenhagen, Denmark
Email: kessler@plan.aau.dk

Carsten Keßler is an Associate Professor for Geoinformatics in the Department of Planning at Aalborg University Copenhagen, Denmark. Before he moved to Copenhagen in 2016, he was an Assistant Professor for Geographic Information Science at Hunter College, City University of New York (CUNY) and Associate Director of the Center for Advanced Research of Spatial Information. He is still an adjunct faculty member in the Earth and Environmental Sciences program at the CUNY Graduate Center. Until fall 2013, he was a post-doc researcher in theSemantic Interoperability Lab at Institute for Geoinformatics (ifgi), University of Münster, Germany. Keßler also finished his PhD on *Context-aware Semanticsbased Information Retrieval* at ifgi in 2010 and has worked as a consultant for the United Nations Office for the Coordination of Humanitarian Affairs in Geneva, Switzerland.

Keßler has published on a range of topics, including geospatial semantics, linked data, context modelling, volunteered geographic information, and location privacy. He was co-chair of the Linked Science workshop series and of the W3C Emergency Information Community Group. His current research focuses on the use of geographic information science methods and spatial simulation techniques to analyse future impacts of global warming under different climate scenarios, particularly heat stress and migration. On the latter topic, he is currently involved in a transdisciplinary project on Global Flows of Migrants and Their Impact on North European Welfare States (FLOW)[1].

http://carsten.io

From Spatial Data to Spatial Information

Not long after the proliferation of data science as a field in its own right—combining methods from disciplines such as statistics, data mining, machine learning, and data visualization—specialized sub-disciplines for application areas emerged, including biomedical, cultural, sports, health, and spatial data science. This division into different "kinds" of data science reminds of the 1990s establishment of different branches of informatics (bioinformatics, health informatics, geoinformatics, etc.), each providing a different angle on the *computational turn* in their respective disciplines. Research agendas were written for each of those new fields and revised over the years, as research made progress and new computational developments opened up new possibilities.

[1] See https://www.flow.aau.dk

In light of the existence of a number of research agendas for GIS, geographical information science, geoinformatics, cyberGIS, geocomputation, and locationbased services (e.g., Rhind, 1988; McMaster & Usery, 2004; Anselin, 2012; Huang et al., 2018), a research agenda for spatial data science would need to be distinctively different from any of those existing research agendas. The identification of the unique research challenges for spatial data science ultimately leads to the question how exactly – besides the limitation of *geo* to certain scales – spatial data science differs from any of those existing fields (and, at the same time, how they differ from each other). The answers to those questions will need an ongoing discussion within and between the respective communities, for which this symposium will provide a promising starting point. Nonetheless, there seem to be some obvious trends that set spatial data science aside from the established subdisciplines, including:

Integration of spatial methods into the regular data science workflow. Analysis of spatial data is becoming more and more common as part of the "regular" data science workflow. Data scientists use software libraries such as *sf* for R or *geopandas* and *PySAL* for Python within their normal working environment, without the requirement for a GIS or a spatial database. These developments promise the establishment of spatial analysis as one of the major pillars of data science, next to statistics, machine learning, etc. However, they also present a new research challenge concerning the usability of those tools, which are often used by data scientists with little training in the opportunities and peculiarities of spatial information and the corresponding analytical tools. These tools hence need to be made more accessible for a user group that is substantially different from traditional GIS users.

Open and Reproducible Science. Looking at academic research, we can already witness an increasing adoption of data science practices for the analysis and mapping of spatial data. The sharing of open source code on public repositories, as well as the development and use of open tutorials and documentation, often using notebooks, has increased substantially over the last years. Moreover, open source languages, particularly Python and R, have subsequently been equipped with impressive capabilities for spatial analysis. In combination, these developments lower the bar for truly reproducible research, whose adoption is currently initiated by various projects and initiatives, but still far from wide acceptance. The development of a more open culture is thus one of the challenges for our field arising from recent developments in (spatial) data science.

From spatial data to spatial information. Ultimately, the goal of spatial data science is still to turn spatial data into insightful, actionable information. The merging of spatial analysis tools into data science therefore opens up new opportunities and challenges for the field, allowing it to grow beyond its classical user base, but likewise demanding an orientation towards a new user group, which is very well trained on the computational side, but may need more guidance on the spatial aspects. If this challenge can be successfully addressed, the field is making a substantial step towards turning spatial into the new normal.

References

Anselin, L. (2012). From SpaceStat to CyberGIS: Twenty years of spatial data analysis software. *International Regional Science Review* 35(2): 131–157.

Huang, H., Gartner, G., Krisp, J. M., Raubal, M., & Van de Weghe, N. (2018). Location based services: ongoing evolution and research agenda. *Journal of Location Based Services* 12(2): 63–93.

McMaster, R. B., & Usery, E. L. (2004). *A research agenda for geographic information science*. CRC Press.

Rhind, D. (1988). A GIS research agenda. *International Journal of Geographical information systems* 2(1): 23–28.