Towards Structured Description of Geospatial Relationships
Panel proposal for NKOS Workshop 2021

Panel

Olha Buchel, New England Complex System Institute, Boston, MA, USA, olha@necsi.edu

Leila Hedayatifar, New England Complex System Institute, Boston, MA, USA, leila@necsi.edu

Pouya Zangeneh, University of Toronto, Toronto, ON,Canada, p.zangene@gmail.com

Alfredo J. Morales-Guzman, Redzone Production Systems, Miami, FL, USA, alfredo.mo.gu@gmail.com

ABSTRACT

Semantic relationships play a crucial role in facilitating knowledge discovery in Linked Open Data. When we talk about geospatial relationships in libraries, we immediately think of gazetteers and thesauri as key Knowledge Organization Systems for geospatial entities. Both of these tools have a limited set of relationships such as is-A, isPartOf, has Parts, and RelatedTo. Hill (2009) also described relationships between spatial footprints (i.e., equals, disjoint, touch, overlaps, cross, within, contains, minimum distance, is contained within) that can enhance knowledge discovery in geospatial Knowledge Organization Systems. However, relationships between footprints were never included in gazetteers because they could be computed from geospatial locations. There was no need to describe and store them in databases.

Recent developments in data science and research on ontologies suggest that the well-established approach to description of geospatial relationships should be revisited (Janowicz, & Keßler, 2008; Zangeneh, & McCabe, 2020; Baclawski et al., 2021). New data-science-driven approaches suggest that we may not only facilitate better information retrieval by adding explicit semantic relationships to geospatial Knowledge Organization Systems, but we can also facilitate the discovery of patterns in geospatial data at multiple scales that can be useful for the development of public policies. This is particularly true about datasets that are not static, and the data are being updated dynamically.

This session will outline role the traditional and nontraditional ways of geospatial relationship descriptions play in the data science today; the construction of knowledge graphs, and analysis of knowledge graphs as complex systems at multiple scales that yield valuable patterns, insights, and results in production of visual artifacts for policy makers in public health (Buchel, et al., 2021), city management (Dong et al., 2020), and industrial management (Zangeneh, & McCabe, 2020).

We will also address the challenges faced by data scientists while working with the open data, and private datasets. The problems are related to data licenses, relationships modeling for further analysis, privacy protection, data provenance, quality, archiving, transformation, sharing, anonymization, aggregation, and bias to name a few. We will discuss standards, specifications,

and best practice guidelines. We plan to stir discussion of priorities in development and implementation of standards for generating better insights for public health, cities, and industries.

Keywords

Geospatial relationships, data standards, research data, open data, data insights, knowledge discovery, knowledge graphs, complex systems.

The Panelists

Short biographies are listed below. The presenters have been playing important roles in the development of peer-reviewed research on public policies, geospatial ontologies, data science, and information visualization. They are engaged in the development of data products (visualizations and analysis) for communities, national, and international institutions.

Olha Buchel is a Postdoctoral Researcher at the New England Complex Systems Institute (NECSI). She holds a PhD in Library and Information Science with specialization in Information Visualization from Western University in Canada. She has extensive working experience with geospatial data and geospatial visualizations. She currently works on data science projects, using machine learning and artificial intelligence algorithms in visualizations. She is also responsible for data collection, data pre-processing and transformation. She develops visualizations for public health, business analytics, and social media analytics. Her visualizations facilitate reasoning and decision making about complex systems. During COVID-19 pandemic, she worked on NSF-funded project related to mobility patterns and optimization of quarantine policies to reduce contagion.

Leila Hedayatifar is a Senior Postdoctoral Researcher at NECSI. She holds a PhD in Physics. She is interested in artificial intelligence and machine learning techniques that are applicable in social and economic systems. A major focus of her current research is understanding the structure of societies and changes in this structure over time, based on differing social and geographic properties and using statistical analysis and agent-based models. She is the PI in a NSF-funded project to optimize the quarantine policies using the mobility patterns of individuals and the severity of COVID-19 contagion in different areas.

Pouya Zangeneh is completing PhD in Engineering at the University of Toronto. His thesis title is "Knowledge Representation and Artificial Intelligence for Management of Socio-Technical Risks in Megaprojects". The project is aimed at capturing the probabilistic dependencies of project variables with an intelligent decision making model and a knowledge base containing

expert inputs and historic data. Therefore, the project is incorporating several topics including "Knowledge Representation", "Graph Data Structures", and "Artificial Intelligence and Probabilistic Graphical Models" all in the context of application in management and forecasting of industrial megaprojects. The research is primarily focused on mining and energy sectors; having the hope of future integration with infrastructures.

Alfredo J. Morales-Guzman was an Assistant Professor at NECSI and Visiting Scholar at MIT Media Lab. He was named Innovator Under 35 by MIT Technology Review en Español. He is passionate in understanding complex, social behaviors by analyzing (big) data from social media and alternative sources, using networks, AI and machine learning. He is an expert in analysis of social dynamics across scales: from global societies down to cities and organizations, and transforming data into visualization, understanding and strategy. He currently works as a Principal Datascience Researcher at Redzone Production Systems in Miami.

References:

Baclawski, K., Bennett, M., Berg-Cross, G., Schneider, T., Sharma, R., Singer, J., & Sriram, R. D. (2021). Ontology summit 2020 communiqué: Knowledge graphs. *Applied Ontology*, (Preprint), 1-19.

Buchel, O., Hedayatifar, L., and Bar-Yam, Y. (2021). The Role of Mobility Patterns in the Formation of Bubbles During the COVID-19 Pandemic. In A. J. Morales, J. Norman and M. Bahrami (Eds.), *COVID-19: a complex systems approach*. Stem Academic Press. pp. 157-172.

Buchel, O., Ninkov, A., Cathel, D., Bar-Yam, Y., Hedayatifar, L.: Strategizing COVID-19 Lockdowns Using Mobility Patterns. arXiv preprint: https://arxiv.org/abs/2012.03284 (2020) (under review by the Royal Society).

Dong, X., Morales, A. J., Jahani, E., Moro, E., Lepri, B., Bozkaya, B., ... & Pentland, A. (2020). Segregated interactions in urban and online space. *EPJ Data Science*, *9*(1), 20.

Hedayatifar, L., Rigg, R. A., Bar-Yam, Y., Morales, A. J.: US social fragmentation at multiple scales. Journal of the Royal Society Interface, 16(159), 20190509 (2019).

Hedayatifar, L., Morales, A. J., & Bar-Yam, Y. (2020). Geographical fragmentation of the global network of Twitter communications. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(7), 073133.

Hill, L. L. (2009). *Georeferencing: The geographic associations of information*. Mit Press.

Janowicz, K., & Keßler, C. (2008). The role of ontology in improving gazetteer interaction. *International Journal of Geographical Information Science*, *22*(10), 1129-1157.

Zangeneh, P., & McCabe, B. (2020). Ontology-based knowledge representation for industrial megaprojects analytics using linked data and the semantic web. *Advanced Engineering Informatics*, *46*, 101164.