Towards Geographic Information Observatories

Krzysztof Janowicz[1], Benjamin Adams[2],
Grant McKenzie[1], and Tomi Kauppinen[3,4]

[1] University of California, Santa Barbara, USA
[2] The University of Auckland, New Zealand
[3] Aalto University School of Science, Finland
[4] University of Bremen, Germany

Abstract. In this work we outline the vision of geographic information observatories, explain how and why GIScience would benefit from broadening its range to studies of the information universe (without necessarily relating it back to the physical universe), how our technology stack puts the GIScience community in a unique position to develop such observatories, and illustrate how these observatories would enable to ask new questions.

Keywords: Information Universe, Information Observatories, Web Science, CyberGIS

1 Introduction

More than two decades ago Geographic Information Science was established as an interdisciplinary research field and since then has significantly contributed to the growth and analysis of spatial data from a wide range of sources such as satellites, mobile devices, and even citizens as sensors. Within the last few years, the resulting *geographic information universe* has entered a phase of exponential growth and today data is available at an ever increasing spatial, temporal, and thematic resolution. Surprisingly, however, the nature, structure, and evolution of this information universe is poorly understood.

Historically, a majority of GIScience research has focused on the relationships between information models and the geographic phenomena they represent. Thus, for instance, researchers have developed methods to compute density surfaces from point observations that can be used as heatmaps for crime analysis, wildlife tracking, and so forth. Interestingly, however, there is a lack of research that expands GIScience towards empirical work on the structure of the geographic information universe itself.

A better understanding of this universe would yield new insights into how geo-information can be utilized, what drives the diversification and diffusion of information, how and why information forms ecosystems, how technological changes will be reflected by the information universe, whether there are *laws* of information science, what spatiotemporal properties of memes are, and so forth. Such research would explicitly include both *observational* and *experimental* approaches to science.

In the following, we will argue for the need for *geographic information observatories* by motivating them in the historical development of GI research, point out existing trends, give examples, and outline challenges.

2 Historic Perspective and Motivation

GIScience was established to study the concepts and theories that underlie Geographic Information Systems and other related technologies and thereby allows to ask more general and foundational questions and to explore topics such as usability, uncertainty, knowledge representation, or even ethical questions using scientific methods instead of a purely engineering driven perspective [2]. Nonetheless, for many years GIScience remained a *support science* which investigates, develops, evaluates, and optimizes the methods, procedures, and tools required by scholars in the broader geosciences including geographers, ecologists, geologists and so forth.

With GIS and GIScience gaining visibility, the developed methods and tools were applied in other domains inducing archeology, economics, and health research. This led to a further broadening of the GIScience realm, namely the notion of *spatial sciences*. The key insight was that many scientific domains require expertise in spatial aspects or study spaces that are closely related to geographic space. This gave birth to *spatial centers* such as spatial@UCSB that serve as nexus for researchers interested in spatial questions. Examples for involved domains include cognitive and brain sciences, religious studies, chemistry, the digital humanities, and so forth. Nonetheless, one could argue that GIScience and spatial science remain supporting sciences that deliver services to other disciplines. In many geography departments, GIScience and related disciplines are still categorized under labels such as *methods*, *techniques*, *computation*, and so forth.[1] It is not widely recognized that GIScience is the information science perspective on geography and the geosciences. It is on the same level as physical and human geography. GIScience does not depend on other domains and branches (of geography) but is a (meta) science in its own right. This, of course, does **not** imply that GI researchers should not closely collaborate with their colleagues in physical and human geography, it just highlights that they have their own agenda as well.

Over the last years, there have been multiple attempts to overcome the narrow perspective outlined above. For instance, Skupin and Fabrikant have investigated *spatialization* methods for non-geographic information visualization [7]. One example for such work is the landscape of a music folksonomy derived by using a self-organizing map.[2] Similar to the case of spatial sciences, the key insight is that research methods from GIScience can be successfully applied to apparently non-spatial domains. The changing perspective on the nature and use of information can also be illustrated by recent work on the foundational

[1] The Annals of the Association of American Geographers, for instance, use *Methods, Models, and GIS* as section name.

[2] See http://cns.iu.edu/docs/research/10-Last.fm.pdf.

aspects of GI research. For example, Couclelis outlines an ontological perspectives on geographic information layers and their semantic contraction [1], while Kuhn discusses core concepts for spatial information [5]. Note that both explicitly focus on characterizing information, not physical reality. Kuhn even argues that the rationale for his selection of proposed core concepts are informed by the available of types of spatial data. To give a final example, the Oxford Internet Institute publishes so-called Internet and information geographies under the slogan *understanding life online*.

3 Geographic Information Observatories

The notion of a *Web Science* introduced by [4] calls for an interdisciplinary approach to study the Web as a large-scale cyber-social-system. Research areas for Web Science include the interaction between social dynamics, creativity, and technologies as enablers, dispersion of information on the Web, provenance, trust and credibility, cyber-crime, new forms of user interaction, motivational topics, the global knowledge graph, and so forth. Along similar lines, Sheth and others proposed *Physical-Cyber-Social Computing* as a holistic study of data and knowledge from physical, cyber, and social environments to provide contextually specific abstractions to human user [6]. To study this new science requires *Web Observatories* as "environment[s] that will enable the next generation of interdisciplinary Web Science research involving mixed methods at a global scale [that] empower researchers by providing a distributed, collaborative, scalable and sustainable online environment to share data, analytical methods and visualization tools to explore the socio-technical evolution of the Web." [8, p.1]

Unsurprisingly most existing realizations of such Web observatories make heavy use of *space and time as fundamental ordering principles* and thus rely on Web mapping, spatial analysis, etc., for their user interfaces and interactions. With respect to the used and envisioned technologies to integrate, display, analyze, and communicate data, those observatories share common ideas with CyberGIS, the Digital Earth, and spatial data infrastructures. So what would geographic information observatories look like, how would they differ from existing work, and why would they be of interest to the GIScience community?

Imagine an observatory that constantly monitors geographic information from different sources and perspectives, using different types of media, spanning multiple themes, attribute-dimensions, resolutions, etc. It would ingest data from a variety of sources including social media, governments, news portals, and Internet traffic statistics. Such an observatory would allow us to ask questions such as whether certain media types and contents show different dispersion patterns, how do new technologies such as *Google Glass* impact the availability of information and add new perspectives, how does the appearance of information items in relation to specific places forecast events, and whether there are *laws* of information science such as the relation between the amount of schema knowledge versus instance-level data [3].

Additionally, it could provide answers to such questions as how and why geo-data and technologies are socially embedded, what types of spatialization techniques are most suitable for non-geographic data, what features of unstructured information are most *geo-indicative*, what is the typical degree-of-separation between various categories of information items and information items about locations and places, how fast is geo-data growing in comparison with other data, how do privacy concerns relate to changes in the resolution of available geo-data, and so on. It is worth noting that a large-scale system such as the information universe has emergent properties that can neither be studied nor predicted by merely analyzing isolated technical and social effects [4]. Many of these highlighted questions are concerned with the nature of information and do not necessarily need to relate back to physical space. Instead, they study the nature of the densely connected information universe as well as *physical-cyber-social system* where humans, technologies, and information form more than the sum of its parts.

The envisioned GI observatories are more than just a rebranding of the Digital Earth or CyberGIS visions. In analogy to observatories in astronomy, oceanography, or volcanology, these GI observatories would form a network of interconnected installations that are open to researchers and the public, and that constantly monitor the information universe at a resolution that would be impossible to achieve for the individual researcher of citizen due to the required hardware and data processing capabilities. Similar to observatories in astronomy, they would have individual 'sensors' and specialize on different domains, i.e., they would monitor different types of signals. In case of interesting events multiple observatories could be targeted at specific regions of the information universe (that may correspond to regions on the surface of the earth). GI observatories would not merely monitor the information universe but would make the full range of spatial analysis and geo-visualization methods publicly available. Thanks to GIScience's active role in fields such as Spatial Data Infrastructures (SDI), CyberGIS, semantics, geo-visualization, spatialization, etc., our community is in a great position to lead such effort. To give a concrete example, multiple GI observatories could monitor social media, news, government data, traffic data, the flow of money, etc, during the 2014 Ukraine conflict and highlight different perspectives on the topic.

From an educational perspective, such observatories are not limited to Cloud-based environments but could actually be physical installations that citizens can visit. This would hopefully lead to a similar kind of visibility, fascination, and public support that we see in the field of astronomy that make people support heavy financial investments for the sake of expanding the frontiers of our knowledge and understanding. In the 21^{st} century, the information universe is equally important as the physical universe.

There are, of course, many interesting challenges that have to be address in order to develop and make use of such GI observatories. In addition to further work on data integration, big geo-data analysis, workflow composition, scalability, uncertainty, spatialization, and so forth, the complex interaction and dynam-

ics between technology, information, and society raises entirely new questions. One of these questions is geo-privacy. While it has been studied before, we are only now beginning to realize what it means that trash bins track our location and movement through cities[3], that geo-information is increasingly stored and not forgotten, and that the new spatial, temporal, and thematic resolutions of the data may change human society forever.

References

1. Couclelis, H.: Ontologies of geographic information. International Journal of Geographical Information Science 24(12), 1785–1809 (2010)
2. Goodchild, M.F.: Geographical information science. International Journal of Geographical Information Systems 6(1), 31–45 (1992)
3. van Harmelen, F.: Keynote: 10 years of semantic web research: Searching for universal patterns. In: The Semantic Web - ISWC 2011, pp. 389–389. Springer (2011)
4. Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., Weitzner, D.: Web science: an interdisciplinary approach to understanding the Web. Communications of the ACM 51(7), 60–69 (2008)
5. Kuhn, W.: Core concepts of spatial information for transdisciplinary research. International Journal of Geographical Information Science 26(12), 2267–2276 (2012)
6. Sheth, A., Anantharam, P., Henson, C.: Physical-cyber-social computing: An early 21st century approach. Intelligent Systems, IEEE 28(1), 78–82 (2013)
7. Skupin, A., Fabrikant, S.I.: Spatialization methods: a cartographic research agenda for non-geographic information visualization. Cartography and Geographic Information Science 30(2), 99–119 (2003)
8. Tiropanis, T., Hall, W., Shadbolt, N., De Roure, D., Contractor, N., Hendler, J.: The web science observatory. IEEE Intelligent Systems 28(2), 100–104 (2013)

[3] See, for instance, http://arstechnica.com/security/2013/08/no-this-isnt-a-scene-from-minority-report-this-trash-can-is-stalking-you/