PhD Topic 9: Exploiting Linked Open Data in Environmental Informatics

First Supervisor	Prof. Dr. A Min Tjoa
Second Supervisor	Prof. Dr. Andreas Frank
Primary Field	Information Semantics, Management and Integration
Secondary Field	Resource Management
Interdisciplinary References	A1, A2, A4

Today, humanity faces a number of significant environmental challenges, such as climate change, energy, pollution, etc., which force us to consider our impact on the natural world and to find more sustainable ways of living. Most such challenges are multidisciplinary in nature and require efficient sharing and exchange of information among many disciplines and knowledge domains in order to address those problems efficiently. In this context, environmental informatics as an interdisciplinary field aims to convert the descriptive and qualitative nature of environmental issues into quantitative and measureable indicators for environmental stakeholders and decision-makers.

In achieving this goal, environmental informatics profits from a number of sophisticated techniques, concepts, and methodologies in different research areas such as informatics, statistical analysis, architecture, urban planning, risk analysis, computation modeling, etc.

One important aspect of environmental solutions is the environmental data, which, as explained above, is usually scattered among multiple knowledge domains. Such data is often either trapped in customized data structures of different institutions or locked in domain-specific databases, making data reuse very difficult, if not impossible.

It is expected that the gap between required and available environmental information can be bridged by building on two main pillars: First, a set of generic and versatile ontologies that serve as a shared standard vocabulary of environmental concepts. Second, the scattered information resources on the Web should be mapped to these ontologies and linked with other data sources.

In this way, the scattered information on the Web can form a global data graph that connects distributed resources and facilitates the discovery of new resources. This approach, which follows a set of best practices for publishing and connecting structured data on the Web, is known as "Linked Data" and has gained momentum among governments, in the academic and

business worlds, and in the public sector over the last few years. Linked Data provides a publishing paradigm in which not only documents, but also data can be first-class citizens of the Web, thereby enabling the extension of the Web with a global data space based on open standards - the Web of Data. Depending on the published data formats and readiness of data providers, the Linked Open Data (LOD) may be modeled, published, and reused in different ways. In this regard, a five-star scheme has been introduced to score the quality of shared LOD.

A number of government initiatives are now supporting the idea of Open Data by publishing high-quality government data in order to facilitate data integration between different environmental stakeholders and government organizations.

The goal of this research is to study the potential of LOD for the specific case of environmental data and create a solid theoretical basis for publishing and sharing Linked Environmental Data based on the concepts of the Semantic Web and LOD. To this end, some novel approaches for capturing and managing the distribution of information should be developed and adapted to the specific needs of environmental problems. A number of challenges, such as interaction paradigms, schema mapping and data fusion, link maintenance, licensing, trust and privacy, and quality assurance should be addressed as well.

References

[1] K. Tochtermann, U. Germany, Knowledge Management and Environmental Informatics, Journal of Universal Computer Science, vol. 6, no. 5 (2000), 517-536.

[2] Heath, T., Hepp, M., and Bizer, C. (eds.). Special Issue on Linked Data, International Journal on Semantic Web and Information Systems (IJSWIS), http://linkeddata.org/docs/ijswis-special-issue

[3] Heath, T., Bizer, C. 2011. Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, ISBN: 9781608454303

[4] Berners-Lee, T., 2009. Linked data. Available at http://www.w3.org/DesignIssues/LinkedData.html