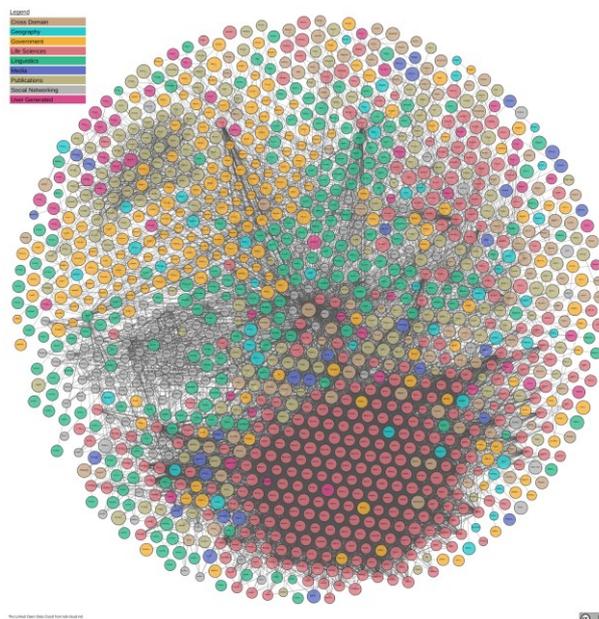# Spatial Linked (Open) Data

Instructors:

- Dr. Erwin Folmer (Kadaster and University of Twente) - erwin.folmer@kadaster.nl (coordinator)
- Stanislav Ronzhin (ITC, University of Twente, Kadaster) - s.ronzhin@utwente.nl
- Dr. Rob Lemmens (ITC, University of Twente) - r.l.g.lemmens@utwente.nl
- Dr. Wouter Beek (Triply, VU University, Kadaster) - wouter@triply.cc

Target audience: Graduate university students in the fields related to (business) information systems/technology and geomatics, staff from survey and mapping agencies; public authorities and interested parties involved in data publishing and consuming.

The Linked Open Data Cloud.

Introduction: Geospatial information has great value for the economy and society alike. Therefore, provision of access to European-wide geospatial data is the ambitious goal of the

Infrastructure for Spatial Information in the European Community (INSPIRE). For more than a decade after the INSPIRE directive came into force, the choice of technology to open up geospatial data resources held at the national level was directly related to the open standards developed by the Open Geospatial Consortium (OGC). Based on the eXtensible Mark-up Language (XML) technology and the Simple Feature Access model, the OGC standards were understandable and worked well for the community of GIS users.

In parallel, in the world of the World Wide Web, major developments by the W3C consortium have taken place resulting in the family of Linked Data standards. Linked Data (LD) technology is proposed as a remedy for solving interoperability problems at technical, organizational and community levels. As approach Linked Data is a promising concept for allowing spatial data to become part of the mainstream World Wide Web. But probably the biggest opportunity is that Linked Data can help with the de-silofication of spatial datasets. By providing the concepts, standards and technology for bridging between datasets, also from different domains, so that we can build powerful (spatial) Knowledge Graphs.

Course objective: After completing this course, the participants will be able to:
- understand the basic concepts of Linked Data
- explain how Linked Data can contribute to interoperability in the spatial domain
- explain the role of ontologies in conceptual modelling
- get insight in the business case of implementing Linked Data (at Kadaster)
- can perform a basic SPARQL query.

The audience will be welcome to come with their own data to discuss with the lecturers about the potential of having their own dataset available as Linked Data.

Course outline: This is an introductory course to Spatial Linked Open Data. Linked Open Data is a standards-based approach for data interoperability. In this course, we will teach the basic theory of Linked Data, and introduce the most important standards such as RDF. More in-depth the topic of data modelling, vocabularies and ontologies will be elaborated as one of the key concepts of Linked Data (module 2). Although the concepts and technology is generic and not specific for spatial data we will discuss in particular the context of spatial data on the web (module 1). The second part of this course is split in a technical module (module 4) and a business module (module 3). The business module will discuss the business case for linked data implementations based on the case study of the Dutch Kadaster, one of the earlier linked data implementations in Europe. The technical module will provide best practices on how to convert data into linked data, and will be practical hands-on creating SPARQL queries.
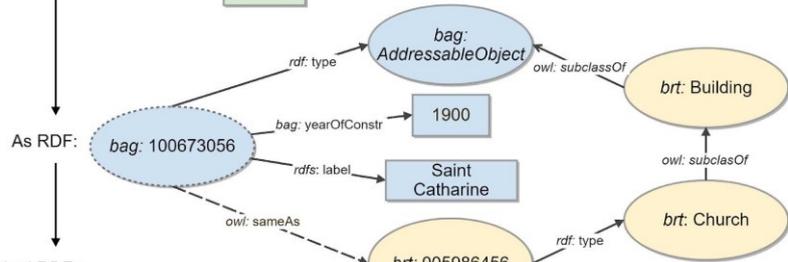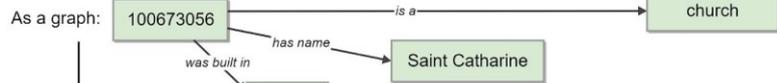
Representing facts about a building from a plain text using RDF. Green rectangles represent a graph made by decomposing a plain text into facts with arbitrary wording. Blue shapes represent the same graph but expressed with standardised vocabularies and URIs. Yellow shapes represent data items from another dataset (BRT) which are linked to blue ones (BAG) forming a part of a KG

.

Module 1: Introduction in Spatial Linked Data - Stanislav Ronzhin

The first module will give the basics of Linked Data technology with a focus on geospatial applications. The technology has a steep learning curve. For this reason, the five-star model of open data will be used as the main framework to introduce the core concepts. The best practices for publishing, retrieving and maintaining Spatial Linked Data on the Web will be introduced and discussed based on real-life experience.

Module 2: Data modelling & Ontologies – Dr. Rob Lemmens

The second module will address the formal modelling of data by means of controlled vocabularies, thesauri and ontologies. Through formal representations, these models provide an understanding of the semantics of the concepts, underlying the data, as part of a knowledge graph. We will present technologies such as UML and OWL and show examples of spatial data models from OGC and INSPIRE and their use in practice.

Module 3: Business Aspects of Linked Data – Dr. Erwin Folmer

The third module will give insights into business reasons to invoke Linked Data implementation. It will discuss the business case of Dutch Kadaster for deploying one of the largest Linked Data implementations so far. Next to the business case, we will also demonstrate potential and realized applications. We will showcase several Data Stories and Data Browsers as well as the latest development regarding the Kadaster Knowledge Graph, and chat-bot application utilizing the Knowledge Graph.

Module 4: Technical Aspects of Linked Data – Dr. Wouter Beek

The fourth and last module will address the technical aspects of working with Linked Geospatial Data. We will introduce the SPARQL query language, which is the standardized way in which Linked Datasets can be accessed by application developers. In addition, we will dive into GeoSPARQL, the OGC standard that extends the SPARQL query language with geospatial functions and features. Finally, we will explore SPARQL federation: a standardized approach for running distributed queries over multiple endpoints. The latter allows information from multiple datasets - possibly published by different organizations - to be combined in new and innovative ways.