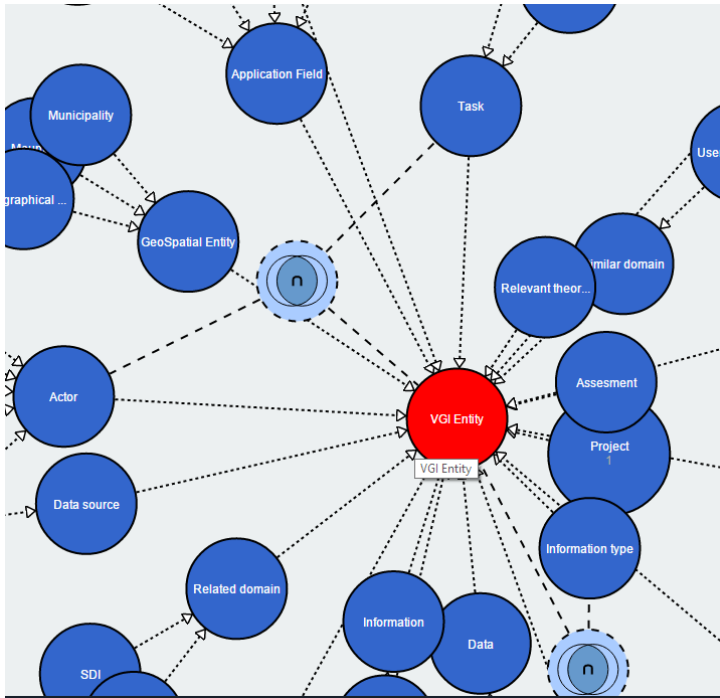


# Linked Data symposium, Eindhoven, 27 June 2016



## ENERGIC Datathon

Rob Lemmens, Frank Ostermann,  
Stanislav Ronzhin

University of Twente, Faculty of Geo-Information Science  
and Earth Observation (ITC), [r.l.g.lemmens@utwente.nl](mailto:r.l.g.lemmens@utwente.nl)

Simon Scheider

Utrecht University

Swarish Marapengopi

GIMA

Ross Purves

University of Zurich



ITC UNIVERSITY OF TWENTE.



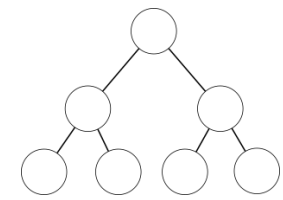
# ENERGIC Project – Objectives - <http://vgibox.eu/>

- Finding new solutions for the exploitation, integration and application of Volunteered Geo-Information (VGI).
  
- Working groups:
  1. Societal and human aspects of VGI,
  2. Spatial data quality and infrastructures and
  3. Data mining, semantics and VGI.
  
- Deliverable: Open and updatable repository of VGI analysis and integration tools and methods, literature and case studies.  
**Ontology serves as semantic backbone.**

# ENERGIC VGI Ontology - Objectives

1. Common understanding of VGI concepts (e.g., for research, outreach and education) – delimiting the field
  - broad coverage, human-readable definitions
2. Create applications and database schema for the development of new VGI applications
  - centered on VGI data structures
3. Usage of ontology of tasks for evaluating of data quality
  - centered on tasks
4. Semantic enrichment of VGI-based systems and indexing literature for VGI source discovery and integration
  - rich set of properties to describe resources
5. Create relationships between VGI and other domains, such as GIS, SDI, etc.

# Semantic enrichment with ENERGIc VGI ontology



## Classes

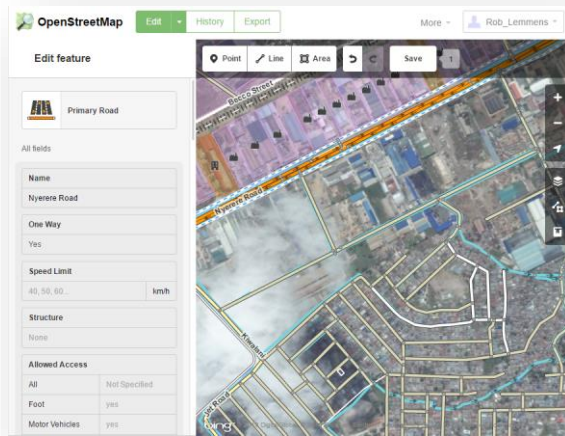
[Accuracy](#) [Active Sensing Task](#) [Activity](#) [Actor](#) [Algorithm](#) [Analysis](#) [Analyst](#) [Annotations](#) [Application](#)  
[Application Field](#) [Article](#) [Assesment](#) [Automated Tagging](#) [Book](#) [Book Chapter](#) [Characterization](#) [Claim](#)  
[Collaborative Human-Computer Analytical Activity](#) [Comparison](#) [Contributor](#) [Daily Activity](#) [Data](#) [Data acquisition](#)  
[Data Analysis](#) [Data description](#) [Data Mining Algorithm](#) [Data quality](#) [Data source](#) [Data type](#) [Dataset Analysis](#)  
[Decision or Policy Making](#) [Definition](#) [Description](#) [Discourse Element](#) [Education](#) [Emergency management](#)  
[End user task](#) [Example](#) [Experimentation](#) [Folksonomy](#) [Geographic Feature Description](#) [Geographic Information](#)  
[Geographical Accuracy](#) [Geographical Object](#) [Geographical precision](#) [Geometry Acquisition](#) [Georef Image](#)  
[Georef Scalar](#) [Georef Structured](#) [Georef Text](#) [GeoreferencedData](#) [GeoSpatial Entity](#) [GIS](#) [Government Agency](#)  
[Hardware](#) [Human Activity](#) [Human sensor](#) [Hypothesis](#) [In-situ sensor](#) [Individual Actor](#) [Information](#)  
[Information type](#) [Institution or Agency](#) [Knowledge Resource](#) [Machine Learning](#) [Mapping](#) [Mapping Activity](#)  
[Measurement Accuracy](#) [Method](#) [Methodology](#) [Mobile application](#) [Monitoring](#) [Mountain](#) [Municipality](#)  
[National Mapping Agency](#) [Navigation](#) [NLP](#) [NonGeoreferenced Data](#) [Ontology](#) [Passive Task](#)  
[Pedestrian Navigation](#) [Positional Accuracy](#) [Precision](#) [Processing](#) [Project](#) [Querying](#)

VGI system: OpenStreetMap



Is of Type

VGI publication



A conceptual model for quality assessment of VGI for the purpose of flood management

Livia Castro Degrossi  
University of São Paulo  
São Carlos, Brazil  
Heidelberg University  
Heidelberg, Germany  
degrossi@icmc.usp.br

João Porto de Albuquerque  
University of Warwick  
Warwick, England  
GIScience Chair, Heidelberg University  
Heidelberg, Germany  
j.porto@warwick.ac.uk

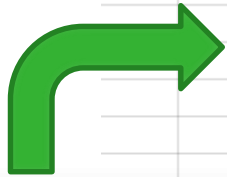
Hongchao Fan, Alexander Zipf  
GIScience Chair, Heidelberg  
University  
Heidelberg, Germany  
hongchao.fan@uni-heidelberg.de,  
zipf@uni-heidelberg.de

### Abstract

Volunteered Geographic Information (VGI) has emerged as a potential source of geographic information for different domains. Despite the many advantages associated with it, such information lacks of quality assurance, since it is provided by individuals with different motivations and backgrounds. In response to this, several methods have been proposed to assess the quality of volunteered geographic information of different platforms. However, there has been little investigation aimed at explaining how cross-platform data could be used for quality

# ENERGIC VGI Ontology Cleaned keyword list

					communicating
					communication
				giscience	community
			urban	walking	community
		positional	accuracy	and	community-oriented
				data	completeness
					completeness
					completeness
					complexity
inherent	spatial		and	temporal	component
			collaborative	cloud	computers
				cloud	computing
					computing
				shared	concepts
					conceptualization
		temporally	accurate	road	conceptualizations
					conditions
					confidence
					confidentiality
					connection
					connections
					connections
				d-sourcing	conrmation
					consciously
					consensusproducing
					consistency
				logical	consistency
					consistency

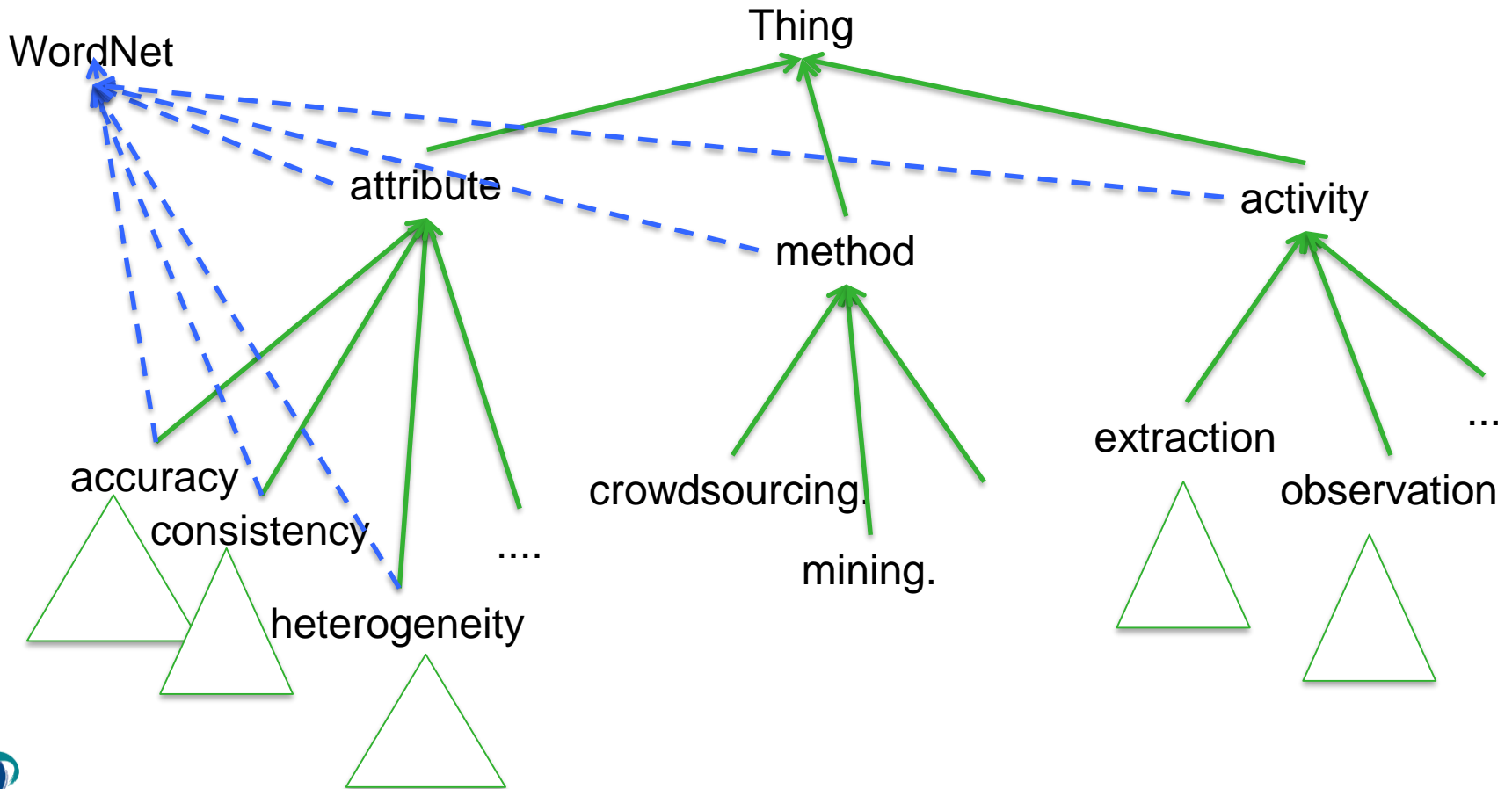


various environments, which might not necessary be caused by or towards these environments. In the following, we extend existing research, and illustrate how social media data can be harnessed to extract people's affective responses to environments. Particularly, we focus on **geotagged photos** in Flickr.

For extracting affective responses from social media data, we apply **sentiment analysis** technique. Sentiment analysis (or **opinion mining**) is a **natural language processing (NLP)** technique, and aims to determine an author's attitudes, opinions or sentiments with respect to the topic written about. Different methods have been proposed for sentiment analysis, among which lexicon-based method is one of the most popular ones. Lexicon-based sentiment analysis employs NLP techniques to tokenize,

# From key terms to organized concepts

- align with WordNet concepts
- add common superclasses from WordNet



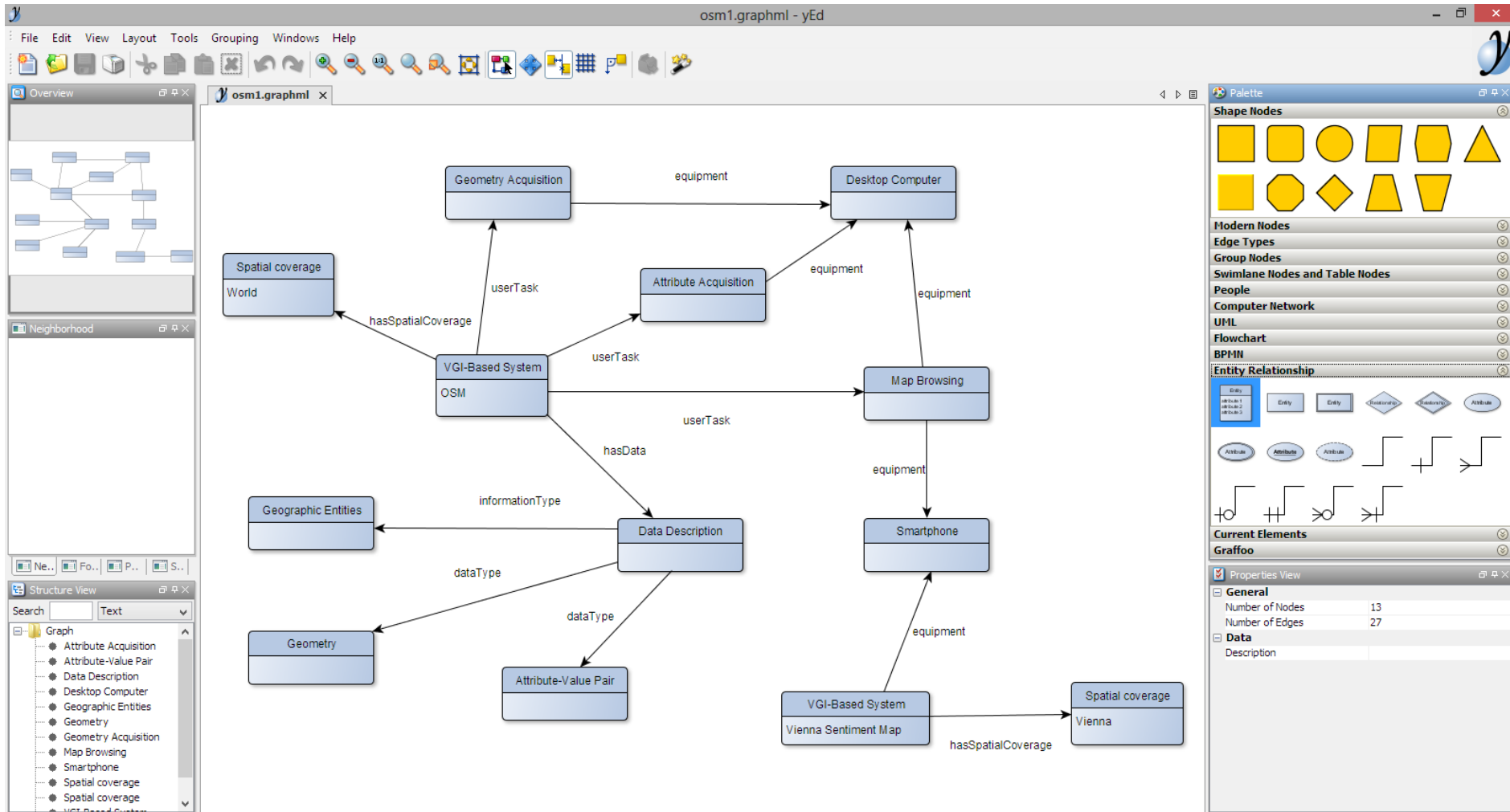
# ENERGIC VGI Ontology Collaborative editing with WebProtege

The screenshot displays the WebProtege web application interface for editing the 'ENERGIC VGI' ontology. The top navigation bar includes 'Project', 'Share', 'Rob Lemmens', and 'Help'. Below this, a tabbed interface shows 'Classes', 'Properties', 'Individuals', 'Changes By Entity', and 'Project Dashboard'. The main workspace is divided into three panels:

- Classes:** A hierarchical tree view on the left showing the ontology structure. The 'Data content class' is selected and highlighted in blue. Other classes include 'Actor', 'Application Field', 'Data quality', 'Data source', 'Sensor', 'Knowledge Resource', 'Method', 'Project', 'Related domain', 'Relevant theoretical concept', 'Similar domain', and 'Task'.
- Class description for Data content class:** A central panel for editing the class. It contains:
  - Display name:** 'Data content class'
  - IRI:** 'http://webprotege.stanford.edu/RyWubZpV7KOhYDNDv1IOZ6'
  - Annotations:** A table with two rows:

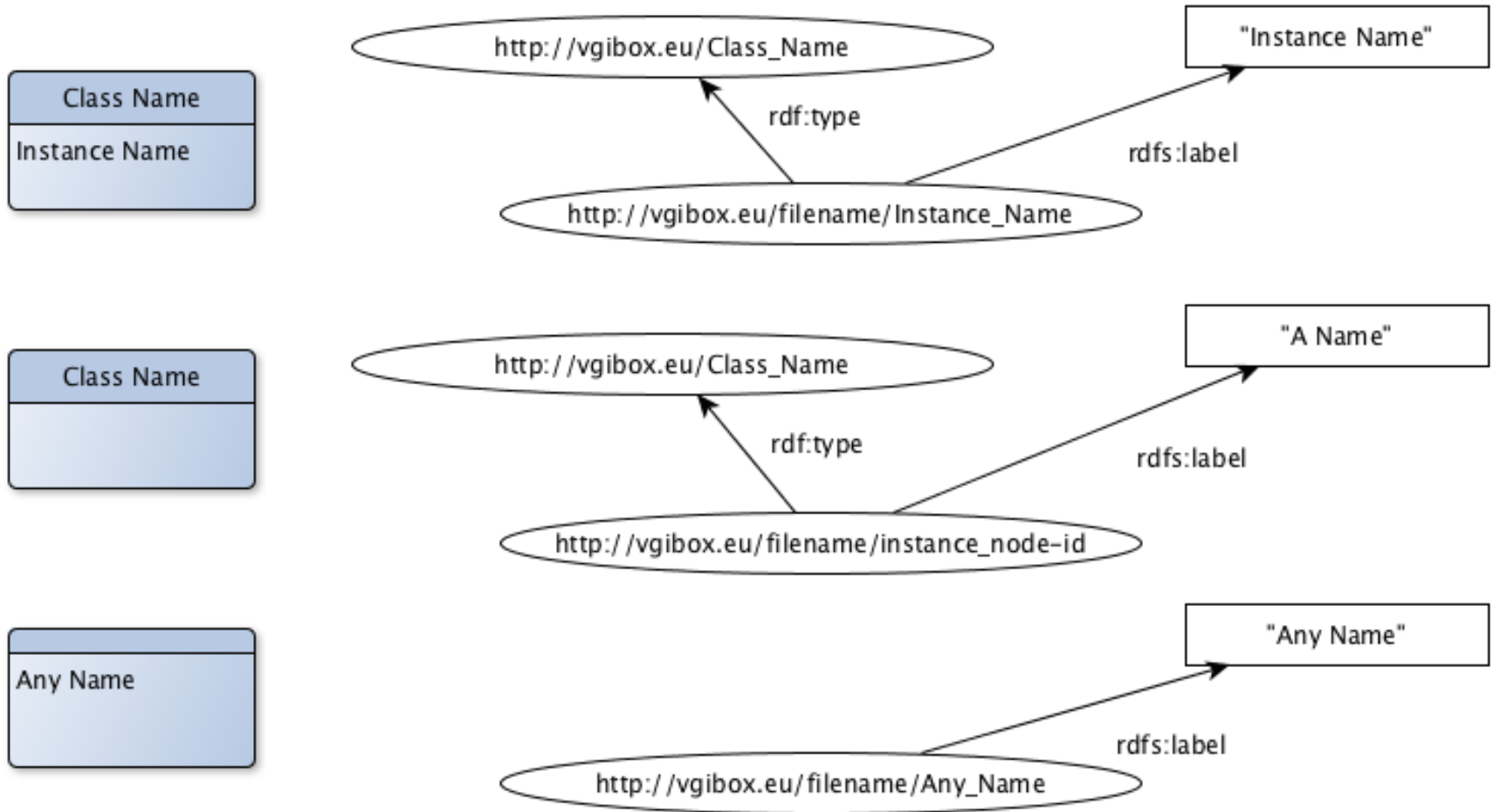
rdfs:label	Data content class	lang	X
rdfs:comment	Link to classes like Building, River, Road, from an existing external ontology	lang	X
  - Properties:** A form with 'Enter property' and 'Enter value' fields and a 'lang' dropdown.
- Discussions for Data content class:** A sidebar on the right for collaborative editing. It features a 'Post new topic' button and a discussion titled 'Or Datatype?' started by 'guest' 9 days ago. The discussion text asks: 'Should we have a class 'Datatype' to describe the different types of data that appear as input and output in VGI processes. Possible subclasses could be:
  - GeoreferencedData
  - Georef Scalar (Nominal, Ordinal, Interval, Ratio)
  - Georef Structured
  - Georef Text
  - Georef Image
  - 
  - NonGeoref Data
  - ..' A response from 'Rob Lemmens' 8 days ago is partially visible.

# ENERGIC VGI Ontology Instantiation with yEd graph editor

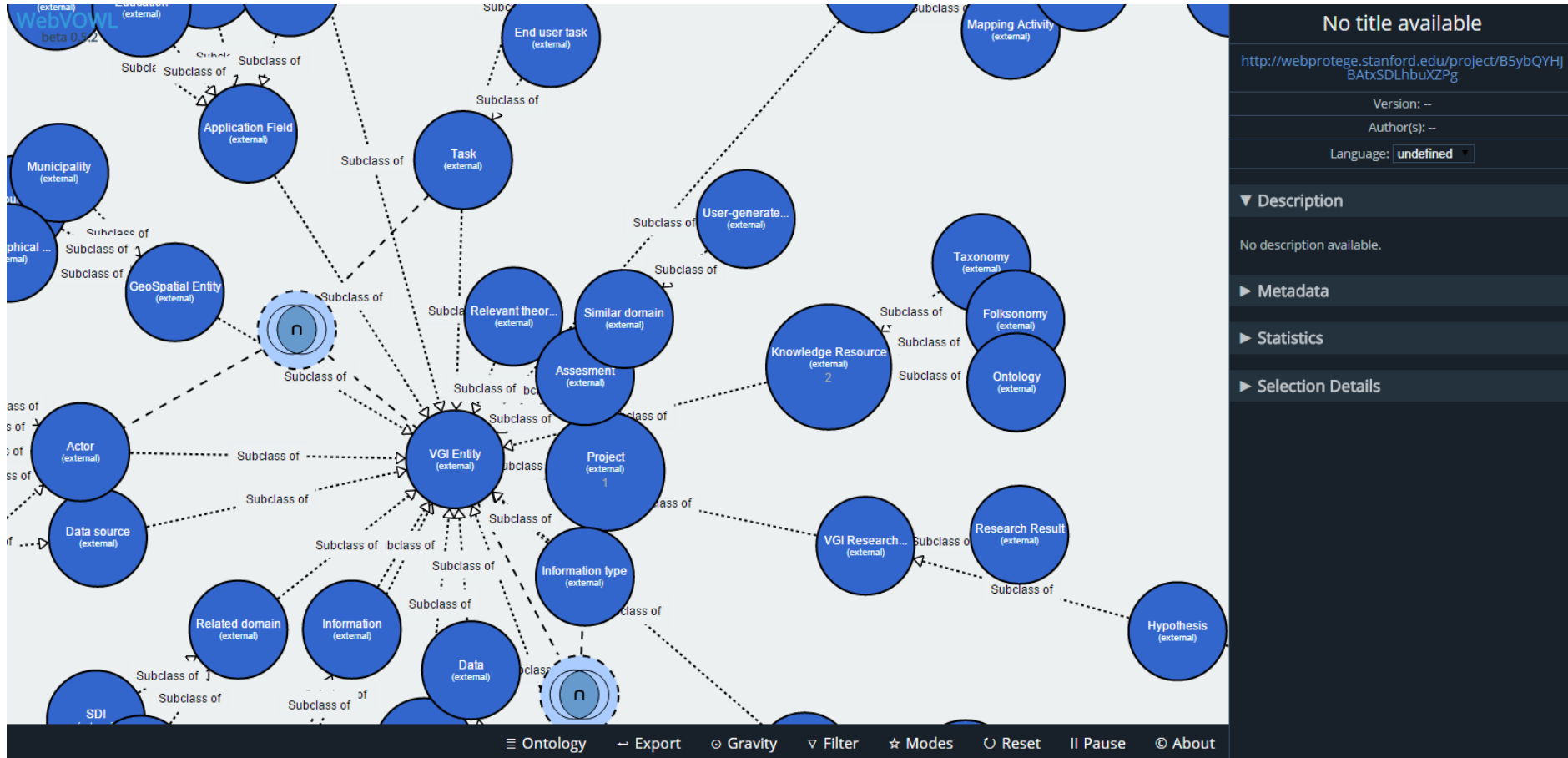




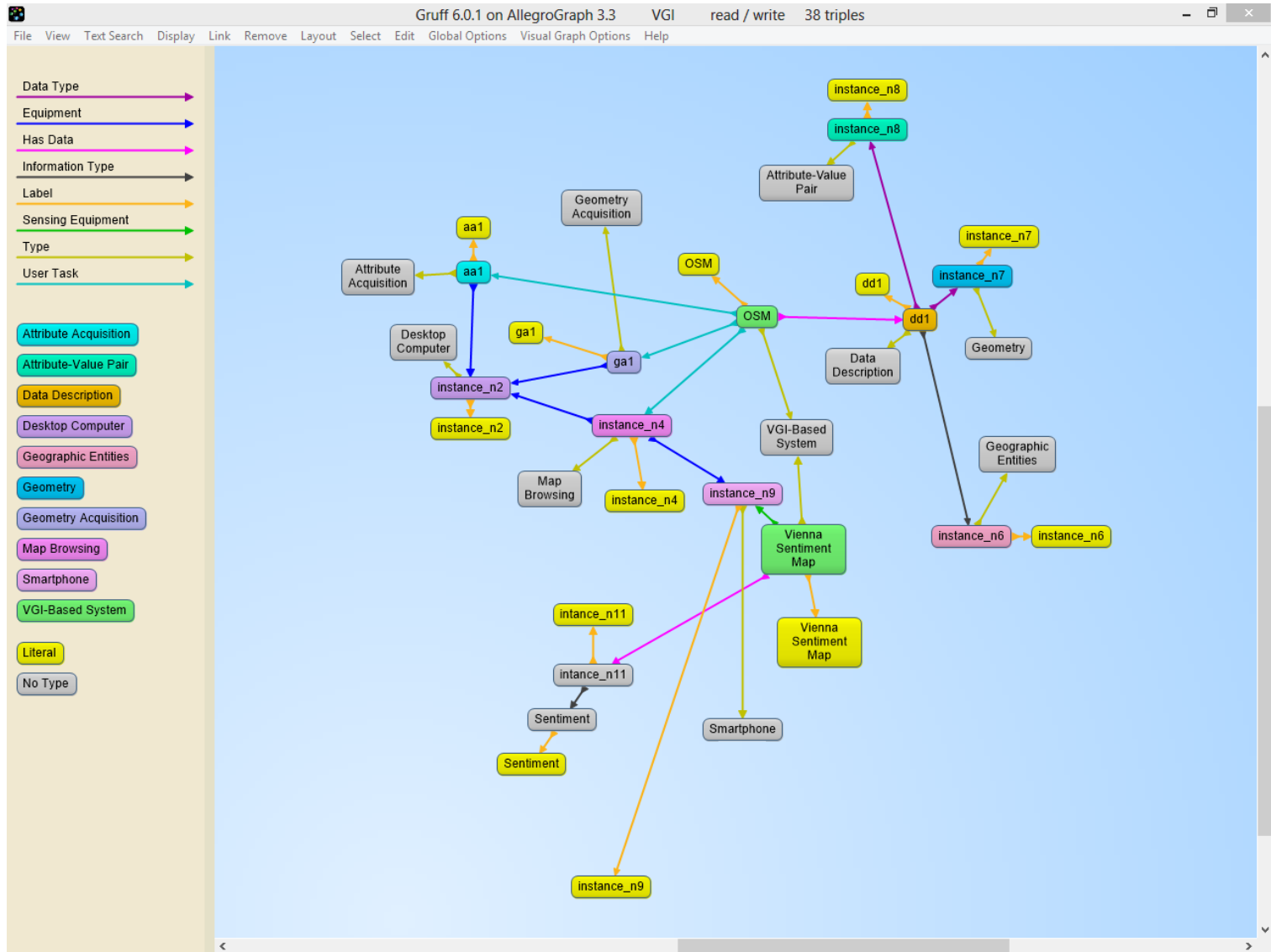
# From E-R diagrams to RDF



# Visualization with WebVOWL



# Visualization with AllegroGraph - Gruff



# Exploratory querying with Linked Data

**SPEX** Spatio-temporal content explorer Endpoint:

Clear Map    Clear Timeline

November 2015

I am looking for:

(2 classes available)

- Things of a kind
- Other

Results (2): [download results as JSON](#)

?VGIbox:VGI-Based_System	?VGIbox:Spatial_coverage	?VGIbox:Map_Browsing	?VGIbox:VGI-Based_System
OSM	World	instance_n4	instance_n9
OSM	World	instance_n4	instance_n2

```

WHERE {
  ?var0 a VGIbox:VGI-Based_System.
  ?var0 VGIbox:hasSpatialCoverage ?var1.
  ?var0 VGIbox:userTask ?var2.
  ?var1 a VGIbox:Spatial_coverage.
  ?var2 a VGIbox:Map_Browsing.
  ?var2 VGIbox:equipment ?var3.
  
```

# ENERGIC Datathon - <http://vgibox.eu/>

Demonstrate the potential of reusing VGI:

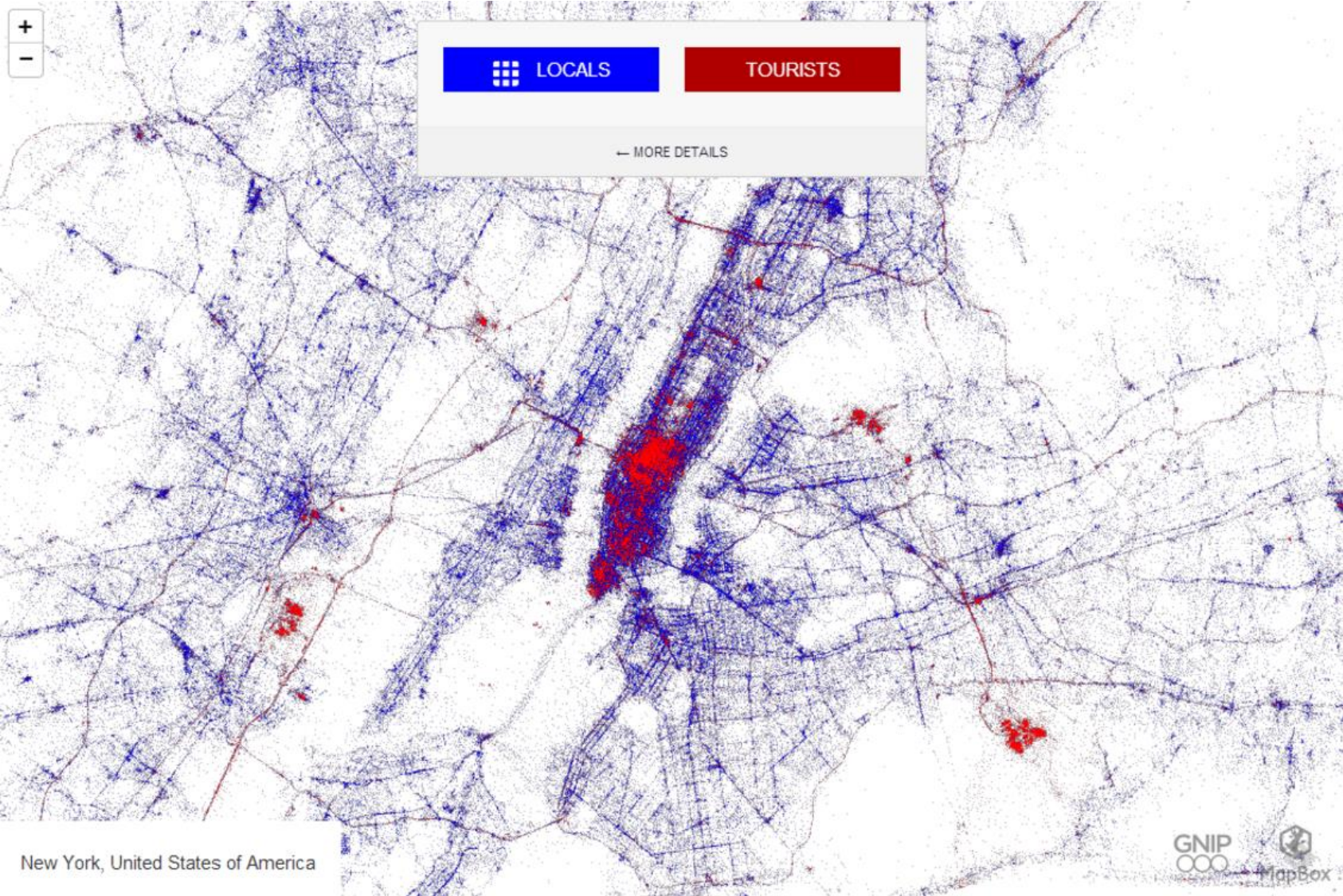
## ■ Challenges:

1. Analyse spatio-temporal VGI sources (Telecom, Tourist, Transport) with other VGI and authoritative data
2. Expanding and test ENERIGIC ontology
  - a. Create relevant VGI system descriptions and detailed VGI publication references based on the ENERIGIC ontology and publish these descriptions online.
  - b. Create and execute (Geo)SPARQL queries.
3. Align the ENERIGIC ontology with other ontologies, such as Geonames, OSM and DBPedia -> create a triple store execute (Geo)SPARQL queries.
4. Triplify VGI and connect to other Linked Data.

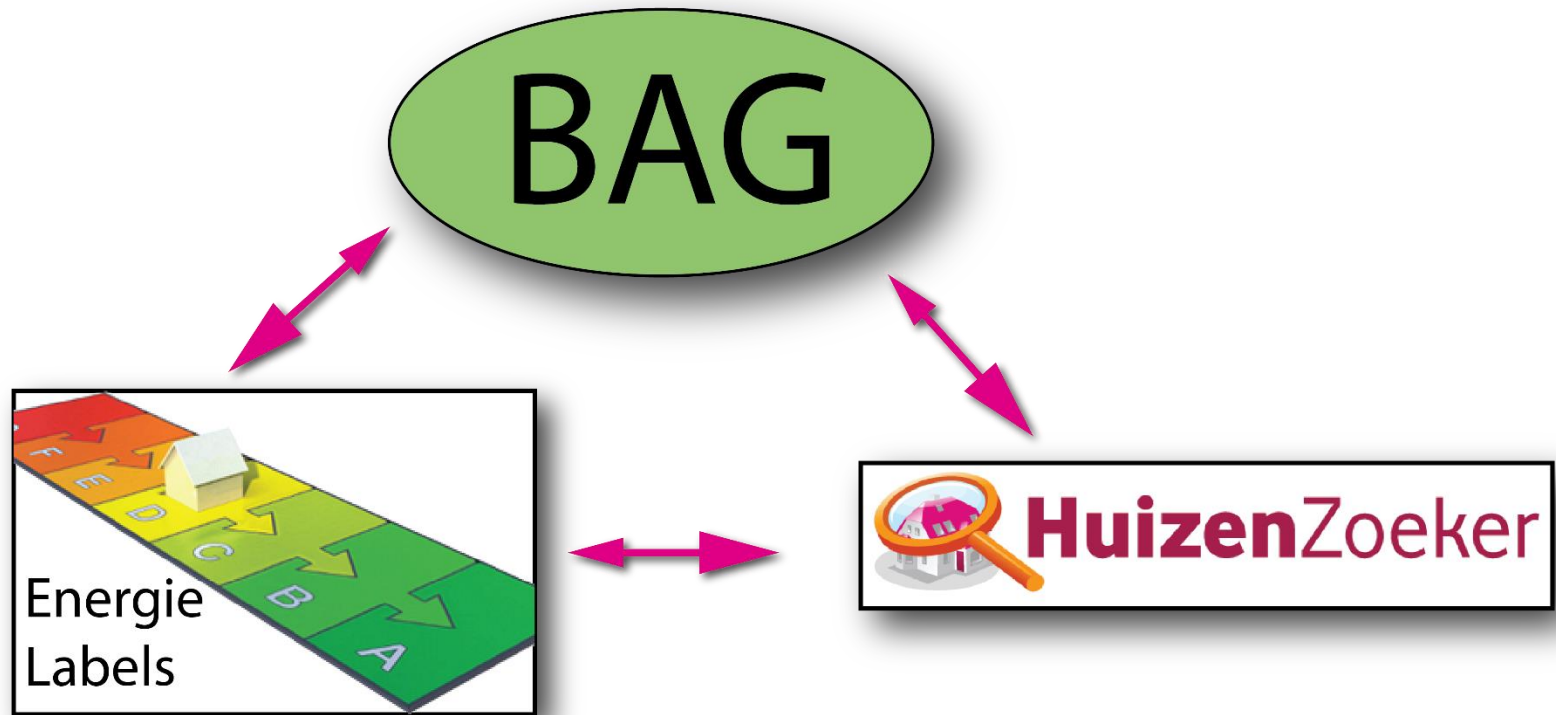
# ENERGIC Datathon - <http://vgibox.eu/>

- Input: VGI ontology and other ontologies, VGI sources
- Deliverable (31 July 2016): Report and recorded pitch on results, sources, methods, novelty.
- Prize: Trip to the final COST Energic meeting in London on 12-13 October 2016.

# VGI analysis



# Use case: BAG + HuizenZoeker + Energie labels



Stanislav Ronzhin (2016)



# Query to select all the 'verblijfsobjecten' with their area in a neighborhood of interest

**SPEX** Spatio-temporal content explorer Endpoint:

```

graph TD
    var3((? var3)) -- "bag:naamOpenbareRuimte" --> OR((? bag:OpenbareRuimte))
    OR -- "bag:gerelateerdeOpenbareRuimte" --> VR((? bag:Verblijfsobject))
    OR -- "bag:hoofdadres" --> NR((? bag:Nummeraanduiding))
    NR -- "bag:huisnummer" --> var4((? var4))
    
```

Map controls: Clear Map, Clear Timeline

Timeline: 22, 23, 24, 25, 26, 27, 28 (September 2015)

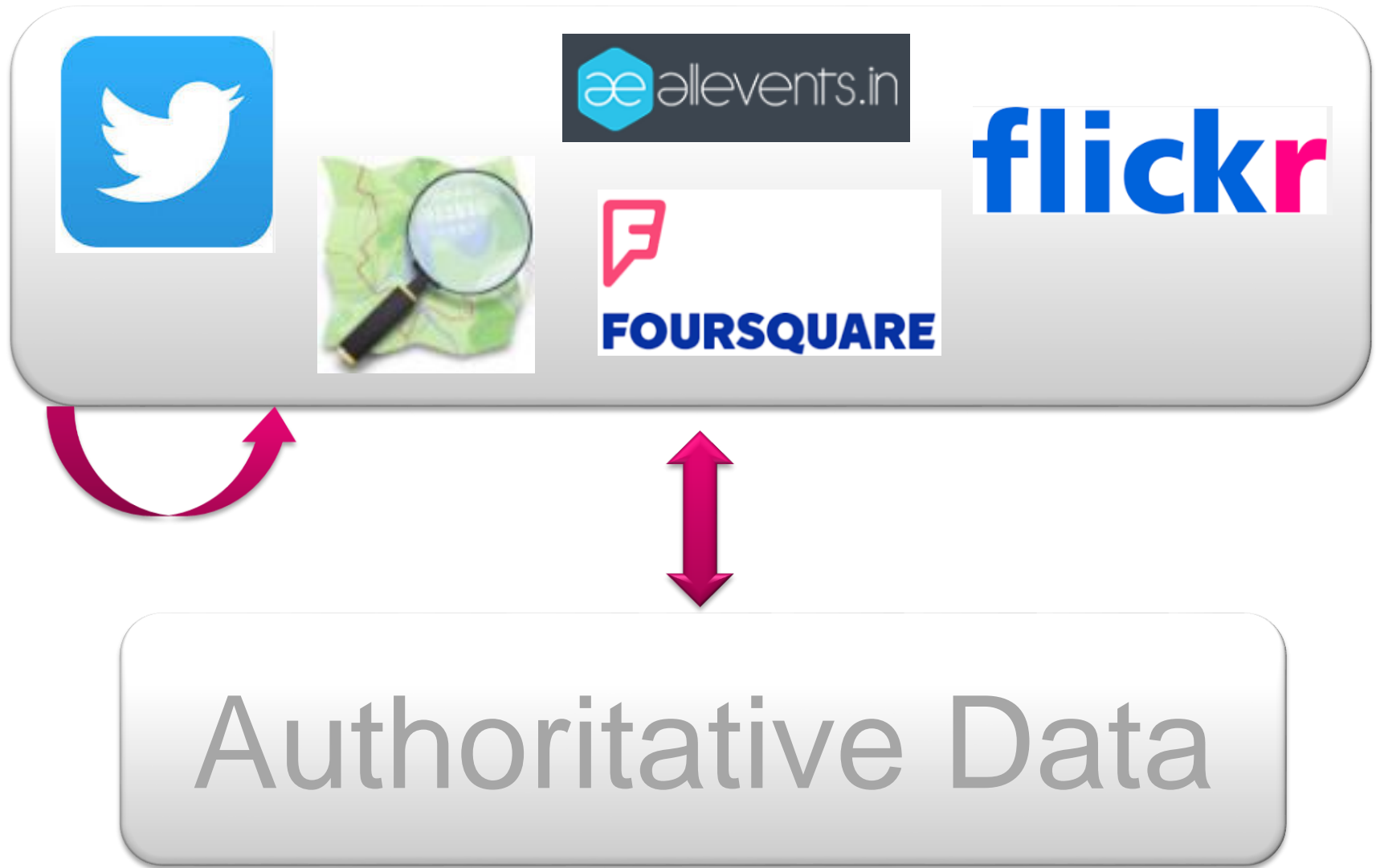
Results (141): [download results as JSON](#)

?bag:Verblijfsobject	?bag:Nummeraanduiding	?bag:OpenbareRuimte	?var3	?var4
0546010000010610	0546200000010610	05463000000001581	Zonneveldstraat	1
0546010000055195	0546200000055195	05463000000000279	Commanderijpoort	8
0546010000011412	0546200000011412	05463000000000279	Commanderijpoort	1
0546010000011414	0546200000011414	05463000000000279	Commanderijpoort	6
0546010000011413	0546200000011413	05463000000000279	Commanderijpoort	2

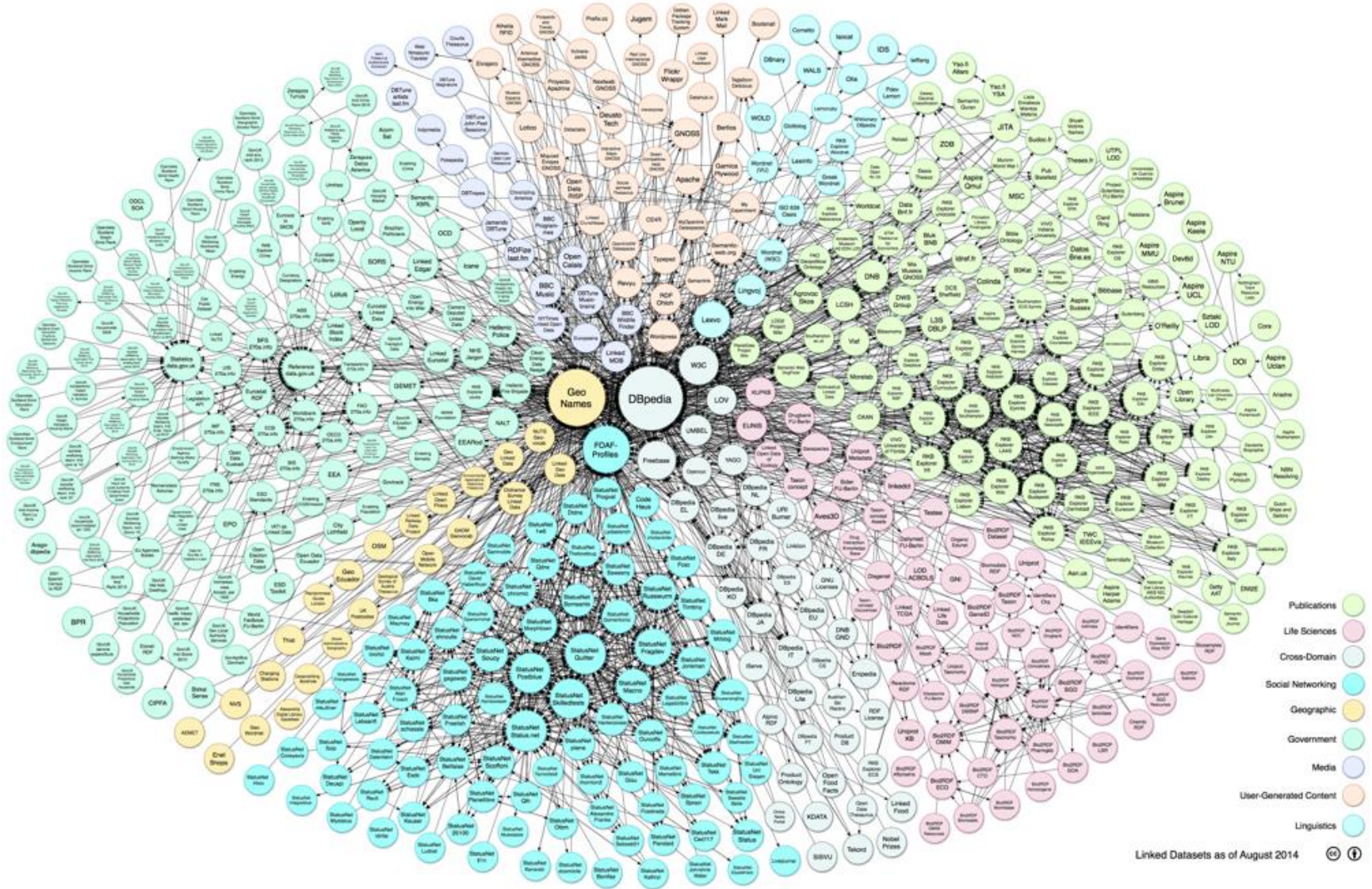
```

Query:
SELECT
DISTINCT
*
WHERE {
?var0 a bag:Verblijfsobject.
?var0 bag:hoofdadres ?var1.
?var1 a bag:Nummeraanduiding.
?var1 bag:gerelateerdeOpenbareRuimte ?var2.
?var1 bag:huisnummer ?var4.
?var2 a bag:OpenbareRuimte.
    
```

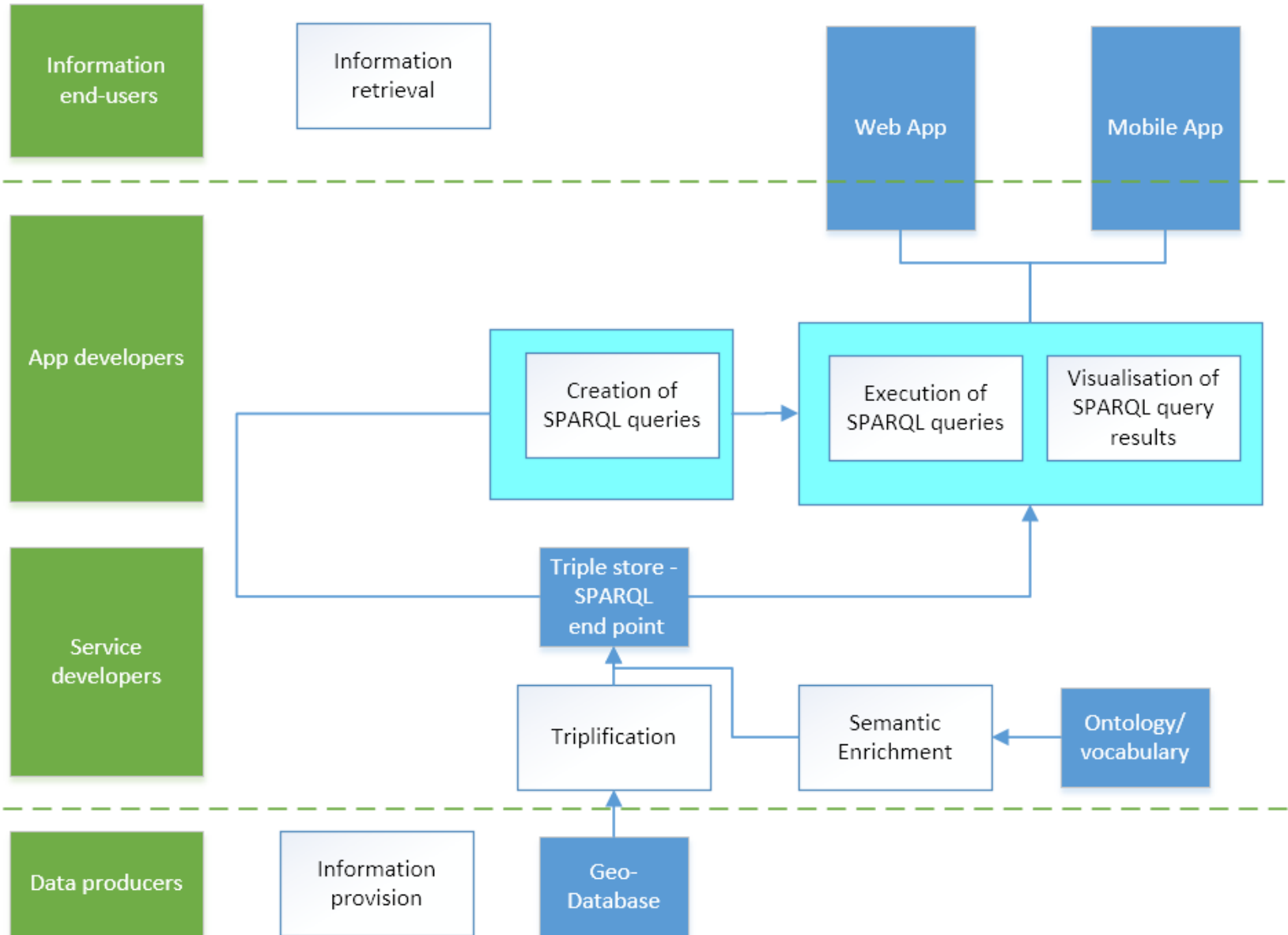
# Combine VGI sources directly and through Linked Data



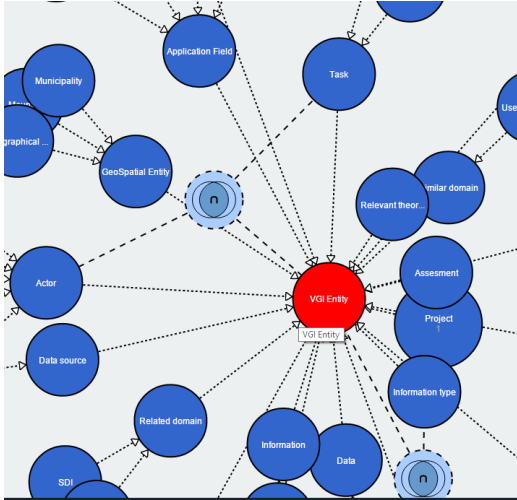
# Linked Open Data cloud



# Workflow - based on Linked Data



# Linked Data symposium, Eindhoven, 27 June 2016



## ENERGIC Datathon

Rob Lemmens, Frank Ostermann,  
Stanislav Ronzhin

University of Twente, Faculty of Geo-Information Science and Earth  
Observation (ITC), [r.l.g.lemmens@utwente.nl](mailto:r.l.g.lemmens@utwente.nl)

Simon Scheider

Utrecht University

Swarish Marapengopi  
GIMA

Ross Purves  
University of Zurich



## Datathon Challenges

<http://vgibox.eu/activities/datathon-challenge/>

## Tips & Tricks

<https://docs.google.com/document/d/1C3B5d7SXjXDQuS6SvBbMCZv6lyZPK E7ywCHSiYbYJcQ/edit?usp=sharing>