Novel indicators for identifying critical *INFRA*structure at *RISK* from Natural Hazards

Deliverable D2.2

GIS Knowledge Base

Primary Author	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF, Dina Sukhobok / SINTEF, Nikolay Nikolov / SINTEF, Antoine Pultier / SINTEF, Arash Khorramhonarnama / SINTEF, Kenneth Gavin / Gavin and Doherty Geosolutions, Karlo Martinović / Gavin and Doherty Geosolutions
WP	2
Submission Date	01/03/2016
Primary Reviewer	Pierre Gehl / UCL
Dissemination Level	PU

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 603960.

Project Information

Project Duration:	1/10/2013 - 30/09/2016
Project Coordinator:	Professor Eugene O' Brien Roughan & O' Donovan Limited eugene.obrien@rod.ie
Work Programme:	2013 Cooperation Theme 6: Environment (Including Climate Change).
Call Topic:	Env.2013.6.4-4 Towards Stress Testing of Critical Infrastructure Against Natural Hazards-FP7-ENV-2013-two stage.
Project Website:	www.infrarisk-fp7.eu

Partners:

Roughan & O' Donovan Limited, Ireland

Eidgenössische Technische Hochschule Zürich, Switzerland.

Dragados SA, Spain.

Gavin and Doherty Geosolutions Ltd., Ireland.

Probabilistic Solutions Consult and Training BV, Netherlands.

Agencia Estatal Consejo Superior de Investigaciones Científicas, Spain.

University College London, United Kingdom.

PSJ, Netherlands.

Stiftelsen SINTEF, Norway.

Ritchey Consulting AB, Sweden.

University of Southampton (IT Innovation Centre), United Kingdom.

Document Information

Version	Date	Description	Primary Author
0.1	02/02/2016	Initial structure (ToC).	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF
0.2	12/02/2016	Initial draft (contents) based on input from all Task T2.2 partners.	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF, Dina Sukhobok / SINTEF, Nikolay Nikolov / SINTEF, Antoine Pultier / SINTEF, Arash Khorramhonarnama / SINTEF
0.3	15/02/2016	Updated draft ready for review by Gavin & Doherty Geosolutions.	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF
0.4	26/02/2016	Added contribution from Gavin and Doherty Geosolutions. New version ready for internal review by UCL.	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF, Kenneth Gavin / Gavin and Doherty Geosolutions, Karlo Martinović / Gavin and Doherty Geosolutions
0.5	29/02/2016	Addressed internal review comments from Pierre Gehl / UCL. Addressed review comments from Brian Bell / BT. Updated sample dataset description based on input from Bryan T. Adey / ETH Zurich.	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF
1.0	01/03/2016	Final formatting and layout.	Dumitru Roman / SINTEF, Brian Elvesæter / SINTEF

Executive Summary

This report describes the knowledge base IT infrastructure, i.e., **Geographic Information System (GIS) Knowledge Base** that has been developed as part of WP2 (with a specific focus on Task T2.2). The report covers the conceptual model and the vocabulary in the Resource Description Framework Schema (RDFS) format that were developed to represent INFRARISK data about infrastructure components and events.

The technical infrastructure for the GIS Knowledge Base is based on DataGraft – a cloud-based service for data transformations and data access. The report describes the transformation and query capabilities of the GIS Knowledge Base using landslides sample dataset as an example. In addition, we have developed a graphical user interface (GUI) application prototype that shows data from the GIS Knowledge Base in a map.

Table of Contents

1 INTRODUCTION

Task T2.2 of Work Package 2 (WP2) aims to apply techniques from the emerging field of Linked Open Data (LOD)[1] to create an interlinked and integrated knowledge base of relevant information about natural disaster events. A knowledge base IT infrastructure is set up for easy access to such information and made available for use to the other WPs in INRARISK and other European projects such as the RAIN project[2].

This report describes the knowledge base IT infrastructure, i.e. **Geographic Information System (GIS) Knowledge Base** that has been developed as part of Task T2.2. The primary potential users of the knowledge base are infrastructure managers. The approach taken was; (i) to provide the framework to describe how such a database should be constructed and (ii) to compile data on a given hazard (in this case landslides) to demonstrate how the database could be used. The experience collected in the database can serve as a case study for the events an infrastructure manager might consider important, and provide them with data of good/bad practices of managing solutions during and after the event. Secondary potential users are researchers (risk management, transportation, civil engineering, natural sciences, etc.) who will benefit from the clear and extensive database.

The remainder of this deliverable is structured as follows:

- Section 2 describes the conceptual model for relating infrastructure failures with natural hazard events.
- Section 3 describes the Resource Description Framework (RDF) vocabulary for representing data in the GIS Knowledge Base.
- Section 4 describes the dataset samples that were collected to populate and test the GIS Knowledge Base development.
- Section 5 introduces DataGraft that was used as a basis to implement the GIS Knowledge Base and describes its data transformation and mapping capabilities.
- Section 6 describes the query capabilities of the GIS Knowledge Base.
- Section 7 describes the graphical user interface (GUI) application prototype developed to show data from the GIS Knowledge Base in a map.
- Section 8 concludes this deliverable.

[1] https://www.w3.org/standards/semanticweb/data
[2] http://rain-project.eu/

2 CONCEPTUAL MODEL

A conceptual object-role model (ORM)[3] was developed. The ORM model defines a conceptual model that relates major global infrastructure failures with natural hazard events. The conceptual model was developed based on a set of interviews and several iterations with the partners involved in WP2. Below we give a brief description of the areas covered by the conceptual model that include:

- Infrastructure components,
- Events (consequences and natural hazards).

2.1 Infrastructure components

An *infrastructure* represents a transport mode, e.g. *road* or *rail*. It has a *name*, *description* and a *geographical feature*. An infrastructure consists of one or more *infrastructure components, e.g. bridge, tunnel, etc.* (see Figure 2-1). Each component has a *name*, *description*, a number of *lanes* and a *geographical feature*. An infrastructure component can be connected to other infrastructure components.

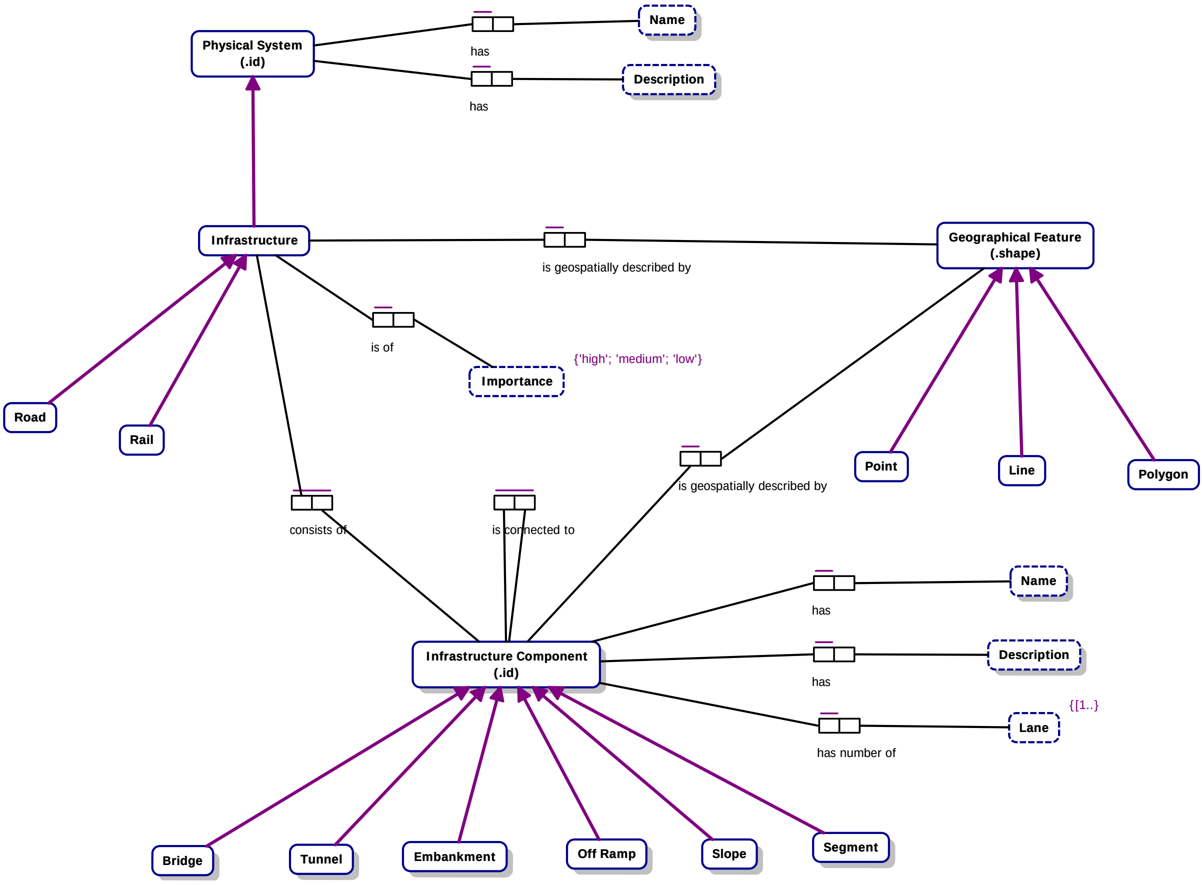

Figure 2-1: Infrastructure component

[3] https://msdn.microsoft.com/en-us/library/aa290383%28v=vs.71%29.aspx

In INFRARISK we focus on critical infrastructure components such as *bridges*, *tunnels*, *embankments*, *off ramps*, *slopes* and *segments (e.g. of a road or rail line)*. Each of these infrastructure component types has their own set of properties as shown in Figure 2-2.

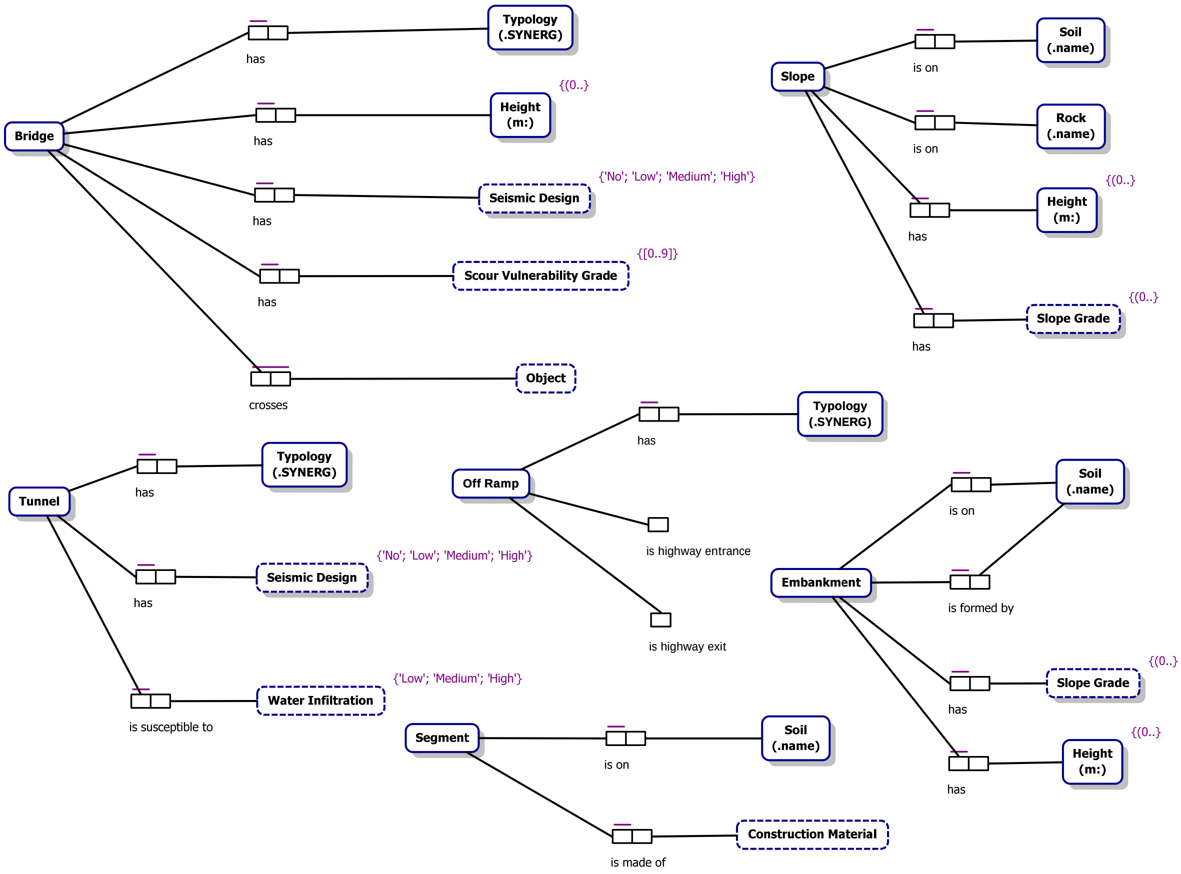

Figure 2-2: Infrastructure component types

2.2 Events

An *event* represents an incident where a *natural hazard* or *infrastructure component failure* has occurred. It has a *name*, *description*, *location*, *date* and *consequence* (see Figure 2-3). An infrastructure component failure concerns the full or partial collapse of an *infrastructure component*.

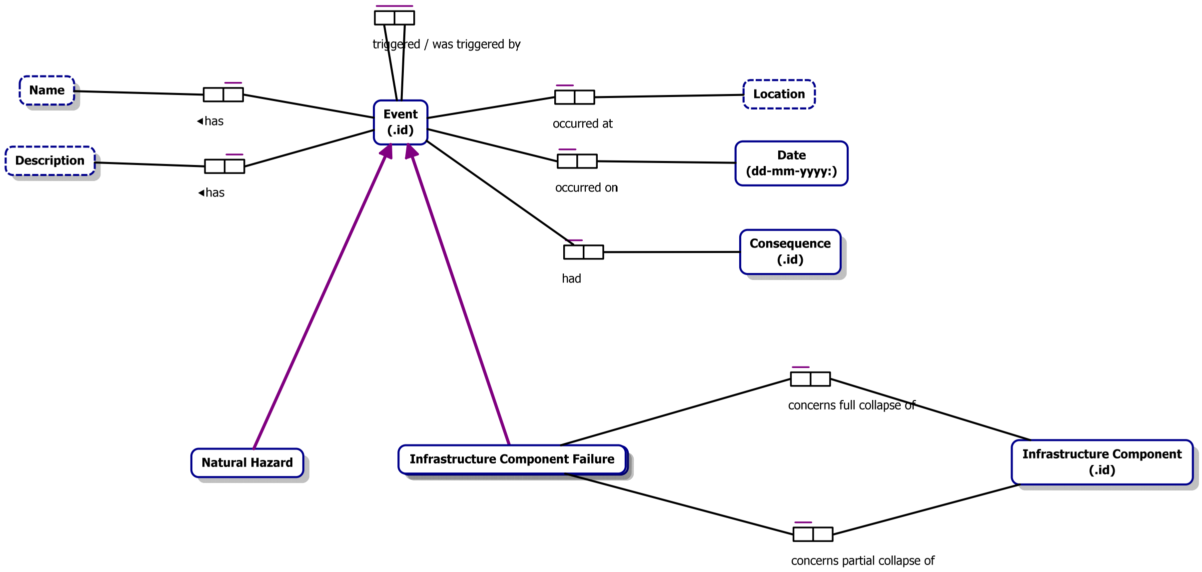

Figure 2-3: Event

A *consequence* represents the expected losses in a specific location as a result of a given *event*. The consequence can be a *monetary loss*, *societal loss* or *usability problem* concerning closure of or reduced traffic on an *infrastructure component* (see Figure 2-4).

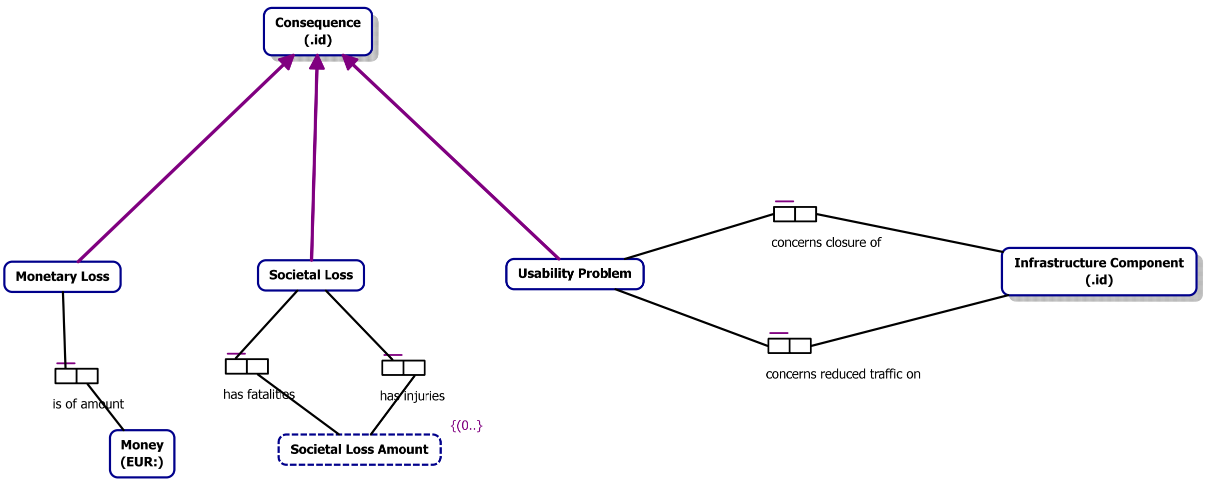

Figure 2-4: Consequence

The conceptual model distinguishes between three types of natural hazard events, namely *earthquakes*, *floods* and *landslides*. The properties for earthquakes and floods are shown in Figure 2-5 and the properties for landslides are shown in Figure 2-6. As can be seen, a landslide can be further classified into a *soil landslide* and a *rock landslide*.

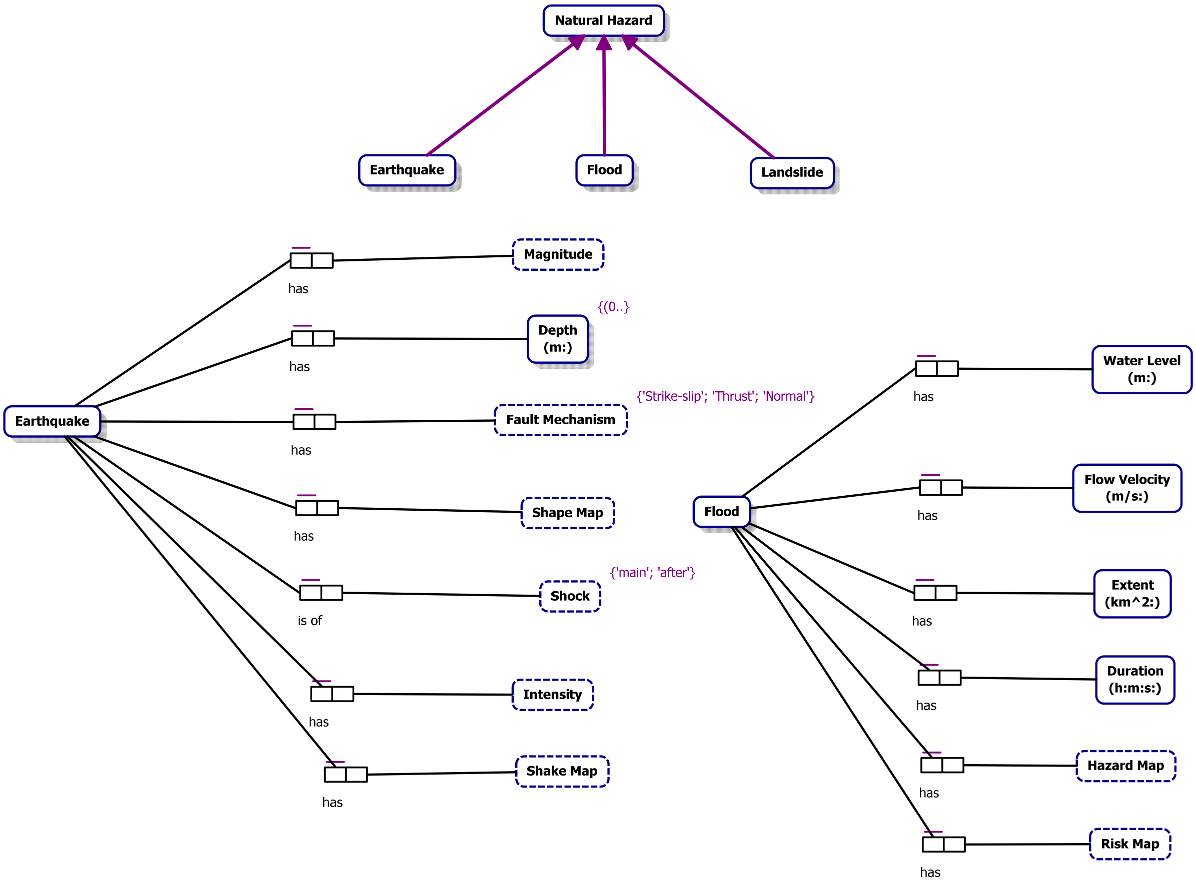

Figure 2-5: Natural hazard types (earthquake and flood)

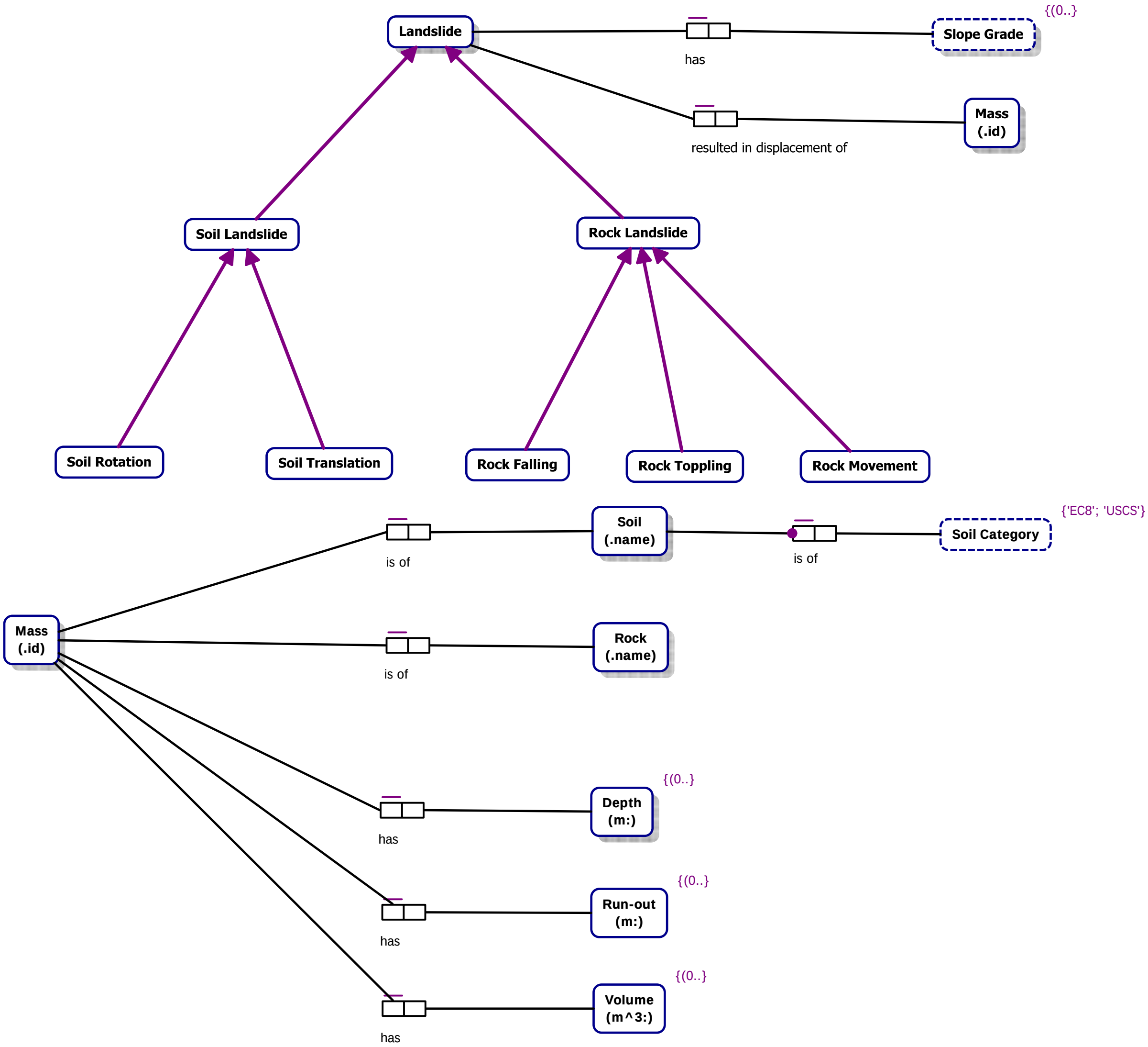

Figure 2-6: Natural hazard types (landslide)

3 RDF VOCABULARY

Resource Description Framework (RDF)[4] is a semantic standard for representing data on the Web. Based on the ORM conceptual model described in Section 2, we defined a Resource Description Framework Schema (RDFS)[5]. The RDF schema provides an RDF vocabulary that can be used to formally represent and exchange data between Web applications.

3.1 Online vocabulary

The RDF vocabulary is available online[6]. It contains two main namespaces which defines a set of terms:

- **rdfs:Resource** which covers the INFRARISK-specific component infrastructure and event terms defined in the ORM conceptual model.
- **geo:Feature** which reuses geographical terms (e.g. *line* and *polygon*) from Schema.org[7]

Figure 3-1 shows the navigable vocabulary documentation online.

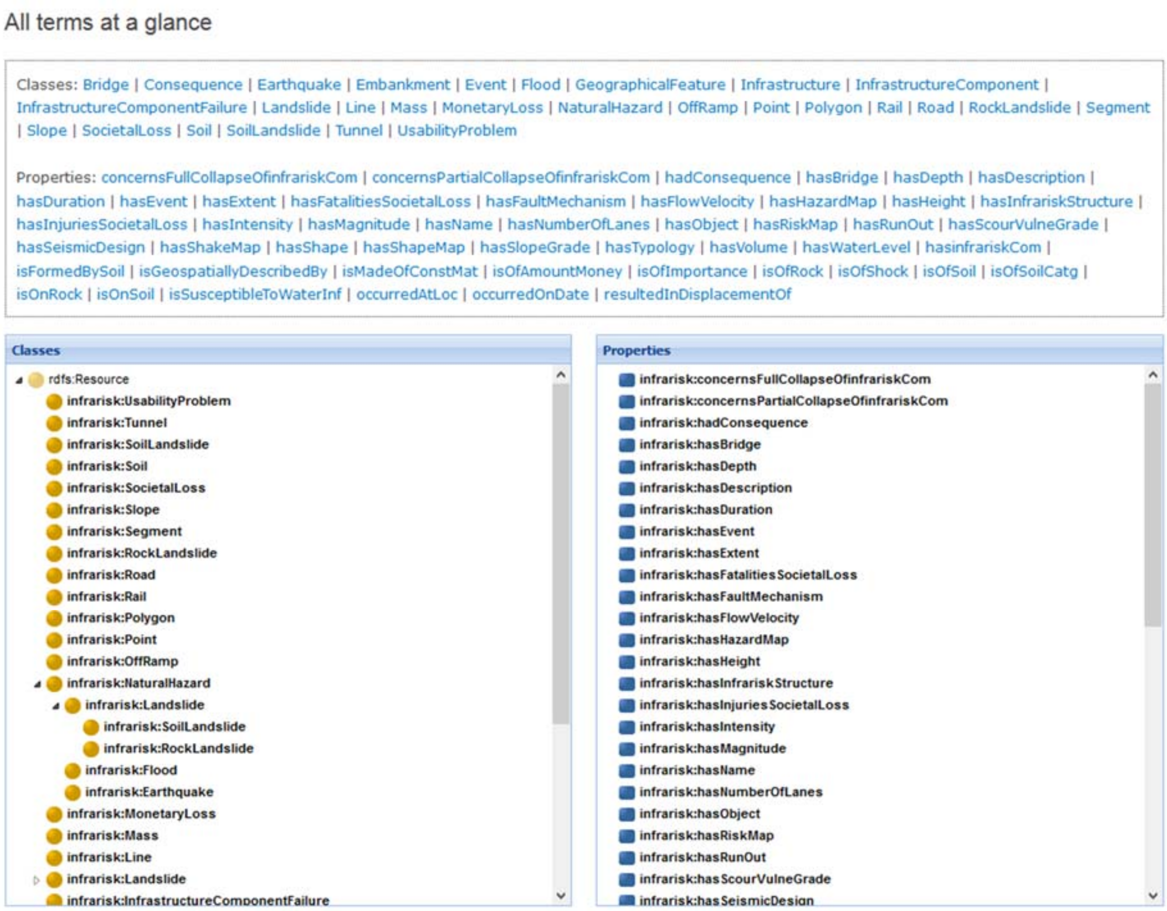

Figure 3-1: INFRARISK RDF vocabulary documentation

[4] https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
[5] https://www.w3.org/TR/2014/REC-rdf-schema-20140225/
[6] http://vocabs.datagraft.net/infrarisk
[7] http://schema.org/

4 DATASET SAMPLES

Infrastructure failures information are available in different formats. For the purpose of this task, we have prepared a number of short Excel spreadsheets and CSV (Comma Separated Values) files with sample data to populate the knowledge base based on the schema (see Section 3) that we have developed.

4.1 Data files

Table 4-1 below lists the sample datasets that we have collected from the INFRARISK WP2 partners.

Table 4-1: Sample datasets

Dataset	Description	Filename
Events #1	Sample data about landslides, bridge failures and road failures in Europe.	ETHZ-Contribution-to-D2.2.xlsx
Floods #1	Sample data about floods in Europe.	database floods yuliya.xls
Floods #2	Sample data about floods in Spain.	EVENTS_DRA.xlxs
Bridge failures #1	Sample data about bridge failures in Europe.	events.xlsx
Bridge failures #2	Sample data about motorway bridge failures in Europe.	events_failures_CSIC_T2.2.xlsx

4.1.1 Events #1 sample dataset

The events #1 sample dataset contains entries of major landslide events (Figure 4-1), bridge failures (Figure 4-2) and road failures (Figure 4-3) in Europe.

	Main diagram													?		?						Consequence diagram								
CSV file:	Infrastructure		Rail	Road	GEOGRAPHICAL FE/	Infrastructure_compone			Slope					Soil		Mass						Cons	Monetary		Societal_Loss			Usability_probl		
input lines from CSV file:	infras	importanc	rail_	road_	ID	Infras	lane	description	slope_l	height	rock_nam	slope_	soil_name	soil_name	soil_categ	mass	depth	rock_nam	run_out	soil_name	volume	cons	mone	mon	soci	has_fa	has_i	usa	conce	conce
data source:					Open Street Map?		open source: reports, articl			after pr	open sour	after p	open source	open sour	open source: d		open sour	open sour	open sour	open sour	open sour	open sourc		open sou		open	open sou		open	open
Hatfield Colliery	1	high	1		line (also, point XY	1	4	Soil cutting ir	1	X	sandstone	X	Alluvium - c	Alluvium -	silt	1	X	sandstone	negligible	Alluvium -	X	1	1	X	1	No	No	1	Yes	No
A85 Glen Ogle	2	medium		1	line (also, point XY	2	2	Natural slope	2	X	semipelit	X	Till and mor	Till and m	till	2	0.6-1.8 m	semipelit	> 400 m	Till and m	50 t	2	2	X	2	No	No	2	No	Yes
Rest and Be Thankful	3	medium		2	line (also, point XY	3	2	Natural slope	3	X	schist	X	Peat and col	Peat and c	colluvium	3	0.6-1.8 m	schist	> 400 m	Peat and c	100 t	3	3	X	3	No	No	3	No	Yes
Harbury tunnel	4	high	2		line (also, point XY	4	2	Soil cutting	4	30 m	mudstone	X	Glaciolacust	Glaciolacu	clay	4	20 m	mudstone	negligible	Glaciolacu	350,000 t	4	4	X	4	No	No	4	Yes	No
Rosyth	5	medium	3		line (also, point XY	5	2	Soil cutting	5	5 m	limestone	31˚	Till	Till	till	5	0.5-1.0 m	limestone	4 m	Till	X	5	5	X	5	No	No	5	No	Yes
St Bees	6	medium	4		line (also, point XY	6	1	Soil cutting	6	19 m	sandstone	19˚	Glaciofluvial	Glaciofluv	granular n	6	0.5-1.5 m	sandstone	> 30 m	Glaciofluvial sand an		6	6	X	6	No	No	6	No	Yes

Figure 4-1: Landslides from the events #1 sample dataset

	Main diagram													?		?						Consequence diagram								
CSV file:	Infrastructure		Rail	Road	GEOGRAPHICAL FE/	Infrastructure_compone			Slope					Soil		Mass						Cons	Monetary		Societal_Loss			Usability_probl		
input lines from CSV file:	infras	importanc	rail_	road_	ID	Infras	lane	description	slope_l	height	rock_nam	slope_	soil_name	soil_name	soil_categ	mass	depth	rock_nam	run_out	soil_name	volume	cons	mone	mon	soci	has_fa	has_i	usa	conce	conce
data source:					Open Street Map?		open source: reports, articl			after pr	open sour	after p	open source	open sour	open source: d		open sour	open sour	open sour	open sour	open sour	open sourc		open sou		open	open sou		open	open
San Benedetto Po bridge	1	medium		1	45.04.27 / 10.56.01	1	2	bridge - RC - 4span - continuous														1	1		1	no	no	1	no	no
Finale Emilia bridge	2	medium		2	44.50.17 / 11.18.40	2	2	bridge - RC - 4span - independent - SS														2	2		2	no	no	2	no	no
Mirandola bridge	3	medium		3	44.52.14 / 11.04.02	3	2	bridge - arch masonry - single span - 25m														3	3		3	no	no	3	no	no
Pontelagoscuro bridge	4	medium		4	44.53.18 / 11.36.29	4	2	bridge - steel deck, rectangular masonry piers, RC abutments - 4span - independent - SS														4	4		4	no	no	4	no	no
Bomporto bridge	5	medium		5	44.43.38 / 11.02.43	5	2	bridge - RC arch - 3span - continuous - SS														5	5		5	no	no	5	no	yes
San Felice sul Panaro bridg	6	medium		6	44.49.35 / 11.08.28	6	2	bridge - RC girder - multi span - SS														6	6		6	no	no	6	no	no
Fossa Station bridge	7	medium		7	42.18.14 / 13.30.12	7	2	bridge - RC - 3span - continuous														7	7		7	no	no	7	yes	no

Figure 4-2: Bridge failures from the events #1 sample dataset

	Main diagram													?		?						Consequence diagram								
CSV file:	Infrastructure		Rail	Road	GEOGRAPHICAL FE/	Infrastructure_compone			Slope					Soil		Mass						Cons	Monetary		Societal_Loss			Usability_probl		
input lines from CSV file:	infras	importanc	rail_	road_	ID	Infras	lane	description	slope_l	height	rock_nam	slope_	soil_name	soil_name	soil_categ	mass	depth	rock_nam	run_out	soil_name	volume	cons	mone	mon	soci	has_fa	has_i	usa	conce	conce
data source:					Open Street Map?		open source: reports, articl			after pr	open sour	after p	open source	open sour	open source: d		open sour	open sour	open sour	open sour	open sour	open sourc		open sou		open	open sou		open	open
Einsiedeln	1	low		1	47.07.42 / 08.50.11	1	1	road														1	1	0	1	no	no	1	yes	no
Hasliberg	2	medium		2	46.44.09 / 08.12.25	2	2	road														2	2	0	2	no	no	2	yes	no
Eggiwil	3	low			46.51.26 / 07.48.41	3		area around Eggiwil														3	3	0	3	no	no	3	no	no
Eggiwil	4	low			46.52.18 / 07.48.36	4		area around Eggiwil														4	4	0	4	no	no	4	no	no
Eggiwil	5	low			46.52.40 / 07.48.14	5		area around Eggiwil														5	5	0	5	no	no	5	no	no
Eggiwil	6	low			46.52.04 / 07.48.45	6		area around Eggiwil														6	6	0	6	no	no	6	no	no
Signau	7	low			46.55.42 / 07.44.15	7		area around Signau														7	7	0,1	7	no	no	7	no	no
Reichenbach im Kandertal	8	medium			46.33.15 / 07.44.32	8		stream														8	8	0,2	8	no	no	8	no	yes
Rudolfstetten-Friedlisberg	9	low			47.22.17 / 08.22.26	9		houses														9	9	0	9	no	no	9	no	no
Aesch bei Birmensdorf	10	medium		3	47.20.12 / 08.26.28	10		roads and houses														10	10	0	10	no	no	10	no	yes

Figure 4-3: Road failures from the events #1 sample dataset

4.1.2 Floods #1 sample dataset

The floods #1 sample dataset contains 362 entries about flood events in Europe (see Figure 4-4).

Number	Country	Detailed Locations	Began	Ended	Duration in Days	Dead	Displaced	Damage (USD)	Type	Main cause	Severity index	flooded area	Affected sq km	peak discharge Q	Peak Water level h	Magnitude (M)**	Centroid X	Centroid Y
1	Albania	Northw estern Albania	15-Mar-13	12-Apr-13	29	0	2800			Heavy Rain	1		3304,87			5,0	19,6182	42,0043
2	Albania	Northw est Albania, central Croatia, parts of Bosnia	4-Jan-10	13-Jan-10	10	0	1500		riverine flood	Heavy Rain	1,5	10500 hectares	60173,47	3.600 m3/sec		6,0	20,0828	41,4664
3	Albania	Southern regions: Gjirokaster, Durres, Elbasan, Lezhe, Berat, Fier. Vlore, also the capital Tirana. Tow ns: Burrel, Rubik. Librazhd district	30-Nov-05	3-Dec-05	4	3	0		riverine flood	Heavy rain	1		10408,03			4,6	20,2064	40,4745
4	Albania	Shkodra district - Obot	4-Dec-04	8-Dec-04	5	0	0	173 000	riverine flood	Heavy rain	1		426,448			3,3	19,5485	42,04
5	Albania	Northern - Lezha and Shkoder regions. areas of Lac, Lezha, Shkodra and Kukes. Southern - districts of Berat, Skrapar, Permet, Tepelena, Gjirokastra, Saranda and	21-Sep-02	8-Oct-02	18	1	9700	17 500 000	riverine flood	Heavy rain	1		11922,12			5,3	19,7738	41,3975
6	Albania	Northw est: Lezhe	20-Dec-97	23-Dec-97	4	0	400		coastal flood	Heavy rain	1		1640,67			3,8	19,7114	41,8429
7	Albania	Northw est Albania: Lezha area	19-Nov-96	21-Nov-96	3	0	50			Heavy Rain	1		1123,04			3,5	19,7026	41,7999
8	Albania	Alaska: Anchorage, Kenai Peninsula, Matanuska-Susitna	21-Sep-95	24-Sep-95	4	4	0	10 000 000	flash flood	Heavy rain	1		17072,03			4,8	19,9654	41,6303
9	Albania	Provinces: Lezhe, Miredita, Laci	19-Aug-95	26-Aug-95	8	5	0			Heavy rain	1		11109,46			4,9	19,9858	41,8101
10	Albania	Kruja, Lac, Lezha, Shkdora, Tropja, Mirdita	17-Nov-92	19-Nov-92	3	7	2000	7 000 000	flash flood	Torrential rain	2		5619,46		up to 1 m inu	4,5	19,5629	42,0387
11	Austria	Graz	16-Jul-09	19-Jul-09	4	0	0		riverine flood	Torrential Rain	1		2765,43			4,0	15,7089	47,4021

Figure 4-4: Floods #1 sample dataset

4.1.3 Floods #2 sample dataset

The floods #2 sample dataset contains 42 entries about floods in Spain (see Figure 4-5).

	Main diagram			Consequence diagram					Event diagram				
CSV file:	Rail	Road	GEOGRAPHICA	Consequence	Monetary Los:	Societal_Loss		Usability_prol	Event				Triggered_Eve
input lines from (	rail_ID	road_ID		consequence_ID	monetary_los	has_fatalities_	has_injuries_a	usability_prot	date	description	location	name	event_ID1
data source:			Open Street N	open source: reports, articles		open source: ı	open source: reports, article		open source: ı	open source: re	open source: ı	open source: reports, article	
FLOOD = Almería (Spain)		1		**Roads cut: A-348 Alcolea AL-4402 Ohanes**	yes			yes	6-12 January 2010	RAIN	Several Locations. Almería (Spain)		river flood
FLOOD = Cádiz (Spain)		1		**Roads cut: A-389 Olvera. A-393 Arcos de la Frontera. 6 Local Roads.**	yes			yes	6-12 January 2010	RAIN	Flooded: Los Barrios, Algeciras, La Línea and Tarifa.		river flood
FLOOD = Córdoba (Spain)		1		**Roads cut: CO-3300 La Rambla, closed by flood. CO-6213 Cabra, closed by mud accumulation.**	yes			yes	6-12 January 2010	RAIN	Several Locations, Córdoba (Spain)		river flood
FLOOD = Granada (Spain)		1		**Roads cut: A-395 A-425 Monachil closed by snow**	yes	**1 = 76 years woman died trapped in her flooded house**		yes	6-12 January 2010	RAIN - SNOW	Valderrubio and other locations, Granada (Spain)		river flood
FLOOD = Jaén (Spain)		1		**Roads cut: CN-323 Carchelejo**	yes			yes	6-12 January 2010	RAIN	Several Locations, Jaén (Spain)		river flood
FLOOD = Jaén (Spain)	1			**Middle distance train with 300 passengers stopped in Villacañas**	yes			yes	6-12 January 2010	RAIN	Villacañas , Jaén (Spain)		river flood

Figure 4-5: Floods #2 sample dataset

4.1.4 Bridge failures #1 sample dataset

The bridge failures #1 sample dataset contains 7 entries about bridge failures in Europe (see Figure 4-6).

	Main diagram								Consequence diagram			
CSV file:	Infrastructure	Rail	Road	GEOGRAPHIC	Infrastructure_component				Consequence	Societal_Loss		Usability_pro
input lines from	importance	rail_ID	road_ID		Infrastructure	name	lane	description	consequence_ID	societal_loss	has_fatalities	concernsClos
data source:				Open Street I	bridge		open source: reports, articles		open source: reports, articles		open source:	open source:
Calva Bridge	low		1	line	yes	central arc of	1	masonry arch bridge, built 1840	links to the other side of the river closed		no	yes
Northside Bridge, Workington.	medium		1	line	yes	Northside Bridge	2	masonry arch bridge, the subject of a Local Government Board Inquiry in 1903	major road from the Low Cloffocks to the north side of the river was closed		1	yes
Navvies Footbridge, Workington	low			line	yes	Navvies Footbridge		masonry arch bridge,built in 1878	Footpath and cycle way, linking to Workington's town centre, was closed		no	yes
Camerton Footbridge, Camerton.	low		1	line	yes	Camerton Footbridge		masonry arch bridge	access to Camerton churge were cut		no	yes
Dock or Harbour Bridge	low	1		line	yes	Dock bridge	1	masonry arch bridge	single track railway and footpath linking the steelworks and the docks were closed		no	yes
Plaka bridge (footbridge)	low		1	line, point co	yes	Plaka bridge		stone bridge	The largest one-arch bridge in Greec and the Balkans,and the third largest one-arch stone bridge in Europe was destroyed		no	

Figure 4-6: Bridge failures #1 sample dataset

4.1.5 Bridge failures #2 sample dataset

The bridge failures #2 sample dataset contains 4 entries about motorway bridge failures in Europe (see Figure 4-7).

	Main diagram							
CSV file:	Infrastructure		Rail	Road	GEOGRAPHICAL FEATURE (.shap	Infrastructure_component		
input lines from CSV file:	infrastructu	importance	rail_ID	road_ID		Infrastructure_compo	lane	description
data source:					Open Street Map?		open source: reports, articles	
Arifiye TEM Overpass	1	high		1	40.709432/30.357475LON	1	4	Motorway bridge,Fourspan, 100 m-long, simply-supported prestressed concrete bridge
Mustafa Inan TEM Viaduc	1	high		1	40.775221/29.902321	2	4	*M*otorway bridge, Box girder bridge and prestressed concrete bridge, Box girder bridge and prestressed concrete bridge, Span lengths 10x40 m, Pier height 90m
Sakarya River TEM Viaduc	1	high		1	40.739304/30.421336	3	4	Motorway bridge, Two parallel bridges with ten simply-supported spans of precast, prestressed delta girders on single column piers with very wide hammer-head bent caps and seat type abutments.
Bolu TEM viaducts	1	high		1	40.775383/31.356440	4	4	Motorway bridge, 59 spans and dual 2.3 km structures, approximately 95% complete and awaiting installation of expansion joints to complete the project at the time of the earthquake, 40 m spans comprised of 7 lines of simply-supported, prestressed concrete box girders (V-girders) seated on pot bearings with stainless steel PTFE-slider interface. The V-girder is a precast, open-box beam with narrow bottom flanges, moderately battered webs, and small top flanges. The cast-in-place (CIP) deck slab is continuous over 10 spans. The piers are single, CIP, octagonal hollow-core reinforced concrete columns, 4.5x8.0 meters in plan dimension with heights varying from 10 m to about 49 m.

Figure 4-7: Bridge failures #2 sample dataset

5 DATAGRAFT TRANSFORMATIONS

We set up the technical infrastructure for the GIS Knowledge Base based on DataGraft[8]. DataGraft is cloud-based service for data transformations and data access (see Figure 5-1). It allows the user to simplify the creation of Linked Data and map the data we have collected into RDF data compliant with the schema that we have developed.

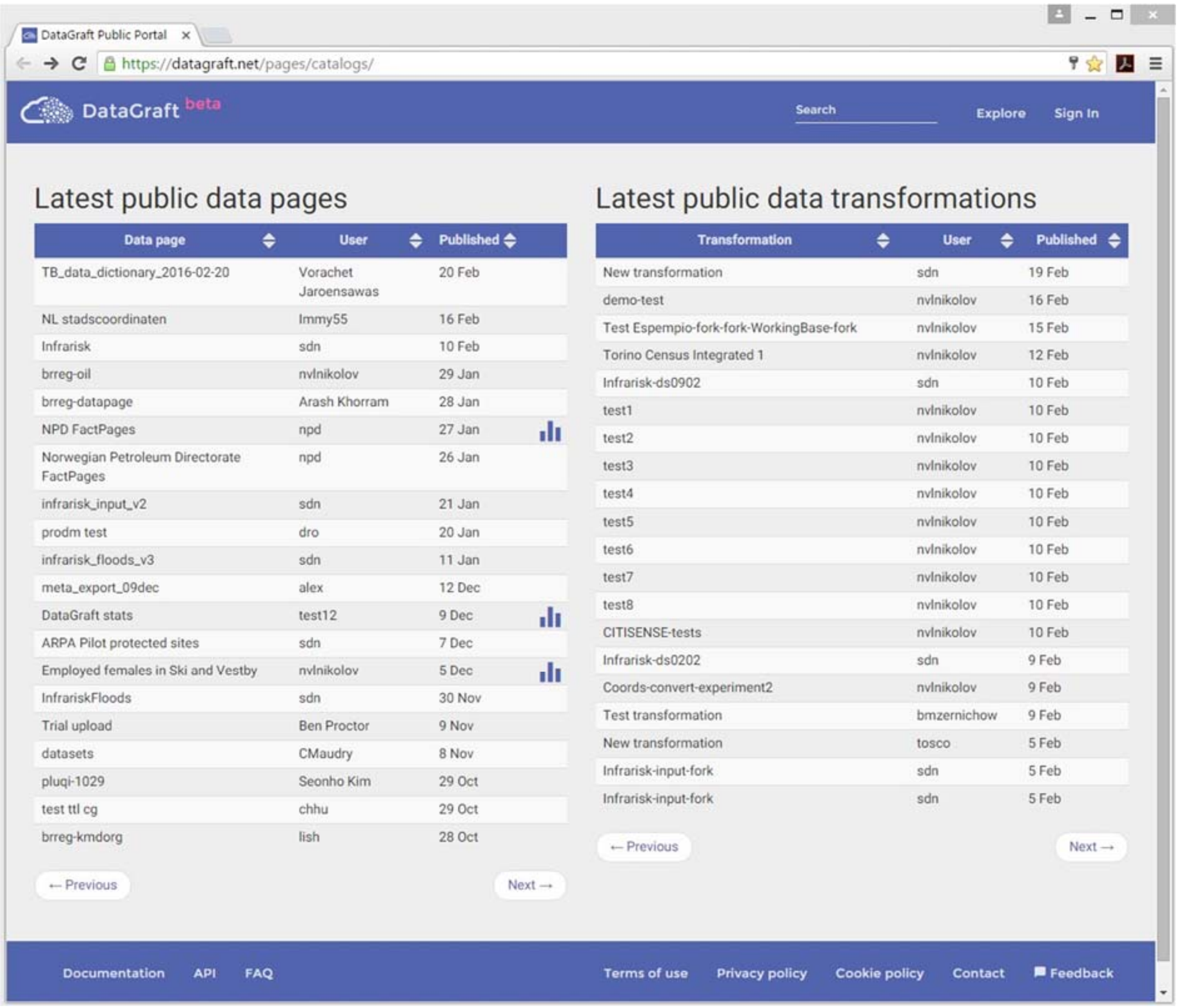

Figure 5-1: DataGraft portal

5.1 Landslides data from the events #1 sample dataset

To illustrate the use we demonstrated the transformation capabilities of DataGraft using the landslides data from the events #1 sample dataset. The original data in tabular format is shown in Figure 5-2 below.

[8] https://datagraft.net

CSV file:	Infrastructure		Rail	Road	GEOGRAPHICAL FEATURE (.shape)	Infrastructure_component			Slope			
Input lines from CSV file:	infrastructure_ID	importance	rail_ID	road_ID		Infrastructure_component_ID	lane	description	slope_ID	height	rock_name	slope_grade
data source:					Open Street Map?		open source: reports, articles			after processing D	open source: dig	after processing
Hatfield Colliery	1	high	1		line (also, point XY coordinates available)	1	4	Soil cutting in immediate vicinity of spoil tip (colliery)	1	X	sandstone	X
A85 Glen Ogle	2	medium		1	line (also, point XY coordinates available)	2	2	Natural slope on the side of roadway	2	X	semipelite	X
Rest and Be Thankful	3	medium		2	line (also, point XY coordinates available)	3	2	Natural slope on the side of roadway	3	X	schist	X
Harbury tunnel	4	high	2		line (also, point XY coordinates available)	4	2	Soil cutting	4	30 m	mudstone and limestone	X
Rosyth	5	medium	3		line (also, point XY coordinates available)	5	2	Soil cutting	5	5 m	limestone	31˚
St Bees	6	medium	4		line (also, point XY coordinates available)	6	1	Soil cutting	6	19 m	sandstone	19˚

Figure 5-2: Landslides sample dataset in tabular format

A DataGraft transformation is defined as a pipeline that consists of two main tasks:

- **Data cleaning** on the original tabular data format
- **Mapping** of the cleaned tabular data format to the RDF schema

5.1.1 Data cleaning functions

In order to map the data entries to the RDF schema we first need to do some data cleaning in DataGraft. Figure 5-3 below shows how we can remove unnecessary rows in the original tabular dataset using the *drop-rows* function in DataGraft.

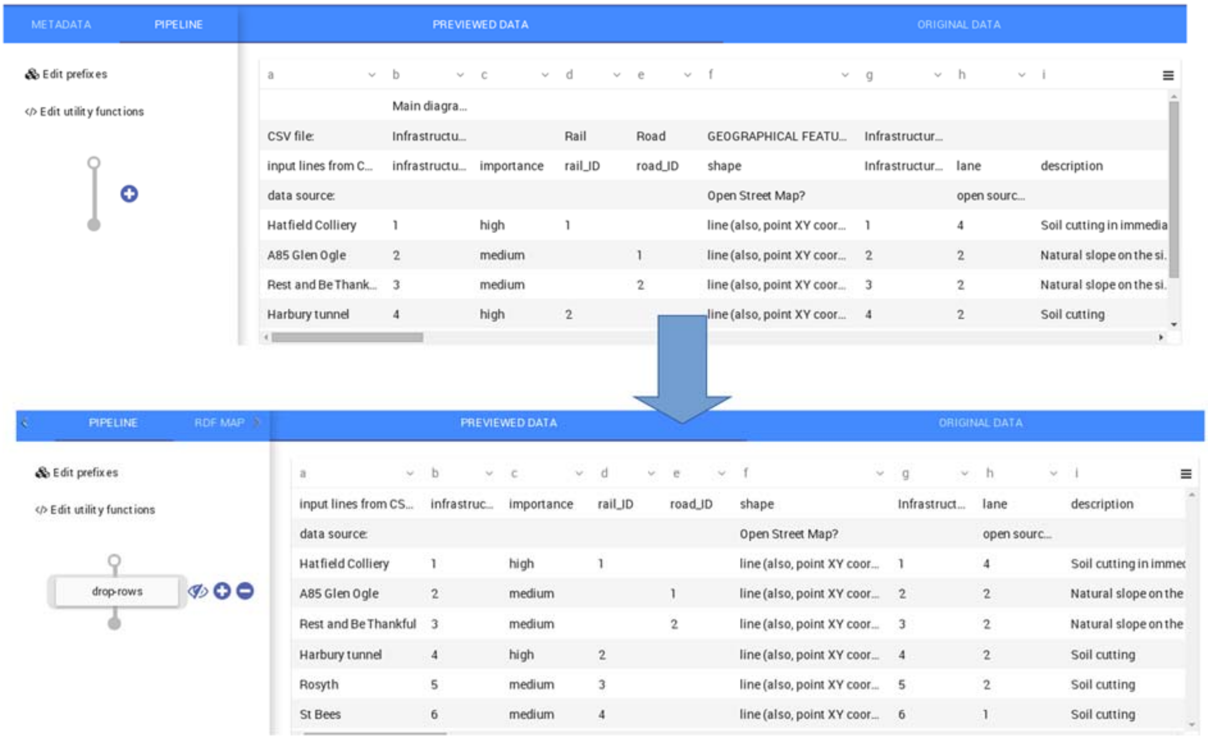

Figure 5-3: Remove unnecessary rows

Next we create a header from the first row and remove unused columns in the original tabular dataset using the *make-dataset* and *columns* functions in DataGraft (see Figure 5-4 where the *description* column has been removed).

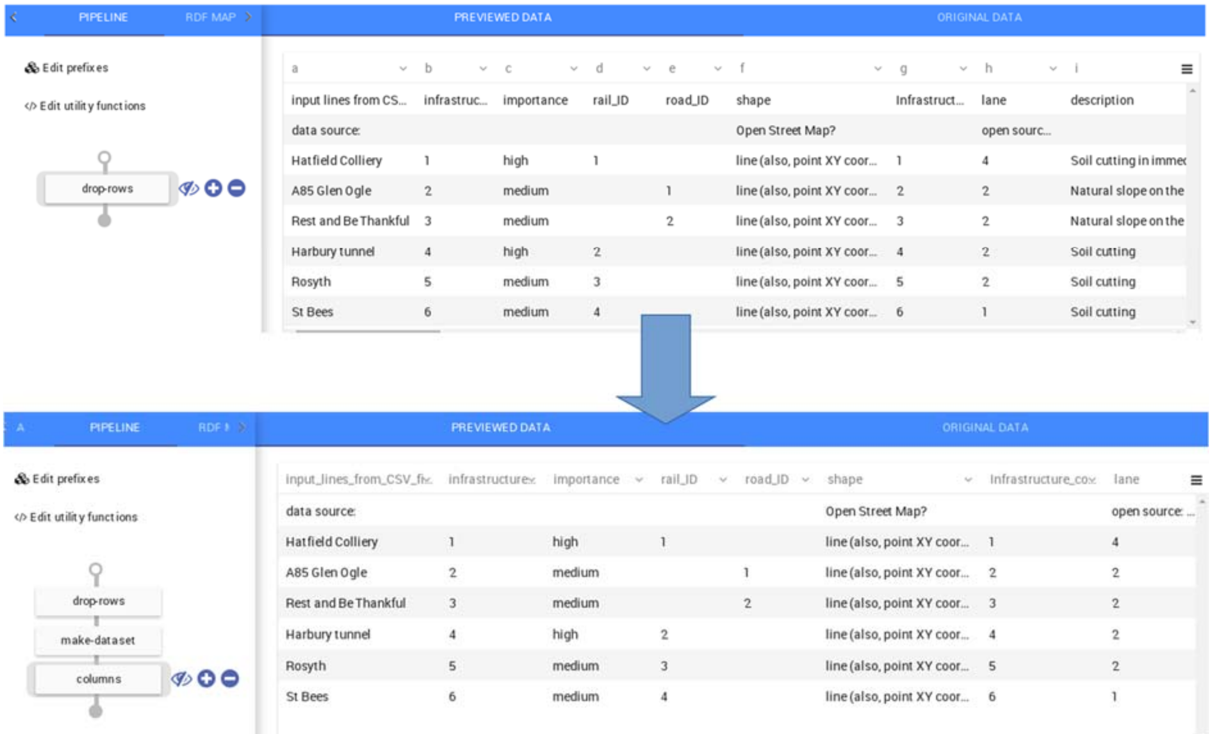

Figure 5-4: Create header from first row and remove unused columns

5.1.2 Mapping functions

After having cleaned up the original tabular dataset we can start specifying the mapping to RDF using the *mapc* function in DataGraft for mapping columns. RDF generation can be done using provided or custom utility functions (see Figure 5-5).

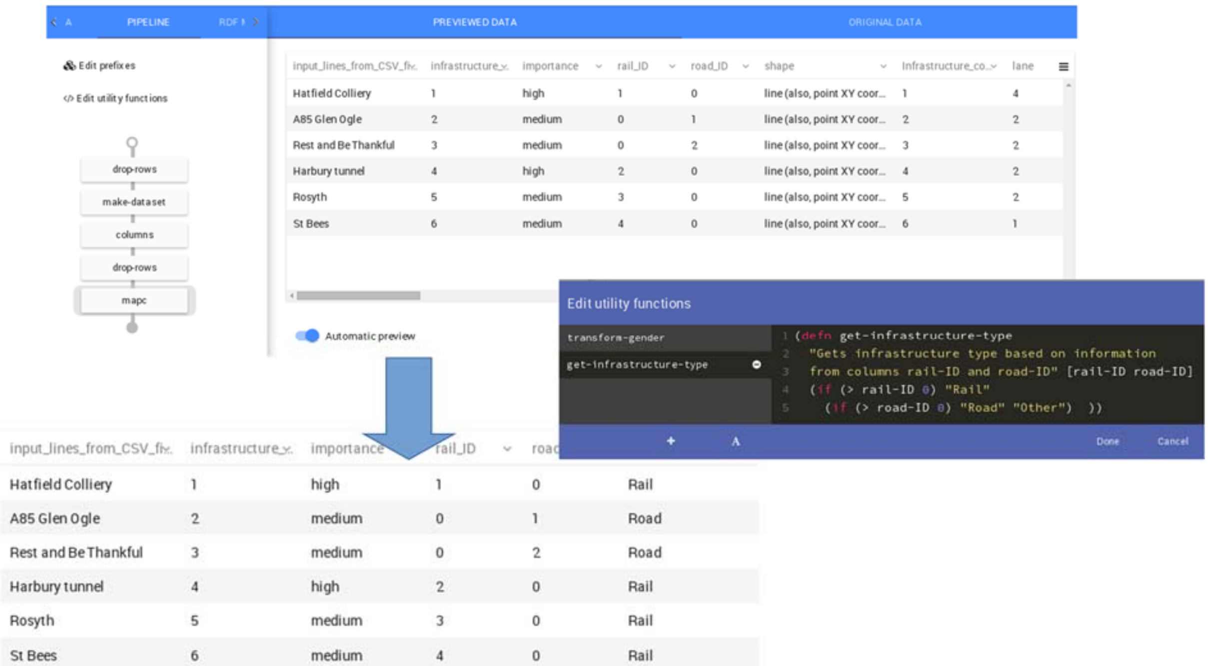

Figure 5-5: Map columns

DataGraft provides a RDF mapping tree view. Figure 5-6 shows the mapping of *infrastructure component* and *landslide*. Similar mappings are defined for all column data in the dataset.

Figure 5-6: RDF mapping in DataGraft

5.1.3 Transforming data

When the pipeline is complete, covering both the data cleaning and RDF mapping tasks, one can run it to transform any source data files (conforming to the original tabular format) into RDF data compliant with the INFRARISK RDF schema. Figure 5-7 shows the complete transformation pipeline for the landslide sample data.

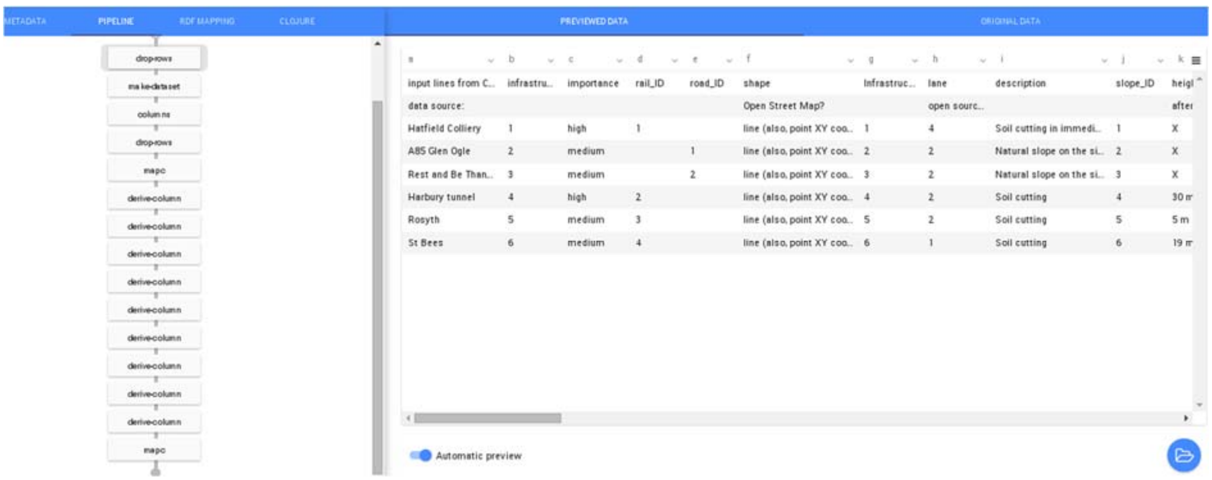

Figure 5-7: Complete transformation pipeline in DataGraft

Figure 5-8 shows the resulting RDF data graph for the landslides sample dataset, with a focused excerpt of the RDF graph on the infrastructure.

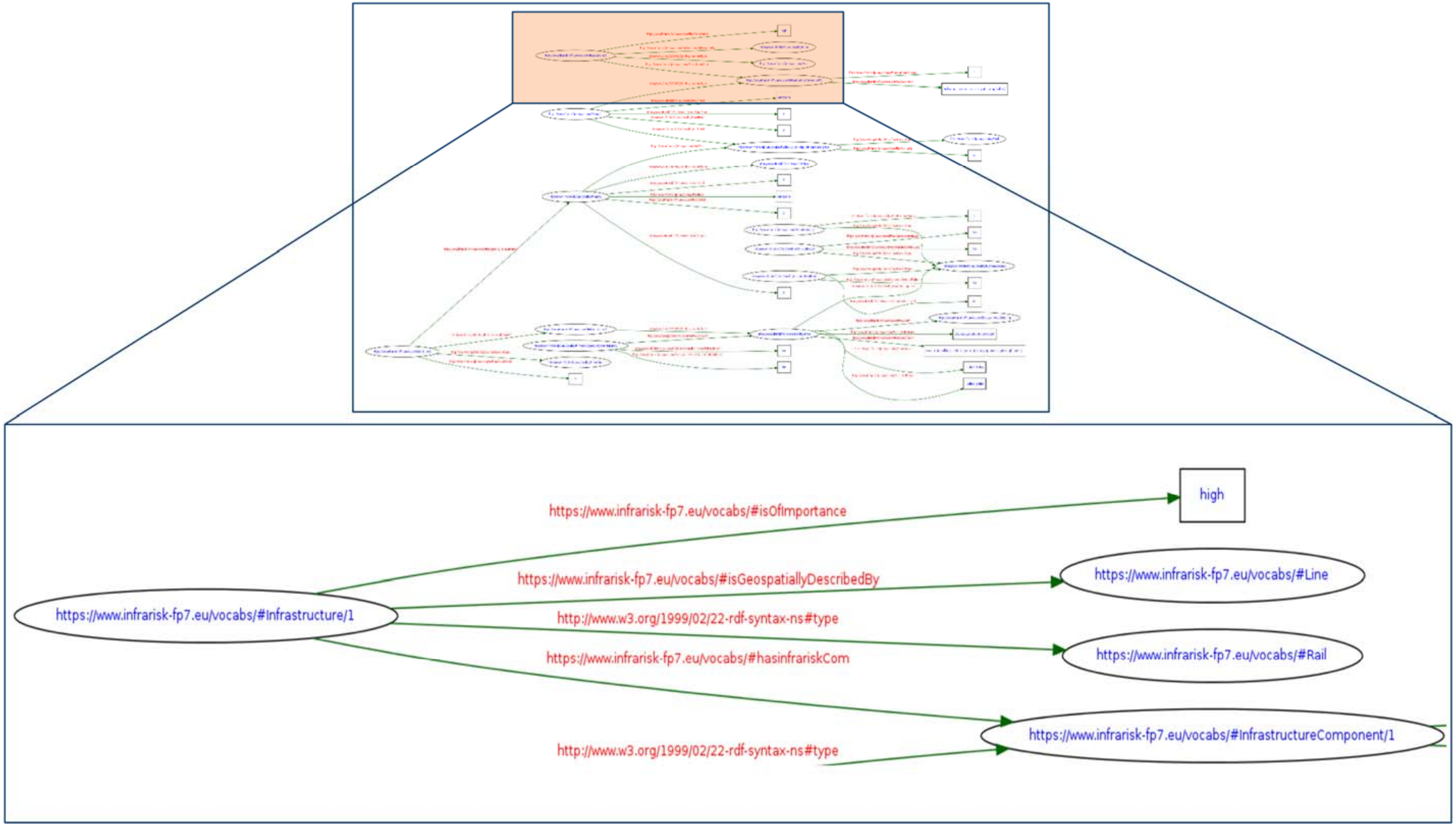

Figure 5-8: Resulting RDF data and excerpt

6 SPARQL QUERIES

Once the data has been transformed one can create a dataset page and query data using the SPARQL query language[9]. The use of SPARQL is primarily targeting data scientists and data-intensive application developers. Section 7 describes the development of a GUI application prototype that illustrates an easier way of navigating and browsing the data for data consumers.

6.1 Landslides sample dataset

Here we use the landslides sample dataset to demonstrate the querying capabilities of DataGraft.

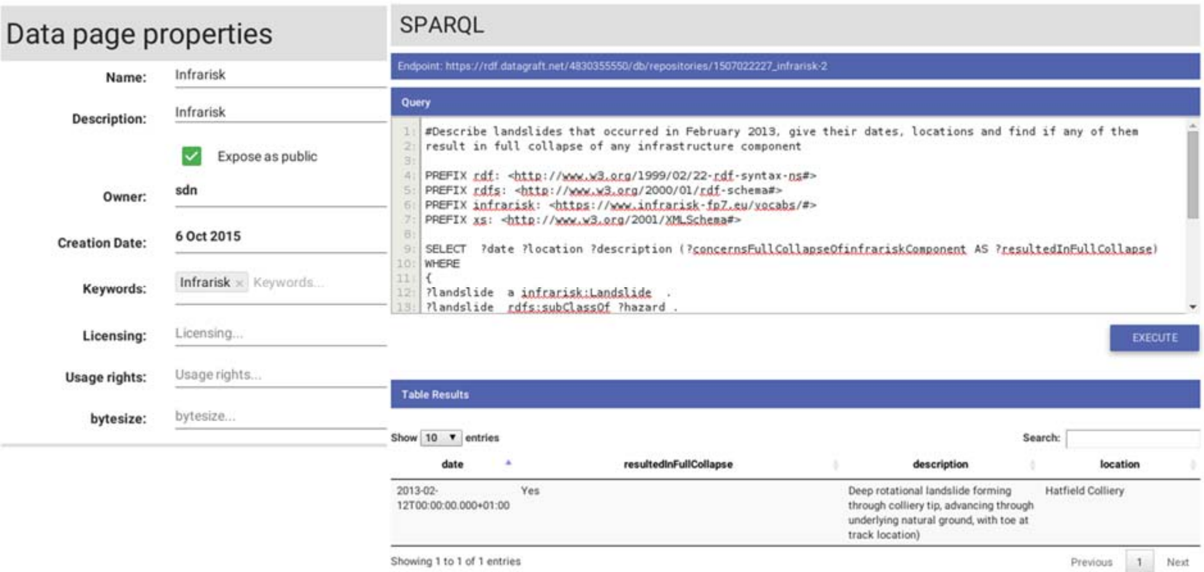

Figure 6-1: Landslides dataset page

6.1.1 Example queries

Table 6-1: Example queries

Query	SPARQL
Describe landslides that occurred in February 2013, give their dates, locations and find if any of them result in full collapse of any infrastructure component	PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX infrarisk: <https://www.infrarisk-fp7.eu/vocabs/#> PREFIX Landslide: <https://www.infrarisk-fp7.eu/vocabs/#Landslide/> PREFIX NaturalHazard: <https://www.infrarisk-fp7.eu/vocabs/#NaturalHazard/> PREFIX xs: <http://www.w3.org/2001/XMLSchema#> SELECT ?date ?location ?description ?concernsFullCollapseOfinfrariskComponent WHERE { ?landslide a infrarisk:Landslide . ?landslide rdfs:subClassOf ?hazard . ?hazard a ?event .

[9] https://www.w3.org/TR/rdf-sparql-query/

	?event infrarisk:occurredOnDate ?date . ?event infrarisk:occurredAtLoc ?location . ?event infrarisk:hasDescription ?description . ?event infrarisk:hadConsequence ?consequence . ?infrastructureComponentFailure rdfs:subClassOf ?event . ?infrastructureComponentFailure infrarisk:concernsFullCollapseOfinfrariskCom ?concernsFullCollapseOfinfrariskComponent FILTER (?date < xs:dateTime("2013-03-01T00:00:00Z") && ?date >= xs:dateTime("2013-02-01T00:00:00Z")) . }

Describe landslides that resulted in displacement of mass containing clay	PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-
	ns#>
	PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
	PREFIX infrarisk: <https://www.infrarisk-
	fp7.eu/vocabs/#>
	PREFIX xs: <http://www.w3.org/2001/XMLSchema#>
	SELECT ?location ?date ?description
	WHERE
	{
	?landslide a infrarisk:Landslide .
	?landslide rdfs:subClassOf ?hazard .
	?hazard a ?event .
	?event infrarisk:occurredOnDate ?date .
	?event infrarisk:occurredAtLoc ?location .
	?event infrarisk:hasDescription ?description .
	?landslide infrarisk:resultedInDisplacementOf ?mass .
	?mass infrarisk:isOfSoil ?soil .
	FILTER regex(STR(?soil),"clay","i") .
	}

7 GUI APPLICATION PROTOTYPE

We have developed a graphical user interface (GUI) application prototype[10] that can be used to show data from the GIS Knowledge Base in a map.

7.1 Graphical user interface (GUI)

This GUI is based on the open source MASTER[11] application developed for the European BRIDGE project[12]. It is a HTML5 application which can be used on smartphones, tablets (see Figure 7-1) and desktop computers.

Figure 7-1: GUI application prototype

In addition to the map view, the application has been integrated with the Google Street View technology. It allows the user to navigate along the roads photographed by Google. This mode provides an interesting alternative for viewing of hazard events (see Figure 7-2).

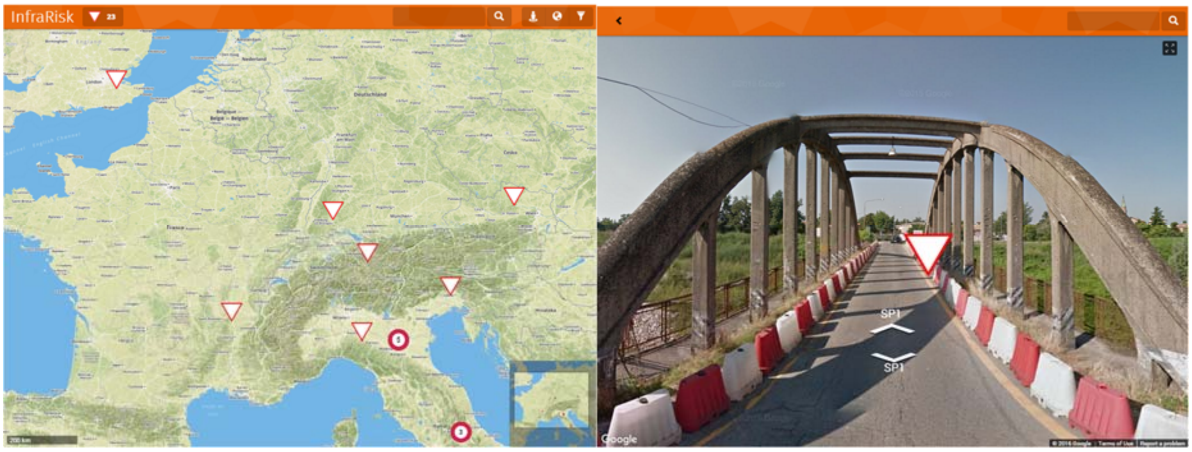

Figure 7-2: Integration with Google Street View

[10] https://infrarisk.datagraft.net/ - supported on Firefox and Chrome. (Support for Internet Explorer expected in the future.)
[11] https://github.com/SINTEF-9012/mobileMaster and https://master-bridge.eu/
[12] http://www.bridgeproject.eu/en

Data about infrastructure components and events that are presented in the GUI (see Figure 7-3) are retrieved from the GIS Knowledge Base. The data is queried from the linked data triple store (used by DataGraft) using SPARQL queries. The output data is formatted using JSON.

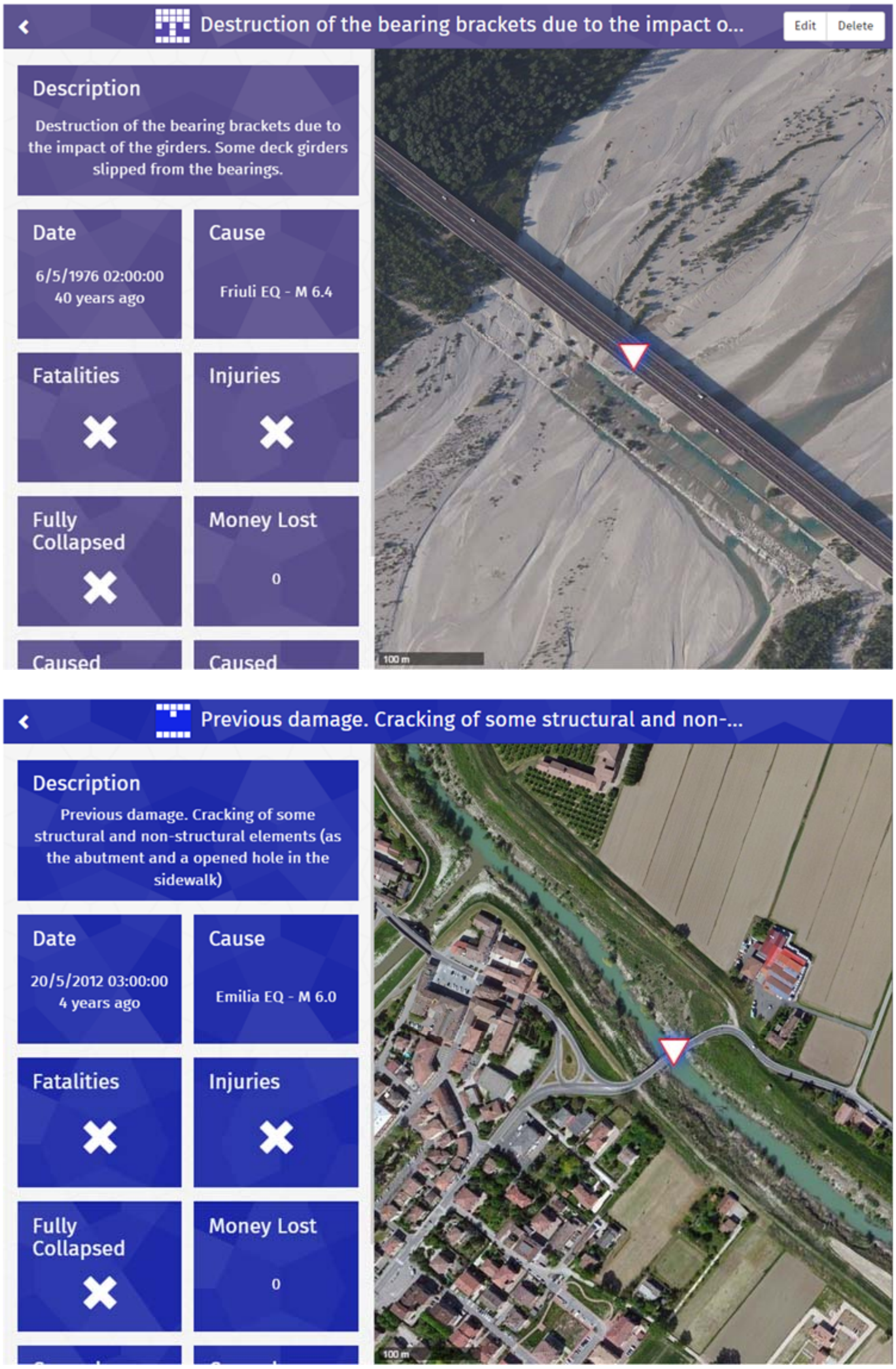

Figure 7-3: Data about infrastructure and events

7.2 Technologies used

The GUI application has been developed using a wide variety of technologies that were selected based on earlier experience and reviews of alternatives. Table 7-1 gives an overview of the technologies that has been used.

Table 7-1: Overview of technologies used in the GUI application prototype

Source	Description
AngularJS	Main framework of the web application
MapBoxAPI	Map provider, based on OpenStreetMap
Leaflet.js	Interactive map library
PruneCluster	Plugin for Leaflet, providing support of large and live datasets
Google Street View	Interactive panoramic views of the streets in the world
Bing Maps APIs	Satellite views
Nokia Here APIs	Geocoding and reverse geocoding
Docker	Software containers framework

8 CONCLUSION

We have set up a technical infrastructure for the **GIS Knowledge Base** based on DataGraft. DataGraft is a cloud-based service for data transformations and data access. We have described the data transformation capabilities, covering data cleaning and mapping functions, and query capabilities of the GIS Knowledge Base using the landslides samples dataset as an example. In addition, we have developed a graphical user interface (GUI) application prototype that can show data from the GIS Knowledge Base in a map.

Future work includes supporting the further population of relevant datasets in the GIS Knowledge Base. This can be done in two ways:

- Providing data spreadsheets with similar structure as those described in Section 4.1. (This can be done by providing the spreadsheets and running the existing transformations directly on DataGraft.) This method is targeting users with domain knowledge and familiar with spreadsheet data.
- Uploading RDF data directly to the GIS Knowledge Base. This can be done via SPARQL update queries. This method is targeting data workers familiar with Linked Data technologies.

ACKNOWLEDGEMENTS

The primary authors would like to acknowledge the support and contribution of other persons (partners) that have contributed to Task T2.2 and this deliverable: Pierre Gehl (UCL), Khaled Taalab (UCL), Pieter van Gelder (PSCT), Yuliya Avdeeva (PSCT), Maria Jose Jimenez (CSIC), Mariano Garcia-Fernandez (CSIC), Bryan T. Adey (ETH), Miguel José Segarra Martínez (Dragados) and Mark Tucker (ROD).