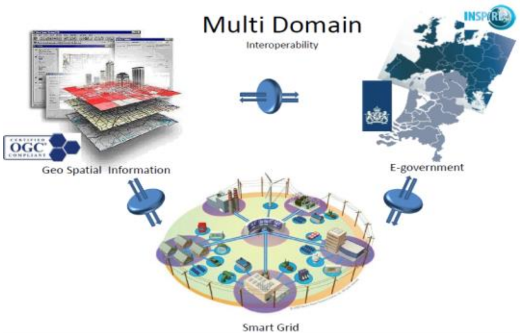


CERISE

Combining Energy and Spatial Information Standards as Enabler for Smart Grids

TKI Smart Grid Project: TKISG01010

D4.2 Linked Data for Crisis management

Werkpakket – 40

Lead partner: TNO

2 September 2015

Versie 1.0 - Final

DOCUMENT INFORMATION	
ID	D4.2 Linked Data for Crisis management
Work package	WP40 Harmonisation
Type	Report
Dissemination	Public
Version	1.0 - Final
Date	2 September 2015
Author(s)	Ramona Roller (TNO) and Jasper Roes (TNO)
Reviewer(s)	Jan Bruinenberg (Alliander)

Table of contents

Summary

The CERISE-SG project (Combining Energy and Geo-information standards as enabler for Smart Grids) focuses on improving interoperability between the geo-, utility-, and e-government domain by connecting sector-specific data definitions and by providing sector-overarching standards for information exchange.

This document presents the outcome of a study into the use of Linked Data during crisis management. It focusses on the use of data in crisis management.

During crisis management sharing of data becomes important when electrical (smart) grids are hit by a power cut. As the electrical grid is important for many parties and societal activities, such as healthcare, schooling, and transport, lots of stakeholders from different sectors are involved in crisis management. For instance electrical grid operators, water boards, and safety regions.

Based on the use case elaborated in deliverable D2.2 a data sharing solution has been developed. The aim of creating a data sharing solution for crisis management was to examine whether Linked Data improves interoperability between the data sharing parties. Results revealed that Linked Data in general and the crisis management ontology in particular provide a very good method to improve interoperability providing that Linked Data will be used more widely by more network operators and water boards.

The study revealed six major benefits of Linked Data compared to conventional technologies when used in the crisis scenario:

- generation of new knowledge,
- handling of dynamic content,
- flexibility,
- compliance to external standards,
- completeness and precision,
- data-centric approach instead of application-centric approach.

Drawbacks include the incompatibility with most existing data management systems as well as privacy and security issues (which are not different of those for conventional technologies). Within the crisis management ontology concepts of electrical supply area and water resistance threshold do not fit exactly to the crisis scenario. Data about electrical supply areas, polders and water levels was inexact or missing.

The suitability of Linked Data should be further tested for other aspects of crisis management. The development of a crisis could be predicted by interlinking domino-effects with response options and a temporal component. The electrical grid might be reconfigured faster and more easily if data about distributed power generating objects were interlinked (e.g. PV panels). Moreover, emergency responses could become more efficient and effective if the individual sub-processes were linked to each other in a meaningful way by mapping them onto an ontology.

Introduction

The project CERISE-SG (Combining Energy and Spatial Information Standards as Enabler for Smart Grids) focuses on interoperability between the geo-, utility and e-government domains, by establishing information links between smart grids and their environment (CERISE, 2015). Obtaining reliable information from various sources is invaluable in order to account for the increasing distributed and dynamic structure of energy management, especially in a crisis situation (Book, Bastiaans & Bruinenberg, 2014).

However, data sharing is complicated since parties use different definitions supported by different standards for their data to manage. For example, HHNK, a regional Dutch water authority, uses the term “kunstwerken” when referring to infrastructural engineering constructions, such as bridges (de Landmeter & van Giessel, 2015). A flood-management novice, however, rather uses terms like “bridge” or “road” to refer to the same concept. In order to help parties share and use this information, the CERISE project seeks to analyze currently used standards, identify as well as overcome gaps, overlaps, and inconsistencies, and finally link proven standards to create a coherent data sharing environment.

Deliverable D2.2 showed that these goals are highly important for crisis management. A use case scenario was developed where two primary stakeholders, Alliander and HHNK, exchange information about water levels and their respective assets (e.g. power stations and water pumping stations) in case of a flooding. Thereby, they would like to localize the affected assets, determine which supply areas are hit by a power cut, and reconfigure the power grid accordingly. Their aims are to re-ensure power supply and reduce the flooding.

Various data sharing technologies can help to achieve these aims, one of them being Linked Data. This report represents CERISE deliverable D4.2 of work package 40 on data harmonization. It will examine the suitability of Linked Data to improve crisis management through smart data. Linked Data describes a method of interlinking data points and data concepts in a structured way on the web (Eckartz & Folmer, 2014). The data can then be read automatically by computers which enables data from different sources to be connected and queried to derive information encoded in the interlinks. Data links will be structured according to a *conceptual model* that defines relevant data concepts and their interrelations.

The target audience of this document is the CERISE team itself and associated committees such as the steering group and the sounding board group. Moreover, all stakeholders of this use case who share and provide data, as well as parties involved and/or interested in Linked Data, data management and crisis management in general are addressed, too. Since this report is highly based on deliverable D2.2 the authors recommend reading D2.2 beforehand.

This report includes six chapters and seven appendices. Chapter 1 describes the crisis management scenario and previous work on data sharing systems for crisis management. Chapter 2 describes the technology of Linked Data and the application within crisis management. Chapter 3 presents the crisis management ontology and its analysis. Chapter 4 describes the data sets and their analysis. Chapter 5 provides the Linked Data analysis and chapter 6 provides the conclusions, discusses the work and describes future work. The appendices describe a hand-on session on Linked Data, a snap shot of the final datasets and five tutorials.

1 Crisis management scenario

1.1 Crisis management scenario

Alliander and HHNK want to share data during a flood-caused power cut. This crisis scenario is described in more depth in deliverable D 2.2. Figure 1 presents the development of the crisis. As a result of a flooding, power station 1 breaks down and therefore can no longer supply electricity to the users in its supply area. Thus, the affected electrical supply area of that flooded power station and the electricity users within it have to be identified. Electrical water pumping station 3 is located in the affected supply area and consequently breaks down, too. Now, counter-measures against the flooding have to be taken, for example by replacing the broken with a functioning pumping station, which involves a reconfiguration of the pumping station grid. Electrical water pumping station 4 is located in a non-affected supply area. If it is close by, it can be used to drain the flooded area, otherwise, a battery or a generator can be used for temporary power supply. In the long run, the electricity grid has to be reconfigured in order to supply electricity to the most vulnerable users (e.g. hospitals, electrical water pumping stations). Once the flood is gone, damaged power assets and pumping stations have to be repaired in order to be put back on the grid.

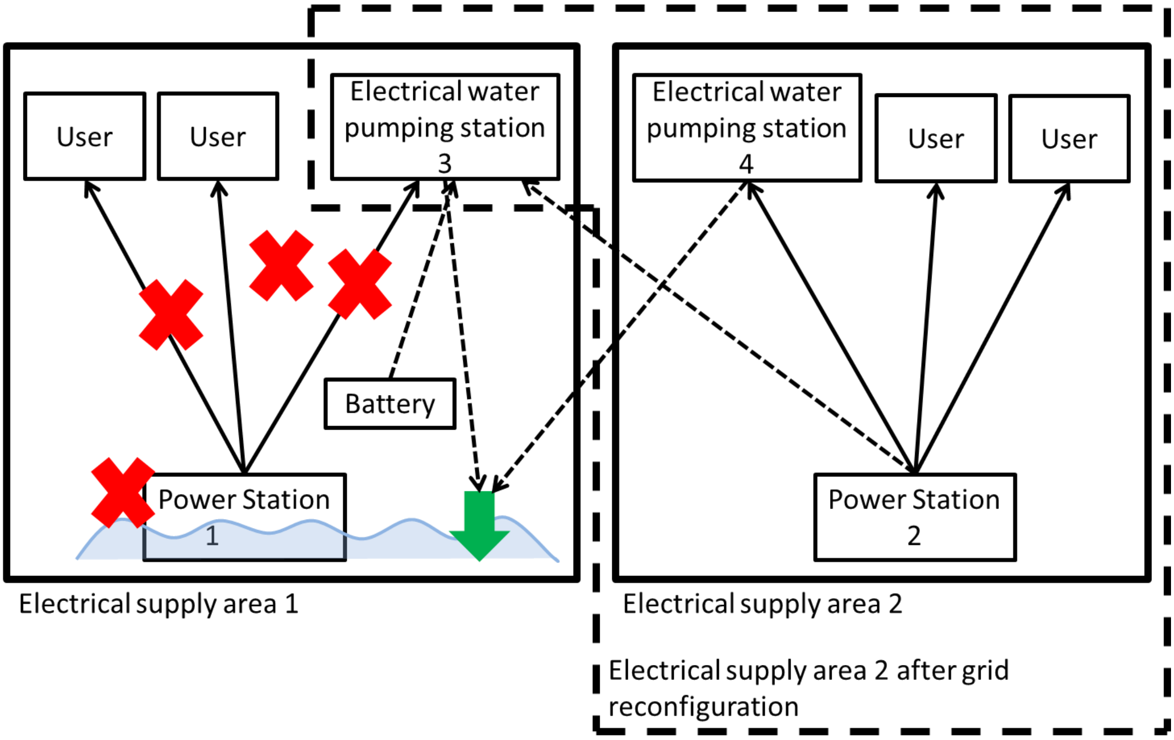


Figure 1: Concepts and processes within the crisis scenario of flood-caused power cut

1.2 Scope

The scope of this data sharing solution is based on stakeholders' requirements as well as organizational constraints, such as the availability of time (Table 1). This provides a clearly defined, but also highly limited scenario. However, for this exploratory pilot study it provides a sufficient thinking framework.

Table 1: Scope of data sharing scenario

Scope of data sharing scenario
Functionality purpose - Identify assets at risk and damaged ones. - Localize affected supply area. - Deal with real-time water level heights.
Durability purpose - Provide future-proof data sharing solution. - Generalize data sharing solution to other domains (e.g. gas). - Data sharing solution should be adopted by more stakeholders.
Small user community Only Alliander and HHNK use data sharing solution. It is therefore easier to adjust it to the stakeholders' IT landscape than the other way round.
Closed application environment Alliander and HHNK only share data with each other. No open data wanted.
Economical purpose Data sharing solution should impose little extra costs on the stakeholders.
Limited time available - Data sharing process has not been running for a sufficiently long period. - Implementation does not exist. - Assessment process must not be too time consuming.
User-centred development Development of data sharing solution is done in close cooperation with Alliander and HHNK and based on their explicit requirements.

1.3 Previous work

Currently, there is a large diversity of crisis management systems in place in the Netherlands. Traditional channels like email or phone are still used although they are time costly and error prone. For example, Alliander and HHNK currently exchange data via USB sticks (Book, Bastiaans & Bruinenberg, 2014). Besides, IT systems such as Eagle One, MultiTeam, and LCMS[1] 2.0 increase in usage (Fan & Zlatanova, 2010; Esri, 2015). In case of a crisis, the safety region determines affected stakeholders and asks them to provide data on their area of responsibility. Stakeholders then prepare their information locally in text plots or map format, which are then layered-up within the central crisis management system (van Dongen et al., 2013). The resulting mash-ups provide the same picture of the crisis to all stakeholders involved, who therefore seem to better cooperate since they all have the same basis of decision-making.

However, these systems do not take crisis-specific requirements for data sharing into account. First, by informing stakeholders only during the crisis about needed data, valuable time is lost (van Dongen, 2013). Each crisis requires different data from one stakeholder, who therefore has to do tedious preparatory work for each crisis. For example, in order to localize broken water pumping stations, a power cut only requires data on electrical pumping stations whereas a flooding requires data on both mechanical and electrical pumping stations.

Second, by preparing their data locally without interaction, stakeholders run the risk of using sector-specific descriptions for their information. These might not be understood by stakeholders from other domains which decreases semantic interoperability. For instance, the HHNK uses the term "kunstwerken" when referring to infrastructural engineering constructions, such as bridges (de Landmeter & van Giessel, 2015). A flood-management novice, however, would probably interpret this label to represent some "art work".

[1] Landelijk crisis management system. All safety regions use this system for data sharing during crises.

Third, since the course of a crisis is often unpredictable, data requirements and the constellation of involved stakeholders might suddenly change. This might lead to an overload of the central crisis management system as it has to integrate constantly changing datasets (Zlatanova et al., 2014).

Fourth, current systems access and display all the data available not differentiating between stakeholders' roles, which require different kinds of information (Zlatanova et al., 2014). This may create an information overload since the human cognitive processing capacity is limited (Endres et al., 2015). For instance, during a flood-caused power cut, the electrical grid operator only has to know the location of affected electricity consumers in order to re-ensure power supply. Buildings damaged by water alone that are still supplied with electricity are irrelevant for this stakeholder.

This deliverable addresses these issues with a novel data sharing solution based on Linked Data. Stakeholders specify the datasets they are willing to share, which are then interlinked in a structured way. This happens *before* the crisis, in order to fully focus on data sharing - not on preparation - during the crisis. Semantic interoperability will be enforced by linking the data to a conceptual model that provides formal definitions of data concepts and defines their interrelations. This model can then be linked to local models of the individual stakeholders to provide a mapping cipher for translating concept definitions from one domain to another. Finally, depending on the crisis situation and the stakeholders' information needs, the data can be specifically queried in order to retrieve relevant information.

To this point, Linked Data has only been sparsely used within crisis management in the Netherlands. The only case known to the authors is a sector-specific application within the fire-fighting domain. The firebrigade of Amsterdam-Amstelland would like to share their data about incidents and fire safety using Linked Open data (van Leeuwen, 2015). Published Data will be linked to an open twitter account, the incident room, and *firebrary*, a conceptual model holding data about fire-fighting terms (Firebrary, 2015). The intention is to inform bypassers how to behave correctly during a fire-related crisis and also to encourage them to inform the firebrigade about new disasters (van Leeuwen, 2015).

This deliverable aims to extend this work by applying Linked Data to cross-sector crisis situations and by assessing the suitability of a novel conceptual model for crisis management

2 Linked Data

2.1 In general

Linked Data is a way to publish and share data on the Web (Eckartz & Folmer, 2014). Each data element becomes a resource on the Web with named links to other data elements and properties. The basic element of this structure is a *triple* consisting of two data resources, *subject* and *object*, whose relation is specified by a *predicate*. For example, the project DBpedia transforms information from Wikipedia into Linked Data and publishes it on the web (Lehman et al., 2015). Structured information from the Wikipedia "infoboxes" containing data on geo-coordinates, dates, categorizations, etc. is extracted and put into a uniform dataset which can be queried. Thus, a data point with the ID "The Hague" (subject) can be linked to the concept "the Netherlands" (object 1) via "isCapitalOf" (predicate 1) as well as to the geo-coordinates 52°5′N, 4°19′E (object 2) via "hasGeoCoordinates" (predicate 2). Consequently, we can retrieve the location and the name of the capital of the Netherlands. As this network gets extended objects become subjects if they are linked to a new object via a predicate (e.g. The Netherlands isOfType Country) (Figure 2).

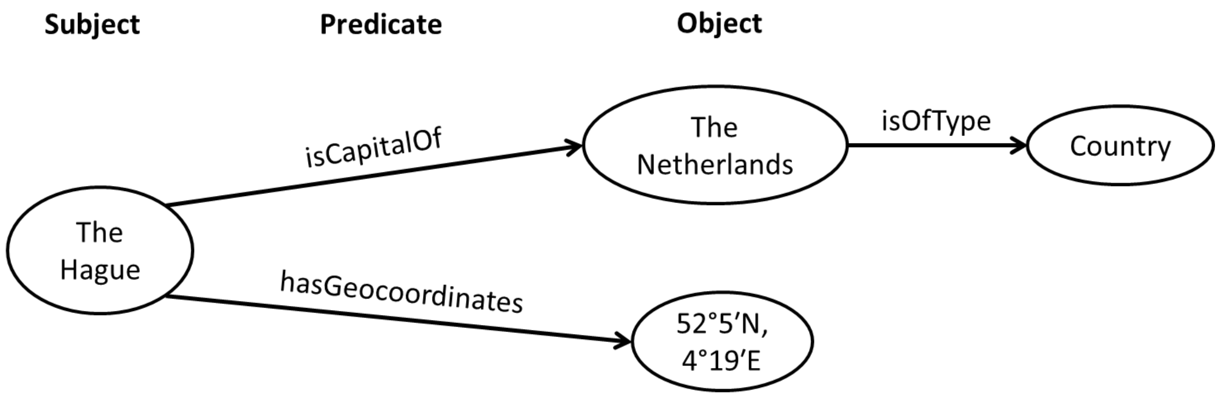


Figure 2: Example of triples

In order to describe the semantics of the data, links are provided to *vocabularies* (also called *ontologies* or *conceptual models*). These are models formalizing common conceptualizations for a specific domain, e.g. within the electricity domain concepts such as voltage or transformer are described within an electricity ontology. Data about IDs and the actual voltage level of different transformers can be linked to these concepts, so that it gets clear which transformer has which voltage level.

By publishing data as Linked Data and annotating it with well-known vocabularies, it becomes much easier for others to understand and use these data. In principle it is possible to annotate the same data with different vocabularies, thus making it possible to 'view' the data from different 'perspectives'.

2.2 Within crisis management

Using Linked Data within the crisis scenario means that data stay at their source (Figure 3). Links are only created when the data is queried to retrieve information, such as the location of broken assets. This information is then released to the query source. Data owners can restrict data access by allowing other stakeholders to only ask pre-specified SPARQL queries. For example, Alliander can allow HHNK to ask which areas are without power but not where electrical assets are located. These restrictions can be done via APIs and allow relevant data to be linked but not to be completely opened and can also solve privacy and security issues.

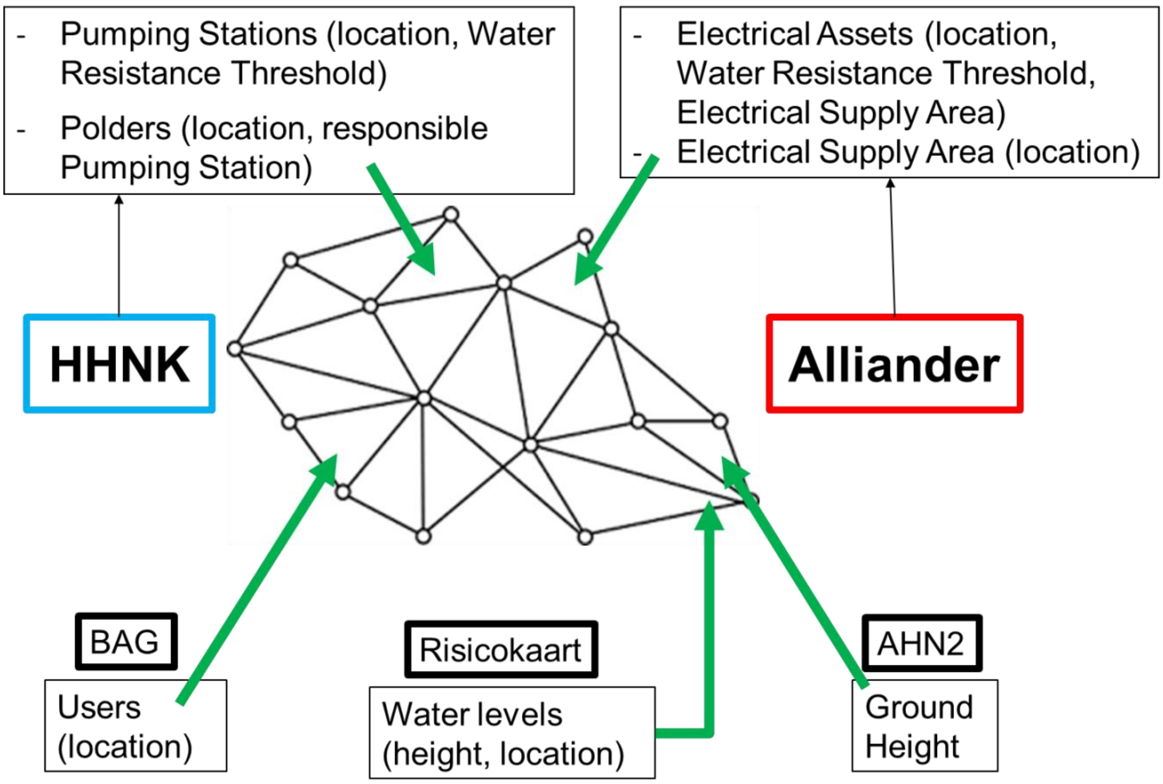


Figure 3: Information exchange between Alliander and HHNK during a flood-caused power cut. Support is provided by secondary data sources

Key concepts of the crisis scenario such as electrical assets, pumping stations, and users (consume electricity) will be interlinked via predicates and also assigned attributes such as name and location. Actual data points will be matched to these attributes. Via predicates they can be combined to answer pressing questions during the crisis (Table 1). In order to answer these questions, those triples describing the required links have to be selected. This triple selection is called a *path* and can be expressed with the query language SPARQL (see Appendix A).

Table 2: Questions in the crisis scenario

Question during the crisis	Required data concepts	Reasoning
Where is the flood?	Current water level and normal water level for the same location.	A point is flooded if the real-time water level is greater than the water level of the normal state at

		that location
Which pumping station is in the vicinity to drain the flooded area?	Location of pumping stations and of flood	Calculate distances between flood location and location of pumping station. Rank the distances.
Which electrical assets are broken down due to the flood?	Water resistance threshold of an electrical asset and the ground height at its location. Retrieve the current water level at this location.	If threshold + ground height < water level then the electrical asset will break down.
Which users are affected by the power cut?	Retrieve location of affected electrical asset and location of its electrical supply area. Retrieve location of users.	If location of electrical supply area = location of user this user will be without power.
Which power station is in the vicinity to replace the broken one?	Location of all electrical assets (broken and functioning).	Calculate distances between broken electrical asset and the functioning ones. Rank the distances.

3 Crisis management ontology

3.1 The ontology

Based on the scope of the data sharing solution deliverable D2.2 identified relevant domain concepts for the crisis management use case and interrelated them in a conceptual model (Figure 4). *Object properties* link domain concepts with each other whereas datatype properties link domain concepts to data points which can be of different types (e.g. string, float, integer).

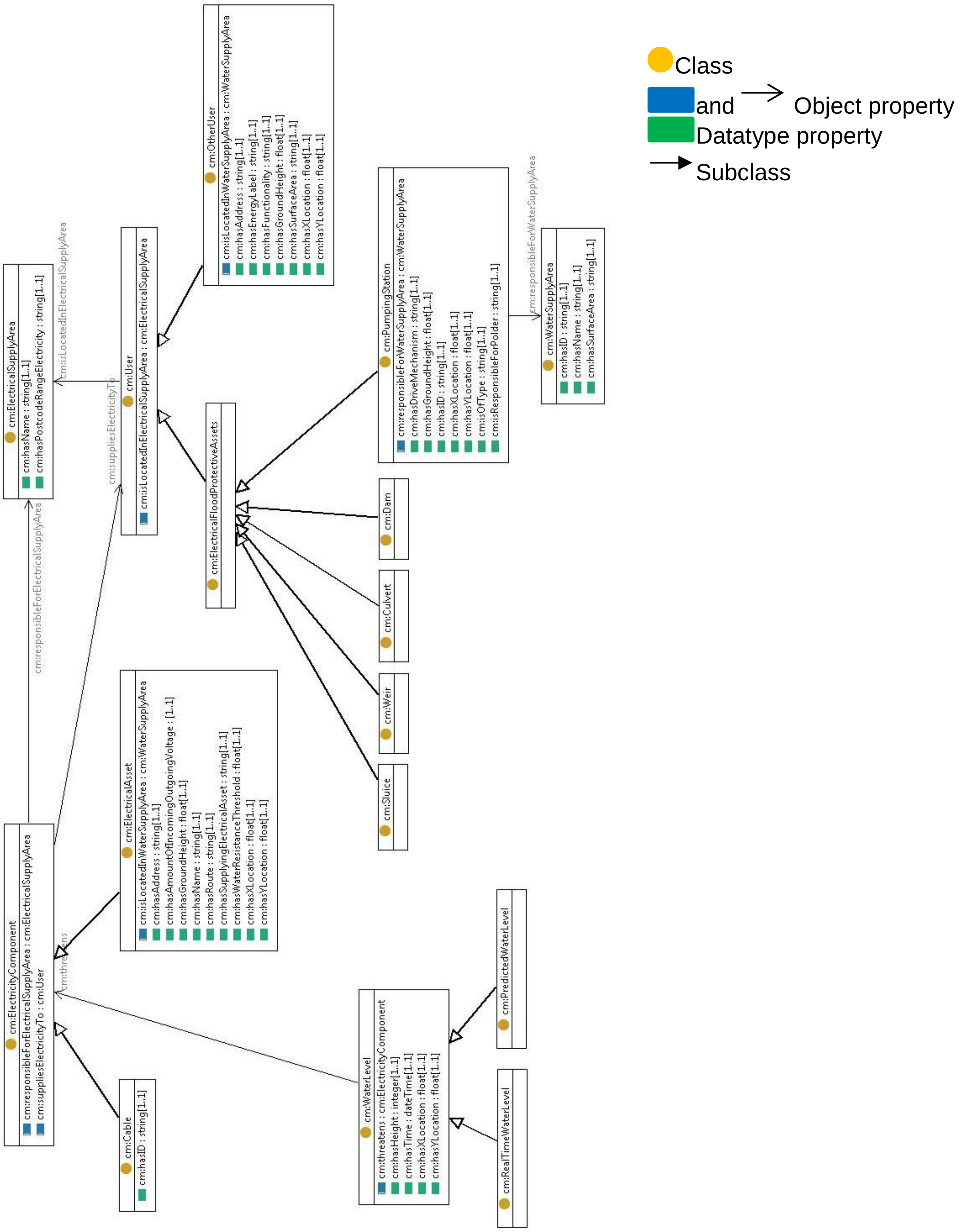


Figure 4: Conceptual model visualizing relationships between domain concepts

3.2 Crisis management ontology analysis

First, the concept of a postcode-based electrical supply area cannot be generally applied to the whole crisis area. Buildings in relatively young towns and cities such as Heerhugowaard have been constructed within a short time period. As a consequence electrical supply paths were set-up at the same point in time. This allowed for highly structured connections between houses and electrical assets so that supplying electrical assets have the minimal distance possible to the addresses they supply power to. This leads to coherent electrical supply areas around the electrical asset (Figure 5).

In contrast, older places like Amsterdam have been constructed over longer periods, meaning that age differences of buildings are higher. Therefore power supply is set-up in a heterogeneous manner. Addresses receiving power from the same electrical asset are spread all over the place and do not form a coherent are that can be divided according to postcodes (Figure 6). Claiming that the electrical supply area of electrical asset 1 has a postcode range from 1704CC to 1704DR is misleading since there might be only one address that lies within one postcode unit (e.g. 1704DA).

It would therefore make more sense to define the electrical supply area of an electrical asset by the exact addresses that are powered by this asset. Data about these connections is available and can be used when privacy issues are respected.

Second, the concept of Water Resistance Threshold does not add true meaning to the crisis management ontology since these data are not tracked by the stakeholders. Alliander provided an estimate of 30 cm for all their electrical assets, but this seems to be too general. In order to increase the power of this concept the stakeholders are advised to track this data in the future for all of their assets.

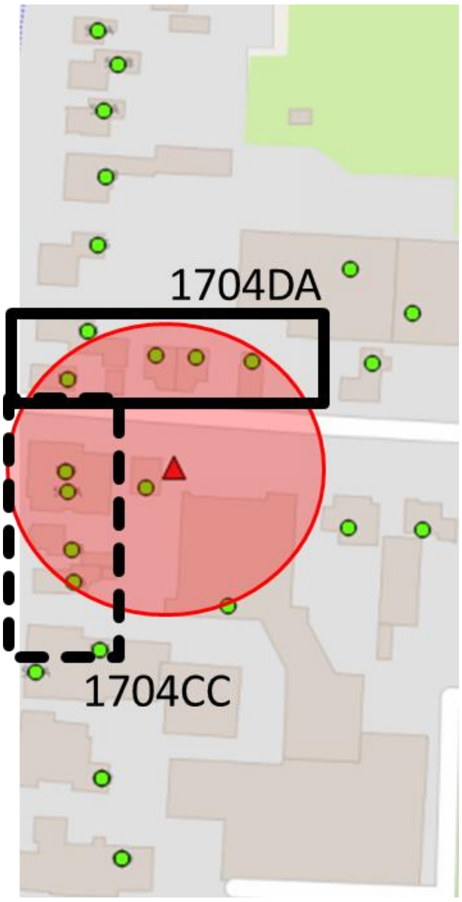


Figure 5: Electrical supply area for homogenously built towns

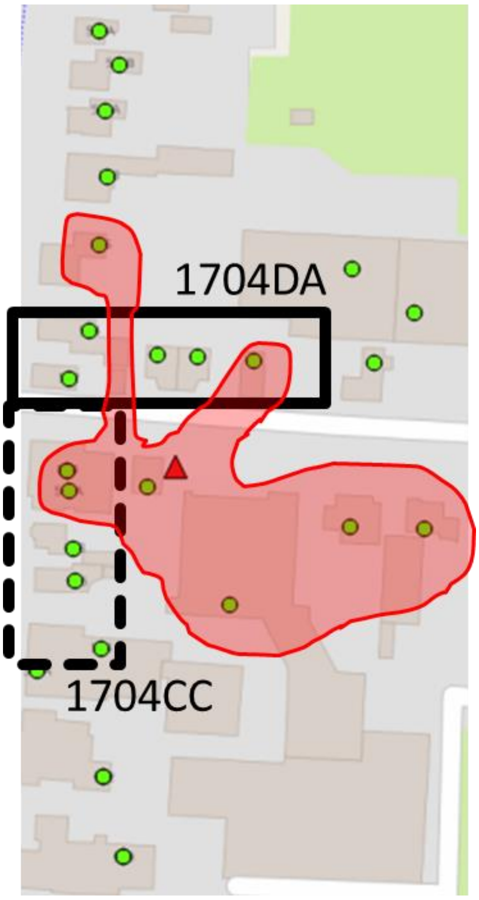


Figure 6: Electrical supply area of heterogeneously built towns

4 Data: from raw to ready-to-use

4.1 Overview

Data sets for this crisis management use case came from four sources: Alliander, HHNK, BAG, AHN2. A summary of the data sets is provided in Table 3.

Table 3: Data sets used for the crisis management use case

	Alliander	**HHNK**	**BAG**	**AHN2**	**Risicokaart**
Content	Characteristics of middle-voltage stations that are located in Holland’s Noorderkwartier.	Characteristics of pumping stations, sluices, dams, weirs and polders that are located in Holland’s Noorderkwartier.	Addresses and function of buildings.	Topographic data indicating the height of a point on the ground surface above sea level.	Overground water height levels.
Format	Shape file	Shape file	Shape file	Shape file	wcs, wms
Programs used for data preparation	qGIS, Excel	qGIS, PostGIS (to spot flaws), Excel	qGIS, Excel	qGIS, Excel	qGIS
Open	No	No	Yes	Yes	Yes
Source	Alliander Private communication	HHNK Private communication	Vrije Universiteit Amsterdam verblijfsobject gebruiksdoel http://geoplaza.vu.nl/ cdm/singleitem/collection/gpz/ id/238#	Pdok AHN2 5m resolution, unfiltered point-cloud data http://geodata. nationaalgeoregister.nl/ahn2/ atom/ahn2_uitgefilterd.xml	Risicokaart Nederland https://data.overheid.nl/data/dataset/waterdiepte-in-risicogebied-risicokaart-nl
Remark	Long data request process.	Location of data points not in line with location of buildings on map.	User-friendly web-interface	Tedious selection process.	Processing in qGIS failed. No water level data received.

4.2 Alliander data

4.2.1 The electrical grid

The electrical grid consists of power-generating and power-transforming stations as well as cables that interlink these stations. Power-generating stations are for example nuclear power stations, wind parks or coal power stations. Of interest for this report are the transformer stations which change the voltage of the incoming current in order to facilitate its further transport. In the Netherlands high-voltage (HV) stations are managed by TenneT. These work within ranges of 380kV to 150kV and supply electricity to middle-voltage (MV) stations, which are operated by Alliander. MV stations work within ranges of 50kV to 220 Volt. Figure 7 shows how the stations with their voltage levels are interlinked.

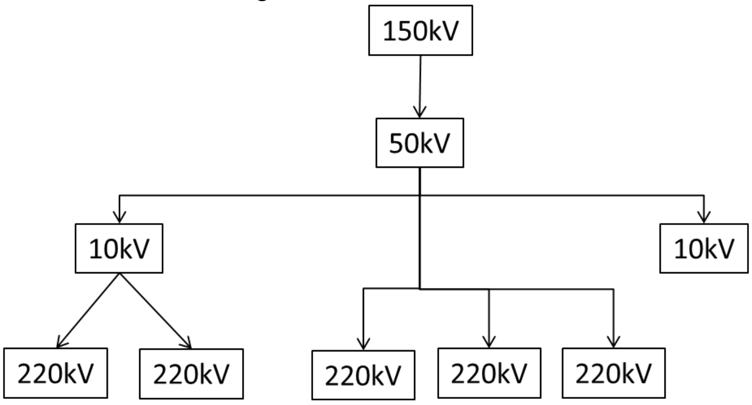


Figure 7: Interlinkage of different transformer stations with different voltage levels.

Electrical stations are connected in a distributed manner, i.e. one station is fed by several stations (Aartsen, Bruinenberg & van Gelder, 2015). This makes the network more robust since broken supplying stations can be replaced by other stations that feed the same lower voltage station. Due to this flexibility connections between stations are not fixed but constantly change in order to adapt to the current situation of the grid. In contrast, if stations are connected in series, like a waterfall, one breakdown causes all the following stations to fall out as well. Figure 8 presents an extract of the electrical grid in the area of Heerhugowaard showing how stations are linked in a distributed manner.

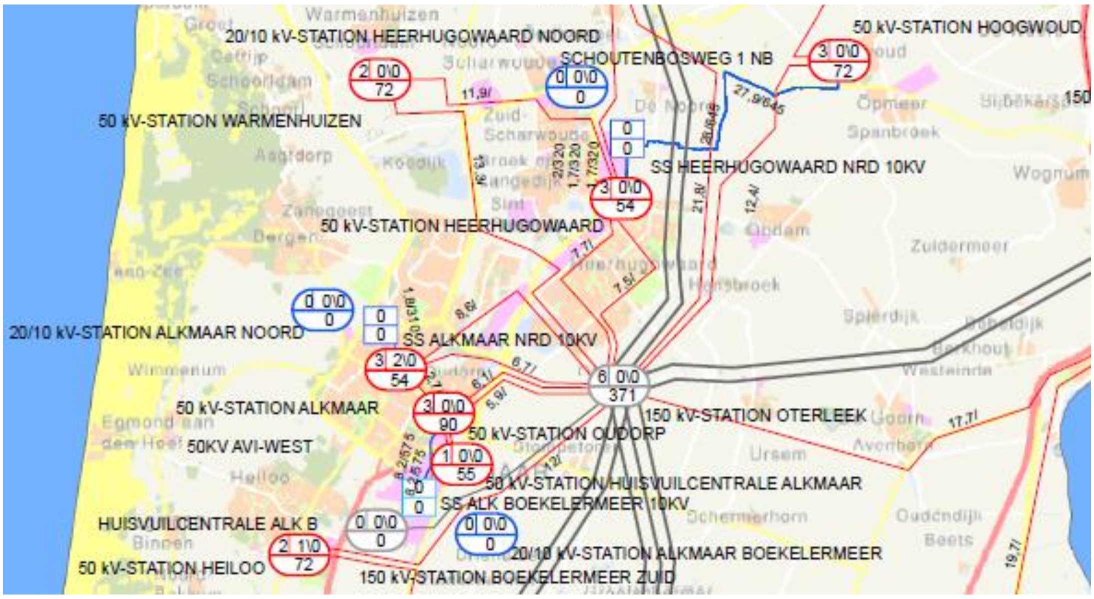


Figure 8: Electrical grid around Heerhugowaard

Supplying MV stations (mostly 10kV) provide current to *routes*. These are groups of smaller stations (400V-220V) from which electricity is directly transported to the end-user. Depending on the current grid configuation supplying stations choose different stations of the route and switch between them. These route stations have an electrical supply area of about 100m radius around them (Aartsen, Bruinenberg and van Gelder, 2015).

4.2.2 The raw data

Alliander provided data about their MV stations (50kV to 220V) with the attributes listed in Table 4. They have been grouped into categories by the authors to facilitate reading.
In total, data for 6929 transformer stations was provided.

Table 4: Alliander raw data

Number	Category and attributes
1	Identification 1. Name of station 2. Case number of building 3. Old number of building
2	Location 1. Address (street name, house number, postcode, town, municipality) 2. X- and y coordinates 3. Geographic region of location
3	Electrical supply area 1. Route: Group of stations that receive electricity from the same supplying station. 2. Supplying HV station

4.2.2.1 Data preparation

Based on the scope of the data sharing scenario (see section 1.2) data were taken from the categories identification, location and electrical supply area. Attributes that were used in the raw data format included: Name of station, Address (street name, house number, postcode, town, municipality) and X- and y coordinates. The electrical supply areas and the ground height of a station had to be calculated from the raw data.

4.2.2.2 Calculating electrical supply areas

Each station in the raw data set was assigned a supplying station and a route where it belonged to. From this information end users within a radius of 100 meters had to be identified. Due to time constraints the electrical supply area was only calculated for the route HHWN 01 SS. It includes 17 stations which are located in the northern part of the town Heerhugowaard and are supplied by the 10kV station SS Heerhugowaard NRD. Figure 9 shows how route HHWN 01 SS is embedded in the electrical grid. Appendix C provides a tutorial for calculating the Electrical Supply Area in qGIS.

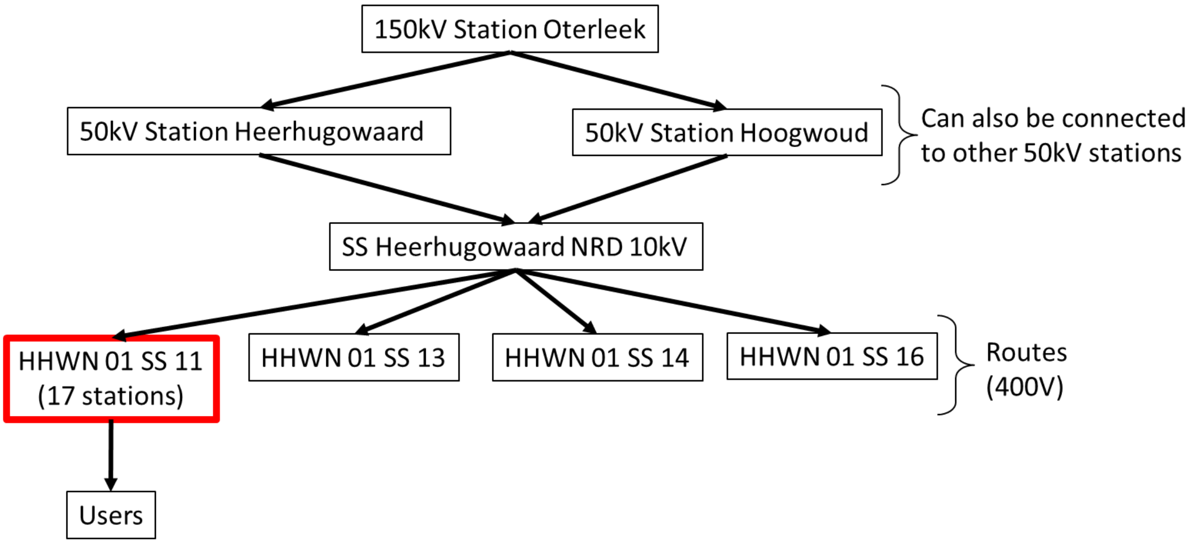


Figure 9: Hierarchical structure of the electrical grid in the area of Heerhugowaard

4.2.2.3 Calculating GroundHeight for each electrical assets

The GroundHeight of an electrical asset refers to how high the ground surface lies above sea level at the location of an electrical asset. These data can be found in the AHN2. Appendix D provides a tutorial to calculate the Ground Height of Electrical assets in qGIS.

4.2.2.4 The final data set

Appendix B shows extracts of the final data sets for Electrical assets and Electrical supply Area. Data for the Water Resistance Threshold (WRT) of these Assets were added later and are based on an educated guess by Alliander (Aartsen, 2015).

4.3 HHNK data

4.3.1 Water facilities

The facilities managed by the HHNK include pumping stations, polders, dams, weirs, culverts and sluices. Dams are barriers to impound water. Weirs are barriers across rivers to alter their flow characteristics. Culverts are structures that allow water to pass under roads or a rail tracks. Sluices are water channels to control the flow of water via a gate. Due to constraint of time and complexity, data about these facilities were not included in the current use case but might well be in a future project.
Of interest here are pumping stations which drain a polder in case it is flooded. A polder is a low-lying tract of land enclosed by barriers (dikes) that cut it off from outside water. One polder can be assigned to one or more pumping stations.
The responsibility area of HHNK is fully covered by polders which contain urban as well as natural habitats.

4.3.2 The raw data

Table 5 and Table 6 list data attributes of pumping stations and polders respectively and their assigned categories. Of these raw data only information on identification is required for this use case. In total, data on 660 pumping stations and 309 polders was provided. Pumping stations on Texel were provided in an additional file which was not included in the analysis in order to reduce complexity.

Table 5: Attributes for Pumping Stations

Number	Category and attributes
1	Identification 1. KGM_ID: Represents the (internal) number of the record in the table. This number is not suitable (not unique) for use as a key code. 2. KGM_IDENT: is unique and used as a key (on a national scale).

Table 6: Attributes of Polders

Number	Category and attributes
1	Identification 1. ID 2. Name 3. Code

4.3.3 Data preparation

Based on the data requirements of the crisis management use case (see section 1.2) data were taken from the category identification. Attributes that were used in the raw data format included: KGM_IDENT, ID (polder) and Name (polder). X- and y-coordinates of pumping stations, their ground height as well as their matching polders had to be calculated.

4.3.4 Calculating x- and y-coordinates and ground height of pumping stations

This process is identical to calculating the GroundHeight for Electrical Assets, except that other data sets were used. Appendix D provides a tutorial for calculating the GroundHeight for Electrical Assets.

4.3.5 Matching pumping stations to the polders

Each pumping station is responsible for one polder. The following calculation was done under the assumption that the pumping station that lies within the polder polygon is responsible. This method is flawed since pumping stations that are located at the borders of polders are not included in this analysis.
This process is identical to calculating the Electrical supply Area of Electrical Assets, except that other data sets were used. Appendix C provides a tutorial for calculating the Electrical supply Area of Electrical Assets.

4.3.6 Final data set

Appendix B shows extracts of the final data sets for Pumping Stations and Polders.

4.4 AHN2 data

4.4.1 Raw data

AHN2 (Actueel Hoogtebestand Nederland versie 2) covers the topography of the Netherlands, i.e. provides the height above sealevel of the ground surface.
The AHN2 can be downloaded on the website of PDOK. The data set is split into tiles, so depending on the required area one must first select the appropriate tiles and then download them individually. For this use case point cloud data were used.

4.4.2 Data preparation

The AHN2 was used to calculate the GroundHeight of Electrical assets, Pumping Stations and Other Users, as well as x-and y-coordinates of pumping stations.

4.4.3 Calculate GroundHeight and x- and y-coordinates

Appendix D provides a tutorial for calculating GroundHeight and x-and y-coordinates of Electrical assets.

4.4.4 Final preparation

AHN2 data is included in the final data sets of Electrical Assets, Pumping Stations and Other Users (Appendix D, GroundHeight and x- and y-coordinate columns).

4.5 BAG data

4.5.1 Raw data

BAG (Basisregistraties adressen en gebouwen) includes data about all kinds of buildings (e.g. public and dwelling spaces, industrial, commercial). Attributes include status, address, surface area, geometry, x-and y-coordinates, year of construction and usage function. The BAG data set can be downloaded on the website of the Vrije Universiteit Amsterdam.

4.5.2 Data preparation

The BAG was used to identify Users within Electrical Supply Areas. Thus, only address and x-and y-coordinates where of interest for this use case. Due to privacy issues, Users within the same electrical supply area were grouped based on postcodes.

4.5.3 Match to User to Electrical supply area

This process is identical to calculating the Electrical supply Area of Electrical Assets, except that other data sets were used. See calculating electrical supply of electrical assets from step 3 onwards (Appendix C).

4.5.4 Final preparation

Appendix B shows extracts of the final data sets for Other Users.

4.6 Risicokaart

Rijkswaterstaat provides a web-service where data from the Risicokaart Nederland can be uploaded into qGIS. In this way data about water levels was supposed to be displayed in qGIS, transformed into a shapefile and finally combined with the BAG data in order to assign a location to each water level. Unfortunately, we did not succeed with the shapefile conversion so that water level data was not available for this use case.
Appendix E describes the process that we followed in detail.

4.7 Data set analysis

Problems were discovered with data about electrical supply area, polders, and water levels.
First, addresses were grouped based on the assumption that an electrical supply area comprises all addresses within a 100m radius of the electrical asset. However, an analysis in qGIS showed that this method leads to overlapping supply areas and addresses that lie outside of any supply area Figure 10. Thus, it is not possible to exactly assign an address to one responsible electrical asset. Further specifying the assumption about the electrical supply area would solve this issue.
Second, polders were matched to the pumping station in charged by visual means. If a pumping station lies within the boundaries of a polder the polder became the responsibility area of this pumping station. However, pumping stations that were located directly on the boundaries of a polder were left out and could not be assigned to any polder) Figure 11. HHNK should provide an explicit pumping station/polder match in their data set.
Third, water level data were not available since the conversion from wms format to shapefile did not work in qGIS. This was very unfortunate as the most basic question within the crisis scenario `Where is the flood` could therefore not be answered.

Figure 10: Overlapping electrical supply areas and address outside electrical supply area

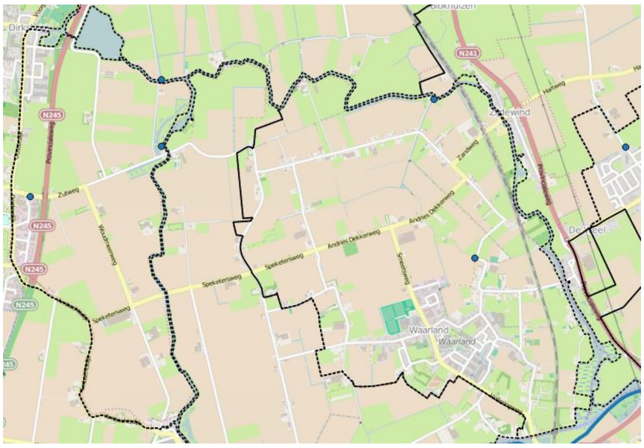


Figure 11: Pumping stations within polders and on polder boundaries

5 Analysis

This section analyzes whether the presented Linked Data-based solution improved data sharing within the crisis scenario. The analysis is split into three parts, assessing the suitability of Linked Data in general, the crisis management ontology and the data sets, respectively. Results are summarized in Table 7.

Table 7: Analysis of Linked-Data based solution

	Suitability for crisis management
Linked Data in general	• Highly suitable due to flexibility and knowledge generation • Benefits are only present if Linked Data becomes dominant technology
Crisis management ontology	• Concept of Electrical Supply Area does not work for old towns (heterogenous construction) (considering the restrictions of the dataset used) • Concept of Water Resistance Threshold is not tracked by stakeholders
Data sets	• 100m radius assumption is over-simplified • Pumping Station – Polder match is over-simplified • Water Level data are missing

5.1 Linked Data analysis

This section analyzes the level of interoperability of the current data sharing solution. Based on the scope of the data sharing scenario (see Section 1.2), measurable concepts were formulated which are summarized in

Table 8 (adapted from Folmer, 2012), this table also contains the advantages and disadvantages of Linked Data in comparison to current data exchange solutions from a generic perspective, and a specific perspective analyzing the developed crisis management ontology. We start with an overview of the seven major benefits of a Linked Data solution compared to a traditional solution, followed by a more detailed analysis in

Table 8.

This analysis revealed six major benefits of Linked Data compared to conventional technologies when used in the crisis scenario:

- generation of new knowledge,
- handling of dynamic content,
- flexibility,
- compliance to external standards,
- completeness and precision,
- data-centric approach instead of application-centric approach.

Table 8: Scope of data sharing scenario

Scope	Measurable concept	Generic	Specific
Functionality purpose	**Knowledge generation** Capability of the data sharing solution to provide novel insights.	Linked Data allows for automated reasoning, since data links are machine-readable and semantics are encoded in these links (Nederstigt et al., 2014). Thus, knowledge that is encoded implicitly within the data links can be extracted. This facilitates decision making since human information processing capabilities are limited and cannot foresee hidden regularities in the datalinks (Endres et al., 2015; Chan and Franklin, 2011).	This advantage is also present in the current crisis management ontology. Power assets that will be at risk for break down can be localized based on their resistance thresholds and predicted water levels and the necessary data is at any moment up-to-date as it is provided by the water board in real-time (instead of extracting the data, sending it to the parties that need the data and then put it in their own database).
	Dynamic content Capability of the data sharing solution to deal with often changing content.	Linked Data is well suitable for dynamic data due to the decentralized publication model (Tummarello et al., 2007). Data can be locally updated by the responsible stakeholder rather than within a centrally controlled database merge as it is often the case with relational models.	Thus, HHNK could make available their real-time data on water levels which are important to determine threatened assets.
	Completeness Extend to which the conceptual model is of sufficient breath, depth and scope for the task at hand.	Linked Data allows sharing parties to constantly adjust the definitions and interrelations in a decentralized manner (Tummarello et al., 2007). This allows them to easily improve the level to which all necessary information elements are covered. Too many or missing information elements will have a negative impact on interoperability (Folmer, 2012). Thus, redundancy in information exchange is reduced. Traditional standards foresee exchanging messages when the relevant event has occurred, e.g. after delivery, an invoice is sent. However, Linked Data rather references information by keeping it at its source and linking it when queries for information are received (Kalcheva, 2015). The covered functions of Linked Data can be extended by adjusting SPARQL queries. In principle, any question can be asked to the data. This acknowledges the unpredictability within a crisis as the user can simply alter his query depending on situational changes.	

	Precision Match between actual representation of functions and information vs. required ones in the conceptual model.	The more accurate data concepts are defined the more powerful Linked Data is since automated reasoning is not hampered by ambiguous concept names. For example, "WaterLevel" might either refer to the overground water height or to the groundwater level. This distinction is very important in the crisis management ontology because only the former concept is needed for the tasks at hand. Since this ontology was developed in cooperation with the stakeholders it was ensured that definitions were sufficiently precise for them. However, if more stakeholders adopt this data sharing solution the level of precision might not be sufficient any more. One of the main drawbacks in this respect is the limited expressiveness of Linked Data. Different concepts can be the same within a certain context but different within another (Kalcheva, 2015). For example, in the staffing industry workers are seen as resources in the same way as power assets are considered to be resources in the utility domain. Therefore the triples "Workers" "sameAs" "Resources" and "PowerAssets" "sameAs" "Resources" are semantically true in the context of the staffing industry and the utility domain respectively. However the concepts "Workers" and "PowerAssets" certainly do not refer to the same concept outside this contexts. This requires an additional specification of the concepts and a precise definition of the contexts within which different concepts are the same (Hammer & McLeod, 1981). For example, the problem above could be fixed by also determining that "Workers" "areATypeOf" "HumanBeings" and "PowerAssets "areATypeOf" "Appliances". By adding these details people and computers can find out that the subjects are both resources but are not identical. This specification might be problematic for data sharing within crisis management as crises situations often change unpredictably and therefore give rise to a vast number of context constellations.	

Durability purpose	**Flexibility** Possibility to extend the capabilities of the datasharing solution for other usage environments.	The triple structure that underlies Linked Data is very flexible. Anyone can always add new facts, i.e., properties to resources (Eckartz & Folmer, 2014). Ontologies can also be easily extend by either specifying new concepts and relationships or by mapping two different ontologies onto each other. This in contrast to data model extension of conventional central solutions which are always complex and time consuming due to the fact that agreement must be reached with many parties and are therefore expensive.	The crisis management ontology can also be easily adapted to new tasks. For example, other grid managers and operators, such as TenneT, who deals with high voltage components, can share data by adding their domain specific concepts to the ontology. This reduces re-use barriers because the use of information from additional data sources does not necessitate changes in the application code (Tummarello et al., 2007).
	External compliance Capability of the data sharing solution to adhere to other data sharing conventions (e.g. ease of mapping onto other ontologies).	Conceptual models that underlie data sharing within Linked Data can easily be mapped onto each other, even if they derive from different domains (Eckartz & Folmer, 2014). Due to the coherent structure that Linked Data provides for data, all data sharing solutions that support this technology comply well with each other. However, non-Linked Data sharing solutions that do not have a conceptual model cannot be related to the Linked Data solution because no inter-model mappings can be realized (Tummarello et al., 2007). In these cases external compliance can only be assured by building a conceptual model from scratch for the non-Linked Data solution and map it onto the Linked Data ontology, i.e. transforming data into Linked Data. A very tedious process.	The crisis management ontology complies very well to external ontologies. For example, it uses similar concepts as the IEC CIM ontology (Common Information Model), a widely accepted standard in the utility domain (Simmins, 2011; IEC 61968, 61970). Due to this overlap in concepts mapping these two ontologies onto each other becomes easier.
	Maturity of technology Amount of expertise in the Linked Data domain and of existing conventions of usage.	Many mature tools, programs, and libraries are available to interact and operate with Linked Data (Alhajj & Rokne, 2014). There are various established best practices for publishing and providing Linked Data on the web, mostly published by the W3C (W3C, 2015). However, it is easy to become overwhelmed by the vast amount of options and novices easily lose the overview (Eckartz & Folmer, 2014). The same accounts for the great amount of available ontologies. There is neither an official register nor recommendations when to use which one. Moreover, the amount of expertise regarding Linked Data, although quickly growing, is still relatively small, since it is a	With respect to this study a new crisis management ontology would not have been created if it had been known that such an ontology already exists.

		relatively young technology. As we consider the availability of expertise part of the maturity of the technology, we can conclude that the Linked Data technology is not fully mature. Thus, the creation and use of Linked Data are currently cumbersome and time consuming.	
Small user community	**Installed based** Connection of the data sharing to the current ICT landscape and business processes of the data sharing parties.	The majority of today's data sharing systems are based on relational data models and do not support Linked Data (Tummarello et al., 2007). They have to be replaced and/or adjusted which is resource consuming since users also have to adapt to this new technology. So it is still doubtful whether there is enough incentive for this change on a national scale in the Netherlands. Nevertheless, Linked Data is only beneficial if it is the dominant solution in a certain context or domain (Tummarello et al., 2007)	The present crisis management ontology is well suited for data exchange between Alliander and HHNK. However, if no more stakeholders join in this solution it will be quickly dropped if it is only compatible with a minority of present data management systems. The potential installed base for the crisis management ontology is: 8 network operators, 22 water boards and 25 safety regions which would provide a steady base for the continuity and availability of the solution.
Closed application environment	**Privacy** Extent to which data are protected from open access.	Linked Data requires data to be published on the web (Berners-Lee, Hendler & Lassila, 2001). Since data is no longer exchanged but interlinked open SPARQL[2] endpoints allow anyone to access the data (Kalcheva, 2015). Currently, these endpoints do not have any security or access policies in place. A web-service built on top of the data, working like an API as defined in the 6-star Linked Data model (Folmer & Verdonk, 2014), might solve this problem. However, its set-up and maintenance can be costly and time consuming.	The present crisis management ontology contains sensitive data about electricity consumption of private households, location of cables and transformers. It has to be protected to ensure privacy, prevent copper theft and hamper tapping substations for private use (Book, Bastiaans & Bruinenberg, 2014). To protect can be done by providing an API instead of an open SPARQL endpoint which can also solve privacy and security issues.
Economical purpose	**Economic benefit** Extend to which data sharing solution provides financial and economic advantages.	As described in "Installed Base" adapting a company's data architecture to be Linked Data-compatible can be very costly. However, in the long run as benefits of flexibility, automated reasoning and dynamic content strive, significant operating cost savings will become prevalent (Steinfield et al., 2011).	

[2] Query language for Linked Data.

6 Conclusion, discussion and future work

6.1 Conclusion

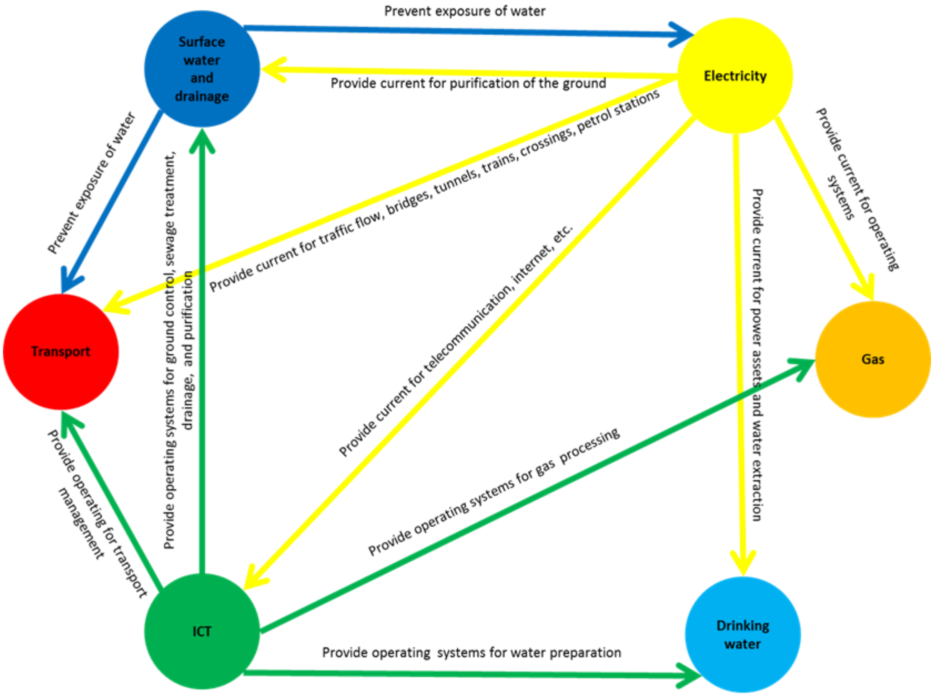
The aim of this study was to create a data sharing solution for crisis management in order to examine whether Linked Data improves interoperability between the data sharing parties. Results revealed that Linked Data in general and the crisis management ontology in particular provide a very good method to improve interoperability providing that Linked Data will be used more widely by network operators and water boards. This report provided evidence for improved interoperability if information is preparing and shared within a common Linked Data-based conceptual framework. Crisis managers will be able to accurately assess the situation and make fast and adequate decisions in order to optimize crisis management. Follow-up project should adopt Linked Data to explore other aspects of crisis management in order to improve ecological validity of research in this domain and the usage of Linked Data to exchange information should be encouraged.
An interesting next step would be to make the proposed HHNK – Liander solution operational to try out the solution in practice. Whilst making the proposed solution operational lessons learned should be taken into account and the 6-star model for Linked Data (Folmer & Verdonk, 2014) should be followed to arrange the availability of the data through an API.

This study revealed the following six major benefits of Linked Data compared to conventional technologies when used in the crisis scenario:

- generation of new knowledge,
- handling of dynamic content,
- flexibility,
- compliance to external standards,
- completeness and precision,
- data-centric approach instead of application-centric approach.

6.2 Discussion and future work

In order to further test the suitability of this Linked Data-based information exchange, future work should focus on the development of an implementation in order to test data sharing in practice. For example, a graphical, map-based web application could show the real-time development of water levels in combination with the location of power- and water assets. Based on the underlying crisis management ontology the application could visualize those assets that are at risk, damaged or have broken down. By showing their respective supply area affected consumers can be localized. Whilst implementing the data sets should be updated and revised and care should be taken to make sure that the data stays up-to-date (by providing a way to update the data).
Moreover, further insights might be gained by extending this data sharing solution to other aspects of crisis management, such as domino effects (van Dongen et al., 2013), grid-reconfiguration (Aarsen et al., 2015), and emergency response processes (Fan & Zlatanova, 2010). By extending the Linked Data data sharing solution the available data becomes more and more rich, whilst preventing more complex solutions based on the conventional ways of exchanging data.
First, in the Western world vital processes of societal life are highly interdependent. For instance, industry depends on electricity for the production of goods and on a functioning transport system (Figure 12)



- failure of internet, mobile and fixed phone traffic, money transactions
- inaccessibility of staff, coordination centres, especially helping centres
- unnavigability of roads, tunnels, bridges
- inaccessibility of areas, people, objects, and infrastructure
- failure of evacuation, aid, supply of vital products (food, fuel, medicines)
- failure of stream and river drainage
- wastewater collection and treatment failure
- environmental and healthcare damages
- failure of drinking water supply in highrise buildings
- failure of lights, heating, fridges and freezers, elevators, radio and internet, petrol pumps, traffic regulation, digital money transfer, milking equipment, ventilation stables, processing of chemical products, medical equipment, etc.
- failure of heating of buildings and of water undercooling
- failure of production processes
- lack of clean drinking water
- lack of water for sanitation, for production processes and for fire fights
- damage for health and production

Figure 12: Key dependencies between sectors (adapted from van Dongen et al., 2013)

How is this complex system affected by a crisis? Which vital sectors are threatened at what point in time and which consequences does this have for humans? Since the electricity and the water management sectors are important factors in society, the current crisis management scenario is well suitable to explore domino-effects during a flood-caused power cut. Linked Data could help to interrelate consequences within and between specific sectors in order to predict the course of the domino effect. By mapping these data onto a time ontology, accuracy of this prediction and ecological validity of the model might improve (Figure 13).

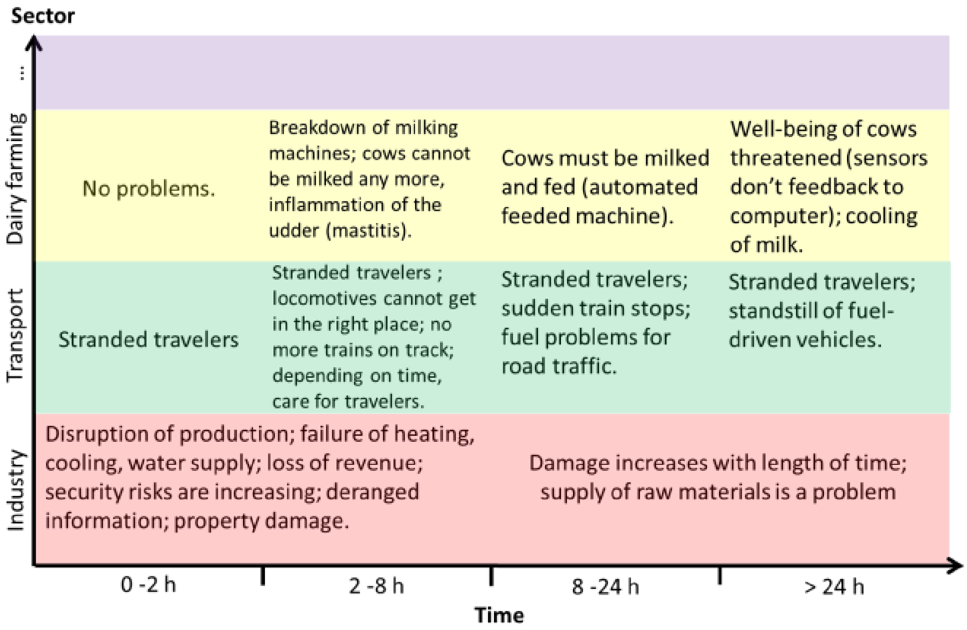


Figure 13: Temporal effect-chain of social disruption after a power cut (adapted from Book, Bastiaans & Bruinenberg, 2014)

Second, an important process in the current crisis management scenario is the reconfiguration of the electrical grid after the breakdown of a power asset. In order to re-ensure electricity supply in the affected area Alliander has to check which additional power generating sources are eligible for this purpose. Solar panels on private roofs represent such a source (Aarsen et al., 2015). They are highly distributed across the country, meaning that it is quite likely that some panels are always in the vicinity of an affected supply area. Moreover, their network represents a very robust source of power generation: if one panel fails others can make up for the individual breakdown. Gathering data on location and production rate of solar panels and linking them to the data sets described in this use case could provide a rich source for improving grid reconfiguration during a crisis.

Although no fixed infrastructure is in place yet to connect solar panels to affected consumers, a mobile system could be flexibly set up. Emergency aggregates could be transported to nearby solar panel sites to be charged there rather than taking them back to the distant headquarter of the grid operator. In the same way electrically powered emergency vehicles such as fire brigade and police cars could be re-charged in this way.

Third, suggestions so far have mainly focused on increasing the amount of interlinked domain-specific data, i.e. information about domains that are affected by a crisis. The aim clearly was to retrieve more information from these interdependencies, leaving out questions such as when and how this specific information will be of use in the whole emergency response process.

By modeling emergency response processes such as fire-fighting, tasks for all actors are well specified (Figure 14, Figure 15). In a follow-up project these process models could be linked to the crisis management ontology. In this way actors know exactly when and where to request certain information from.

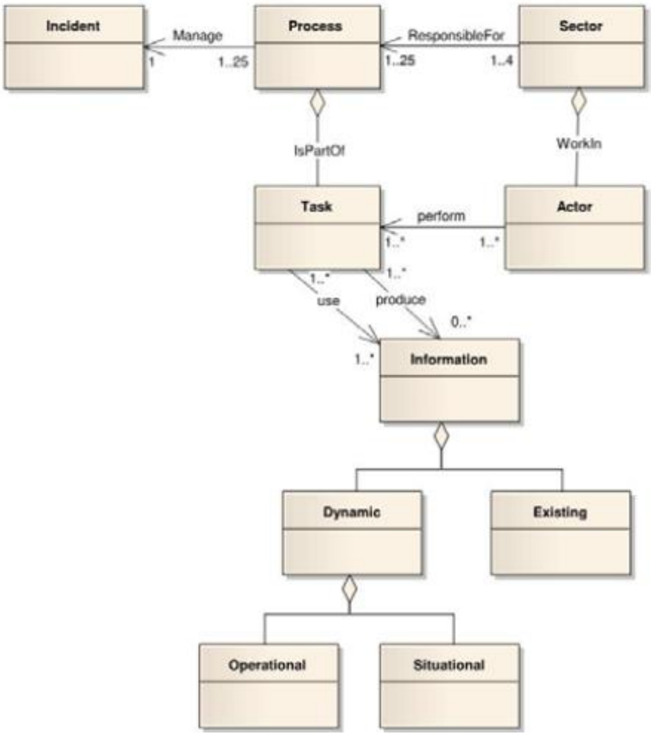


Figure 14: Conceptual model of general emergency response adapted from Zlatanova, 2010)

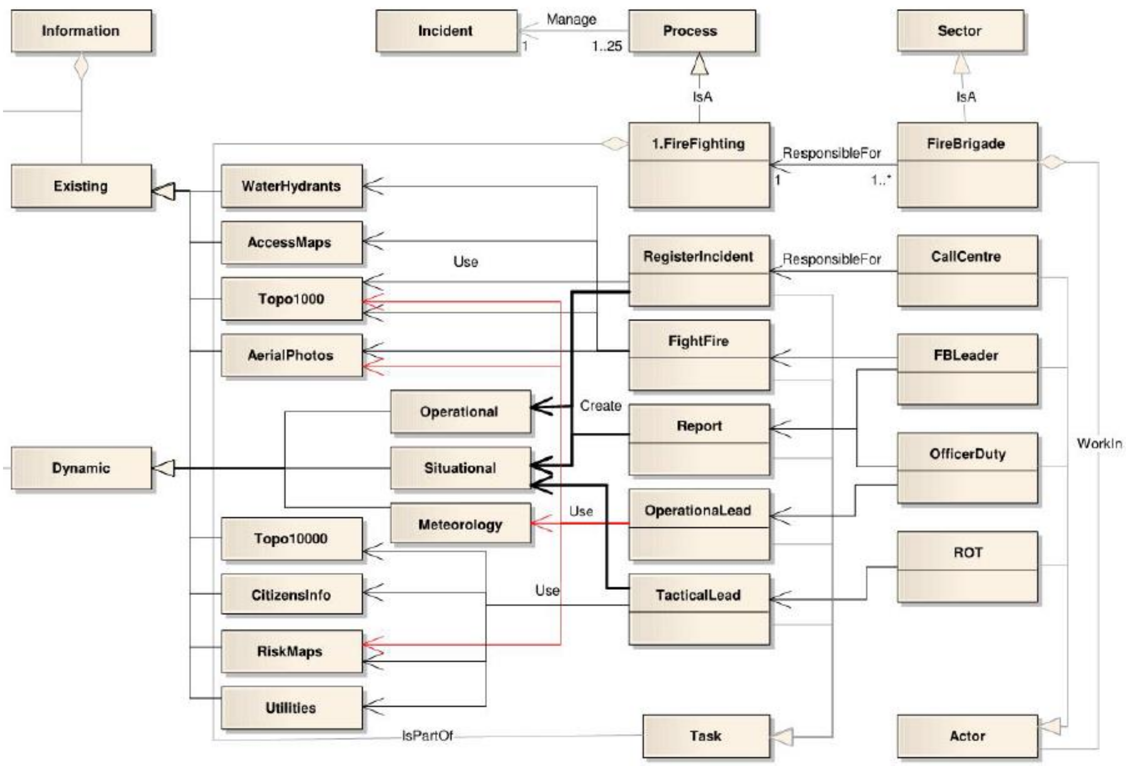


Figure 15: Modelling of a specific emergency response tasks: fire fighting (taken from Zlatanova, 2010).

References

Aarsen, R., Ramkisoen, M., Janssen, M., Quak, W. & Verbree, E., 2015. Installed Base Registration of Decentralized Solar Panels. Submitted to Gi4DM workshop at ISPRS GEOSPATIAL WEEK 2015.

Aartsen, R. Bruinenberg J. & van Gelder, J., 2015. Electrical grid discussion. Private communication.

ACIR, 2005. De vrijblijvendheid voorbij. Retrieved from: http://www.infopuntveiligheid.nl/Publicatie/DossierItem/68/778/acir-rapport-commissie-hermans-de-vrijblijvendheid-voorbij.html on 28th April 2015.

Alhajj, R. & Rokne, J., 2014. *Encyclopedia of Social Network Analysis and Mining,* Springer Science Business Media, New York, pp. 811-813.

Berners-Lee, T., Hendler, J. & Lassila, O., 2001. The semantic web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American 284, pp. 34-43.

Book, M. Bastiaans, S. & Bruinenberg, J., 2014. Crisis management discussion. Private communication.

CERISE, 2015. CERISE-SG: Alignment of the information infrastructure architecture of utility companies with related architectures: enlarging circles (for smart grids). Retrieved from http://www.cerise-project.nl/index.php?lang=en&Itemid=103 on 23rd April.

Chan, S.W.K. & Franklin, J., 2011. A text-based decision support system for financial sequence prediction. *Decision Support Systems,* 52, pp. 189-198.

de Landmeter, R. & van Giessel, K., 2015. Crisis management discussion. Private communication.

Eckartz S. & Folmer, E., 2014. Stappenplan: Best practice voor het publiceren van Linked Data. Retrieved from http://www.pilod.nl/wiki/BoekTNO/stappenplan om 28th April 2015.

Endres, M.J., Houpt, J.W., Donkin C. & Finn, P., 2015. Working Memory Capacity and Redundant Information Processing Efficiency*. Frontiers in Psychology: Quantitative Psychology and Measurement,* 6, pp. 1-18.

Esri, 2011. Uitleg over het nieuwe Landelijk Crisismanagement Systeem. Retrieved from http://www.esri.nl/nieuws-media/nieuws-gis/uitleg-over-het-nieuwe-landelijk-crisismanagement-systeem on 23rd April 2015.

Fan, Z. & Zlatanova, S., 2010. Exploring ontology potential in emergency management. Presented at the conference on geomatics for crisis management. Torino, Italy.

Firebrary, 2015. Firebrary Home. Retrieved from http://www.firebrary.com/nl/ on 23rd April, 2015.

Folmer, E. & Verdonk, Y., 2014. Linked Data in beeld. Retrieved from http://www.geonovum.nl/sites/default/files/Linked%20Data%20in%20beeld-2014-Locked.pdf

Hammer, M. & McLeod, D., 1981. Database description with SDM: A semantic database model. *ACM Transactions on Database Systems,* 6(3), pp. 351-386.

Kalcheva, K., 2015. A model to showcase the potential for LD application within (financial) business processes. In press.

Lehman, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P.N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S. & Bizer, C., 2015. Dbpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal,* 6(2), pp. 167-195.

Nederstigt, L.J., Aanen, S.S., Vandic, D., & Frasincar, F., 2014. FLOPPIES: A Framework for Large-Scale Ontology Population of Product Information from Tabular Data in E-commerce Stores. *Decision Support Systems.* 59, pp. 296-311.

Simmins, J.J., 2011. The impact of PAP 8 on the Common Information Model (CIM). Power Systems Conference and Exposition (PSCE), IEEE/PES, Phoenix, AZ.

Steinfield, C.W., Markus, M.L., & Wigand, R.T., 2011. Cooperative Advantage and Vertical Information System Standards: An Automotive Supply Chain Case Study. (IEEE), pp. 1-10.

Tummarello, G., Delbru, R., & Oren, E., 2007. Sindice.com: Weaving the Open Linked Data. ISWC07ASWC07 Proc. 6th Int. Semantic Web 2nd Asian Conf. Asian Semantic Web Conf. pp. 552-565.

van Dongen, K., Stolk, D., Weber, L. & de Lange, M., 2013. Informatiepreparatie overstromingsrisico's en domino-effecten: Een verkenning. Retrieved from http://www.infopuntveiligheid.nl/Publicatie/DossierItem/10/5488/informatiepreparatie-overstromingsrisicos-en-domino-effecten-een-verkenning.html on 28th April 2015.

van Leeuwen, B., 2015. Realtime Linked Open Data - Brandweer Amsterdam-Amstelland, PiLOD Conference, Utrecht, Netherlands. Retrieved from http://www.pilod.nl/w/images/2/29/20150402_Waternet_Bart_van_Leeuwen.pdf on 28th April 2015.

Waterschappen, 2015. Ontdek ons. Retrieved from http://www.waterschappen.nl/ontdek-ons/ on 28th April.

World Wide Web Consortium (W3C), 2015. Linked Data. Retrieved from http://www.w3.org/standards/semanticweb/data on 24th April 2015.

Zlatanova, S. 2010. Formal modelling of processes and tasks to support use and search of geo-information in emergency response. *13th Annual International Conference and Exhibition on geospatial Information Technology and Applications,* Gurgaon, India, pp. 1-10.

Zlatanova, S. Ghawana, T., Kaur, A. & Neuvel, J., 2014. Integrated Flood Disaster Management and Spatial Information: Case Studies of Netherlands and India. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XL-8, pp. 147-154.

A Hands-on session: Applying Linked Data to crisis management data

A.1 Populating the ontology with data

This section explains how to link actual data sets to the crisis management ontology. By mapping the data onto the ontology, the latter is filled with instances (TopBraid Composer term). This attaches a semantic value to the data which can then be queried using SPARQL. The raw data of the current use case was in shape file format initially. After having prepared the data in the geo-program qGIS (see Appendix C), the shape files were exported into csv format and finally converted into excel. The following step-by-step-guide describes the populating process with this file format. Two semantic web-tools were used for this process: Web-Karma and TopBraid Composer. Tutorials can be found in Appendix F and G respectively. Table 9 compares the two tools.

Table 9: Comparison of semantic web-tools

	Web-Karma	**TopBraid Comoser**
Benefits	Free	• Data is not published • One tool option
Drawbacks	• Data is published • Video tutorials (“Karma has learned this from previous sessions and suggests the right option.”)	• Standard or Maestro edition needed • Misleading feedback (“Instances have been uploaded to the ontology.”)
Pitfalls	• .owl format required (Protégé)	• Specific format of data set (1st column label has to be equivalent with class name) • Get the prefix right!

A.2 SPARQL queries for crisis management ontology

In order to answer the relevant questions during the crisis (Section 2.2), SPARQL queries had to be formulated (see Table 10). Since data on water levels was missing these queries were never validated. They show how we would have proceeded if data had been available.

Table 10: SPARQL queries

Question during the crisis	**Reasoning**	**SPARQL query**

Where is the flood?	A point is flooded if the real-time water level is greater than the water level of the normal state at that location	**SELECT** ?cm:RealTimeWaterLevel, ?q, ?r
		WHERE {
		?cm:RealTimeWaterLevelrdfs:hasHeight ?o .
		?cm:NormalWaterLevel rdfs:hasHeight ?p.
		?cm: RealTimeWaterLevel rdfs: hasXLocation ?q.
		?cm:RealTimewaterLevel rdfs:hasYLocation ?r.
		}
		FILTER (?o > ?p)
		}

Which pumping station is in the vicinity to drain the flooded area?	Calculate distances between flood location and location of pumping station.
	Rank the distances.
		SELECT ?u
		WHERE {
		?cm:RealTimeWaterLevel rdfs:hasHeight ?o .
		?cm:NormalWaterLevel rdfs:hasHeight ?p.
		?cm: RealTimeWaterLevel rdfs: hasXLocation ?q.
		?cm:RealTimewaterLevel rdfs:hasYLocation ?r.
		
		?cm:PumpingStation rdfs:hasXLocation ?s.
		?cm:PumpingStation rdfs:hasYLocation ?t
		}
		FILTER (?o > ?p)
		(SQURT((?q - ?s)^2 + (?r - ?t)^2) = ?u)
		}

Which electrical assets are broken down due to the flood?	If threshold + ground height < water level then the electrical asset will break down.	**SELECT** ?ElectricalAsset
		WHERE {
		?cm:ElectricalAsset rdfs:hasGroundHeight ?k .
		?cm:ElectricalAsset
		rdfs:hasWaterResistanceThreshold ?l.
		?cm: RealTimeWaterLevel rdfs: hasHeight ?m.
		}
		FILTER (?l + ?k < ?m)
		}

Which users are affected by the power cut?	If location of electrical supply area = location of user this user will be without power	**SELECT** ?n, ?i
		WHERE {
		cm:ElectricalAsset rdfs:hasGroundHeight ?k .
		cm:ElectricalAsset rdfs:hasWaterResistanceThreshold ?l.
		cm:RealTimeWaterLevel rdfs: hasHeight ?m.
		
		cm:ElectricalSupplyArea rdfs:hasName ?n.
		cm:ElectricalAsset rdfs:hasName ?h.
		cm:ElectricalSupplyArea rdfs:hasPostcodeRangeElectricity
		?i
		}
		FILTER (?l + ?k < ?m)
		}

Which power station is in the vicinity to replace the broken one?	Calculate distances between broken electrical asset and the functioning ones.
	Rank the distances
		SELECT ?a
		WHERE {
		cm:ElectricalAsset rdfs:hasGroundHeight ?k .
		cm:ElectricalAsset rdfs:hasWaterResistanceThreshold ?l.
		cm:RealTimeWaterLevel rdfs: hasHeight ?m.
		
		cm:ElectricalAsset rdfs:hasGroundHeight ?o .
		cm:ElectricalAsset rdfs:hasWaterResistanceThreshold ?p.
		cm:RealTimeWaterLevel rdfs: hasHeight ?q.
		
		?b rdfs:hasXLocation ?d.
		?b rdfs:hasYLocation ?e.
		?c rdfs:hasXLocation ?f.
		?c rdfs:hasYLocation ?g.
		}
		FILTER (?l + ?k < ?m) #store result in variable ?b
		FILTER (?o + ?p > ?q) #store result in variable ?c
		SQURT((?d - ?f)^2 + (?e - ?g)^2) = ?a

A.3 Ontology-ontology mappings

Mapping the crisis management ontology to other existing ontologies means that relationships between concepts of both ontologies are specified. For example, the concept “electrical asset” of the crisis management ontology can be mapped via the predicate “sameAs” to the concept “transformer station”, which might be part of another ontology.

There are two main benefits of inter-onbtology mappings. First, mappings facilitate understanding since unknown terms in a new ontology can be “translated” to familiar terms of known ontologies. For example, within Alliander and the utility domain in general the “Common Information Model” (CIM) is a widely accepted standard for defining terms and inter-relations within the utility domain. Currently, CIM is published in UML format. However, with the CIM-tool it can be converted into an ontology in ttl-format. This ontology could then be linked to the crisis management ontology. Deliverable 4.1 of WP40 describes this process in detail and present examplatory mappings between CIM and the ebif ontology, another CERISE-ontology. The same process could also be applied to the crisis management ontology.

Second, mappings ensure a wider use of the ontology since it gets accessible via multiple sources. In the same way, highly interlinked web-pages are more likely to have higher number of visitors than les sinter-linked sites.

B Final Data sets

The following tables show extracts of the final data sets used to populate the crisis management ontology.

Table 11: Final data set for Electrical Assets. WRT = Water Resistance Threshold

Name	GroundHeight	Xlocation	Ylocation	StreetName	HouseNumber	Postcode	TownName	RouteOfElectricalAsset	SupplyingHighVoltageAssetOfElectricalAsset	AmountOfIncomingOutgoingVoltageOfElectricalAsset	WRT
GV ZIRKOONSTR SORTIVA VD6	-0.26700	112524,307	513189,579	BOEKELERDIJK	13	1812LV	ALKMAAR	01 GVB 04	GV ZIRKOONSTRAAT SORTIVA	10kV/0.4kV	0,3
GV ZIRKOONSTR SORTIVA VD1	-0.39600	112446,913	513302,806	BOEKELERDIJK		1812RN	ALKMAAR	01 GVB 04	GV ZIRKOONSTRAAT SORTIVA	10kV/0.4kV	0,3

Table 12: Final data set for Electrical Supply Area

Name	PostcodeElectricity1	PostcodeElectricity2	PostcodeElectricity3	PostcodeElectricity4	PostcodeElectricity5	PostcodeElectricity6
AS ALTONSTRAAT 15 NST	1704CC					
ENTIUS 1	1704CP	1704BG				

Table 13: Final data set for Pumping Stations

ID	TypeOfPumpingStation	DriveMechanismOfPum	GroundHeight	Xlocation	YLocation	ResponsibleForPolder
KGM-A-1100		electrisch	-399.100	124914	503430	Beemster
KGM-A-1903		electrisch	-446.400	129425	500327	Purmer landelijk no

Table 14: Final data set for Polders

IDPolder	NamePolder	SurfaceAreaOfPolder
GGA_GAF_01	Berkmeer	71103916
GGA_GAF_02	Veenhuizen	11120257

Table 15: Final data set for Other user

FunctionalityOfOtherUserBuilding	Street	Postcode	TownOfOtherUser	SurfaceAreaOfOtherUs	EnergyLabelOfOtherUs	GroundHeightOfOtherU	XLocationOfOtherUser	YLocationOfOtherUser
woonfunctie	Groote Geldebosch 34	1704CP	Heerhugowaard	129 m2	geen label	-153.900	118748	524812
woonfunctie	Groote Geldebosch 32	1704CP	Heerhugowaard	129 m2	geen label	-153.300	118746	524804

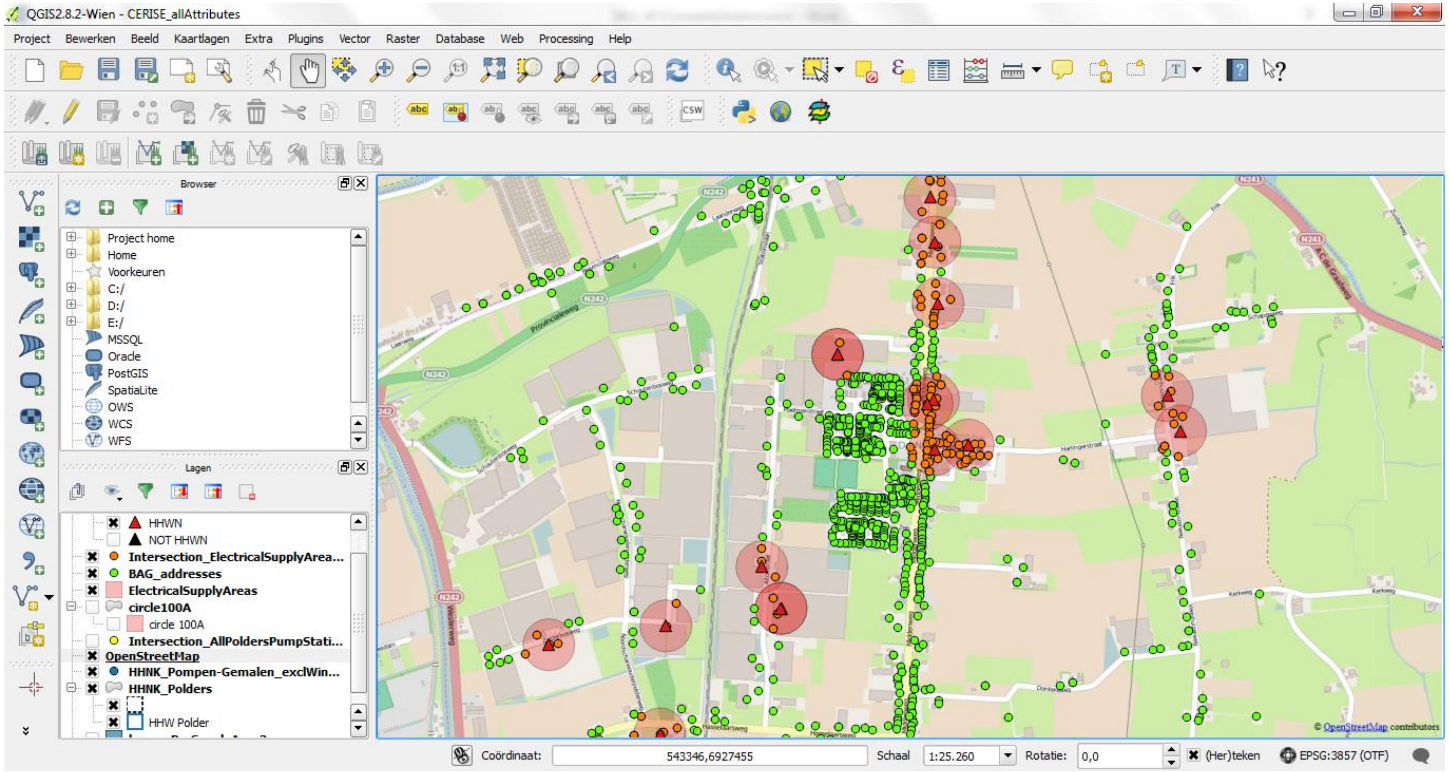


Figure 16: Final data sets mapped onto a street map

C Tutorial to calculate Electrical Supply Area in qGIS

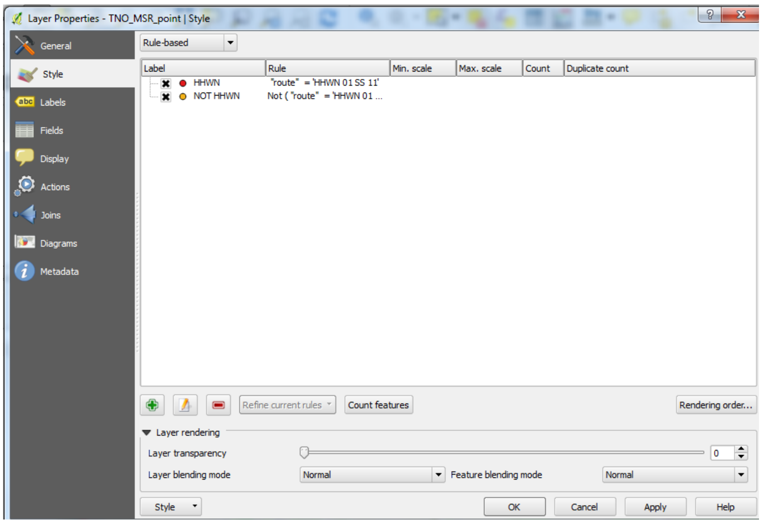
Step-by-step guide to calculating electrical supply area. This tutorial consists of 7 main steps, each containing one or more sub steps.

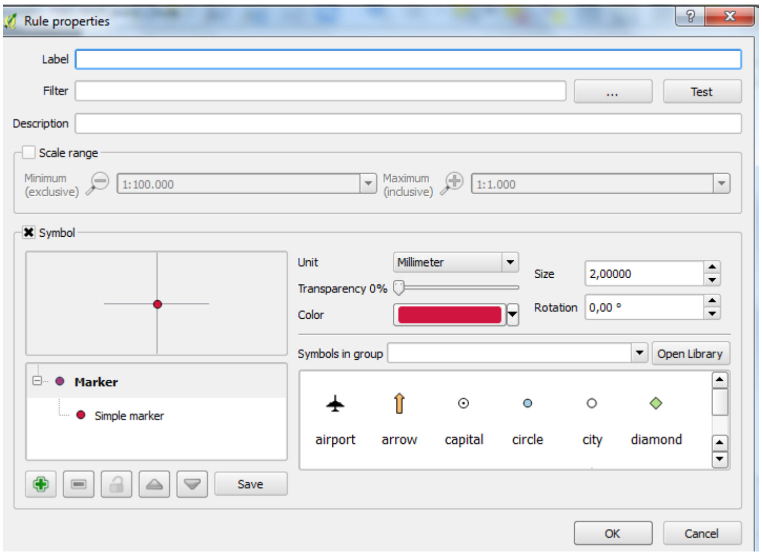
Step 1: Preparation

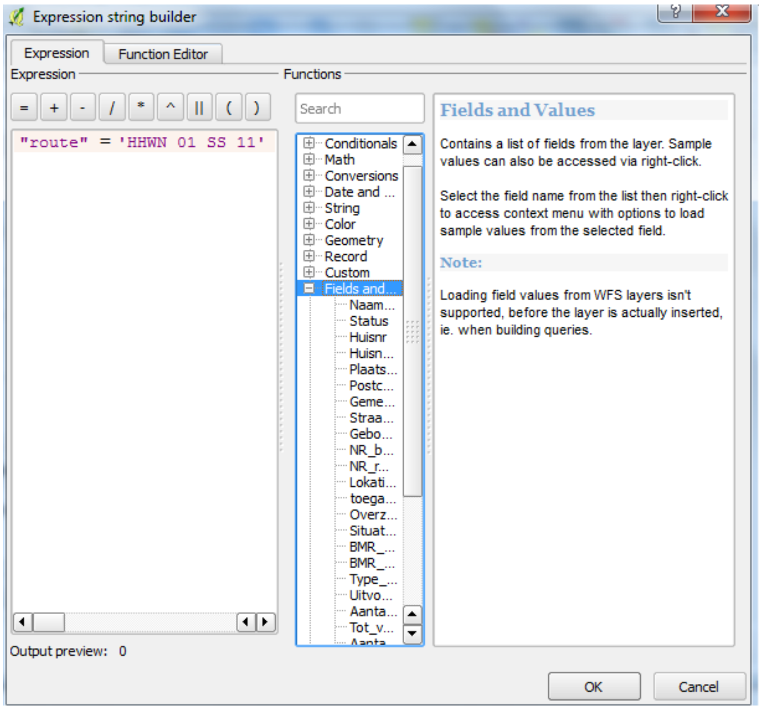
1. Upload Alliander shapefile, a map (e.g. OpenStreet map, plugin needed), and BAG shapefile as three separate layers into qGIS.

Step 2: Select stations of route HHWN 01 SS 11.

1. Open style tab of Alliander file.
2. Select “rule-based” → “+” → label rule (e.g. HHWN) → describe rule →“…”.
3. Select “routes” from “fields and values menu”.
4. Type in rule: “route”= ‘HHWN 01 SS 11’ or NOT (“route”= ‘HHWN 01 SS 11’) (select all rows where the route value is HHWN 01 SS 11 or not HHWN 01 SS 11).
5. RESULT: HHWN 01 SS 11 are now a subcategory of the Alliander layer and displayed with a different symbol and colour on the canvas.

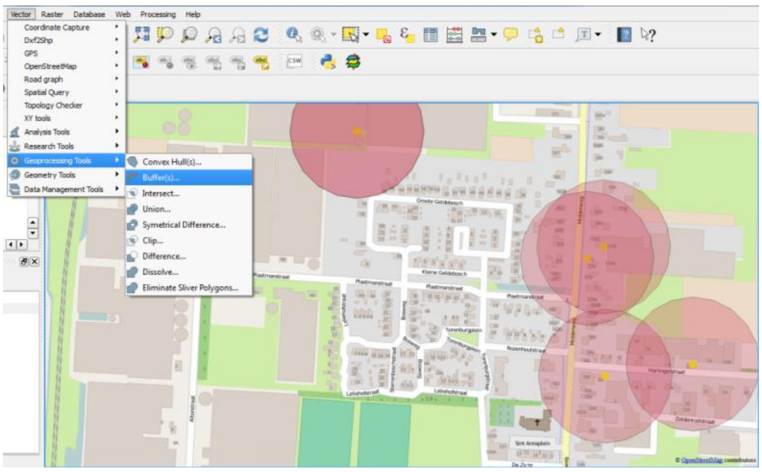


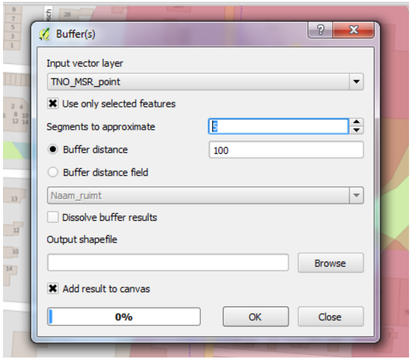




Step 3: Draw circles around stations to visualize electrical supply areas.

1. Highlight filtered HHWN 01 SS 11 stations.
2. Select Vector → geoprocessing tools → buffer.
3. Specify buffer distance (100m).
4. Save new circle layer as shape file.
5. RESULT: New shape file is created in a new qGIS layer.
6. Buffer tool always use the input layer's Coordinate Reference System (CRS) units. Make sure that new shape file and input file (Alliander data) use the same CRS.





Step 4: Graphically select addresses within electrical supply areas.

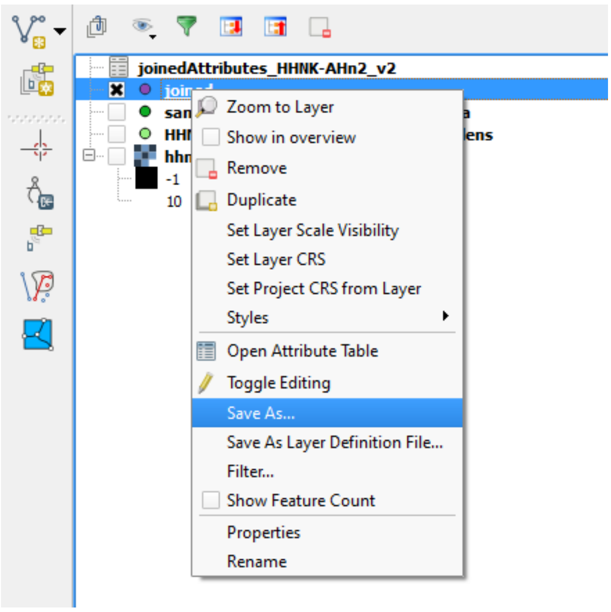
1. Select Vector → research Tools → select by location.
2. Select features in: BAG (address data set).
3. that intersect features in: Electrical supply areas (circle polygons around power stations).
4. tick appropriate boxes (intersection, overlap, touch, etc.).
5. Choose "creating new selection".
6. RESULT: points within circle polygons will appear in a different colour, corresponding rows in the BAG attributes table will be highlighted.
7. http://gis.stackexchange.com/questions/61753/how-to-select-features-within-a-polygon-from-another-layer

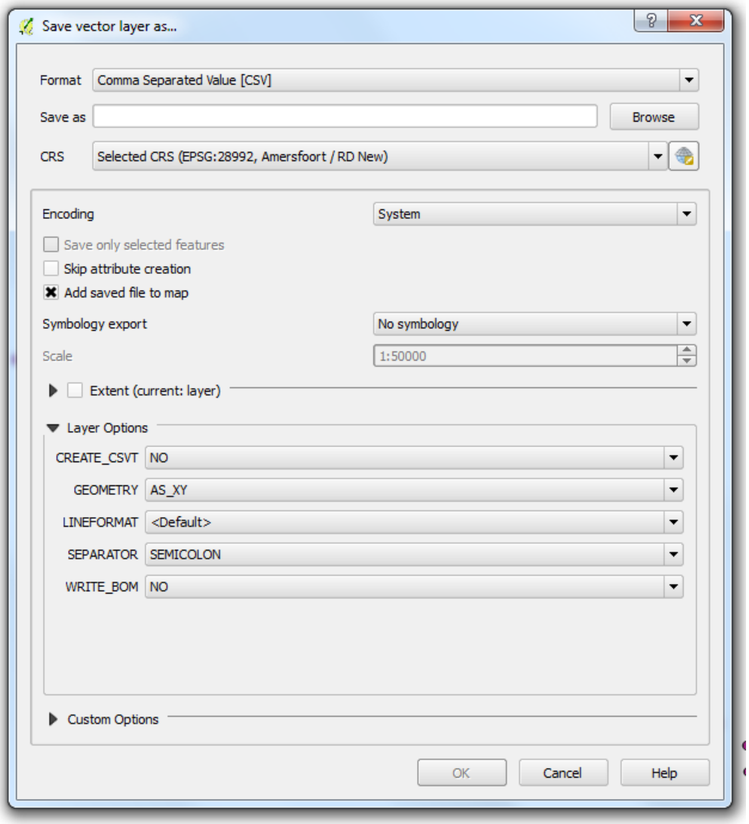
Step 5: Extract intersections between addresses and electrical supply areas as shapefile.

1. Make sure that data from step 3 is still selected in BAG attributes table.
2. Select Vector → Geoprocessing tools → intersect.
3. Input Vector layer: BAG (tick box “only selected features”).
4. Intersect layer: ElectricalSupplyAreas.
5. Do not swap input and intersect layer fields (otherwise intersection file doesn't contain any data).
6. RESULT: New shapefile is created as an additional layer. On canvas points within circle polygons are marked in a different colour.
7. http://gis.stackexchange.com/questions/61753/how-to-select-features-within-a-polygon-from-another-layer

Step 6: Exporting shape files into csv.

1. Right-click on shape file layer, select Save as → csv
2. Choose semicolon as delimiter NOT comma





Step 7: Clustering addresses per electrical supply area

The newly created data set assigns each address to the electrical supply area of one power stations. Since this violates privacy and security issues, addresses have to be clustered into postcode ranges. This can be done with the sort function in excel.

D Tutorial for calculating the GroundHeight

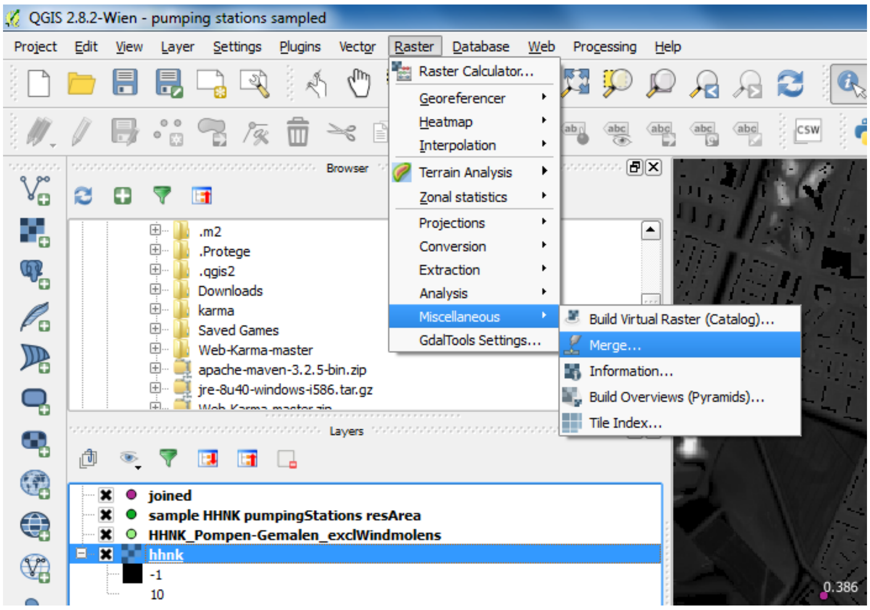
Step-by-step guide to calculating GroundHeight of Electrical Assets, Pumping Stations and Other Users. This tutorial consists of 5 main steps, each containing one or more sub steps.

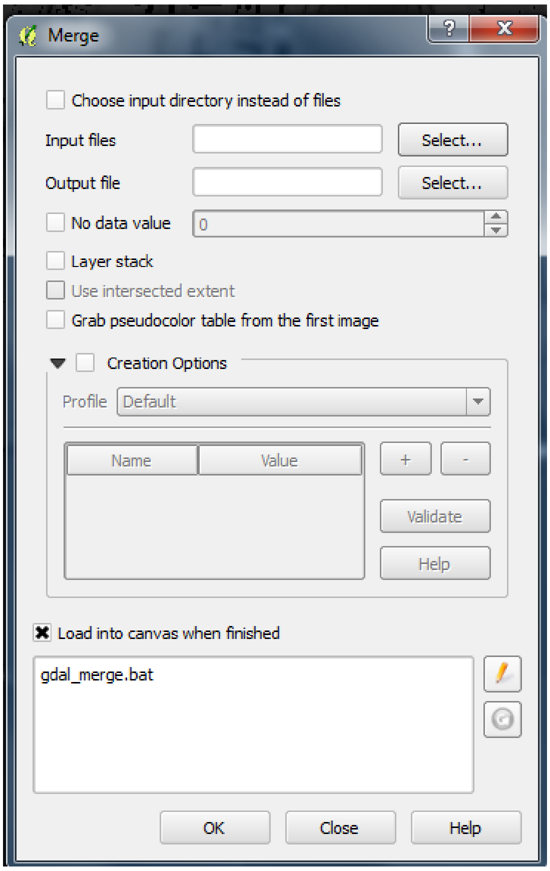
Step 1: Preparation

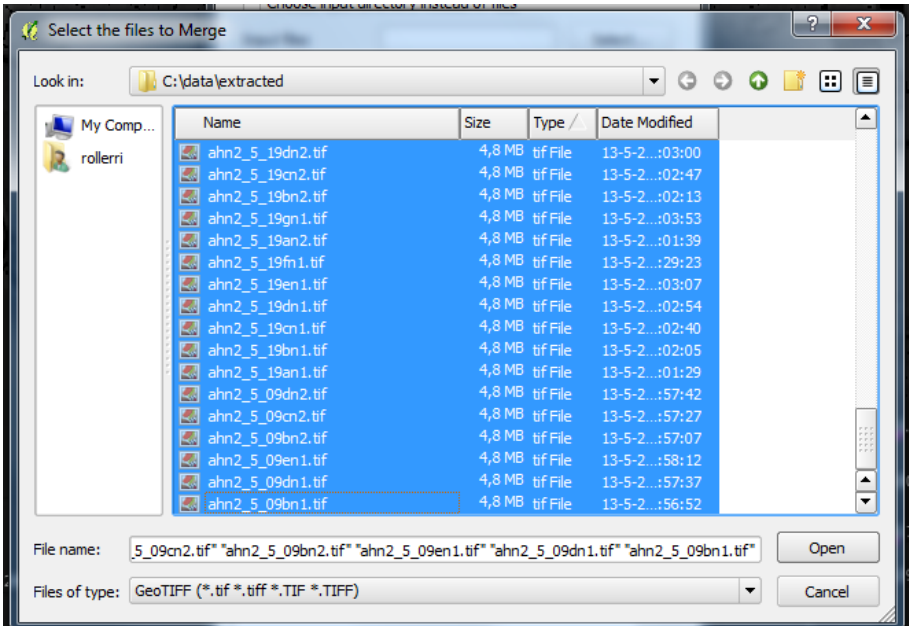
1. Upload Alliander shapefile as a new layers into qGIS.
2. Save all AHN2 TIF files in the same directory with a short path (e.g. C:\data_AHN2).
3. Install point sampling plugin.

Step 2: Merge separate AHN2 TIF files

1. Select Raster → Miscellaneous → Merge.
2. Input file: Select all AHN2 files in directory that was specified in step 0.
3. Output file: Name new file.



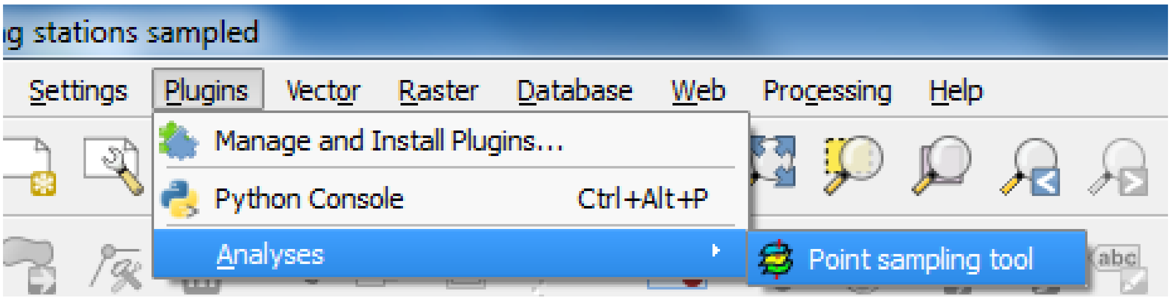


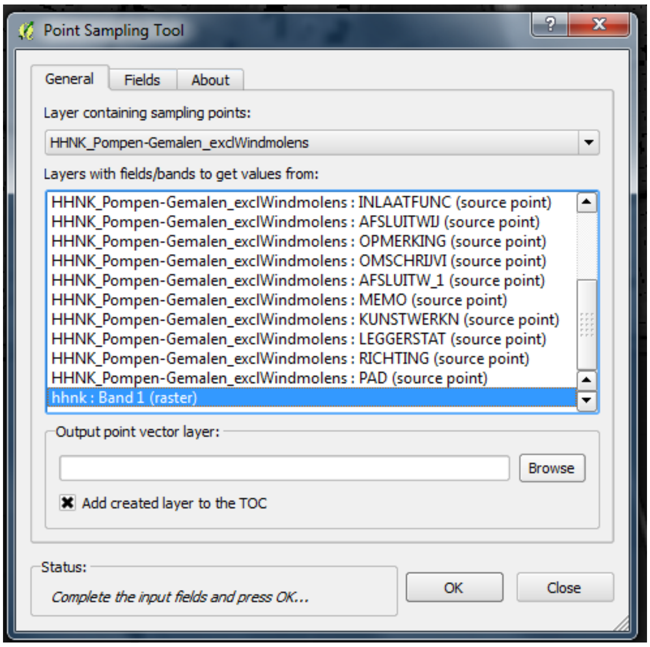


Step 3: Sample Alliander shape file and merged AHN2 file

1. Select Plugins → Analyses → Point sampling tool.
2. Layer containing sampling points: Alliander file.
3. Layers […] to get values from: AHN2 merged file (at the bottom of the list).
4. Output point vector layer: select name and directory for new sample file.
5. RESULT: Attributes table of new sample file contains merged column (for Groundheight value).

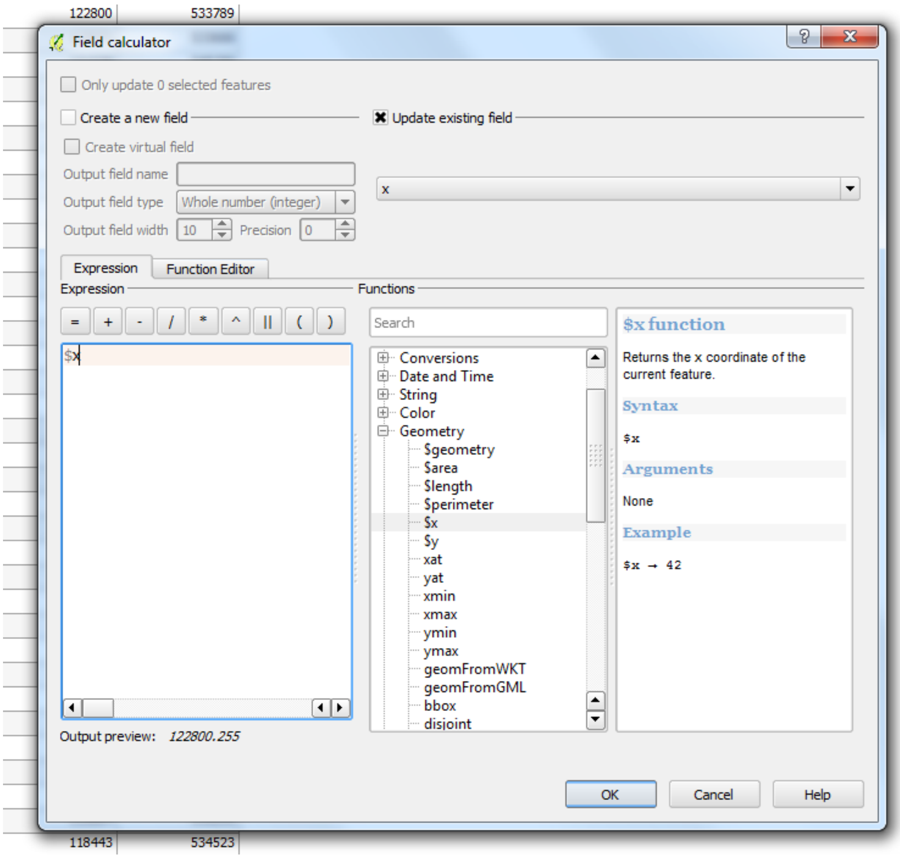
6. Alliander and merged AHN2 layer have to be selected in the layer view beforehand in order to appear in the drop-down lists.





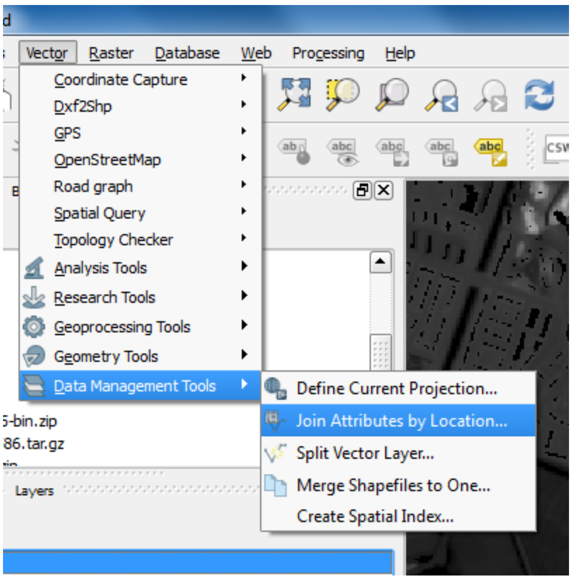
Step 4: Calculate new values (e.g. x- and y coordinates of location of power stations)

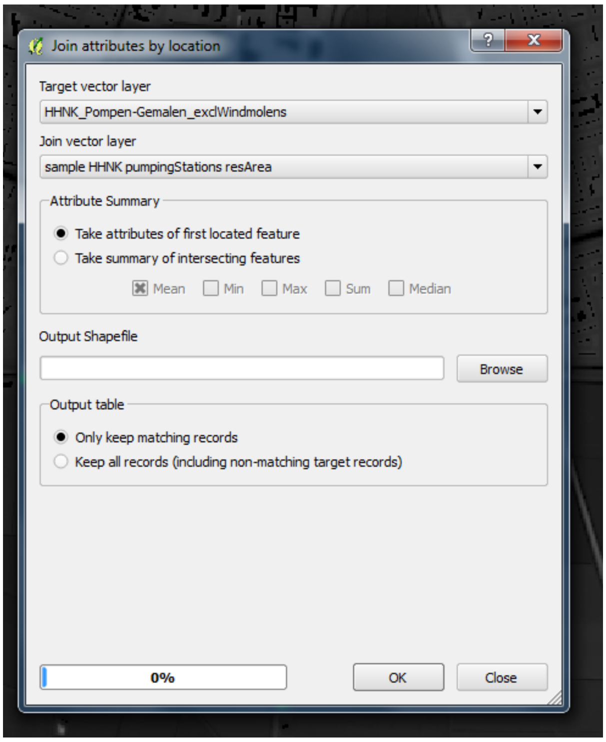
1. Open attributes table of sample file, open field calculator.
2. Create new field, name it XLocation.
3. Select Geometry → $x (double click) → okay.
4. RESULT: Column of XLocation will be filled with data.



Step 5: Create joined file of HHNK shape file and sample file.

1. Select Vector → Data Management Tools → Join Attributes by Location.
2. Target vector layer: Alliander file.
3. Join vector layer: sample file.
4. RESULT: Newly created joined file is added as an additional layer in qGIS. Its attributes table includes 3 more columns: hhnk for the height value, x- and y coordinates.



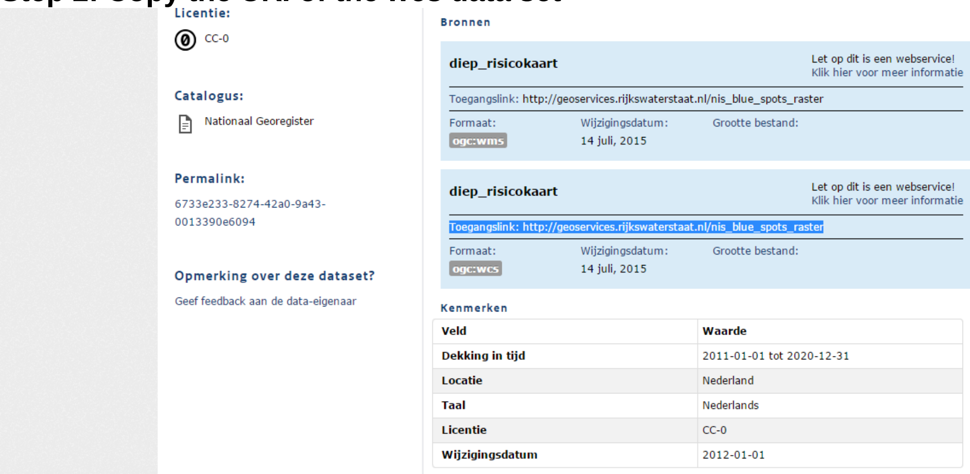

Join attributes by location
Target vector layer
HHNK_Pompen-Gemalen_exclWindmolens
Join vector layer
sample HHNK pumpingStations resArea
Attribute Summary
Take attributes of first located feature
Take summary of intersecting features
Mean
Min
Max
Sum
Median
Output Shapefile
Browse
Output table
Only keep matching records
Keep all records (including non-matching target records)
0%
OK
Close

E Attempt to retrieve water level data in qGIS

This tutorial consists of 9 steps.

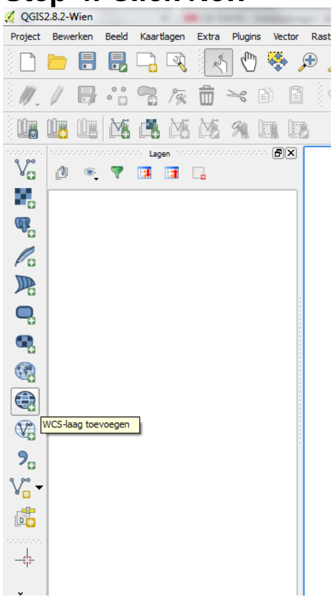
Step 1: Go to the data.overheid website and access the risicokaart web/service. https://data.overheid.nl/data/dataset/waterdiepte-in-risicogebied-risicokaart-nl

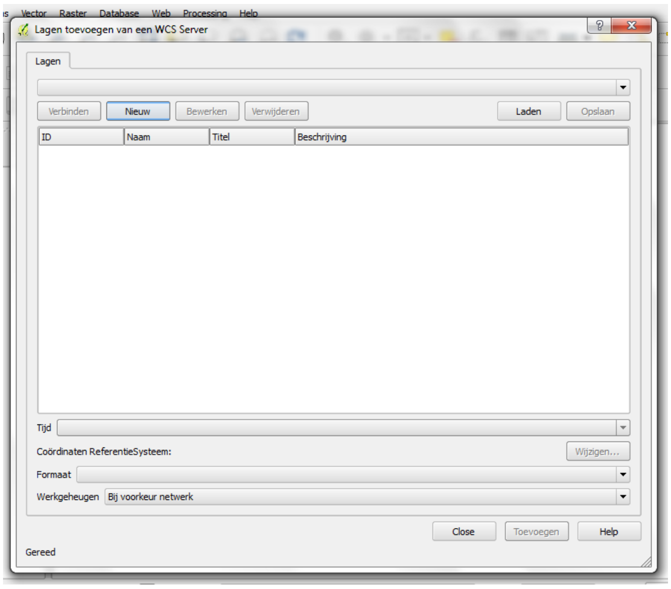
Step 2: Copy the URI of the wcs data set



Step 3: In qGIS open the wcs layer

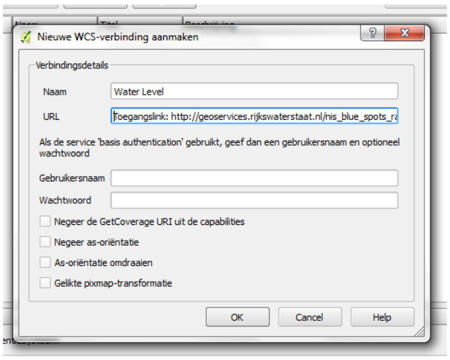
Step 4: Click New



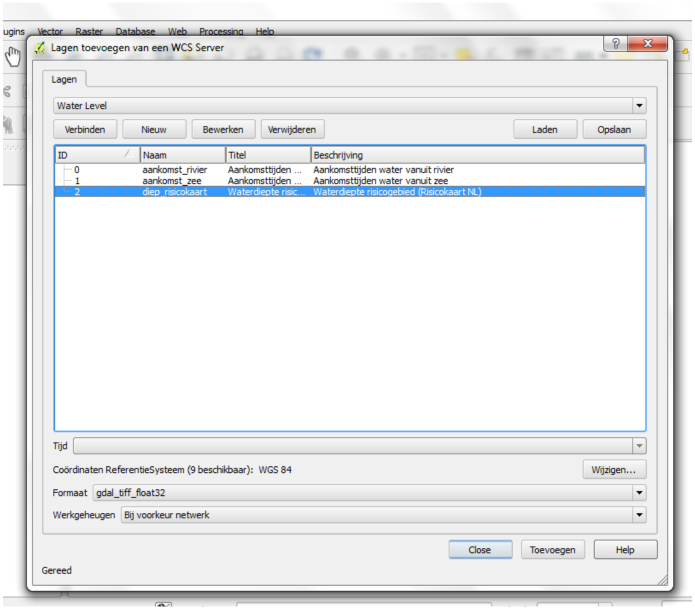


Step 5: Provide a name for the new layer in qGIS)e.g. Water Level'

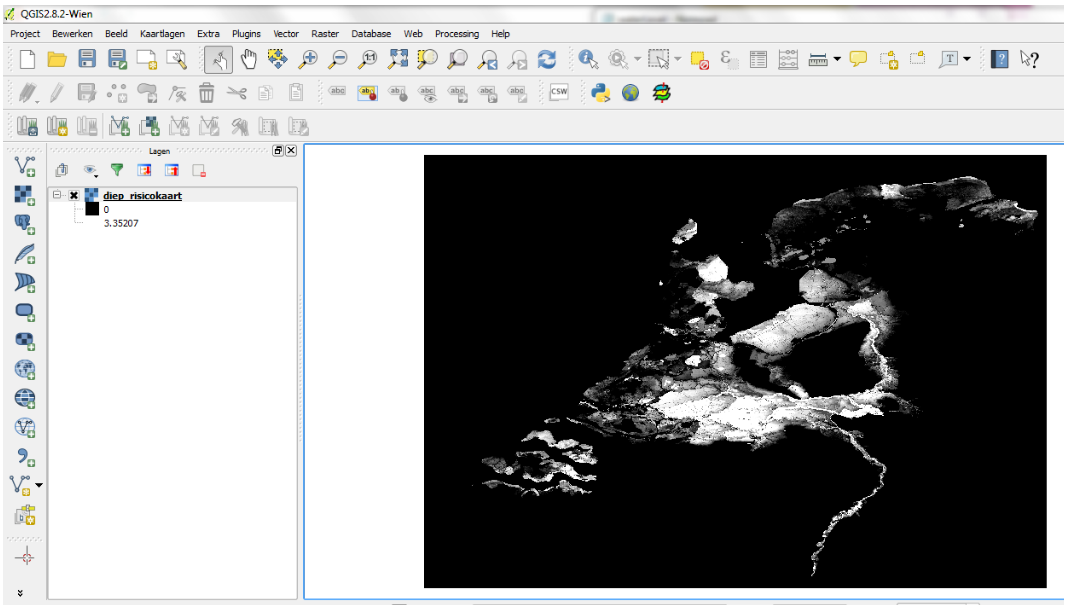
Step 6: Paste the URI from the web/service in the URL field and click okay



Step 7: Click Connect and select diep risicookart from the available options, click add



Step 8: A new layer appears in qGIS



Step 9: If we now try to join the BAG data and the water level data)cf. joining AHN2 and BAG) we receive an empty attribute table. We uploaded the BAG as a second layer and followed the steps in Appendix D. The attributes table of the sample file is then empty rather than showing data of water levels.

F Tutorial for populating the ontology with Web-Karma

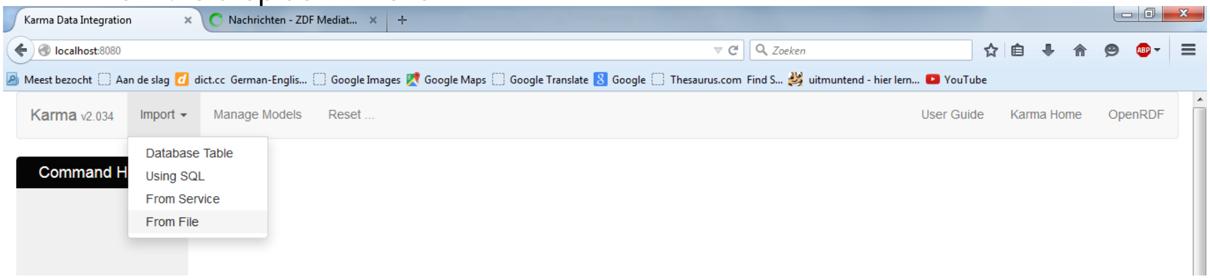
This tutorial uses the data integration tool Web-Karma http://www.isi.edu/integration/karma/. It is based on the official ontology-data mapping tutorial on the web-karma website: using karma to map museum data to the CIDOC CRM ontology. This tutorial consists of 9 main steps, each containing one or more sub steps.

Step 1: Files and storage directory

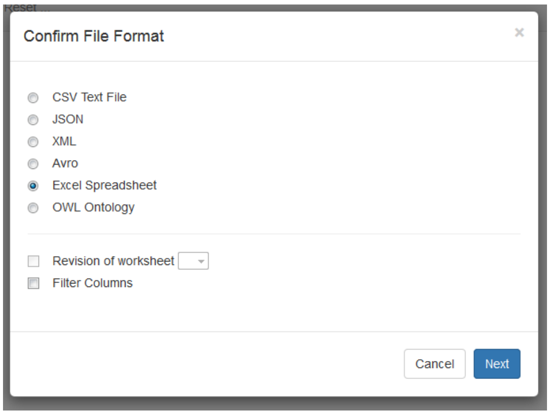
1. Save all data files and the ontology file in the karma-web-services folder so that karma can easily find them.
2. Save ontology file as .owl, since web-karma has difficulties with .ttl files. I created the ontology in TopBraid (.ttl file) and exported it then to Protégé where I saved it as RDF/XML file. This provides the ontology file with the required .owl format. NOTE: TopBraid cannot read .owl even if you have not changed the original TopBraid ontology in Protégé.
3. Example:
   - Our data contains 4 properties of water pumping stations: ID, x-coordinate, y-coordinate, height of the ground surface at the location of a pumping station.
   - The data are stored in an excel file. Each column contains datapoints for one property respectively.
   - Our ontology specifies relationships between electrical grid components (e.g. power stations), electrical flood protective assets (e.g. pumping stations) and overground water levels.
   - The path to our data file is C:/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/joinedAttributes_HHNKpumpstations-AHN2_v1_hxyID/
   - The path to our ontology file is C:/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/CERISE_model1_v3

Step 2: Import data file

1. In web-karma import data file by clicking on “Import” and selecting “From File” from the drop down menu



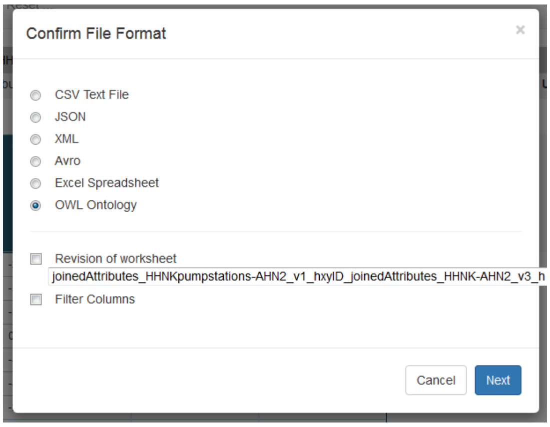
2. Select the appropriate file format. In our case it’s excel. Then click “next”.



3. Click “Import” to confirm that only the first 1000 rows of your data set will be displayed.

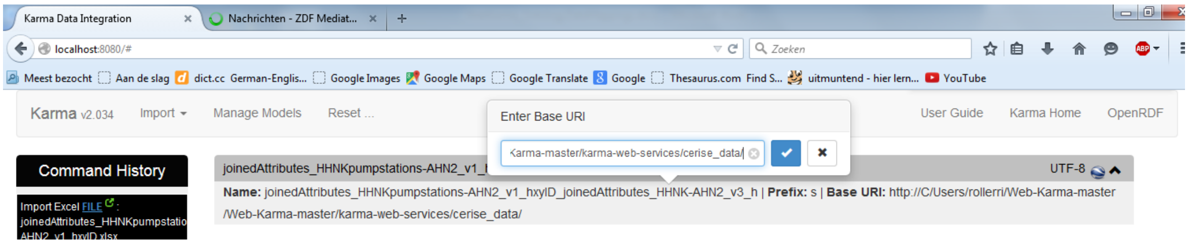
Step 3: Import ontology file

1. Import the ontology file by clicking on “Import”and selecting “From File”from the drop down menu.
2. Select “OWL Ontology” as file format. Then click “Next”.



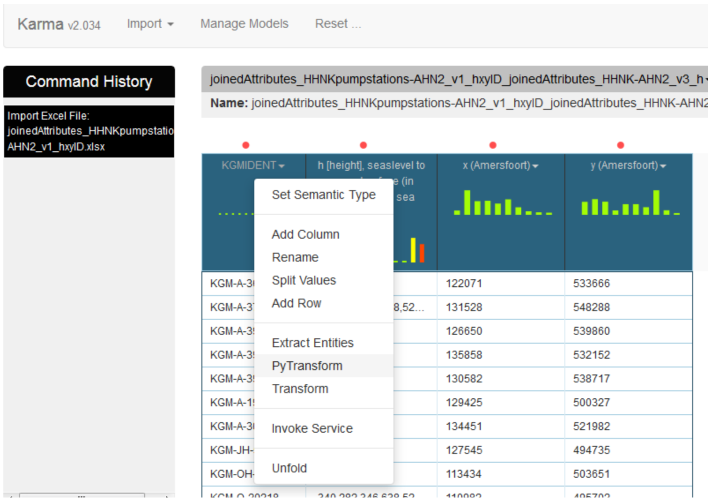
Step 4: Change base URI

1. Change the base URI of the mapping output by clicking on “Base URI” in the upper part on right side of the scrreen.
2. Use the URI of the karma-web services folder where the data file and the ontology file are stored. In our example it is C:/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/

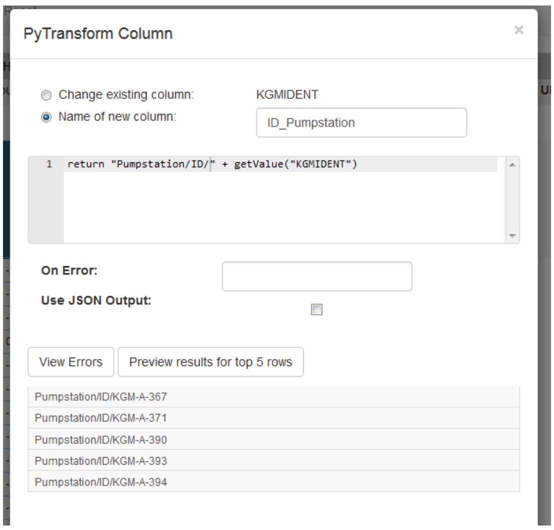


2. Create URIs for all individual datapoints with the Python Transform function

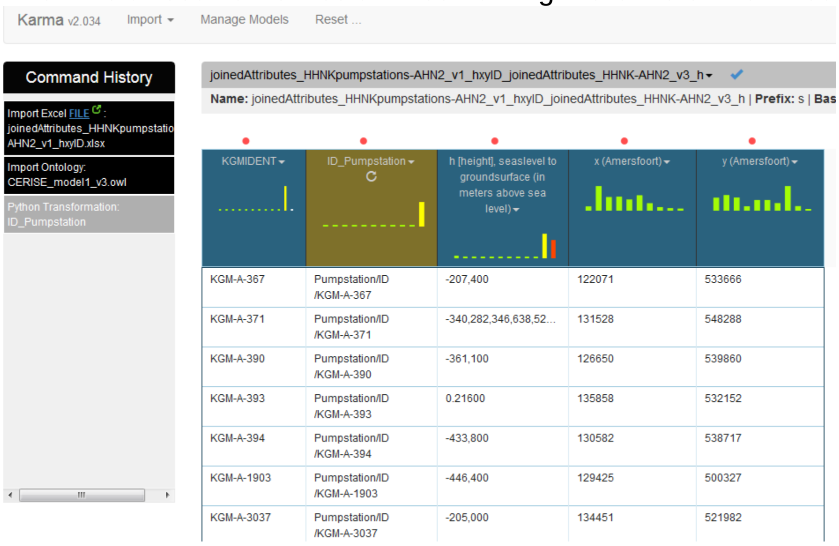
1. Click on the arrow on the right side of the first column's name and select "Py Transform" from the drop down menu. In our example the first column is "KGMIDENT" which represents the ID of pumping stations.



2. Choose a name for the new column which will contain the URIs. In our example we named it "ID_Pumpstation".
3. Use the getValue function in Python to retrieve each datapoint value and to define the URI. Add some meaningful label in quotationmarks and a + between return and getValue to define the URI. In our example it is "Pumpstation/ID/" + The datapoint value is retrieved from the variable in brackets behind the getValue function. In our example it is getValue ("KGMIDENT")
4. Click "Preview results for top 5 rows" to check for errors. Then click save.

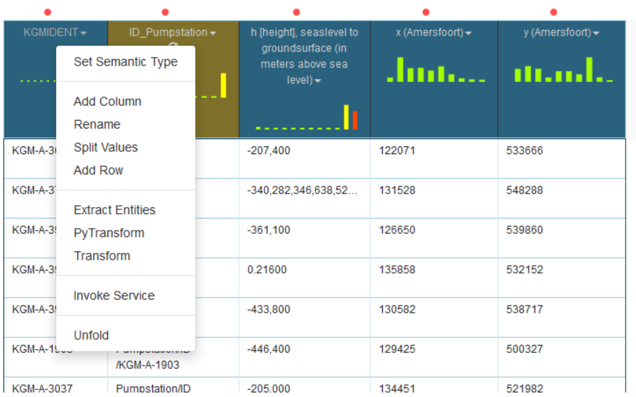


5. As a result a new column containing the URIs of the ID datapoints is created.

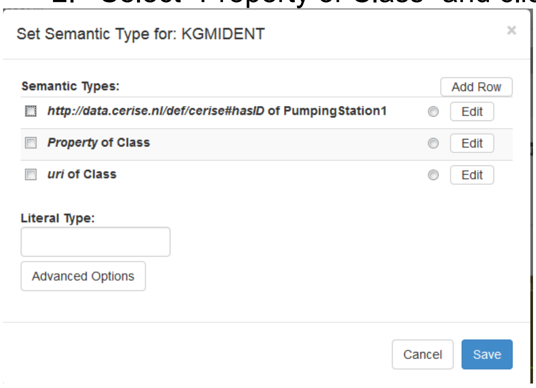


Step 6: Map ontology to data

1. Click on the arrow on the right side of the first column’s name. Select “Set Semantic Type”from the drop down menu.

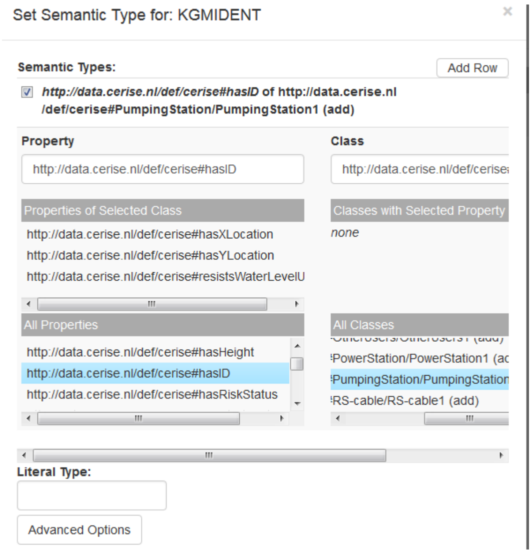


2. Select “Property of Class” and click “Edit” in the same line.

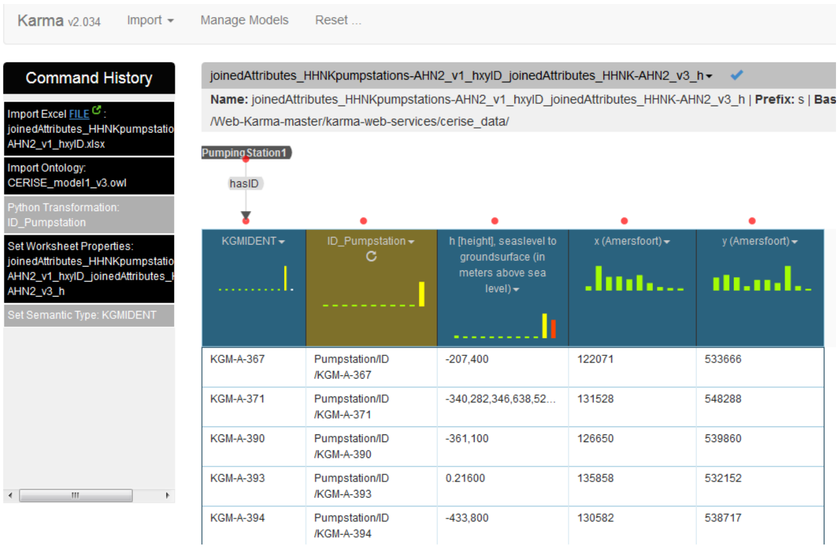


3. On the left hand side we specifiy the property that defines the link between a class of the ontology and the data points. The class can be specified on the right hand side.

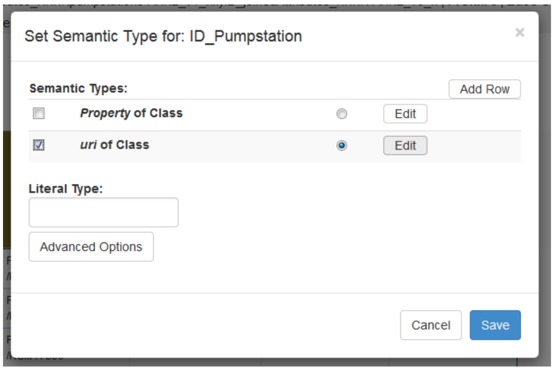
4. As the first column contains the ID of pumping stations the appropriate property is “hasID”, which we selct from the suggestions under “All Properties”. The suggestion feature all properties of the ontology that we uploaded initially.
5. For the class we select “PumpingStation” since all our data describes properties of pumping stations.
6. The selected property and class appear in the white spaces under the labels “Property”and “Class”respectively. Then click save.



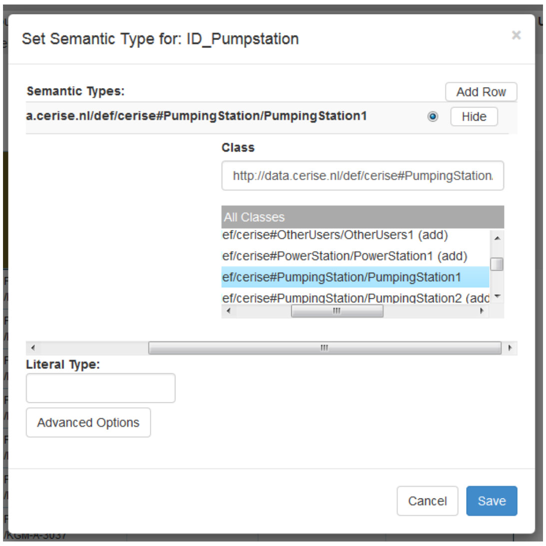
7. Our first mapping is visualized in the workspace: PumpingStation1 hasID KGMIDENT.



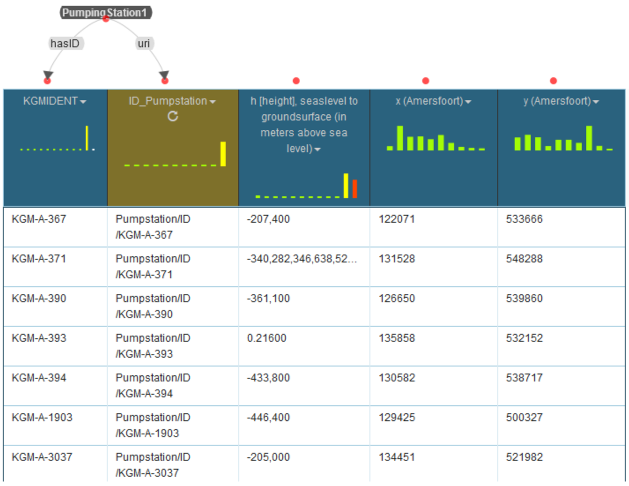
8. We then define the link between PumpingStation and the URI column which we created shortly before.
9. Click on the arrow to the right of the 2nd column’s label. Select “Set Semantic Type” from the drop down menu. Select “uri of Class” and click “Edit” in the same line.



10. Select the appropriate class from the drop down menu. In our example it is "PumpingStation1"



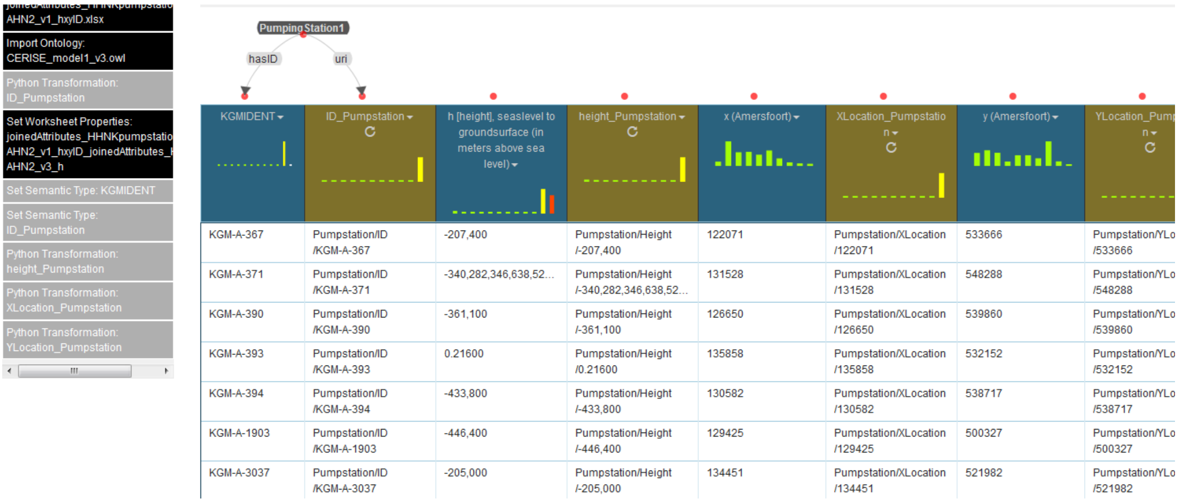
11. The result of the 2nd mapping are displayed in the workspace.



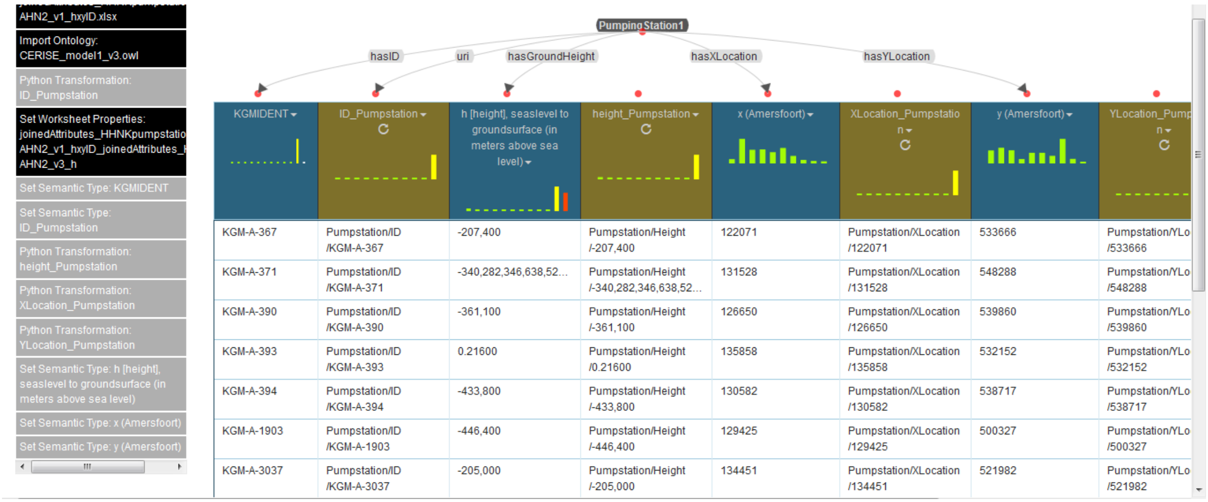
KGMIDENT	ID_Pumpstation	h [height], seaslevel to groundsurface (in meters above sea level)	x (Amersfoort)	y (Amersfoort)
KGM-A-367	PumpstationID /KGM-A-367	-207,400	122071	533666
KGM-A-371	PumpstationID /KGM-A-371	-340,282,346,638,52...	131528	548288
KGM-A-390	PumpstationID /KGM-A-390	-361,100	126650	539860
KGM-A-393	PumpstationID /KGM-A-393	0.21600	135858	532152
KGM-A-394	PumpstationID /KGM-A-394	-433,800	130582	538717
KGM-A-1903	PumpstationID /KGM-A-1903	-446,400	129425	500327
KGM-A-3037	PumpstationID /KGM-A-3037	-205,000	134451	521982

Step 7: More mappings

1. Repeat step 5. For all remaiing columns in order to transform the datapoints into URIs.
2. For each column transformation a new column containing the URIs appears.

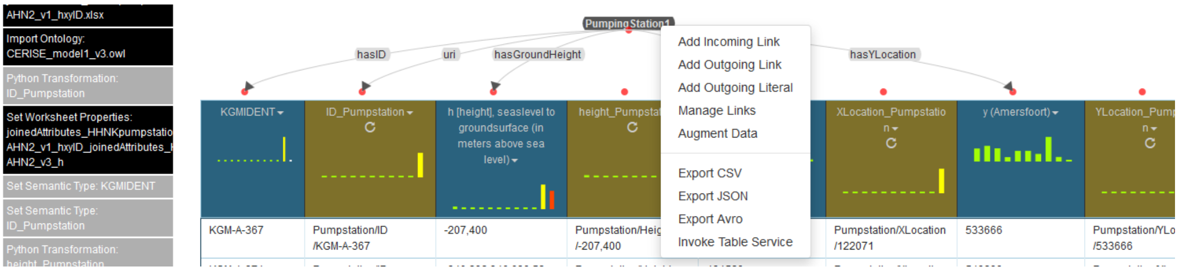


3. Repeat step 6 for all property columns. Each time select an appropriate property from the drop down menu to map the selected class (in our example PumpingStation1) to the data column.

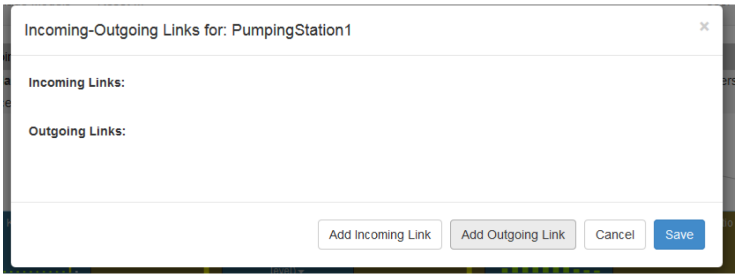


Step 8: Add additional classes

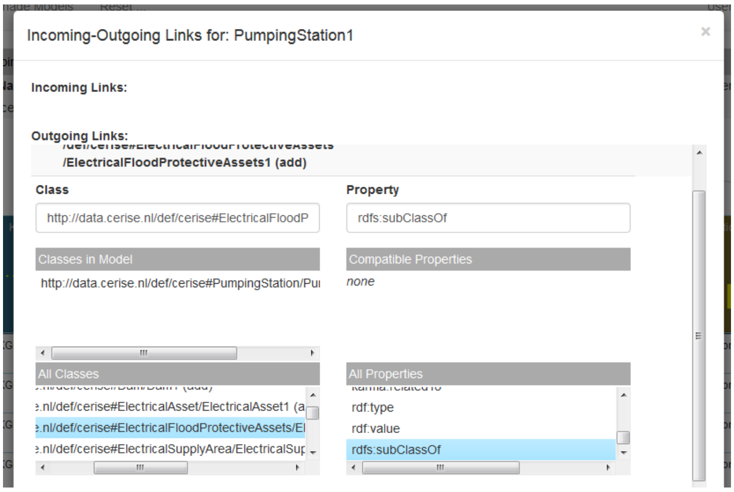
1. Our data set only contained properties for the PumpingStation class. However, our ontology includes more class which can be indirectly mapped to the data by interlinking them with the PumpingStation class.
2. Click on the “PumpingStation” button and select “mange links”from the drop down menu.

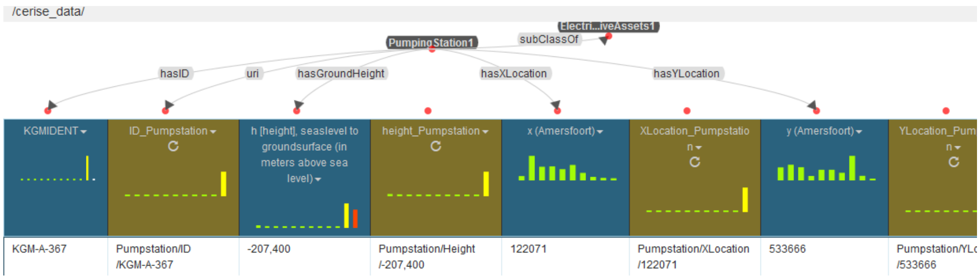


3. You can now specify whether to add an incoming Link or and outgoing link. In our example an incoming link would could state: Class 2 linking property”

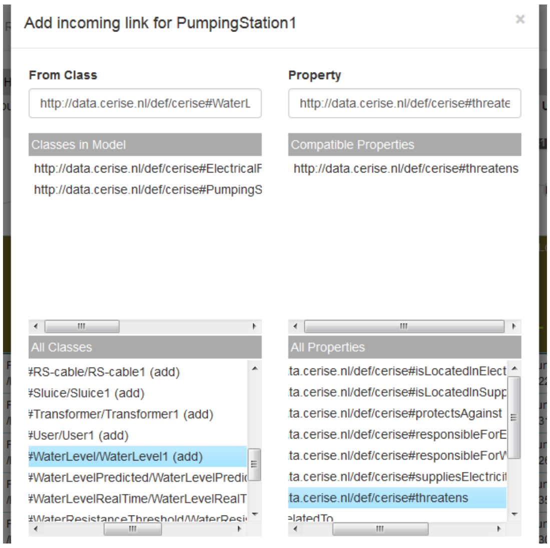


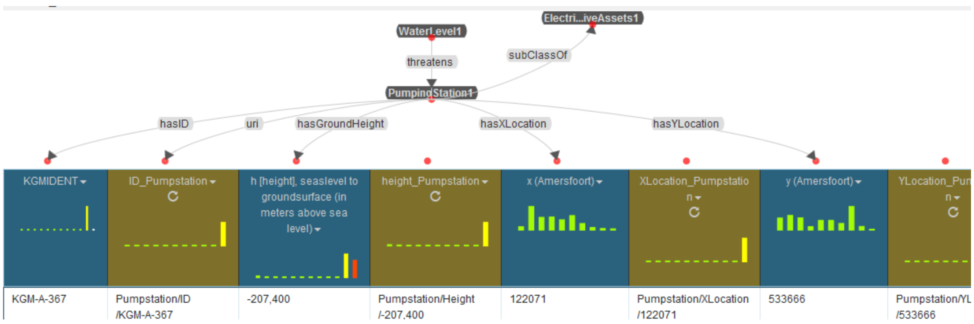
4. To add an outgoing link to our mapping we select a second class (e.g. "Electrical Flood ProtectiveAsset") as a target and an appropriate property that links the pumping station class to this 2nd class. E.g. PumpingStation is a subClass of ElectricalFloodProtectiveAssets.



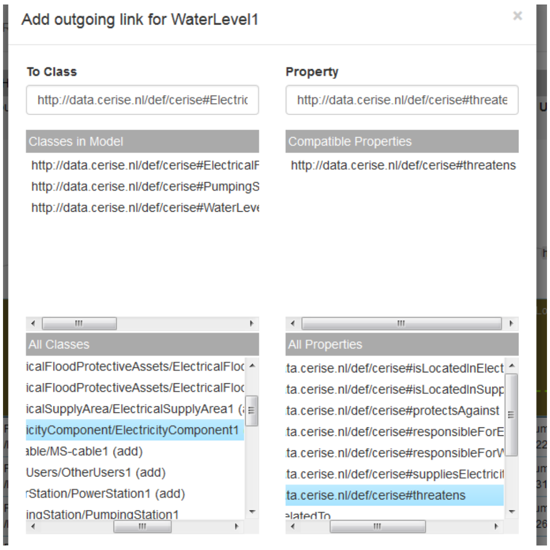


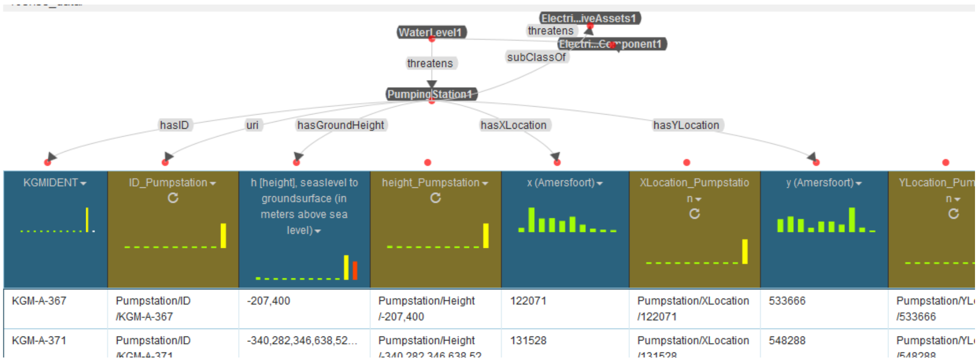
5. In the same we can define an incoming link where a second class is linked to the PumingStation class by a property.
6. In our example waterLevel threatens PumpingStation.





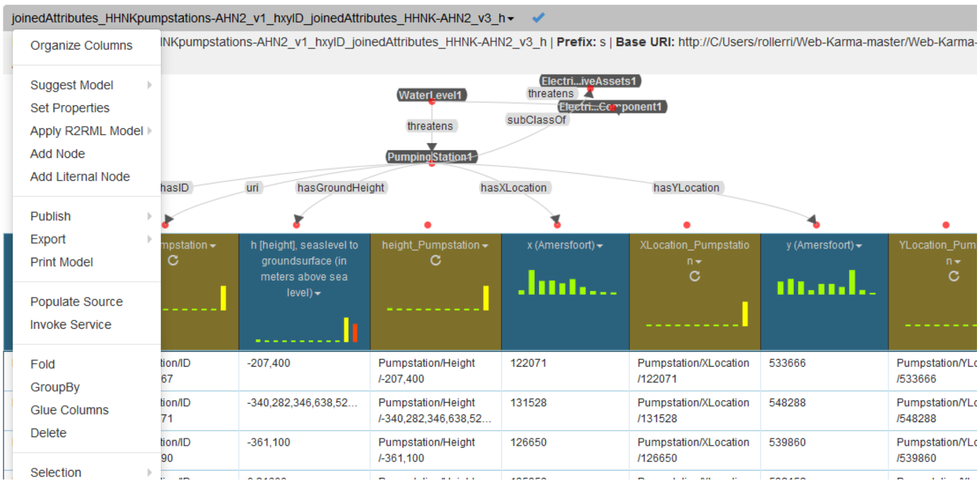
7. The following screenshots provide examples of possible other links
8. WaterLevel threatens ElectricityComponent

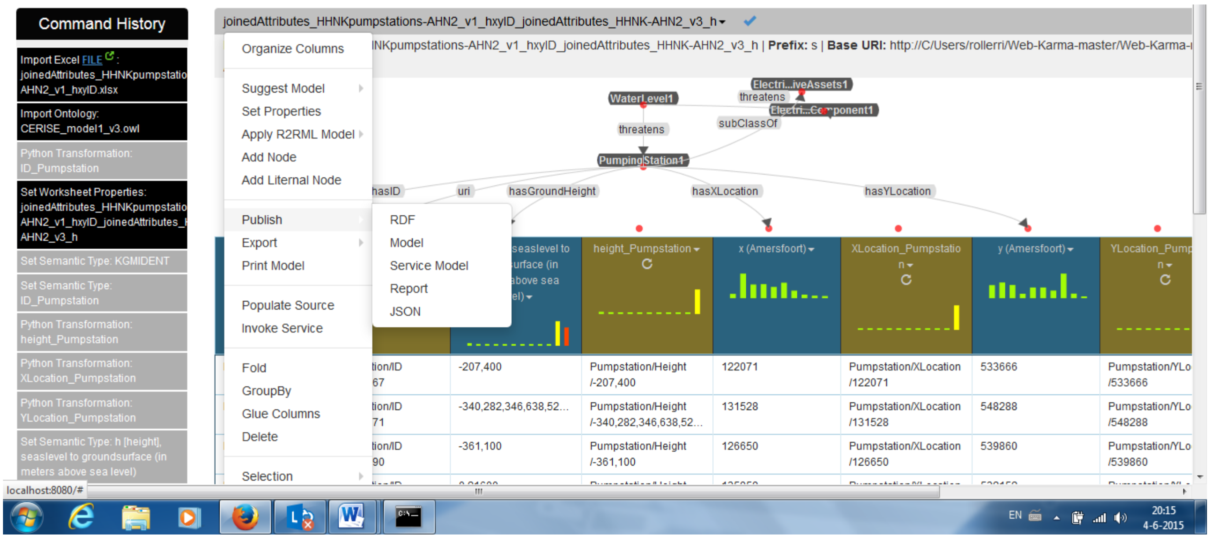


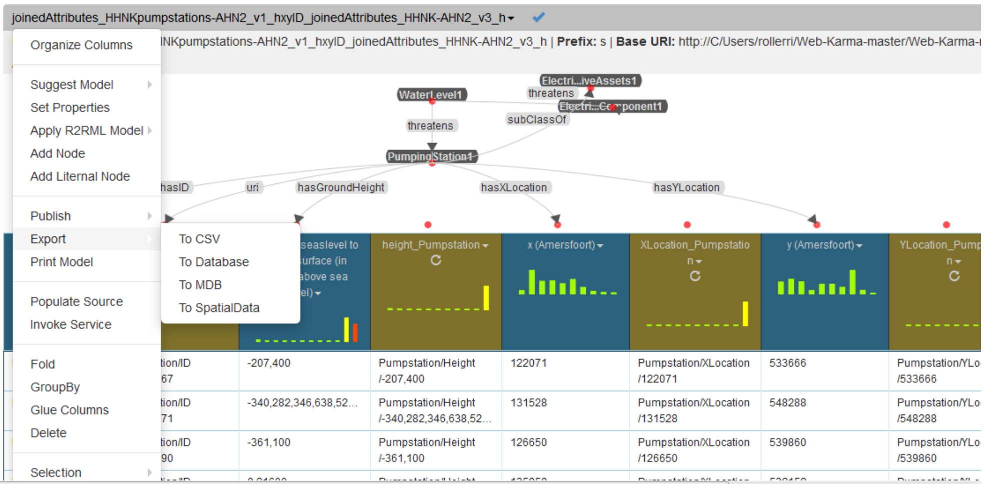


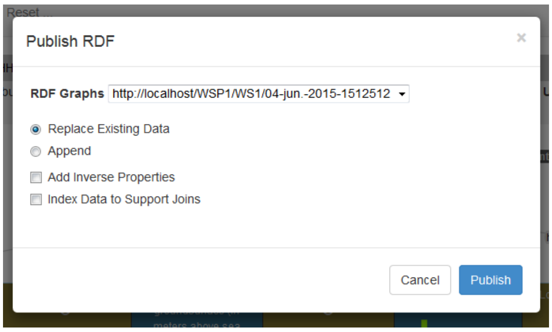
Step 9: Publish and/ or export mappings

1. In order to publish or export the mappings click on the arrow to the right of the data file's name in the dark grey bar. In our example it is "joinedAttributes_HHNKpumpstations-AHN2_v1_hxyID_joinedAttributes_HHNK-AHN2_v3_h
2. Either select "Publish" (e.g. JSON) or "Export" (csv) from the drop down menu. Other formats are also available.

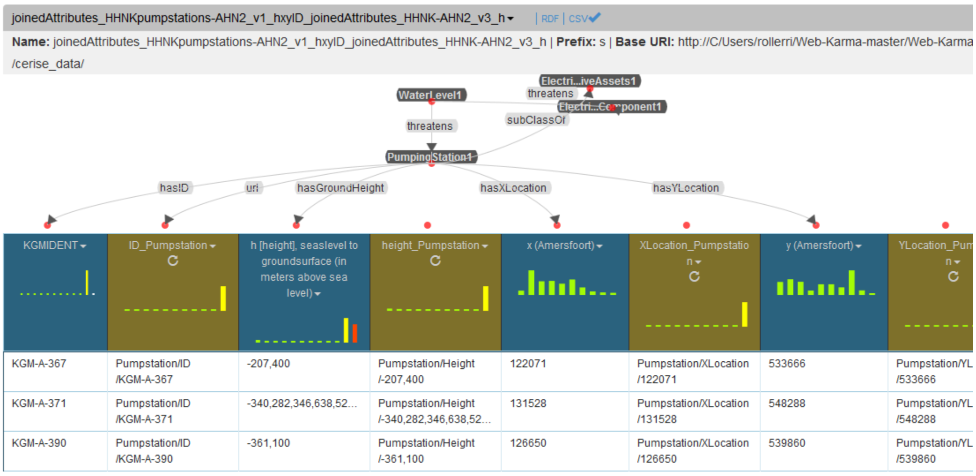




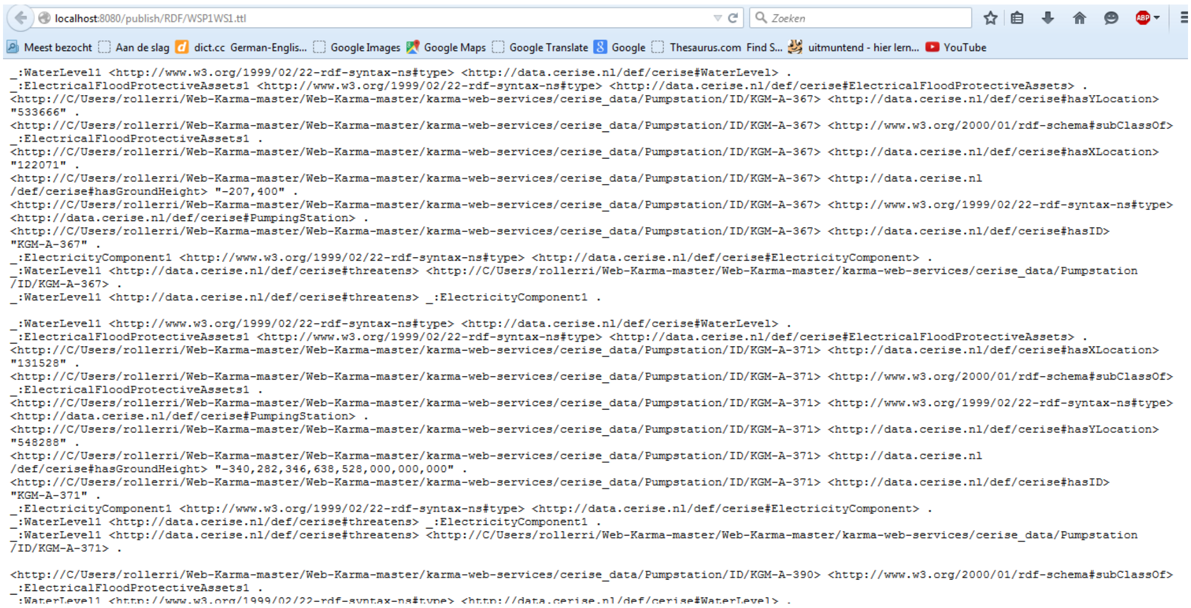




3. As a result of the publishing/ export blue. RDF and CSV hyperlinks will show up on the right side of the dark grey bar.
4. If you click on either of these hyperlinks you will be lead to the respective format storing the mapped data.



5. RDF data result



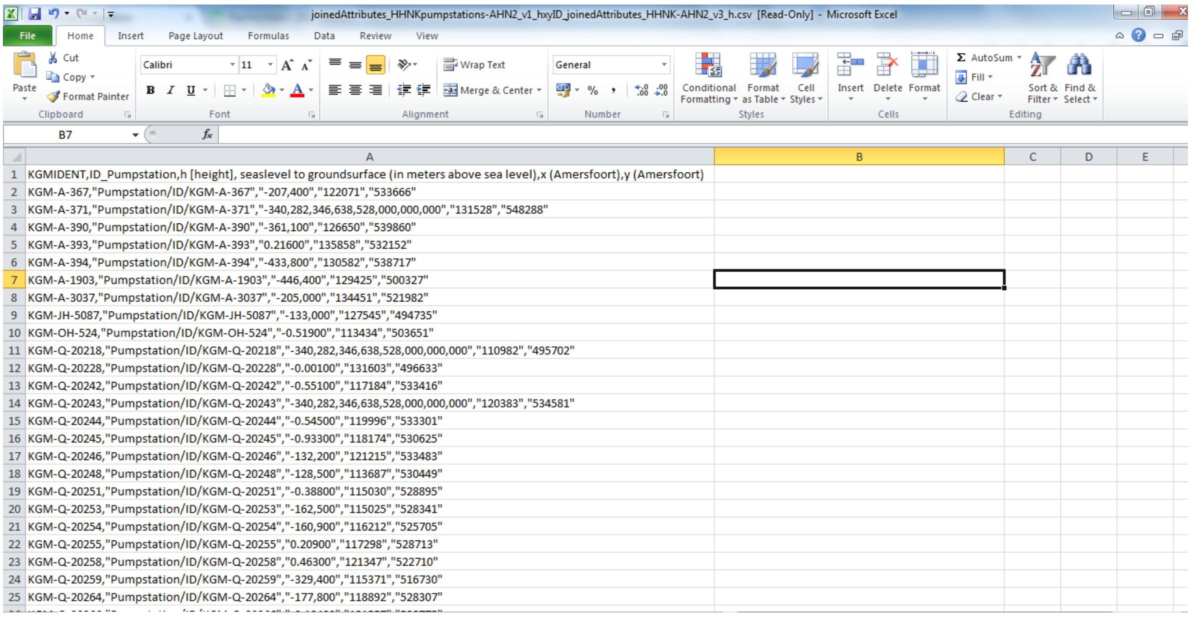

```
localhost:8080/publish/RDF/WSP1WS1.ttl

_:WaterLevel1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#WaterLevel> .
_:ElectricalFloodProtectiveAssets1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#ElectricalFloodProtectiveAssets> .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://data.cerise.nl/def/cerise#hasYLocation>
"533666" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://www.w3.org/2000/01/rdf-schema#subClassOf>
_:ElectricalFloodProtectiveAssets1 .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://data.cerise.nl/def/cerise#hasXLocation>
"122071" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://data.cerise.nl
/def/cerise#hasGroundHeight> "-207,400" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://data.cerise.nl/def/cerise#PumpingStation> .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-367> <http://data.cerise.nl/def/cerise#hasID>
"KGM-A-367" .
_:ElectricityComponent1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#ElectricityComponent> .
_:WaterLevel1 <http://data.cerise.nl/def/cerise#threatens> <http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation
/ID/KGM-A-367> .
_:WaterLevel1 <http://data.cerise.nl/def/cerise#threatens> _:ElectricityComponent1 .

_:WaterLevel1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#WaterLevel> .
_:ElectricalFloodProtectiveAssets1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#ElectricalFloodProtectiveAssets> .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://data.cerise.nl/def/cerise#hasXLocation>
"131528" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://www.w3.org/2000/01/rdf-schema#subClassOf>
_:ElectricalFloodProtectiveAssets1 .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://data.cerise.nl/def/cerise#PumpingStation> .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://data.cerise.nl/def/cerise#hasYLocation>
"548288" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://data.cerise.nl
/def/cerise#hasGroundHeight> "-340,282,346,638,528,000,000,000" .
<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-371> <http://data.cerise.nl/def/cerise#hasID>
"KGM-A-371" .
_:ElectricityComponent1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#ElectricityComponent> .
_:WaterLevel1 <http://data.cerise.nl/def/cerise#threatens> _:ElectricityComponent1 .
_:WaterLevel1 <http://data.cerise.nl/def/cerise#threatens> <http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation
/ID/KGM-A-371> .

<http://C/Users/rollerri/Web-Karma-master/Web-Karma-master/karma-web-services/cerise_data/Pumpstation/ID/KGM-A-390> <http://www.w3.org/2000/01/rdf-schema#subClassOf>
_:ElectricalFloodProtectiveAssets1 .
_:WaterLevel1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://data.cerise.nl/def/cerise#WaterLevel> .
```


6. CSV result



joinedAttributes_HHNKpumpstations-AHN2_v1_hxyID_joinedAttributes_HHNK-AHN2_v3_h.csv [Read-Only] - Microsoft Excel

	A
1	KGMIDENT,ID_Pumpstation,h [height], seaslevel to groundsurface (in meters above sea level),x (Amersfoort),y (Amersfoort)
2	KGM-A-367,"Pumpstation/ID/KGM-A-367","-207,400","122071","533666"
3	KGM-A-371,"Pumpstation/ID/KGM-A-371","-340,282,346,638,528,000,000,000","131528","548288"
4	KGM-A-390,"Pumpstation/ID/KGM-A-390","-361,100","126650","539860"
5	KGM-A-393,"Pumpstation/ID/KGM-A-393","0.21600","135858","532152"
6	KGM-A-394,"Pumpstation/ID/KGM-A-394","-433,800","130582","538717"
7	KGM-A-1903,"Pumpstation/ID/KGM-A-1903","-446,400","129425","500327"
8	KGM-A-3037,"Pumpstation/ID/KGM-A-3037","-205,000","134451","521982"
9	KGM-JH-5087,"Pumpstation/ID/KGM-JH-5087","-133,000","127545","494735"
10	KGM-OH-524,"Pumpstation/ID/KGM-OH-524","-0.51900","113434","503651"
11	KGM-Q-20218,"Pumpstation/ID/KGM-Q-20218","-340,282,346,638,528,000,000,000","110982","495702"
12	KGM-Q-20228,"Pumpstation/ID/KGM-Q-20228","-0.00100","131603","496633"
13	KGM-Q-20242,"Pumpstation/ID/KGM-Q-20242","-0.55100","117184","533416"
14	KGM-Q-20243,"Pumpstation/ID/KGM-Q-20243","-340,282,346,638,528,000,000,000","120383","534581"
15	KGM-Q-20244,"Pumpstation/ID/KGM-Q-20244","-0.54500","119996","533301"
16	KGM-Q-20245,"Pumpstation/ID/KGM-Q-20245","-0.93300","118174","530625"
17	KGM-Q-20246,"Pumpstation/ID/KGM-Q-20246","-132,200","121215","533483"
18	KGM-Q-20248,"Pumpstation/ID/KGM-Q-20248","-128,500","113687","530449"
19	KGM-Q-20251,"Pumpstation/ID/KGM-Q-20251","-0.38800","115030","528895"
20	KGM-Q-20253,"Pumpstation/ID/KGM-Q-20253","-162,500","115025","528341"
21	KGM-Q-20254,"Pumpstation/ID/KGM-Q-20254","-160,900","116212","525705"
22	KGM-Q-20255,"Pumpstation/ID/KGM-Q-20255","0.20900","117298","528713"
23	KGM-Q-20258,"Pumpstation/ID/KGM-Q-20258","0.46300","121347","522710"
24	KGM-Q-20259,"Pumpstation/ID/KGM-Q-20259","-329,400","115371","516730"
25	KGM-Q-20264,"Pumpstation/ID/KGM-Q-20264","-177,800","118892","528307"

Conclusion

Ontology-data mappings work really well in web-karma and the visualization is easy to understand. However, it is only suitable if the data used can be published openly as RDF on the web. For the crisis management use case of the CERISE project web-karma was unsuitable as the Alliander data may not be openly published online.

G Tutorial for populating the ontology with TopBraid Composer

This tutorial uses the semantic web tool TopBraid Composer (TBC). This tutorial consists of 3 main steps, each containing one or more sub steps.

Step 0: Preparation
1. Make sure you either have TBC Standard or Maestro version.
2. Convert data shape files into excel. Each file contains properties for one class only. Column labels represent names of these properties.
3. Add a new first column in the excel file and label it with the name of the class that the whole file is assigned to.
4. Copy and paste values of one property into the first column (e.g. ID values). These values will become the names of instances. You can of course also come up with new values.
5. Save the excel files as tab-delimited text files.

Step 1: Import the data
1. Open the ontology in TopBraid Composer.
2. In the Navigator-view click on the folder the ontology is stored in. In this way you select the parent folder for the imported data.
3. Select File → Import → Import tab-delimited spreadsheet file.
4. Select data file via the “Browse” button.
5. Tick Import to the current ontology. Click Next.
6. Tick Don’t import special characters (or another option if that is more suitable).

Step 2: Specify matches between column labels in the dataset and property names in the ontology.
1. Make sure that values in “Property in Ontology” column EXACTLY match the names of the properties in the ontology (case sensitive). In the default version “Property in Ontology” values have the same names as columns in the data set. If you keep the default version, new properties will be added to the current ontology but no data will be attached to them.
2. In the My Class field type the prefix of the ontology followed by the class name (e.g. cm:ElectricalAsset). The default value is without prefix.

Step 3: View instances
1. In the Classes view select the class you have populated with data.
2. Open Instance view to see all the uploaded data points.
3. If you click on one data point its values for all specified properties appear in the Resource form.

Conclusion
By populating the ontology with data they are assigned a semantic value. Data is automatically stored in RDF format but not published on the web. They can now be queried with SPARQL (open SPARQL editor in TBC) to derive novel relationships between data points. Besides tab-delimited spreadsheet files TBC can also import data from other file formats. Of high interest for the crisis management use case is the option “Create Connection to RDBMS source (using D2RQ)”. It links the ontology to a data base which might be relevant to retrieve real-time water levels.