DELFT UNIVERSITY OF TECHNOLOGY

THESIS PROPOSAL

Integrated modeling of utility networks in the urban environment

Androniki Pavlidou
StudentID: *5267536*

1st Supervisor: Giorgio Agugiaro
2nd Supervisor: Jantien Stoter
External supervisor: Jan van Voorst Vader

Delft University of Technology

Contents

1 Introduction **3**

2 Related work **4**
- 2.1 INSPIRE network model . . . 4
- 2.2 CityGML Utility Network ADE . . . 5
- 2.3 Industrial Foundation Classes -IFC . . . 5
- 2.4 ESRI Geometric Network . . . 6
- 2.5 Model for Underground Data Definition and Integration - MUDDI . . . 7
- 2.6 Kabel- en Leiding Informatie Portaal -KLIC . . . 8

3 Research questions **10**
- 3.1 Objectives . . . 10
- 3.2 Scope of research . . . 10

4 Methodology **11**

5 Time planning **12**

6 Tools and datasets used **13**
- 6.1 Datasets . . . 13
  - 6.1.1 Thematic analysis . . . 14
  - 6.1.2 Preliminary suggestions . . . 22
- 6.2 Tools . . . 23

1 Introduction

The different types of underground utility networks are of particular importance for the proper functioning of cities and their sustainable development. Therefore the need for detailed utility network data models is revealed in order to represent, exchange and store their spatial information, with the ultimate goal to manage their relation to other network systems as well as to integrate them, and better understand the interaction between city entities and utility networks [1].

Despite the development that has taken place in the mapping and depiction of city objects from 2D to 3D, the corresponding information related to the underground condition is still absent or quite limited. Regarding the latter definition, we mean the utility networks related to urban infrastructure (e.g. sewage, electricity, water, gas pipes or other public commodities). Taking into account the rapid development of smart cities and the need for further sustainable development of their urban environment, it is revealed the necessity of an accurate location and map of the existing underground condition of utility networks. The expansion of the urban fabric incidental to excavations and trenches, results in the demand for knowledge of the existing underground utility networks, aiming to avoid interruption of services and other costly damages [2]. Considering the complexity of the below-ground networks, their mapping should be implemented in a way that would allow for changes and keep their spatial and thematic information up-to-date [2].

The representation and mapping of the Underground Utility Network (UNN) information should be carried out with respect to existing standards. The main reason for this is the achievement of interoperability and to tackle issues concerning the wide range of different forms. OGC CityGML is operating properly behind UNN applications and specifically by using the Application Domain Extension (ADE) for Utility Networks, CityGML Utility Network ADE [3]. CityGML represents an open international standard, which allows for an integrated representation of all urban entities at city scale [4]. Having the CityGML Utility Network ADE extension, it is given the chance to both model and represent in 3D supply as well as disposal infrastructures, and at the same time it is supported the integrated representation of the various utility networks [5]. This extension ensures the addition of new properties to existing standard CityGML classes as well as of the new objects types. In addition to this model, other models have been developed (Chapter 2) and are constantly evolving to achieve this goal; the integration of underground networks with the above-ground city objects.

The model of the utility networks should be accompanied by the corresponding functional characteristics concerning their usage as well as the link with the above-ground city objects. This connection is of vital importance since the UUN are associated directly with the (above-ground) city objects and specifically with the infrastructure of the cities they belong to [6]. Once the connection has been achieved it is then possible to use the entire model for better understanding as well as management of both above- and under-ground space, which eventually allows for better city planning, making the invisible structures become visible [[7], [8]]. Finally, detailed utility network data models can be further used in cases such as storm drainage, water, electricity, energy planning, as well as facility management [1].

2 Related work

2.1 INSPIRE network model

For the representation of networks the Infrastructure or Spatial Information in Europe (INSPIRE) has developed its model. This model provides information describing the process of designing and integrating networks into existing infrastructure. The Utility networks model belongs to the *family* of Generic Network Model but it has some extra information for each network element (e.g., UtilityLinkSequence, UtilityNode and UtilityNetwork) that is accompanied with its specific application schema [9]. However, this kind of model has some open issues that need to be confronted. One issue is that utility networks are considered as separated objects and they do not take into account the other networks. This results in partial development of the underground relationships. Additionally, given the main concern for achieving the interoperability of the provided datasets, the use of standards for legal reference or contact information as well as the harmonization of the temporal information are other issues should be covered. Finally, the most important issue trying to be covered is the need to host 3D data, which means the ability to integrate the utility networks in 3D [9] (Figure 1). Currently only 2D topological relationships between the networks are supported.

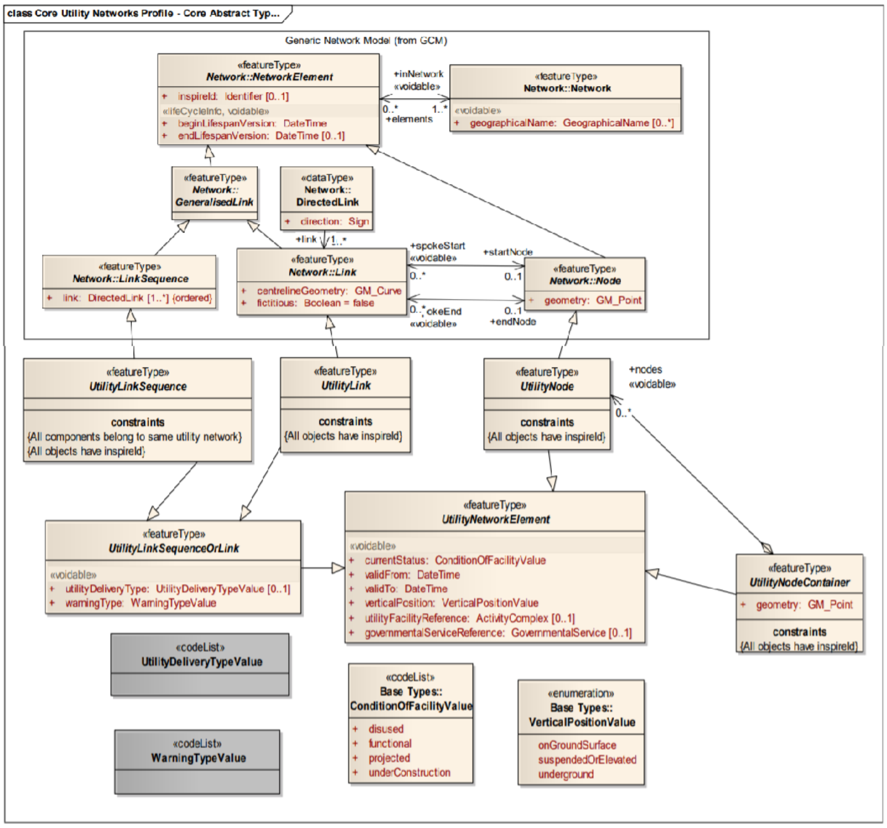


Figure 1: INSPIRE network model

Source: *https://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_US_v3.0rc2.pdf*

2.2 CityGML Utility Network ADE

As it is stated from Becker (et al. 2011) [10], it is of vital importance to integrate the utility network models into their 3D urban context in order to make them functional for multiple uses. For that purpose they proposed a CityGML application domain extension, Utility Network ADE. In this model there are specific classes and relationships representing the network entities used in the various types of utility networks and commodities [11]. With CityGML has been accomplished the definition of a conceptual model that allows for representation, storage and exchange of virtual 3D city models [12]. The ADE for the utility networks has contributed to the addition of new properties to existing CityGML classes that allow their integration.

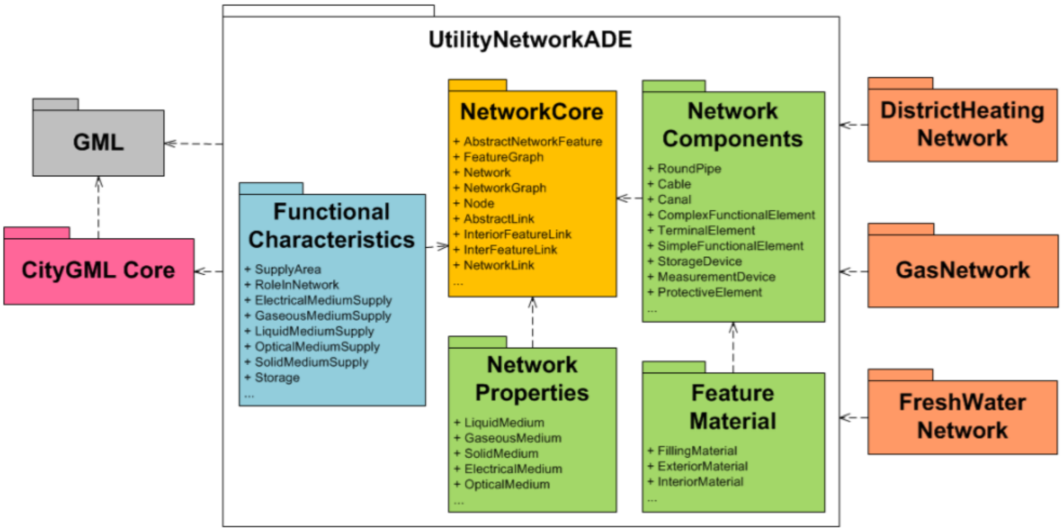


Figure 2: CityGML Utility Network ADE

Source:
https://www.asg.ed.tum.de/en/gis/projects/citygml-utility-network-ade/

2.3 Industrial Foundation Classes -IFC

Industrial Foundation Classes is a schema developed for exchanging data concerning building information model (BIM). Although it is characterised as a description of the built asset industry, it also supports utility networks. However, in this case those networks are limited only to the level of buildings, thus it considers the building service system (inside building networks) [13]. IFC includes a connectivity concept, which allows the representation of the relationship between the different network objects that are categorized according to their role in the network. The limitation of this model, as far as the integration of city scale network is concerned, is the fact that it is restricted to model utilities at the building level [1].

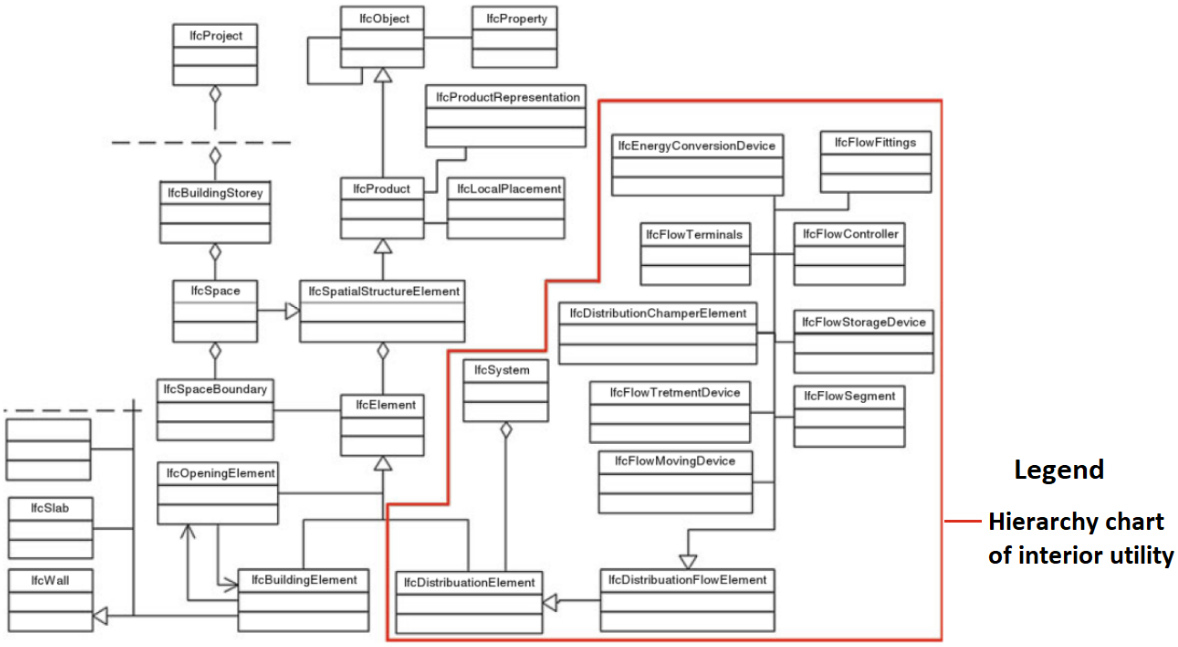


Figure 3: IFC inside building utility network schema

Source: Advances in 3D Geo-Information Sciences edited by Thomas H. Kolbe, Gerhard König, Claus [1]

2.4 ESRI Geometric Network

In ESRI Geometric Network a set of connected edges and junctions is used in order to represent and model the behavior of a typical network infrastructure (above- and below- ground). In more details, there are geodatabase feature classes that support the definition of the geometric network and the corresponding rules that provide information about the way that the resources flow through the geometric network (physical connectivity and logical relation of the network) [14]. Although this model allows for the representation of the linkage between different network objects, it lacks the topographic representation of those objects in 3D [1] (Figure 4).

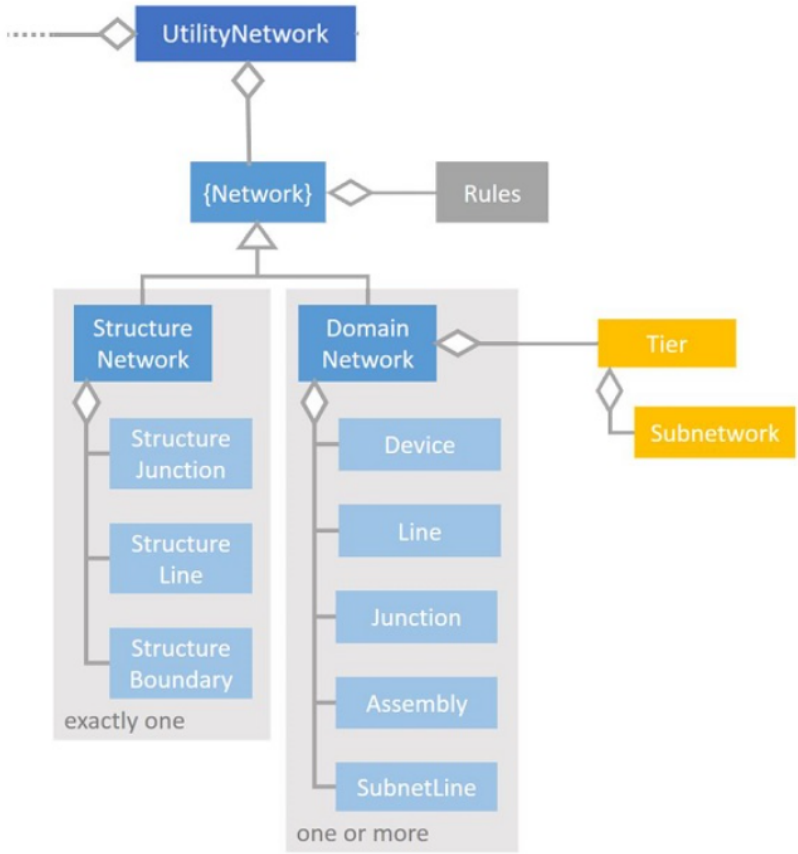


Figure 4: ESRI Geometric Network

Source: Esri Utility Network – Simple and Scalable Network Data Model https://sspinnovations.com/blog/esri-utility-network-simple-and-scalable-network-data-model/

2.5 Model for Underground Data Definition and Integration - MUDDI

In addition to the above mentioned models, MUDDI (Model for Underground Data Definition and Integration) model is intended to serve as a basis for integrating underground data from multiple sources, systems, and schemas [15]. As the previous models, MUDDI consists of individual entities that serve different cases. The aforementioned entities are further specialized in the case of utility networks in order to support real-world utility network types as well as their attributes [15]. MUDDI core model consists of various packages and features including the ones concerning networks, where network entities and relationships are be described (Figure 5). MUDDI is still in active development and focuses on the extension of the original classes.

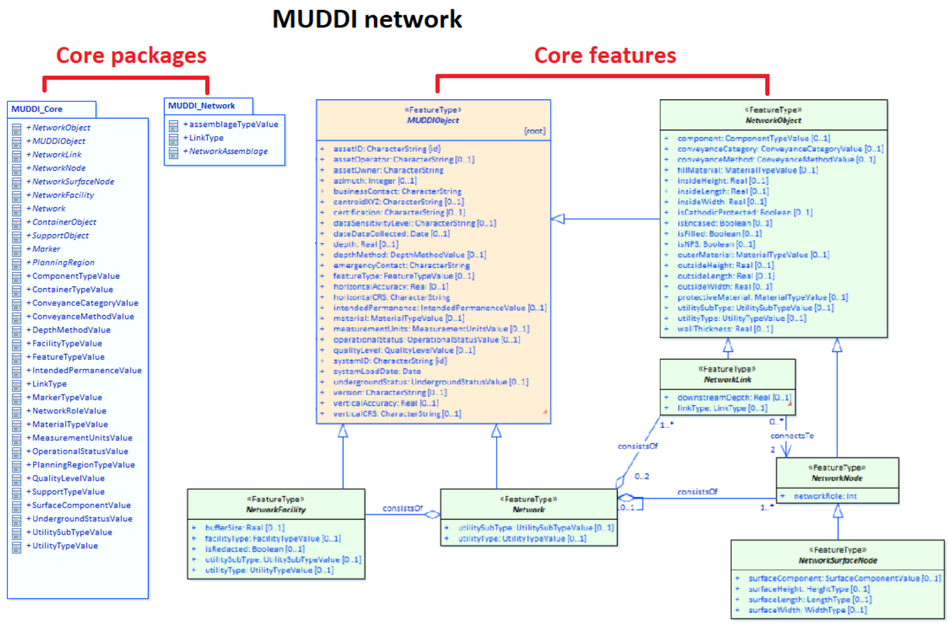


Figure 5: MUDDI utility network environment

Source: Model for Underground Data Definition and Integration Engineering Report http://www.opengis.net/doc/PER/MUDDI

2.6 Kabel- en Leiding Informatie Portaal -KLIC

In the Netherlands a domestic system for the presentation of underground networks has been developed, which follows certain standards set by the country's cadastre. In more details, Dutch cadastre manages the national KLIC facility that includes detailed information about the location of the existed cables and pipelines (IMKL). Apart from their location it is provided the possibility to visualise them through an online environment. This is achieved within the framework of the Information Exchange of above-ground and underground Networks and Networks Act (WIBON) [16]. The IMKL offers a common conceptual framework for data exchange that works on both cables as well as pipelines. The information model is based on NEN3610 and therefore part of the NEN3610 family. NEN3610 is a basic geo-information model that has been developed for semantic coordination in the Netherlands and connection to international standards [17]. In this case, there is not an individual model and the main purpose of this tools is to illustrate the exact location of the pipes without focusing on the integration of the detected networks. In general terms, KLIC model can be considered as an extension of INSPIRE model. The semantic core of IMKL, however, includes extra information for the Dutch user applications, thus only the information have been added extra to the INSPIRE Utility networks model are shown (geometry and network topology of INSPIRE Utility are excluded) [18].

An example of the UML diagram used in KLIC case is illustrated in Figure 6.

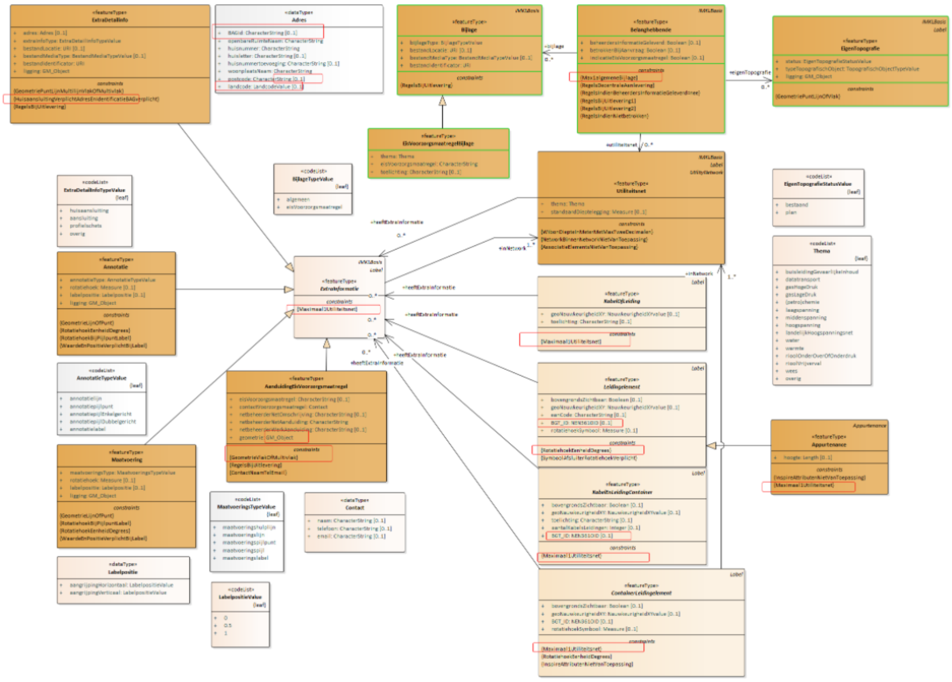


Figure 6: IMKL semantic core for WIBON application.The elements have attributes and a link with additional information.

Source:
https://geonovum.github.io/imkl2015-review/IMKL2.x/1-dataspecificatie/#uml-diagrammen

3 Research questions

3.1 Objectives

The main research question for this thesis is:

- *How is it possible to model utility networks in 3D, integrated with the above-ground objects, such that they can be suitable for multiple uses?*

The aim of this research will be to study the possibilities of implementing the integration in the third dimension of the underground networks using available tools and following the international standards that have been developed for this purpose. Specifically, the desired 3D map will be obtained either by directly utilizing the data of underground networks in the third dimension (if available) or by determining a way to extract the three-dimensional information by combining data derived from different sources. While having the 3D information available it will be examined the possibility of a direct connection of the underground network with the above-ground objects based on their internal relationships (e.g. sewage network connected with the houses, electricity network connected with both houses and city furniture, such as lamps). Considering this another relevant sub-questions derives:

- *How to represent a direct connection with the above-ground condition?*
- *Is it possible to achieve that connection?*

3.2 Scope of research

The thesis will focus on creating a 3D geo-referenced map of a subset of existing utility networks (e.g. sewage) in an area of the TU Delft campus, observing their current condition, and proposing a strategy to harmonize and integrate the existing information by means of existing data models -KLIC, the CityGML Utility Network ADE.

4 Methodology

In order to develop a comprehensive methodology for achieving the desired goal, a quantitative and qualitative analysis of the data is needed (Chapter 6). In more detail, examining the data sets and based on the corresponding completeness needed for a 3D integration, the following steps are to be followed:

1. **Retrieve and understand the provided information**: The data needs to be processed to match an application and its requirements. To do this, the first stage must analyze them qualitatively and quantitatively to better understand their content and related properties (attributes). If the information provided is incomplete or contains topological errors, then appropriate actions should be taken to improve the corresponding quality (i.e. spatial operations).

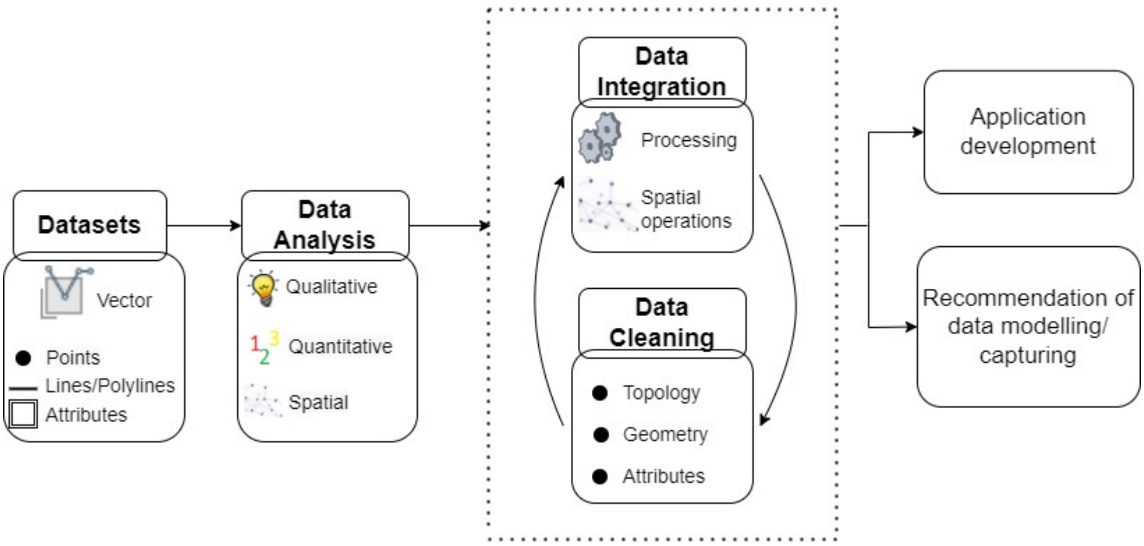
2. **Decide the type of application (use case) will be held to fix the incorrect topology**: To start using the data, an application needs to be decided. In our case, this application will concern the three-dimensional representation of underground utility networks (network integration) and the possibility of connect the with the above-ground objects (e.g. houses) and/or simulating their actual use (i.e. the flow of the water inside pipes).

3. **Evaluate the effort needed to clean up the dataset (based on the above-mentioned application)**: For the implementation of the above application, it is necessary to modify the given data to match its use. To do so, and based on their current situation (see 6), the time needed to be invested on their optimization must be evaluated.

4. **Requirement Analysis - data preparation**: The next crucial step is to define the use case requirements with the ultimate goal of determining a sufficient system design. At this stage, the software to be used for the best possible management, processing, and storage of data will be determined as well.

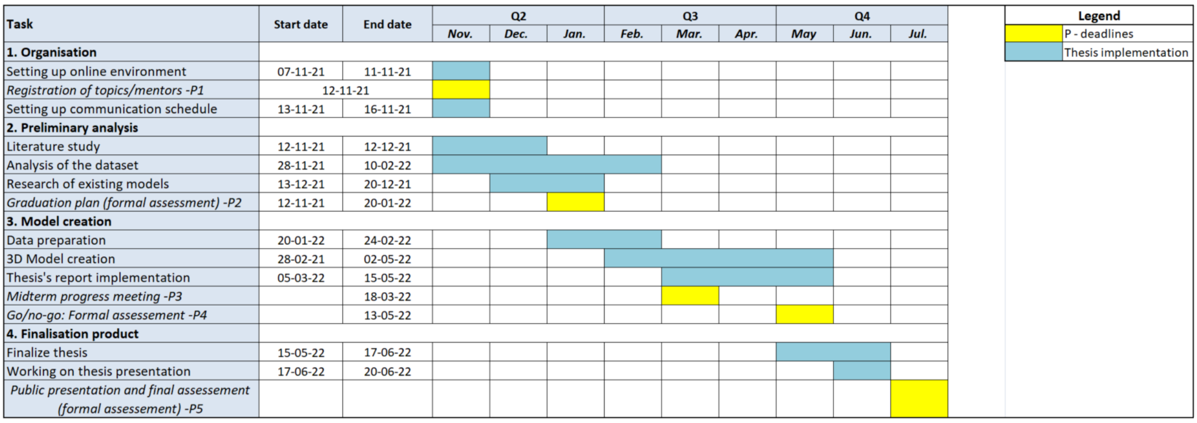
5. **Develop the application**: By completing the above steps it is possible to implement the application, following the available standards, in order to be compatible with the various formats.

Based on the main research question and the possibility to represent the datasets in 3D, other steps may be added to the methodology. Those steps will concern recommendations for their modelling and/or capturing, and how to standardize them, to be able to get integrated for any relevant application.



5 Time planning

In order to have an effective organization of the thesis, a project schedule was created, which depicts the working periods and the respective stages of the dissertation (Figure 7)



Task	Start date	End date	Q2			Q3			Q4		
			Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.
1. Organisation											
Setting up online environment	07-11-21	11-11-21									
Registration of topics/mentors -P1	12-11-21										
Setting up communication schedule	13-11-21	16-11-21									
2. Preliminary analysis											
Literature study	12-11-21	12-12-21									
Analysis of the dataset	28-11-21	10-02-22									
Research of existing models	13-12-21	20-12-21									
Graduation plan (formal assessment) -P2	12-11-21	20-01-22									
3. Model creation											
Data preparation	20-01-22	24-02-22									
3D Model creation	28-02-21	02-05-22									
Thesis's report implementation	05-03-22	15-05-22									
Midterm progress meeting -P3		18-03-22									
Go/no-go: Formal assessement -P4		13-05-22									
4. Finalisation product											
Finalize thesis	15-05-22	17-06-22									
Working on thesis presentation	17-06-22	20-06-22									
Public presentation and final assessement (formal assessement) -P5											

Legend	
	P - deadlines
	Thesis implementation

Figure 7: Graduation calendar

It is noted that during the year some dates may change, while the end dates for presentations have not been set yet.

During the implementation of the thesis, weekly meetings are scheduled to take place in order to resolve problems/ questions and to provide update for its development. The meeting will be held, mainly, with the main supervisor dr. Giorgio Agugiaro, while when necessary the second supervisor, dr. Jantien Stoter will be present. Additionally, meetings with the external supervisor will be held when needed, that are related, mainly, to the content of the datasets and problems concerning their completeness/ quality.

6 Tools and datasets used

6.1 Datasets

For the implementation of the present thesis, a set of datasets was provided. Those datasets concern different types of underground utility networks, accompanied by the corresponding cables and/ or pipes as well as nodes, representing utility networks elements. The data that will be used at the first experimental stage for the thesis implementation (use case), will be a set of vector data concerning a subset of the underground facilities of the Delft Technical University and specifically they will concern sewer network. However, a preliminary study was made for all datasets. The subset concern will be used in the end concern:

- Sewage (vector type: line)
- Sewage (vector type: point)
- Electricity

This selection was made for convenience reasons and more flexible data processing. However, depending on the results of the application, the rest data will be used. The vector data are geo-referenced in the Dutch reference system which is the Amersfoort / RD New Netherlands. The attributes of that reference system are:

- **Geodetic CRS:** Amersfoort
- **Datum:** Amersfoort
- **Ellipsoid:** Bessel 1841
- **EPSG code:** 28992

Before starting the development and creation of any application (and model), it is necessary to study in detail the provided data, in order to understand their content. Any set of data is distinguished for its completeness, both qualitatively and quantitatively. Starting with the quality of the data, we refer to the elements that characterize them; metadata (transcription) that give a description of their content as well as to their completeness (e.g. clean data free of useless and irrelevant information- or even wrong) [19]. In our case the metadata was not present and an extensive study of their attribute table was implemented in combination with the visual comparison in the environment of QGIS. During this examination attention was paid to all the included units accompanied by their units of measurement (if available). Taking into account that a network model represents reality the corresponding units must conform to reality. The qualitative analysis of the data includes, also, their organization based on their content or internal connection, in order to be able to retrieve them. Having completed these steps, the process of getting acquainted with the material has essentially begun and therefore it is possible to start doing some preliminary coding in order to process them so that they serve our purpose and the dataset reliable and valid according to existed standards [19]. As far as the quantitatively analysis of the data is concerned, we refer to the process followed to collect as well as analyze numerical data (Tables 1, 2). In more details, the attention this time will be paid mostly on finding patterns, predictions, as well as cause-effect relationships between the variables being studied [20].

6.1.1 Thematic analysis

In order to familiarize and analyze the dataset, QGIS version 3.16 and FME Workbench version 2021.1.1 were used. Specifically the provided data concerned apart from sewage pipes and cables, networks concerning electricity, gas transmission, telecommunication network (signal cables) as well as other particular points that are relevant to the starting and ending point of the pipe or other important element of the networks. Some examples of the provided dataset and their content are shown in Figures 8, 16 and 17 where the different colors illustrate the different types and uses of the pipes and/or points. For each dataset, the corresponding attribute table was examined in order for its content to be understood as well as to detect possible inconsistencies.

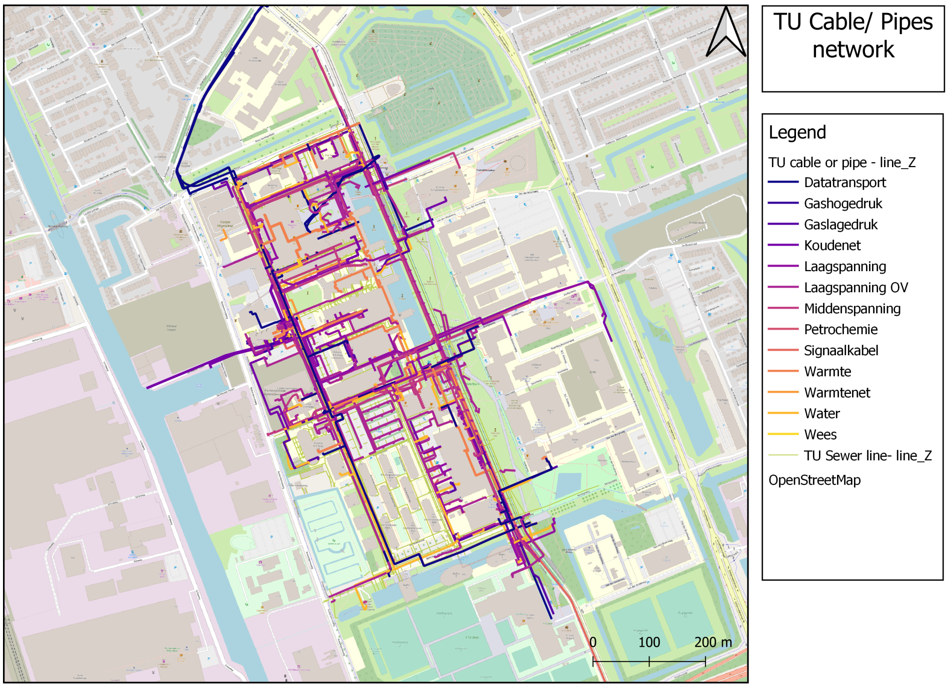
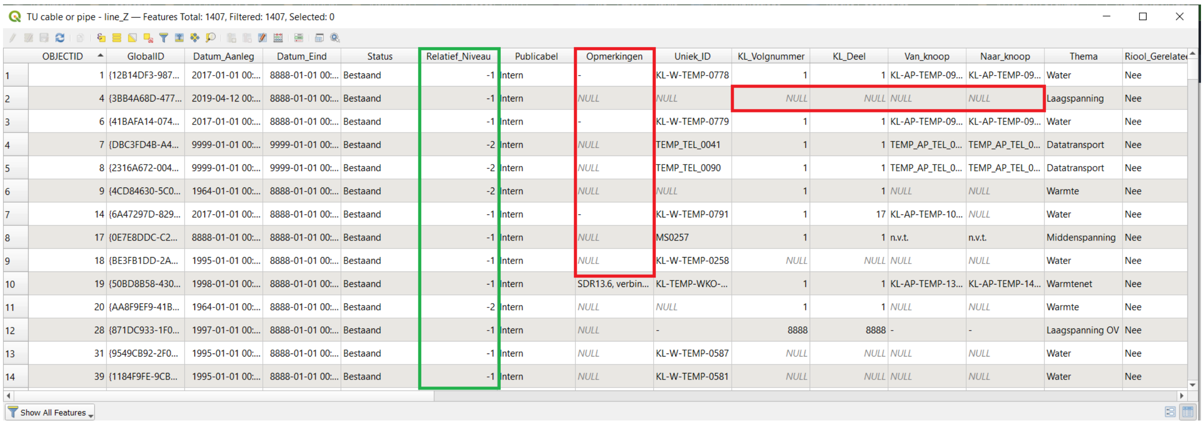


Figure 8: TU cable/ pipes network

In Figure 8 are depicted the different types of cables and pipes existed in the study area, except from the network datasets used in this thesis and located in the TU Delft campus. Specifically, there are thirteen (13) types concerning: *data transport, gas (high and low pressure), cooling and heating network, low and medium voltage, petrochemical, drink water and 'left over' pipes*. The attribute table of the present layer contains information about the current status of the underground utilities (active or not), their relative height in relation to a reference surface (which is unknown), data concerning their starting and ending point, their type, length and width as well as other construction information (e.g cable type, dimensions etc.). However, these recordings are not all up-to-date, while others are generally not available or even unknown (Figure 9). In addition, the unit of measurement of the data that are physical quantities (measurable) is not distinguished from the table, something that makes more difficult to perceive their real size, while in the case of different units between the different quantities, confusion is created (Figure 9).



TU cable or pipe - line_Z — Features Total: 1407, Filtered: 1407, Selected: 0

	OBJECTID	GlobalID	Datum_Aanleg	Datum_Eind	Status	Relatief_Niveau	Publicabel	Opmerkingen	Uniek_ID	KL_Volgnummer	KL_Deel	Van_knoop	Naar_knoop	Thema	Riool_Gerelatee
1	1	{12B14DF3-987...	2017-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	-	KL-W-TEMP-0778	1	1	KL-AP-TEMP-09...	KL-AP-TEMP-09...	Water	Nee
2	4	{3BB4A68D-477...	2019-04-12 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	NULL	NULL	NULL	NULL	NULL	Laagspanning	Nee
3	6	{41BAFA14-074...	2017-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	-	KL-W-TEMP-0779	1	1	KL-AP-TEMP-09...	KL-AP-TEMP-09...	Water	Nee
4	7	{DBC3FD4B-A4...	9999-01-01 00:...	9999-01-01 00:...	Bestaand	-2	Intern	NULL	TEMP_TEL_0041	1	1	TEMP_AP_TEL_0...	TEMP_AP_TEL_0...	Datatransport	Nee
5	8	{2316A672-004...	9999-01-01 00:...	9999-01-01 00:...	Bestaand	-2	Intern	NULL	TEMP_TEL_0090	1	1	TEMP_AP_TEL_0...	TEMP_AP_TEL_0...	Datatransport	Nee
6	9	{4CD84630-5C0...	1964-01-01 00:...	8888-01-01 00:...	Bestaand	-2	Intern	NULL	NULL	1	1	NULL	NULL	Warmte	Nee
7	14	{6A47297D-829...	2017-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	-	KL-W-TEMP-0791	1	17	KL-AP-TEMP-10...	NULL	Water	Nee
8	17	{0E7E8DDC-C2...	8888-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	MS0257	1	1	n.v.t.	n.v.t.	Middenspanning	Nee
9	18	{BE3FB1DD-2A...	1995-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	KL-W-TEMP-0258	NULL	NULL	NULL	NULL	Water	Nee
10	19	{50BD8B58-430...	1998-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	SDR13.6, verbin...	KL-TEMP-WKO-...	1	1	KL-AP-TEMP-13...	KL-AP-TEMP-14...	Warmtenet	Nee
11	20	{AA8F9EF9-41B...	1964-01-01 00:...	8888-01-01 00:...	Bestaand	-2	Intern	NULL	NULL	1	1	NULL	NULL	Warmte	Nee
12	28	{871DC933-1F0...	1997-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	-	8888	8888	-	-	Laagspanning OV	Nee
13	31	{9549CB92-2F0...	1995-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	KL-W-TEMP-0587	NULL	NULL	NULL	NULL	Water	Nee
14	39	{1184F9FE-9CB...	1995-01-01 00:...	8888-01-01 00:...	Bestaand	-1	Intern	NULL	KL-W-TEMP-0581	NULL	NULL	NULL	NULL	Water	Nee

Show All Features

Figure 9: Example of the attribute table content (in red is given an example of empty records and in green the absence of the unit of measurement)

From the above attribute table it can be observed that many records are empty (NULL) or they contain invalid information (e.g. "8888", which is not a valid value). Traversing all the attribute tables from the provided datasets it is noted that only 2/6 datasets contain some fully completed information while the rest are incomplete (Table 1).

Table 1: Data statistics

Dataset/ completeness	Total number of rows	Rows with complete information	Rows with incomplete information
TU Cable or Pipe line Z	1407	-	1407
TU Cable element point Z	912	34	878
TU Cable pipeline Z	1407	-	1407
TU Sewer knot point Z	1986	-	1986
TU Sewer line Z	2080	19	2061
TU pipe element point Z	1038	-	1038

It is noted that the above table strictly presents the information that is available or not in the datasets, without separating its correlation with those (information) related to our application. The information that is considered directly related to the present application is the length/shape of the underground pipelines, the coordinates of the corresponding point elements, their date that indicates if the data is up to date, their type as well as its use.

Until now the attention was paid on the completeness of the dataset regarding the contained records of the corresponding attribute table. Moving the analysis further to the extraction of the 3D information, it was observed that only a few records of the datasets provided indeed a 3D information (Z value-Z coordinate). Specifically, with the help of FME software, the coordinates included in the geometry of the datasets were extracted and by implementing a statistical analysis, Table 2 was created.

Table 2: 3D vs 2D information

Dataset/ completeness	Total number of rows	Missing Z coordinate	Zero Z coordinate	Existed Z coordinate	Percentage of existed 3D information	Percentage of existed 2D information
TU Cable or Pipe line Z	1407	0	1361	46	***3.27 %***	***96.73%***
TU Cable element point Z	911	0	909	2	***0.22%***	***99.78%***
TU Cable pipeline Z	1407	0	1361	46	***3.27 %***	***96.73%***
TU Sewer knot point Z	1986	0	1954	32	***1.61%***	***98.39%***
TU Sewer line Z	2080	2	2048	30	***1.44%***	***98.46%***
TU pipe element point Z	1037	0	1005	33	***3.18%***	***96.91%***

From the above table (Table 2) it can be observed that datasets cannot be characterized as 3D, on the contrary, they are mainly 2D with some exceptions. Still, given the lack of information about the units of measurement, the reference surface as well as their concurrent validity, the information provided as Z coordinate cannot be considered completely reliable. An example of that no-reliability can be seen in Figure 10, where is illustrated a part of the *TU Sewer knot point* attribute table and a non-logical value concerning Z information (height regarding 3D information).

OBJECTID	GlobalID	Type_Riool	Afvoertype	Lengte	Breedte	Hoogte_Put	_indices{0}.x	_indices{0}.y	_indices{0}.z
2	**{5163D424-B57E-41CE-BABE-000E85AA582E}**	**Kolk-Straatkolk**	**Hemelwater**	**0**	**0**	**-0.75300002**	**85059.4551**	**446438.1341**	**32**
12	{31C236F4-FD43-4007-BFA8-008EBC20D1DF}	Kolk-Trottoirkolk	Hemelwater	0	0	-0.47299999	85256.3561	445984.2761	0
23	{3BBD5EE3-E4B5-43AF-87FD-00EB234ACFC0}	Inlaat	Hemelwater	0	0	8888	85106.0547	446323.5937	0
25	{4F4C0F95-33FC-4850-8ADF-010941EDACFC}	Kolk-Straatkolk	Hemelwater	0	0	-0.648	85249.0681	446060.8961	0
28	{323D91E4-BDAF-4F9C-8035-012E80305849}	_onbekend	Hemelwater	8888	8888	8888	85214.6631	446475.3391	0
30	{92F65493-F938-4E0E-A931-0132EDCB44E0}	Uitlaat	Oppervlaktewater	8888	8888	8888	85253.6971	445797.7731	0
33	{FC3E60DE-FD09-4D94-A044-01A09173FD16}	Inlaat	Hemelwater	0	0	8888	85156.0001	446524.8751	0
36	{D0D3D407-E2D5-490B-87CA-01DC19DA5EA8}	Inlaat	Hemelwater	0	0	8888	85210.7735	446270.4063	0
39	{24E8E27D-7AA3-4DF5-82D1-02017C412657}	Inspectieput	Hemel_en_grondwater	8888	8888	-0.59100002	85409.3341	445857.6181	0
42	{92B1218B-4753-4DD2-B2E9-0223E54FE7EA}	Inspectieput	Hemel_en_grondwater	8888	8888	-0.78100002	85291.0231	446459.8451	0
44	{4C700D66-4634-45AB-B13B-022BFF5E1C57}	Controleput	Vuilwater	315	315	0	85163.5461	446044.5261	0
46	{0DFB47CF-016D-453F-AB22-0239B8350226}	Eindkap	Vuilwater	0	0	0	85038.8253	446446.2061	0
47	{F5F526C3-3A96-40F6-BB17-0243D6F94429}	Inspectieput	Hemel_en_grondwater	1000	1000	-0.76599997	85333.6501	446060.1207	0

Figure 10: No reliable Z (height) information

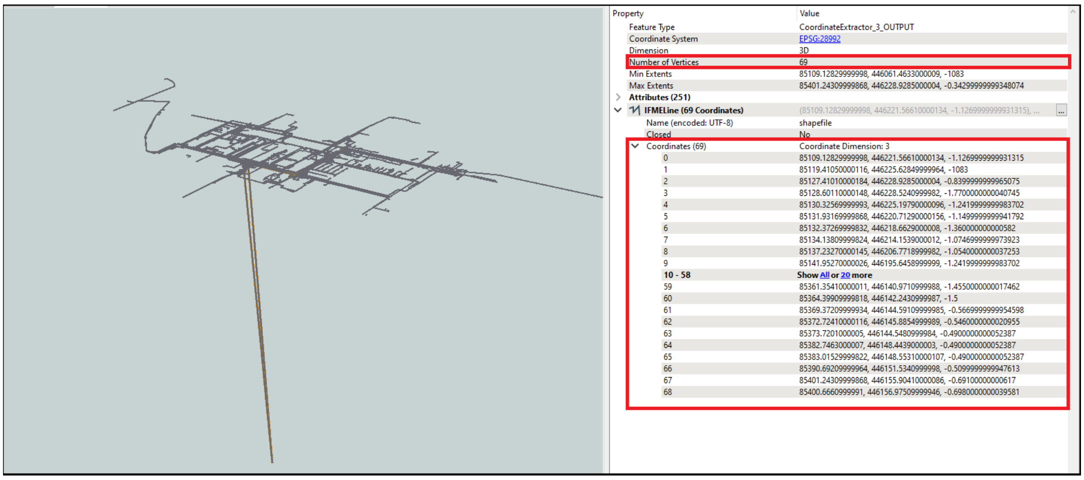


Figure 11: Incorrect topology

In Figure 11 it can be observed that the geometry of the network as it is, is not correct. Specifically, the point that stands out (lower level) and corresponds to the incorrectly written coordinate Z = 32 (m) can not be located at this level, taking into account the overall form of the network. At this point, this incorrect information that existed on the data should be either skipped or fixed based on the frequency it appears on them. On the image below, another example of 3D information is given, where it can be noticed that a pipe is located indeed below pipes of the same use. However, since the majority of the rest of the pipes are located at the same level, this sole part may be 'left-over' or is placed at the correct underground location. In any case, while missing the knowledge about the reference surface and the exact length of the pipes (that could be a means of distinction between actual pipes and left-overs), only assumptions can be made.

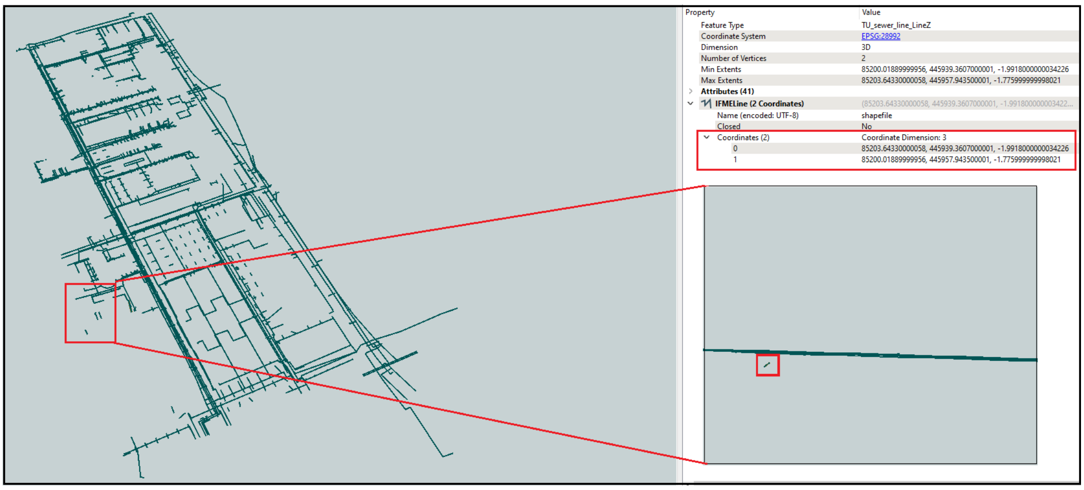


Figure 12

In more detail regarding the data set concerning the sewage line and their attributes, apart from the absense of full three-dimensional information, there is also vague information regarding their size. Comparing the columns that concern that attribute (length and shape columns: of its attribute table) are presented, in most cases, negligible differences, but there are also cases where the size is unknown. This leads to the selection of one of the two pieces of information, the one that is more complete and will help to correct the topology of the data. Additionally, as far as the connectivity between sewer lines and points is concerned, it is noted that although the *TU Sewer line Z* and *TU Sewer knot point Z* datasets are presented to be connected (see Figure 18) the corresponding attributes are not internally associated (in terms of primary and/or foreign keys), even though they share some common attributes (starting-ending points). This implies the impossibility of (the direct) possible use of a database for their processing (DBMS) by using the interdependence relationships, based on primary and/or foreign keys. Instead, spatial queries and operations will be implemented to modify the data in order to adopt a form that will allow their integration.

Continuing the research regarding the height information, it should be mentioned that apart from the Z-coordinate corresponding to the height information, on the datasets some semantic height information was apparent. This information relates to a reference surface (which, however, is not known), as well as to points concerning starting and ending points (Table 3).

Table 3: Semantic Height information

Dataset/ height informaion	**Total number of records**	**Relative level**	**Height (Z-coordinate)**	**Height bottom**	**Height top**	**Height begin**	**Height end**
Unknown information							
TU Sewer knot point Z	1986	553	1953	1780	803	-	-
TU Sewer line Z	2080	-	2048	-	-	1431	1499
Known information							
TU Sewer knot point Z	1986	1433	34	206	1183	-	-
TU Sewer line Z	2080	2080	32	-	-	649	581

Table 3 presents the amount of available information about the height. However, it cannot be used without concerns since a description of the reference surface or reference point is absent. Moreover, in the same data set, this information can not be combined to reach a conclusion about the exact third dimension as at first glance they do not show any correlation or it does not make sense. An example is given on Table 4, where different pipes (of the same use) share the same relative height and Z-coordinate, which is zero, and thus the distinction of their location is not possible. Taking into account the topology as well as the use of the pipes those data cannot be considered correct. To combine the information for our purpose assumptions need to be made in order to follow a fixed reference surface and based on it to make any calculations.

Networks can be characterised by a set of lines (polylines) continuously connected to each other, which has a starting and ending point (nodes) (e.g. a building and/or a physical receiver such as channels, sea). This continuity should be accompanied by the correct topology. In more detail, each depicted line of a network should be accompanied by the appropriate topological information, which should correspond to reality and to make sense. For example, unconnected lines (”left overs”) that can be shown in Figure 13, without a description of their existence and use can be consider as inconsistencies on the datasets. Other examples are given in Figures 14 and 15.

Figure 13: Example of unconnected lines

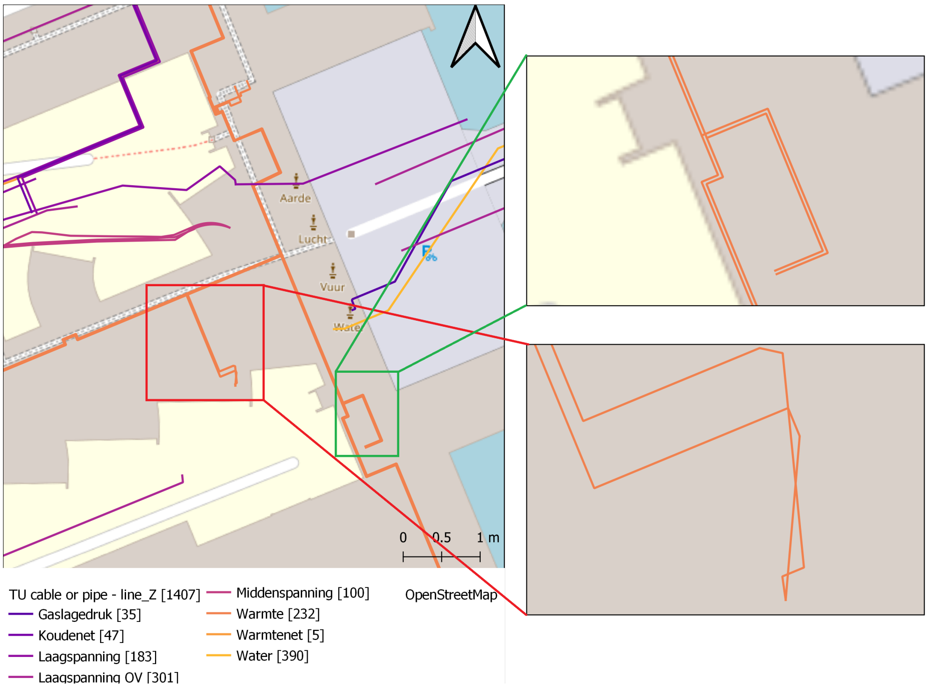


Figure 14: Example of existed topological condition

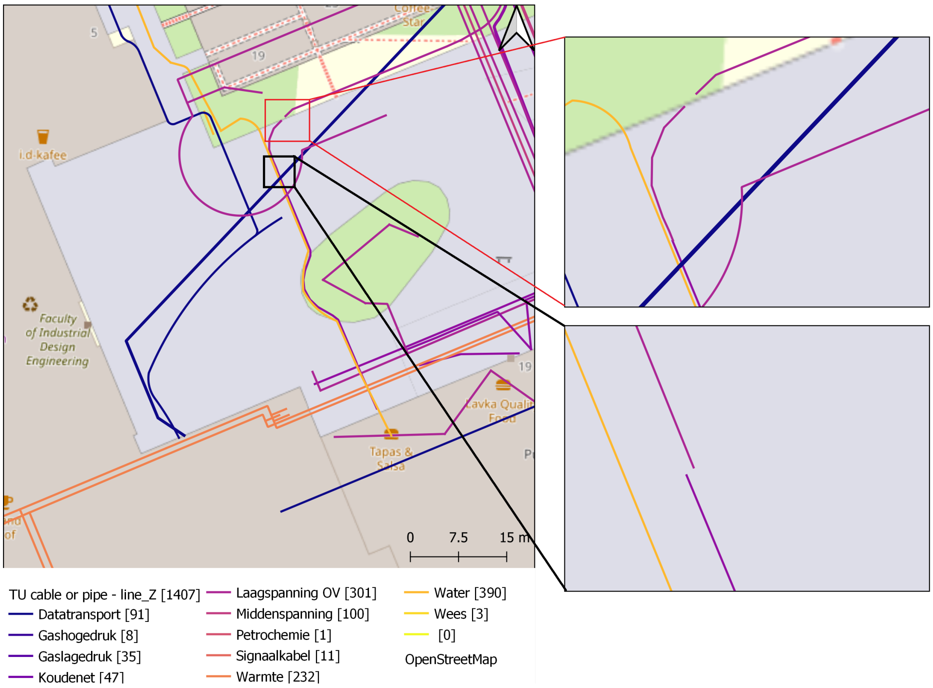


Figure 15: Example of discontinuity

Individual illustrated objects may be due to incorrect topological information, old data sets, incomplete data or pipes of other uses (for example, pipes that collect rainwater from the roofs of buildings and deposit it in the ground) that should not be included as part of the underground network, since their use does not connected directly with the use of a utility network. Similar examples to those mentioned above exist in the case of the points, which represent inspection pits, street, sidewalk gully and other, as well as in the case of the different sewer types (Figure 16 and 17).

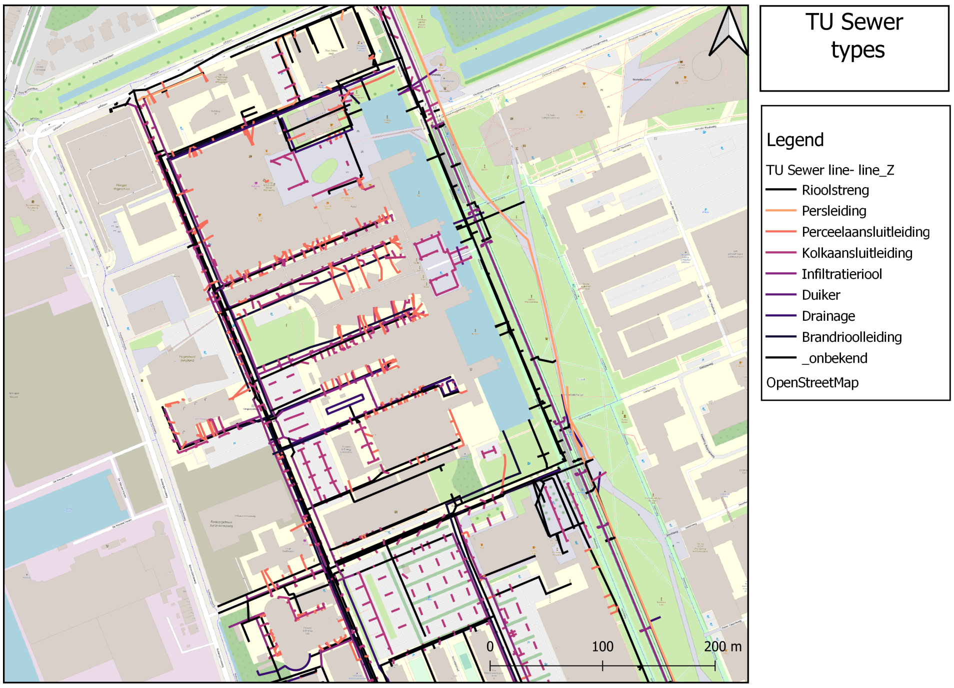


Figure 16: TU Sewer types

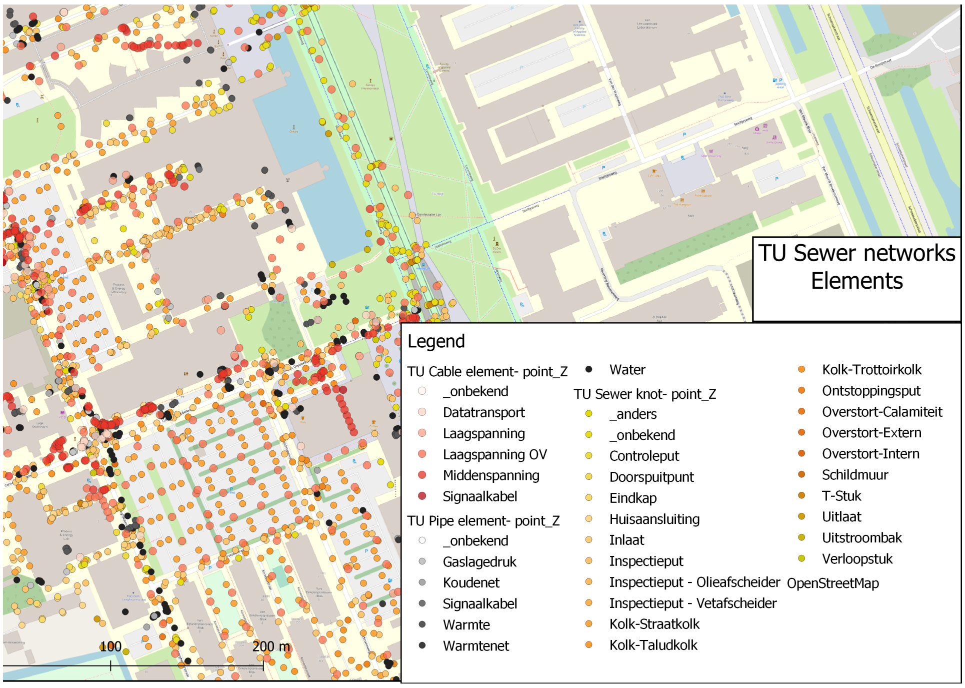


Figure 17: TU Sewer network elements

An example of no direct connection between the lines and the corresponding point is shown in Figure 18. This results from either incorrect information of the data or it is indeed the reality and the point is not yet used or is connected with the network. In Figure 18 are given two examples one with a non-connected point (top right) and one with correctly connected points (bottom right).

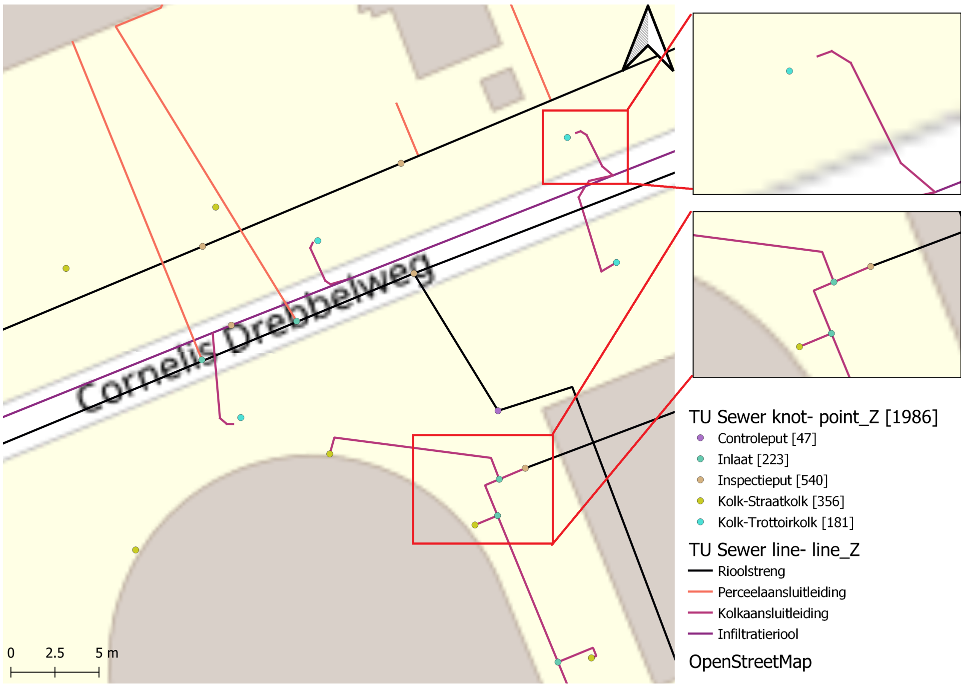


Figure 18: Example of not connected points with lines

By comparing the connection of the edges with the corresponding nodes (in terms of their use), as well as their height (depth) information, other inconsistencies were raised. In more details, it was observed that apart from zero (0) Z value, the relative height (relative to an unknown surface), is the same for all, which means that they all are connected and it is wrongly illustrated or passing through each other (Figure 19, Table 4).

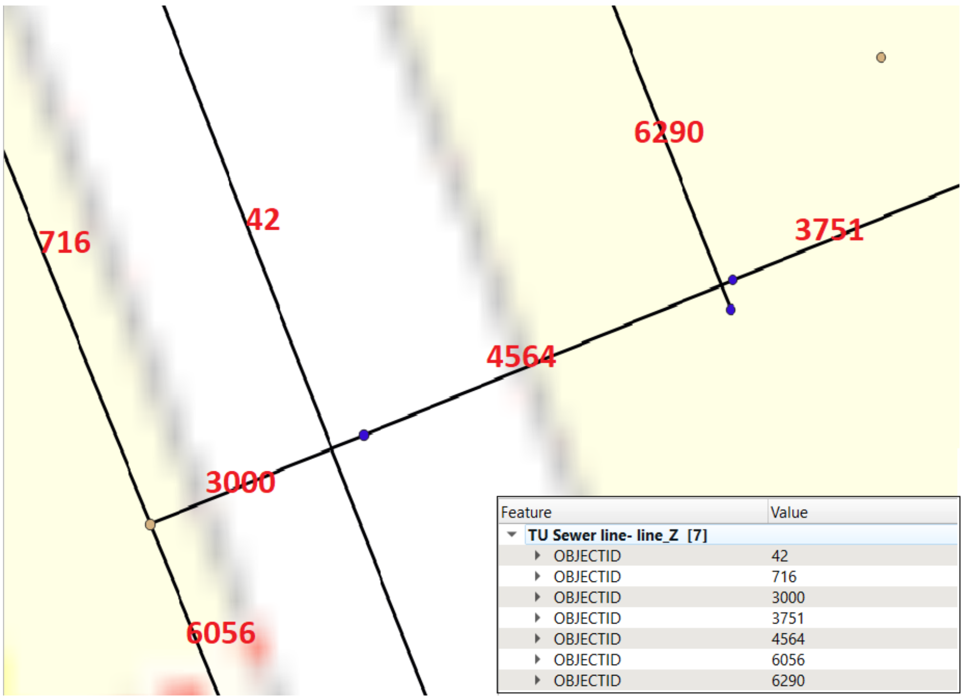


Figure 19: Example of overlay

Table 4: Height information -Rioolstreng (sewer line)

OBJECTID	Relatief_N	Hoogte_Beg	Hoogte_Ein	Shape_Leng	_indices{0}.x	_indices{0}.y	_indices{0}.z
42	-1	-2.3900001	-2.42000008	23.62438378	85098.3291	446315.8661	0
716	-1	-3.92000008	-3.92000008	20.46712762	85096.4621	446315.2421	0
3000	-1	-3.86999989	-3.86999989	1.963902747	85106.3581	446297.1351	0
3751	-1	-3.69000006	-3.83999991	6.019727982	85115.1441	446300.5321	0
4564	-1	-3.83999991	-3.86999989	3.400113234	85109.5291	446298.3621	0
6056	-1	-3.92000008	-4.05999994	68.47499436	85104.5251	446296.4301	0
6290	-1	-2.18000007	-2.21000004	17.97241122	85102.5689	446314.7067	0

From the above table (4) it can be observed, as mentioned above the common height information. Although it make sense to have a stable reference surface, when it is combined with a constant zero value for depth for an underground network, confusion is created. However, from the table it can noticed that the starting and ending point for each line (Hoogte_Beg and Hoogte_Ein accordingly) has the same value (or with fractional differences).

6.1.2 Preliminary suggestions

Considering the data and taking into account the existing inconsistencies in terms of topology, geometry and their attributes, processes need to be done in order to correct them and bring them to a more usable form.

Starting with the topology of the data in order for a clearer view of their reliability to be obtained, a Digital Terrain Model (DTM) of the study area can be used. It would be beneficial to match the geometry of the vector information with the one provided in the raster and compare if they fit together. This DTM can be a raster (GeoTIFF) or can be derived from a point cloud from which a triangulated irregular network (TIN) will be created (surface model). The 3D surface map in combination with the 2D features results in draped features, where the 3D information can be extracted. However, taking into account that the vector features correspond to the underground condition, the Z-value should be extracted using a combination of techniques and assumptions. For example, a reference surface can be considered to be at a constant distance from the ground and to be used as a point of dependence for the depths. This surface could be derived by calculating the average of the available Z information of the datasets. Apart from the DTM, the 3D information can be derived by combining the existed information with the one comes from the network service providers (NSP). As depending on the legislation of each country, it is necessary for the NSP to follow some specifications for the construction of the corresponding network. These specifications among others include regulations about the accepted depth that the pipes of different uses should have below ground. Another useful tool could be the KLIC application (Chapter 2) available from the country's Cadaster, where information about the location of the third parties networks (including the sewage- study case) can be obtained.

Since both the exact 3D coordinates as well as the corresponding semantic height information is absent, incomplete, or unknown, another source of data can be useful (e.g. underground plans provided by the Municipality of Delft) and by combining different information we can come to the identification of similarities between the data and enrich the existing ones with extra (relevant to the third-dimension) information. However, in addition to the lack of information about heights (semantic height information and Z-coordinate), which are a prerequisite for creating a 3D map, there were, as mentioned previously, other inconsistencies, concerning the attributes of the data as well as their geometry. At this point, spatial operations need to be implemented in order for the final network to correspond to reality and to have a logical connection of the displayed features. In this case, for example, spatial operations can be considered the connection of underground pipes to the points of the ground (underground and/or above-ground) with which they are connected due to usage, as well as the possible union of pipes with the same use, using a search cycle with a constant tolerance. This will result in a simplified dataset suitable for a test case. The attention was paid more to the implementation of spatial operation since while examining the attributes of the data there were no keys (primary-foreign) that allow for internal connections (interdependence relationships) and the datasets should be handled separately.

Having completed the 3D mapping of the underground utility networks the next step is their connection with the above-ground objects. To do so, a 3D city model of the study area is about to be used with the goal to connect the network with the above-ground connection point (e.g. Building). It is noted that from the external environment not all objects will be used but those related to the networks. The connection between the objects (underground-above-ground) will be modelled by using their location (coordinates).

6.2 Tools

A combination of software needs to be used to display, edit/process, develop a model as well as to save/store it. Specifically, QGIS software, as well as Feature Manipulation Engine (FME), will be used for the processing, enriching, and display of the data. In addition, access for the ArcGIS Pro software has been given and it is likely to be used for possible utilization of the tool it offers (ArcGIS Network Analyst extension), relative to utility network management. For the same purpose, it is not excluded to use the library of Python NetworkX (for more information [21]). Moreover, 3D City Database or a simpler structured database is possible to be used in order to store, represent, and manage the virtual 3D model will be created, on top of a standard spatial relational database (available in `https://www.3dcitydb.org/3dcitydb/downloads/`, [22]). However, since we are working with geometries, a relational database system is possible to be used for handling the different layers available. With the help of PostGIS extension and its facilities it is possible to handle apart from 2-dimensional geometries, also 3D.

References

[1] I. Hijazi, "Use cases and their requirements on the semantic modeling of 3d supply and disposal networks," 02 2017.

[2] P. G. Tabarro, J. Pouliot, R. Fortier, and L.-M. Losier, "A webgis to support gpr 3d data acquisition: A first step for the integration of underground utility networks in 3d city models," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W7, pp. 43–48, 2017.

[3] E. Vishnu and S. Saran, "Semantic modeling of utility networks implementation of use cases for dehradun city," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-5, pp. 139–145, 11 2018.

[4] I. Boates, G. Agugiaro, and A. Nichersu, "Network modelling and semantic 3d city models: Testing the maturity of the utility network ade for citygml with a water network test case," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-4, pp. 13–20, 2018.

[5] D. I. H. Prof. Thomas H. Kolbe, Dr. Tatjana Kutzner, "Citygml utility network ade." https://www.asg.ed.tum.de/en/gis/projects/citygml-utility-network-ade/.

[6] Kutzner, Tatjana and Ihab Hijazi, and Thomas H. Kolbe, "Semantic modelling of 3d multi-utility networks for urban analyses and simulations: The citygml utility network ade.," *International Journal of 3-D Information Modeling (IJ3DIM)*, vol. 7, pp. 1–34, 6 2018.

[7] J. Yan, S. W. Jaw, R. van Son, K. Soon, and G. Schrotter, "Three-dimensional data modelling for underground utility network mapping," vol. XLII-4, pp. 711–715, 09 2018.

[8] G. Jiang, X. Zhou, J. Li, and H. Chen, "A cable-mapping algorithm based on ground-penetrating radar," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 10, pp. 1630–1634, 2019.

[9] I. T. W. G. Utility and governmental services, "D2.8.iii.6 inspire data specification on utility and governmental services – draft guidelines."

[10] T. Becker, C. Nagel, and T. H. Kolbe, "Integrated 3d modeling of multi-utility networks and their interdependencies for critical infrastructure analysis," 2011.

[11] T. Adolphi, C. Nagel, and T. Kolbe, *Semantic 3D Modeling of Multi-Utility Networks in Cities for Analysis and 3D Visualization*, pp. 41–62. 10 2013.

[12] O. G. Consortium, "Citygml." https://www.ogc.org/standards/citygml, 2021.

[13] P. J. M. van Oosterom, *Advances in 3D geoinformation systems*, pp. 96–98. 2008.

[14] ESRI, "What are geometric networks?." https://desktop.arcgis.com/en/arcmap/10.3/manage-data/geometric-networks/what-are-geometric-networks-.htm, 2017a.

[15] J. Lieberman, "Model for underground data definition and integration (muddi) engineering report." http://www.opengis.net/doc/PER/MUDDI-ER, 2019.

[16] GEONOVUM, "Information model cables and pipes." https://www.geonovum.nl/geo-standaarden/informatiemodel-kabels-en-leidingen, 2020.

[17] GEONOVUM, "Nen 3610 basic model for information models." https://www.geonovum.nl/geo-standaarden/nen-3610-basismodel-voor-informatiemodellen, 2020.

[18] L. R. Herman van den Berg and G. Paul Janssen, "Information model cables and pipes (imkl) - version 2.0." https://geonovum.github.io/imkl2015-review/IMKL2.x/1-dataspecificatie/#uml-diagrammen, 2021.

[19] L. A. and L. D., "Qualitative research analysis." The NIHR RDS for the East Midlands/ Yorkshire & the Humber, 2007.

[20] VoXco, "What is quantitative research?." https://www.voxco.com/blog/quantitative-research/, 2020.

[21] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using networkx," in *Proceedings of the 7th Python in Science Conference* (G. Varoquaux, T. Vaught, and J. Millman, eds.), (Pasadena, CA USA), pp. 11 – 15, 2008.

[22] T. U. München, "3dcitydb database." https://www.3dcitydb.org/3dcitydb/downloads/, 2021.