Doctoral theses at NTNU, 2021:135

Knut Jetlund

Harmonizing and linking conceptual models of geospatial information

Technologies for information modelling in GIS, ITS and BIM

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Engineering
Department of Manufacturing and Civil
Engineering

NTNU
Norwegian University of
Science and Technology

Knut Jetlund

Harmonizing and linking conceptual models of geospatial information

Technologies for information modelling in GIS, ITS and BIM

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2021

Norwegian University of Science and Technology
Faculty of Engineering
Department of Manufacturing and Civil Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Manufacturing and Civil Engineering

ISBN 978-82-326-6601-0 (printed ver.)
ISBN 978-82-326-5703-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:135

Printed by NTNU Grafisk senter

"Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?"

– T. S. Elliot: The Rock, 1934

ACKNOWLEDGEMENTS

"I'm going on an adventure!"
– Bilbo Baggins in The Hobbit, by J.R.R. Tolkien.

This journey into the world of academics has been loaded with hard work, discoveries, knowledge extension and lots of experiences. I want to express gratitude to the many who have motivated, supported and helped me find the way there and back again.

First, the research has been jointly funded by the Norwegian Research Council and the Norwegian Public Roads Administration (NPRA).

Foremost among the personal contributors, I would like to thank my supervisors at the Norwegian University of Science and Technology (NTNU), Associate professor Erling Onstein and Associate professor Lizhen Huang, for their excellent guidance through unknown academic territories. Besides, Professor Rune Hjelsvold deserves much gratitude for guiding me into the Semantic Web technologies. You have all been invaluable for building the knowledge needed for research and academic writing.

Special thanks to my leaders in the NPRA for giving me this opportunity to extend my horizons and contribute to an improved digitalization and information flow for topics that are vital for our agency. John Mikalsen opened the gates by suggesting that I could start this journey, while Grete Storløpa and Per Morten Lund gave the permissions needed. Besides, Dr Trond Arve Haakonsen has been my internal mentor and valuable support, especially in navigating the initial formal and informal processes.

I would also like to thank my colleagues at the NPRA and NTNU and my colleagues within International standardization activities, for professional and informal discussions.

My most tremendous gratitude goes to my wife Ragnhild and our children Ingrid, Mari and Eivind for their understanding and support and keeping me in touch with the real world throughout this journey. My wife Ragnhild deserves special gratitude for her patience and company at the home office during these months of Covid-19 restrictions.

Knut Jetlund

Ridabu, January 2021

ABSTRACT

"He who believes that higher education alone gives wisdom and prudence, he has not even seen the beginning of the path of wisdom."
– Thor Heyerdahl.

Structured representations of phenomena from the real world in a digital geospatial environment are essential for developing, maintaining, and using the built and natural environment. In the real world, phenomena relate to, influence and are influenced by other phenomena through their location, shape and extent. These geospatial characteristics and relations must also be described in the digital environment.

Three of the key application domains for handling the real world in a digital geospatial environment are Geographic Information Systems (GIS), Intelligent Transport Systems (ITS) and Building Information Modelling (BIM). The three application domains have distinct but related roles: GIS applications are primarily used for analysis and presentation of the natural and built environment; BIM applications are used for planning, developing, constructing and maintaining the built environment; while applications and systems for ITS use the geospatial information for planning and controlling transportation. Despite the distinct roles, many of the same real-world phenomena are described and used in all three application domains. Therefore, exchange and reuse of information between application domains, life cycle stages and stakeholders should be possible. Changes to the digital environment first arise in the BIM domain but based on the existing situation described in GIS. The changes to the built and natural environment need to be updated in GIS and ITS after real-world construction. Besides, authoritative information for legal and safe navigation needs to be shared from authorities' GIS databases to ITS.

Stakeholders from each of the three application domains have developed conceptual models of geospatial information where phenomena from the real world are described in application-specific contexts. Less concern has been put on common modelling approaches and information use across application domain. Consequently, there is a lack of a shared understanding and interpretation of how the real world is represented in information models, and fundamental differences between the information models are obstacles for information exchange across domain borders.

The research presented in this thesis has strived towards improved interoperability between GIS, ITS and BIM through a joint approach for representations of the real world in conceptual models of geospatial information. The exploratory research was conducted through a state-of-the-art study on information modelling in the three application domains, followed by experiments with prototypes that could improve the interoperability. The thesis is built around six appended articles that describe individual parts of the research.

The state-of-the-art analysis found that different approaches and technologies have been used for information modelling in the three application domains, but that there are similarities and possibilities for harmonization and linking. Improved syntactic interoperability would be possible if information models for all three application domains were described in the same conceptual language and with a harmonized approach for information modelling. The Unified Modelling

Language (UML) and Model-Driven Architecture (MDA) have been used in all three application domains and were identified as candidate technologies for a common approach. The use of UML and MDA for GIS as standardized by ISO/TC 211 was considered the most structured and complete approach, despite several identified deficiencies. Therefore, it was suggested that a harmonized approach for information modelling in GIS, ITS, and BIM could be based on an improved version of the UML and MDA approach defined by ISO/TC 211. A prototype of a structure of UML profiles and two prototypes of information models showed how the approach could be implemented:

- A prototype structure of UML profiles was founded on a core geospatial UML profile with improvements of the UML profiles defined in ISO 19103 and ISO 19109. Community conceptual UML profiles were suggested for use in platform-independent information models for BIM and ITS, while encoding profiles for general and community encodings were suggested for use in platform-specific models.
- A prototype of a generic information model for the information exchange from GIS to ITS was developed according to the suggested approach, based on existing information models from the GIS and ITS application domains. Test implementations of the prototype showed that the generic model could be used to exchange information according to different feature catalogues, such as the ITS standard GDF. Minor modifications of the GDF model were required to comply with the harmonized information modelling approach.
- A prototype of the BIM standard IFC following the harmonized approach showed that transformations from the existing EXPRESS model to UML according to ISO/TC 211 standards were possible, and that core concepts from IFC and models from ISO/TC 211 standards could be linked. Some additional semantics were needed in the UML model for implementation in EXPRESS.

The research showed that improved syntactic interoperability could be achieved by describing information models from all three application domains according to the harmonized approach. The semantic interoperability could be improved by using the same core concepts in distinct information models. Concepts for primitive and fundamental datatypes should be reused from external and generic models, while more specific concepts for geometry and location referencing should be defined in one harmonized model for use in all three application domains.

While an improved syntactic and semantic interoperability could be achieved with the harmonized modelling approach, a full harmonization of information models would not be appropriate. Given the distinct roles of the three application domains in the digital geospatial environment, the information models need to describe the real world in different contexts. Therefore, it was suggested that improved semantic interoperability could be achieved by applying Semantic Web technologies to link and map concepts. The research showed that conversions from UML models to OWL ontologies used for linking and mapping were possible. The transformed ontologies were not optimized for use in the Semantic Web though, due to fundamental differences between UML and OWL modelling approaches. Additional semantics were described in a UML profile for encoding in OWL in the prototype structure and could improve the conversions, but fully optimized ontologies would require manual editing.

The research presented in this thesis has shown that the combination of harmonization and linking could improve the syntactic and semantic interoperability between information models for GIS, ITS and BIM. A harmonized modelling approach would be the fundament for achieving the improved interoperability.

SAMMENDRAG

"Den som tror at høy utdannelse alene gir visdom og klokskap,
han har ikke engang sett begynnelsen på visdommens vei."
– Thor Heyerdahl.

Strukturerte representasjoner av fenomener fra den virkelige verden i et digitalt geografisk miljø er sentralt for utvikling, vedlikehold og bruk av det bygde så vel som det naturlige miljøet. Fenomener i den virkelige verden forholder seg til, påvirker og påvirkes av andre fenomener gjennom beliggenhet, form og omfang. De geografiske egenskapene og relasjonene må også beskrives i det digitale miljøet.

Geografiske informasjonssystemer (GIS), Intelligente transportsystemer (ITS) og Bygningsinformasjonsmodellering (BIM) er tre av de viktigste applikasjonsdomenene for håndtering av den virkelige verden i et digitalt geografisk miljø. De tre applikasjonsdomenene har forskjellige, men beslektede roller: Applikasjoner for GIS brukes primært til analyse og presentasjon av det naturlige og bygde miljøet. BIM-applikasjoner brukes til å planlegge, utvikle, konstruere og vedlikeholde det bygde miljøet, mens applikasjoner og systemer for ITS bruker geografisk informasjon for planlegging og kontroll av transport. Til tross for de forskjellige rollene er mange av de samme fenomenene fra virkeligheten representert og brukt i alle de tre applikasjonsdomenene. Det burde derfor være mulig å utveksle og gjenbruke informasjon mellom applikasjonsdomener, livssyklusstadier og interessenter. Endringer i det digitale miljøet oppstår først i BIM-domenet, men basert på den eksisterende situasjonen beskrevet i GIS. Endringene i det bygde og naturlige miljøet må oppdateres i GIS og ITS etter konstruksjon i virkeligheten. Dessuten må autoritativ informasjon for lovlig og sikker navigering deles fra myndigheter sine GIS-databaser til ITS.

Interessenter fra hvert av de tre applikasjonsdomenene har utviklet konseptuelle modeller for geografisk informasjon der fenomener fra den virkelige verden er beskrevet i applikasjonsspesifikke kontekster. Det har blitt lagt mindre vekt på felles modelleringsmetoder og bruk av informasjon på tvers av applikasjonsdomener. Følgelig mangler det en felles forståelse og tolkning av hvordan den virkelige verden er representert i informasjonsmodeller, og grunnleggende forskjeller mellom informasjonsmodellene begrenser mulighetene for informasjonsutveksling på tvers av domenegrenser.

Forskningen som presenteres i denne avhandlingen har tilstrebet en forbedret interoperabilitet mellom GIS, ITS og BIM gjennom en felles tilnærming for representasjoner av den virkelige verden i konseptuelle modeller av geografisk informasjon. Den undersøkende forskningen ble utført ved en analyse av status på informasjonsmodellering i de tre applikasjonsdomenene, etterfulgt av eksperimenter med prototyper som kunne forbedre interoperabiliteten. Avhandlingen er bygget rundt seks artikler som beskriver individuelle deler av forskningen.

Statusanalysen viste at ulike tilnærminger og teknologier har blitt brukt for informasjonsmodellering i de tre applikasjonsdomenene, men også at det er likhetstrekk og muligheter for harmonisering og kobling. En forbedret syntaktisk interoperabilitet ville være mulig hvis informasjonsmodeller for alle de tre applikasjonsdomenene ble beskrevet med samme konseptuelle språk og med en harmonisert tilnærming til informasjonsmodellering. Teknologiene

Unified Modeling Language (UML) og Model-Driven Architecture (MDA) har blitt brukt i alle de tre applikasjonsdomenene og ble identifisert som kandidatteknologier for en felles tilnærming. Bruk av UML og MDA for GIS slik det er standardisert av ISO/TC 211 ble ansett som den mest strukturerte og komplette tilnærmingen, til tross for flere identifiserte mangler. Derfor ble det foreslått at en harmonisert tilnærming til informasjonsmodellering i GIS, ITS og BIM kunne baseres på en forbedret versjon av UML- og MDA-tilnærmingen som er definert av ISO/TC 211. En prototype for en struktur av UML-profiler og to prototyper av informasjonsmodeller viste hvordan tilnærmingen kan implementeres:

- En prototype for en struktur av UML-profiler ble foreslått å være basert på en kjerne-UML-profil for geografisk informasjon, med forbedringer av UML-profilene definert i ISO 19103 og ISO 19109. Domenespesifikke konseptuelle UML-profiler ble foreslått for bruk i plattformuavhengige informasjonsmodeller for BIM og ITS, mens profiler for generelle og felles implementasjonsteknologier ble foreslått for bruk i plattformspesifikke modeller.
- En prototype av en generell informasjonsmodell for utveksling av informasjon fra GIS til ITS ble utviklet i henhold til den foreslåtte tilnærmingen, basert på eksisterende informasjonsmodeller fra GIS og ITS. Testimplementasjoner av prototypen viste at den generiske modellen kunne brukes til å utveksle informasjon i henhold til forskjellige objektkataloger, for eksempel ITS-standarden GDF. Noen mindre endringer av GDF-modellen var nødvendige for at den skulle være i samsvar med den harmoniserte tilnærmingen for informasjonsmodellering.
- En prototype for å beskrive BIM-standarden IFC i henhold til den harmoniserte tilnærmingen viste at det var mulig å transformere fra den eksisterende EXPRESS-modellen til UML i henhold til ISO/TC 211-standarder, og at grunnleggende konsepter fra IFC og modeller fra ISO/TC 211-standarder kunne kobles. Noe ekstra semantikk var nødvendig i UML-modellen for implementasjon i EXPRESS.

Forskningen viste at forbedret syntaktisk interoperabilitet kunne oppnås ved å modellere informasjonsmodeller fra alle de tre applikasjonsdomenene i henhold til en harmonisert tilnærming. Dessuten kunne den semantiske interoperabiliteten forbedres ved å bruke de samme grunnleggende konseptene i forskjellige informasjonsmodeller. Konsepter for primitive og grunnleggende datatyper burde hentes fra eksterne og generiske modeller, mens mer spesifikke konsepter for geometri og stedfesting burde defineres i en harmonisert modell for bruk i alle de tre applikasjonsdomenene.

Forbedret syntaktisk og semantisk interoperabilitet kunne oppnås med den harmoniserte tilnærmingen, men en full harmonisering av informasjonsmodeller ville ikke være hensiktsmessig. Informasjonsmodellene har behov for å beskrive den virkelige verden i forskjellige perspektiver, gitt de forskjellige rollene til de tre applikasjonsdomenene i det digitale geografiske miljøet. Derfor ble det foreslått at en forbedret semantisk interoperabilitet kunne oppnås ved å anvende teknologier fra den semantiske webben for kobling og transformasjon av konsepter. Forskningen viste at konvertering fra UML-modeller til OWL-ontologier som kunne brukes til lenking og transformasjon var mulig. De konverterte ontologiene var imidlertid ikke optimalisert for bruk i den semantiske webben, på grunn av grunnleggende forskjeller mellom modelleringstilnærmingene til UML og OWL. Ytterligere semantikk beskrevet i henhold til en UML-profil for implementering i OWL kunne forbedre konverteringene, men fullt optimaliserte ontologier ville kreve manuell redigering.

Forskningen presentert i denne oppgaven har vist at kombinasjonen av harmonisering og kobling kan forbedre den syntaktiske og semantiske interoperabiliteten mellom informasjonsmodeller for GIS, ITS og BIM. En harmonisert tilnærming for informasjonsmodellering vil være grunnlaget for å oppnå forbedret interoperabilitet.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AAA	Anyone can say anything about any topic
ADAS	Advanced Driver Assistance Systems
AIM	Asset Information Model
BIM	Building Information Modelling
bSI	buildingSMART International
CEN	European Committee for Standardization (Comité européen de normalisation)
CEN/TC 278	CEN Technical Committee 278 – Intelligent transport systems
CGP	Core Geospatial Profile
CP	Conference proceedings
CWA	Closed World Assumption
DATEX	Data exchange between traffic and travel information centres
DIKW	Data, information, knowledge and wisdom
GDF	Geographic Data Files
GIS	Geographic (Geospatial) Information Systems
GML	Geography Markup Language
HD	High Definition
IBM	International Business Machines
IEC	International Electrotechnical Commission
IFC	Industry Foundation Classes
INSPIRE	Infrastructure for spatial information in the European community
ISO	International Organization for Standardization
ISO/TC 184 SC 4	ISO Technical Committee 184 – Automation systems and integration, Sub Committee 4 – Industrial data
ISO/TC 204	ISO Technical Committee 204 – Intelligent transport systems
ISO/TC 211	ISO Technical Committee 211 – Geographic information/Geomatics
ISO/TC 59 SC 13	ISO Technical Committee 59, Sub Committee 13 – Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM)
ITS	Intelligent Transportation Systems
JSON	JavaScript Object Notation
LCIM	Level of Conceptual Interoperability Model
MDA	Model-Driven Architecture
MOF	Meta Object Facility
MRS	Media Record Structure
NPRA	Norwegian Public Roads Administration
NDS	The Navigation Data Standard

NTNU	Norwegian University of Science and Technology
NVDB	The Norwegian National Road Database
OCL	Object Constraint Language
OGC	Open Geospatial Consortium
OMG	Object Management Group
OWA	Open World Assumption
OWL	Web Ontology Language
PIM	Platform-Independent Model (in MDA)
BIM-PIM	Project Information Model (PIM in ISO 19650)
PM	Platform Model
PSM	Platform-Specific Model
RAS	Rules for application schemas
RDF	Resource Description Framework
RM-ODP	Reference Model of Open Distributed Processing
SDI	Spatial Data Infrastructure
SP	Standardization project
SPARQL	Simple Protocol and RDF (Resource Description Framework) Query Language
SQL	Structured Query Language
SRQ	Sub research question
STEP	STandard for the Exchange of Product data
TN-ITS	Transport Network for ITS
TPEG	Transport Protocol Experts Group
UML	Unified Modelling Language
W3C	World Wide Web Consortium
WFS	Web Feature Service
XMI	XML Metadata Interchange Format
XML	eXtensible Markup Language

Table of Contents

1 INTRODUCTION

"We can only see a short distance ahead, but we can see plenty there that needs to be done."
– Alan Turing.

1.1 BACKGROUND

1.1.1 Digital twins in the digital geospatial environment

The digital transformation known as Industry 4.0 is changing processes for developing, maintaining and using the natural and built environment (Boje et al., 2020, Ustundag and Cevikcan, 2017). Structured digital information has become vital for processes through the whole lifecycle of built constructions such as infrastructures for road transportation, from planning and design to construction, operation, maintenance, analysis, use and demolition.

Physical and intangible phenomena from the real world and their characteristics are represented in digital twins, where information can be analyzed and combined, changes can be foreseen, and future development can be planned (Beetz et al., 2020, Hetherington and West, 2020, Boje et al., 2020, Fjeld, 2020, Evans et al., 2019). A single digital twin represents a limited part of the real world (the physical twin) and is prepared for serving a purpose. Contrary, an ecosystem where a multitude of digital twins are brought together in an integrated digital environment can cover larger portions of the real world, give a broader view and serve multiple purposes (Gilbert et al., 2020, Beetz et al., 2020, Hetherington and West, 2020).

Constructions in the built environment have a location, shape and extent in the real world, through which they have spatial relations to, influence, and are influenced by other phenomena from the natural and built environment. Therefore, the geospatial context is vital for the digital representation of the natural and built environment in the integrated digital environment, too (Beetz et al., 2020, Gilbert et al., 2020). In a digital geospatial environment, the digital twins representing the natural and built environment are described in a geospatial context, where the representation of real-world phenomena includes their location, shape and extent.

During the lifecycle of road infrastructures, processes involve several application domains where the geospatial context is essential. Among these are the application domains of Geographic Information Systems (GIS), Intelligent Transport Systems (ITS) and Building Information Modelling (BIM). The applications utilize information from the digital geospatial environment concerning road networks and associated restrictions, road equipment, events, and the surrounding natural and built environment. The existing digital geospatial environment describes the real world and forms the basis for planning and design for future development. The designed new digital geospatial environment is used for managing machines and personnel to build the physical road infrastructure (Statens vegvesen, 2015). In operation and maintenance of roads, the updated digital environment is the foundation for operator contracts, budgets, statistics, environmental and other analysis, and planning and documentation of the physical operations (Statens vegvesen, 2014). For the road users, information from the digital geospatial environment is crucial for route planning, fleet management, notification of events, driver assistance and automated driving (The European Commission C-ITS Deployment Platform, 2017, 2016).

1.1.2 GIS, ITS and BIM

Geographic (Geospatial) Information Systems (GIS) and road databases

The term "Geographic Information Systems" (GIS) was introduced in the 1960s for systems for capturing, storing, analyzing and visualizing information that describes a part of the Earth (Kresse and Danko, 2011). ISO 19101-1:2002 defines GIS as an "*information system dealing with information concerning phenomena associated with location relative to the Earth*" (ISO/TC 211, 2014). Later, the term "geospatial" has been introduced as an alternative to "geographic", with a broader perspective. The two terms are mostly used as synonyms when referring to GIS and the information handled in a GIS (Kresse and Danko, 2011).

GIS technologies with applications, database systems and geospatial information are used for various disciplines, including databases for road networks and road-related geospatial information. Public road authorities like the NPRA manage information about the road network and road-related features in road databases and applications based on technologies and standards from the GIS application domain. The Norwegian National Road Database (NVDB) (Statens vegvesen, 2020b) is one example of such road databases. NVDB is a centralized database with a navigable road network, information about restrictions and other road network properties, road equipment and events. NVDB and similar road databases in other countries are essential tools for authorities and operators for planning, developing, operating and maintaining roads. Besides, the information is essential for other users; for example, road users and traffic planners who need authoritative information for legal and safe navigation (Borzacchiello et al., 2016, NPRA and SINTEF, 2008). Furthermore, information on traffic volumes, noise data and accidents is essential for municipalities and others for planning building sites, industrial areas, roads and other elements in the built environment. While information about utilities along the road network, when combined with information from other actors, forms a data set of national importance for safety, energy and handling of extreme weather (Kommunal- og moderniseringsdepartementet, 2020).

Intelligent Transport Systems (ITS)

The term "Intelligent Transport Systems" (ITS) is a collective term for a broad range of *"information and communication technologies and services applied to transport and the related infrastructure"* (Appel et al., 2018), whose purpose is to provide information to users of the transport infrastructure and to monitor and control traffic (Statens vegvesen, 2018, 2020a). ITS technologies are implemented in the physical infrastructure, vehicles and mobile units.

ITS technologies have a strong potential for improving road safety and efficiency in road transport and are expected to play an essential role in the future of connected and automated driving. Some examples are Advanced Driver-Assistance Systems (ADAS) for advising and warning vehicles and drivers, such as adaptive cruise control, lane-keeping assistants and warnings about road works, approaching emergency vehicles and challenging weather conditions. Another group of applications is vehicle-automation systems used to moderately control vehicles for secure driving, such as speed adaption systems and platooning, or for fully controlling vehicles in case of an emergency (Statens vegvesen, 2018, The European Commission C-ITS Deployment Platform, 2017, 2016).

Systems for automated driving need to combine a range of applications that rely on geospatial information to advise and control vehicles in safe and legal navigation. The applications need to collect, store, access, understand and use detailed geospatial information about the road network and road-related features such as restrictions and road conditions, coming from a combination of vehicle sensors, sensors in the infrastructure, and pre-processed datasets. (Paul et al., 2017, The European Commission C-ITS Deployment Platform, 2017, 2016, Statens vegvesen, 2018, Zang et al., 2017, Jomrich et al., 2017).

Building Information Models and Modelling (BIM)

The application domain of BIM was introduced in the late 1980s and early 1990s. The abbreviation BIM has been used in parallel for the two terms "Building Information Modeling" and "Building Information Models". ISO 19650-1:2018 defines Building Information Modelling as "*use of a shared digital representation of a built asset to facilitate design, construction and operation processes to a reliable basis for decisions"* (ISO/TC 59/SC 13, 2018b). Building Information Models are digital models of the built asset, referred to as Asset Information Models (AIM) and Project Information Models (BIM-PIM[1]) in ISO 19650-1:2018.

While the initial scope of BIM was limited to information about buildings, the scope has been extended over the last years also to include infrastructures, known as BIM for Infrastructure. With BIM for Infrastructure, the technology and project management has been brought into planning, construction, operation and maintenance of infrastructures such as roads and railways. BIM for road infrastructure includes digital geospatial representations of roads, utilities, road equipment and the surrounding environments (buildingSMART International, 2019b). BIM for infrastructure raises additional concerns for the geospatial context, as infrastructure projects extend over large geographic areas and relate more to other features from the natural and built environment than individual buildings do. The NPRA describes the planning and construction of road infrastructures based on BIM technologies in their Handbook V770 (Statens vegvesen, 2015).

1.1.3 Roles and relations in the digital geospatial environment

The three application domains of GIS, ITS and BIM have distinct but related roles in the digital geospatial environment, as illustrated in Figure 1. GIS applications have a primary purpose of analyzing and presenting the existing natural and built environment, while BIM applications are used to plan, develop, construct, and maintain the built environment. Finally, ITS applications and systems use information from the digital geospatial environment to plan and control transportation.

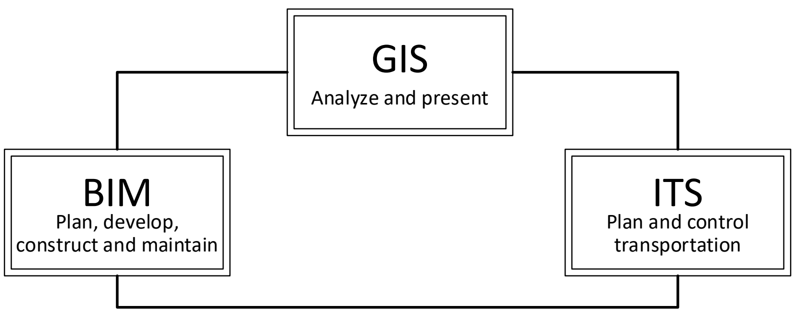


Figure 1: Roles and relations in the digital geospatial environment.

[1] The abbreviation BIM-PIM is used in this thesis for PIM as defined in ISO 19650, in order to avoid confusion with the abbreviation PIM as used MDA.

While the roles of GIS, ITS and BIM are distinct and the real world is modelled in different perspectives, many of the real-world features and concepts they handle are the same. Therefore, reuse of information across application domain borders should be possible, given that there exists a common understanding of how the digital geospatial environment represents the real world.

A change in the physical environment, such as a new road, is initially designed in a BIM-PIM for road infrastructure. The complete digital geospatial environment in the BIM-PIM describes the road project and the modified surrounding environment. The basis for designing the new environment is the existing digital geospatial environment described in GIS datasets, while the updated information in as-built BIM-PIMs is needed for updating GIS datasets after completed development (Statens vegvesen, 2015). On the ITS side, automated driving systems need the updated situation for safe and correct navigation in the modified environment. One specific example is the digital representation of a guardrail along a new road. The guardrail will first be designed in a BIM-PIM in the planning stage for the road and will later be constructed with the support of BIM processes. After the construction, GIS datasets and High Definition (HD) maps for ITS need to be updated with the feature representation. The distinct roles of GIS, ITS and BIM require specific information about the guardrail in each application domain, but the physical feature and many characteristics are the same.

Besides the digital representation of physical features, systems for automated driving in the ITS application domain need a navigable digital network with authoritative information about restrictions for legal and safe navigation. The core authoritative data source for the digital road network and restrictions are the authorities' geospatial road databases (Borzacchiello et al., 2016, The European Commission C-ITS Deployment Platform, 2017, 2016, NPRA and SINTEF, 2008).

1.2 Problem statement

The capability to communicate and exchange information between computer systems is known as interoperability (IEC, 2019, ISO/IEC JTC 1, 2015). Successful communication of information representing the real world requires that all partakers know how to interpret the exchanged data into useful information. They need a shared understanding, gained from the description of the real world in information models (Schenck and Wilson, 1994, Hitzler et al., 2012, Zhao et al., 2011).

Great efforts have been put into development and standardization of information models in the application domains of GIS, ITS and BIM by international standardization organizations and industry stakeholders. The information models describe real-world features and concepts in a context defined by the application domain's role in the digital geospatial environment. Less concern has been put on the use of information across application domain. Furthermore, different technologies have been used for information modelling, and the technologies have been used in various ways in the different application domains, even within each domain (ISO/TC 59/SC 13, 2020, ISO/TC 211, 2020a). As a result, there are fundamental differences between the representations of real-world phenomena. A physical object or a restriction in the road network, its location and extent, can be described in significantly different ways. Furthermore, formats for storing and exchanging information are developed for use and exchange within a specific application domain and not for information exchange between domains. <u>The differences between information models are obstacles for the common understanding of interpretation rules required for information exchange</u>.

1.3 RESEARCH MOTIVATION

Research on geospatial information in GIS, ITS and BIM has shown the potential and challenges for interoperability. In particular, many studies have suggested transformations between the GIS model CityGML and the BIM model IFC (Zhu et al., 2018, Liu et al., 2017, Song et al., 2017). Others have studied specific data sets and project models for geospatial information in ITS (Chen et al., 2011, Richter and Scholz, 2017, Jomrich et al., 2017, Borzacchiello et al., 2016). However, the studies have focused on specific data sets and information models for one or two application domains, while little research has been found on the integration of core concepts and approaches for information modelling from all three application domains.

Therefore, the motivation for this thesis was to find ways to achieve improved interoperability between GIS, ITS and BIM through a shared understanding of the core concepts for representations of the real world in information models.

Given the distinct roles of GIS, ITS and BIM in the digital geospatial environment, interoperability through full harmonization of information models would not be appropriate. Therefore, several studies have discussed the use of Semantic Web technologies for linking models and transforming information as a supplemental approach to harmonization of information models (Hor et al., 2016, Hbeich and Roxin, 2020, Roxin and Hbeich, 2019, Luiten et al., 2019, Luiten et al., 2016, Luiten et al., 2017).

In order to achieve the best possible integration, this thesis studies the combination of harmonizing information models and linking by applying Semantic Web technologies.

1.4 RESEARCH QUESTIONS

The problem statement and research motivation indicated a need for further research on approaches for improved interoperability between conceptual models for GIS, ITS and BIM. Therefore, the main research question for this thesis was:

How can approaches and technologies for information modelling be applied for harmonization and linking of conceptual models of geospatial information from the three application domains of GIS, ITS and BIM?

Figure 2 shows the scope of the main research question, with harmonization between core information models from the different domains and linking where harmonization is not possible.

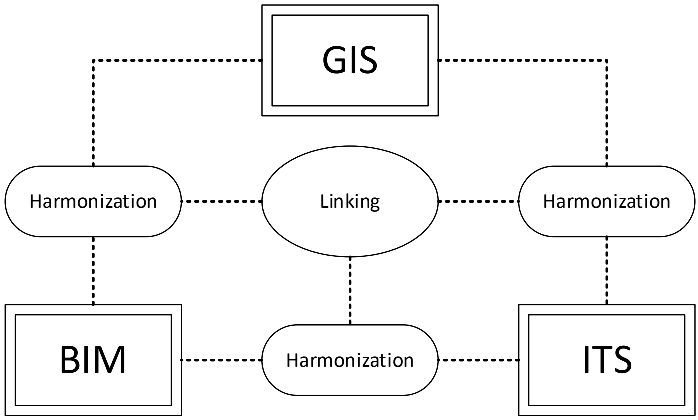


Figure 2. The scope of the main research question.

The main research question was supported by five sub research questions (SRQ), which in sum were intended to cover information modelling in the three application domains, the roles they have in the digital geospatial environment, and the relations between them.

The first sub research question laid the foundation for understanding the state of the art of information modelling in the three application domains, as well as the relevant context for geospatial information in each domain:

SRQ1: What approaches and technologies have been used for modelling geospatial information in GIS, ITS and BIM?

Based on the knowledge about the state of the art from SRQ1, the remaining sub research questions studied approaches for improving interoperability by harmonizing and linking information models. Figure 3 illustrates the scopes of sub research questions 2, 3, 4 and 5.

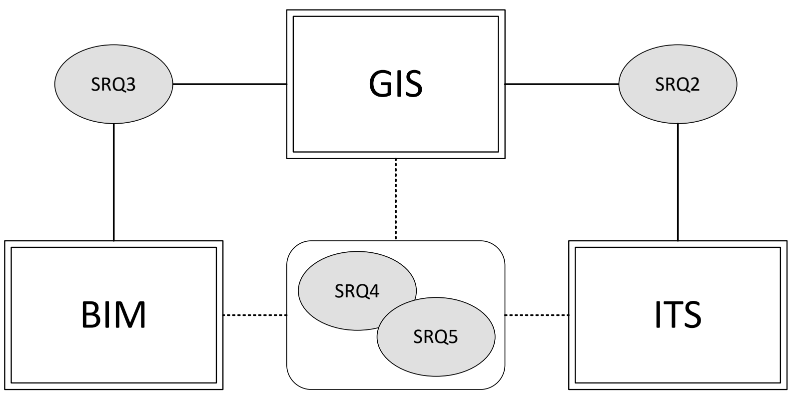


Figure 3. The scopes of sub research questions (SRQ) 2-5.

The second sub research question concerned interoperability between information models for GIS and ITS, in the context of information exchange from authorities' GIS databases to geospatial databases for ITS:

SRQ2: How can models for exchange of geospatial information from road and mapping authorities to geospatial databases for ITS be improved?

The third sub research question concerned the interoperability between information models for GIS and BIM by studying the potential for integration and linking:

SRQ3: How can information models and semantics for implementation technologies for BIM be integrated and linked with GIS standards?

The fourth sub research question concerned a common approach for information modelling in all three domains, while the fifth sub research question concerns the implementation and linking of models of geospatial information in the Semantic Web:

SRQ4: How can information models and semantics for implementation technologies for GIS, ITS and BIM be integrated into a joint modelling approach?

SRQ5: How can UML models of geospatial information be implemented as OWL Ontologies, for linking and mapping by applying Semantic Web technologies?

1.5 RESEARCH SCOPE

This thesis studies conceptual models of geospatial road-related information in GIS, ITS and BIM, the interoperability between the models, and possibilities for improved information exchange. The research scope covers approaches and technologies applied for describing the real world in information models and how they can be used to improve interoperability through harmonization and linking.

Issues concerning transformations between distinct representations of solid and volume geometries have been studied by many researchers, e.g., Deng et al. (Deng et al., 2016) and Donkers et al. (Donkers et al., 2016). Likewise, transformations between different location referencing methods have been studied by, e.g., CEN Technical Committee 278 for ITS (CEN/TC 278) (CEN/TC 278, 2018c, b). This thesis's scope includes how geometries and location references are described in information models, but not issues concerning transformations between different representations.

1.6 STRUCTURE OF THE THESIS

This thesis is based on a collection of appended articles which are put into a common context in order to answer the research questions. The thesis contains six chapters:

Chapter one introduces the background and motivation for the research and defines the research questions. Chapter two sets the frame of reference for the research by describing the theory behind information modelling and relevant information modelling technologies. Chapter three presents the scientific approach with the methods and materials used for the research. Chapter four summarizes findings from the appended articles, while the results are discussed for each sub research question in chapter five. Finally, conclusions are presented in chapter six.

1.7 APPENDED ARTICLES AND PUBLICATIONS

1.7.1 Appended articles

The thesis is supported by the following publications, which are appended at the end of the thesis.

Conference proceeding 1

Jetlund, K. (2018). "Experiences and challenges with standards for location referencing from the GIS and ITS domains." 25th ITS World Congress, Copenhagen, Denmark, 17-21 September 2018 (Jetlund, 2018a).

The conference proceeding presented an overview of relevant standards and specifications for geospatial information in GIS and ITS, identified challenges for the interoperability and suggested future research.

Author contribution: Carried out all work for the conference presentation and proceedings.

Article 1

Jetlund, K., E. Onstein and L. Huang (2019). "Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model." Isprs International Journal of Geo-Information 8(3): 141 (Jetlund et al., 2019b).

Article 1 studied the interoperability between GIS and geospatial databases for ITS. A prototype for information exchange from road databases to ITS databases for route planning and navigation was developed and tested.

Author contribution: Developed the concept and methodology, performed analysis, investigations, data curation and visualizations. Wrote the original draft and carried out the review process. Co-authors contributed to methodology, validation, editing and the review process.

Article 2

Jetlund, K. (2018). "IMPROVEMENTS IN AUTOMATED DERIVATION OF OWL ONTOLOGIES FROM GEOSPATIAL UML MODELS." Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-4: 283-290 (Jetlund, 2018b).

Article 2 investigated methods for transforming UML models of geospatial information to OWL for use in the Semantic Web. Improvements were suggested for conversion methods as well as UML models.

Author contribution: Carried out all work for the article and the conference presentation.

Article 3

Jetlund, K., E. Onstein and L. Huang (2019). "Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies." ISPRS International Journal of Geo-Information 8(9): 365 (Jetlund et al., 2019a).

Article 3 was a further development of article 2. An adapted UML profile and adapted rules for UML modelling of geospatial information for improved OWL implementation were presented and tested.

Author contribution: Developed the concept and methodology, performed analysis, investigations, data curation and visualizations. Wrote the original draft and carried out the review process. Co-authors contributed to methodology, validation, editing and the review process.

Article 4

Jetlund, K., E. Onstein and L. Huang (2020). "IFC Schemas in ISO/TC 211 compliant UML for improved interoperability between BIM and GIS." ISPRS International Journal of Geo-Information 9(4) (Jetlund et al., 2020).

Article 4 studied the interoperability between GIS and BIM. The IFC information model was transformed into a UML model according to GIS standards and linked with core GIS concepts. The model was tested through implementation schemas for both domains.

Author contribution: Developed the concept and methodology, performed analysis, investigations, data curation and visualizations. Wrote the original draft and carried out the review process. Co-authors contributed to methodology, validation, editing and the review process.

Article 5

Jetlund, K. (2020). "A STRUCTURE OF UML PROFILES FOR MODELLING OF GEOSPATIAL INFORMATION IN GIS, ITS AND BIM." ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. VI-4/W1-2020: 101-108 (Jetlund, 2020).

Article 5 studied the interoperability between GIS, ITS and BIM by suggesting and testing a structure of UML profiles for implementation in all three application domains.

Author contribution: Carried out all work for the article and the conference presentation.

1.7.2 Other relevant contributions

Besides the appended articles, the studies have included contributions to other relevant work, listed below:

Standardization project 1

CEN/TC 278, "CEN/TS 17268:2018 Intelligent transport systems — ITS spatial data — Data exchange on changes in road attributes." 2018, CEN: Brussels, Belgium (CEN/TC 278, 2018a).

The technical specification defines a UML model and GML implementation schemas to exchange road information from road authorities to map providers for ITS.

Author contribution: Acted as the group's expert in geographic data standards. Responsible for the UML model and resources for XML Implementation. Contributed to writing and review of the Technical Specification along with co-experts.

Standardization project 2

CEN/TC 278, "Intelligent transport systems — Location Referencing Harmonisation for Urban ITS — State of the art and guidelines." 2018 (CEN/TC 278, 2018c).

The Technical Report presented an overview of location referencing methods and how they had been used in ITS standards.

Author contribution: Acted as the group's expert in location referencing. Wrote the chapter on location referencing. Contributed to writing and review of the Technical Report along with co-experts.

CEN/TC 278, "Intelligent transport systems — Location Referencing Harmonisation for Urban-ITS — Part 2: Translation methods." 2018 (CEN/TC 278, 2018b).

The Technical Specification describes transformations between location referencing methods used in ITS standards.

Author contribution: Acted as the group's expert in location referencing. Contributed to writing and review of the Technical Specification along with co-experts.

Standardization project 3

ISO/TC 211, "ISO CD/TR 19169 Geographic Information — Gap-analysis: To map and describe the differences between the current GDF and ISO/TC211 conceptual models to suggest ways harmonize and resolve conflicting issues." 2020, ISO: Geneva, Switzerland (ISO/TC 211, 2020a).

The Technical Report was a joint work between the ISO Technical Committees for ITS (ISO/TC 204) and Geographic information/Geomatics (ISO/TC 211). The ITS standard GDF was compared with ISO/TC 211 standards for GIS, including identifying gaps and suggesting actions for improved interoperability.

Author contribution: Contributed to writing and reviewing the Technical Report along with co-experts.

Standardization project 4

ISO/TC 59/SC 13, "ISO DTR 23262.2 GIS (Geospatial) / BIM interoperability." 2020, ISO: Geneva, Switzerland (ISO/TC 59/SC 13, 2020).

The Technical Report was a joint work between the ISO Technical Committee for BIM (ISO/TC 59 SC 13) and ISO/TC 211. Technical barriers between the GIS and BIM domains were investigated, and new work items for improved interoperability were suggested.

Author contribution: Contributed to writing and reviewing the Technical Report along with co-experts.

Handbook Chapter 1

Coetzee, S., R. Plews, J. Brodeur, J. Hjelmager, A. Jones, K. Jetlund, R. Grillmayer and C. Wasström (2019). "Standards - Making Geographic Information Discoverable, Accessible and Usable for Modern Cartography." Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management. J. Döllner, M. Jobst and P. Schmitz. Cham, Springer International Publishing: 325-344 (Coetzee et al., 2019).

The handbook chapter described standards development for geospatial information, resources available for implementing standards, and implementations examples.

Author contribution: Wrote the clause on UML models. Contributed to writing and review of the handbook chapter along with co-experts.

Handbook Chapter 2

Jetlund, K. and B. Neuhäuser, "GIS for Transportation." Springer Handbook of Geographic Information (Jetlund and Neuhäuser, (Awaiting publication)).

The handbook chapter described the use of geospatial information in transportation, including the theory of navigable digital transport network models and location referencing methods. Examples of databases and services with geospatial information from transport authorities, open data sources and commercial map providers were presented.

Author contribution: Developed concept and wrote the handbook chapter in cooperation with the co-author.

2 THEORY

"If language is not correct, then what is said is not what is meant;
if what is said is not what is meant, then what ought to be done
remains undone."
– Confucius.

2.1 DATA, INFORMATION, KNOWLEDGE AND WISDOM

2.1.1 The DIKW Hierarchy

This thesis studies information modelling. Therefore, the concept of information needs to be defined, including its relations to, and differences from three other concepts: data, knowledge and wisdom. These concepts have been discussed since the ancient Greek philosophers Plato and Aristoteles, and later by Descartes, Kant and others. In more recent times, the discussion has moved from the philosophic perspective into how the concepts are used in information technologies (Rowley, 2007, Jennex, 2017).

One of the first researchers to describe the hierarchy of Data, Information, Knowledge and Wisdom (DIKW) was Ackoff (Ackoff, 1989), who defined five levels, ranging from data through information, knowledge and understanding to wisdom. Other researchers disputed that understanding is a distinct level and described the DIKW Hierarchy (Rowley, 2007, Bellinger et al., 2004) as a pyramid where data is at the lower level, followed by information and knowledge, and wisdom at the pinnacle, as illustrated in Figure 4. Each level of the hierarchy, starting from the top, relies on the level below: wisdom is gained from knowledge; knowledge is gained from information; information is gained from data. The wisdom level has later been disputed in the context of machine learning and automated decision making (Hoppe et al., 2011, Jennex, 2017).

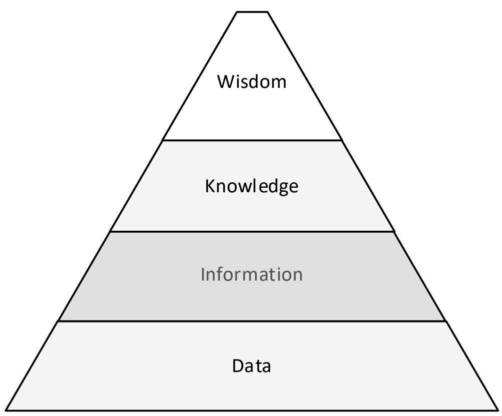


Figure 4. The DIKW Hierarchy. Adapted from Rowley (Rowley, 2007).

Bellinger et al. (Bellinger et al., 2004) suggested that understanding is a part of the transition from one level in the hierarchy to the level above. They described the transition in a graph where understanding and connectedness are the axes, as illustrated in Figure 5. By adding more connections, more understanding can be gained, and from more understanding, more connections can be added. A transition from one level to the next is gained by increased understanding and connectedness.

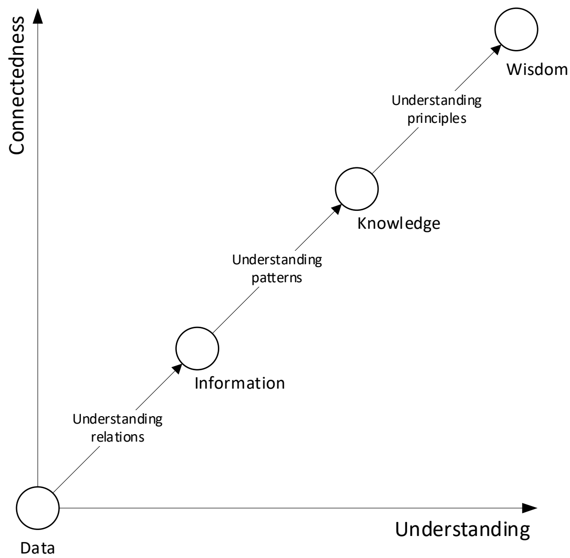


Figure 5. The DIKW graph. Adapted from Bellinger et al. (Bellinger et al., 2004).

2.1.2 Data

Data is the founding concept upon which all other concepts in the DIKW Hierarchy depend. Schenk and Wilson (Schenck and Wilson, 1994) and Ackoff (Ackoff, 1989) simply stated that data are symbols. Other researchers suggested that data are discrete, objective facts or observations, which are unorganized and unprocessed, and do not convey any specific meaning (Rowley, 2007). While ISO/IEC 2382:2015 (ISO/IEC JTC 1, 2015) defines data as "*reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing*". In short: data represent information.

The digital representations of data are values that represent the observed properties of objects and events. Two examples of digital data related to road traffic are the value "90" in a network dataset or identified by the vehicle sensors from a sign plate along the road, and an observation of the value "2" from a temperature sensor.

2.1.3 Information

While data are the raw facts about something, information is the result of understanding or interpreting the data in a specific context. Schenck and Wilson (Schenck and Wilson, 1994) described information as *"data put into context"*, while the formal definition of information in ISO/IEC 2382:2015 (ISO/IEC JTC 1, 2015) is "*knowledge concerning objects, such as facts, events,* *things, processes, or ideas, including concepts, that within a certain context has a particular meaning*". Ackoff (Ackoff, 1989) stated that information answers questions that begin with words such as "who", "what", "where" and "how many". Besides, Ackoff stated that the difference between data and information is functional, not structural: Both concepts represent the properties of objects and events, but the information is processed to be more useful.

The digital representation of information is the values from data presented in a context where they describe specific properties and have specific data types. The value "90" in the example above is more valuable if it is put into the context of speed limits with km/h as the unit of measure. Likewise, the temperature observation becomes useful by adding degree Celsius as the unit of measure and stating that it is outdoors.

2.1.4 Knowledge

The Oxford English Dictionary defines knowledge as a broad term that includes facts and assumptions acquired through education or experiences. Furthermore, knowledge in a specific domain is the fundament for making decisions and taking actions (Oxford Dictionaries, 2019). Ackoff (Ackoff, 1989) stated that knowledge answers questions that begin with "how-to", while Bellinger et al. (Bellinger et al., 2004) suggested that knowledge is the appropriate collection of information. ISO/IEC 2382:2015 (ISO/IEC JTC 1, 2015) defines knowledge as "*collection of facts, events, beliefs, and rules, organized for systematic use*".

Knowledge can be gained by combining pieces of information through relations with meaning and understanding patterns. While digital data and information represent the properties of objects and events, knowledge is more likely to be represented in knowledge graphs with relations between pieces of information (Regoczei and Hirst, 1992).

For the speed limit example, an automated vehicle can gain knowledge by combining the information about the identified speed limit with the semantics of the concept speed limit. The vehicle will then know how fast it is allowed to drive. While the temperature information can be combined with information that characterizes challenging conditions and give the vehicle knowledge about possible ice on the road, from which the vehicle will know that it should reduce the speed from the legal to a safe level.

2.1.5 Wisdom

Wisdom is the final level in the DIKW hierarchy and has mostly been considered a uniquely human state. Ackoff (Ackoff, 1989) and Bellinger et al. (Bellinger et al., 2004) suggested that wisdom is an evaluated understanding based on a personal judgement. Likewise, Rowley (Rowley, 2007) found definitions stating that "*wisdom is accumulated knowledge that allows one to understand how to apply concepts from one domain to new situations or problems, and the ability to plan for the future*" and "*wisdom is a very elusive concept. It has more to do with human intuition, understanding, interpretation and actions, than with systems*". Other researchers discussed whether the wisdom level should be in the hierarchy, and suggested it should be related to intelligence in a more complex view of knowledge management adapted for machine learning and automated decision making (Hoppe et al., 2011, Jennex, 2017).

2.2 INFORMATION MODELLING

2.2.1 Information models and ontologies

Interpretation rules that describe how data shall be interpreted in a given context are needed to extract information from data. Such interpretation rules are defined in information models (Schenck and Wilson, 1994). Schenck and Wilson (Schenck and Wilson, 1994) defined an information model as *"… a formal description of types of ideas, facts and processes which together form a model of a portion of interest of the real world and which provides an explicit set of interpretation rules."* Similarly, Zhao et al. (Zhao et al., 2011) defined information models as "… *a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse*". From the two definitions, and Schenk and Wilson's definitions of data and information, information models represent objects from the real world, the relations between them, and constraints, rules and operations needed to specify the objects and their behavior.

Hitzler et al. (Hitzler et al., 2012) defined ontologies for the Semantic Web as "*… a set of precise descriptive statements about some part of the world (usually referred to as the domain of interest or the subject matter of the ontology)"*. Likewise, Noy and McGuinness (Noy and McGuinness, 2001) stated that *"an ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them".* These definitions are equivalent to the definitions of an information model and show that information models and ontologies have equivalent purposes. In this thesis, the term information model is considered to include ontologies, while the term ontology will be used specifically for information models for the Semantic Web.

Another related concept is vocabulary, defined in ISO 1087:2019 (ISO/TC 37/SC 1, 2019) as a "*terminological dictionary that contains designations and definitions from one or more domains or subjects*". There is no clear distinction between the three concepts, but typically, information models and ontologies are more complex than vocabularies (Hitzler et al., 2012). While a vocabulary defines a concept and its meaning, the information model or the ontology will implement the concept in a modelling language.

Noy and McGuinness (Noy and McGuinness, 2001) listed several reasons for developing ontologies, which also apply to information models: To share a common understanding of the structure of information among people or software agents; enable reuse of domain knowledge; make domain assumptions explicit; separate domain knowledge from operational knowledge; and analyze domain knowledge. The development of an information model is not the goal itself; it is a tool for defining information for others to reuse, including humans and machines.

Models for specification and development of software systems have defined the role of information models in the bigger picture (Lankhorst, 2009). The Reference Model of Open Distributed Processing (RM-ODP) (Kerry, 1995, ISO et al., 2020) defines five viewpoints for the specification of a complex system: the enterprise, information, computational, engineering and technology viewpoints, as illustrated in Figure 6. ISO 19101:2014 (ISO/TC 211, 2014) states that the concern of information models is to see the real world from the information viewpoint, which concerns the information and any constraints on the use and interpretation of that information.

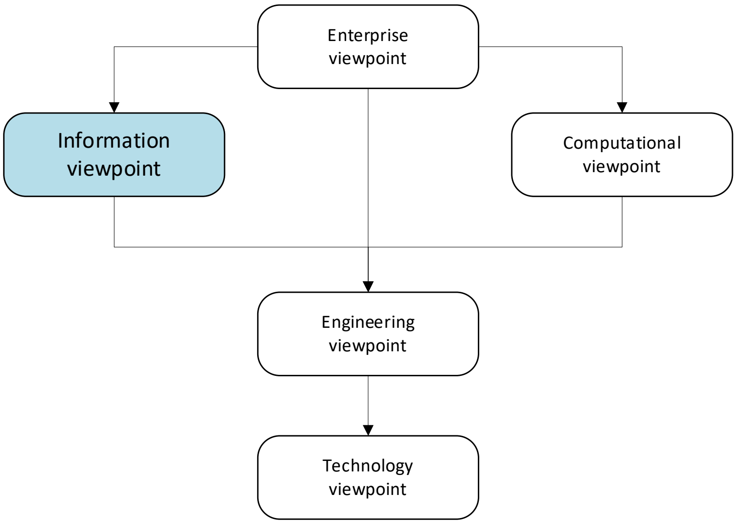


Figure 6. The RM-ODP Viewpoints focused on the information viewpoint. Adapted from ISO 19101-1:2014 (ISO/TC 211, 2014).

The 4+1 view of software architecture defined by Kruchten (Kruchten, 1995, Kruchten, 2004) is another model that addresses information modelling as part of a system, with five different views, as illustrated in Figure 7: the logical view; the process view; the development view; the physical view; and the use case view. While all five views are vital for software development, Miles and Hamilton (Miles and Hamilton, 2006) defined the logical view as the main view for information modelling. In the logical view, the system is "*decomposed into a set of key abstractions, taken (mostly) from the problem domain, in the form of objects or object classes*" (Kruchten, 1995).

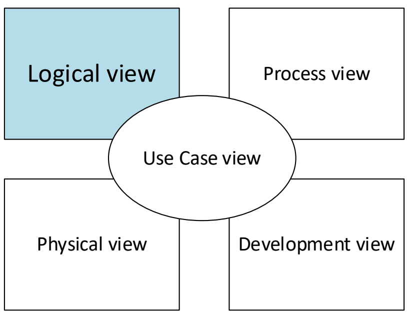


Figure 7: the 4+1 view model of software architecture, focused on the logical view. Adapted from Kruchten (Kruchten, 2004).

2.2.2 The Universe of Discourse and Context

Of course, one single information model cannot describe every concept from the real world. Therefore, the definitions of information models and ontologies state that they describe a *"portion of interest"*, *"domain of discourse"* or *"some part"* of the real world. Likewise, ISO 19101-1:2014 defines the term universe of discourse as a "*view of the real or hypothetical world that includes everything of interest*" (ISO/TC 211, 2014). The universe of discourse includes objects and their characteristics and can, for example, be a specific group of objects, such as buildings, rivers, roads or protected sites, or it can be all objects related to an application, such as road transport or asset management.

Furthermore, a perception of the real world will be distinct and depending on the scope, or context, of the model. Schenk and Wilson (Schenck and Wilson, 1994) stated that "*A scope also defines a context in which the model items reside, thus providing a specific viewpoint in which the items are defined*". The context for models of geospatial information will typically be the geospatial context, or it may be more specific such as "geospatial information for road navigation". The latter will put other requirements on how a road network is represented than a pure cartographic representation.

One of the fundamental steps of information modelling is to define the universe of discourse and the context it shall be perceived within. Noy and McGuinness (Noy and McGuinness, 2001) described seven steps of ontology development, where the first step was to determine the domain (universe of discourse) and scope (context). Likewise, Schenk and Wilson (Schenck and Wilson, 1994) and ISO 19103:2015 (ISO/TC 211, 2015) defined an initial phase where the scope and context are identified.

Figure 8 illustrates the universe of discourse as a defined part of the real world, perceived in a specified context, in order to be defined in a conceptual model.

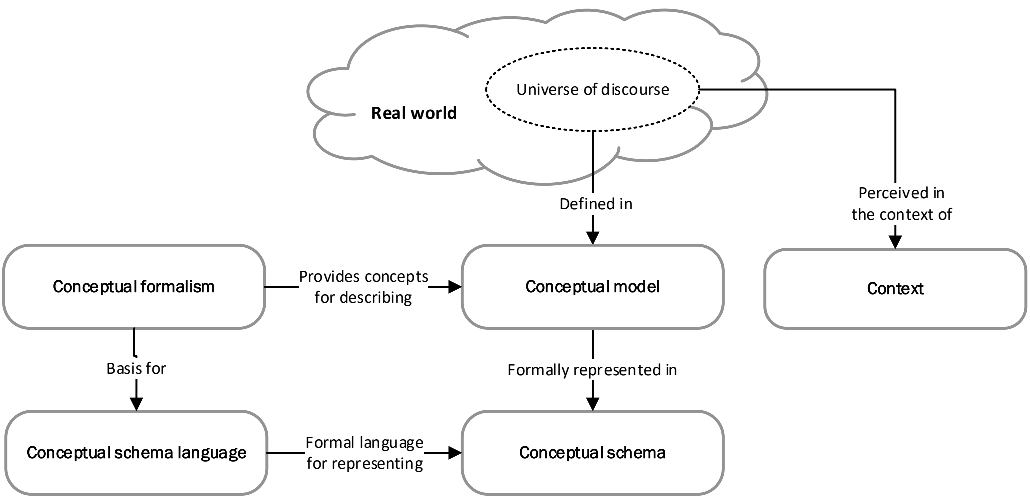


Figure 8. From the real world to conceptual models. Adapted from ISO 19101-1:2014 (ISO/TC 211, 2014).

2.2.3 Conceptual models and conceptual schemas

Pilone and Pitman (Pilone and Pitman, 2005) defined modelling as: "*a means to capture ideas, relationships, decisions, and requirements in a well-defined notation that can be applied to many different domains*". Figure 8 illustrates how ISO 19101-1:2014 (ISO/TC 211, 2014) describes the process of capturing concepts from the universe of discourse and representing it in a formal language.

In Figure 8, the universe of discourse is defined in a conceptual model. ISO 19101-1:2014 (ISO/TC 211, 2014) defines a conceptual model as a "*model that defines concepts of a universe of discourse*", while Schenk and Wilson (Schenck and Wilson, 1994) defined a conceptual information model as a model that is "*independent of any particular instantiation form*". From the two definitions, a conceptual model shall describe the concepts from the defined universe of discourse, and it shall be independent of implementation technology. The independency of implementation technologies is one of the basic principles for Platform Independent Models (PIM) in Model-Driven Architecture (MDA) (Object Management Group, 2014a).

Furthermore, the conceptual model is formally represented in a conceptual schema, defined in ISO 19101-1:2014 (ISO/TC 211, 2014) as the "*formal description of a conceptual model*". A formal language for conceptual modelling is needed to create such a formalized description, known as a conceptual schema language. Just like human communication is based on languages with vocabularies and grammatical rules, a description of the real world in information models for machine-to-machine communication needs a defined language. Well-known examples of conceptual schema languages are the Unified Modelling Language (UML) (Object Management Group, 2017), EXPRESS (ISO/TC 184 SC 4, 2004) and the Web Ontology Language (OWL) (Hitzler et al., 2012). These languages are further described in section 2.3.

Conceptual schema languages are based on a conceptual formalism, defined as "*set of modelling concepts used to describe a conceptual model*" in ISO 19101-1:2014 (ISO/TC 211, 2014). The modelling concepts are defined in metamodels, such as the Meta Object Facility (MOF) meta-meta model (Object Management Group, 2019) and the UML metamodel (Object Management Group, 2017).

2.3 CONCEPTUAL SCHEMA LANGUAGES

2.3.1 Graphical and lexical representations

ISO 19101:2014 (ISO/TC 211, 2014) and Schenk and Wilson (Schenck and Wilson, 1994) have defined conceptual schema languages as graphical or lexical, or as combinations with both categories of representations. A graphical representation applies icons and connecting lines between them, while a lexical representation applies words and character symbols to describe the model.

A graphical representation is useful for human communication and discussions, and any modelling process will most often include sketching of graphical representations. Miles and Hamilton (Miles and Hamilton, 2006) described the graphical representations as windows into the model. However, to be machine-interpretable, information models need to have more semantics than what is shown in the graphical representation. Furthermore, the content must apply to formalized rules that can be understood by machines, and it must be provided in a native, machine-readable format (IEC, 2019). For this purpose, the lexical representation is more useful.

Schenk and Wilson (Schenck and Wilson, 1994) found that a combination of graphical and lexical representations had significant benefits for developing an unbiased and complete model.

2.3.2 The Unified Modelling Language (UML)

UML (Object Management Group, 2017) was first introduced in 1997, after several years of discussions, development and harmonization of three different modelling methods (Miles and Hamilton, 2006). The Object Management Group (OMG) is responsible for the standardization and management of UML, currently at revision 2.5.1, published in December 2017 (Object Management Group, 2017), while ISO has adopted UML as ISO/IEC 19505:2012 part 1 and 2 (ISO/IEC JTC 1/SC 7, 2012a, b), currently at UML version 2.4.1. UML is a widely used language for system development and information modelling. It has been the selected conceptual language for GIS standards in ISO/TC 211 since 1998 (ISO/TC 211, 2020c) and is also extensively used for domain-specific GIS models developed by, for example, the Open Geospatial Consortium (OGC). Besides, UML has been used for information modelling of the ITS standard GDF since 2001 (Shibata, 2018), developed by ISO/TC 204; and over the last years for the infrastructure extensions of the BIM standard IFC, developed by buildingSMART International (bSI) (buildingSMART International, 2019c).

UML is structured around a set of standardized graphical representations (diagrams) that illustrate parts of a model according to Kruchten's 4+1 view (Kruchten, 1995), while the complete model is defined through standardized semantics that may not be visible in diagrams. Three types of diagrams are the basis for information modelling: The class diagram with classes, attributes, operations, and relations between classes; the package diagram with packages that form groups of objects and relations between them; and the object diagram for instances of elements in the model, which is typically used to test and describe how the concepts from the model will work in practice.

The conceptual formalism for UML is defined through metaclasses for core information concepts in the UML metamodel (Object Management Group, 2017), founded on the MOF meta-metamodell (Object Management Group, 2019). Information models described in UML can have semantics for classes, attributes and relations with standardized element properties, for example, name, definition and type. Specialized concepts, semantics and restrictions for UML used in a specific domain can be formalized in UML profiles through the stereotype mechanism, which defines extensions of UML metaclasses. Stereotypes can have properties for additional semantics, represented as tagged values, and constraints that restrict the concept. The UML specification (Object Management Group, 2017) requires that a stereotype shall extend one or more UML metaclasses and shall not exist independently. Furthermore, extensions defined by UML profiles shall remain conformant with the UML metamodel. These requirements are vital for further analysis of UML usage in this thesis.

UML does not have implicit semantics for constraints on elements in information models but can be extended with constraints written in the Object Constraint Language (OCL) (Object Management Group, 2014b). OCL was initially developed by IBM and has been adopted and developed further by the OMG. ISO has also adopted OCL as an International standard, currently at OCL version 2.3.1, standardized in ISO/IEC 19507-1:2012 (ISO/IEC JTC 1/SC 7, 2012b).

With the combination of standardized graphical representations and additional non-graphical semantics, UML can be considered a combination of a graphical and lexical language. However, UML models are described internally in specialized tools for UML modelling, and not in a machine-readable and application-independent format. Therefore, the models need to be converted to an external implementation format such as XML Schema to become machine-readable. Besides, the OMG has specified the XML Metadata Interchange (XMI) format (Object Management Group, 2015b) for exchange of UML models. Several researchers have found limitations in using XMI as an exchange format. For example, Kutzner (Kutzner, 2016) pointed at complex mapping rules and differing implementations as critical challenges.

2.3.3 EXPRESS

The EXPRESS modelling language was developed as a part of the ISO 10303 STEP series of standards. The STEP standards are developed by the ISO Technical Committee 184 for automation systems and integration, Sub Committee 4 for industrial data (ISO/TC 184 SC 4) (ISO, 2020). Formal development of the STEP standards started in 1984, focusing on software for product engineering, construction and manufacturing. At that time, there was no standardized functional language for information modelling. Therefore, EXPRESS was developed as part 11 of ISO 10303 (Object Management Group, 2015a). The first EXPRESS version was standardized in 1994, and the current version from 2004 is the second edition (ISO/TC 184 SC 4, 2004).

EXPRESS is a lexical language with core information concepts and advanced semantics for describing rules and constraints for classes, attributes and relations. The graphical language EXPRESS-G is specified in the standard but limited to a subset of the language. The language EXPRESS-I for describing instances has been developed as part 12 of ISO 10303 (ISO/TC 184 SC 4, 1997). Two parts in the ISO 10303 series describe implementation from EXPRESS models, including specifications of the formats as well as a mapping from the EXPRESS model to implementation serializations: Part 21 describes the STEP exchange structure (ISO/TC 184 SC 4, 2016), while part 28 describes an implementation based on XML Schema (ISO/TC 184 SC 4, 2007). Furthermore, a specific standard for converting EXPRESS models to the UML exchange format XMI has been developed as part 25 of ISO 10303 (ISO/TC 184 SC 4, 2005). Finally, an EXPRESS metamodel based on the MOF meta-meta model (Object Management Group, 2019) was developed in 2005 and standardized by the OMG in 2015 (Object Management Group, 2015a).

Like UML, EXPRESS has been used for a range of domains. More than 300 major information models specified in EXPRESS were known in 2005, especially in the manufacturing industry (Object Management Group, 2015a). An essential use of EXPRESS is the BIM standard IFC (buildingSMART International, 2019a, ISO/TC 59/SC 13, 2018a). ISO/TC 211 standards for GIS were also specified in EXPRESS for a short period in the late 1990s (ISO/TC 211, 2020c).

2.3.4 Resource Description Framework (RDF) and Web Ontology Language (OWL)

OWL (Hitzler et al., 2012) is the language for describing information and knowledge in ontologies for the Semantic Web. The Semantic Web is the more structured version of the World Wide Web where information and knowledge are provided in triples and graphs that can be accessed, understood, and shared independently of applications.

The basic framework for the Semantic Web is the Resource Description Framework (RDF) (Manola et al., 2014), standardized by the World Wide Web Consortium (W3C) (World Wide Web Consortium, 2020). RDF has been a recommendation from the W3C since 1999 and is currently at version 1.1 (Manola et al., 2014). In RDF, statements are described in triples consisting of a subject, a predicate and an object. Subjects and objects are resources that can be anything from a concrete physical phenomenon to an abstract concept, and the predicate describes the connection between the subject and the object. A subject in one triple may be an object in another triple, and that way, a set of triples form a graph of knowledge.

OWL was adopted as a W3C recommendation in 2004 and is currently in the second edition (OWL 2) (Hitzler et al., 2012). OWL is based on RDF and uses the same principles for statements in triples. While RDF describes statements about individuals, OWL describes the structure of the information in ontologies. Like RDF, OWL is a pure lexical language, as there is no standardized graphical representation. Applications for ontology development may have functionality for presenting triples graphically, but each application has its own presentation rules.

Like other conceptual schema languages, OWL can describe classes and hierarchies of classes, relationships, properties and constraints. However, as OWL is based on RDF and uses the same principles for describing statements in triples, information can be combined from triples into knowledge graphs, which is a clear distinction between modelling in OWL and modelling in UML or EXPRESS. This ability to extend information into knowledge by inferencing or reasoning is one of the central characteristics of The Semantic Web (Hitzler et al., 2012, Manola et al., 2014).

To combine concepts for information modelling in OWL and UML, the OMG has developed metamodels for RDF and OWL, based on the MOF meta-meta model (Object Management Group, 2019), including a UML profile for RDF and OWL (Object Management Group, 2014c).

2.4 MODELLING APPROACHES

2.4.1 Model-Driven Architecture (MDA)

Model-Driven Architecture (MDA) (Object Management Group, 2020) is a framework for software development and information modelling specified by the OMG. One of the fundamental principles in MDA is to develop models that are independent of any implementation technology (platform), categorized as Platform-Independent Models (PIM). Models with semantics for implementation in specific technologies are categorized as Platform-Specific Models (PSM), while implementation schemas are categorized as Platform Models (PM). By applying transformations, PSMs can be derived and further elaborated from PIMs, and implementable resources (PMs) can be derived from PSMs.

Models in MDA are organized in architectural layers according to their level of abstraction. There is a close relationship between the level of abstraction and the dependency on domains and implementation technologies. The most abstract models eliminate details and are scoped for a broader set of systems, independent of specific domains and platforms. Some examples of abstract models are general concepts for geometry, location referencing, time and network topology. The abstract concepts are realized and used in less abstract models that are more detailed and scoped for specific domains such as buildings and transport networks. Several levels of abstraction can exist for one category of models, for example, the two levels of abstraction for PIMs defined in ISO 19103:2015 (ISO/TC 211, 2015): Abstract Conceptual Schemas and Application (Domain) Conceptual Schemas.

Figure 9 illustrates an MDA structure with two levels of abstraction for PIMs: abstract conceptual models and domain models. The two domain models for Transport Networks and Buildings realize and use concepts from the abstract models for Geometry, Time and Network. Furthermore, the domain models are transformed to PSMs and extended with semantics for three different technologies: GML, SQL and JSON. Finally, PMs for the distinct technologies are derived from the PSMs through transformations.

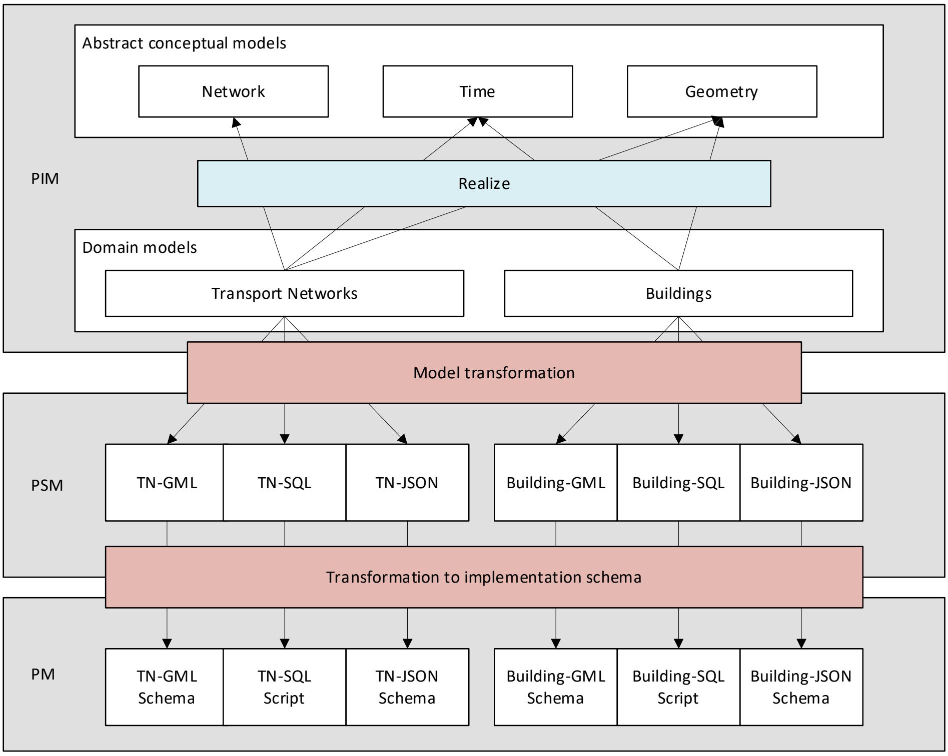


Figure 9. MDA levels of abstraction and transformations.

2.4.2 Closed and Open World Assumptions

The Closed World Assumption (CWA) and the Open World Assumption (OWA) are two distinct approaches for representing the universe of discourse in information models. Information modelling based on the CWA assumes that the universe of discourse is wholly represented in the information model within the given context. Every relevant aspect of the universe of discourse is assumed to be known, and information is considered wrong if it does not fit the model. Contrary, an information model following the OWA describes the real world only as known at the time and in the given context. More information outside the current knowledge may be added, even conflicting information based on different views on the real world. No conclusions that assume all information is available can be drawn, and nothing is true or false unless explicitly stated. (Allemang and Hendler, 2011, Hart and Dolbear, 2016).

OWA and the ability for anyone to say anything about any topic (AAA) is fundamental for the Semantic Web. Therefore, ontologies described in OWL need to adopt the OWA. Anyone may add additional characteristics as properties and classifications of objects besides what is defined in the ontology, so restrictions must be applied carefully. The representation in the ontology is a minimum classification that may be extended by anyone. Contrary, information modelling in UML and EXPRESS are founded on the CWA. The models describe a complete representation of the universe of discourse, with only the specified object types and their characteristics, and with specified restrictions. According to the CWA, the representation of a real-world object is an abstraction of the real-world object limited to the specified context (Hart and Dolbear, 2016).

2.5 INTEROPERABILITY

The term "interoperability" has been used in many contexts and with different definitions. Essendorfer et al. (Essendorfer et al., 2017) compared several definitions and referred to interoperability as "*the ability of two or more entities to exchange and use resources following a well-defined and agreed process to achieve a common goal*". ISO/IEC 2382:2015 (ISO/IEC JTC 1, 2015) has defined interoperability as *"capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units"*. The definition in ISO/IEC 2382:2015 has been adopted by ISO/TC 211 for use in GIS standards (ISO/TC 211, 2020b). Roxin and Hbeich (Roxin and Hbeich, 2019) considered GIS and BIM as systems comprising of several parts, where interoperability is achieved if the parts of the system and the overall system cooperate seamlessly in order to reach a common goal or function.

In this thesis, the term "interoperability" is used in the context of information models for geospatial information. Like in Roxin and Hbeich's research (Roxin and Hbeich, 2019), GIS, ITS and BIM are considered as systems comprised of several parts. Interoperability of information models is reached when the systems and their parts can communicate, transfer and use information based on each system's information model without the need for a user to interact.

The Levels of Conceptual Interoperability Model (LCIM) was first proposed by Tolk and Maguira (Tolk and Muguira, 2003), and have later been elaborated into a model with seven levels of conceptual interoperability, ranging from no interoperability to full conceptual interoperability (Roxin and Hbeich, 2019, Essendorfer et al., 2017, Axelsson, 2020, Kubicek et al., 2011). Table 1 lists the seven levels described by Wang et al. (Wang et al., 2009).

Table 1. The LCIM levels of conceptual interoperability. Adapted from Wang et al. (Wang et al., 2009).

Level	**Layer name**	**Premise**	**Information defined**
6	Conceptual	Common conceptual model	Assumptions, constraints
5	Dynamic	Common execution model	Effect of data
4	Pragmatic	Common workflow model	Use of data
3	Semantic	Common reference model	Meaning of data
2	Syntactic	Common structure	Structure of data
1	Technical	Common communication protocol	Bits and bytes
0	None	No connection	None

The levels of technical, syntactic and semantic interoperability concern the description of the information, while the levels of pragmatic, dynamic, and conceptual interoperability consider the context and use of each system's information in the interoperating systems. Therefore, the levels of syntactic and semantic interoperability are particularly relevant within this thesis's scope, while the levels of pragmatic, dynamic and conceptual interoperability are considered for further research.

Technical interoperability concerns the ability to communicate and exchange data (symbols) through physical interfaces or connections at hardware and software level. Syntactic interoperability concerns the structure and syntax of information by standard data formats for communication and exchange, such as XML. Veltman (Veltman, 2001) described four challenges for syntactic interoperability: identifying all the elements in each system; establishing rules for structuring these elements; creating crosswalks between equivalent elements; and agreeing on equivalent rules to bridge different cataloguing and registry systems. In the context of this thesis, the syntactic interoperability concerns the different languages for conceptual modelling, how they are used and how information models can be mapped between them.

Semantic interoperability concerns the understanding of terms and expressions used in communication and exchanged information. The International Electrotechnical Commission (IEC) has defined semantic interoperability as *"the ability of two or more assets (e.g. agents, machines, systems) to exchange and understand each other's data correctly"* (IEC, 2019). Veltman (Veltman, 2001) defined semantic interoperability as *"the ability of information systems to exchange information on the basis of shared, pre-established and negotiated meanings of terms and expressions"*. Essendorfer et al. (Essendorfer et al., 2017) referred to standard terms, concepts and metadata models as examples of semantic interoperability. Janowicz et al. *(Janowicz et al., 2010)* pointed at the lack of semantic interoperability due to "*the lack of meaningful descriptions of the actual content"* as a source for misunderstandings and incorrect use of geospatial information. Roxin and Hbeich (Roxin and Hbeich, 2019) pointed at different vocabularies and the lack of equivalencies between these as challenges for the semantic interoperability between GIS and BIM. In the context of this thesis, semantic interoperability concerns the use of common concepts in the different systems and the common understanding of equivalent concepts.

The pragmatic interoperability level extends the semantic interoperability by understanding the context where the information is used, while at the dynamic interoperability level, the systems can adapt to changes in the other systems, due to changes in context over time. Finally, at the conceptual interoperability level, the interoperating systems have a complete understanding of the other systems' conceptual models and are entirely aware of each system's processes, contexts, and modelling assumptions.

3 SCIENTIFIC APPROACH

"Computer science inverts the normal. In normal science, you're given a world, and your job is to find out the rules. In computer science, you give the computer the rules, and it creates the world."
– Alan Curtis Kay.

3.1 RESEARCH DESIGN

This thesis applies qualitative exploratory research to study information models and technologies for information modelling. Qualitative exploratory research has been described as useful for areas where there is a need to examine existing concepts to gain new knowledge and iteratively adjust the direction of the research along the way (Fossey et al., 2002, Hoepfl, 1997, Rowley, 2002, Yin, 2018).

The research design is the action plan for the path from questions to conclusions and shall ensure a clear view of what shall be achieved from the research (Rowley, 2002). To reach this goal, the research for this thesis was designed in an iterative process of five steps, adapted from Jassim (Jassim, 2019) and illustrated in Figure 10. The process was applied to each of the sub research questions to answer the main research question.

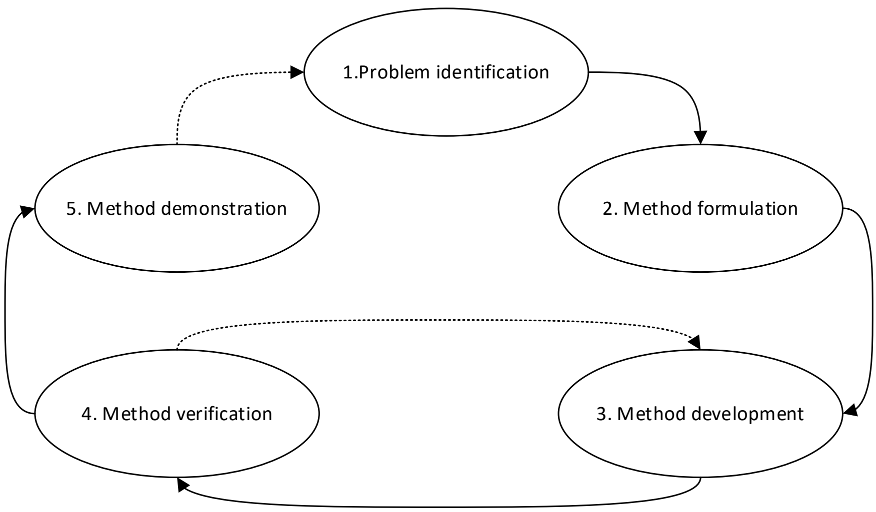


Figure 10. The process for designing the research.

The first step of the research design was the problem identification through initial studies of the state of the art. Knowledge of the state of the art was established by studies of existing information models and applied technologies for information modelling, supported by literature reviews of previous research. The state-of-the-art-knowledge provided the fundament for identifying the problem and specifying research questions.

In order to answer each research question, the second step was to formulate the research method. The selected method for the exploratory research was to perform further surveys through documentation reviews for SRQ1, while SRQ2-5 would be studied through experiments with prototypes. Document review surveys are fit for answering "what" questions like SRQ1 (Bowen, 2009, Rowley, 2002, Yin, 2018), while experiments are fit for answering the "how" questions in SRQ2-5 (Rowley, 2002, Yin, 2018). The iterative development of prototypes is an acknowledged approach for research striving towards developing new technical solutions (Golfarelli and Rizzi, 2011, Peng, 2005, Rasmussen et al., 2017).

The selected method was developed for each sub research question in the third step. For SRQ1, this included identifying relevant information models and searching for related research. For SRQ2-5, the prototypes were designed and developed. The applicability of the developed methods was tested in step four, which included a possibility for new input to the development step and iterations on development and testing. Finally, the results were demonstrated and described in articles in step five. As illustrated in Figure 10, the design process included the possibility for a full new iteration starting with an improved problem identification based on new knowledge from the research.

3.2 Methods

3.2.1 Literature review

The literature review applied for answering SRQ1 included studies of two main types of documentation: (1) Documentation of existing information models from the three application domains and (2) academic literature from research articles, project reports and thesis. The documentation was identified and studied through an iterative process, illustrated in Figure 11.

Existing information models and documentation of how they have been used were found by browsing standards catalogues from ISO, CEN and other standardization stakeholders, and by regular keyword-searches on the World Wide Web. Academic literature was found by applying combinations of keywords and filters in three academic search engines and bibliographic databases: Oria, Web of Science and Google Scholar. Research has found significant performance differences between academic search systems, concluding that no single search system is perfect (Gusenbauer and Haddaway, 2020). On the one hand, Oria and Web of Science have functionality for advanced filtering and could return few and relevant hits. Contrary, Google Scholar have less functionality for filtering but have been identified as the most comprehensive academic search engine (Gusenbauer, 2019), and found literature that was not included in the results from the two other search systems.

The search results were filtered by manual inspection of the findings. Furthermore, experiences from the inspection were used to further refine the searches for the inclusion of other relevant literature and exclusion of non-relevant literature. Inspection of cited literature (backward search) and literature that cited the selected literature (forward search) identified additional relevant literature and were used for extended searches based on additional keywords. Finally, the identified literature was studied, compared, and discussed to determine the state of the art.

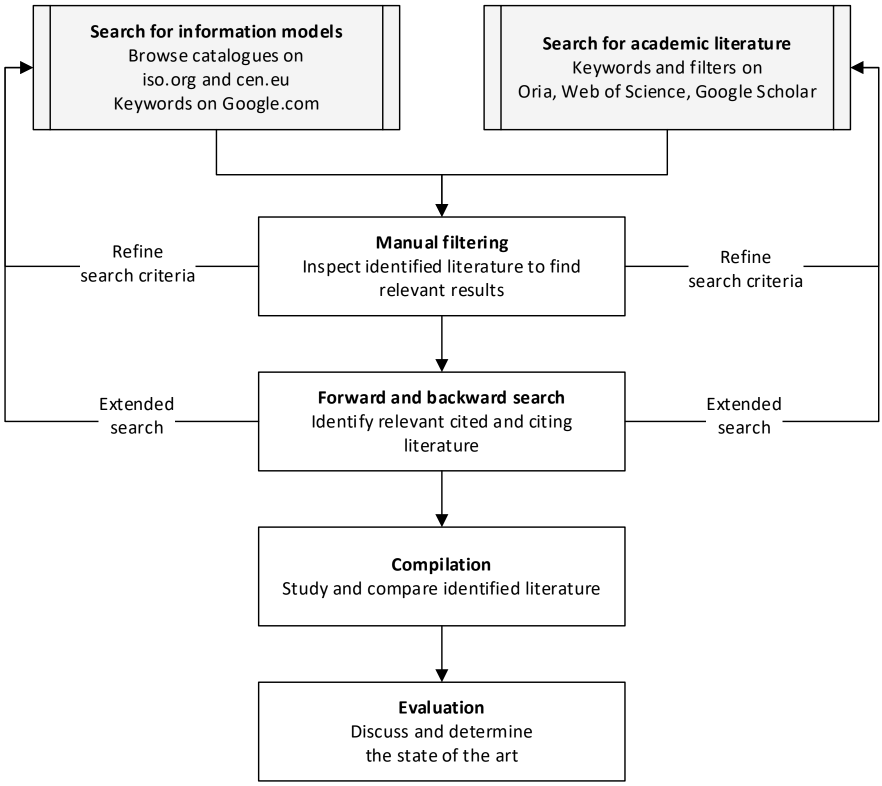


Figure 11. The method for literature review.

3.2.2 Experiments

The experiments for SRQ2-5 were performed according to an adaption of the iterative prototype experiment cycle described by Tronvoll et al. (Tronvoll et al., 2017), illustrated in Figure 12. First, existing information models, technologies and modelling approaches found in the literature review for SRQ1 were studied in detail in step one and then compared and evaluated in step two. The findings from step two lay the foundation for defining the purpose of the prototype in step three.

The iterative process started by designing and building the prototype in steps four and five (a). Like in the research described by Tronvoll et al. (Tronvoll et al., 2017), the definition of a prototype in this thesis is quite open. They defined a prototype as "*an approximation of the product along one or two dimensions*". The prototypes developed for this thesis were UML models, UML profiles and conversion scripts.

In parallel to the prototype development, the test environments were defined in step five (b), with principles for testing the prototypes through implementations and comparisons with existing solutions, performed in step six. The results from tests were analysed in step seven and served as input to improvements of the prototypes and refined test environments in new iterations, leading to the final evaluation and discussion of the results in step eight.

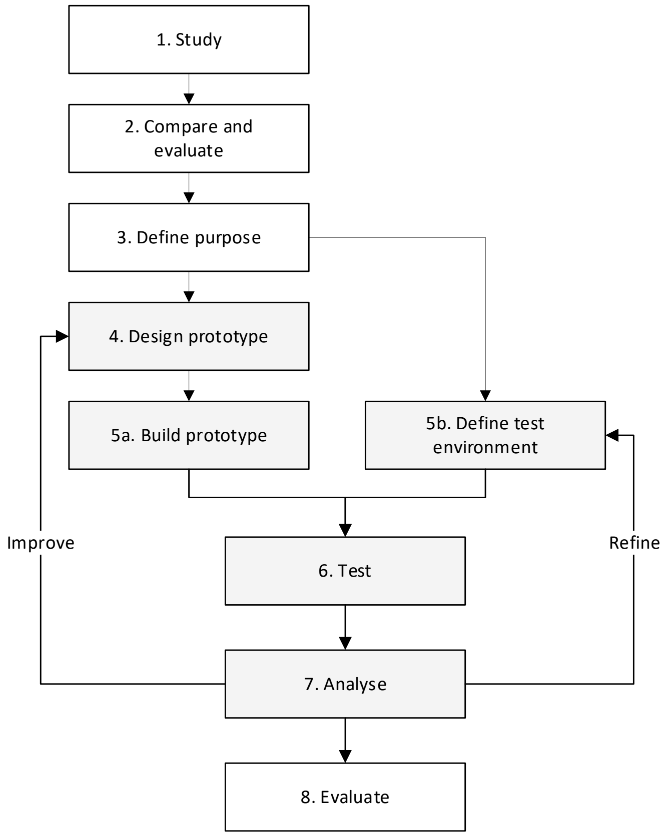


Figure 12. The iterative prototype experiment cycle.

3.3 RESEARCH PROCESS

The main research question and the five sub research questions were answered through the appended articles and contributions to other relevant work. As illustrated in Figure 13, there are close relations between the research questions. Several articles and related works contributed to answering several questions.

Findings from the **state-of-the-art analysis for SRQ1** were described in all six appended articles as well as other works. A study of approaches and technologies for information modelling in GIS and ITS was presented in Conference proceeding 1 (Jetlund, 2018a). The results were followed up in Article 1 (Jetlund et al., 2019b), Article 5 (Jetlund, 2020) and Standardization project 3 with the joint work between ISO/TC 211 and ISO/TC 204 (ISO/TC 211, 2020a). Approaches for information modelling in GIS and BIM were studied in Article 4 (Jetlund et al., 2020) and followed up in Article 5 (Jetlund, 2020) and also in Standardization project 4 with the joint work between ISO/TC 211 and ISO/TC 59/SC 13 (ISO/TC 59/SC 13, 2020). Besides, the use of Semantic Web technologies for geospatial information was studied in Article 2 (Jetlund, 2018b) and followed up in Article 3 (Jetlund et al., 2019a).

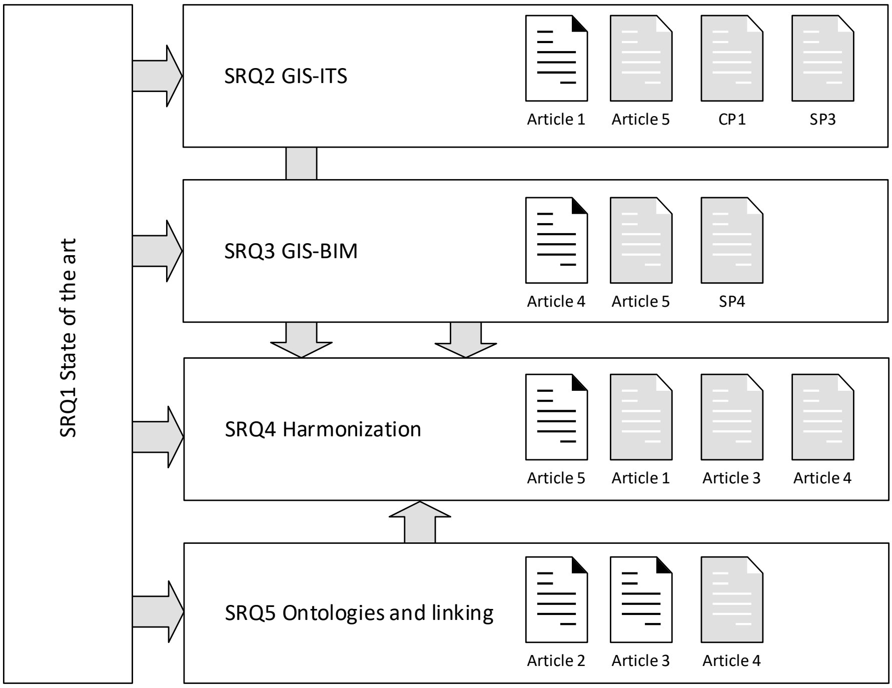


Figure 13. The research process. CP=Conference proceedings, SP=Standardization project.

SRQ2 asked for improved models for exchange of geospatial information from road and mapping authorities to geospatial databases for ITS. Existing solutions for exchanging information were discussed in Conference proceeding 1 (Jetlund, 2018a) and in Article 1 (Jetlund et al., 2019b). The main contribution to SRQ2 was the prototype for an improved UML model for information exchange, which was described, validated and demonstrated in Article 1 (Jetlund et al., 2019b). The prototype was used as the foundation for a proposed future ITS standard in Standardization project 3 with the joint work between ISO/TC 211 and ISO/TC 204 (ISO/TC 211, 2020a). A shared structure of UML profiles described in Article 5 (Jetlund, 2020) included profiles for the proposed solution.

SRQ3 asked for integration and linking of information models and semantics for implementation technologies for BIM with GIS standards. Information models for the two domains were studied in Standardization project 4 with the joint work between ISO/TC 211 and ISO/TC 59/SC 13 (ISO/TC 59/SC 13, 2020) and Article 4 (Jetlund et al., 2020). The main contribution to SRQ3 was the prototype for representing the IFC EXPRESS information model for BIM in the UML modelling framework from GIS standards, with links between core concepts from the two domains. The prototype was described and tested through implementation schemas in Article 4 (Jetlund et al., 2020). The structure of UML profiles for models of geospatial information described in Article 5 (Jetlund, 2020) included profiles for implementation in EXPRESS.

SRQ4 asked how information models and semantics for implementation technologies for GIS, ITS and BIM could be integrated into a joint modelling approach. A joint approach for GIS and ITS information models was suggested and tested in Article 1 (Jetlund et al., 2019b), while a joint approach for GIS and BIM was demonstrated in Article 4 (Jetlund et al., 2020). Article 3 (Jetlund et al., 2019a) described requirements for models that shall be implemented in OWL. The main contribution to SRQ4 was the suggested structure of UML profiles for all three domains, which was suggested, tested and demonstrated in Article 5 (Jetlund, 2020).

SRQ5 asked how UML models of geospatial information can be implemented as OWL Ontologies, for linking and mapping by applying Semantic Web technologies. The main contributions to SRQ5 were presented in Article 2 (Jetlund, 2018b) and Article 3 (Jetlund et al., 2019a), which studied rules for transformation from UML to OWL and suggested improvements. A prototype for an extended UML profile was suggested in Article 3 (Jetlund et al., 2019a) and demonstrated for sample cases. The use of Semantic Web technologies for discovery, linking and mapping was studied in Article 2 (Jetlund, 2018b), Article 3 (Jetlund et al., 2019a) and Article 4 (Jetlund et al., 2020).

4 SUMMARY OF FINDINGS FROM APPENDED ARTICLES

> *"Data is a precious thing and will last longer than the systems themselves."*
> *– Sir Tim Berners-Lee.*

4.1 CP1: A COMPARISON OF INFORMATION MODELS FOR GIS AND ITS

Title:

Experiences and challenges with standards for location referencing from the GIS and ITS domains (Jetlund, 2018a)

Purpose:

The purpose of Conference proceeding 1 was to study the state of the art for International standards and specifications for geospatial information in GIS and ITS, identify challenges for interoperability and suggest further research activities.

Findings:

The conference proceeding found that International standards for GIS and ITS have mostly been developed in application-specific silos and with little focus on interoperability between the two domains. Standards from ISO/TC 211 are the core International standards for GIS, while standards from ISO/TC 204 and CEN/TC 278 are the core International and European standards for geospatial information for ITS.

ISO/TC 211 standards are described in UML models according to an MDA with four levels of abstraction, as illustrated in Figure 14. The models are based on standardized rules for the use of UML and are maintained in a harmonized model repository where concepts are reused between the distinct models. Standardized rules for deriving implementation schemas from UML models are described in standards and implemented in conversion tools. The core models and rules from ISO/TC 211 standards have been essential for developing interoperable information models within the GIS domain. Specifications from the OGC are based on ISO/TC 211 standards and are widely implemented, and so are national and regional standards such as the INSPIRE specifications in Europe. Road and mapping authorities maintain geospatial information relevant to ITS, such as road networks, restrictions, and road equipment, in applications and databases based on ISO/TC 211 standards.

Like standards from ISO/TC 211, the ITS standards from ISO/TC 204 and CEN/TC 278 are also described in UML models. However, there are no overall modelling rules and no harmonized UML model for ITS standards, which is an obstacle for interoperable standards. Different rules for modelling and conversions to implementation schemas have been applied, and concepts are not reused between models. Therefore, service providers and end-users will need to deal with different information structures and exchange formats for related information.

The findings suggested that improved interoperability between GIS and ITS information models could be reached by adapting standard modelling rules and developing information models based on a joint UML profile. Furthermore, future research activities should include standards that have been developed by other standardization actors within the two application domains.

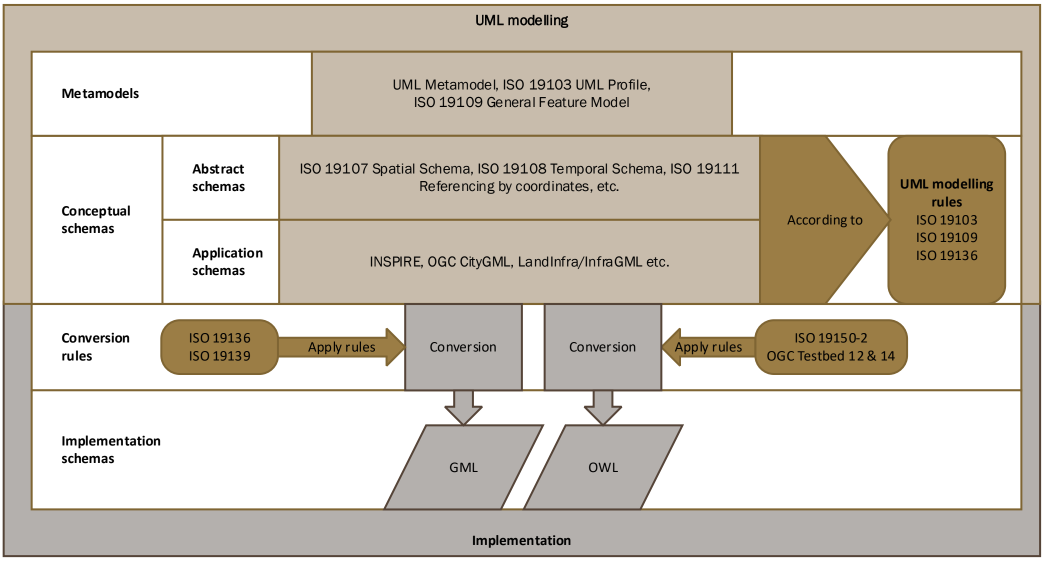


Figure 14. The levels of abstraction in MDA according to ISO 19103. From Article 3 (Jetlund et al., 2019a).

4.2 Article 1: A generic model for information exchange from GIS to ITS

Title:

Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model (Jetlund et al., 2019b)

Purpose:

Article 1 followed up on the findings from Conference proceeding 1 with further studies and an evaluation of solutions for modelling and exchanging geospatial information in the two application domains of GIS and ITS. Article 1 aimed to improve methods for such information exchange, which was done by developing, testing, and evaluating a prototype model.

Findings:

New vehicles are supplied with a range of sensors from which they create local geospatial knowledge and assist the drivers in different levels of automation. However, the local knowledge must be combined with pre-processed information covering larger areas for route planning and navigation under challenging conditions. Authoritative information needed for legal and safe navigation should be provided from road and mapping authorities' systems, based on GIS standards, to ITS databases for route planning and navigation, as illustrated by the horizontal arrows in Figure 15.

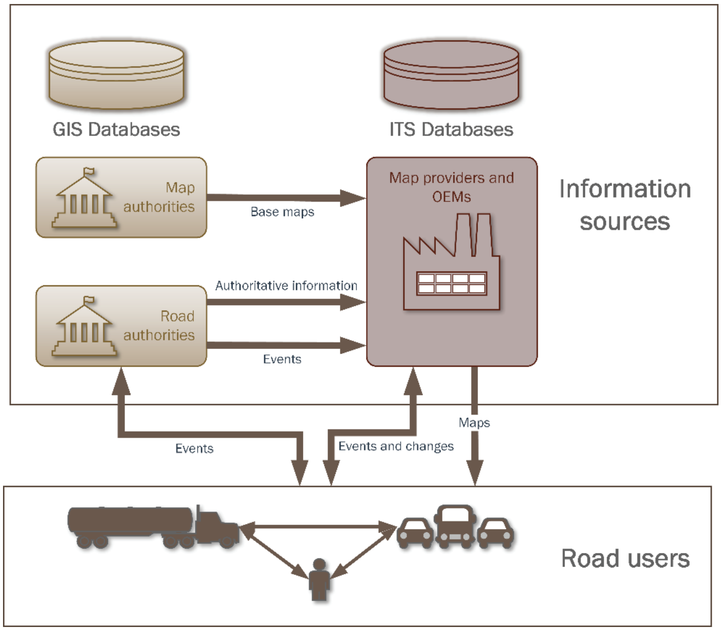


Figure 15. Parts of the flow of geospatial information in Intelligent Transport Systems (ITS). From Article 1 (Jetlund et al., 2019b).

The article identified ten candidate standards and specifications for the exchange of road-related geospatial information and evaluated them against six requirements specified to study two main topics: The usability of the solutions for exchanging information from GIS databases; and the flexibility for handling a continuously evolving real world. Table 2 shows a summary of the analysis, where the evaluation is generalized to a simple "yes" or "no" for each requirement.

Table 2. Evaluation of studied solutions. Req = Requirement. From Article 1 (Jetlund et al., 2019b).

Solution	**Req 1: ISO/TC 211 MDA**	**Req 2: GIS exchange format**	**Req 3: Feature catalogue**	**Req 4: Feature catalogue exchange model**	**Req 5: Network model**	**Req 6: Generic feature exchange model**
GDF	No	No	Yes	Yes	Yes	Yes
NDS Open Lane Model	No	No	Yes	No	Yes	No
OpenDRIVE	No	No	Yes	No	Yes	No
TN-ITS	Yes	Yes	No	No	No	Yes
OpenTNF	No	Yes	No	Yes	Yes	Yes
INSPIRE TN	Yes	Yes	Yes	No	Yes	No
CityGML	Yes	Yes	Yes	No	No	No
LandInfra/ InfraGML	Yes	Yes	Yes	No	No	No
DATEX II	No	No	Yes	No	No	No
TPEG2	No	No	Yes	No	No	No

The evaluation showed that none of the studied solutions met all requirements. Four solutions were considered promising candidates that could be the basis of a new and improved solution: GDF, INSPIRE-TN, TN-ITS and OpenTNF. A prototype for an improved solution was developed, where the goal was to fulfil all six requirements. The prototype's core was a generic feature model that could be used for exchanging any information according to a feature catalogue. The exchange model could be kept stable, while the feature catalogue's content could be maintained with more flexibility. Besides, the feature model could be used with different feature catalogues. Figure 16 shows the prototype's central concepts, while Figure 17 shows the generic Feature Model.

The prototype was tested and validated in a use case with two different feature catalogues (INSPIRE-TN and GDF) and was found to be a candidate solution for improved information exchange from GIS to ITS. The INSPIRE-TN model could be used directly in the prototype, while minor changes were needed for the GDF model.

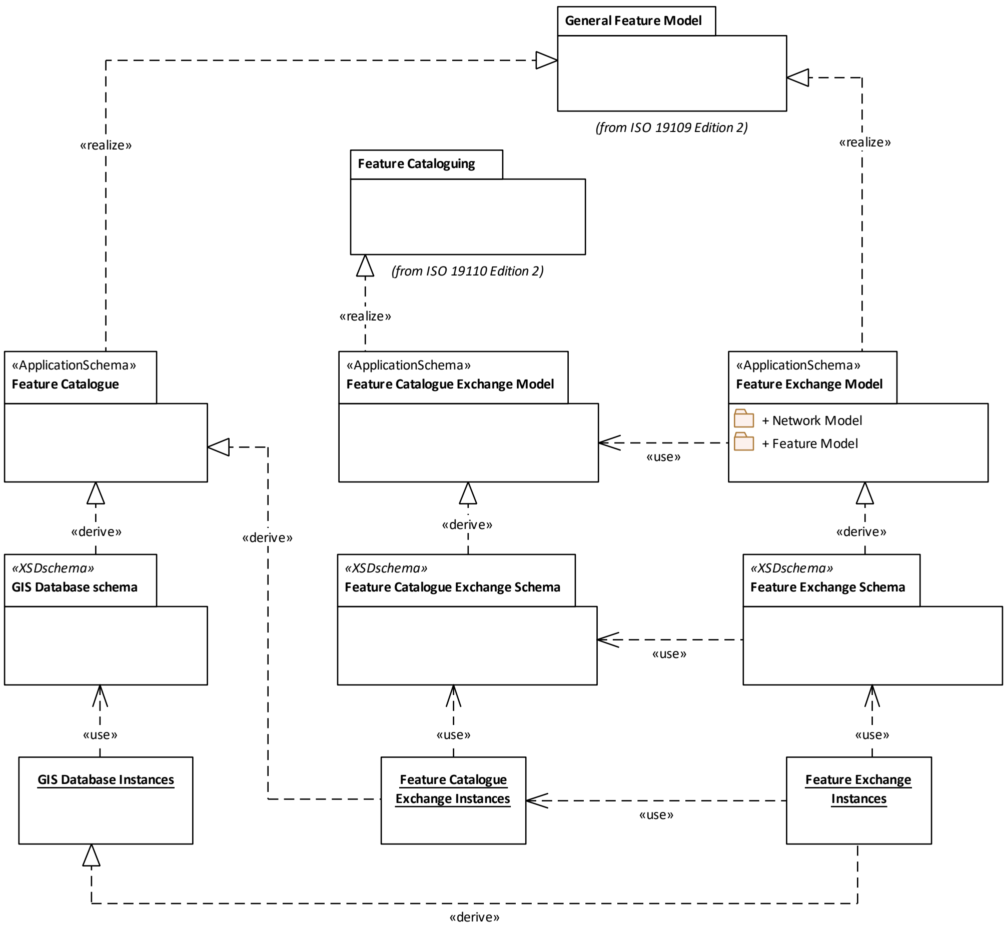


Figure 16. The concepts of the prototype. From Article 1 (Jetlund et al., 2019b).

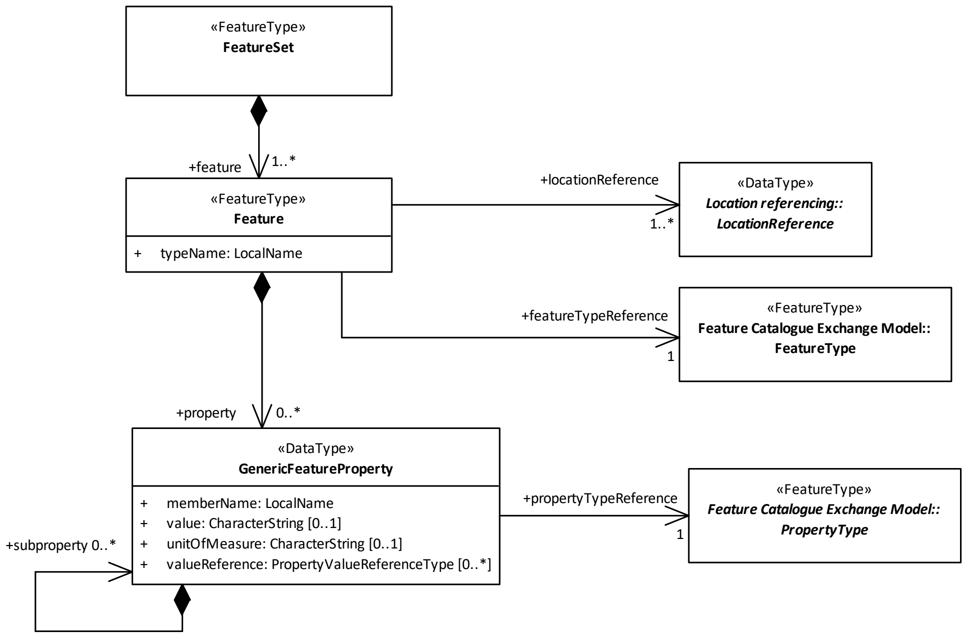


Figure 17. The Feature Model in the prototype. From Article 1 (Jetlund et al., 2019b).

4.3 Articles 2 and 3: Conversions from geospatial UML models to OWL

Titles:

Article 2: Improvements in automated derivation of OWL ontologies from geospatial UML models (Jetlund, 2018b)

Article 3: Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies (Jetlund et al., 2019a)

Purpose:

The purpose of Article 2 and 3 was to study and suggest improvements to the implementation of UML models of geospatial information in OWL, to enable sharing and linking of geospatial information in the Semantic Web. Article 2 studied transformation methods from UML to OWL and suggested improvements and further research. Article 3 followed up on the findings from Article 2 with more detailed studies of transformation methods and suggested methods for overcoming identified challenges.

Findings:

The articles found that vast amounts of geospatial information had been established according to information models based on ISO/TC 211 standards and made available on the World Wide Web through domain-specific web services for geospatial information. However, the amount of geospatial information available as RDF data according to OWL ontologies for the Semantic Web was still limited. Researchers had found that geospatial information and information models prepared for use in the Semantic Web could increase the value of the information. Discovery in Spatial Data Infrastructures (SDIs) could be improved, and, most important, spatial and non-spatial information could be linked and reused between stakeholders and domains.

Fundamental differences between UML and OWL were identified related to how the two languages approach information modelling. While OWL and the Semantic Web are based on the Open World Assumption (OWA), UML models are by nature following the Closed World Assumption (CWA). Furthermore, UML focuses on information structure, while OWL focuses on the meaning of concepts, described with set theory and description logic.

Despite the differences, the articles found that conversions are possible and had been described in standards and related research. ISO 19150-2 and reports from OGC testbeds defined rules for conversions according to the ISO/TC 211 MDA approach illustrated in Figure 14. Simple conversions between core concepts from the two languages were unanimously defined through the conversion rules, while some UML restrictive concepts were more challenging to describe in OWL. The resulting ontology's purpose would be fundamental for how conversion challenges should be handled.

Article 3 proposed solutions for addressing two specific conversion challenges in order to achieve better ontologies: Properties owned by individual classes in UML versus global properties in OWL, and reuse of external concepts. An extended UML profile and adapted conversion rules were described, and tests and demonstrations showed how the modified GDF information model from article 1 could be implemented in OWL based on the proposed solutions.

4.4 Article 4: IFC schemas as an ISO/TC 211 compliant UML model

Title:

IFC Schemas in ISO/TC 211 compliant UML for improved interoperability between BIM and GIS (Jetlund et al., 2020)

Purpose:

Article 4 aimed to improve the syntactic and semantic interoperability between information models for GIS and BIM. The syntactic interoperability was addressed by transforming the IFC information model, described in EXPRESS, into a prototype UML model according to ISO/TC 211 standards. The semantic interoperability was addressed by linking and mapping core concepts from IFC and ISO/TC 211 standards in the prototype.

Findings:

The article studied and compared the two modelling languages UML and EXPRESS, how they have been used for information modelling for GIS and BIM, and presented a pattern for conversion from IFC EXPRESS schemas to UML models according to ISO/TC 211 standards. The conversion pattern covered core conversions between EXPRESS and UML, as well as the specific use of EXPRESS for IFC and UML for GIS standards. The conversion pattern was applied on the IFC Road EXPRESS model and resulted in a prototype UML model which was tested for implementation by deriving implementation schemas for EXPRESS and GML.

The conversion pattern and the resulting prototype UML model were compared to the IFC-UML model from bSI, who plan to move from EXPRESS to UML as the original modelling language for

IFC. While the goal of the conversion pattern was to relate the IFC model to core ISO/TC 211 concepts, the IFC-UML model had defined more specific semantics for EXPRESS, independent of ISO/TC 211 standards. However, the conversion pattern could be adapted to handle conversion from IFC-UML as an alternative to EXPRESS. Figure 18 shows an example of a part of the converted IFC model.

The prototype UML model was used to relate core concepts in IFC and ISO/TC 211 standards to improve semantic interoperability. The IFC model's classes were defined as realizations of UML metaclasses from the General Feature Model defined in ISO 19109. Furthermore, basic datatypes for date, time and measure were linked between the IFC model and ISO/TC 211 standards. Linking geometry datatypes from IFC and ISO/TC 211 standards was a more complex issue. Many direct links could be defined, but several geometry types were defined on only one side of the comparison. Swept and constructive solids are defined only in IFC and were kept out of the analysis. The studies showed that harmonization and linking of core concepts were essential for the semantic interoperability.

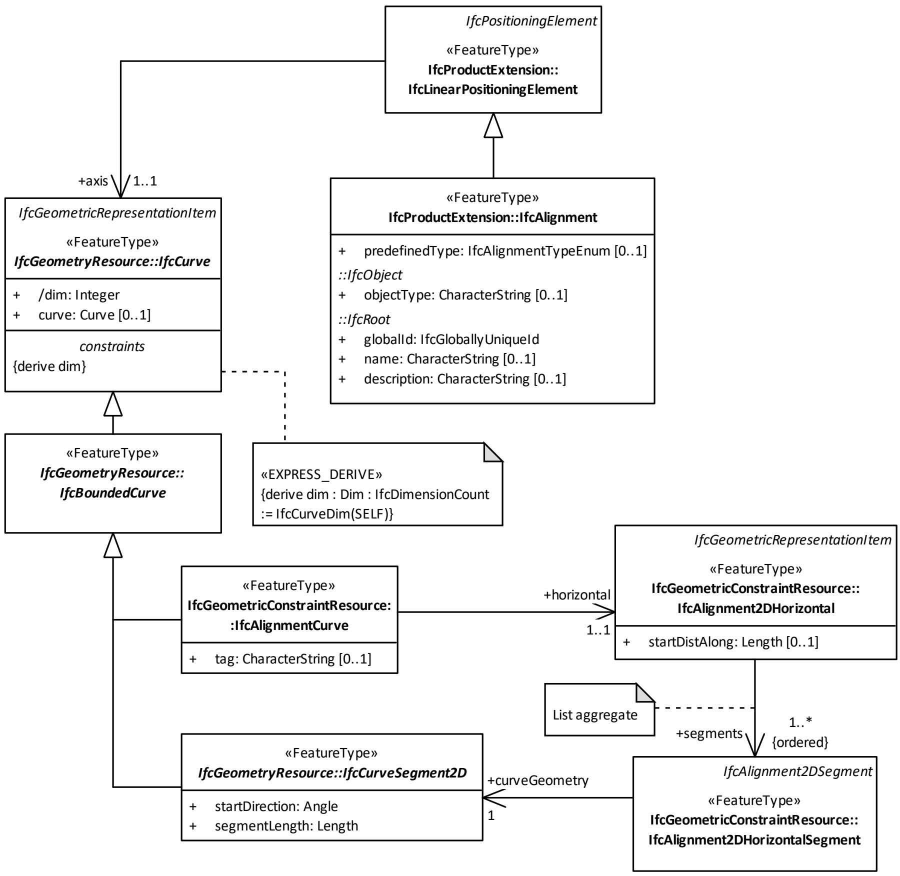


Figure 18. UML diagram showing a part of the converted IFC model. From Article 4 (Jetlund et al., 2020).

4.5 Article 5: A structure of UML profiles for GIS, ITS and BIM

Title:

A structure of UML profiles for modelling of geospatial information in GIS, ITS and BIM (Jetlund, 2020)

Purpose:

The purpose of Article 5 was to improve the syntactic and semantic interoperability between information models for GIS, ITS and BIM, which was addressed by establishing a joint framework for information modelling. Findings from articles 1 to 4, and related research, had shown that the harmonized approach could be based on MDA and UML.

Findings:

The article presented a prototype structure of UML profiles for GIS, ITS and BIM, developed from related research and findings from article 1 to 4. The structure included base and community profiles for PIMs and PSMs based on UML profiles and specifications from the ISO/TC 211 Standards ISO 19103, ISO 19109 and ISO 19136, and incorporated findings from articles 1 to 4. The structure is shown in Figure 19.

The basis of the prototype structure was the Core Geospatial Profile (CGP) that adapted and improved the UML profiles from ISO 19103 and ISO 19109. The general encoding profile for GML was based on ISO 19136, while the general encoding profile for OWL was based on the findings from article 3. Community specific conceptual and encoding profiles were suggested for IFC, GDF and DATEX II. The IFC EXPRESS encoding profile was based on findings from article 4.

The prototype was tested and demonstrated for implementation in the UML modelling software Enterprise Architect, which have been used for information modelling in all three application domains.

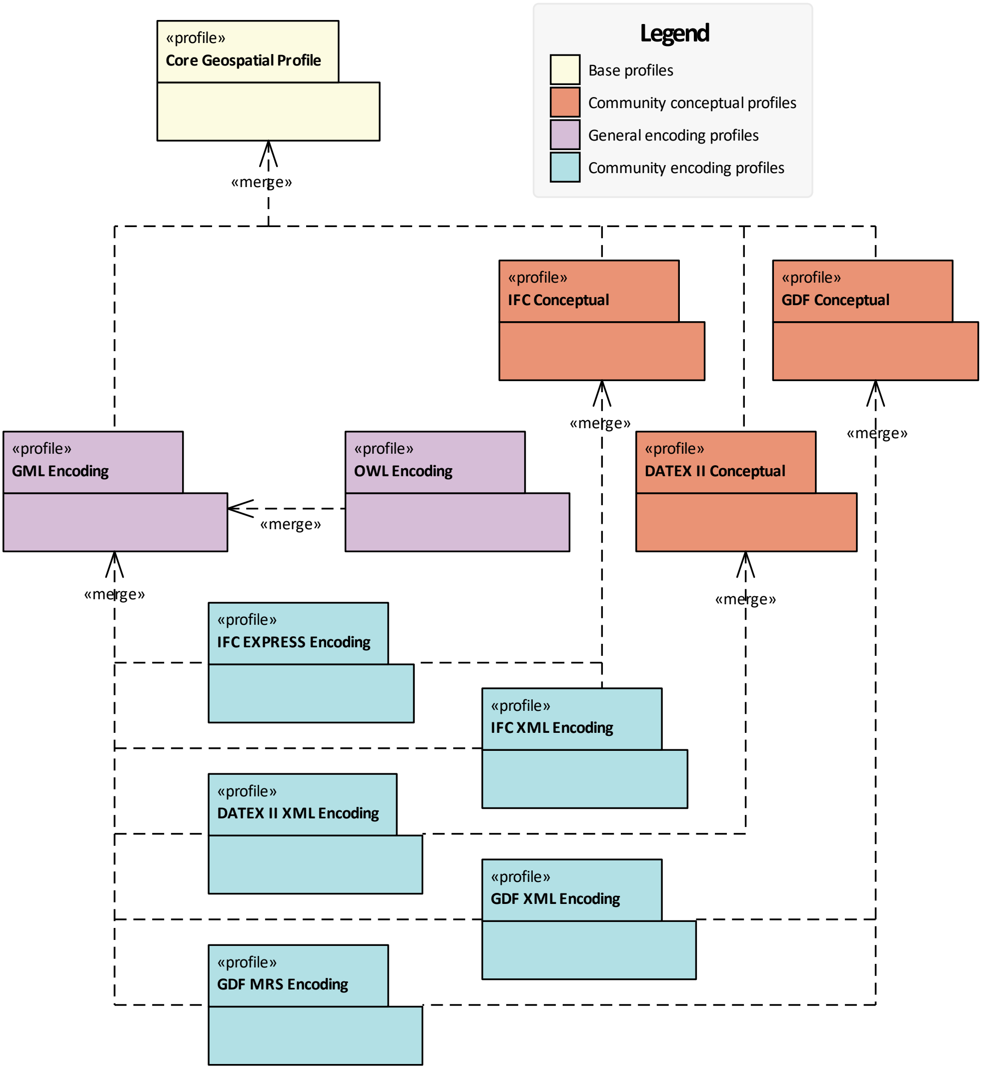


Figure 19. The structure of UML Profiles. From Article 5 (Jetlund, 2020).

5 RESULTS AND DISCUSSION

"Any fool can know. The point is to understand."
– Albert Einstein.

5.1 SRQ1: STATE OF THE ART FOR MODELLING APPROACHES AND TECHNOLOGIES

SRQ1: What approaches and technologies have been used for modelling geospatial information in GIS, ITS and BIM?

The first sub research question was addressed in all the appended articles, with a summary in Article 5. The studies showed that standards and specifications from all three application domains describe rules for modelling of geospatial road-related information with various approaches and technologies. Relevant information models have been developed by technical committees under ISO and CEN and by consortiums of industry stakeholders. Besides, related research has suggested models and solutions for interoperability. Table 3 summarizes the identified approaches and technologies from the findings in the appended articles.

Table 3. Studied specifications and technologies.

Domain	**Stakeholder**	**Standards or specifications**	**Technologies**	
			Information modelling	**Implementation**
GIS	ISO/TC 211	ISO 19101, ISO 19103, ISO 19109, ISO 19136, ISO 19150, ISO 19107, ISO 19108, ISO 19110, ISO 19111, ISO 19112, ISO 19148	ISO/TC 211 UML and MDA	GML OWL
GIS	OGC	CityGML, LandInfra/InfraGML	ISO/TC 211 UML and MDA	GML JSON
GIS	INSPIRE	Inspire Transport Networks	ISO/TC 211 UML and MDA	GML OWL
ITS	ISO/TC 204	GDF	UML and XML	XML MRS
ITS	ISO/TC 204	TPEG 2	TPEG 2 UML and MDA	XML Binary
ITS	CEN/TC 278	DATEX II	DATEX II UML and MDA	XML
ITS	CEN/TC 278	TN-ITS	ISO/TC 211 UML and MDA	GML
ITS	Navigation Data Standard Association	Navigation Data Standard	Database script	SpatiaLite
ITS	Triona et al.	OpenTNF	UML	GeoPackage
ITS	VIRES	OpenDRIVE	XML	XML
BIM	bSI ISO/TC 59 SC 13	IFC (ISO 16739)	EXPRESS and UML	XML STEP OWL

The studies showed that the modelling of geospatial information in the GIS application domain is based on a modular set of standards from ISO/TC 211, as illustrated in Figure 14. The standards describe a specific use of UML with profiles, rules for modelling, and an MDA approach with four levels of abstraction. Rules for conversions from UML models to implementation schemas have been defined for implementation in GML and OWL. Furthermore, core datatypes and concepts for geometry and location referencing are defined. The ISO/TC 211 set of standards describe the core concepts, while standards and specifications from other stakeholders have defined specific information models for implementation. Some relevant examples are the OGC specifications CityGML and LandInfra/InfraGML; the INSPIRE Transport Networks specification; and the ITS specification TN-ITS. Figure 20 shows a simplified view of the MDA approach defined in ISO/TC 211 standards.

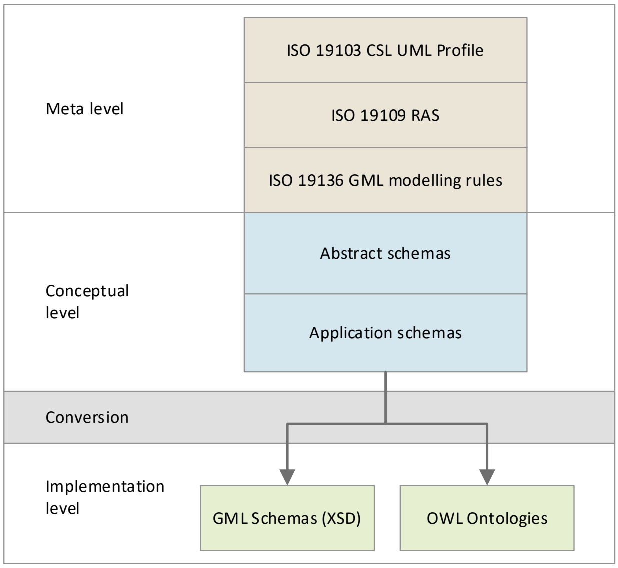


Figure 20. MDA for GIS as defined in ISO/TC 211 standards. From Article 5 (Jetlund, 2020).

The studies found no overall approach and rules for modelling of geospatial information in the ITS domain. Articles 2 and 5 showed that the standardization organizations ISO and CEN had applied UML and MDA in various ways for ITS standards, while industry standards for ITS have been developed in implementation-specific technologies such as XML schemas and database scripts.

The ISO/TC 204 standard GDF has been identified as a primary model of road-related geospatial information for ITS. Findings in articles 2 and 5 showed that GDF is described in UML and is partly but not fully compliant to ISO/TC 211 modelling rules. Joint work between ISO/TC 211 and ISO/TC 204 recommended developing a future version of GDF based on UML and MDA according to ISO/TC 211 standards. Two series of ITS standards for dynamic information have defined MDA approaches and rules for specific UML use similar to the ISO/TC 211 approach: The ISO TPEG2 series and the CEN DATEX II series. The MDA approaches in TPEG2 and DATEX II are illustrated in Figure 21.

Articles 4 and 5 found that core concepts for models of geospatial information in the BIM application domain have been defined in the IFC standard, developed by bSI and formalized by ISO/TC 59 SC 13. Unlike the modular ISO/TC 211 standards, the current IFC model contains all concepts in one single model, from core concepts to domain-specific classes. IFC was initially described in EXPRESS, but bSI is working towards the next generation of IFC where UML will be applied in an MDA approach similar to the ISO/TC 211 approach, and with a more modular structure. The planned MDA approach for IFC is illustrated in Figure 22.

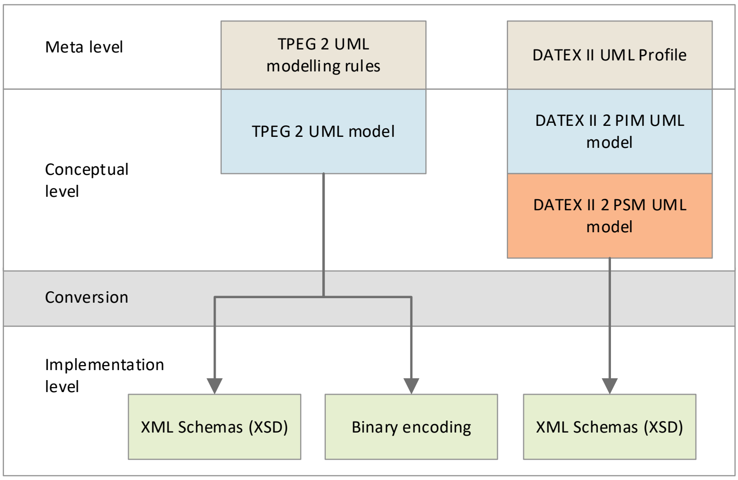


Figure 21. MDA in TPEG 2 and DATEX II. From Article 5 (Jetlund, 2020).

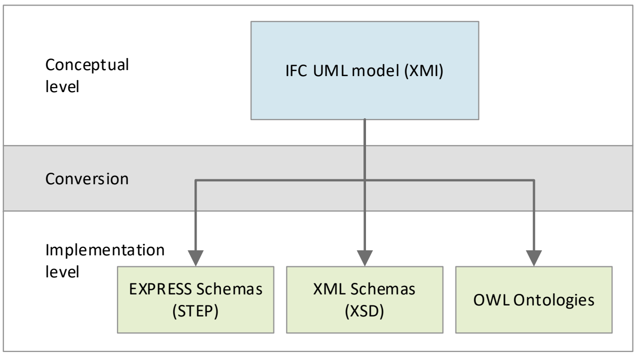


Figure 22. The planned MDA approach for IFC. From Article 5 (Jetlund, 2020).

Even though different approaches and technologies have been applied for modelling geospatial information in the three application domains, the findings in the appended articles indicated that there are similarities and possibilities for harmonization and linking. Primarily, the core models in all three application domains have been described in UML or similar conceptual schema languages. Transformations between the modelling approaches and conceptual languages were shown in articles 1, 4 and 5 and could improve the syntactic interoperability.

The information modelling approach in GIS was identified as the most structured and complete approach for modelling geospatial information. The defined UML profiles, specified levels of abstraction, modelling and conversion rules and the modular structure are advantages compared to approaches in the two other application domains. Articles 1 and 4 showed that the concepts from ISO/TC 211 standards could be used for information models in ITS and BIM. However, there are deficiencies in the ISO/TC 211 approach too. Therefore, Article 5 suggested modifications of the ISO/TC 211 UML profiles in order to be consistent with the main UML metamodel and specification, while Articles 3, 4 and 5 suggested modifications of the UML profiles in order to improve implementation in OWL and EXPRESS.

Related research on modelling of geospatial information in the three application domains indicated that there had been little research on interoperability between the core concepts from distinct information models. Geospatial information for ITS had been studied for specific data sets and project models within either the GIS or ITS application domain, while less research was found on integration between the two application domains. Among the exceptions were European studies of the provision of road information from authorities to ITS map providers, which focused on TN-ITS and INSPIRE Transport Networks as potential solutions. For the interoperability between GIS and BIM, most of the identified studies focused on integration between IFC and OGC specifications, and not on the core concepts. Furthermore, most studies focused on BIM for buildings, and only a few included the IFC extensions for infrastructure. The findings indicated that more research on the integration of core concepts from distinct information models could contribute to improved interoperability between GIS, ITS and BIM.

5.2 SRQ2: Information exchange from GIS to ITS

SRQ2: How can models for exchange of geospatial information from road and mapping authorities to geospatial databases for ITS be improved?

The second sub research question was addressed in Conference proceeding 1, articles 1 and 5, and in Standardization project 3. Article 1 found that base maps and authoritative information should be provided from GIS databases maintained by authorities to ITS databases for route planning and navigation, as illustrated in Figure 15. In order to achieve an efficient exchange from GIS databases, information models should be based on ISO/TC 211 modelling approaches, and exchange formats that are well-known in the GIS application domain should be used. Furthermore, as the real world is continuously evolving, and different stakeholders may provide information based on different feature catalogues, the exchange models should have a high degree of flexibility.

The main answer to the sub research question was given in Article 1 by six requirements for an efficient exchange model and a prototype for an improved model which fulfilled the requirements. Article 5 described how the improved solution could be implemented in a framework of UML Profiles for geospatial information. Ten existing solutions were evaluated in Article 1. While none of the studied solutions met all requirements, concepts from four solutions were used as the fundament for the prototype. Table 4 lists the requirements and how they were addressed in the prototype, while Figure 16 and Figure 17 shows the structure of the prototype.

The prototype was tested and validated in Article 1, in a use case where road or mapping authorities wished to set up one single service that could provide road-related geospatial information for different user groups in the ITS domain. In the use case, the information should be provided according to the information models in GDF and INSPIRE Transport Networks. The case study results showed that the prototype could be used with different feature catalogues modelled according to ISO 19109.

The INSPIRE Transport Networks model could be used directly in the prototype, while only minor modifications were needed for the GDF information model to make it compliant with ISO 19109.

Table 4. Requirements and how they were addressed in the prototype.

Requirement	**Prototype approach**
Req 1: ISO/TC 211 MDA Solutions for exchanging information from GIS databases should be based on the ISO/TC 211 MDA approach to avoid complex transformations.	The application schemas in the prototype were based on MDA according to ISO/TC 211 standards, mainly the General Feature Model from ISO 19109 and the abstract conceptual feature catalogue model from ISO 19110. Implementation schemas were derived from the application schemas.
Req 2: GIS exchange format: An exchange format from the GIS domain such as GML or GeoPackage is essential for avoiding additional conversions.	The prototype used the GML exchange format. With the ISO/TC 211 MDA approach, implementation schemas for other formats such as GeoPackage may also be derived.
Req 3: Feature catalogue: Unambiguous descriptions of real-world features and properties in a feature catalogue are essential for a common understanding of the exchanged information.	The prototype did not contain a feature catalogue of its own but could implement any feature catalogue modelled as an application schema according to ISO 19109. Selected feature types from the feature catalogues from INSPIRE and GDF were implemented in the case study.
Req 4: Feature catalogue exchange model: A feature catalogue exchange model is needed for sharing the classifications of the real world described in the feature catalogue.	The prototype contained a Feature Catalogue Exchange Model and a derived Feature Catalogue Exchange Schema for exchange in GML
Req 5: Network model: A navigable digital network is the foundation of route planning and navigation and must include geometry and topology of the network elements, and mechanisms for relating features to the network.	The Feature Exchange Model in the prototype contained a Network Model based on the INSPIRE model.
Req 6: Generic feature exchange model: Combining a feature catalogue exchange model and a generic feature exchange model gives the advantage of keeping the feature exchange model stable, while the feature catalogue can be modified. Besides, the feature exchange model may also be used with different feature catalogues.	The Feature Exchange Model in the prototype contained the Feature Model, which was a generic feature model based on the TN-ITS model. An implementation schema for exchanging network and features in GML was derived from the model.

A comparison between the prototype and existing solutions showed that the prototype covered and improved concepts for exchanging feature catalogues, road features and simple network topology. The prototype was considered a candidate for improved information exchange of authoritative information from road authorities’ GIS databases to ITS databases. It was also suggested as input to a future GDF version based on ISO/TC 211 standards in Standardization project 3.

However, several existing solutions had more complex road network models, with distinct levels of topology and information for lane-level navigation. These issues were identified as possible future improvements of the prototype. Besides, an even more complex representation of road network and restrictions with area and volume geometries in addition to links will be fundamental for automated driving. Models are under development in the ITS application domain with the belt concept in part 2 of GDF (ISO/TC 204, 2020), and in the GIS application domain with traffic spaces in CityGML version 3.0 (Kutzner et al., 2020, Beil et al., 2020). A future solution for exchanging authoritative information will need to handle these complex geometric representations of the road network and restrictions. In order to support such information exchange, the authorities' databases will also need to maintain authoritative information according to more complex models, for example, based on CityGML version 3.0.

5.3 SRQ3: Integration of models for BIM and GIS

SRQ3: How can information models and semantics for implementation technologies for BIM be integrated and linked with GIS standards?

The third sub research question was addressed in articles 4 and 5 and discussed in Standardization project 4. Article 4 found IFC, which has been modelled in EXPRESS, to be the core information model for BIM. Contrary, ISO/TC 211 standards based on UML were found to be the foundation for GIS information models. The results from sub research question 1 indicated that the modelling approach applied for GIS is more structured and complete than the BIM approach. Therefore, sub research question 3 was answered in Article 4 by developing and testing a pattern for converting IFC from EXPRESS to UML according to ISO/TC 211 standards and linking core concepts in IFC to concepts from ISO/TC 211 standards.

The results from Article 4 showed that the conversion pattern maintained most of the EXPRESS concepts that have been used in IFC, but some concepts needed a more complex and controlled conversion. Some additional semantics were needed in the UML model to maintain EXPRESS structure. Article 5 described how the semantics could be included in an IFC EXPRESS Encoding UML Profile in a framework of UML Profiles for geospatial information. A comparison between the converted model and the IFC-UML model developed by bSI for future versions of IFC showed that the conversion pattern could be adapted to handle conversions from IFC-UML as an alternative to EXPRESS. However, a better solution for syntactic and semantic interoperability would be to use UML according to ISO/TC 211 standards for the original IFC model.

The conversion pattern concerned the syntactic interoperability, while the semantic interoperability was addressed by linking core concepts from IFC to ISO/TC 211 standards. The results in Article 4 showed that primitive datatypes and core datatypes for date, time and measures defined in the IFC model could be linked to equivalent datatypes from ISO/TC 211 standards. However, better semantic interoperability would be possible if models in both application domains reused basic datatypes from more general models and ontologies instead of domain-specific models for GIS and BIM. Better semantic interoperability for geometry and location referencing datatypes would be possible if original datatypes were defined in one of the models and reused in both, as well as in other models of geospatial information. Unlike the more general datatypes, datatypes for geometry and location referencing could be owned originally by a harmonized model with core concepts for information modelling in GIS and BIM.

The prototype UML model was tested for implementation by deriving implementation schemas for EXPRESS and GML. The results in Article 4 showed that EXPRESS implementation was possible through a reverse conversion pattern, supported by the extra semantics defined in the EXPRESS Encoding Profile. The implementation in GML was possible by applying standard conversions according to ISO 19136, supported by the established links between core concepts. The links between core concepts, and in particular the mapping of geometry and location referencing data types, were vital for the GML implementation. Incomplete harmonization and linking of data types proved to be a challenge for implementing the IFC model in GML.

5.4 SRQ4: A JOINT MODELLING APPROACH

SRQ4: How can information models and semantics for implementation technologies for GIS, ITS and BIM be integrated into a joint modelling approach?

The fourth sub research question was addressed articles 1, 3, 4 and 5, with the main contribution in Article 5. The results from sub research question 1 indicated that the modelling approach applied for GIS is more structured and complete than ITS and BIM approaches. Furthermore, the results from Article 1 indicated that the GDF standard for ITS could be modelled according to the ISO/TC 211 MDA approach, while Article 4 showed that the IFC model for BIM could be transformed to a UML model based on ISO/TC 211 MDA. Therefore, Article 5 suggested that the answer to sub research question 4 could be a structure of UML profiles founded on ISO/TC 211 MDA, as illustrated in Figure 19. A list of recommended actions for formalizing the structure was described, including revising existing profiles and developing new formal profiles. The recommendations are listed in Table 5.

Table 5. Recommended actions for formalizing UML profiles in the structure. From Article 5 (Jetlund, 2020).

UML Profile	Recommended actions
CGP	Revise the UML profiles in ISO 19103 and ISO 19109.
GML Encoding Profile	Define a formal UML profile in ISO 19136.
OWL Encoding Profile	Define a formal UML profile in ISO 19150-2.
IFC Conceptual Profile	No actions are needed; the CGP can be used.
IFC EXPRESS Encoding Profile	Define a formal profile for encoding in EXPRESS.
GDF Conceptual Profile	No actions are needed; the CGP can be used.
GDF Encoding Profiles	Define formal profiles for encoding in GDF XML and MRS.
DATEX II Conceptual and Encoding Profiles	Define a two-way mapping between concepts in the CGP and the GML Encoding Profile.
TPEG 2 Conceptual and Encoding Profiles	Define formal profiles for the conceptual model and the encodings, based on rules in ISO 21219.
Transmodel and NeTEx	Define a formal profile for specific concepts from UML use in existing models.

The results from Article 5, supported by articles 1, 3 and 4, indicated that the suggested structure could be implemented in UML modelling software and applied for information modelling in the three application domains. The approach for adapting existing information models into the structure would depend on the structure of the original model. Transformations are always concerned with the risk of losing information or expressiveness. Therefore, model adaption by adding more semantics could be a better approach than model transformation.

Integration of information models from all three application domains into one joint modelling approach, as suggested in Article 5, would contribute to a shared understanding of how the real world is described in information models. Core concepts could be defined in joint abstract models and reused in community-specific models, as described in articles 1 and 4. As a result, the information models would become more harmonized, and syntactic and semantic interoperability would be improved. However, a full harmonization would not be possible and not appropriate, given the distinct roles of GIS, ITS and BIM in the digital geospatial environment. An integration across application domain would still require a linking and mapping between domain-specific concepts. Still, the quality and accuracy of links depend upon how much concepts from the distinct models differ on semantics. Harmonized models that use the same core concepts are better prepared for linking than heterogeneous models with internal and specific core concepts. Therefore, harmonizing the modelling approaches and core concepts is essential for enabling linking and mapping information models.

5.5 SRQ5: Implementation as OWL ontologies

SRQ5: How can UML models of geospatial information be implemented as OWL Ontologies, for linking and mapping by applying Semantic Web technologies?

The fifth sub research question was addressed in Articles 2, 3, 4 and 5, where article 2 and 3 gave the main contribution. Article 4 found that Semantic Web technologies for linking and mapping had been considered a promising approach for improved interoperability between heterogeneous information models. For information models initially described in UML, the information models would need to be converted into OWL ontologies. Furthermore, the datasets described by the information models would need to be converted to graphs of RDF triples according to the converted ontologies.

Articles 2 and 3 found that conversions from UML according to ISO/TC 211 standards to OWL and export from GIS datasets to RDF were possible, despite the fundamental differences between UML and OWL. Challenges were identified for the conversion of some restrictions from the closed UML world to the open OWL world. Standards and related research had described various conversions that maintained the restrictions to various degrees. Besides, the conversion from UML to OWL resulted in ontologies that described the real world according to a UML structure, while ontologies developed in OWL would be structured differently, focusing on the meaning of concepts. Finally, reuse of existing external concepts is a good practice for ontology development in order to enable interoperability. The findings in articles 2 and 3 showed that UML models had often defined internal concepts instead of reusing equivalent concepts from existing ontologies. A consequence of the issues mentioned above was that, although the resulting ontologies were syntactically and structurally correct, they were not optimized for use in the Semantic Web.

To improve the conversion, a prototype for an extended UML profile and adapted conversion rules for encoding in OWL was developed, tested and demonstrated in Article 3. The results indicated that the conversions described in existing methods could be improved by applying the suggested UML profile and adapted conversion rules. Two conversion issues were primarily handled by additional semantics in the UML models: UML properties owned by individual classes could be converted to global OWL properties or properties shared by several classes, and concepts from the UML models could be linked to concepts defined in other ontologies.

Articles 2 and 3 found the purpose of the ontology to be essential for whether CWA restrictions from UML should be maintained in OWL, to what degree the structure from the UML models should be strictly maintained, and to what degree new ontologies should be designed manually in OWL. The articles identified three levels of information flow and defined different rules for maintaining restrictions and structure depending on the level:

1. Ontologies for use only in the Semantic Web, in which case restrictions and the original structure from the UML models can be disbanded. The ontologies should be based on the OWA as far as possible and existing concepts from other ontologies should be used when appropriate. Furthermore, OWL principles for describing the real world should be followed rather than UML principles. In many cases, an ontology developed in OWL would be preferred instead of conversion from UML.
2. Unidirectional information exchange to the Semantic Web, where the original geospatial information is maintained in datasets according to UML models and converted to RDF according to OWL ontologies for use in the Semantic Web. Like for the first level of information flow, few restrictions from the UML model are needed in the ontologies, as they will primarily be used for describing the information. Likewise, some externally defined concepts can replace internal concepts, as strict maintenance of the structure is not vital. Therefore, conversions from UML to OWL do not need to maintain all restrictions and can include mapping to external concepts.
3. Bidirectional information exchange, where information can be registered, maintained and developed in both worlds, and exchanged in both directions. For this purpose, the OWL ontologies and the original UML model should be as identical as possible. The information must be valid according to the CWA-based UML models to make sure it can be imported into repositories defined by the UML models. Therefore, conversions need to maintain restrictions and the original structure as far as possible.

Model-driven conversions from UML to OWL need to consider the different levels of information flow and apply conversions appropriate for the selected level. Therefore, conversion rules and software implementing the rules need to be configurable.

Although the most optimized ontologies for use in the Semantic Web might be achieved by development in OWL instead of conversions from UML, it is not necessarily the best solution for ontologies of geospatial information. A decoupling from existing applications and information models, as described in level 1 above, is not a likely situation in the short term. Specialized applications from the individual application domains have complex and advanced functionality which cannot straightforwardly be based on OWL and RDF. Especially, describing and handling complex geometry has been identified as a challenge in OWL and RDF.

The most discussed purpose found in related research was a unidirectional exchange, as described in level 2. For such use, it has also been suggested to retain complex structures and geometries in geospatial databases made available by web services for geospatial information (WFS) and apply links between the simplified RDF dataset and the original web service. This approach is equivalent to what is known as polyglot persistence in database research.

Ontologies based on all three levels of information flow can be used for linking and mapping by applying linksets. However, if the ontology aims to enable import from RDF into a repository based on a UML model, the ontology needs to be according to level 3. Figure 23 illustrates a process for mapping by linkset, and how ontologies according to the different levels of information flow can be applied.

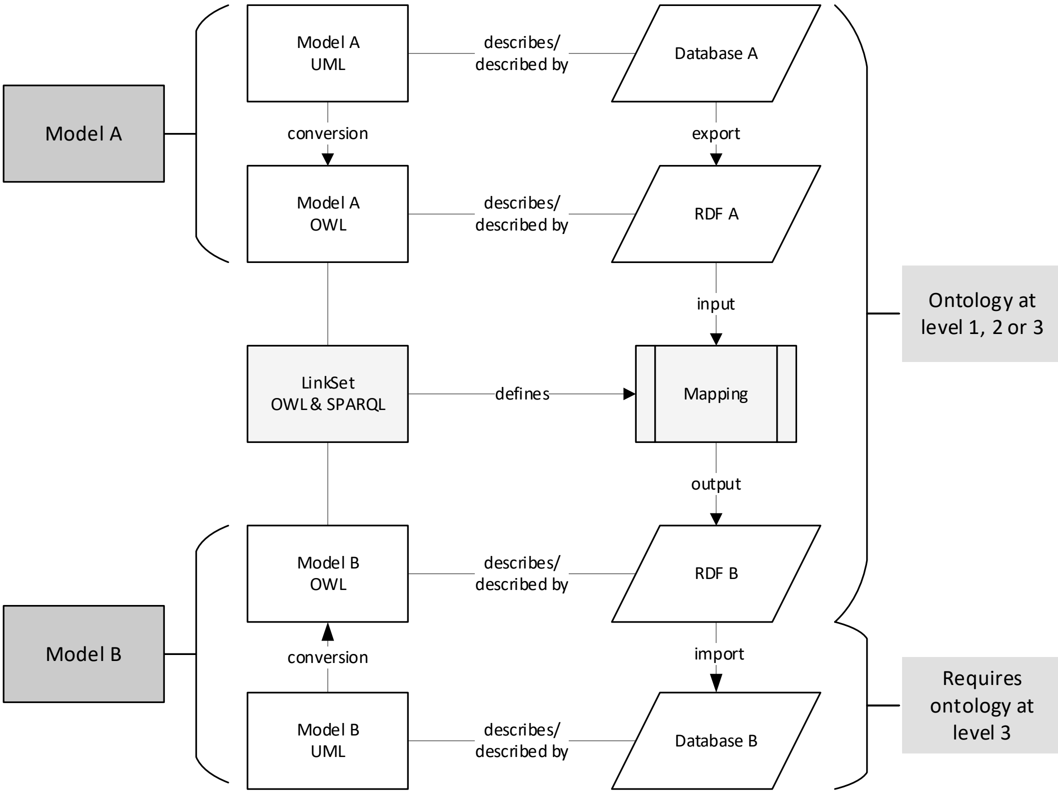


Figure 23. Mapping by linkset.

6 CONCLUSIONS AND FURTHER WORK

"The first law of geography: everything is related to everything else, but near things are more related than distant things."
– Waldo R. Tobler.

6.1 THEORETICAL AND PRACTICAL CONTRIBUTIONS

The main research question of this thesis, which was answered through a set of sub research questions, was**: How can approaches and technologies for information modelling be applied for harmonization and linking of conceptual models of geospatial information from the three application domains of GIS, ITS and BIM**?

The most important outcome of the research presented in this thesis is the development and testing of a harmonized approach for modelling of geospatial information in GIS, ITS and BIM, founded on MDA and a structure of UML profiles. Improved syntactic and semantic interoperability between GIS, ITS and BIM could be achieved by describing models from the three application domains in the same conceptual schema language, based on a joint modelling approach and with common core concepts.

The state-of-the-art analysis showed that all three application domains had applied UML and some degree of MDA for information modelling. Furthermore, the results indicated that the modelling approach described and applied in ISO/TC 211 standards in the GIS domain was the most complete and structured approach for modelling geospatial information. The formalized UML profiles, specific rules for UML modelling, defined levels of abstraction, modularization and conversion rules for implementation are primary advantages of the GIS approach. The approach has been widely adopted for modelling geospatial information world-wide in the GIS application domain and other domains. The research showed that the existing core information models IFC and GDF from the BIM and ITS application domains could be transformed from their current modelling approaches to be compliant with a modified GIS approach. However, deficiencies were identified for the GIS approach as well. Therefore, improvements were suggested, and the harmonized approach was based on a modification of the MDA approach described in ISO/TC 211 standards.

Investigating and reusing existing concepts is an essential foundation for semantic interoperability and is recommended as one of the initial steps of information modelling. In this thesis, such reuse of core concepts across application domain borders was made possible by following the harmonized modelling approach, where core concepts were defined at a joint abstract conceptual level and reused in domain-specific models. Information models from the three application domains have defined internal primitive datatypes and fundamental datatypes for general concepts such as date, time and measure. These should instead be reused from more generic and not domain-specific vocabularies. For concepts used for describing geometry and location referencing, the distinct concepts described in current models should be harmonized into one model and reused in all three application domains. The research in this thesis showed that many concepts from the BIM model IFC could be harmonized and linked with concepts from ISO/TC 211 standards, but that a full harmonization could not be achieved without significant changes in fundamental concepts.

While a full harmonization would give complete interoperability between models of geospatial information for GIS, ITS and BIM, it would not be the best solution for optimized use in each domain. The three application domains have different roles in the digital geospatial environment and need to describe features from the real world in different contexts, with different representations. For example, a physical guardrail along a road may be represented in GIS as a feature with a simple line geometry, supported by height information and other characteristics needed for, e.g., planning snow removal and calculating noise pollution. In BIM, the same guardrail would be a complex feature, consisting of several parts that are connected internally and to other features, with volume geometries and product characteristics needed for, e.g., foundation and maintenance. In ITS, the guardrail could be represented as a wall along the road, with information describing the possibilities for hindering an unexpected maneuver or blocking an emergency maneuver. Therefore, a full harmonization would not be appropriate, and further improvements of the semantic interoperability must be achieved by linking and mapping heterogeneous information models.

Semantic Web technologies with linksets described as OWL ontologies and mapping rules in SPARQL queries have been considered a promising approach for linking and mapping heterogeneous information models. However, they require that the information models are made available as OWL ontologies. The research in this thesis showed that UML models based on the harmonized modelling approach could be converted into OWL ontologies by applying conversion rules. However, several deficiencies were identified for existing conversion rules, especially related to the different modelling approaches for UML and OWL. The resulting ontologies could be improved by introducing an extended UML profile as part of the harmonized modelling approach, supported by improved conversion rules. The converted ontologies were still not fully optimized for use in the Semantic Web, but they were fit for linking and mapping to achieve better semantic interoperability. Generally, any conversion from UML to OWL must be configured to fit the ontology's purpose and maintain restrictions and structure according to that purpose.

Given the potential of linking and mapping with Semantic Web technologies, a possible solution for improved syntactic and semantic interoperability could also be to adopt OWL ontologies as the original modelling approach for all three application domains. OWL ontologies as the original information models would make conversions from UML to OWL superfluous, enable more reuse of external concepts directly in original models, and enable optimized ontologies for use in the Semantic Web. However, the research for this thesis indicated that OWL and RDF lack some functionality for modelling and handling complex geometries. Furthermore, applications for GIS, ITS, and BIM have advanced functionality that depends more upon the structure of information as described in UML and EXPRESS than on the meaning of concepts as focused on in OWL. The standardized graphical representations in UML and EXPRESS have has also been pointed out by other researchers as an advantage, especially for UML. Finally, the research showed that large amounts of geospatial information have been established and maintained based on existing UML and EXPRESS models. For these models and the information based upon them, conversions to OWL and RDF would be a more natural step towards the Semantic Web than modelling directly in OWL. Still, OWL ontologies optimized for use in the Semantic Web may replace the closed worlds of UML and EXPRESS in the longer term.

The combination of harmonization and linking presented in this thesis has considerable potential for improving the syntactic and semantic interoperability between models of geospatial information from the three application domains of GIS, ITS and BIM. The studied information models cannot be wholly harmonized, but they should be as harmonized as possible. Unnecessary transformations should be avoided, as any transformation leads to a loss of information. Models that are based on the same modelling approach and use the same core concepts are more similar, and fewer transformations are needed. Furthermore, the links needed for transformation can be defined more precisely between models based on a common information modelling approach. Therefore, harmonizing the modelling approaches and core concepts is essential for improved semantic interoperability by linking and mapping information models.

Tobler (Tobler, 1970) defined his first law of geography as "everything is related to everything else, but near things are more related than distant things". In a digital world with multiple representations of real-world things, Tobler's first law can be modernized to say that "every digital representation of a thing can be linked to any other digital representation of the same thing, but links between similar representations are more precise than links between dissimilar representations".

6.2 LIMITATIONS AND FURTHER RESEARCH

The exploratory research on information modelling presented in this thesis was conducted at a core conceptual level. Consequently, the research has limitations concerning real implementations in the three application domains. Therefore, future research should include experiments on real implementations of data sets according to information models founded on the harmonized approach, and information exchange across domain borders.

The research has also revealed several opportunities for further research based on the findings in the thesis, as described below.

The research showed how a selection of concepts could be harmonized or linked across application-domain borders. Primitive and core datatypes were harmonized and linked between IFC and ISO/TC 211 standards. Further studies should investigate how core datatypes from less application-specific models can be used for information models in all three application domains according to the harmonized approach. Furthermore, shared models for geometry and location referencing concepts are crucial for improved interoperability of geospatial information. Further studies on this topic should focus on harmonization of concepts and development of one standard model for geometry and location referencing for use in the three application domains, as well as other domains. Identified challenges from the research presented in this thesis should be included, such as complex volume geometries representing constructive solids in BIM and traffic spaces in ITS, and how linear located information can be combined with these concepts.

The optimization and use of OWL ontologies of geospatial information for different purposes also needs to be studied further. The research presented in this thesis showed how UML models based on the harmonized approach could be transformed into OWL ontologies, but also that there are remaining challenges for the conversions. On the one hand, the challenges concern how to maintain UML restrictions in OWL. While on the other hand, the question is how to optimize information models of geospatial information for use in both the Semantic Web and traditional applications.

The systems that are supposed to interoperate consist of information models and applications and organizations that need to understand the other systems' context and dynamics. Other levels of conceptual interoperability are needed for applications and organizations to achieve actual interoperability of real implementations. Mainly, the pragmatic and dynamic levels, as described by, e.g., Axelsson (Axelsson, 2020), are vital and need to be investigated further with a basis in the syntactic and semantic interoperability described in this thesis.

Finally, the three application domains of GIS, ITS and BIM are only three of several application domains in an ecosystem of digital twins where spatial and non-spatial data are brought together. Other stakeholders such as Life-Cycle Analysis (LCA), Smart Cities and other transportation modes than road transportation play vital roles too. Different stakeholders' roles in such integrated digital environments are essential for pragmatic and dynamic interoperability. Therefore, the roles of the three application domains and other domains in the integrated digital environment need to be investigated further.

7 REFERENCES

"If I have seen further, it is by standing upon the shoulders of giants."
– Sir Isaac Newton.

Ackoff, RL, 1989. From data to wisdom. *Journal of applied systems analysis*, 16 (1):3-9.

Allemang, D, and Hendler, JA, 2011. *Semantic web for the working ontologist : effective modeling in RDFS and OWL*. 2nd ed. ed. Burlington, Mass.: Elsevier.

Appel, K, Aakre, A, and Kronborg, P, 2018. ITS TERMINOLOGY. Oy Nord Print Ab.

Axelsson, J, 2020. Achieving System-of-Systems Interoperability Levels Using Linked Data and Ontologies. *INCOSE International Symposium*, 20-22 July 2022.

Beetz, J, van Berlo, L, Borrman, nA, et al., 2020. Enabling an Ecosystem of Digital Twins. A buildingSMART International Positioning Paper. https://www.buildingsmart.org/wp-content/uploads/2020/05/Enabling-Digital-Twins-Positioning-Paper-Final.pdf.

Beil, C, Ruhdorfer, R, Coduro, T, et al., 2020. Detailed Streetspace Modelling for Multiple Applications: Discussions on the Proposed CityGML 3.0 Transportation Model. *ISPRS International Journal of Geo-Information*, 9 (10):603.

Bellinger, G, Castro, D, and Mills, A, 2004. Data, information, knowledge, and wisdom. https://www.systems-thinking.org/dikw/dikw.htm (10 January 2021).

Boje, C, Guerriero, A, Kubicki, S, et al., 2020. Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, 114:103179, doi: 10.1016/j.autcon.2020.103179.

Borzacchiello, MT, Boguslawski, R, and Pignatelli, F, 2016. Improving accuracy in road safety data exchange for navigation systems: European Union Location Framework Transportation Pilot. http://publications.jrc.ec.europa.eu/repository/handle/JRC104569.

Bowen, GA, 2009. Document analysis as a qualitative research method. *Qualitative research journal*, 9 (2):27.

buildingSMART International, 2019a. IFC Specifications Database. buildingSMART International, London, UK. https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/ (10. January 2021).

buildingSMART International, 2019b. Infrastructure Room. buildingSMART International, London, UK. https://www.buildingsmart.org/standards/rooms/infrastructure/ (10. January 2021).

buildingSMART International, 2019c. NextGen-IFC. buildingSMART International, London, UK. https://github.com/buildingSMART/NextGen-IFC (10. January 2021).

CEN/TC 278, 2018a. CEN/TS 17268:2018 Intelligent transport systems — ITS spatial data — Data exchange on changes in road attributes. CEN: Brussels, Belgium.

CEN/TC 278, 2018b. Intelligent transport systems — Location Referencing Harmonisation for Urban-ITS — Part 2: Translation methods. CEN: Brussels, Belgium.

CEN/TC 278, 2018c. Intelligent transport systems — Location Referencing Harmonisation for Urban ITS — State of the art and guidelines. CEN: Brussels, Belgium.

Chen, S, Tan, J, Claramunt, C, et al., 2011. Multi-scale and multi-modal GIS-T data model. *Journal of Transport Geography*, 19 (1):147-161.

Coetzee, S, Plews, R, Brodeur, J, et al., 2019. Standards - Making Geographic Information Discoverable, Accessible and Usable for Modern Cartography. In *Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management*, edited by Döllner, Jobst and Schmitz, 325-344. Cham: Springer International Publishing.

Deng, Y, Cheng, JC, and Anumba, C, 2016. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, 67:1-21.

Donkers, S, Ledoux, H, Zhao, J, et al., 2016. Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Transactions in GIS*, 20, doi: 10.1111/tgis.12162.

Essendorfer, B, Kerth, C, and Zaschke, C, 2017. Adaptation of interoperability standards for cross domain usage. *SPIE Defense+ Security*.

Evans, S, Savian, C, Burns, a, et al., 2019. Digital twins for the built environment. https://www.theiet.org/media/4719/digital-twins-for-the-built-environment.pdf.

Fjeld, TMB, 2020. Digital Twin - Towards a joint understanding within the AEC/FM sector. Master, Faculty of Engineering, Department of Civil and Environmental Engineering, Norwegian University of Science and Technology.

Fossey, E, Harvey, C, Mcdermott, F, et al., 2002. Understanding and Evaluating Qualitative Research. *Australian & New Zealand Journal of Psychiatry*, 36 (6):717-732, doi: 10.1046/j.1440-1614.2002.01100.x.

Gilbert, T, Rönsdorf, C, Plume, J, et al., 2020. Built environment data standards and their integration: an analysis of IFC, CityGML and LandInfra. Open Geospatial Consortium: Wayland, MA, USA, buildingSMART International: London, UK,, https://www.ogc.org/docs/discussion-papers.

Golfarelli, M, and Rizzi, S, 2011. Data warehouse testing: A prototype-based methodology. *Information and Software Technology*, 53 (11):1183-1198, doi: 10.1016/j.infsof.2011.04.002.

Gusenbauer, M, 2019. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, 118 (1):177-214.

Gusenbauer, M, and Haddaway, NR, 2020. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11 (2):181-217, doi: 10.1002/jrsm.1378.

Hart, G, and Dolbear, C, 2016. *Linked data: a geographic perspective*: CRC Press.

Hbeich, E, and Roxin, A, 2020. Linking BIM and GIS Standard Ontologies with Linked Data. *LDAC2020 - 8th Linked Data in Architecture and Construction Workshop (17-19 June 2020)*, Virtual.

Hetherington, J, and West, M, 2020. The pathway towards an Information ManagementFramework. A 'Commons' for Digital Built Britain. Digital Built Britain.

Hitzler, P, Krötzsch, M, Parsia, B, et al., 2012. OWL 2 Web Ontology Language Primer (Second Edition). https://www.w3.org/TR/owl2-primer/ (10. January 2021).

Hoepfl, MC, 1997. Choosing qualitative research: A primer for technology education researchers. *Journal of Technology Education*, Volume 9 Issue 1 (fall 1997).

Hoppe, A, Seising, R, Nurnberger, A, et al., 2011. Wisdom - the blurry top of human cognition in the DIKW-model? *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-2011)*, 2011-01-01.

Hor, A-H, Jadidi, A, and Sohn, G, 2016. BIM-GIS INTEGRATED GEOSPATIAL INFORMATION MODEL USING SEMANTIC WEB AND RDF GRAPHS. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 3.

IEC, 2019. Semantic interoperability: challenges in the digital transformation age. International Electrotechnical Commission, https://basecamp.iec.ch/download/iec-white-paper-semantic-interoperability-challenges-in-the-digital-transformation-age-en/.

ISO, 2020. ISO/TC 184/SC 4 Automation systems and integration - Industrial data. https://www.iso.org/committee/54158.html (10 January 2021).

ISO, IEC, and ITU-T, 2020. The Reference Model of Open Distributed Processing (RM-ODP). http://www.rm-odp.net/#More%20about%20RM-ODP (10 January 2021).

ISO/IEC JTC 1, 2015. ISO/IEC 2382:2015 Information technology -- Vocabulary. ISO: Geneva, Switzerland, https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:ed-1:v1:en.

ISO/IEC JTC 1/SC 7, 2012a. ISO/IEC 19505-1:2012 Information technology -- Object Management Group Unified Modeling Language (OMG UML) -- Part 1: Infrastructure. ISO: Geneva, Switzerland.

ISO/IEC JTC 1/SC 7, 2012b. ISO/IEC 19505-2:2012 Information technology — Object Management Group Unified Modeling Language (OMG UML) — Part 2: Superstructure. ISO: Geneva, Switzerland.

ISO/TC 37/SC 1, 2019. ISO 1087:2019 Terminology work and terminology science — Vocabulary. ISO: Geneva, Switzerland.

ISO/TC 59/SC 13, 2018a. ISO 16739-1:2018 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema. ISO: Geneva, Switzerland.

ISO/TC 59/SC 13, 2018b. ISO 19650-1:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) - Information management using building information modelling - Part 1: Concepts and principles. ISO: Geneva, Switzerland.

ISO/TC 59/SC 13, 2020. ISO/DTR 23262.2 GIS (Geospatial) / BIM interoperability. ISO: Geneva, Switzerland.

ISO/TC 184 SC 4, 1997. ISO/TR 10303-12:1997 Industrial automation systems and integration -- Product data representation and exchange -- Part 12: Description methods: The EXPRESS-I language reference manual. ISO: Geneva, Switzerland.

ISO/TC 184 SC 4, 2004. ISO 10303-11:2004 Industrial automation systems and integration — Product data representation and exchange — Part 11: Description methods: The EXPRESS language reference manual. ISO: Geneva, Switzerland.

ISO/TC 184 SC 4, 2005. ISO/TS 10303-25:2005 Industrial automation systems and integration — Product data representation and exchange — Part 25: Implementation methods: EXPRESS to XMI binding. ISO: Geneva, Switzerland.

ISO/TC 184 SC 4, 2007. ISO 10303-28:2007 Industrial automation systems and integration — Product data representation and exchange — Part 28: Implementation methods: XML representations of EXPRESS schemas and data, using XML schemas. ISO: Geneva, Switzerland.

ISO/TC 184 SC 4, 2016. ISO 10303-21:2016 Industrial automation systems and integration -- Product data representation and exchange -- Part 21: Implementation methods: Clear text encoding of the exchange structure. ISO: Geneva, Switzerland.

ISO/TC 204, 2020. ISO/DIS 20524-2 Intelligent transport systems — Geographic Data Files (GDF) — GDF5.1 — Part 2 : Map data used in automated driving systems , Cooperative ITS , and multi-modal transport. ISO: Geneva, Switzerland.

ISO/TC 211, 2014. ISO 19101-1:2014 Geographic Information — Reference model — Part 1: Fundamentals.

ISO/TC 211, 2015. ISO 19103:2015 Geographic Information — Conceptual schema language. ISO: Geneva, Switzerland.

ISO/TC 211, 2020a. ISO CD/TR 19169 Geographic Information — Gap-analysis: To map and describe the differences between the current GDF and ISO/TC211 conceptual models to suggest ways harmonize and resolve conflicting issues. ISO: Geneva, Switzerland.

ISO/TC 211, 2020b. ISO/TC 211 Multi-Lingual Glossary of Terms. https://github.com/ISO-TC211/TMG/releases (10 January 2021).

ISO/TC 211, 2020c. ISO/TC 211 Resolutions. https://committee.iso.org/sites/tc211/home/resolutions.html (10 January 2021).

Janowicz, K, Schade, S, Bröring, A, et al., 2010. Semantic enablement for spatial data infrastructures. *Transactions in GIS*, 14 (2):111-129.

Jassim, HSH, 2019. Assessing Energy Use and Carbon Emissions to Support Planning of Environmentally Sustainable Earthmoving Operations. PhD, Luleå University of Technology.

Jennex, ME, 2017. Big Data, the Internet of Things, and the Revised Knowledge Pyramid. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 48 (4):69-79, doi: 10.1145/3158421.3158427.

Jetlund, K, 2018a. Experiences and challenges with standards for location referencing from the GIS and ITS domains. *25th ITS World Congress*, Copenhagen, Denmark, 17-21 September 2018.

Jetlund, K, 2018b. Improvements in automated derivation of OWL ontologies from geospatial UML models *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-4:283-290, doi: 10.5194/isprs-archives-XLII-4-283-2018.

Jetlund, K, 2020. A structure of UML profiles for modelling of geospatial information in GIS, ITS and BIM. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, VI-4/W1-2020:101-108, doi: 10.5194/isprs-annals-VI-4-W1-2020-101-2020.

Jetlund, K, and Neuhäuser, B, (Awaiting publication). GIS for Transportation. In *Springer Handbook of Geographic Information*.

Jetlund, K, Onstein, E, and Huang, L, 2019a. Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies. *ISPRS International Journal of Geo-Information*, 8 (9):365, doi: https://doi.org/10.3390/ijgi8090365.

Jetlund, K, Onstein, E, and Huang, L, 2019b. Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model. *Isprs International Journal of Geo-Information*, 8 (3):141, doi: https://doi.org/10.3390/ijgi8030141.

Jetlund, K, Onstein, E, and Huang, L, 2020. IFC Schemas in ISO/TC 211 compliant UML for improved interoperability between BIM and GIS. *ISPRS International Journal of Geo-Information*, 9 (4), doi: https://doi.org/10.3390/ijgi9040278.

Jomrich, F, Sharma, A, Rückelt, T, et al., 2017. Dynamic Map Update Protocol for Highly Automated Driving Vehicles. *Proceedings of the International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*.

Kerry, R, 1995. Reference model of open distributed processing (RM-ODP): Introduction. In *Open distributed processing*, 3-14. Springer.

Kommunal- og moderniseringsdepartementet, 2020. Det offentlige kartgrunnlaget (DOK). https://www.regjeringen.no/no/tema/plan-bygg-og-eiendom/plan--og-bygningsloven/plan/veiledning-om-planlegging/plankartsiden/det-offentlige-kartgrunnlaget-dok/id2470662/.

Kresse, W, and Danko, D, 2011. *Springer handbook of geographic information*: Springer Berlin Heidelberg.

Kruchten, P, 2004. *The rational unified process: an introduction*: Addison-Wesley Professional.

Kruchten, PB, 1995. The 4+ 1 view model of architecture. *IEEE software*, 12 (6):42-50.

Kubicek, H, Cimander, R, and Scholl, HJ, 2011. Layers of Interoperability. In *Organizational Interoperability in E-Government*, 85-96. Springer Berlin Heidelberg.

Kutzner, T, 2016. Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles. PhD, Technische Universität München.

Kutzner, T, Chaturvedi, K, and Kolbe, TH, 2020. CityGML 3.0: New Functions Open Up New Applications. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88 (1):43-61, doi: 10.1007/s41064-020-00095-z.

Lankhorst, M, 2009. Viewpoints and Visualisation. In *Enterprise Architecture at Work: Modelling, Communication and Analysis*, 153-197. Berlin, Heidelberg: Springer Berlin Heidelberg.

Liu, X, Wang, X, Wright, G, et al., 2017. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 6 (2):53.

Luiten, B, Böhms, M, Alsem, D, et al., 2019. Asset information management using Linked Data for the life-cycle of Roads. *Delft: CRC Press/Balkema*.

Luiten, B, Böhms, M, O'Keeffe, A, et al., 2016. A hybrid linked data approach to support asset management. *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016)*, Delft, The Netherlands, 16-19 October 2016.

Luiten, B, O'Keeffe, A, Stolk, S, et al., 2017. INTERLINK D4. Principles for a European Road OTL.

Manola, F, Miller, E, and McBride, B, 2014. RDF 1.1 Primer. https://www.w3.org/TR/rdf11-primer/ (10 January 2021).

Miles, RR, and Hamilton, K, 2006. *Learning UML 2.0*. First ed: O'Reilly.

Noy, NF, and McGuinness, DL, 2001. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory, Stanford University, CA, USA. https://protegewiki.stanford.edu/wiki/Ontology101 (10 January 2021).

NPRA, and SINTEF, 2008. ROSATTE: ROad Safety ATTributes exchange infrastructure in Europe: Deliverable D1.1 State of the Art. https://www.tn-its.eu/storage/uploads/documents/2018/09/02/ROSATTE-D11-State-of-the-art-v10.pdf.

Object Management Group, 2014a. Model Driven Architecture (MDA) Guide rev. 2.0. Object Management Group: Needham, MA, USA, https://www.omg.org/mda/.

Object Management Group, 2014b. The Object Constraint Language Version 2.4. Object Management Group: Needham, MA, USA.

Object Management Group, 2014c. Ontology Definition Metamodel, Version 1.1. Object Management Group: Needham, MA, USA, https://www.omg.org/spec/ODM/1.0/About-ODM/.

Object Management Group, 2015a. Reference Metamodel for the EXPRESS Information Modeling Language. Object Management Group: Needham, MA, USA.

Object Management Group, 2015b. XML Metadata Interchange (XMI) Specification Version 2.5.1. Object Management Group: Needham, MA, USA, https://www.omg.org/spec/XMI.

Object Management Group, 2017. Unified Modelling Language Specification Version 2.5.1. Object Management Group: Needham, MA, USA, https://www.omg.org/spec/UML.

Object Management Group, 2019. OMG Meta Object Facility (MOF) Core Specification. Object Management Group: Needham, MA, USA.

Object Management Group, 2020. MDA. Object Management Group: Needham, MA, USA. https://www.omg.org/mda/ (10 January 2021).

Oxford Dictionaries, 2019. Oxford English Dictionary. https://en.oxforddictionaries.com (10 January 2021).

Paul, A, Chilamkurti, N, Daniel, A, et al., 2017. *Intelligent Vehicular Networks and Communications*: Elsevier.

Peng, ZR, 2005. A proposed framework for feature-level geospatial data sharing: a case study for transportation network data. *International Journal of Geographical Information Science*, 19 (4):459-481.

Pilone, D, and Pitman, N, 2005. *UML 2.0 in a Nutshell*: O'Reilly.

Rasmussen, MH, Hviid, CA, and Karlshøj, J, 2017. Web-based topology queries on a BIM model. *LDAC2017–5th Linked Data in Architecture and Construction Workshop*.

Regoczei, SB, and Hirst, G, 1992. Knowledge and knowledge acquisition in the computational context. In *The psychology of expertise*, 12-25. Springer.

Richter, A, and Scholz, M, 2017. Deploying guidelines and a simplified data model to provide real world geodata in driving simulators and driving automation. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Rowley, J, 2002. Using case studies in research. *Management Research News*, 25 (1):16-27, doi: 10.1108/01409170210782990.

Rowley, J, 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*, 33 (2):163-180.

Roxin, A, and Hbeich, E, 2019. Semantic interoperability between BIM and GIS–review of existing standards and depiction of a novel approach. *In 36th CIB W78–Information Technology for Construction.*

Schenck, D, and Wilson, P, 1994. *Information modelling : The EXPRESS way*. New York: Oxford University Press.

Shibata, J, 2018. ISO TC204 WG3 : ITS Database Technology PROGRESS REPORT as of March 12 , 2018.

Song, Y, Wang, X, Tan, Y, et al., 2017. Trends and Opportunities of BIM-GIS Integration in the Architecture, Engineering and Construction Industry: A Review from a Spatio-Temporal

Statistical Perspective. *ISPRS International Journal of Geo-Information*, 6:397, doi: 10.3390/ijgi6120397.

Statens vegvesen, 2014. Håndbok R610: Standard for drift og vedlikehold av riksveger. Statens vegvesen, https://www.vegvesen.no/_attachment/61430/binary/964067.

Statens vegvesen, 2015. Håndbok V770: Modellgrunnlag - Krav til grunnlagsdata og modeller. Statens vegvesen, https://www.vegvesen.no/_attachment/395908/binary/1098509?fast_title=H%C3%A5ndbok+V770+Modellgrunnlag.pdf.

Statens vegvesen, 2018. ITS-strategi for Statens vegvesen 2018-23 Et vegkart mot fremtidens transportsystem. Statens vegvesen, https://www.vegvesen.no/_attachment/2198857/binary/1241615?fast_title=ITS-strategi+2018-2023.pdf.

Statens vegvesen, 2020a. Intelligente transportsystemer (ITS) i Statens vegvesen. https://www.vegvesen.no/fag/trafikk/its (10 January 2021).

Statens vegvesen, 2020b. Nasjonal vegdatabank (NVDB). https://www.vegvesen.no/fag/teknologi/nasjonal+vegdatabank (10 January 2021).

The European Commission C-ITS Deployment Platform, 2016. C-ITS Platform: Final Report. The European Commission, https://ec.europa.eu/transport/sites/transport/files/themes/its/doc/c-its-platform-final-report-january-2016.pdf.

The European Commission C-ITS Deployment Platform, 2017. C-ITS Platform Final report Phase II. The European Commission, https://ec.europa.eu/transport/sites/transport/files/2017-09-c-its-platform-final-report.pdf.

Tobler, WR, 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46:234-240, doi: 10.2307/143141.

Tolk, A, and Muguira, JA, 2003. The levels of conceptual interoperability model. *Proceedings of the 2003 fall simulation interoperability workshop*.

Tronvoll, SA, Elverum, CW, and Welo, T, 2017. Prototype Experiments: Strategies and Trade-offs. *Procedia CIRP*, 60:554-559, doi: 10.1016/j.procir.2017.01.049.

Ustundag, A, and Cevikcan, E, 2017. *Industry 4.0: managing the digital transformation*: Springer.

Veltman, KH, 2001. Syntactic and semantic interoperability: new approaches to knowledge and the semantic web. *New Review of Information Networking*, 7 (1):159-183.

Wang, W, Tolk, A, and Wang, W, 2009. The levels of conceptual interoperability model: applying systems engineering principles to M&S. *arXiv preprint arXiv:0908.0191*.

World Wide Web Consortium, 2020. W3C. https://www.w3.org/ (10 January 2021).

Yin, RK, 2018. *Case study research and applications*: Sage.

Zang, A, Li, Z, Doria, D, et al., 2017. Accurate vehicle self-localization in high definition map dataset. *Proceedings of the 1st ACM SIGSPATIAL Workshop on High-Precision Maps and Intelligent Applications for Autonomous Vehicles*, Redondo Beach, CA, USA, 7–10 November 2017.

Zhao, Y, Brown, R, Kramer, TR, et al., 2011. Practices of Information Modeling. In *Information Modeling for Interoperable Dimensional Metrology*, 21-52. London: Springer London. Original edition, Information Modeling for Interoperable Dimensional Metrology.

Zhu, J, Wright, G, Wang, J, et al., 2018. A Critical Review of the Integration of Geographic Information System and Building Information Modelling at the Data Level. *ISPRS International Journal of Geo-Information*, 7:66, doi: 10.3390/ijgi7020066.

Conference proceeding 1

Experiences and challenges with standards for location referencing from the GIS and ITS domains

25th ITS World Congress, Copenhagen, Denmark, 17-21 September 2018

Paper ID EU-SP1445

Experiences and challenges with standards for location referencing from the GIS and ITS domains

Knut Jetlund[1*]

1. Norwegian University of Science and Technology, and Norwegian Public Roads Administration, Norway,

knut.jetlund@vegvesen.no

Abstract

Intelligent Transport Systems (ITS) for driver assistance, from basic route planners to systems for autonomous driving, depend on geospatial road information from various sources. One important source is authoritative information from road authorities concerning road networks, restrictions, traffic information, etc. ISO and CEN standardization committees in the domains of geographic information (GIS) and ITS have developed several standards with models, location referencing methods and exchange formats for geospatial information. Road authorities are maintaining information in systems based on GIS standards, while the exchange formats and the receiving systems on the ITS side are based on standards from the ITS domain. The senders and receivers of data will need to handle the different conceptual models and exchange formats from standards in the two domains. This paper presents an overview of relevant standards and specifications, identifies challenges related to interoperability, and suggest future research.

Keywords:

Geospatial, standards, interoperability

Introduction

Geospatial information representing road networks and restrictions have been used in applications for route planning and vehicle navigation for several decades. With the expanding use of smart mobile devices and sensors over the last years, geospatial information and applications for navigation has become more and more important, and are now an integrated part of our daily life. This is also true for vehicles, where most new vehicles include sensors and systems for route planning and navigation, as well as advanced driver assistance systems (ADAS) for sign recognition, lane keeping assistance, speed alert etc. and even more advanced systems for autonomous driving. All these systems, from basic route planning services to systems for autonomous driving, depend on geospatial knowledge from a variety of sources. Commercial map providers have created and delivered data sets for route planning and navigation for a long time, and are extending their products to support autonomous driving. New vehicles contain a range of sensors, and are able to create their own map of the surroundings, and to upload data to vehicle manufacturers, who again may deliver the data to other vehicles. Road authorities maintain information with road network, restrictions, road equipment, traffic information, events etc. Even though a large amount of data is being captured from individual vehicles and by private map providers, the road authorities that set the speed limits, turn restrictions, weight limitations etc. are the main authoritative source for restrictions in the road network, as well as information about planned roadwork, closed roads due to weather or accidents etc. The authorities are also likely to have information for new roads before the road opens, and for rural areas where fewer vehicles have travelled. The systems for route planning, navigation, ADAS, C-ITS and autonomous driving will need to combine data from all these sources into the knowledge needed for safe navigation. To achieve this, each source must provide information in a way that systems can understand and validate, both in terms of what the data represents, and of the location references provided. For this purpose, standardization and harmonization is a crucial element.

The goal of this paper is threefold: First, to develop a comprehensive overview of standards and specifications for geospatial road information, including a walk-through of location referencing methods defined in the standards. Second, to identify possible challenges for interoperability and information exchange based on the overview. Third, to suggest future research that may improve interoperability and information exchange.

Related work

The INSPIRE Directive [1], the ITS Directive [2] and the Delegated Regulation 2015/962 [3] are foundations for the sharing of geospatial road information in Europe. Through the directives, requirements are given for what information shall be shared from public authorities, while the delegated regulation ties the two directives together, and specifies that static information according to the ITS Directive shall be compatible with the specifications established by the INSPIRE Directive, and in particular the theme Transport Networks [4]. Several perspectives of the ITS Directive are discussed in [5], among them the availability, accuracy and re-use of public sector data. The need for harmonized specifications for the data that shall be provided is emphasized, and the authors point at the INSPIRE specification for Transport Networks as one possibility. One major difference between INSPIRE and the ITS Directive is pointed out: While the INSPIRE directive does not impose an obligation upon EU member states to collect specific spatial information, this can be required under the ITS Directive.

The European Strategy on Cooperative Intelligent Transport Systems (EU C-ITS) [6] points out C-ITS applications that are to be deployed all over Europe, based on the work in the EU C-ITS Platform [7; 8]. The platform introduced a priority list of services that were considered mature enough to be included from day 1, among them several hazardous location notifications (road works, weather conditions, emergency vehicle approaching etc.) and signage applications (in-vehicle signage, in-vehicle speed limits etc.). Furthermore, a list of the next prioritized services for ‘day 1.5’ was also introduced. The focus of the work with the platform was primarily services for information exchange, but future development against fully connected, cooperative and automated vehicles was studied as well in phase II of the platform. The applications for C-ITS are depending on several standards, developed by several standardization actors, and standards that form the foundation for C-ITS applications are discussed in [9; 10].

The ROSATTE Project [11-13] with its descendants TN-ITS and the EULF Transportation Pilot [14] developed a specification for exchange of changes on road information. This specification has been further developed by CEN/TC 278 to become a CEN Technical Specification. The purpose of the projects and the specification has been to enable mechanisms for exchanging changes in road data from road authorities to commercial map providers. Services based on the ROSATTE specification were first implemented in Sweden and Norway, and have also been implemented in several other European countries, while even more countries are planning to implement services in the near future.

The CEN/TC 278 Project team 1703 has worked on location referencing harmonization, with a special focus on ITS in the urban environment. The project team has studied and documented the different location referencing methods that are specified in International and European standards and use cases where they may be applied, and described methods for transformation between the location referencing methods [15; 16].

The needs and benefits of updated authoritative information on road networks and restrictions for route planning and navigation has been documented in several projects and articles. The European HeavyRoute project [17; 18] focused on route guidance for heavy goods vehicles. One of the findings was that route planning for heavy vehicles need additional information compared to what is normally included in data sets and services for route planning. Restrictions related to height, width and weight are essential, and a recommended route should be a combination of the allowable routes based on restrictions, services like lodging, and a cost effective route based on driver safety and comfort, environmental effects etc. A study on route planning for heavy vehicles has also been conducted in [19], where one finding was that a substantial amount of processing and data cleaning was needed.

A study of interoperability of public transport datasets from multiple authorities was conducted in [20], with a goal of lowering the cost of adopting data from different providers in route planning services. A main approach was the use of Semantic Web technologies to combine information from diverse sources.

Standardization actors

International standardization of geospatial information relevant for the ITS domain is mainly done by two committees within ISO: ISO/TC 211 – Geographic information/Geomatics [21], and ISO/TC 204 – Intelligent Transport Systems [22]. Within ISO/TC 204, standards from WG 3 (ITS Database Technology) are particularly relevant for the scope of this paper, and some work from WG 10 (Traveller Information Systems) as well. CEN and ETSI develop European standards, with the CEN Technical Committee 278 responsible for standardization in the ITS domain [23]. Several CEN/TC 278 working groups are developing standards that specifies the exchange of geospatial road information, in particular WG7 (ITS Spatial Data), and WG8 (Road Traffic Data).

I addition to ISO, CEN and ETSI, several other actors develop relevant standards, among them The Open Geospatial Consoritum (OGC), Car-to-Car Communicational Consortium (C2C-CC), TomTom

International, and Navigation Data Standards (NDS). The scope of this paper is on the main official standards developed by ISO/TC 211, ISO/TC 204 and CEN/TC 278, but standards from other actors may be included in further research.

Table 1 – Standardization actors

	Domain	
Scope	**GIS**	**ITS**
Official International Standards	ISO/TC 211	ISO/TC 204
Official European Standards		CEN/TC 278, ETSI TC ITS
Industry/consortiums	OGC	C2C-CC, TomTom, NDS

Conceptual modelling approaches

The standardization work in ISO/TC 211 is based on the principles of Model Driven Architecture (MDA), where application schemas and implementation schemas are derived from conceptual models. The foundation for this is the profile of The Unified Modelling Language (UML), described in ISO 19103 and ISO 19109. It has been important for ISO/TC 211 to create interoperable standards that reuse already defined concepts, and most of the ISO/TC 211 standards are developed with conceptual UML models based on ISO 19103. The models are maintained in a common harmonized UML model, where concepts defined in one standard are reused in other standards. ISO/TC 211 also maintain a common glossary of terms, to make sure terms are defined once, and reused in other standards. Implementation schemas for ISO/TC 211 models and from other UML models based on the standards are derived directly from the UML models, based on rules for conversion from UML to implementation schemas in ISO 19136 and ISO 19139. Tools for this conversion have been developed by software vendors, with ShapeChange [24] as the most widely used tool.

The approach with a common conceptual schema language, a common harmonized UML model and common rules for conversion to implementation schemas has been successful for the creation of interoperable standards in the GIS domain. The standards from ISO/TC 211 are widely used in applications, databases and exchange formats for geospatial information, and are the foundation for a wide range of national and regional application schemas, e.g. the INSPIRE data specifications.

In the ITS domain, the TPEG2 (Traffic and travel information via transport protocol experts group, generation 2) and DATEX II (Data exchange specifications for traffic management and information) series of standards are based on specific profiles of UML, and TPEG2 also describe rules for conversion from UML to XML in part 4. The ISO 17572 Location referencing series use UML as well, but not based on a specific profile. ISO 14825 Geographic Data Files (GDF) 5.0 was developed following ISO 19103 and ISO 19019, but this has not been followed up for the new version 5.1. There is no common UML Profile for all ISO/TC 204 and CEN/TC 278 standards, and there is no common harmonized UML model either.

Table 2 – Foundation standards for conceptual modelling of geospatial information

Standard	Name
Object Management Group	
MDA	Model Driven Architecture (MDA)
UML	Unified Modelling Language (UML)
ISO/TC 211	
ISO 19103:2015	Conceptual schema language
ISO 19109:2015	Rules for application schema
ISO 19136:2007	Geography Markup Language (GML)
ISO/DTS 19139-1	Metadata — XML schema implementation — Part 1: Encoding rules (revision of ISO/TS 19139:2007)
ISO/TC 204	
ISO/TS 21219-2:2014	TPEG2 — Part 2: UML modelling rules
ISO/TS 21219-4:2015	TPEG2 — Part 4: UML to XML conversion rules
CEN/TC 278	
prEN 16157-1	DATEX II — Part 1: Context and Framework

Standards for location referencing

Standards for geospatial road information describe several different location referencing methods, and are also defining the terms 'location' and 'location referencing' in different manners. The CEN/TC 278 Project team 1703 has studied the definitions in different standards, and suggested a common definition where location referencing is used to describe a location within a location referencing system (LRS), according to a location referencing method (LRM) [15; 16]. A single location in the real world, e.g. the location of a road link, a traffic sign, a speed limit or an accident may be described with different location references, based on different LRMs and LRSs defined in standards, and in registries based on these.

ISO/TC 211 have a range of standards for geographic information, all except one within the ISO 19100 series. Five of these standards define the framework for location referencing. ISO 19111 defines the concepts of coordinate reference systems (CRS), coordinate systems, datums, conversions, projections and more. These concepts are the foundation for e.g. the widely used EPSG Geodetic Parameter Registry where specific CRSs are defined and identified with a EPSG code [25], and for referencing locations by coordinates in other ISO/TC 211 standards. A pair or a triple of coordinates is not sufficient to identify a location; one will also need to know the CRS. Therefore, exchange of coordinate based location references must include information about the CRS, or the sender and the receiver of the location reference must have agreed in advance on what CRS to use. Closely related to ISO 19111 is the standard ISO 6709, which defines how geographic coordinates shall be represented. ISO 19112 defines the concepts for location referencing based on geographic identifiers, where the geographic identifiers have representational geometries that relate the location reference to the real world. ISO 19107 defines the spatial schema, with data types for geometries ranging from simple points and lines to complex parameter based shapes and solids, all referenced to the concepts from ISO 19111. ISO 19148 defines the conceptual model for linear referencing, with methods for describing positions of points and segments with one-dimensional coordinates along links, e.g. roads. An advantage with linear referencing is that the basic road network may be held stable, while changes in characteristics such as speed limits, road width etc. are referred to the network. Linear referencing is widely used in network datasets for transportation and utilities, and is also used in the ITS domain.

Table 3 – ISO/TC 211 Standards for location referencing

Standard	Name
ISO/CD 19111	Referencing by coordinates (revision of ISO 19111:2007)
ISO 6709:2008	Standard representation of geographic point location by coordinates
ISO/DIS 19112	Spatial referencing by geographic identifiers (revision of ISO 19112:2003)
ISO/DIS 19107	Spatial schema (revision of ISO 19107:2003)
ISO 19148:2012	Linear referencing

The core ISO/TC 204 standard for location referencing is ISO 17572, with the core concepts for location referencing in part 1. The standard specifies two basic methods for location referencing: Pre-coded location references, described in part 2, and dynamic location referencing in part 3. In addition, a new part 4 for location referencing at lane level is under development. The concept of pre-coded location references is that the sender and the receiver of a location reference have access to identical location databases. A commonly used pre-coded method referred to from ISO 17572 part 2 is the ALERT-C protocol, defined in ISO 14819-3. This method is used for exchanging traffic messages through RDS-TMC. Linear referencing has also been classified as pre-coded location referencing. Dynamic location referencing as described in ISO 17572 part 3 has been developed to enable exchange of road data and road related data between road network datasets, while compensating for differences in the datasets. The AGORA-C method is a dynamic location referencing method defined in ISO 17572 part 3, another method is the open standard OpenLR, developed by TomTom. The foundation of dynamic location referencing is a set of road characteristics and a simplified network geometry. Both AGORA-C and OpenLR use characteristics such as road name, form of way and functional road class. The messages containing the geometry and the characteristics is encoded at the sender side, while the receivers decodes the message and matches the information with their own dataset. Due to differences between the network datasets, some errors are likely to occur. However, the ROSATTE Project tested and documented good results for matching in several test areas in Europe [26].

In addition to the standards from WG 3, ISO 21219 TPEG2 from WG 10 is relevant for location referencing. TPEG2 is a multi-part standard, where several parts describe mechanisms for location referencing: Part 7 describes the basic element location referencing container, while more specific location referencing methods are described in part 21, which describes geographic location referencing, and in part 22, which describes location referencing using OpenLR.

Table 4 – ITS Standards and specifications for location referencing

Standard	Name
ISO 17572	Location referencing for geographic databases
ISO 17572-1:2015	Part 1: General requirements and conceptual model
ISO 17572-2:2015	Part 2: Precoded Location References (Precoded Profile)
ISO 17572-3:2015	Part 3: Dynamic Location References (Dynamic Profile)
ISO/WD 17572-4	Part 4 Lane-level location references (lane-level profile)
ISO 14819-3:2013	Traffic and travel information messages via traffic message coding — Part 3: Location referencing for Radio Data System — Traffic Message Channel (RDS-TMC) using ALERT-C
OpenLR	TomTom OpenLR
ISO/TS 21219-7:2017	TPEG2 — Part 7: Location referencing container (TPEG2-LRC)
ISO/PDTS 21219-21	TPEG2 — Part 21: Geographic location referencing (TPEG-GLR)
ISO/TS 21219-22:2017	TPEG2 — Part 22: OpenLR location referencing (TPEG2-OLR)

Other relevant ISO/TC 204 Standards

Several other standards developed by ISO/TC 204 WG 3 are also concerning geospatial information, primarily but not limited to standards described here. ISO 14825 describes the conceptual and logical data model for geographic databases, including the GDF 5.0 format for data exchange. A revised and extended version of GDF is under development as ISO 20524 – GDF 5.1. While GDF 5.0 mainly deals with applications for navigation systems, version 5.1 extends the scope to include requirements from applications for C-ITS, multi-modal transportation and systems for automated driving. Part 2 also introduces the Belt-concept for automated driving. ISO 14296 defines models for static map data for use in C-ITS, while the development of ISO 22726 will standardize models for map data for automated driving, and also define the concepts of static, semi-static and semi-dynamic data. ISO 19297 focuses on the sharing of various geospatial databases for use in ITS services, currently with the framework under development in part 1.

Table 5 – Other relevant ISO/TC 204 standards

Standard	Name
ISO 14825:2011	Geographic Data Files (GDF) — GDF5.0
ISO/DIS 20524-1	Geographic Data Files (GDF) GDF5.1 — Part 1 : Application independent map data shared between multiple sources
ISO/WD 20524-2	Geographic Data Files (GDF) — GDF5.1 — Part 2 : Map data used in automated driving systems , Cooperative ITS , and multi-modal transport

ISO 14296:2016	Extension of map database specifications for applications of cooperative ITS
ISO/PWI 22726	Dynamic events and map database specifications for applications of automated driving systems, cooperative ITS, and advanced road/traffic management systems
ISO/CD 19297-1	Shareable Geospatial Databases for ITS Applications — Part 1: Framework

ISO/TC 211 and ISO/TC 204 Joint work

The two committees ISO/TC 204 and ISO/TC 211 have mainly been working with standards within their own domain, but there have also been initiatives for cooperation in areas where the two domains intersect. ISO/TC 211 have a range of standards for location based services (LBS), mainly ISO 19132, ISO 19133 and ISO 19134. These standards have existed for some years, and at the same time the technologies for LBS and ITS have developed rapidly, with mobile phones and sensors on vehicles and infrastructure. Based on this, the future of these standards have been discussed in an ad hoc group, where also standards from the ITS domain have been discussed [27].

Table 6 – ISO/TC 211 standards for location based services

Standard	Name
ISO 19132:2007	Location-based services — Reference model
ISO 19133:2005	Location based services — Tracking and navigation
ISO 19134:2006	Location based services — Multimodal routing and navigation

The ISO/TC 211 standards are widely used in application and database systems in GIS domain, including maintenance of geospatial information representing road networks and restrictions. Mapping authorities and road authorities are likely to maintain their geospatial information in systems based on ISO/TC 211 standards [28]. To be able to use such information in the ITS domain, standards for geographic information and standards for ITS need to communicate. For this purpose, a joint taskforce (JTF) between ISO/TC 204 WG 3 and ISO/TC 211 WG 10 (Ubiquitous public access) has been established, where the rationale is to learn more about each other’s standards, models and concepts, and to start working together and create standards based on a common understanding.

European specifications for exchange of geospatial road information

The data specifications, the models and the exchange format specifications for INSPIRE themes are based on ISO/TC 211 standards. UML is used for modelling, based on ISO 19103 and ISO 19109, and GML as described in ISO 19136 is the exchange format. The specification for Transport Networks [4] describes a multimodal transport network for road, rail, air and water transport, with a common model and specified models for each transportation mode. The specification contains a model with elements described as links, nodes and areas, and network properties that are localized on the network element with linear referencing. The model for linear referencing was developed prior to ISO 19148, based on the same concepts, but limited to only one of the methods defined in the standard.

Two standardization projects from CEN/TC 278 are particularly relevant for this paper. The specification developed through the ROSATTE Project [13] and later known as TN-ITS [14] have been further developed to become a CEN Technical Specification for data exchange on changes in road attributes. The specification is based on ISO/TC 211 standards, with an ISO 19109 compliant application schema, and a GML schema for exchange based on ISO 19136. Furthermore, the specification includes models from the DATEX II standard for validity periods and vehicle classifications, and from ISO 14823 (Graphic data dictionary) for traffic sign classification. Several possible location referencing methods are provided, where at least one is mandatory for each road feature: Geometry according to ISO 19107, dynamic location referencing with AGORA-C or the OpenLR specification, or pre-coded location referencing; either as a URI to a predefined location, or through linear referencing. The model for linear referencing is compliant to ISO 19148 and the linear referencing model used in INSPIRE Transport Networks [2].

While the TN-ITS specification is focusing on changes in static road data, the DATEX II standard has its focus on dynamic data for traffic and travel. DATEX II is referred to from the Delegated Regulation 2015/962 [3] as a preferred method for providing real-time information. DATEX II is a multi-part standard, at the time consisting of four parts: Part 1 defines the context and framework. Part 3 describes the most commonly used publication for traffic information messages, while part 7 describes common classes and data types required for publishing information. Part 2 describes the model for location referencing, with several possible location referencing methods, including TPEG, AGORA-C, OpenLR, geometry according to ISO 19107, and linear referencing according to ISO 19148.

Table 7 – European specifications for exchange of geospatial road information

Standard	Name
TN-ITS	Data exchange on changes in road attributes
prEN 16157-2	DATEX II – Part 2: Location referencing
prEN 16157-3	DATEX II – Part 3: Situation Publication
prEN 16157-7	DATEX II – Part 7: Common data elements

Discussion

The previous sections of this paper has presented an overview of standards for geospatial road information. This section will discuss possible challenges for interoperability and information exchange based on the overview.

Several actors from the GIS and ITS domains have developed standards concerning geospatial road information. Most of the standards have been developed in separate silos, based on different use of the UML language, without reusing concepts already defined by other actors, and without harmonizing with standards from other actors. This has led to different conceptual models for the same real-world phenomena, different models for describing locations and different exchange formats. Road and mapping authorities maintain road networks, restrictions and other geospatial information that are very relevant for use in the ITS domain. This information is maintained in systems based on standards for geographic information, and the lack of harmonized and interoperable standards is challenging for reusing this information in the ITS domain, both for authorities providing information, and for the receivers. One practical example is the differences between TN-ITS and DATEX II. Services based on these specifications are likely to be provided by the same authorities and used by the same receivers: TN-ITS for changes in static data like speed limits, and DATEX II for dynamic data like road closures. However, the two services will need to be implemented in different ways, as the structure of phenomena and location references are modelled in different ways, and with different encoding in the XML exchange format.

The common modelling rules and the harmonized model for standards in the GIS domain has been an important foundation for the use of standards and exchange of information within that domain. In the ITS domain, the lack of common modelling rules and also the lack of a harmonized model can be considered a weakness for the development of interoperable standards. Furthermore, common modelling rules and harmonized models covering both domains would probably improve interoperability and information exchange both within the ITS domain and between the two domains.

The work in the JTF between ISO/TC 204 and ISO/TC 211 may improve this situation for future work on standardization, through a joint work for standards that touches both domains. However, many existing standards that are already implemented also need to be handled. Interoperability between these standards can be established through a common information model and transformation rules between this model and the original standards, possibly based on Semantic Web technologies as described in [20].

Several LRMs are specified in standards from the two domains. Linear referencing, ALERT-C, AGORA-C and OpenLR are all closely related to the road network. However, they are mainly referring to a road segment in terms of positions along the road; they are not considering the different lanes that may exist across a road. For ADAS and autonomous driving, location referring at this level of detail will be important. Two of the standards mentioned in previous sections are working on models for location referencing at lane level: ISO 17572 part 4 is developing a method for lane-level location referencing, while the Belt-concept in GDF 5.1 part 2 is handling the lanes as belt polygons. More work on location referencing at lane level will probably be done in the future, and this may be a possible arena for a closer cooperation on standards development between the GIS and ITS domains.

Conclusions and research recommendations

Standards covering geospatial road information and location referencing of information have been developed by several actors from the domains of GIS and ITS. The lack of cooperation and common modelling rules both between the domains and within the ITS domain has led to challenges for interoperability and information exchange. The JTF between ISO/TC 204 and ISO/TC 211 is expected to improve this situation for future work through a closer cooperation, and in addition, common modelling rules and common information models would improve interoperability even more. It should be studied if it is possible to establish common information models for standards based on different UML profiles, through transformation between profiles or to a common profile. The approach of using technologies from the Semantic web to combine information from various sources, as described in [20], should also be studied further, to investigate if it is a possible solution for combining different information models.

The studies in this paper have been limited to standards from ISO/TC 204, ISO/TC 211 and CEN/TC 278. There are also standards from other standardization actors where geospatial information is essential, and the studies should be extended to include these standards as well.

The rapidly growing development of technologies for ADAS and autonomous vehicles will lead to new requirements for location referencing and spatial accuracy for geospatial road information, e.g. location referencing at lane level and in belts. Further studies should look into of how locations can be referenced more precisely and accurate, and how the different location referencing methods can handle the requirements.

Finally, as the vehicles are also collecting large amounts of geospatial information with their sensors, it should be studied how standards can contribute to the use of this information for improving the road authorities' databases.

References

1. The European Commision. (2007). *DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).*
2. The European Commission. (2010). *DIRECTIVE 2010/40/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 7 July 2010*.
3. The European Commission. (2015). *COMMISSION DELEGATED REGULATION (EU) 2015/962 of 18 December 2014 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide real-time traffic information services*.
4. INSPIRE. (2010). *INSPIRE Data Specification on Transport Networks – Guidelines*.
5. Vandezande, N., and Janssen, K. (2012). *The ITS Directive: More than a timeframe with privacy concerns and a means for access to public data for digital road maps?* Computer Law & Security Review 28, 416-428.
6. The European Commission. (2016). *A European strategy on Cooperative Intelligent Transport Systems, a milestone towards cooperative, connected and automated mobility*.
7. The European Commission. (2017). *C-ITS Platform Final report Phase II*.
8. The European Commission. (2016). *C-ITS Platform: Final Report*.
9. Festag, A. (2014). *Cooperative intelligent transport systems standards in Europe*. IEEE communications magazine 52, 166-172.
10. Chen, L., and Englund, C. (2014). *Cooperative ITS - EU standards to accelerate cooperative mobility*. 2014 International Conference on Connected Vehicles and Expo (ICCVE), 681-686.
11. Svensk, P., and Wikström, L. (2013). *Georeferencing methods-A study on how to deal with georeferencing in the ROSATTE Implementation Platform*.
12. T'Siobbel, S., Landwehr, M., Mahiou, R., and Wevers, K. (2011). *ROSATTE: ROad Safety*

ATTributes exchange infrastructure in Europe: deliverable D4.1: Description of applicable and viable data integration methods.

13. Wikström, L., Landwehr, M., Bock, H., Nasr, A., Wevers, K., Svensk, P.-O., T'Siobbel, S., Boterbergh, B., Blaive, L., Hamish, K., et al. (2009). *ROSATTE: ROad Safety ATTributes exchange infrastructure in Europe: deliverable D3.1: Specification of data exchange methods*.
14. Borzacchiello, M.T., Boguslawski, R., and Pignatelli, F. (2016). *Improving accuracy in road safety data exchange for navigation systems: European Union Location Framework Transportation Pilot*.
15. CEN/TC 278. (2018). *Intelligent transport systems — Location Referencing Harmonisation for Urban-ITS — Part 2: Translation methods*.
16. CEN/TC 278. (2018). *Intelligent transport systems — Location Referencing Harmonisation for Urban ITS — State of the art and guidelines*.
17. Adesiyun, A., and Ihs, A. (2009). *Heavy Route D6.4 Final summary report*.
18. Sjögren, L., and Landwehr, M. (2008). *Heavy route D2.1 Available road databases and relevant data*.
19. Hess, S., Quddus, M., Rieser-Schüssler, N., and Daly, A. (2015). *Developing advanced route choice models for heavy goods vehicles using GPS data*. Transport Res E-Log 77, 29-44.
20. Colpaert, P. (2017). *Publishing transport data for maximum reuse*. PhD Thesis, Universiteit Gent.
21. ISO/TC 211. *ISO/TC211 Geographic information/Geomatics*. Available from: https://committee.iso.org/home/tc211.
22. ISO/TC 204. (2017). *ITS Standardization Activities Activities of ISO/TC204 2017*.
23. CEN/TC 278. *CEN/TC278*. Available from: http://www.itsstandards.eu.
24. Interactive instruments GmbH. *ShapeChange*. Available from: http://shapechange.net.
25. International Association of Oil and Gas Producers. *EPSG Geodetic Parameter Registry*. Available from: https://www.epsg-registry.org.
26. Rainer, S., and Jacek, F. (2010). *ROSATTE: ROad Safety ATTributes exchange infrastructure in Europe: D5.4 Aggregated test report including detailed test reports*.
27. ISO/TC 211. (2017). *Report from the Ad hoc group on Location based services*.
28. Kresse, W., Danko, D.M., and Fadaie, K. (2011). *Standardization*. In Springer Handbook of Geographic Information, W. Kresse and D.M. Danko, eds. Berlin, Heidelberg, Springer Berlin Heidelberg), pp 245-271.

Article 1

Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model

International Journal of
Geo-Information

MDPI

Article

Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model

Knut Jetlund [1,2,*], Erling Onstein [2] and Lizhen Huang [2]

1 Norwegian Public Roads Administration, 2318 Hamar, Norway
2 Department of Manufacturing and Civil Engineering, Norwegian University of Science and Technology, 2821 Gjøvik, Norway; erling.onstein@ntnu.no (E.O.); lizhen.huang@ntnu.no (L.H.)
* Correspondence: knut.jetlund@vegvesen.no; Tel.: +47-95962863

Received: 31 January 2019; Accepted: 10 March 2019; Published: 14 March 2019

Abstract: This study aims to improve interoperability between Geographic Information Systems (GIS) and geospatial databases for Intelligent Transport Systems (ITS). Road authorities maintain authoritative information for legal and safe navigation in GIS databases. This information needs to be shared with ITS databases for route planning and navigation, and for use in combination with local knowledge from vehicle sensors. Current solutions for modelling and exchanging geospatial information in the domains of GIS and ITS have been studied and evaluated. Limitations have been pointed out related to usability in the GIS domain and flexibility for representing an evolving real world. A prototype for an improved information exchange model has been developed, based on ISO/TC 211 standards, Model Driven Architecture (MDA), and concepts from the studied solutions. The prototype contains generic models for feature catalogues and features, with implementation schemas in the Geography Markup Language (GML). Results from a case study indicated that the models could be implemented with feature catalogues from the ITS standard ISO 14825 Geographic Data Files (GDF) and the INSPIRE Transport Networks specification. The prototype can be a candidate solution for improved information exchange from GIS databases to ITS databases that are based on the Navigation Data Standard.

Keywords: geographic information systems; intelligent transport systems; information model; model driven architecture

1. Introduction

1.1. Geospatial Information in Intelligent Transport Systems (ITS)

Detailed geospatial information that represents road networks and the surrounding road environment is a critical component for route planning and navigation with Intelligent Transport Systems (ITS) [1]. Advanced Driver Assistance Systems (ADAS) and systems for autonomous driving depend on geospatial information from a variety of sources, as illustrated in Figure 1.

Modern vehicles contain a range of sensors from which they create local geospatial knowledge, and can share this information with map providers, Original Equipment Manufacturers (OEMs), and other road users. However, the local knowledge is neither sufficient for route planning nor for local navigation under challenging conditions, such as fog or snow-covered roads, and must be combined with geospatial information from pre-processed databases covering larger areas [2,3].

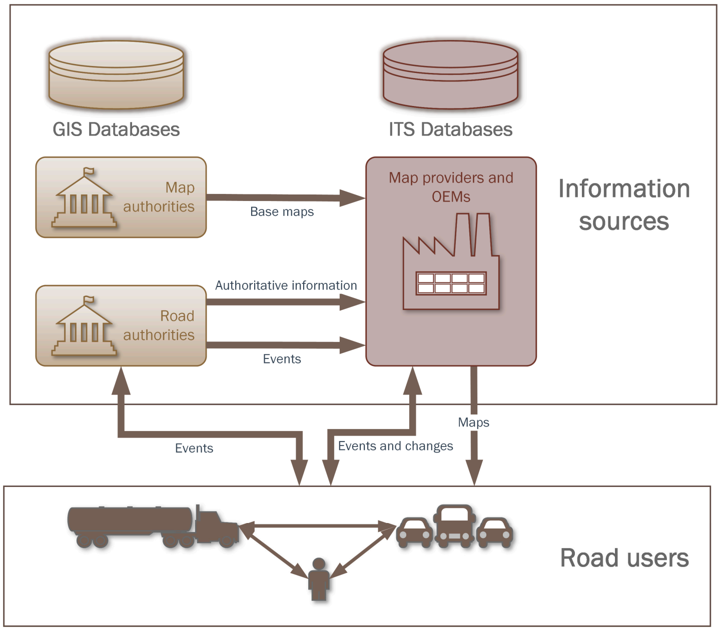


Figure 1. Parts of the flow of geospatial information in Intelligent Transport Systems (ITS).

Commercial map providers and OEMs have created and delivered ITS databases and services for route planning and navigation for several decades and are extending their products to support ADAS and autonomous driving with High Definition (HD) maps. The databases are maintained through data capture from both professional mapping vehicles and private vehicles [4].

However, road authorities are the authoritative source for information needed for legal and safe navigation in the road network. Road authorities set regulations for road use and navigation and maintain geospatial information representing the road network, regulations, events, conditions, and road equipment, often in applications and databases from the domain of Geographic Information Systems (GIS). Furthermore, road authorities may have geospatial information for rural areas where fewer vehicles have travelled, as well as Building Information Models (BIM) with geospatial information representing new roads before the roads are opened. Meanwhile, mapping authorities, of course, maintain base maps of roads and the surrounding environment in GIS databases.

The users of ITS applications for route planning and navigation rely on accurate and updated geospatial information for the complete knowledge needed for legal and safe navigation. Map providers and OEMs need reliable and harmonized mechanisms that can provide them with information from authorities for further sharing with the road users [5], and for simulation and testing [6,7]. Sharing information from authorities may improve the data quality of ITS databases for route planning and navigation and thereby improve public safety, reduce the risk of damage to infrastructure and improve the quality of mobility services. To enable the flow of information, models that describe the real world and specifications for information exchange are needed [5,8].

1.2. Information Modelling

A foundation for information exchange is a common understanding of how data represents the real world by means of a standard way of information modelling [9]. The most often used language for information modelling is Unified Modelling Language (UML), developed by the Object Management Group (OMG) [10,11], who have also developed Model Driven Architecture (MDA) [12]. In MDA, conceptual models are developed at different levels of abstraction, independent of specific implementations. Implementation schemas for file and database formats are derived from the conceptual models, following rules for conversion. MDA gives flexibility for modifying and extending the models and ensures that conceptual models and implementation schemas match each other.

ISO/TC 211 and the Open Geospatial Consortium (OGC) are the main actors concerning international standardization in the GIS domain, and they have based their work on general information technology from UML and MDA [13]. ISO 19103 [14] describes a UML profile, modelling rules and an MDA approach, and ISO 19109 [15] describes a General Feature Model for geospatial information. Other ISO/TC 211 and OGC standards and specifications are based on ISO 19103 and ISO 19109. The UML models are maintained in a common harmonized UML model repository, where concepts defined in one standard are reused in other standards. Rules for conversion from UML to implementation schemas for the standardized Geography Markup Language (GML) format are described in ISO 19136 [16], and implementation schemas for other formats may also be derived in similar ways. Figure 2 illustrates the levels of abstraction in MDA according to ISO 19103.

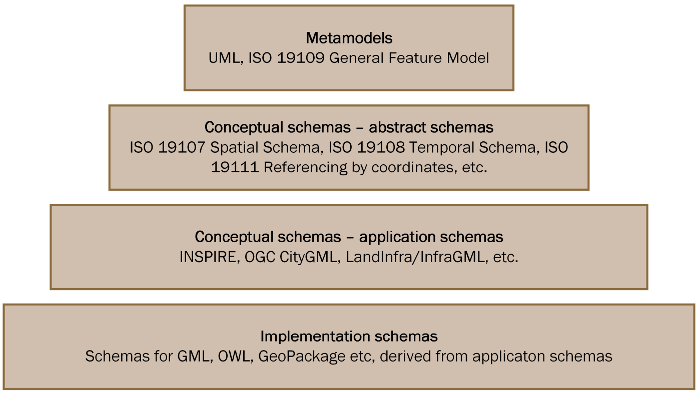


Figure 2. The levels of abstraction in model driven architecture according to ISO 19103 [14].

The approach based on MDA and UML has been successful for developing interoperable standards in the GIS domain. The standards are widely adopted in applications and national and regional information models, including systems and databases for geospatial road-related information maintained by road and mapping authorities. Some examples of specifications based on ISO/TC 211 standards are the European INSPIRE specification on Transport Networks [17] and the OGC specifications CityGML [18] and LandInfra/InfraGML [19–21].

In the ITS domain, the Navigation Data Standard (NDS) and its free part, the Open Lane Model [22], is the standard for storing geospatial road-related information in ITS databases, while ISO 14825 Geographic Data Files (GDF) [23] is the official ISO standard for information exchange between databases with geospatial information for road navigation. Road and mapping authorities maintain some of the information described in these standards in GIS databases. However, the descriptions in the ITS standards are based on other information modelling principles than ISO/TC 211 standards. For the information exchange from road and mapping authorities' GIS databases to ITS databases, the information requirements described in standards from the ITS domain must be described according to GIS standards and specifications.

Due to the lack of harmonized standards for the GIS and ITS domains, little information is currently exchanged between the two domains, and the exchange is likely to be based on proprietary solutions from each data owner. There has been some progress in Europe though, through the European Union Location Framework Transportation Pilot [5] and the TN-ITS GO project [24], where several European countries have implemented or plan to implement the specification TN-ITS [25]. However, the specification covers only a limited set of regulations for road use and navigation.

1.3. Research Question

The research question in this study is how to improve existing methods for exchange of geospatial information from road and mapping authorities' GIS databases to geospatial ITS databases, illustrated by the three right-headed horizontal arrows in Figure 1.

2. Materials and Methods

2.1. Research Design

The research described in this paper was based on five steps, illustrated in Figure 3. The state of the art of solutions for information modelling and exchange, in and between the domains of GIS and ITS, was studied and evaluated in step one and two. A prototype for an improved information exchange model was developed in step three. Schema files for implementing the prototype in the GML format were derived from the UML models in step four, following the conversion rules from ISO 19136 [16] and by using the conversion software ShapeChange [26]. Finally, the model was tested and demonstrated in a case study with different feature catalogues in step five.

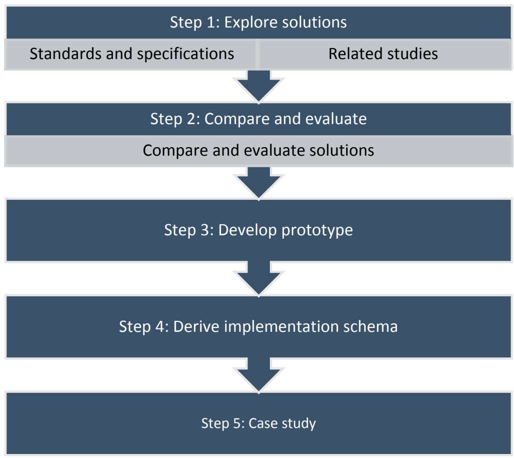


Figure 3. Steps in the research.

2.2. State of the Art

2.2.1. Literature Search

Literature presumed relevant for answering the research question was identified by using the search engines Oria and Google Scholar. Searches were based on keywords related to GIS such as "geospatial information," "geographic information," and "ISO/TC 211," combined with keywords related to ITS such as "Intelligent Transport Systems," "GDF," and "Navigation Data Standard." Searches in standards catalogues identified additional standards and specifications. A shorter list of literature was derived from the search results by studying titles and abstracts, and finally, the most relevant literature was selected by more detailed studies of the content. The results from the literature search included research articles, handbooks, standards, and specifications describing relevant solutions for information exchange.

2.2.2. Overview of Standards and Specifications

Several actors from the domains of GIS and ITS are involved in standardization concerning the modelling and exchange of geospatial road-related information. Of primary interest for this study are some official standards and specifications from the International Organization for Standardization (ISO) and the European Committee for Standardization (CEN), as well as open standards and specifications from industry actors such as the Navigation Data Standard (NDS) Association and OGC. Table 1. presents an overview of identified standards and specifications.

Table 1. Identified standards and specifications for geospatial road-related Information.

Standard/Specification	Title
GDF	ISO 14825:2011 Geographic Data Files (GDF) – GDF 5.0 [23] ISO/DIS 20524-1 Geographic Data Files (GDF) – GDF5.1 – Part 1: Application independent map data shared between multiple sources [27] ISO/CD 20524-2 Geographic Data Files (GDF) – GDF5.1 – Part 2: Map data used in automated driving systems, cooperative ITS, and multi-modal transport [28]
NDS Open Lane Model	Navigation Data Standard (NDS) Open Lane Model version 1.0 [22]
OpenDRIVE	OpenDRIVE version 1.5 [29]
TN-ITS	FprCEN/TS 17268 data exchange on changes in road attributes [25] Transport Network Intelligent Transport Systems (TN-ITS)
OpenTNF	Open Transport Network Format (OpenTNF) 1.0 [30]
INSPIRE TN	INSPIRE Data Specification on Transport Networks [17]
CityGML	OGC City Geography Markup Language (CityGML) Encoding Standard version 2.0 [18]
LandInfra/InfraGML	OGC LandInfra and InfraGML 1.0 Encoding Standard [19–21]
DATEX II	CEN/TS 16157 DATEX II data exchange specification for traffic management and information [31–34]
TPEG2	ISO/TS 21219 Traffic and Travel Information (TTI) via Transport Protocol Expert Group, Generation 2 (TPEG2) [35–38]

2.2.3. Geographic Data Files (GDF)

GDF version 5.0 [23] is the ISO standard for exchange of information between databases with geospatial information for road navigation, and is referred to by several authors as the core standard for this purpose. Among them is [8], where the need for standards for the digital road network supporting automated driving is emphasized. GDF describes the network, regulations, and the surrounding road environment with a generic model (the feature model), a general topology model and catalogues with more specific models of features, relationships, and attributes. A new version of GDF is under development, with an extended scope that includes requirements from applications for Cooperative Intelligent Transport Systems (C-ITS), multi-modal transportation, and systems for automated driving [27,28]. Actions were taken in the development of GDF version 4.0 to harmonize the standard with ISO/TC 211 standards, but the work was not completed [39,40].

2.2.4. Navigation Data Standard (NDS) Open Lane Model

The Navigation Data Standard (NDS) is developed by the Navigation Data Standard Association, with a broad representation of OEMs and providers of services, and is considered an essential standard for geospatial ITS databases. While GDF is developed for information exchange, the scope of NDS is storage of vehicle navigation data in a size-efficient and standardized way. The NDS Open Lane Model [22] is an open specification that contains a limited set of features and functions from NDS, covering similar contexts as GDF. Of the 16 building-blocks in the complete NDS, the Open Lane Model is limited to three core blocks for navigation and map display: routing with the network model; lanes with the information for navigation at lane-level; and shared data with metadata [4].

2.2.5. OpenDRIVE

The OpenDRIVE specification [29] is an open XML file format for road networks and is described as the de facto standard for driving simulation [7]. The specification defines a navigable road network with centerlines, lanes, and road features, including marking, signals, and more. Some examples of the use of the OpenDRIVE specification are described in [6,7,41].

2.2.6. Transport Network Intelligent Transport Systems (TN-ITS)

The European TN-ITS specification [25] was developed in cooperation between road authorities and commercial map providers in the ROSATTE [42] and TN-ITS [5] projects, and was later standardized by CEN/TC 278. The scope of the specification is the exchange of changes in static road-related information, while the network and topology are outside of the scope. The information model is generic, with types and content of features and properties defined in extendible external code lists.

2.2.7. Open Transport Network Format (OpenTNF)

The open specification OpenTNF [30] describes a generic model for exchange of transport network related information. The specification contains a network model, a generic model of road features and road feature properties, and a model of feature catalogues. OpenTNF is described in [43] as a candidate model for exchanging road asset information between stakeholders in road construction and maintenance.

2.2.8. INSPIRE Transport Networks (INSPIRE TN)

The European INSPIRE Directive [44] defines the framework for spatial information for a range of themes in Europe. The transport networks (TN) theme [17] describes a multi-modal transport network based on the INSPIRE Generic Network Model (GNM) [45], with specific network elements and network properties for each transportation mode. The need for harmonized specifications for information that shall be provided according to the European ITS Directive [46,47] are discussed in [48], and the authors point at the INSPIRE TN specification as one possibility. INSPIRE TN is also described in [43] as a candidate model for the exchange of transport networks information outside of Europe.

2.2.9. CityGML and LandInfra/InfraGML

OGC has developed a range of specifications based on ISO/TC 211 standards. Two specifications that are relevant for this paper are CityGML [18], which defines a model for 3D city models, and LandInfra/InfraGML [19–21], which defines models for land and civil engineering infrastructure facilities, including road infrastructure. A model for detailed modelling of streets by extending CityGML is described in [49].

2.2.10. DATEX II and TPEG2

The scope of the European data exchange specification for traffic management and information (DATEX II) series of standards is dynamic information for traffic and travel. The European Commission has referred to DATEX II as the preferred method for providing real-time information according to the ITS Directive [46,47]. The scope of the Transport Protocol Expert Group, Generation 2 (TPEG2) series of standards is the broadcasting of traffic and travel information, mainly dynamic information. As the TPEG2 and DATEX II series of standards overlap to some degree, rules for converting DATEX II messages to TPEG2 messages have been defined through the EasyWay project [50]. A model for combining dynamic traffic information with the network from an NDS dataset is described in [51].

2.2.11. Other Research and Solutions

Several research projects have worked with specific models and solutions for geospatial ITS databases. Possible models for HD maps are described at an overview level in [1,2]. Conceptual models for transportation networks on different topology levels similar to GDF are described in [52,53], while [54] describes a conceptual UML model for multimodal transport. None of these describes mechanisms for information exchange. Solutions for dynamic update of ITS databases in vehicles from server databases are described in [3,55,56], and ref. [3] includes an overview of research on incremental map updates as well. A model for lane-level navigation is described in [57], and ref. [7] describes a simplified model for conversion between different ITS formats, such as OpenDRIVE and NDS. Finally, a specification of elements that make up a model for automated driving systems is described in [58].

2.3. Evaluation of Solutions

Identified relevant solutions were compared and evaluated against a set of requirements, listed in Table 2. The purpose of the requirements was twofold: Firstly, to evaluate the usability of the specifications for exchanging information from GIS databases, and secondly to evaluate the flexibility of the specifications for exchanging information representing an evolving real world.

Table 2. Requirements used in the evaluation of solutions.

Requirement	Description
ISO/TC 211 Model Driven Architecture (MDA)	The ISO/TC 211 MDA approach is the foundation of information modelling in the GIS domain. Adapting an information model based on a different approach for use in GIS may be a complicated task that requires fundamental changes. Therefore, solutions for exchanging information from GIS databases should be based on the ISO/TC 211 MDA approach.
GIS exchange format	Using a familiar exchange format from the GIS domain, such as GML or GeoPackage, is vital to avoid additional conversions and, thereby, to enable efficient exchange from GIS databases.
Feature catalogue	Advanced Driver Assistance Systems (ADAS) and systems for automated driving need a range of features for legal and safe navigation, including regulation features such as speed limits and other features such as lane dividers. Unambiguous descriptions of these classifications of the real world in a feature catalogue are essential for a common understanding of the exchanged information.
Feature catalogue exchange model	A feature catalogue exchange model is needed for sharing the classifications of the real world described in the feature catalogue with the users of the information. Furthermore, as the real world is changing (e.g., if a new valid value for speed limits is introduced, a new type of sign or a new kind of access regulation is introduced), the feature catalogue must also be modified and shared with receivers of the exchanged information. Using a feature catalogue exchange model for exchanging the feature catalogue together with the features gives the flexibility of maintaining a dynamic feature catalogue outside of the primary standard or specification. Contrary to this, a standardized feature catalogue in an ISO standard cannot be modified without revising the standard.
Network model	A navigable digital network is the foundation of route planning and navigation and must have both geometry and topology of the network elements, together with mechanisms for relating features to the network.
Generic feature exchange model	Feature exchange can either be based on specific feature models precisely as they are modelled in the feature catalogues, or on a generic feature model that refers to external descriptions of features and their properties in feature catalogues. The combination of a feature catalogue exchange model and a generic feature exchange model gives the advantage of keeping the main feature exchange model stable while the feature catalogue is modified, and the feature exchange model may also be used with different feature catalogues.

Table 3 presents an evaluation of studied solutions against the requirements described in Table 2. Only standards and specifications are included, as none of the other solutions described in Section 2.2.11 is mature enough to be considered. The evaluation is generalized to a simple yes or no for each requirement, to simplify the presentation.

Table 3. Evaluation of studied solutions.

Solution	ISO/TC 211 MDA	GIS Exchange Format	Feature Catalogue	Feature Catalogue Exchange Model	Network Model	Generic Feature Exchange Model
GDF	No	No	Yes	Yes	Yes	Yes
NDS Open Lane Model	No	No	Yes	No	Yes	No
OpenDRIVE	No	No	Yes	No	Yes	No
TN-ITS	Yes	Yes	No	No	No	Yes
OpenTNF	No	Yes	No	Yes	Yes	Yes
INSPIRE TN	Yes	Yes	Yes	No	Yes	No
CityGML	Yes	Yes	Yes	No	No	No
LandInfra/InfraGML	Yes	Yes	Yes	No	No	No
DATEX II	No	No	Yes	No	No	No
TPEG2	No	No	Yes	No	No	No

The analysis shows that none of the solutions meets all requirements. The GDF standard covers four out of six requirements, but as it is not based on ISO/TC 211 MDA and does not use an exchange format from the GIS domain, it cannot easily be implemented in a GIS. Like GDF, neither the NDS Open Lane Model nor OpenDRIVE is based on ISO/TC 211 MDA, and besides, they are missing a feature catalogue model and a generic exchange model. Nevertheless, the NDS Open Lane Model is essential for the further work, as it defines the model of the ITS database to which the information shall be exchanged. For DATEX II and TPEG2, the situation is different, as they are based on similar MDA approaches as ISO/TC 211 standards and may be converted to ISO/TC 211 MDA. However, they contain only the feature catalogue, and their main scope is limited to dynamic information.

Of the solutions that are based on ISO/TC 211 MDA, INSPIRE TN is a promising candidate, with a network model and feature catalogue described according to ISO/TC 211 MDA and with implementation schemas in the GML format. However, INSPIRE TN does not have a feature catalogue exchange model or a generic feature exchange model. Another candidate is TN-ITS with a generic feature exchange model described according to ISO/TC 211 MDA, but the specification does not have a network model, a feature catalogue, or a feature catalogue exchange model. However, both INSPIRE and TN-ITS may be extended to meet the remaining requirements. OpenTNF is also a promising candidate, even though it is not based on ISO/TC 211 MDA. The specification includes a network model, a feature catalogue exchange model and a generic feature exchange model and uses familiar GIS formats for exchange (SQLite and OGC GeoPackage). It is based on a conceptual UML model, uses some concepts from ISO/TC 211 standards, and the models may be modified to ISO/TC 211 MDA without fundamental changes.

The two last specifications based on ISO/TC 211 MDA are CityGML and LandInfra/InfraGML. The main scope of these specifications is not information for navigation, but rather visualization, planning, construction, and asset management. The CityGML specification is specific on this issue and refers to GDF for network topology. However, the specifications do contain road centerlines and road equipment that are needed for navigation, and most important, the information from construction projects may be delivered to road authorities and map providers for further preparation in network databases. Information exchange between CityGML, LandInfra/InfraGML, and ITS databases will probably be essential for maintaining HD maps in the future.

3. Results

3.1. A Generic Model for the Exchange of Road-Related Geospatial Information

As none of the studied solutions meets all of the requirements from Table 2, a prototype of an improved solution was developed. The goal was to support the information exchange illustrated by the three right-headed horizontal arrows in Figure 1 by fulfilling all requirements from Table 2. Furthermore, the prototype should enable the exchange of feature and relationship types defined in ISO 14825 GDF, as this is described as the core model for geospatial information for ITS applications. Exchange of feature types described in other feature catalogues should also be supported, such as INSPIRE TN. The prototype was based on concepts from the most promising solutions from the evaluation: GDF, INSPIRE-TN, TN-ITS, and OpenTNF.

The concepts of the prototype are illustrated in Figure 4, with three application schemas based on ISO 19109: the Feature Catalogue, the Feature Catalogue Exchange Model, and the Feature Exchange Model, and implementation schemas that are derived from the application schemas.

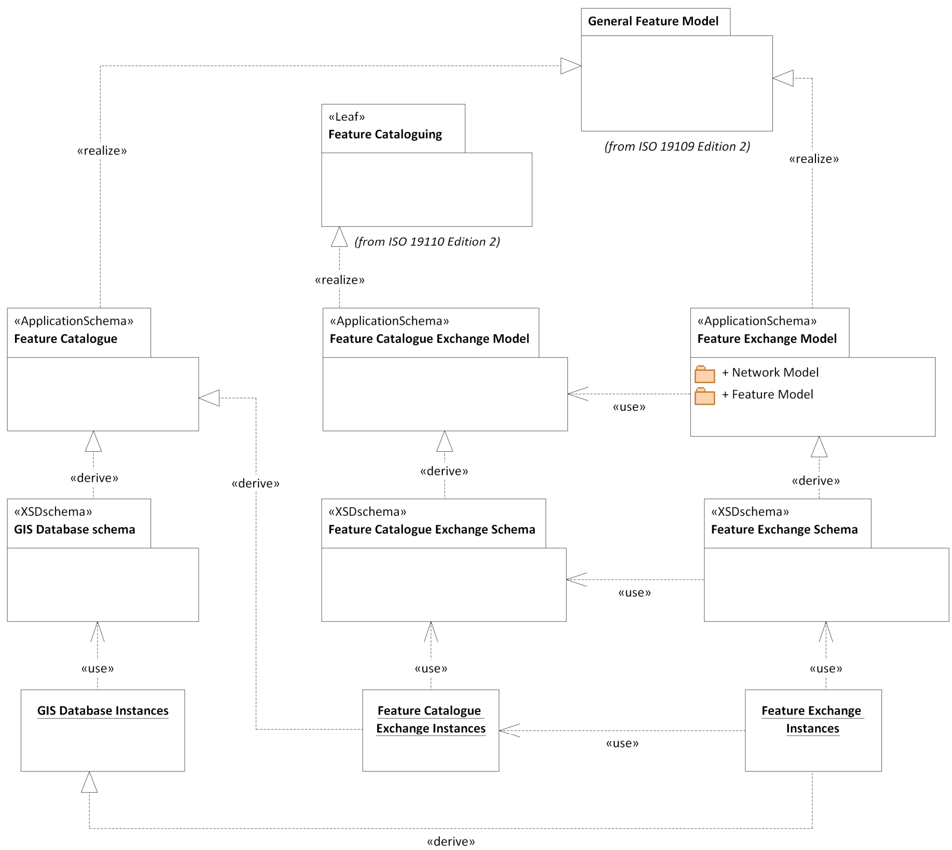


Figure 4. Prototype concepts.

The Feature Catalogue application schema contains specific models for individual feature types. A feature type is defined in ISO 19101 as a class of features having common characteristics, while a feature is an abstraction of a real-world phenomenon [59]. Similar to this, GDF defines a feature as a representation of a real-world geographic object [23]. In the scope of ITS databases for route planning and navigation, a feature type will be a class of features related to the road, such as speed limits and other regulations, or road equipment and the surrounding road environment, such as signs and railings. As one of the goals was to enable the exchange of information based on different and dynamic feature catalogues, no specific feature types were defined in the prototype. Instead, any feature catalogue with feature types modelled in application schemas according to ISO 19109 can be used.

An implementation schema for a GIS database can be derived from the Feature Catalogue application schema, and the original GIS Database Instances (the records in the database) can be stored and maintained in a GIS database according to the implementation schema. The original GIS database may also have a different implementation schema (e.g., for a national road database), and the instances are then converted to feature types in the specified feature catalogue for exchange (e.g., to the INSPIRE TN feature types).

The Feature Catalogue Exchange Model defines the structure for exchange of a feature catalogue. This approach enables the use of dynamic feature catalogues, and the receiver of the information will have access to an updated version of the feature catalogue. Implementation schemas are derived from the application schema, while the Feature Catalogue Instances (the representation of each feature type and their characteristics) are derived from the Feature Catalogue application schema for exchange.

The Feature Exchange Model application schema defines the structure of the actual instances that shall be exchanged and consists of two main parts: The Network Model, with a navigable network, and the Feature Model, with the generic model of features and their characteristics. Each Feature Exchange Instance refers to a Feature Catalogue Exchange instance that defines the model for the feature. Implementation schemas are derived from the application schema, while the Feature Exchange Instances are derived from the GIS Database instances.

GML was used as the implementation format for both the Feature Catalogue Exchange Model and the Feature Exchange Model in the study. Implementation schemas for other exchange formats (e.g., GeoPackage), may also be derived from the application schemas.

3.2. Feature Catalogue Exchange Model

ISO 19110 [60] describes an abstract conceptual model of a feature catalogue for geospatial information. The model is a realization of the General Feature Model from ISO 19109, and defines concepts for the feature catalogue, feature types in the catalogue, and feature type characteristics such as attributes and association roles. As ISO/TC 211 MDA, with reuse of concepts from ISO/TC 211 standards, is the first requirement in Table 2, the Feature Catalogue Exchange Model in the prototype was developed as a realization of the abstract model from ISO 19110. The model is illustrated in Figure 5.

The core concept is the class FeatureCatalogue, where an instance must include one or more feature types as instances of the class FeatureType. Attributes of feature types are instances of the class FeatureAttribute and can be either local attribute types for a specific feature type, or global attribute types for use on several feature types in the feature catalogue. Attributes for identifiers and names are examples of properties that may be reused on several feature types. The content of an attribute can either be a simple value, a value from a code list, or a structured value based on a data type. Code lists are instances of the class ValueDomain, code list values are instances of the class ListedValue, and structured values are instances of the class DataType. Associations between feature types are instances of the class FeatureAssociation and must be related to two instances of the class AssociationRole. Like attributes, association roles are also defined as a property of feature types. As FeatureAssociation is a subtype of FeatureType, an association may have attributes, as in the Relationship concept in GDF.

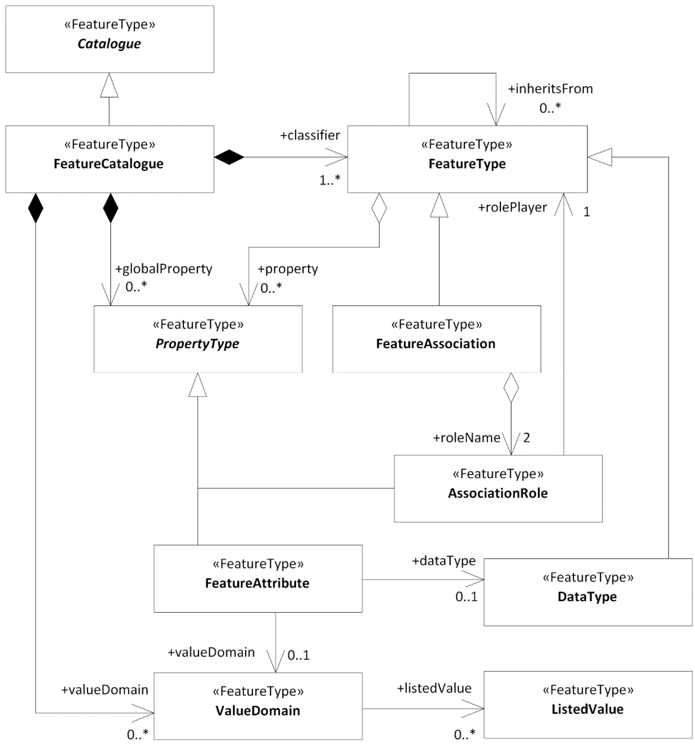


Figure 5. Overview of the Feature Catalogue Exchange Model in the prototype.

The Feature Catalogue Exchange Model contains some modifications from the abstract model in ISO 19110. First, all aggregations and composite associations are set unidirectional, to avoid duplication of information. The navigable direction is set from the "whole" to the "part" (e.g., from FeatureType to FeatureAttribute), which fits the hierarchical structure in an XML implementation. For a relational database, they might instead be set navigable in the opposite direction. Second, a specific class in ISO 19110 for describing inheritance relations between feature types has been replaced with a self-association on the class FeatureType, to simplify the model. Third, the concepts for connecting global properties to feature classes have been simplified. Finally, the classes for value domains and data types have been added to the model based on experiences from the case study. These classes do not exist in ISO 19110.

3.3. Feature Exchange Model

The network model in the INSPIRE GNM is based on a link and node structure, which is a common way of modelling a network in GIS [61]. Furthermore, the network models in OpenTNF and the INSPIRE GNM are very similar, and the TN-ITS specification also refers to INSPIRE for network representation. The Network Model part of the Feature Exchange Model in the prototype was, therefore, developed with the INSPIRE GNM as a foundation. The model is illustrated in Figure 6.

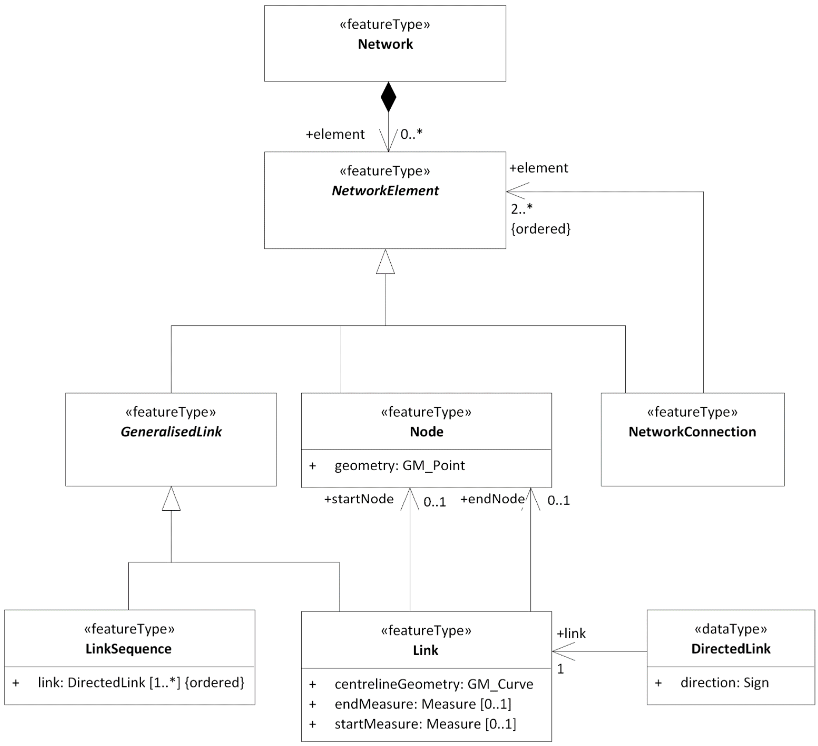


Figure 6. The Network Model in the prototype.

The core concept of the Network Model is the class Network, which is an aggregation of the abstract class NetworkElement, with non-abstract subclasses for Link, Node, LinkSequence, and NetworkConnections. The class NetworkConnection is used for connecting several networks (e.g., for connecting networks in different countries, counties, or municipalities, or even networks representing different transportation modes). The classes that represent the actual network elements are Link, which represents the centerline geometry of road segments, and Node, which represents the point geometry of the intersections where links meet. Furthermore, an ordered set of Link instances can be described as an instance of the LinkSequence class. While each link usually goes from intersection to intersection, a link sequence can represent longer sections. Each Link instance must have a specified direction in the LinkSequence, relative to the direction of the centerline geometry.

The Network Model part of the prototype has two modifications from the INSPIRE GNM. First, the INSPIRE GNM is an abstract model that is further specialized into implementable classes in the INSPIRE theme models, such as the Road Network Model in the INSPIRE TN specification. This specialization is not necessary for the generic approach in the prototype, and instead, the classes from the INSPIRE GNM were set as non-abstract. Secondly, all associations were set unidirectional in this model as well, in order to avoid duplication of information.

The second part of the Feature Exchange Model is the Feature Model that describes the generic concepts for exchanging features and their characteristics. Table 3 shows that generic exchange models are defined in OpenTNF, TN-ITS, and GDF. Of these, the TN-ITS specification is the only one that is based on ISO/TC 211 MDA. The Feature Model in the prototype was, therefore, developed based on the TN-ITS specification. The model is illustrated in Figure 7.

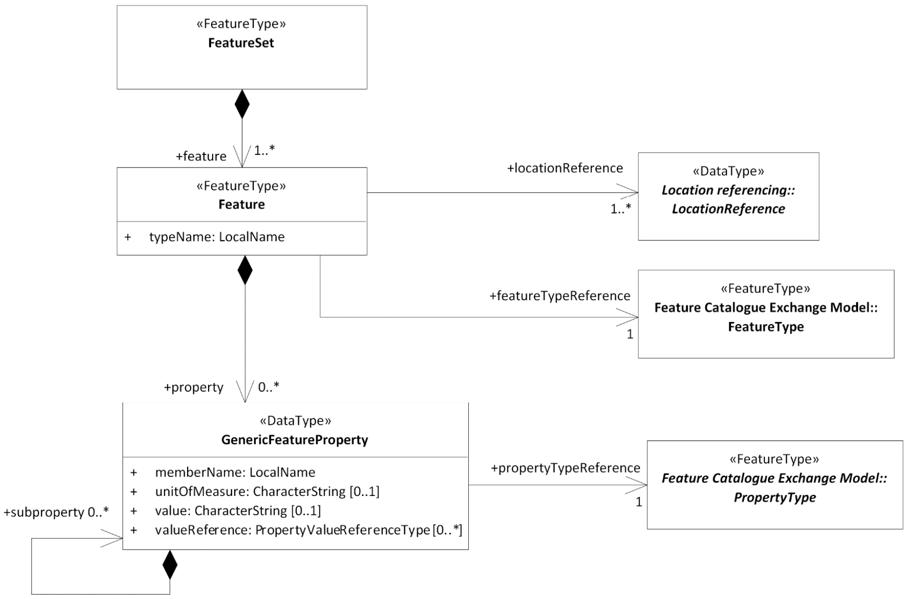


Figure 7. The Feature Model in the prototype.

The core concept in the Feature Model is the FeatureSet class, where an instance must include at least one Feature instance. A Feature instance must have one or more location references (e.g., coordinate-based geometry, linear referencing, or a dynamic location reference, as defined in TN-ITS [25]). Furthermore, a Feature instance can have properties as instances of the GenericFeatureProperty class. As defined in the Feature Catalogue Exchange Model, a property can be an attribute with a simple value, such as a speed limit value; a code list value, such as the type of railing; a structured value, based on a data type; or it can be an association to another feature.

The Feature Model is generic, without specific attributes or association for any specific feature type. Instead, mandatory associations to FeatureType and PropertyType instances in the Feature Catalogue are used for defining the model of each Feature and GenericFeatureProperty instance. This generic concept gives flexibility for exchanging any feature type with any attribute type or association, given that it is described as Feature Catalogue Exchange Instances that are derived from a Feature Catalogue.

4. Case Study

4.1. Purpose and Workflow

The use case for the case study was that road or mapping authorities wish to set up one single service that can provide geospatial road-related information for different user groups in the ITS domain. One group of information receivers are map providers that need updated geospatial information according to the GDF Standard, for updating ITS databases that are based on the NDS Open Lane Model. Another group are public transport planners that need an updated road network with network properties according to the INSPIRE TN specification. The purpose of the case study was to test and validate the prototype based on this use case and to demonstrate how it can be implemented with feature types from different feature catalogues.

The case study was conducted in five steps, as illustrated in Figure 8. A selection of feature types was chosen from the GDF standard as the core standard for geospatial databases for ITS, while a selection of feature types from the INSPIRE TN specification was chosen to represent information specified by other feature catalogues. The specific processes and results for each feature catalogue are further described in Sections 4.2 and 4.3.

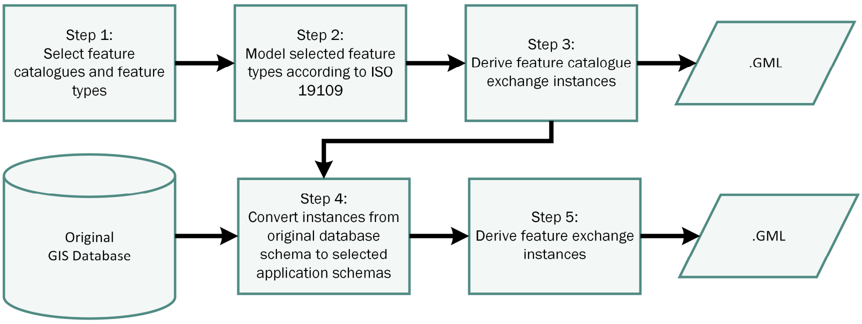


Figure 8. Steps in the case study.

The derivation of Feature Catalogue Instances from the Feature Catalogue application schemas was solved with a script developed in the UML software, Enterprise Architect [62]. This step was the implementation of the dashed “derive” arrow between the Feature Catalogue application schema and the Feature Catalogue Instances in Figure 4, and the result was GML files with Feature Catalogue Instances according to the Feature Catalogue Exchange Schema.

The original GIS database was a national road database where information is maintained according to an internal database schema. Instances were converted from the internal database schema to the application schema with the ETL (Extract, Transform, Load) software, FME [63], and exported to GML files as Feature Exchange Instances. This last step was the implementation of the dashed “derive” arrow between GIS Database Instances and Feature Exchange Instances in Figure 4.

GML Application Schemas and GML files from the case study are available online at https://github.com/jetgeo/ITSGML.

4.2. INSPIRE TN

Five feature types were selected from the INSPIRE TN specification: The road network elements RoadLink, RoadNode, and RoadLinkSequence, and the network properties SpeedLimit and FunctionalRoadClass, as shown in Figure 9. These feature types were selected as they were considered particularly important information for route planning and navigation. The models were already conformant with ISO 19109, and the complete feature catalogue for the INSPIRE TN Specification could be derived to Feature Catalogue Exchange Instances in GML format without modification.

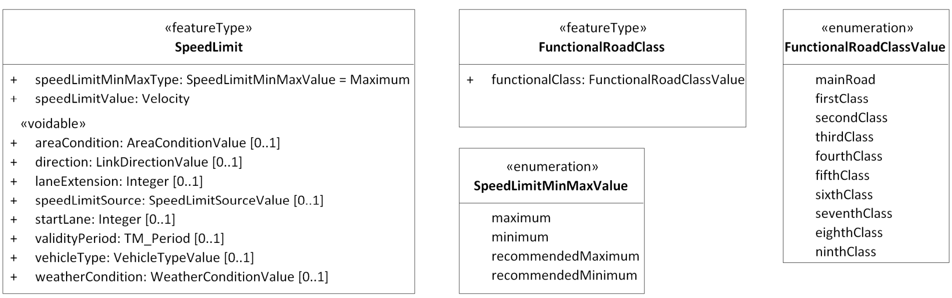


Figure 9. INSPIRE SpeedLimit and FunctionalRoadClass.

An extract of the derived Feature Catalogue Exchange GML is shown in Figure 10 with the instance for the FunctionalRoadClass feature type. Inheritance of characteristics from the TransportPropery class in the UML model is represented by the inheritsFrom tag and a link to the instance for that class. As shown in the UML model in Figure 9, FunctionalRoadClass has one attribute (functionalClass) with values from the code list FunctionalRoadClassValue. This code list is referred to in the GML file in Figure 10 with the valueDomain tag.

```
<itsgml:classifier>
    <itsgml:FeatureType gml:id="INSPIRE_TN.FunctionalRoadClass">
        <itsgml:typeName>FunctionalRoadClass</itsgml:typeName>
        <itsgml:isAbstract>false</itsgml:isAbstract>
        <itsgml:inheritsFrom xlink:href="INSPIRE_TN.gml#INSPIRE_TN.TransportProperty" />
        <itsgml:property>
            <itsgml:FeatureAttribute gml:id="INSPIRE_TN.FunctionalRoadClass.functionalClass">
                <itsgml:memberName>functionalClass</itsgml:memberName>
                <itsgml:cardinality> … </itsgml:cardinality>
                <itsgml:valueType>FunctionalRoadClassValue</itsgml:valueType>
                <itsgml:valueDomain xlink:href="INSPIRE_TN.gml#INSPIRE_TN.FunctionalRoadClassValue" />
            </itsgml:FeatureAttribute>
        </itsgml:property>
    </itsgml:FeatureType>
</itsgml:classifier>
```

Figure 10. Extract of the INSPIRE TN Feature Catalogue Exchange GML file.

Extracts of the Feature Exchange GML file with instances of the feature types FunctionalRoadClass and SpeedLimit are shown in Figures 11 and 12 References to definitions of feature types and attributes are represented by the featureTypeReference and propertyTypeReference tags. In the examples, these references are set to a local file (TransportNetworks.GML), but for real implementation, they would likely be set to a URL instead.

```
<itsgml:Feature gml:id="vegvesen.no.nvdb.568225673_2">
    <itsgml:typeName>FunctionalRoadClass</itsgml:typeName>
    <itsgml:validFrom>2018-08-31</itsgml:validFrom>
    <itsgml:beginLifespanVersion>2018-08-31T00:00:00</itsgml:beginLifespanVersion>
    <itsgml:featureTypeReference xlink:href="TransportNetworks.gml#TransportNetworks.FunctionalRoadClass"/>
    <itsgml:locationReference> … </itsgml:locationReference>
    <itsgml:locationReference> … </itsgml:locationReference>
    <itsgml:property>
        <itsgml:GenericFeatureProperty>
            <itsgml:memberName>functionalClass</itsgml:memberName>
            <itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
            <itsgml:propertyTypeReference xlink:href="TransportNetworks.gml#TransportNetworks.FunctionalRoadClass.functionalClass"/>
            <itsgml:valueReference xlink:href="TransportNetworks.gml#TransportNetworks.FunctionalRoadClassValue.thirdClass"/>
        </itsgml:GenericFeatureProperty>
    </itsgml:property>
</itsgml:Feature>
```

Figure 11. Extract of the INSPIRE TN Feature Exchange GML file with an instance of the FunctionalRoadClass feature type.

The FunctionalRoadClass instance in Figure 11 represents a road section classified as functional road class "thirdClass", described by the valueReference tag. The SpeedLimit instance in Figure 12 represents a road section with a maximum speed limit of 50 km/h. The type of speed limit (maximum) is described by the valueReference tag, while the speed limit is described by the value tag.

```xml
<itsgml:Feature gml:id="vegvesen.no.nvdb.88306515_0">
    <itsgml:typeName>SpeedLimit</itsgml:typeName>
    <itsgml:validFrom>1980-01-01</itsgml:validFrom>
    <itsgml:beginLifespanVersion>1980-01-01T00:00:00</itsgml:beginLifespanVersion>
    <itsgml:featureTypeReference xlink:href="TransportNetworks.gml#TransportNetworks.SpeedLimit"/>
    <itsgml:locationReference> … </itsgml:locationReference>
    <itsgml:locationReference> … </itsgml:locationReference>
    <itsgml:property>
        <itsgml:GenericFeatureProperty>
            <itsgml:memberName>speedLimitMinMaxType</itsgml:memberName>
            <itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
            <itsgml:propertyTypeReference xlink:href="TransportNetworks.gml#TransportNetworks.SpeedLimit.speedLimitMinMaxType"/>
            <itsgml:valueReference xlink:href="TransportNetworks.gml#TransportNetworks.SpeedLimitMinMaxValue.maximum"/>
        </itsgml:GenericFeatureProperty>
    </itsgml:property>
    <itsgml:property>
        <itsgml:GenericFeatureProperty>
            <itsgml:memberName>speedLimitValue</itsgml:memberName>
            <itsgml:propertyStructure>simple</itsgml:propertyStructure>
            <itsgml:propertyTypeReference xlink:href="TransportNetworks.gml#TransportNetworks.SpeedLimit.speedLimitValue"/>
            <itsgml:value>50</itsgml:value>
            <itsgml:unitOfMeasure>km/h</itsgml:unitOfMeasure>
        </itsgml:GenericFeatureProperty>
    </itsgml:property>
</itsgml:Feature>
```

Figure 12. Extract of the INSPIRE TN Feature Exchange GML file with an instance of a SpeedLimit feature type.

4.3. GDF

From GDF, the feature types RoadElement and Junction and the relationship ProhibitedManoeuvre were selected as examples of network and network regulations, and the feature type PedestrianCrossing was selected as an example of information related to the network. The models had to be modified slightly to conform with ISO 19109, primarily by using stereotypes and datatypes defined in ISO 19103 and ISO 19109.

Figure 13 shows the ISO 19109 compliant model of GDF RoadElement, Junction, and Manoeuvre, where the network feature types RoadElement and Junction are modelled as subtypes of the general network elements Link and Node from the Network Model part of the Generic Feature Exchange Model. Figure 14 shows the ISO 19109 compliant model of GDF PedestrianCrossing, with a reduced number of attributes for the case study.

The modified extract of the feature, relationship, and attribute catalogues from GDF was derived to Feature Catalogue Exchange Instances in GML format. Figure 15 shows an extract of the Feature Catalogue Exchange GML file with the instance for the abstract Manoeuvre feature type with associations to RoadElement and Junction, and the ProhibitedManoeuvre feature type with an inheritance from Manoeuvre.

Figures 16 and 17 show extracts of the Feature Exchange GML file with feature instances of the feature types ProhibitedManoeuvre and PedestrianCrossing.

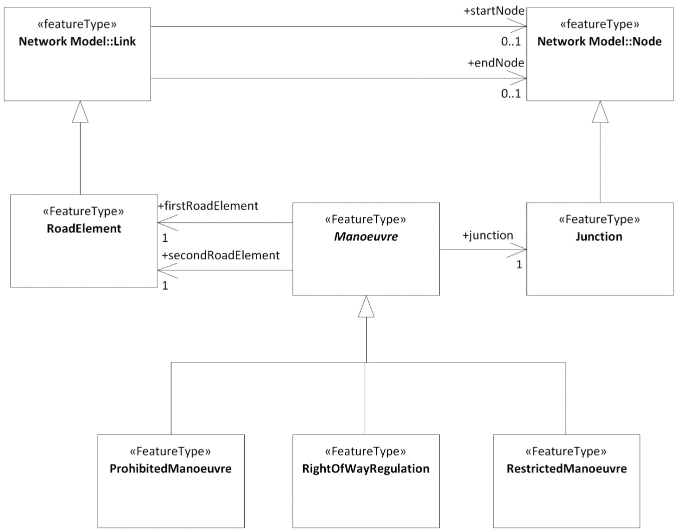


Figure 13. ISO 19109 compliant model of GDF 5.0 RoadElement, Junction, and Manoeuvre.

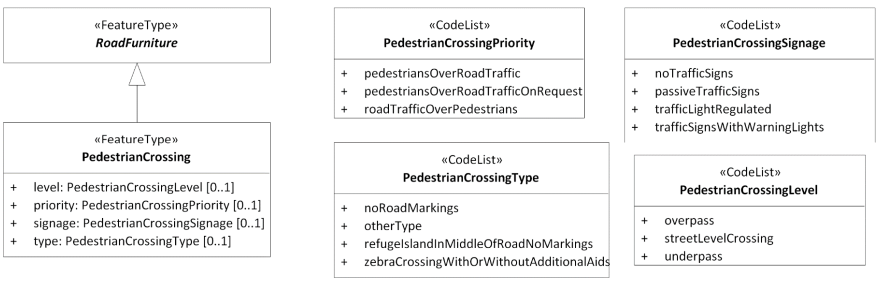


Figure 14. ISO 19109 compliant model of GDF 5.0 PedestrianCrossing.

```
<itsgml:classifier>
    <itsgml:FeatureType gml:id="GDF.Manoeuvre">
        <itsgml:typeName>Manoeuvre</itsgml:typeName>
        <itsgml:isAbstract>true</itsgml:isAbstract>
        <itsgml:property>
            <itsgml:AssociationRole gml:id="GDF.AssociationRoles.Manoeuvre.RoadElement.firstRoadElement"> ...
            </itsgml:AssociationRole>
        </itsgml:property>
        <itsgml:property>
            <itsgml:AssociationRole gml:id="GDF.AssociationRoles.Manoeuvre.RoadElement.secondRoadElement"> ...
            </itsgml:AssociationRole>
        </itsgml:property>
        <itsgml:property>
            <itsgml:AssociationRole gml:id="GDF.AssociationRoles.Manoeuvre.Junction.junction"> ...
            </itsgml:AssociationRole>
        </itsgml:property>
    </itsgml:FeatureType>
</itsgml:classifier>
<itsgml:classifier>
    <itsgml:FeatureType gml:id="GDF.ProhibitedManoeuvre">
        <itsgml:typeName>ProhibitedManoeuvre</itsgml:typeName>
        <itsgml:isAbstract>false</itsgml:isAbstract>
        <itsgml:inheritsFrom xlink:href="GDF.gml#GDF.Manoeuvre" />
    </itsgml:FeatureType>
</itsgml:classifier>
```

Figure 15. Extract of the GDF Feature Catalogue Exchange GML file.

```
<itsgml:Feature gml:id="vegvesen.no.nvdb.114169011_0">
	<itsgml:typeName>ProhibitedManoeuvre</itsgml:typeName>
	<itsgml:validFrom xsi:nil="true"/>
	<itsgml:beginLifespanVersion>1950-01-01T00:00:00</itsgml:beginLifespanVersion>
	<itsgml:featureTypeReference xlink:href="GDF.gml#GDF.ProhibitedManoeuvre"/>
	<itsgml:locationReference> … </itsgml:locationReference>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>firstRoadElement</itsgml:memberName>
			<itsgml:propertyStructure>association</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.AssociationRoles.Manoeuvre.RoadElement.firstRoadElement"/>
			<itsgml:valueReference xlink:href="Network.gml#vegvesen.no.nvdb.rl.704520_7"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>secondRoadElement</itsgml:memberName>
			<itsgml:propertyStructure>association</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.AssociationRoles.Manoeuvre.RoadElement.secondRoadElement"/>
			<itsgml:valueReference xlink:href="Network.gml#vegvesen.no.nvdb.rl.705183_4"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>junction</itsgml:memberName>
			<itsgml:propertyStructure>association</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.AssociationRoles.Manoeuvre.Junction.junction"/>
			<itsgml:valueReference xlink:href="Network.gml#vegvesen.no.nvdb.rn.721031"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
</itsgml:Feature>
```


Figure 16. Extract of the GDF Feature Exchange GML file with an instance of the ProhibitedManoeuvre feature type.


```
<itsgml:Feature gml:id="vegvesen.no.nvdb.159605145_0">
	<itsgml:typeName>PedestrianCrossing</itsgml:typeName>
	<itsgml:validFrom>2008-04-08</itsgml:validFrom>
	<itsgml:beginLifespanVersion>2008-04-08T00:00:00</itsgml:beginLifespanVersion>
	<itsgml:featureTypeReference xlink:href="GDF.gml#GDF.PedestrianCrossing"/>
	<itsgml:locationReference> … </itsgml:locationReference>
	<itsgml:locationReference> … </itsgml:locationReference>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>priority</itsgml:memberName>
			<itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.PedestrianCrossing.priority"/>
			<itsgml:valueReference xlink:href="GDF.gml#GDF.PedestrianCrossingPriority.pedestriansOverRoadTrafficOnRequest"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>signage</itsgml:memberName>
			<itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.PedestrianCrossing.signage"/>
			<itsgml:valueReference xlink:href="GDF.gml#GDF.PedestrianCrossingSignage.trafficLightRegulated"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>type</itsgml:memberName>
			<itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.PedestrianCrossing.type"/>
			<itsgml:valueReference xlink:href="GDF.gml#GDF.PedestrianCrossingType.zebraCrossingWithOrWithoutAdditionalAids"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
	<itsgml:property>
		<itsgml:GenericFeatureProperty>
			<itsgml:memberName>level</itsgml:memberName>
			<itsgml:propertyStructure>listedValue</itsgml:propertyStructure>
			<itsgml:propertyTypeReference xlink:href="GDF.gml#GDF.PedestrianCrossing.level"/>
			<itsgml:valueReference xlink:href="GDF.gml#GDF.PedestrianCrossingLevel.streetLevelCrossing"/>
		</itsgml:GenericFeatureProperty>
	</itsgml:property>
</itsgml:Feature>
```


Figure 17. Extract of the GDF Feature Exchange GML file with an instance of the PedestrianCrossing feature type.

5. Discussion

The comparison and evaluation of standards and specifications for geospatial road-related information in Table 3 shows that none of the studied solutions meets all the requirements described in Table 2. A prototype for an improved solution was, therefore, developed and implemented in a case study. In Table 4, the prototype is evaluated against the requirements from Table 2. The evaluation indicates that the prototype meets all requirements used in the study.

Table 4. Evaluation of the prototype.

Requirement	Description	Requirement Fulfilled
ISO/TC 211 MDA	The application schemas in the prototype are based on MDA according to ISO/TC 211 standards, particularly the General Feature Model from ISO 19109 and the abstract conceptual feature catalogue model from ISO 19110. Implementation schemas are derived from the application schemas.	Yes
GIS exchange format	The exchange format used in the prototype is GML, which is the standardized GIS exchange format defined in ISO 19136. With the MDA approach, implementation schemas for other formats, such as GeoPackage, may also be derived.	Yes
Feature catalogue	The prototype does not contain a feature catalogue of its own, but can implement any feature catalogue modelled as an application schema according to ISO 19109. Selected feature types from the feature catalogues from INSPIRE TN and ISO 14825 GDF were implemented in the case study. The INSPIRE TN model was used directly, while the GDF model was modified to be compliant with ISO 19109.	Yes
Feature catalogue exchange model	The prototype contains the Feature Catalogue Exchange Model and a derived feature catalogue exchange schema for exchange in GML.	Yes
Network model	The Feature Exchange Model in the prototype contains the Network Model based on the INSPIRE GNM.	Yes
Generic feature exchange model	The Feature Exchange Model in the prototype contains the Feature Model, which is a generic feature model based on the TN-ITS model. An implementation schema for the exchange of network and features in GML was derived from the model.	Yes

In Table 5, concepts from the Feature Catalogue Exchange Model are compared to models from two of the studied solutions: The overall conceptual data model from GDF and the Transport Object Property Catalogue from OpenTNF. The comparison shows that the prototype covers the concepts for feature cataloguing defined in the two solutions, but with some improvements. First, the catalogue concept is not defined in GDF. Instead, the GDF standard is the catalogue. Furthermore, GDF does not have the concept of value domains. Global attribute types are not handled directly in OpenTNF, but rather through the concept of global value domains that may be related to many attribute types. Finally, inheritance relations are neither described in the feature catalogue models for GDF nor for OpenTNF, which means that every feature type must be described with all their properties individually.

The Network Model part of the Feature Exchange Model in the prototype is based on the models from INSPIRE, which are similar to the model in OpenTNF. In addition to these, Table 3 shows that the GDF standard and the NDS Open Lane Model contain network models as well. Table 6 presents a comparison of concepts from the Network Model in the prototype and the network models from INSPIRE, OpenTNF, GDF, and NDS Open Lane Model. The comparison shows that the two solutions from the ITS domain (GDF and NDS Open Lane Model) do not have the link sequence concept. The link and node concepts exist in all solutions in the table, which indicates that the Network Model in the prototype covers the fundamental concepts required for the exchange of network elements to ITS databases. However, the complexity described in GDF with different topology types and different levels of feature representation is not covered by the prototype. Furthermore, network information at lane-level is described both in the NDS Open Lane Model and in GDF 5.1 Part 2, but is not covered by the prototype. These issues may be topics for further studies.

Table 5. Comparison of concepts from feature catalogue exchange models.

Prototype	GDF	OpenTNF
Feature catalogue	GDF (the standard)	Catalogue
Feature type	Feature	Property object type
Feature association	Relationship	Property object type
Feature attribute	Attribute	Property object property type
Association role	Role in relationship	Property object property type
Value domain	Not defined	Value domain
Listed values	Attribute data type code list	Valid value
Data type	Composite attribute	Value domain and structured value domain property type
Global attribute types	All attribute types are global	All attribute types are local, but with global value domains
Inheritance	Not defined	Not defined

Table 6. Comparison of concepts from network models.

Prototype	INSPIRE	OpenTNF	GDF	NDS Open Lane Model
Link	Link (GNM) Road link (TN)	Link	Non-planar topo line feature, Road (level 2) Road element (level 1) Edge (level 0)	Link, Road geometry line
Node	Node (GNM) Road node (TN)	Node	Non-planar topo point feature Intersection (level 2) Junction (level 1) Node (level 0)	Intersection
Link sequence	Link sequence (GNM) Road link sequence (TN)	Link sequence	Not defined	Not defined

The Feature Model part of the Feature Exchange Model in the prototype is based on the feature model in TN-ITS. In addition to these, Table 3 shows that GDF and OpenTNF also contain generic feature exchange models. The Feature Model in GDF is particularly essential, as GDF is the core standard for exchanging information for use in ITS databases. A comparison of the concepts in the Feature Model in the prototype and the equivalent models in TN-ITS, OpenTNF, and GDF are presented in Table 7.

The comparison in Table 7 shows that the Feature Model in the prototype covers concepts described in other solutions, including the Feature Model in GDF. This implies that the Feature Model in the prototype covers the fundamental concepts required for the exchange of features to ITS databases. The prototype has the advantage of direct associations between features and properties and their models in the Feature Catalogue. GDF refers to class IDs that must be looked up in the standard, while TN-ITS refers to code lists without any further model details.

Table 7. Comparison of concepts from generic feature models.

Prototype	TN-ITS	OpenTNF	GDF
Feature	Road feature	Property object	Feature, Relationship
Association to location reference	Association to location reference	Network reference, Direct location reference	Node, Edge, Face
Association to Feature Type	Attribute: "type", referring to code list	Attribute: Property object type	Reference to feature class ID
Feature property	Road feature property	Property	Attribute
Association to Property Type	Attribute: "type", referring to code list	Embedded in attribute XML	Reference to attribute class ID

The prototype was tested for implementation in the case study, by using the feature catalogue from INSPIRE TN and a modified extract of the feature, relationship, and attribute catalogues from GDF. The results from the case study show that the prototype may be used with different feature catalogues if these are modelled according to ISO 19109. For the selected features from GDF, only minor modifications were needed, which indicates that the models of features, relationships, and attributes described in GDF can be modified according to ISO 19109. The prototype may then be used for exchanging features and their characteristics as described in GDF from GIS databases.

6. Conclusions

This study has analyzed existing standards and specifications for describing and exchanging road-related geospatial information from GIS to ITS databases, based on six requirements. The analysis reveals that none of the studied solutions meets all six requirements, but that several are promising candidates for further development of an improved solution. The GDF standard, the INSPIRE Transport Networks specification, the TN-ITS specification and the OpenTNF specification were considered as the most promising, while the NDS Open Lane Model was considered essential, as it defines the model of an ITS database.

A prototype for an improved solution was developed in the study, aiming to cover all six requirements and to enable the exchange of information described both by the GDF standard and by other feature catalogues. The prototype is based on the MDA principles from ISO 19103 and 19109 and is modelled in UML, with derived implementation schemas in the GML exchange format. Three application schemas conformant to ISO 19109 define the framework of the prototype: The Feature Catalogue, the Feature Catalogue Exchange Model based on ISO 19110, and the Feature Exchange Model with the Network Model and the Feature Model. The models are generic, and no specific feature types are defined in the Feature Catalogue. Instead, feature types from GDF and other solutions may be used. A case study was conducted to test and validate the prototype, with selected feature types from the INSPIRE Transport Networks specification and the GDF standard. The results from the case study show that feature types modelled according to ISO 19019 may be used directly in the model, and that feature types from the GDF standard may be used with minor modifications.

A comparison between the prototype and the most promising solutions from the initial studies indicates that the prototype covers the concepts defined in other solutions while having several improvements. Compared to GDF, which is considered the primary standard for exchanging geospatial information for use in ITS databases, the prototype is more suitable for implementation in GIS applications, and has more flexible handling of feature catalogues. The prototype may be a candidate for improved information exchange from road authorities' GIS databases to ITS databases based on the NDS Open Lane Model. However, as it is a prototype, it is not complete and may be extended to cover topology levels and lane-level navigation, as well as implementation in other formats, such as GeoPackage.

Author Contributions: Conceptualization, Knut Jetlund; Methodology, Knut Jetlund, Erling Onstein and Lizhen Huang; Validation, Knut Jetlund; Formal Analysis, Knut Jetlund; Investigation, Knut Jetlund; Data Curation, Knut Jetlund; Writing—Original Draft Preparation, Knut Jetlund; Writing—Review & Editing, Erling Onstein and Lizhen Huang; Visualization, Knut Jetlund; Supervision, Erling Onstein and Lizhen Huang.

Funding: This research received no external funding.

Acknowledgments: NTNU Open Access publishing funds covers the article processing charges.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zang, A.; Chen, X.; Trajcevski, G. High definition maps in urban context. *Sigspatial Spec.* **2018**, *10*, 15–20. [CrossRef]
2. Zang, A.; Li, Z.; Doria, D.; Trajcevski, G. Accurate vehicle self-localization in high definition map dataset. In Proceedings of the 1st ACM SIGSPATIAL Workshop on High-Precision Maps and Intelligent Applications for Autonomous Vehicles, Redondo Beach, CA, USA, 7–10 November 2017; p. 2.
3. Jomrich, F.; Sharma, A.; Rückelt, T.; Burgstahler, D.; Böhnstedt, D. Dynamic Map Update Protocol for Highly Automated Driving Vehicles. In Proceedings of the 3rd International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2017), Porto, Portugal, 22–24 April 2017.
4. Kleine-Besten, T.; Behrens, R.; Pöchmüller, W.; Engelsberg, A. Digital Maps for ADAS. In *Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 647–661.
5. Borzacchiello, M.T.; Boguslawski, R.; Pignatelli, F. *Improving Accuracy in Road Safety Data Exchange for Navigation Systems: European Union Location Framework Transportation Pilot*; European Union: Brussels, Belgium, 2016.
6. Richter, A.; Friedl, H.; Scholz, M. Beyond OSM–Alternative Data Sources and Approaches Enhancing Generation of Road Networks for Traffic and Driving Simulations. In Proceedings of the SUMO2016–Traffic, Mobility, and Logistics, Berlin, Germany, 23–25 May 2016; pp. 23–25.
7. Richter, A.; Scholz, M. Deploying guidelines and a simplified data model to provide real world geodata in driving simulators and driving automation. *Transp. Res. Part F Traff. Psychol. Behav.* **2017**. [CrossRef]
8. Grubmüller, S.; Plihal, J.; Nedoma, P. Automated Driving from the View of Technical Standards. In *Automated Driving*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 29–40.
9. Morales, L.G.; Orrell, T. Data Interoperability: A Practitioner's Guide to Joining up Data in the Development Sector. Available online: http://www.data4sdgs.org/sites/default/files/services_files/Interoperability%20-%20A%20practitioner%E2%80%99s%20guide%20to%20joining-up%20data%20in%20the%20development%20sector.pdf (accessed on 14 March 2019).
10. Object Management Group. *Unified Modelling Language Specification Version 2.5.1*; Object Management Group: Needham, MA, USA, 2017.
11. Miles, R.R.; Hamilton, K. *Learning UML 2.0*, 1st ed.; O'Reilly: Sebastopol, CA, USA, 2006.
12. Object Management Group. *Model Driven Architecture (MDA) Guide, rev. 2.0.*; Object Management Group: Needham, MA, USA, 2014.
13. Coetzee, S.; Plews, R.; Brodeur, J.; Hjelmager, J.; Jones, A.; Jetlund, K.; Grillmayer, R.; Wasström, C. Standards—Making Geographic Information Discoverable, Accessible and Usable for Modern Cartography. In *Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management*; Döllner, J., Jobst, M., Schmitz, P., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 325–344. [CrossRef]
14. ISO/TC 211. *ISO 19103:2015 Geographic Information—Conceptual Schema Language*; ISO: Geneva, Switzerland, 2015.
15. ISO/TC 211. *ISO 19109:2015 Geographic Information—Rules for Application Schema*; ISO: Geneva, Switzerland, 2015.
16. ISO/TC 211. *ISO 19136:2007 Geographic Information—Geography Markup Language (GML)*; ISO: Geneva, Switzerland, 2007.
17. INSPIRE. *INSPIRE Data Specification on Transport Networks–Guidelines*; INSPIRE Thematic Working Group Transport Networks: Brussels, Belgium, 2014.

18. OGC. *OGC City Geography Markup Language (CityGML) Encoding Standard Version 2.0*; Open Geospatial Consortium: Wayland, MA, USA, 2012.
19. OGC. *OGC InfraGML 1.0—Part 0–LandInfra Core—Encoding Standard*; Open Geospatial Consortium: Wayland, MA, USA, 2017.
20. OGC. *OGC InfraGML 1.0—Part 3–Alignments—Encoding Standard*; Open Geospatial Consortium: Wayland, MA, USA, 2017.
21. OGC. *OGC InfraGML 1.0—Part 4–LandInfra Roads—Encoding Standard*; Open Geospatial Consortium: Wayland, MA, USA, 2017.
22. Navigation Data Standard. *Open Lane Model Documentation*; Navigation Data Standard (NDS) e.V.: Wolfsburg, Germany, 2016.
23. ISO/TC 204. *ISO 14825:2011 Intelligent Transport Systems—Geographic Data Files (GDF)—GDF5.0*; ISO: Geneva, Switzerland, 2011.
24. ERTICO—ITS Europe. TN-ITS GO. Available online: https://tn-its.eu/tn-its-go (accessed on 15 January 2019).
25. CEN/TC 278. *CEN/TS 17268:2018 Intelligent Transport Systems—Its Spatial Data—Data Exchange on Changes in Road Attributes*; CEN: Brussels, Belgium, 2018.
26. Interactive Instruments GmbH. ShapeChange. Available online: http://shapechange.net (accessed on 15 January 2019).
27. ISO/TC 204. *ISO/DIS 20524-1 Intelligent Transport Systems—Geographic Data Files (GDF) GDF5.1—Part 1: Application Independent Map Data Shared between Multiple Sources*; ISO: Geneva, Switzerland, 2017.
28. ISO/TC 204. *ISO/CD 20524-2 Intelligent Transport Systems—Geographic Data Files (GDF)—GDF5.1—Part 2: Map Data Used in Automated Driving Systems, Cooperative ITS, and Multi-Modal Transport*; ISO: Geneva, Switzerland, 2017.
29. OpenDRIVE. *OpenDRIVE Format Specification*; rev. 1.4; VIRES Simulationstechnologie GmbH: Bad Aibling, Germany, 2015.
30. OpenTNF.org. OpenTNF White Paper Version 1.0. Available online: http://www.opentnf.org (accessed on 15 January 2019).
31. CEN/TC 278. *EN 16157-1:2018 Intelligent Transport Systems—DATEX II Data Exchange Specifications for Traffic Management and Information—Part 1: Context and Framework*; CEN: Brussels, Belgium, 2018.
32. CEN/TC 278. *FprEN 16157-2 Intelligent Transport Systems—DATEX II Data Exchange Specifications for Traffic Management and Information—Part 2: Location Referencing*; CEN: Brussels, Belgium, 2018.
33. CEN/TC 278. *EN 16157-3:2018 Intelligent Transport Systems—DATEX II Data Exchange Specifications for Traffic Management and Information—Part 3: Situation Publication*; CEN: Brussels, Belgium, 2018.
34. CEN/TC 278. *EN 16157-7:2018 Intelligent Transport Systems—DATEX II Data Exchange Specifications for Traffic Management and Information—Part 7: Common Data Elements*; CEN: Brussels, Belgium, 2018.
35. ISO/TC 204. *ISO/TS 21219-2:2014 Intelligent Transport Systems—Traffic and Travel Information (TTI) via Transport Protocol Experts Group, Generation 2 (TPEG2)—Part 2: UML Modelling Rules*; ISO: Geneva, Switzerland, 2014.
36. ISO/TC 204. *ISO/TS 21219-3:2015 Intelligent Transport Systems—Traffic and Travel Information (TTI) via Transport Protocol Experts Group, Generation 2 (TPEG2)—Part 3: UML to Binary Conversion Rules*; ISO: Geneva, Switzerland, 2015.
37. ISO/TC 204. *ISO/TS 21219-4:2015 Intelligent Transport Systems—Traffic and Travel Information (TTI) via Transport Protocol Experts Group, Generation 2 (TPEG2)—Part 4: UML to XML Conversion Rules*; ISO: Geneva, Switzerland, 2015.
38. ISO/TC 204. *ISO/TS 21219-7:2017 Intelligent Transport Systems—Traffic and Travel Information (TTI) via Transport Protocol Experts Group, Generation 2 (TPEG2)—Part 7: Location Referencing Container (TPEG2-LRC)*; ISO: Geneva, Switzerland, 2017.
39. Van Essena, R.; Hiestermannb, V. "X-GDF"—The ISO model for geographic information for ITS. In Proceedings of the ISPRS Workshop on Service and Application of Spatial Data Infrastructure, Hangzhou, China, 14–16 October 2004.
40. Shibata, J. *ISO TC204 WG3: ITS Database Technology PROGRESS REPORT as of March 12, 2018*; ISO: Geneva, Switzerland, 2018.

41. Althoff, M.; Urban, S.; Koschi, M. Automatic Conversion of Road Networks from OpenDRIVE to Lanelets. In Proceedings of the 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), Singapore, 31 July–2 August 2018; pp. 157–162.
42. Wikström, L.; Landwehr, M.; Bock, H.; Nasr, A.; Wevers, K.; Svensk, P.-O.; T'Siobbel, S.; Boterbergh, B.; Blaive, L.; Hamish, K.; et al. *ROSATTE: ROad Safety ATTributes Exchange Infrastructure in Europe: Deliverable D3.1: Specification of Data Exchange Methods*; ERTICO - ITS Europe: Brussels, Belgium, 2009.
43. Niestroj, M.G.; McMeekin, D.A.; Helmholz, P. Overview of standards towards road asset information exchange. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-4*, 443–450. [CrossRef]
44. European Parliament and the Council of the European Union. Directive 2007/2/EC of the European Parliament and of the Council of the European Union of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *EUR-LEX Off. J. Eur. Union* **2007**, *30*, 270–283.
45. INSPIRE. *INSPIRE Data Specifications—Base Models—Generic Network Model*; INSPIRE Drafting Team "Data Specifications": Brussels, Belgium, 2013.
46. European Parliament and the Council of the European Union. Directive 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport. *EUR-LEX Off. J. Eur. Union* **2010**, *207*, 1–13.
47. European Parliament and the Council of the European Union. Commission Delegated Regulation (EU) 2015/962 of 18 December 2014 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide real-time traffic information services. *EUR-LEX Off. J. Eur. Union* **2014**, *157*, 21–31.
48. Vandezande, N.; Janssen, K. The ITS Directive: More than a timeframe with privacy concerns and a means for access to public data for digital road maps? *Comput. Law Secur. Rev.* **2012**, *28*, 416–428. [CrossRef]
49. Beil, C.; Kolbe, T.H. CityGML and the streets of New York-A proposal for detailed street space modelling. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-4/W5*, 9–16. [CrossRef]
50. TISA Executive Office; EasyWay. General Description of the Easyway/Tisa Demonstration Functionality. Available online: http://www.easyway-its.eu (accessed on 15 January 2019).
51. Kwoczek, S.; Di Martino, S.; Rustemeyer, T.; Nejdl, W. An architecture to process massive vehicular traffic data. In Proceedings of the 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, Krakow, Poland, 4–6 November 2015; pp. 515–520.
52. Zhi-Ren, S.; Min, H.; Hai-Bin, W. A conceptual multi-level data model for road networks. In Proceedings of the Fifth International Conference on Intelligent Computation Technology and Automation (ICICTA), Zhangjiajie, China, 12–14 January 2014; pp. 712–715.
53. Liu, C.; Jiang, K.; Xiao, Z.; Cao, Z.; Yang, D. Lane-level route planning based on a multi-layer map model. In Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–7.
54. Chen, S.; Tan, J.; Claramunt, C.; Ray, C. Multi-scale and multi-modal GIS-T data model. *J. Transp. Geogr.* **2011**, *19*, 147–161. [CrossRef]
55. Sharma, A. Differential Map Updates for Highly Automated Driving and Enhanced Driver Assistance Services. Master's Thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2016.
56. Runde, M. Strategies for Automated Updating of Road Networks for Driving Simulations. Master's Thesis, Westfälische Wilhelms-Universität Münster, Münster, Germany, 2017.
57. Barsi, A.; Poto, V.; Somogyi, A.; Lovas, T.; Tihanyi, V.; Szalay, Z. Supporting autonomous vehicles by creating HD maps. *Prod. Eng. Arch.* **2017**, *16*. [CrossRef]
58. Czarnecki, K. *Operational World Model Ontology for Automated Driving Systems—Part 1: Road Structure*; Waterloo Intelligent Systems Engineering Lab (WISE) Report; Waterloo Intelligent Systems Engineering Lab: Waterloo, ON, Canada, 2018.
59. ISO/TC 211. *ISO 19101-1:2014 Geographic Information—Reference Model—Part 1: Fundamentals*; ISO: Geneva, Switzerland, 2014.
60. ISO/TC 211. *ISO 19110:2016 Geographic Information—Methodology for Feature Cataloguing*; ISO: Geneva, Switzerland, 2016.
61. Campbell, J.E.; Shin, M. *Essentials of Geographic Information Systems*; FlatWorld Pubislding: Boston, MA, USA, 2017.

62. Sparx Systems Pty Ltd. Enterprise Architect Version 14. Available online: https://sparxsystems.com/products/ea/index.html (accessed on 15 January 2019).
63. Safe Software Inc. FME 2018.1. Available online: https://www.safe.com/fme/fme-desktop/ (accessed on 15 January 2019).

Article 2

Improvements in automated derivation of OWL ontologies from geospatial UML models

IMPROVEMENTS IN AUTOMATED DERIVATION OF OWL ONTOLOGIES FROM GEOSPATIAL UML MODELS

Knut Jetlund [1]

[1]Norwegian University of Science and Technology, Gjøvik, Norway, and The Norwegian Public Roads Administration - knut.jetlund@vegvesen.no

Commission IV, WG IV/OGC

KEY WORDS: Ontologies, UML, Geospatial, MDA, Standards, OGC, ISO/TC 211

ABSTRACT:

Standards from ISO/TC 211 are the foundation for modelling a universe of discourse in a geospatial context. UML models based on the standards, and in particular based on the UML profile defined in ISO 19103, have been developed and implemented in applications and databases for a wide range of geospatial information, from international to national and agency level. Amounts of information has been collected, maintained and made available based on the models, but mainly through specific services and exchange formats for geospatial information. To make the models and the information available in The Semantic Web, the geospatial UML models need to be transformed from UML to OWL ontologies, and the information needs to be transformed from UML-based structures to RDF triples. This paper investigates methods for transforming UML models of geospatial information to OWL ontologies, identifies challenges, suggest improvements and identifies needs for further research. Several methods for automated transformation from geospatial UML models to OWL handle basic concepts, but some concepts and context-closed restrictions from UML cannot be directly transformed to the open world of The Semantic Web. None of the analysed methods handles all of these issues, and suggested improvements include combining and improving transformation rules, as well as modifications in the UML models. To what degree and how these issues need to be handled will depend on whether the scope of the ontologies is to simply present geospatial information on The Semantic Web, or if they shall be used in a bidirectional information exchange.

1. INTRODUCTION

Most of the information on the World Wide Web is available as documents and images in formats like HTML, PDF or JPEG, or in databases that are accessed by special applications. Humans can combine the information, make assumptions and extend knowledge by reading and understanding documents and looking at tables and maps, even if the documents and databases are structured in different ways, and even if different terms are used for the same phenomena. For processing by machines however, the information must have a formal structure and explicit semantic. The Semantic Web provides the framework for describing information in structures that machines can use to understand and share information, and reuse it independently of applications.

The basic framework for information modelling on The Semantic Web is the Resource Description Framework – RDF, in which the information is described with triples and graphs. A triple consists of an object, a predicate and a subject, where objects and subjects are resources that can be anything from a concrete physical phenomena to an abstract concept, and the predicate describes the connection between the object and the subject. An object in one triple may be a subject for another triple, and a set of triples form a graph of information. The Web Ontology Language – OWL is the main framework for describing ontologies, built on top of RDF and using the same principles with triples and graphs.

The geospatial aspect of information on the Web is important for many use cases, e.g. navigation, travel, advertising etc. The ISO Technical Committee 211 – ISO/TC 211 and the Open Geospatial Consortium – OGC have both been working on standardization of geospatial information since 1994, individually and in cooperation. The work is based on the ISO/TC 211 approach to modelling information, described in the standards ISO 19103 (ISO/TC 211, 2015a) and ISO 19109 (ISO/TC 211, 2015b). The ISO/TC 211 approach is to perceive a part of the real world, known as a universe of discourse, limit the perception to a closed context of geographic[1] information, and classify feature types (classes) and properties (attributes) according to this perception. Figure 1, from ISO 19109, illustrates the ISO/TC 211 approach.

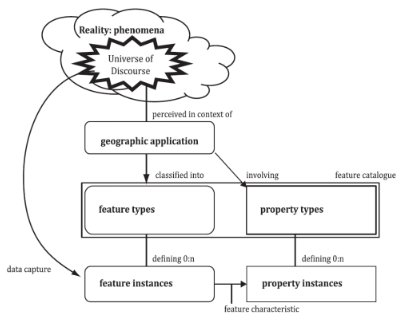


Figure 1 The process for modelling geographic information, from ISO 19109.

The foundation for information modelling in the ISO/TC 211 approach is the Unified Modelling Language – UML,

[1] ISO/TC 211 use the term "geographic information" (ISO/TC 211, 2014), while OGC use the term "geospatial information" (OGC, 2018). In this paper, the terms "geographic" and "geospatial" are considered equivalent.

formalized through the UML profile defined in ISO 19103 with extensions in ISO 19109, and the General Feature Model – GFM defined in ISO 19109. UML is both a graphical and a lexical modelling language, and has become the most common language used for modelling information and software applications (Miles and Hamilton, 2006). The graphical view presented in UML diagrams is very useful for human communication, while more semantics needed for machine processing is described lexically. ISO 19103 and ISO 19109 contain specific rules of how the mechanisms of UML shall be used to add semantics for automated generation of implementation schemas in e.g. XML, based on the principles of Model Driven Architecture – MDA.

The ISO/TC 211 approach for UML modelling of geospatial information has been used for a wide range of applications and information models in the domain of geospatial information. These are models at agency level, national level, regional level and international level, and large amounts of geospatial information have been collected and maintained based on these models. One important set of models and data is the European INSPIRE Directive, (European Commission, 2007), which defines several spatial themes that are described in information models according to ISO 19109. For this purpose, INSPIRE has also defined specific information modelling rules in the Generic Conceptual Model (INSPIRE, 2013). Another important set of models are the application schemas developed by OGC, such as CityGML and InfraGML.

Geospatial information have been published on the Web in various forms for many years, both as web services and as download services. The ISO/TC 211 and OGC standards Web Map Service – WMS and Web Feature Service – WFS have been used extensively over the last 10-15 years, and Spatial Data Infrastructures – SDIs provide portals and catalogue services for searching and accessing information from several stakeholders. The services are mainly targeting geospatial applications and their users, and not so much The Semantic Web. However, mapping authorities in some countries have started to publish geospatial information for The Semantic Web, e.g. Ordnance Survey in the United Kingdom and Ordnance Survey in Ireland. Furthermore, OGC and W3C established The Spatial Data on the Web Working Group in cooperation in 2015 (Tand, van den Brink, and Barnaghi, 2017).

The purpose of this paper is to analyse methods for transformation of ISO/TC 211 conformant UML models of geospatial information to OWL ontologies for publication on The Semantic Web. Furthermore, based on the analysis, to identify possible challenges, and suggest improvements and further research.

2. LITERATURE REVIEW

The Object Management Group – OMG specification Ontology Definition Metamodel (Object Management Group, 2009) defines a metamodel and a UML profile for OWL, and describes transformations between UML and OWL in general. Transformations have been discussed in several articles as well, with similar approaches, but with differences in how complex UML characteristics they handle. Some articles, such as (Ferreira and Manuel, 2007, Gasevic et al., 2004, Gherabi and Bahaj, 2012) cover mainly classes, attributes and simple associations, while more complex methods described in e.g. (Bourahla and Belghiat, 2012b, a, Xu et al., 2012, Bahaj and Bakkas, 2013, Hajjamy et al., 2016) also cover generalization, abstract classes, compositions and multiplicity.

The ISO/TC 211 standard ISO 19150, part 2 (ISO/TC 211, 2015c), defines standardized rules for transforming ISO 19103 and ISO 19109 compliant UML models to OWL. The ISO/TC 211 Group for Ontology Management – GOM (ISO/TC 211, 2018) have developed technologies for deriving ontologies from ISO/TC 211 compliant UML models, based on the rules in ISO 19150-2. For INSPIRE UML models, specific guidelines have been developed for transforming from UML to OWL (ARE3NA project, 2017).

Several research articles describe transformations of geospatial UML models to OWL. Some of the articles are early studies on the subject, written before the ISO/TC 211 and INSPIRE communities started to work on ontologies, and present some of the challenges and the benefits of preparing geospatial UML models for The Semantic Web. The use of ontologies for translation between data sources for land cover information was described already in 1999 in (Stuckenschmidt et al., 1999). Transformation of some core ISO and OGC UML models to ontologies are described in (Probst, Bibotti, and Pazos, 2004), while (Russomanno, Kothari, and Thomas, 2005) describe the transformation from UML to OWL for the OGC Specification SensorML. A similar transformation for the ISO/TC 211 and OGC standard for observation and measurement is described in (Probst, Gordon, and Dornelas, 2006) and later in (Cox, 2013, Cox, 2017). In (Buccella et al., 2011), an ontology based on the core ISO/TC 211 standards ISO 19107 and 19109 is described, while (Zedlitz and Luttenberger, 2012) discuss differences between UML and OWL and model transformation at a meta level, with reference to the UML Profile in ISO 19103.

One of the early articles on the use of geospatial information in The Semantic Web is (Egenhofer, 2002), where the concept of the Semantic Geospatial Web and research issues needed to enable it is described. In particular, two issues are pointed out: The need for a method for querying on geospatial characteristics, and the need for methods for enabling geospatial data sources for use in The Semantic Web. (Kolas, Hebeler, and Dean, 2005) points out the way towards the Sematic Geospatial Web by suggesting an architecture of five types of geospatial ontologies: A base geospatial ontology, a geospatial service ontology, a filter ontology, domain ontologies and feature data source ontologies.

A state of art overview of methodologies for querying geospatial information on The Semantic Web is described in (Battle and Kolas, 2012), where the query language for geospatial information on The Semantic Web – GeoSPARQL (Perry and Herring, 2012) is introduced and described. Several articles show the practical use of geospatial information on The Semantic Web. The three articles (Aditya and Kraak, 2007, Klien, 2007, Lutz and Kolas, 2007) describe the potential of The Semantic Web for discovery in SDIs. Geospatial information is accessed as linked data from WFS in (Hietanen, Lehto, and Latvala, 2016), while (van den Brink et al., 2014) and (Patroumpas et al., 2015) describe a transformation of both UML models and data in GML format to OWL and RDF. (Karan, Irizarry, and Haymaker, 2015) describe and demonstrate how Semantic Web technologies can be used to integrate and query data sets from the GIS and BIM domains, while the ongoing INTERLINK Project (Luiten et al., 2017) uses Semantic Web technologies for combining geospatial information from different domains and stakeholders.

3. TRANSFORMING MODELS FOR GEOSPATIAL INFORMATION TO OWL

3.1 Fundamental differences

Several articles, e.g. (Noy and Klein, 2004, Kiko and Atkinson, 2008, ARE3NA project, 2017), point out differences between UML and OWL that are important to be aware of. One fundamental difference is the assumptions of an open or a closed world. The Open World Assumption – OWA and the possibility for anyone to say anything about anything is an important part of information modelling for The Semantic Web, while UML models based on the ISO/TC 211 approach are limited to the closed context of geographic information, and are assumed complete in that context. To preserve the Closed World Assumption – CWA of the original model, the transformation may need to include some restrictions. Furthermore, ontologies are reusing and extending other ontologies, including ontologies from other domains. UML models based on the ISO/TC 211 approach are reusing concepts from other models as well, but mainly limited to models from the domain of geospatial information. Finally, the logic in ontologies is based on set theories, with set-based class relations such as disjoint, union, intersect and equivalent. UML is not based on the same logic, but some restrictions from UML models must be translated to these kinds of relations in OWL.

3.2 Packages

UML packages correspond to OWL ontologies, and the transformation is described as straightforward, where the package name becomes the ontology name in (Object Management Group, 2009), (Bourahla and Belghiat, 2012a) and in ISO 19150-2. The INSPIRE Guidelines states that a package stereotyped as application schema according to ISO 19109 shall be converted into a single ontology with name and namespace derived from the tagged value "xmlns" on the UML package.

3.3 Classes

The concepts of *class*, *generalization* and *inheritance* exists in both UML and OWL. A UML class is simply transformed to an OWL class, while a UML generalization is transformed to an OWL *subclassOf* axiom.

The UML concept of *abstract classes* defines classes that shall not have instances, and is often used in generalizations where the abstract class is a *superclass* used for defining attributes, associations and operations that are common to all subclasses. The concept does not exist in OWL, and must be handled by using other mechanisms. ISO 19150-2 introduce an annotation property "*isAbstract*" for this purpose, and the INSPIRE Guidelines use this property as well. However, there are no rules connected to the property, so it may still be possible to create instances of the abstract class. (Zedlitz and Luttenberger, 2012) suggest to use the *DisjointUnion* axiom for abstract classes and subclasses, but as stated in the article, this will not prohibit creating instances of the abstract class directly.

Method	Described by
isAbstract	ISO 19150-2, INSPIRE
DisjointUnion	(Zedlitz and Luttenberger, 2012)

Table 1. Methods for transforming abstract classes

Multiple inheritance, where a class is a subclass of more than one class, is sometimes used in UML, but is a problematic issue for implementation in e.g. XML. The rules for conversion from UML models to implementation schemas in ISO 19136 (ISO/TC 211, 2007) do not support multiple inheritance, and ISO 19103 recommends that multiple inheritance is avoided unless really needed. However, multiple inheritance is still used in some UML models for geospatial information, e.g. some INSPIRE models (INSPIRE, 2010). Multiple *subclassOf* predicates is very common in OWL, and the INSPIRE guidelines states that multiple inheritance shall be handled with multiple *subclassOf* predicates. In (Hajjamy et al., 2016), a more specific representation is applied, with the subclass as an *intersection* of the superclasses, to make sure the subclass follows restrictions from all of its superclasses.

3.4 Data types

Data types for attributes in UML can be classified as primitive and complex. *Primitive data types* are types with atomic values, such as integer, string and boolean. A set of primitive data types are defined in ISO 19103, and according to both ISO 19150-2 and the INSPIRE Guidelines, these are mapped to equivalent XML Schema (XSD) data types and referred to as *DatatypeProperties*. The same approach is followed by (Hajjamy et al., 2016).

Complex data types have an internal structure. They have own attributes and associations, possibly with complex data types too. ISO 19103 defines a range of such data types: for measures (angle, length etc.) that require a unit of measure in addition to the value; name types; record types; and more. Other standards based on ISO 19103 defines other and more domain specific data types. There are some differences in how complex data types are transformed according to the rules in ISO 19150-2 and the INSPIRE Guidelines. ISO 19150-2 have a short list of data types from ISO 19103 that are mapped to equivalent XSD data types and referred to as *DatatypeProperties*, while all other data types are converted to OWL Classes, and referred to as *ObjectProperties*. The INSPIRE Guidelines focus more on reusing existing ontologies, and have a longer list of predefined mappings of commonly used data types, from both ISO 19103 and other ISO Standards. These are mapped to equivalent or similar datatypes from both XSD and other ontologies. As stated in the guidelines, this may lead to a loss of information when transforming from GML to RDF and back, because of differences between the UML data types and the external data types. Of the other mappings described in articles, only (Hajjamy et al., 2016) describe mapping of complex data types, and they are then converted to OWL classes referred to as ObjectProperties.

Method	Described by
Mapping a few, else new class	ISO 19150-2
Mapping all similar, else new class	INSPIRE

Table 2. Levels of mapping to existing data types

3.5 Enumerations

Enumerations in UML are data types that define complete lists of possible values. Attributes using the enumeration can only have values defined in the list, and no other value. OWL has the equivalent axiom *oneOf*, which can be used to restrict both DatatypeProperties (*DataOneOf*) and ObjectProperties

(*ObjectOneOf*). Both ISO 19150-2, the INSPIRE Guidelines, (Zedlitz and Luttenberger, 2012) and (Hajjamy et al., 2016) describe the transformation of UML enumerations to DatatypeProperties with the oneOf axiom and a collection of values from the UML enumeration. However, the INSPIRE Guidelines state that only enumerations with self-describing codes that have an obvious meaning shall be represented with the oneOf axiom. In other cases, the enumeration shall be handled as a separate *SKOS Concept Scheme*, and the range for the attribute shall refer to the generic class *skos:Concept*. A *seeAlso* statement is added with the URL to the SKOS Concept Scheme, to describe where the list is to be found, but without any actual binding. This is the same approach as used for code lists in the INSPIRE Guidelines.

3.6 Code lists

The ISO/TC 211 UML profile in ISO 19103 defines code lists as flexible enumerations, meaning that there can be other values than those described in the list. This is an important issue, as one cannot expect that all possible values are described in the model or the ontology, they may be described elsewhere. The INSPIRE Guidelines use the same approach for code lists as described for enumerations without self-describing codes. ISO 19150-2 also uses SKOS Concept Schemes for code lists, but with a closer binding than the approach in INSPIRE. The code list is defined both as a class, a concept scheme and a collection, where the class is a subclass of skos:Concept. The binding of the code list and the attribute is not described in the standard, but in the ISO/TC 211 official ontologies, the attributes are bound to the code list class with the *allValuesFrom* axiom. This close binding excludes the possibility for additional values described in other SKOS Concept Schemes.

(Zedlitz and Luttenberger, 2012) describe a different approach. They use a *UnionOf* axiom to define a union, and let the union include an *OneOf* list with the code list values, and any other value, defined with a standard XML Schema expression that is also used for code lists in GML.

Method	Described by
SKOS and allValuesFrom	ISO 19150-2
SKOS	INSPIRE
DataUnionOf and any value	(Zedlitz and Luttenberger, 2012)

Table 3. Handling of code lists

3.7 Union

The ISO/TC 211 UML profile in ISO 19103 defines unions as types with a list of several alternative datatypes, where one and only one shall be used for an attribute value. A union is similar to an enumeration, except that the values in the list are data types, not literals. This is a different meaning of union than the set-based union in OWL, where a union contains every individual contained in at least one of the classes in the union. (Zedlitz and Luttenberger, 2012) describe two possible methods for transforming a UML union to OWL. The first method cover the situation where all members in the UML union can be transformed to either DataProperties or ObjectProperties. They define a property (either data or object) for each member, and an additional property that all members are a subproperty of, and which has a cardinality of exactly one (*ExactCardinality*). By using this property as range for attributes from the UML model, only one of the members of the UML union can be selected. However, because of the Open World Assumption, this method does not avoid the use of other properties that are not members of the union. The other method cover situations where some of the members in the union can be transformed to DataProperties, and some to ObjectProperties. A class is defined for each member of the union, and these classes are set as disjoint from each other. Each of the new classes is set to be equal to a set of exactly one of the current union member. The downside of this solution is that it gets much more complex with many axioms.

In ISO 19150-2, a UML union is simply transformed to an OWL union. As long as only one member is assigned to each UML attribute, this will give the correct representation. However, with the OWL Union, several members from the union can be assigned for the same instance of the UML attribute, which breaks the rules for a UML union. The INSPIRE Guidelines have a fourth approach, where each member of the union is transformed to a property. For each of these new properties, an intersection with all union members is defined, where cardinality expressions are used to define that only this property can have a value. A union expression combines all the intersections, but because of the cardinality restrictions in each intersection, choosing one of them will exclude the others. So only one of the transformed properties in the union may be used. Like the second method from (Zedlitz and Luttenberger, 2012), this method is quite complex, but it seems to maintain the purpose of a UML union.

Method	Described by
Union	ISO 19150-2
Intersection and union	INSPIRE
Subproperty and ExactCardinality	(Zedlitz and Luttenberger, 2012)
Disjoint classes	(Zedlitz and Luttenberger, 2012)

Table 4. Methods for transforming UML Unions

3.8 Attributes and association roles

UML has two ways of describing further characteristics of a class: As *attributes* with primitive or complex data types, or as *associations* to other classes. Both of these are similar to properties in OWL: In principle, attributes with simple (primitive) data types are equivalent to DataProperties, while attributes with complex data types and association roles are equivalent to ObjectProperties, as they refer to another class. However, there is one fundamental difference: While a UML class is the single owner of its attributes and associations, properties in OWL are globally defined and may be assigned to any class. Several classes in a UML model may have attributes or association roles that are identical on all classes, having identical name, data type and definition, or they may be almost identical. Furthermore, several classes may have an attribute or an association role with identical name, but different data type and/or definition. When UML attributes and association roles shall be transformed to OWL properties, these issues need to be handled. Identical attributes and roles should preferably be handled as one global property, assigned to the respective classes, while attributes and roles with identical names, but different data type and/or definition need to be handled as separate properties, with different identifiers. The attribute or role name alone may lead to duplicate properties with different meanings.

The referred articles, standards and guidelines handle these issues to various degrees, and with various approaches. (Gasevic et al., 2004, Gherabi and Bahaj, 2012, Bahaj and Bakkas, 2013, Hajjamy et al., 2016) do not refer to these issues at all, and just perform a simple transformation to properties. While (Xu et al., 2012, Zedlitz and Luttenberger, 2012, Cox, 2013, Cox, 2017) only reflect over the issues without proposing a method to solve them, and (Buccella et al., 2011) performs a similarity matching to identify global concepts. The rule in ISO 19150-2 is to add the class name as a prefix, and thereby create unique properties from all attributes and association roles.

The INSPIRE Guidelines takes a more advanced approach, striving to achieve globally scoped properties and reuse of existing properties when possible. Properties that have identical or close to identical meaning (but not necessarily identical name) shall be converted to properties with a global scope, and properties that are similar to properties already defined in other ontologies shall be converted to those properties. The guidelines recognize that the process of identifying these properties require a manual review. For all other properties, the class name is added as a prefix, following the rules from ISO 19150-2.

Method	Described by
Prefix	ISO 19150-2
Manual matching	INSPIRE
Similarity matching	(Buccella et al., 2011)

Table 5. Methods for attribute globalization

3.9 Associations

An association in UML is a relationship between two classes, is similar to an attribute, and is implemented in the same manner as attributes in e.g. XML. The transformation of a simple association from UML to OWL is done by creating an ObjectProperty with one class as the *domain* and the other class as the *range* of the property. This approach has been followed in all the referred articles, and also in ISO 19150-2 and the INSPIRE Guidelines.

A UML aggregation, also known as shared aggregation, is a more specific relationship, where the associated class is a part of the main class (the whole). This kind of association does not add any actual semantics to the model; it just describes a closer relationship. Aggregation is described as an association type in several articles, but (Bahaj and Bakkas, 2013) is the only article that describe a method for maintaining the aggregation in the OWL model. They have created a hierarchy of properties representing relationship types, with association at the top, and aggregation and composition as subproperties. Every association from the UML model is transformed to a subproperty of one of these properties. This way, possible semantics may be added to the different relationship types. ISO 19150-2 maintains information about the relationship type by adding an annotation with *aggregationType.*

A UML composition, also known as composite aggregation, is a stronger relationship between two classes. An instance of a class, related to another class through a composition, can only take part in one composition, i.e. the part instance can only be related to one whole instance at the same time. Often, but not always, the part will also be deleted if the related whole is deleted.

ISO 19150-2 handles compositions in the same manner as aggregations, with the aggregationType annotation. Composition is also described in several articles, but (Bahaj and Bakkas, 2013) and (Hajjamy et al., 2016) are the only articles that describe a way of maintaining it. In (Bahaj and Bakkas, 2013), this is done with the described hierarchy of relationship types, but no semantics is added to the composition property. (Hajjamy et al., 2016) describe a method for maintaining the restrictions by setting the property that represent the composition as *InverseFunctional*, which means that the associated class can only be linked to one class through this property. Furthermore, they also add restrictions saying that the composition cannot be from the related class itself (*Irreflexive*) and that the composition cannot be applied in the other direction (*Asymmetric*).

Method	Described by
aggregateType	ISO 19150-2, INSPIRE
Property hierarchy	(Bahaj and Bakkas, 2013)
InverseFunctional	(Hajjamy et al., 2016)

Table 6. Methods for transforming compositions

4. DISCUSSION

4.1 The scope of geospatial ontologies

Even though both The Semantic Web and digital geospatial information has existed for many years, the use of geospatial information in The Semantic Web is still limited. Large amounts of geospatial information are available on the Web, but almost solely in domain specific Web services or download services from the GIS domain. One example of this is the European INSPIRE Geoportal, with more than 58000 data sets available through WMS or WFS, and more than 37000 data sets available for download, but no information or information models available RDF or OWL. However, as described in several research articles, these services and the information provided by them can still be used as resources in The Semantic Web. Metadata may be converted to RDF and used for querying and discovery. The application schemas provided by the web services may be converted to ontologies, and data in GML format may be converted to RDF, on request or as complete exports. Both (van den Brink et al., 2014), (Hietanen, Lehto, and Latvala, 2016) and (Patroumpas et al., 2015) describe how such functionality can be built as extensions to existing SDIs, and thereby making the large amounts of geospatial information available for The Sematic Web.

(Noy and McGuinness, 2001) states that the first step in ontology development shall be to determine the scope of the ontology. The scope of geospatial ontologies is important for deciding how well the transformation of UML models to ontologies shall maintain the closed world-based concepts and restrictions from the UML model, and for deciding if similar but not identical concepts from existing ontologies can be reused. At least three possible main scopes are relevant for geospatial ontologies:

1. Enable geospatial information from GIS applications for The Semantic Web, by unidirectional information exchange.
2. Enable bidirectional exchange of information between GIS applications and The Semantic Web
3. Replace GIS databases and applications with triple stores and query engines in The Semantic Web.

The third scope is not a likely situation a short term, mainly because of the complexity and advanced functionality that exists in GIS applications and databases that have been developed specifically for handling complex geometries and topologies, and operations on these. As discussed in e.g. (Tand, van den Brink, and Barnaghi, 2017), it will not be convenient to replace all of this in triple stores and query engines. The second scope is applied in the INTERLINK project (Luiten et al., 2017), where information will be exchanged bi-directionally between application domains, stakeholders and lifecycle phases, but where the original information models are mainly developed directly in OWL. This is a more likely scope for the ontologies, in which case more restrictions will need to be included in OWL. However, the most discussed scope, e.g. in the INSPIRE Guidelines and in (Tand, van den Brink, and Barnaghi, 2017), is to store the information in GIS databases and transform it to RDF. With this scope, less strict transformations can be applied.

4.2 Transformation issues

The studies indicate that the transformation methods handle the basic UML concepts *classes*, *generalizations*, *primitive data types*, *enumerations* and *simple associations* similarly and acceptable. For other UML concepts, several approaches have been used, and the level of complexity needed will depend on the scope of the ontologies. If all restrictions shall be maintained, the transformation need to be more complex.

The concept of *abstract classes* is widely used in UML models, but only briefly handled in the transformation methods. None of the methods fully maintain the concept of abstract classes; it may still be possible to create instances of the classes in RDF. A solution that may maintain the concept better is to transform only the properties of the abstract class to OWL and not the class itself, and then assign the properties to each subclass in OWL instead. This would also be closer to implementations in databases, where the abstract class itself will not be implemented. This approach has not been discussed in any of the articles.

Reusing data types and classes from existing ontologies is a fundamental part of The Semantic Web. All *primitive data types* from ISO 19103 and some *complex data types* can be transformed directly to XML Schema data types. This is a simple mapping for primitive data types, while for complex data types there is a question on whether or not the data types are identical, and how identical they need to be. The INSPIRE Guidelines have specified mapping to more existing data types than ISO 19150-2, including mappings to similar but not identical types, which may lead to information loss due to minor differences. An approach that may improve this in general is to reuse more existing data types in the original UML model. The method described by (van den Brink et al., 2014), where elements that shall be linked to existing ontologies are tagged in the UML model, may be used to support this approach.

The methods for handling *Code lists* according to ISO 19150-2 and the INSPIRE Guidelines both describe the use of complete SKOS Concept Schemes. However, there are some weaknesses in the methods. The use of the allValuesOf axiom in the official ISO/TC 211 ontologies excludes values that are not in the defined code lists. This breaks the concept of a code list as an open enumeration. The solution described in (Zedlitz and Luttenberger, 2012), with a union of defined values and any other value, defined with a standard XML Schema expression, seems to solve this better. The INSPIRE Guidelines recommend to use separate SKOS Concept Schemes also for *enumerations* that do not have obvious meaning. This seems good for being able to describe the values in the enumerations, but it removes the restriction that lies in the concept of an enumeration as a closed list.

The handling of a UML *union* in OWL is a complex issue that is solved with different methods in ISO 19150-2, the INSPIRE Guidelines and in (Zedlitz and Luttenberger, 2012). The most complete method seems to be the one described in the INSPIRE Guidelines, but it is also quite complex. One method that may simplify this is to use the ObjectOneOf axiom, similar to the way enumerations are handled.

For *attributes and association roles*, a main question is how to handle attributes or association roles that are identical for several classes, how to handle those that have almost identical names and meaning, and those that have identical names but different meaning. The approach in ISO 19150-2, with class name as prefix for all properties, makes all properties globally unique, but excludes reuse of properties globaly. The INSPIRE Guidelines suggests to harmonize properties internally in the ontology, and with external ontologies, to enable reuse, which will make the ontologies differ from the UML model. Once again, this is acceptable with the scope of unidirectional information exchange. However, a weakness is that the harmonization must be done manually, but string-matching algorithms, as described in (Buccella et al., 2011) may support the process. An additional approach, as used in (Hietanen, Lehto, and Latvala, 2016), would be to improve the UML models from the ontologies, and thereby achieve harmonization between the UML and OWL representations. Including tagging of attributes and association roles in the UML model that shall be linked to existing or similar internal attributes or association roles, as discussed for complex data types and in (van den Brink et al., 2014), would also be an improvement. Furthermore, a use of properties and subproperties for almost identical attributes is also a method that may be used.

Of the described methods for transformation of *associations*, only (Hajjamy et al., 2016) include a solution for maintaining the semantics of a composition. However, both the aggregationType annotation that is used in ISO 19150-2 and the INSPIRE Guidelines, and the association type hierarchy that is used in (Bahaj and Bakkas, 2013) may be extended to include such semantics. Combining one of these approaches with the method from (Hajjamy et al., 2016) is a possible solution that may improve the handling of association restrictions.

5. CONCLUSIONS AND RESEARCH RECOMMENDATIONS

5.1 Conclusions

In this paper, methods for transforming UML models of geospatial information has been studied and discussed. This subject, and methodologies for enabling geospatial information from SDIs for The Semantic Web, has been studied for many years and in several projects. Still, geospatial information is mainly available in domain specific Web services for geospatial information in SDIs, and not as OWL and RDF. Research indicates that a method for enabling models and information from SDIs for The Semantic Web can be to extend existing Web services and download services with transformation functionality to OWL and RDF.

The information models for geospatial information are mainly developed using UML, based on standards from ISO/TC 211 and OGC. An important part of enabling geospatial information for The Semantic Web will be to transform these UML models to OWL ontologies. Methods and rules for transforming from UML to OWL in general, and in particular for UML models of geospatial information, has been studied in several research articles, and standardized rules for the transformation have been developed in ISO/TC 211, as well as guidelines in INSPIRE. The review in this paper indicates that most of the transformation is straight forward, but some fundamental differences between UML and OWL must be handled. Three fundamental differences are particularly important: First, the UML models represents a closed world in a geospatial context, while OWL operates in an open world. The UML concepts of abstract classes, enumerations, unions and aggregations represent restrictions in the model that need special handling if they shall be maintained in the ontology. This has been done in diverse ways in the methods that has been studied, but none of them fully maintain all restrictions. Second, attributes and associations in UML are uniquely owned by each class, while properties in OWL are globally scoped, and one property may be used for many classes. Several classes in UML may have identical attributes and association roles; almost identical attributes and association roles; or attributes and association roles with identical name but different meaning. All of these need to be handled in the transformation to OWL, and several approaches have been used here as well, each with strengths and weaknesses. Third, an important principle in ontology development is reusing elements from existing ontologies, while UML models mainly use elements internally in the model. For this matter, mapping to existing ontologies have been defined in several ways.

An important question to be answered for the transformations is whether the ontologies shall be used for unidirectional or bidirectional information exchange. In the case of unidirectional information exchange, the need for maintaining restrictions from the UML model is not so important, as the information shall only be transformed from UML-based databases to RDF. In the case of bidirectional information exchange, where information shall also be transformed back to UML-based databases, instances may possibly be created in RDF. Maintaining the restrictions from the UML models is then more important, or else instances may be created that are illegal according to the UML model.

5.2 Research recommendations

Based on the discussion and the conclusions, two main topics have been identified as candidates for further research:

The semantics of some UML concepts cannot be directly transformed to OWL concepts, in particular abstract classes; code lists; unions; and aggregations. Possible approaches and improvements have been discussed in this paper, and should be studied further and tested in implementations, by extending the technologies developed by ISO/TC 211 GOM (ISO/TC 211, 2018).

For attributes and associations, and for complex data types, there is a question of how to handle singular versus globally owned properties, and the reuse of elements from existing ontologies. Some improvements have been discussed in this paper, including methods for reusing and marking elements from external ontologies in the UML models, and methods for defining global properties in the UML models. Experiences from the work on ontologies may be brought into the work on UML models, to overcome some of the challenges. Further research on these issues should include methods for defining global and external elements in the UML models, and methods for similarity matching with internal and external elements.

ACKNOWLEDGEMENTS

The author would like to acknowledge the support from Rune Hjelsvold, Erling Onstein and Lizhen Huang with this article.

REFERENCES

Aditya, Trias, and Menno-Jan Kraak, 2007. A Search Interface for an SDI: Implementation and Evaluation of Metadata Visualization Strategies. *Transactions in GIS*, 11 (3):413-435, doi: 10.1111/j.1467-9671.2007.01053.x.

ARE3NA project, 2017. Guidelines for the RDF encoding of spatial data. accessed 2018-07-04. http://inspire-eu-rdf.github.io/inspire-rdf-guidelines/.

Bahaj, Mohamed, and Jamal Bakkas, 2013. Automatic Conversion Method of Class Diagrams to Ontologies Maintaining Their Semantic Features. *International Journal of Soft Computing and Engineering (IJSCE)*:65-69.

Battle, R., and D. Kolas, 2012. Enabling the geospatial Semantic Web with Parliament and GeoSPARQL. *Semantic Web*, 3 (4):355-370, doi: 10.3233/Sw-2012-0065.

Bourahla, Mustapha, and Aissam Belghiat, 2012a. Transformation of UML models towards OWL ontologies. *2012 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, SETIT 2012*.

Bourahla, Mustapha, and Aissam Belghiat, 2012b. UML Class Diagrams to OWL Ontologies: A Graph Transformation based Approach. *International Journal of Computer Applications*, 41 (3):41-46, doi: 10.5120/5525-7566.

Buccella, A., A. Cechich, D. Gendarmi, F. Lanubile, G. Semeraro, and A. Colagrossi, 2011. Building a global normalized ontology for integrating geographic data sources. *Computers & Geosciences*, 37 (7):893-916, doi: 10.1016/j.cageo.2011.02.022.

Cox, S. J. D., 2017. Ontology for observations and sampling features, with alignments to existing models. *Semantic Web*, 8 (3):453-+, doi: 10.3233/SW-160214.

Cox, Simon, 2013. An explicit OWL representation of ISO/OGC Observations and Measurements. *SSN@ ISWC*:1-18.

Egenhofer, Max J, 2002. Toward the semantic geospatial web. *Proceedings of the 10th ACM international symposium on Advances in geographic information systems*.

European Commission, 2007. DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *Official Journal of the European Union*.

Ferreira, David De Almeida, and Alberto Manuel, 2007. UML to OWL Mapping Overview An analysis of the translation process and supporting tools. *7th Conference of Portuguese Association of Information Systems (CAPSI'06)*.

Gasevic, Dragan, Dragan Djuric, Vladan Devedzic, and Violeta Damjanovi, 2004. Converting UML to OWL ontologies. *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters - WWW Alt. '04*.

Gherabi, N., and M. Bahaj, 2012. A New Method for Mapping UML Class into OWL Ontology. *International Journal of Computer Applications*:5-9.

Hajjamy, O. E., K. Alaoui, L. Alaoui, and M. Bahaj, 2016. Mapping UML to OWL2 ontology. *Journal of Theoretical and Applied Information Technology*, 90 (1):126-143.

Hietanen, E., L. Lehto, and P. Latvala, 2016. Providing Geographic Datasets as Linked Data in Sdi. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B2:583-586, doi: 10.5194/isprs-archives-XLI-B2-583-2016.

INSPIRE, 2010. INSPIRE Data Specification on Transport Networks – Guidelines.

INSPIRE, 2013. INSPIRE Generic Conceptual Model.

ISO/TC 211, 2007. ISO 19136:2007 Geographic information — Geography Markup Language (GML).

ISO/TC 211, 2014. ISO 19101-1:2014 Geographic Information — Reference model — Part 1: Fundamentals.

ISO/TC 211, 2015a. ISO 19103:2015 Geographic Information — Conceptual schema language.

ISO/TC 211, 2015b. ISO 19109:2015 Geographic information — Rules for application schema.

ISO/TC 211, 2015c. ISO 19150-2:2015 Geographic information — Ontology — Part 2: Rules for developing ontologies in the Web Ontology Language (OWL).

ISO/TC 211, 2018. ISO/TC 211 Group for Ontology Management. accessed 2018-07-04. https://github.com/ISO-TC211/GOM.

Karan, Ebrahim P, Javier Irizarry, and John Haymaker, 2015. BIM and GIS integration and interoperability based on semantic web technology. *Journal of Computing in Civil Engineering*, 30:4015043.

Kiko, K, and Colin Atkinson, 2008. A detailed comparison of uml and owl. *Technical Report 4, Department of Mathematics and Computer Science, University of Mannheim*:1-58.

Klien, Eva, 2007. A Rule-Based Strategy for the Semantic Annotation of Geodata. *Transactions in GIS*, 11 (3):437-452, doi: 10.1111/j.1467-9671.2007.01054.x.

Kolas, D, J Hebeler, and M Dean, 2005. Geospatial semantic web: Architecture of ontologies. *Geospatial Semantics, Proceedings*.

Luiten, Bart, Aonghus O'Keeffe, Sander Stolk, Lars Wikström, and Matthias Weise, 2017. INTERLINK D4. Principles for a European Road OTL.

Lutz, Michael, and Dave Kolas, 2007. Rule-Based Discovery in Spatial Data Infrastructure. *Transactions in GIS*, 11:317- 336.

Miles, Russ (Russell), and Kim Hamilton, 2006. *Learning UML 2.0*. First edit ed: O'Reilly.

Noy, Natalya F, and Michel Klein, 2004. Ontology evolution: Not the same as schema evolution. *Knowledge and information systems*, 6:428-440.

Noy, Natalya F., and Deborah L. McGuinness, 2001. Ontology Development 101: A Guide to Creating Your First Ontology. *Stanford Knowledge Systems Laboratory*:25, doi: 10.1016/j.artmed.2004.01.014.

Object Management Group, 2009. Ontology Definition Metamodel, Version 1.0.

OGC, 2018. Glossary of terms. accessed 2018-07-10. http://www.opengeospatial.org/ogc/glossary.

Patroumpas, K., N. Georgomanolis, T. Stratiotis, M. Alexakis, and S. Athanasiou, 2015. Exposing INSPIRE on the Semantic Web. *Journal of Web Semantics*, 35:53-62, doi: 10.1016/j.websem.2015.09.003.

Perry, Matthew, and John Herring, 2012. OGC GeoSPARQL-A geographic query language for RDF data.

Probst, Florian, Fernando Bibotti, and Andres Pazos, 2004. Connecting ISO and OGC Models to the Semantic Web | OGC Network. *The Third International Conference on Geographic Information Science: Extended Abstracts and Poster Summaries, Adelphi, MD, USA,*.

Probst, Florian, Andrew Gordon, and Ivan Dornelas, 2006. OGC Discussion paper: Ontology-based representation of the OGC observations and measurements model. In *Institute for Geoinformatics (ifgi)*: Citeseer.

Russomanno, David J., Cartik R. Kothari, and Omoju a. Thomas, 2005. Building a Sensor Ontology: A Practical Approach Leveraging ISO and OGC Models. *Proceedings of International Conference on Artificial Intelligence*.

Stuckenschmidt, H, U Visser, G Schuster, and T. Vögele, 1999. Ontologies for geographic information integration. *Proceedings of the workshop intelligent methods for handling enviromental information: Special aspects of processing space and time*.

Tand, Jeremy, Linda van den Brink, and Payam Barnaghi, 2017. Spatial Data on the Web Best Practices. In *W3C Working Group Note*.

van den Brink, Linda, Paul Janssen, Wilko Quak, and Jantien Stoter, 2014. Linking spatial data: automated conversion of geo-information models and GML data to RDF. *International Journal of Spatial Data Infrastructures Research*, 9:59-85, doi: 10.2902/.

Xu, Z. M., Y. Y. Ni, W. J. He, L. L. Lin, and Q. Yan, 2012. Automatic extraction of OWL ontologies from UML class diagrams: a semantics-preserving approach. *World Wide Web-Internet and Web Information Systems*, 15 (5-6):517-545, doi: 10.1007/s11280-011-0147-z.

Zedlitz, Jesper, and Norbert Luttenberger, 2012. Transforming Between UML Conceptual Models And OWL 2 Ontologies. *Proceedings of the Terra Cognita Workshop on Foundations Technologies and Applications of the Geospatial Web at the 11th International Semantic Web Conference ISWC 2012*.

Article 3

Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies

International Journal of
Geo-Information

MDPI

Article

Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies

Knut Jetlund [1,2,*], Erling Onstein [2] and Lizhen Huang [2]

1 Norwegian Public Roads Administration, 2318 Hamar, Norway
2 Department of Manufacturing and Civil Engineering, Norwegian University of Science and Technology, 2821 Gjøvik, Norway
* Correspondence: knut.jetlund@vegvesen.no; Tel.: +47-9596-2863

Received: 28 June 2019; Accepted: 19 August 2019; Published: 22 August 2019

check for updates

Abstract: This study aims to improve the implementation of models of geospatial information in Web Ontology Language (OWL). Large amounts of geospatial information are maintained in Geographic Information Systems (GIS) based on models according to the Unified Modeling Language (UML) and standards from ISO/TC 211 and the Open Geospatial Consortium (OGC). Sharing models and geospatial information in the Semantic Web will increase the usability and value of models and information, as well as enable linking with spatial and non-spatial information from other domains. Methods for conversion from UML to OWL for basic concepts used in models of geospatial information have been studied and evaluated. Primary conversion challenges have been identified with specific attention to whether adapted rules for UML modelling could contribute to improved conversions. Results indicated that restrictions related to abstract classes, unions, compositions and code lists in UML are challenging in the Open World Assumption (OWA) on which OWL is based. Two conversion challenges are addressed by adding more semantics to UML models: global properties and reuse of external concepts. The proposed solution is formalized in a UML profile supported by rules and recommendations and demonstrated with a UML model based on the Intelligent Transport Systems (ITS) standard ISO 14825 Geographic Data Files (GDF). The scope of the resulting ontology will determine to what degree the restrictions shall be maintained in OWL, and different conversion methods are needed for different scopes.

Keywords: geographic information systems; information model; model driven architecture; unified modelling language; web ontology language

1. Introduction

1.1. Geospatial Information

This article presents novel research continued from the study presented in [1].

Geospatial information is a vital part of the knowledge about real-world features and events. Sharing and reusing geospatial information and non-spatial information from heterogeneous sources is of paramount importance for domains such as Smart Cities [2]; disaster management [3,4]; construction and asset management [5–7]; and Life Cycle Assessment (LCA) [8]. One specific example is the domain of Intelligent Transport Systems (ITS), where sensors in vehicles and road-side equipment collect and share vast amounts of data about weather conditions, traffic events, regulations and road environments. The collected data depends on location references to become valuable information and needs to be combined with geospatial information from other sources in order to build the knowledge needed for legal and safe navigation [9].

Standards and specifications developed by the International Standardization Organization, Technical Committee 211 (ISO/TC 211) and the Open Geospatial Consortium (OGC) provide the foundation for structured geospatial information. Local, national and regional authorities, agencies and organizations worldwide have applied the standards for collecting and maintaining large amounts of structured geospatial information covering a wide range of purposes. Information models, databases and applications for geospatial information are based on standards from ISO/TC 211 and OGC [9,10].

Modelling of geospatial information according to ISO TC/211 and OGC standards is a process where a portion of the real world—known as a universe of discourse—is perceived in a context of geographic application and defined in a conceptual model. ISO/TC 211 uses the term "geographic information", while OGC uses the term "geospatial information". The terms "geographic" and "geospatial" are considered equivalent in this article. The conceptual model is represented in a conceptual schema, formalized by the Unified Modeling Language (UML) [11]. Figure 1 illustrates the process.

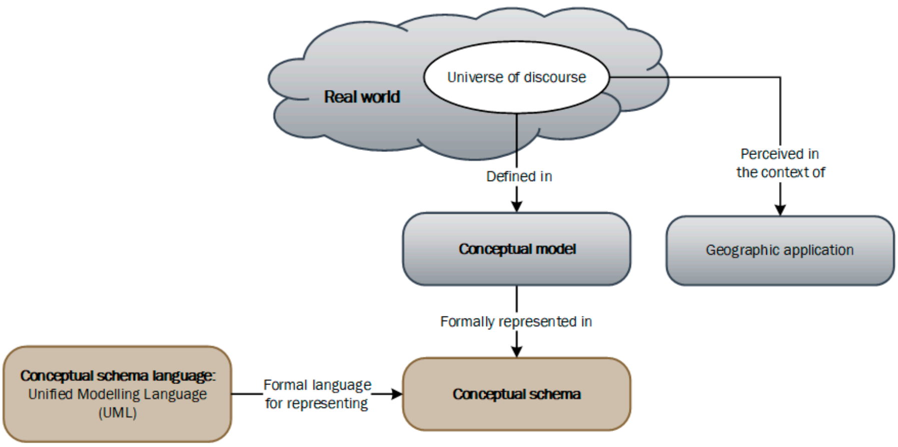


Figure 1. The ISO/TC 211 process of information modelling. Adapted from [12,13].

According to Model Driven Architecture (MDA) [14], the conceptual schemas shall be independent of any implementation technology but can be converted to implementation schemas in database and exchange formats. The core standards ISO 19103 [13] and ISO 19109 [12] define the ISO/TC 211 UML profile, the General Feature Model (GFM) for geospatial information and rules for semantics needed for conversion to implementation schemas. ISO 19136 [15] defines specific UML modelling rules and conversion rules for implementation in the exchange format GML. Figure 2 illustrates the four levels of abstraction in MDA according to ISO 19103: metamodels; abstract conceptual schemas; conceptual application schemas; and implementation schemas. Furthermore, the figure illustrates how conceptual schemas are converted to implementation schemas in the two formats GML and OWL.

Structured geospatial information based on ISO/TC 211 and OGC standards has mainly been stored and maintained in relational databases with extensions for geometry, topology and geospatial operations, and has been accessed through specialized GIS software [16]. Over the last 10 to 15 years, service standards such as Web Map Services (WMS) [17] and Web Feature Services (WFS) [18] have been used extensively to share geospatial information and information models on the World Wide Web. Spatial Data Infrastructures (SDI) provide portals and catalogue services for searching and accessing information from different service providers. However, even though WMS and WFS are based on general IT standards, they are specific standards for geospatial information, with limited possibilities for linking the information with other sources on the World Wide Web [19,20].

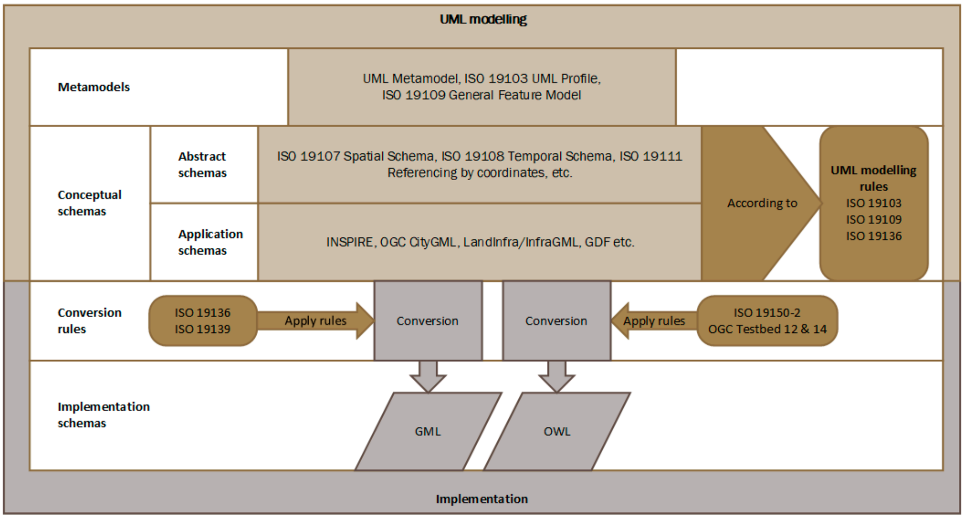


Figure 2. UML modelling and conversion to implementation schemas according to ISO/TC 211 and OGC standards.

1.2. Geospatial Information in the Semantic Web

The Semantic Web provides the concepts for structured and machine-readable descriptions of any kind of information on the World Wide Web, independent of applications and domains. Information is described in the Resource Description Framework (RDF) [21] as graphs of triples with subject, predicate and object. Information models are described as ontologies in the Web Ontology Language (OWL) [22], which is based on RDF.

Enabling information models and geospatial information based on ISO/TC 211 and OGC standards in RDF and OWL can bring more value to the domain of GIS and other domains as well as the Semantic Web in general [23]. Semantic Web technologies can be used to improve discovery in SDIs [24–29]. Authoritative geospatial information can be combined with less structured geospatial information from crowdsourcing, citizen science and sensors [20,30], and Semantic Web concepts for reasoning may be used to derive new knowledge [31–33]. Most importantly, geospatial information can be easier accessed from outside the GIS domain, linked with non-spatial information and reused between stakeholders and domains.

The availability of standardized geospatial information on the Semantic Web is still limited, but some effort has been put into the issue by OGC and W3C, who established The Spatial Data on the Web Working Group in cooperation in 2015 [34]. ISO/TC 211 developed the standard ISO 19150, part 2 [35] with rules for conversion from UML to OWL. The ISO/TC 211 Group for Ontology Management (GOM) [36] derived OWL ontologies for most ISO/TC 211 standards following the rules from ISO 19150-2. Further work has been done in INSPIRE [37] and OGC [38,39]. The Ordnance Survey in Ireland has published more than 50,000,000 spatial objects in their Prime2 database [40–42]. Mapping authorities in several other countries have published data sets or researched the subject as well, including, but not limited to; the United Kingdom [43], The Netherlands [44,45], Turkey [31], Spain [46] and Greece [47].

Some of the OWL ontologies for the published data sets have been converted from UML models, while others have been developed directly in OWL, parallel to existing UML models.

1.3. Contribution and Research Questions

This study investigates conversions from conceptual UML models according to ISO/TC 211 and OGC standards to OWL ontologies, as illustrated in Figure 2. The study aims to improve the process and achieve more precise implementation in OWL. Two research questions have been studied:

1. What are the primary challenges in conversions from UML models of geospatial information to OWL ontologies?
2. How can conversion challenges be overcome with adapted rules for UML modelling?

The scope of the study is limited to information models and conversion from UML to OWL. Challenges related to managing geospatial information in the Semantic Web, such as spatial indexing, complex geometries and operations are considered out of scope.

2. Materials and Methods

2.1. Literature Search

The primary purpose of the literature search was to identify relevant literature for answering the research questions. The results include literature where UML and OWL were compared and literature that described conversions from UML models to OWL ontologies. Besides, literature that described use cases and experiences with geospatial information in the Semantic Web was considered relevant for a broader view on the issue.

Combinations of keyword sets, listed in Table 1, were used in the search portals Oria, Web of Science and Google Scholar. Combinations of sets 1–3 were used in the first search for literature about geospatial UML models and OWL ontologies, preferably in the title: "1 and 2"; "2 and 3"; and "1 and 2 and 3", while sets 4 and 5 were used for refining the searches. Finally, keyword set 6 was used in combination with keyword sets 3 and 4 to find literature about structured geospatial information on the Semantic Web.

Table 1. Keyword sets for the literature search.

	Keyword Set	Purpose—Literature Mentioning
1.	"UML"	The abbreviation UML
2.	"OWL" OR "Ontology" OR "Ontologies"	The abbreviation OWL or ontologies in general.
3.	"Geospatial" OR "Geographic" OR "Spatial" OR "GIS"	Terms for geospatial information.
4.	("ISO AND "211") OR "OGC"	The geospatial standardization actors ISO/TC 211 and OGC.
5.	"Mapping" OR "Conversion" OR "Transformation"	Terms for conversion processes.
6.	"Semantic Web" OR "Linked Data"	Terms for the Semantic Web and Linked Data.

The initial search results were refined by studying titles and abstracts, while the final selection was made by studying the content. Inspection of cited literature (backward search) and literature that cited the selected literature (forward search) identified additional relevant literature. Besides, relevant standards, specifications and reports were found through searches in standards catalogues and on the World Wide Web. The final results from the searches are presented in Table 2.

Table 2. Selected literature per subject and literature type.

Subject	Literature Type					
	Book or Book Section	Conference Paper or Proceedings	Journal Article	Report or Standard	Web Page	Total
Geospatial information in the Semantic Web	4	8	34	4	3	53
Comparing UML and OWL	10	2	8	1	1	22
UML to OWL conversions in general	1	4	9	0	0	14
UML to OWL conversions for geospatial information	1	5	6	2	1	15

Results from the literature search on geospatial information in the Semantic Web are the foundation for the overall description in the introduction of this article. Results from the searches on UML and OWL are further discussed in the subsequent state-of-the-art study.

2.2. State of the Art

2.2.1. Comparing UML and OWL

Fundamental differences that are essential for conversions between UML and OWL were pointed out in our state-of-the-art study presented in [1]. One fundamental difference in modelling approaches is the assumption of an open or closed world. OWL and the Semantic Web is founded on an Open World Assumption (OWA) and the possibility for anyone to say anything about anything (AAA). An information model following the OWA describes the real world only as it is known at the time, and more information which is outside of the current knowledge may be added. No conclusions that are assuming all information is available can be drawn, and nothing is true or false unless explicitly stated [48]. UML is following a Closed World Assumption (CWA), where the information model is assumed to be a complete description of the real world in a given context (e.g., specific use of geographic information). A feature type in an ISO/TC 211 compliant UML model is a classification of real-world phenomena with common characteristics, perceived in the context of a geographic application. A feature instance in a data set shall be of one single feature type, with characteristics as single property types, as illustrated in Figure 3. Contrary, an individual instance in an open world of OWL can be linked to several classes and have a flexible set of properties [23]. Because of these different assumptions, UML models have implicit restrictions that may need to be specified if they shall be maintained in OWL.

Another key principle in the Semantic Web is linking, reusing and extending concepts from other ontologies—including ontologies from other domains. UML models are defined in a closed environment with less flexibility for linking to external concepts. UML models developed according to ISO/TC 211 standards are reusing concepts from other models, but the reuse is mainly limited to models from the domain of geospatial information.

One fundamental logical difference is that OWL is based on set theory and description logic, with set-based class constructors such as union, intersect and complement (e.g., an OWL class may be constructed as the union of two other classes, meaning that it contains all instances of the two classes). UML does not have set-based class constructors, but some of the implicit restrictions in UML models must be translated to class constructors in OWL. Furthermore, OWL classes and properties are individuals which may be queried in the same manner as, and in combination with, the individuals in the data set. A UML model must be realized as an implementation schema (e.g., XML or a database schema), and there is a clear distinction between the schema and the instances in the data set. Finally, the concepts for class properties are different in the two technologies: while each class in UML is the single owner of its attributes and associations, properties are individual concepts and globally defined in OWL, and may be assigned to any class.

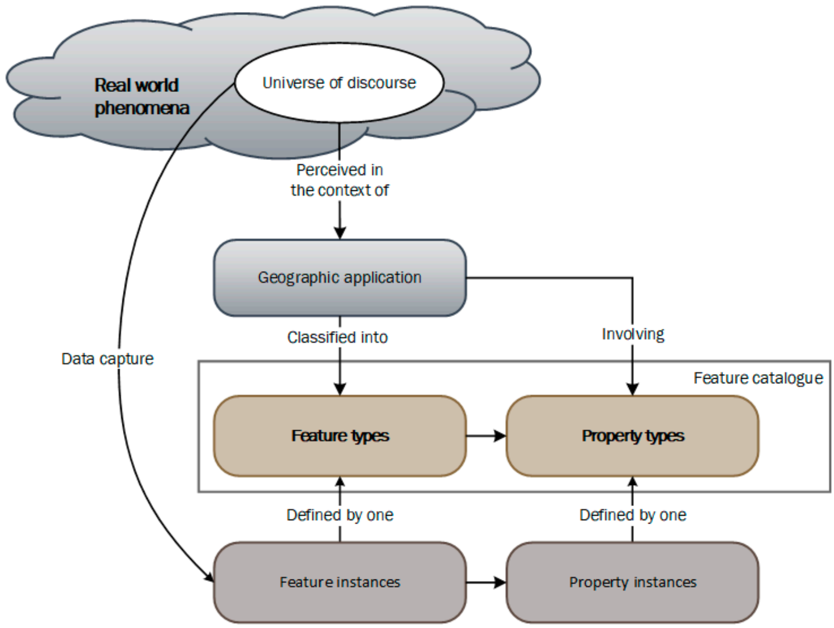


Figure 3. From the real world to data. Adapted from [12].

Despite the differences between UML and OWL, some effort has been put in developing UML profiles for applying UML as a tool for OWL ontology development [49–61]. The motivation has been to exploit the standardized graphical notation in UML—which is very useful for human communication. Some applications for ontology development include a graphical presentation of the ontologies, but there is no standardized graphical view equivalent to UML class diagrams. Most prominent among the UML profiles for OWL is the Ontology Definition Metamodel (ODM) [60], developed by the Object Management Group (OMG). The ODM contains a formal metamodel and UML profile for RDF and OWL.

A study that compared modelling of geospatial information in UML based on ISO/TC 211 MDA and OWL Ontologies was described in [31]. The UML-based model was considered the most robust and prepared model for use and linking with other datasets within SDIs, whereas the OWL-based model provided better support for sharing, discovery and linking with any other information on the Semantic Web. The authors of [31] suggested that both approaches should be used, but did not discuss whether models based on one approach should be the original and the other in a derived model, or whether the models should be developed and maintained in parallel.

2.2.2. Conversions from UML to OWL

The state-of-the-art study of conversions from geospatial UML models to OWL ontologies presented in [1] was extended for this article with work described in additional sources and more detailed analysis of conversion rules. The main source of the studies is the conversion rules defined in ISO 19150, part 2 [35]. The AR3NA project refined and modified the rules from ISO 19150-2 into guidelines for RDF encoding of geospatial information and models defined according to the INSPIRE Directive in Europe [37]. The OGC Testbeds 12 [38] and 14 [39] focused on specific geospatial UML to OWL issues and particularly on the implementation of rules in the conversion software ShapeChange [62]. Finally, several studies described conversions between UML and OWL in general [63–69] while other described specific conversions of UML models of geospatial information to OWL [70–72].

Table 3 summarizes rules for conversion from UML to OWL for fundamental concepts used in UML models for geospatial information. Further details are discussed in [1] and in subsequent sections.

Table 3. Summary of conversion rules.

UML Concept	OWL Concept	Conversion Rule Specification
Package	Ontology	Name and structure as in UML [35]. Name and structure from tagged values [37,38].
Class	Class	Direct conversion. Subclass of AnyFeature [35]. Mapping to external classes [38].
Class generalization	subclassOf	Direct conversion.
Abstract class	Not existing	isAbstract annotation property [35,37,38]. DisjointUnion axiom [72].
Primitive data type	DatatypeProperty	Matched to XSD Datatypes.
Structured data type	DatatypeProperty or ObjectProperty	Mapping to a few external types, else new class [35]. Mapping to specified external types, else new class [38]. Mapping to all similar external types, else new class [37].
Spatial data types	Data types defined in ISO 19107 [73] and GeoSPARQL [74]	Data types defined in the ISO 19107 ontology [35]. Mapping to GeoSPARQL data types [3,28,29,34,37,44]. Extending GeoSPARQL [31,41,42].
Enumeration	DataOneOf	Direct conversion. SKOS Concept Scheme [37,38].
Code lists	Several options	SKOS and allValuesFrom [35]. SKOS [37,38]. DataUnionOf and any value [72].
Union	Several options	Union [35]. Intersections and Union [37,38]. Flattening [38] Subproperty and ExactCardinality [72]. DisjointClasses [75].
Attribute and association role	Property	Simple conversion [68,69]. Globalization by similarity matching [71]. Globalization by prefix [35,38]. Global attributes with domain AnyFeature [35]. Globalization by manual matching [37,38]. Mapping to external properties [38,67].
Simple association	Domain and range	Direct conversion.
Aggregation	Not existing	Hierarchy of properties [66]. aggregationType annotation property [35].
Composition	Not existing	Hierarchy of properties [66]. InverseFunctional [67]. aggregationType annotation property [35].

2.2.3. Packages

The package concept in UML corresponds to an OWL ontology. However, packages in UML may also be used for an informal structure that does not need to be reflected in the implementation. ShapeChange can use tagged values on the packages to define the output structure, following recommendations from OGC Testbed 12.

2.2.4. Classes

The class concept exists with similar semantics in UML and OWL—including generalization and inheritance between classes. ISO 19150-2 specifies that classes defined as feature types according to ISO 19109 shall be declared as subclasses of the class AnyFeature from the ISO 19109 ontology.

OGC Testbed 12 discussed mapping from classes in UML models to equivalent classes already defined in OWL ontologies. The mapping was implemented in ShapeChange and can be defined in a configuration file.

Abstract classes in UML are used in generalizations, in which the abstract class is a superclass whose purpose is to define attributes, associations and operations that are common to all subclasses. The abstract class shall not have instances in an implementation, which is a type of restriction that does not exist in OWL. As discussed in [1], none of the identified conversion rules maintains the implicit restriction of the abstract class concept.

2.2.5. Data Types

Data types for attributes in UML may be equal or similar in models from different domains (e.g., an integer value is an integer value regardless of whether the model describes geospatial information or any other kind of information). As discussed in [1], ISO 19150-2 defines a mapping from a set of data types to XML Schema types. The INSPIRE Guidelines emphasize more reuse of data types from existing ontologies with an extended list of predefined mappings of commonly used data types. ShapeChange has implemented a configurable mapping that can be set individually for each model.

2.2.6. Spatial Data Types

The standard ISO 19107 [73] defines specific data types for geometry and topology, while ISO 19125-1 [76] defines a simplified profile of ISO 19107. Data types from ISO 19107 and ISO 19125-1 are reused in other ISO/TC 211 and OGC standards as well as other models based on the standards. ISO 19150-2 specifies that the ISO 19107 ontology—which has been derived from the ISO 19107 UML model—shall be used for spatial attributes in application schemas. A list of valid spatial object types is provided in the standard.

The OGC GeoSPARQL [74] specification includes a vocabulary for describing geometry in RDF and OWL. The INSPIRE Guidelines map data types from ISO 19107 to the generic GeoSPARQL Geometry class or one of its subclasses, which are ontology representations of the data types from either ISO 19125-1 or ISO 19136 (GML). This approach was used in several studies and projects for the publication of geospatial information in the Semantic Web as well. ShapeChange is configurable and can support several approaches. For practical use, there should not be any difference between using data types from GeoSPARQL or the ISO 19107 ontology, as they are initially based on the same conceptual model from ISO 19107.

2.2.7. Enumerations

ISO 19150-2 defines the conversion of UML enumerations to the “OneOf” axiom with a collection of values from the UML enumeration. OGC Testbed 12 discussed an alternative conversion where enumerations were treated as code lists. Both approaches have been implemented in ShapeChange. The INSPIRE Guidelines follow the rule from ISO 19150-2 for enumerations with self-describing codes that have a distinct meaning. In other cases, the enumeration shall be handled as code lists.

2.2.8. Code Lists

The code lists concept is explicitly defined in ISO 19103. Code lists are flexible enumerations, meaning that there can be other values than those described in the list. ISO 19150-2 defines the code list in OWL as three parts based on SKOS concept schemes. OGC Testbed 12 discussed several methods for handling code lists in OWL and focused on the management of the lists as external references outside of the ontology. The INSPIRE Guidelines follow the approach from OGC Testbed 12 with separate external SKOS concept schemes for the code lists. The ontologies include the URL to the SKOS Concept Scheme, but without any formal binding.

An alternative approach was described in [72], with a union of a “OneOf” list with the code list values and any other value, defined with a standard XML Schema expression.

2.2.9. Unions

The union concept is explicitly defined in ISO 19103 and is different from the set-based union class constructor in OWL. A union in UML models according to ISO 19103 is a list of several alternative datatypes, where one and only one shall be used for an attribute value. ISO 19150-2 defines the conversion of a UML union directly to an OWL union of classes. OGC Testbed 12 pointed out several weaknesses with this conversion rule and described a different approach that is also used in the INSPIRE Guidelines. They used a combination of intersections where only one member could have a value, combined in a union that combined all the intersections. Two other approaches where the union was flattened were described as well. They believed the flattened approaches would work better in Semantic Web software.

2.2.10. Attributes and Association Roles

UML classes can have two kinds of properties: attributes or associations to other classes. Both kinds of properties are equivalent to OWL properties and are handled as such in the studied conversion rules. However, as discussed in [1], one fundamental difference requires special treatment: properties in OWL are individual concepts that are globally scoped, while UML properties are unique within each class. Several classes in a UML model may have identical or almost identical properties, or they may have properties with an identical name, but different data type or definition. Identical UML properties should be handled as one global OWL property, while UML properties with identical names but different data type or definition need to be handled as separate OWL properties.

The core rule for globalization of UML properties in ISO 19150-2 is to add the class name as a prefix to the property name, which makes all properties locally defined for each class, but globally unique. Reused attributes can alternatively be defined without prefix, and it is suggested to use the generic class "AnyFeature" as the domain. The INSPIRE Guidelines supplies the rules with more reuse of existing properties from external ontologies. UML properties with identical or close to identical meaning are converted to OWL properties with a global scope, and UML properties that can be matched to properties already defined in other ontologies are converted to those properties in OWL. Similarity matching, as described in [71], may be used to reduce the manual work involved in the process of identifying potential global and external properties.

The issue of property globalization was discussed in OGC Testbed 12 as well, with four approaches that were implemented in ShapeChange: all properties locally defined with class name prefix as in 19150-2; all global properties defined in one UML class; global properties defined in a configuration file; or global properties identified by globally unique names. Besides, ShapeChange can be configured to map UML properties to externally defined OWL properties, as discussed in [44].

2.2.11. Associations

Aggregations and composite associations add more semantics to UML models. An UML aggregation defines one class as a part of a main class (the whole) but does not add any restrictions to the model. The aggregation merely informs about a closer relationship. A UML composition defines a more restricted relationship between two classes. An instance of a part class in a composition can only be related to one whole instance. The restriction defined by a composition does not exist directly in OWL, and, as discussed in [1], none of the identified conversion rules maintains the restriction.

3. Results

3.1. Semantics in UML Models for Implementation as OWL Ontologies

Model-driven conversion from conceptual UML models to implementation schemas as illustrated in Figure 2 relies on a specific use of concepts and inclusion of specific semantics in the UML models. Rules defined in ISO 19103, ISO 19109 and ISO 19136 state how general UML concepts shall be used for models of geospatial information, supplemented with specialized UML stereotypes and tagged values for additional semantics. The standards ISO 19103 and ISO 19109 defines general rules for UML models of geospatial information, while ISO 19136 describes specific rules for UML models that shall be implemented as GML schemas. However, specific rules for semantics that could support conversion to OWL ontologies are not defined in any of the identified standards.

Most of the studies described in the state-of-the-art study focused on rules for conversion from UML models to OWL ontologies. Less attention was put on how to develop UML models that are prepared for implementation in OWL. The issue was briefly discussed in [28], where they suggested that conversion challenges could be overcome by improving the original UML models. Semantics for linking UML attributes to externally defined OWL properties were suggested in [44], while OGC Testbed 12 discussed and implemented conversion rules in ShapeChange, based on specific tagged values. The UML profiles for OWL described in [49–61] specified how UML could be used to develop RDF and OWL ontologies specifically. However, the profiles differ significantly from the profile in ISO 19103 and cannot be applied to models according to ISO/TC 211 standards. Furthermore, the MDA approach in ISO/TC 211 standards is to develop conceptual models that can be implemented in several formats, not models scoped for one single implementation format.

Two of the core challenges for the conversion from a closed UML world to an open OWL world can potentially be overcome by adding more semantics to UML models: defining attributes in UML for implementation as global properties in OWL, and linking internal UML classes, data types and properties to existing external OWL concepts.

3.2. Extended UML Profile for Geospatial Information

Additional semantics in UML models are handled by profiles that extend concepts in the UML metamodel. The primary construct to be used in profiles is the stereotype. Stereotypes extend existing metaclasses and can specify semantics as properties for UML concepts, referred to as tagged values [11]. The formal UML profile for geospatial information is defined in ISO 19103. ISO 19109 and ISO 19136 describe additional semantics as well—but not as formal profiles. Furthermore, the INSPIRE Generic Conceptual Model [77] describes specific tagged values for INSPIRE information models. Finally, the conversion rules developed in OGC Testbed 12 and implemented in ShapeChange use some tagged values not specified in standards.

We suggest an extended UML profile for geospatial information that includes semantics needed for conversion to OWL. The profile is based on ISO 19103, ISO 19109 and ISO 19136, and includes extensions from INSPIRE and ShapeChange. Figure 4 illustrates the extended UML profile, while the tagged values that are added in the profile are described in Table 4. Subsequent sections describe how the extended semantics can be utilized for improved conversion from UML to OWL.

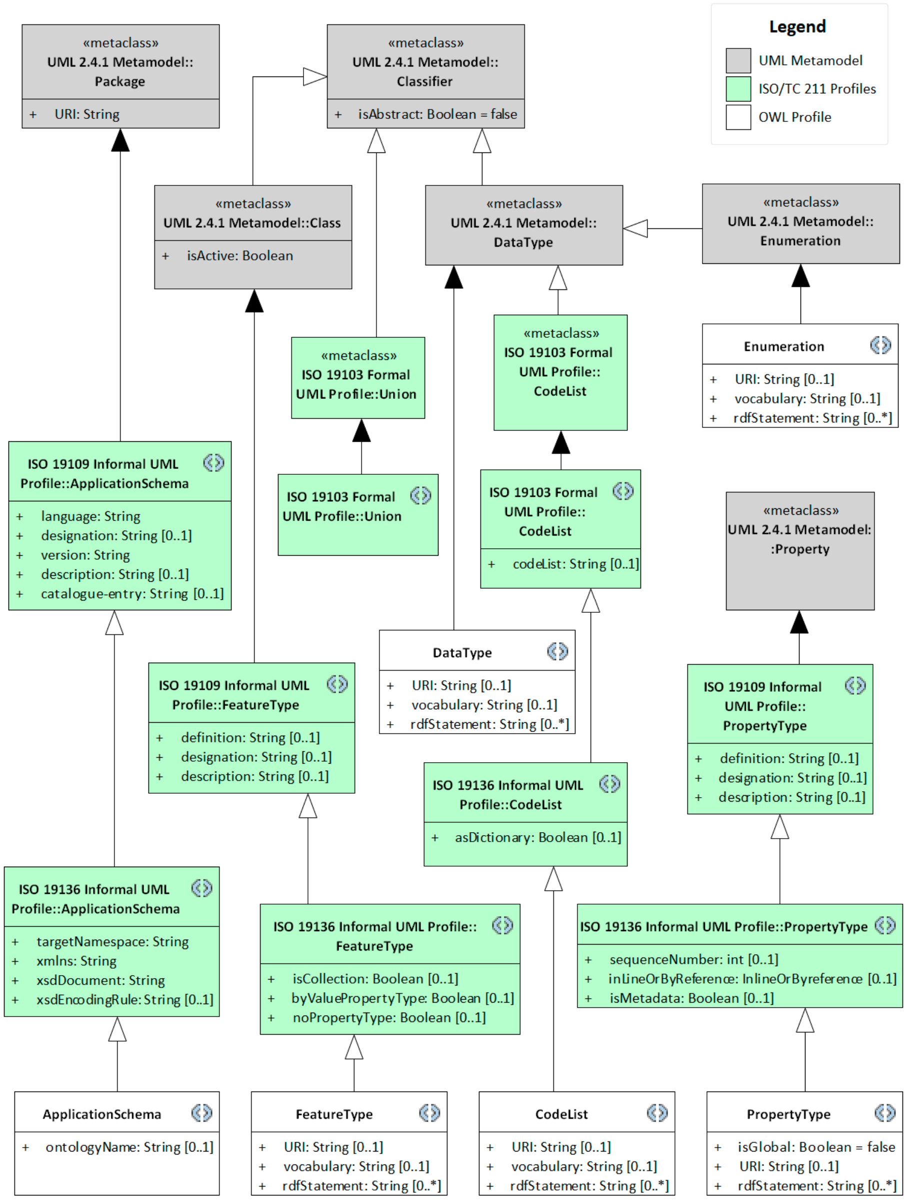


Figure 4. Extended UML Profile for geospatial information.

Table 4. Tagged values in the extended UML profile.

Tagged Value	Extended Concepts	Description
ontologyName	Package (ApplicationSchema)	Ontology name, if different from both package name and xsdDocument.
URI	Class (FeatureType) DataType Enumeration CodeList PropertyType	URI for the element, for internal and external references.
vocabulary	Class (FeatureType) DataType Enumeration CodeList	Reference to an external vocabulary that will replace the internal concept in an OWL implementation. Specified for INSPIRE and implemented in ShapeChange.
rdfStatement	Class (FeatureType) DataType Enumeration CodeList PropertyType	One or more RDF statements for linking internal and external concepts.
isGlobal	PropertyType	Specifies if the scope of a property is local or global. Default set to false.

3.3. Global Properties in UML

The rules for converting UML attributes to OWL properties, defined in ISO 19150-2 and OGC Testbed 12, can be used to ensure that OWL properties are globally unique. Each OWL property will either be assigned to one specific class (locally scoped) or any class (globally scoped). However, the rules do not cover the situation where an attribute in UML is reused for several but not all classes—which is a typical situation in UML.

Figure 5 shows examples of reuse in a UML Model based on the ITS standard ISO 14825 Geographic Data Files (GDF) [78]. The model has been modified to become conformant to ISO 19109, and to fit for implementation in OWL. The example contains two abstract superclasses and five implementable classes: The abstract class "RoadFurniture" with subclasses "TrafficSign", "PedestrianCrossing" and "Lighting" and the abstract class "PublicTransportFeature" with subclasses "StopPoint" and "RoutePoint". Two attributes are reused in several classes: The attribute "displayClass" is used in the classes "TrafficSign" and "PedestrianCrossing", while the attribute "accessibility" is used in the two classes "StopPoint" and "PedestrianCrossing". These attributes should be globally unique properties in an OWL implementation, assigned only to the classes they are part of in the UML model.

One approach to avoid duplication of attributes in UML classes is to define superclasses where the reused attributes are defined, with generalization associations from each class to the superclass. However, this is not a viable solution for the example of reuse in Figure 5, as each of the two attributes "displayClass" and "accessibility" shall only be used in two out of five classes. One would need to have multiple combinations of inheritance from one superclass with the "displayClass" attribute and one with the "accessibility" attribute. A larger model with many combinations of reuse would become an intricate spider web of superclasses and generalizations.

OGC Testbed 12 discussed another approach for handling reused attributes, as an issue for future improvement of the conversion rules. They suggested that such attributes could be identified with tagged values in UML and handled in OWL with a union of all classes that have the attribute in the UML model. Our solution follows up the suggestion from OGC Testbed 12 and combines it with one of the approaches that are implemented in ShapeChange, which is the use of a specific class for common attribute concepts. We suggest using a specific and abstract class that contains the original description of global attributes, as illustrated in Figure 6. The class is called "AttributeCatalogue" in the example model and is simply a container for global attributes. The class shall not have any instances in an implementation.

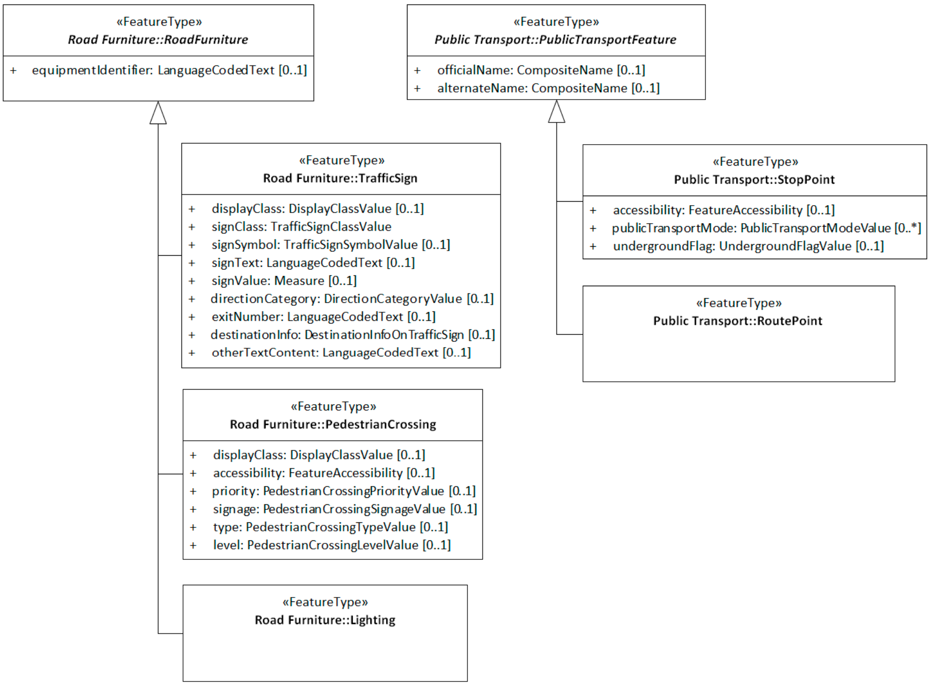


Figure 5. Examples of classes based on ISO 14825 Geographic Data Files (GDF) with reused attributes.

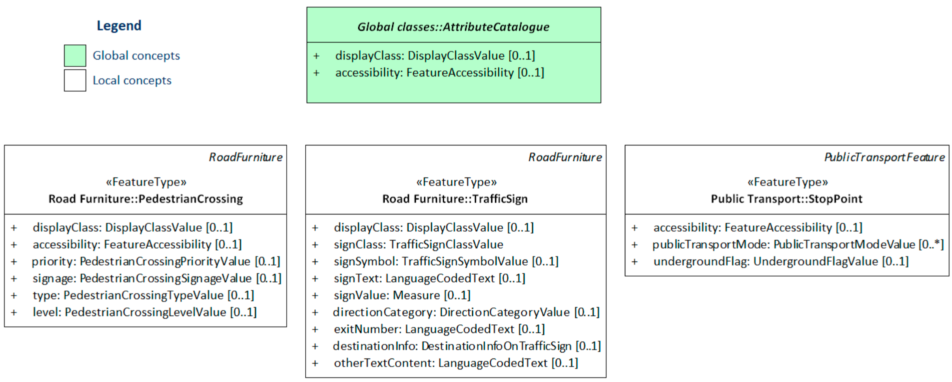


Figure 6. "AttributeCatalogue" and classes with reused attributes.

Each global attribute must be duplicated on the classes where it shall be reused as UML has no way of connecting attributes from one class to another class. Global attributes are assigned with two tagged values in the "AttributeCatalogue" class as well as in the classes where they are reused, as illustrated in Table 5. The tagged value "isGlobal" identifies whether the attributes are global or not, while "URI" stores the global identifier that connects the original and the reused copies. A script can keep all reused copies of the attributes updated from the originals in "AttributeCatalogue", by referring to the "URI" tagged value.

The conversion process from UML to OWL can use the tagged values to identify the reuse of global attributes. The attributes are converted to OWL properties and assigned with a domain, where the domain is a union of the UML classes that reuse the attributes. The conversion in this study was done manually, but a future improvement would be to implement the conversion in ShapeChange, as suggested in OGC Testbed 12.

Figure 7 shows an extract of the OWL ontology with the domain assignment for the two attributes "displayClass" and "accessibility". Figure 8 shows a graphical view of the three classes with the two superclasses, the two unions and properties.

Table 5. Tagged values for attributes from Figure 6.

Attribute Name	Tagged Values	
	isGlobal	**URI**
AttributeCatalogue.displayClass	true	gdf#displayClass
AttributeCatalogue.accessibility	true	gdf#accessibility
PedestrianCrossing.displayClass	true	gdf#displayClass
PedestrianCrossing.accessibility	true	gdf#accessibility
PedestrianCrossing.priority	false	
PedestrianCrossing.signage	false	
PedestrianCrossing.type	false	
PedestrianCrossing.level	false	
TrafficSign.displayClass	true	gdf#displayClass
TrafficSign.signClass	false	
TrafficSign.signSymbol	false	
TrafficSign.signText	false	
TrafficSign.signValue	false	
TrafficSign.directionCategory	false	
TrafficSign.exitNumber	false	
TrafficSign.destinationInfo	false	
TrafficSign.otherTextContent	false	
StopPoint.accessibility	true	gdf#accessibility
StopPoint.publicTransportMode	false	
StopPoint.undergroundFlag	false	

```
<!-- gdf#accessibility -->
<owl:ObjectProperty rdf:about="gdf#accessibility">
  <rdfs:range rdf:resource="gdf#FeatureAccessibility" />
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <rdf:Description rdf:about="gdf#PedestrianCrossing" />
        <rdf:Description rdf:about="gdf#StopPoint" />
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<!-- gdf#displayClass -->
<owl:ObjectProperty rdf:about="gdf#displayClass">
  <rdfs:range rdf:resource="gdf#DisplayClassValue" />
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <rdf:Description rdf:about="gdf#PedestrianCrossing" />
        <rdf:Description rdf:about="gdf#TrafficSign" />
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
```

Figure 7. OWL domain assignment for reused properties.

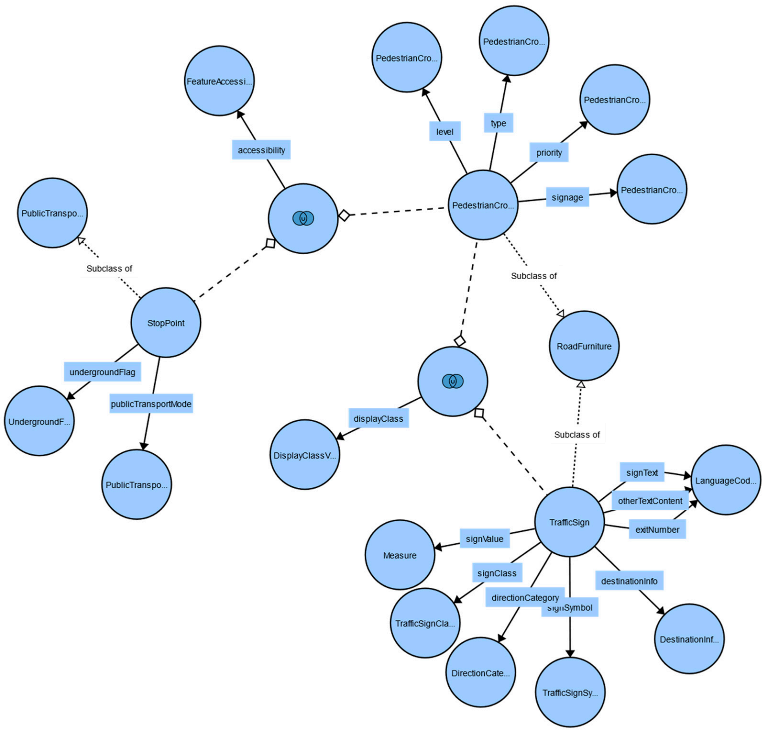


Figure 8. Graphical view of the OWL implementation of Figure 6 (generated with WebVOWL [79]).

3.4. Linking to External Concepts

Mapping of data types and classes to existing external concepts was discussed in OGC Testbed 12 and has been implemented in ShapeChange with configurable settings where the classes that shall be mapped are specified in configuration files. A method for linking UML concepts to external concepts was suggested in [44], where elements were linked to existing ontologies with tagged values. They implemented the conversion in ShapeChange through an extension. Our solution follows up the method from [44] and includes mapping to external concepts as well as linking. The tagged value "vocabulary" specifies the URI of an external concept that the internal concept shall be mapped to, while "rdfStatement" can contain an RDF statement that links the internal concept to an external concept. The tagged value "vocabulary" was specified in the INSPIRE Generic Conceptual Model [77] and is already implemented in ShapeChange rules.

Figure 9 shows the class "TrafficSign" from Figure 6 with the data type "LanguageCodedText", which is the data type for three attributes in the class: "signText", "exitNumber" and "otherTextContent". The data type "LanguageCodedText" is designed for the language-specific representation of a text, which is a common issue in information models from many domains. One example of a data type that resembles "LanguageCodedText" is the class "GeographicName" from the INSPIRE Specification for Geographical Names [80]. The OWL representation of "GeographicName" was discussed in the AR3NA project [81] where the INSPIRE Guidelines for the RDF encoding of spatial data [37] was developed. The Guidelines states that the class "GeographicName" can be simplified to the concept "rdfs:Literal" if only the string is needed, and to the subclass "rdfs:langString" if the language is needed. A related approach was used in [44], where the RDF statement "owl:equivalentProperty = rdfs:label" was added to a "name" attribute as a tagged value. The range of the "rdfs:label" property is "rdfs:Literal", which links the "name" attribute to the "rdfs:Literal" and "rdfs:langString" concepts.

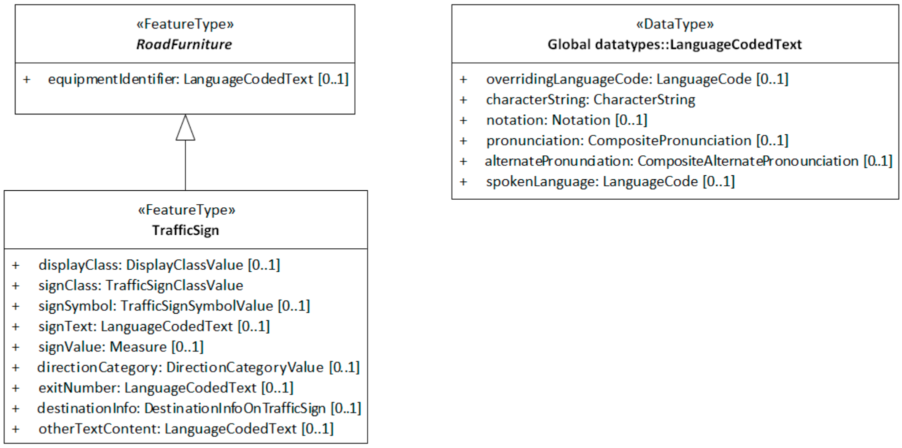


Figure 9. The class "TrafficSign" and the data type "LanguageCodedText".

Links from the language coded attributes and the data type "LanguageCodedText" in Figure 9 to general RDFS concepts can be specified at two levels: Each attribute can be linked to the property "rdfs:label" as in [44], while the data type can be linked to the class "rdfs:Literal". A link from the data type will be more generic and link all attributes with this data type to the external concept. A link from each attribute will reduce the number of levels needed to link the specific model to the external concept but will also require more work on specifying the RDF statements. Furthermore, links at the attribute level may also be used to separate attributes that can be linked from those that need more specific characteristics from the internal data type. Examples of RDF statements for linking attributes and data types to external concepts are shown in Table 6, while Figure 10 shows the implementation in OWL.

The "equivalent" statements in Table 6 and Figure 10 imply that classes or properties represent the same concepts but are not necessarily equal. If the data type had been equal to an external concept, the link could be specified tighter with the equality statement "sameAs". Other RDF statements such as "subclassOf" or "subPropertyOf" may also be specified, depending on the relation between the internal and external concepts.

Table 6. The tagged value "rdfStatement" for classes and attributes from Figure 9.

Class or Attribute Name	Tagged Value "rdfStatement"
LanguageCodedText	owl:equivalentClass rdf:resource = "rdfs:Literal"
TrafficSign	
TrafficSign.displayClass	
TrafficSign.signClass	
TrafficSign.signSymbol	
TrafficSign.signText	owl:equivalentProperty rdf:resource = "rdfs:label"
TrafficSign.signValue	
TrafficSign.directionCategory	
TrafficSign.exitNumber	owl:equivalentProperty rdf:resource = "rdfs:label"
TrafficSign.destinationInfo	
TrafficSign.otherTextContent	owl:equivalentProperty rdf:resource = "rdfs:label"

```
<!-- gdf#LanguageCodedText -->
 <owl:Class rdf:about="gdf#LanguageCodedText">
      <owl:equivalentClass rdf:resource="rdfs:Literal"/>
 </owl:Class>
 </owl:ObjectProperty>
 <!-- gdf#signText -->
 <owl:ObjectProperty rdf:about="gdf#signText">
      <owl:equivalentProperty rdf:resource="rdfs:label"/>
      <rdfs:range rdf:resource="gdf#LanguageCodedText"/>
      <rdfs:domain rdf:resource="gdf#TrafficSign"/>
 <!-- gdf#exitNumber -->
 <owl:ObjectProperty rdf:about="gdf#exitNumber">
      <owl:equivalentProperty rdf:resource="rdfs:label"/>
      <rdfs:range rdf:resource="gdf#LanguageCodedText"/>
      <rdfs:domain rdf:resource="gdf#TrafficSign"/>
 </owl:ObjectProperty>
 <!-- gdf#otherTextContent -->
 <owl:ObjectProperty rdf:about="gdf#otherTextContent">
      <owl:equivalentProperty rdf:resource="rdfs:label"/>
      <rdfs:range rdf:resource="gdf#LanguageCodedText"/>
      <rdfs:domain rdf:resource="gdf#TrafficSign"/>
 </owl:ObjectProperty>
```

Figure 10. OWL Implementation of the tagged value "rdfStatement".

An alternative to defining equality with the "sameAs" statement is to specify reuse of external concepts in the UML model. The internal concept will then be a copy of the external concept, included in the UML model only to make the model complete for implementation in other formats than OWL. The code lists "LanguageCode" for the attribute "overridingLanguageCode" in the data type "LanguageCodedText" is an example of a data type that might be reused from external concepts. The code list is defined in an annex of the GDF standard and contains three-letter language codes according to the ISO standard ISO 639-2 [82]. The Library of Congress is the registration authority for ISO 639-2, and have made available a vocabulary for the language codes [83]. A link to the external vocabulary can be added to the UML model with the tagged value "vocabulary". Table 7 shows the tagged values for the code list.

Table 7. Tagged value "vocabulary".

Class Name	Tagged Value "Vocabulary"
LanguageCode	http://id.loc.gov/vocabulary/iso639-2
TrafficSignClassValue	https://git.io/fjwVy
TrafficSignSymbolValue	https://git.io/fjwVS

Two attributes in the class "TrafficSign" have data types that are candidates for the reuse of external concepts as well. These are the attributes "signClass" with code list "TrafficSignClassValue" and "signSymbol" with code list "TrafficSignSymbolValue". The standard ISO 14823 [84] defines a data dictionary for traffic signs with sign classes (service categories) and symbols (pictogram category).

These classifications should be reused in other models with traffic signs. There is no official vocabulary derived from the standard, but SKOS Concept Schemes has been developed for this study [85]. Links can be added to the UML model in the tagged value "vocabulary". Table 7 shows the tagged values for the two code lists.

Figure 11 shows the implementation in OWL for the three code lists. The OWL implementation has been done manually for the example in this study. However, conversion from UML code lists to vocabulary references in OWL are implemented in ShapeChange and have been used for the INSPIRE ontologies. The ranges of the properties are set to the generic concept "skos:Concept", while the statement "seeAlso" provides a link to the vocabulary, like in the INSPIRE ontologies.

```
<!-- gdf#overridingLanguageCode -->
<owl:ObjectProperty rdf:about="gdf#overridingLanguageCode">
    <rdfs:domain rdf:resource="gdf#LanguageCodedText"/>
    <rdfs:range rdf:resource="skos:Concept"/>
    <rdfs:seeAlso rdf:resource="http://id.loc.gov/vocabulary/iso639-2"/>
</owl:ObjectProperty>
<!-- gdf#signClass -->
<owl:ObjectProperty rdf:about="gdf#signClass">
    <rdfs:domain rdf:resource="gdf#TrafficSign"/>
    <rdfs:range rdf:resource="skos:Concept"/>
    <rdfs:seeAlso rdf:resource=" https://git.io/fjwVy"/>
</owl:ObjectProperty>
<!-- gdf#signSymbol -->
<owl:ObjectProperty rdf:about="gdf#signSymbol">
    <rdfs:domain rdf:resource="gdf#TrafficSign"/>
    <rdfs:range rdf:resource="skos:Concept"/>
    <rdfs:seeAlso rdf:resource=" https://git.io/fjwVS"/>
</owl:ObjectProperty>
```

Figure 11. OWL Implementation of the tagged value "vocabulary".

4. Discussion

4.1. The Scope of Ontologies for Geospatial Information

The first and fundamental step in ontology development is to determine the domain and scope of the ontology. Defining the scope includes identifying what the ontology shall be used for and what type of questions the ontology shall answer [86]. For conversion from UML to OWL, the scope of the ontology is essential for deciding whether all restrictions from the closed UML world need to be maintained in the open OWL world, and for deciding whether or not similar concepts from existing ontologies can be reused.

We identified three main scopes for geospatial ontologies in [1] and will refer to them in further discussions as levels of information flow for which the ontologies may be used:

1. Use in Semantic Web technology and applications only.
2. Unidirectional information exchange from GIS applications to the Semantic Web.
3. Bidirectional information exchange between GIS applications and the Semantic Web.

The first level implies a decoupling from existing GIS models and applications, in which case restrictions are little needed, and the ontologies can comply with the OWA. This is not a likely situation for most of the structured geospatial information on a short term basis, as GIS databases and applications have complex and advanced functionality that will not be easily replaced with Semantic Web applications [23,34].

The second level covers the most discussed scopes (e.g., in [31,34,37]), where the original information is maintained in GIS databases and transformed to RDF for publication on the Semantic Web. Some restrictions must be applied to describe the information distinctly, but the most complex restrictions are not needed, as the ontologies will only describe existing information.

The third level implies the most complete conversion and a strict need for maintenance of restrictions. Most restrictions from UML should be maintained in the OWL implementation to ensure that the information is valid according to the closed schema in GIS databases. An example of bidirectional exchange was described in [5,6], where information was exchanged between application domains, stakeholders and lifecycle phases for construction and asset management.

Model-driven conversions from UML to OWL need to consider the different levels of information flow and must be configured according to the scope of the ontology.

4.2. Challenges for Conversions from UML to OWL

The first research question in this study asked for the primary challenges in conversions from UML models of geospatial information to OWL ontologies. The results from the state-of-the-art study indicate that conversion rules handle the basic concepts packages, classes, generalizations, primitive data types, spatial data types, enumerations and simple associations consistent and thorough. Conversion of abstract classes, unions, compositions and code lists are handled with several methods that maintain the UML restrictions to various degrees. The choice of conversion method for these concepts will depend on the scope of the resulting ontologies and the level of information flow. Maintaining all restrictions require more complicated conversion methods and is not necessary for all levels. Table 8 describes conversion challenges and discusses possible solutions for the different levels of information flow.

4.3. Rules for UML Modelling

The second research question in the study asked for adapted rules for UML modelling that could improve the conversion from UML models to OWL ontologies. The state-of-the-art study indicates that no standards have defined specific rules for semantics needed in UML for conversion to OWL. Two specific challenges have been pointed out in this study as candidates for improvements: conversion from UML attributes to global properties in OWL and linking to existing external concepts. We suggest an extended UML profile for geospatial information, which includes semantics for improved conversion from UML to OWL. Furthermore, Table 9 discusses rules and recommendations to ensure better implementation of UML models of geospatial information in OWL.

Table 8. Conversion challenges.

Concept	Level of Information Flow	Description	Discussion
Abstract classes	1 and 2	No restrictions needed.	The information given by the "isAbstract" annotation in ISO 19150-2 will be satisfactory.
	3	Restrictions must prevent instances of the classes.	None of the described conversion rules maintains the restrictions. Only the properties from the abstract class should be implemented in OWL, and not the class itself. The domain of the properties should then be set to a union of the subclasses, similar to our suggested solution for reused properties.
Unions	1 and 2	No restrictions needed.	The OWL "union" defined in ISO 19150-2 will be satisfactory.
	3	Restrictions are needed to define valid data types.	Conversion of UML unions may be the most complex challenge for implementation in OWL, and several methods have been suggested. The method described in the INSPIRE Guidelines and OGC Testbed 12 maintains the restrictions of a union, but it has a complicated representation in OWL. It was questioned in OGC Testbed 12 whether Semantic Web software could handle the complexity of the solution. Two alternative and simplified solutions were suggested in OGC Testbed 12 as well. Existing literature is not clear regarding what situations the different solutions should be used for. Further studies may be needed to achieve experiences from implementations and give recommendations for when to use different solutions.
Compositions	1 and 2	No restrictions needed.	The informative "aggregationType" annotation defined in ISO 19150-2 will be satisfactory.
	3	Restrictions are needed to ensure that a part instance is related to only one whole instance.	A stricter conversion with an "InverseFunctional" restriction was suggested in [67], while [66] defined a hierarchy of association types. Combining the approaches from [66] and [67] with the annotation from ISO 19150-2 may be a better solution for maintaining the implicit restrictions.
Code lists	1 and 2	No restrictions needed.	The open approach defined in OGC Testbed 12 and implemented in INSPIRE, with an informative reference to a vocabulary statement, will be satisfactory.
	3	Restrictions should refer to valid predefined values in the model or an external vocabulary, but also be open for using other values.	The ISO/TC 211 ontologies exclude additional values, while the INSPIRE ontologies have a very open approach. The approach in INSPIRE is very flexible as the resource can be anything, but the flexibility comes with the cost of reduced possibilities for the direct use of the predefined values. A possible improvement is to combine the union approach from [72] with a closer binding to clearly defined external SKOS Concept Schemes.

Table 9. Rules and recommendations for UML modelling.

Concept	Level of Information Flow	Description	Discussion
Global properties	All	Attributes that are identical in several UML classes should be converted to global properties in OWL.	Original definitions of such attributes should be maintained in one specific and abstract UML class. The class is called "AttributeCatalogue" in this study. The global attributes are reused in individual classes as copies of the original from "AttributeCatalogue". The class shall not be implemented in OWL, but the attributes in the class are implemented as global properties. A specific class for attributes that can be reused in other classes is a valuable approach independently of implementation in OWL as well. Models become easier to understand and implement when identical characteristics of different real-world features are defined in a harmonized manner.
	All	Identification of globally defined attributes.	The tagged value "isGlobal" shall identify the attribute as global. The tagged value "URI" shall be used to uniquely identify each attribute and link originals and reused copies.
	3	Restrictions must ensure that properties are assigned to specific classes.	Properties shall be linked to specific classes through a domain which is restricted to a union of the involved classes.
	All	Names are converted to URIs in OWL. UML property names are unique within each class, while OWL properties are globally scoped.	Properties (attributes and association roles) should have unique names within a UML package that shall be implemented as an ontology. This recommendation is stated in ISO 19109 with identifier /rec/general/property-name as well.
	All	Names are converted to URIs in OWL, and URIs are not always treated as case-sensitive.	Names of properties and classes should be non-case-sensitive unique. An example from the UML model used in the study is that the code lists have been given a suffix "Value" (e.g., "DisplayClassValue" for the attribute "displayClass").
Reuse of external concepts	All	Reuse of existing concepts is a vital part of information modelling for the Semantic Web [86].	Information modelling in UML is conducted in a closed environment and depends on concepts available in the model. However, reuse of existing concepts is a good practice that should be applied to UML models as well. Existing concepts can be duplicated in the UML models, and links to existing externals vocabularies can be added as tagged values.
	All	Links to external concepts.	The tagged value "rdfStatement" shall be used for linking internal and external concepts through a valid RDF statement.
	All	Mapping to external concepts.	The tagged value "vocabulary" shall be used for identifying the URI of external concepts that the internal concept shall be mapped to.
	3	Precise concepts are needed.	Mapping to external concepts should be done with care at this level. The mapping might lead to a loss of information in an exchange, due to differences between UML data types and external data types.

5. Conclusions and Further Work

This study has analyzed methods and rules for conversion from UML models of geospatial information to implementation as OWL Ontologies. Information models for geospatial information are developed using UML, based on standards from ISO/TC 211 and OGC. A fundament for enabling geospatial information for the Semantic Web will be to convert the UML models to OWL ontologies. A state-of-the-art study of conversion rules as described in standards and research indicates that basic concepts from UML can be converted to OWL. However, UML models of geospatial information are developed in a closed world in a geospatial context, while a core principle for OWL is the Open World Assumption (OWA). The UML concepts of abstract classes, unions, compositions and code lists represent closed world restrictions that must be treated specially to be maintained in the ontology. None of the existing conversion methods fully maintain these restrictions. Possible improvements have been suggested in this study and may be followed up in further studies.

Furthermore, attributes and associations in UML are uniquely owned by each class, while properties in OWL are globally scoped and may be used by many classes. This study suggests that global attributes should be treated specially in UML models, and specifies the use of tagged values to add semantics for converting them to global properties in OWL. Finally, an essential principle in ontology development is an extensive reuse of concepts from existing ontologies, while UML models mainly use concepts that are internally defined in the model. For this matter, mapping to existing ontologies has been defined in several ways. This study suggests a specific use of tagged values to add semantics to UML models for linking and mapping internal UML concepts to external OWL concepts.

The scope of the ontology is essential for the precision needed in the conversion. Ontologies that shall only be used within the Semantic Web or for publishing information from geospatial databases on the Semantic Web do not need to maintain all restrictions from UML and can apply mapping to external concepts. Ontologies that shall be used for updating geospatial databases with content from the Semantic Web need to be more precise and maintain both restrictions and specific internal concepts, to avoid instances that are invalid according to the UML model. Conversion settings that can be configured according to the scope of the ontology is a necessary fundament for a model-driven implementation as OWL ontologies.

Author Contributions: Conceptualization, K.J.; methodology, K.J., E.O. and L.H.; validation, K.J.; formal analysis, K.J.; investigation, K.J.; data curation, K.J.; writing—original draft preparation, K.J.; writing—review and editing, E.O. and L.H.; visualization, K.J.; supervision, E.O. and L.H.

Funding: This research received no external funding.

Acknowledgments: Rune Hjelsvold provided invaluable support to the state-of-the-art-study for this article. NTNU Open Access publishing funds cover the article processing charges.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jetlund, K. Improvements in Automated Derivation of owl ontologies from geospatial uml models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2018**, *XLII-4*, 283–290. [CrossRef]
2. Dangol, A.; Dewaelheyns, V.; Steenberghen, T. Why Geospatial Linked Open Data for Smart Mobility? In *REAL CORP 2016–SMART ME UP! How to Become and How to Stay a Smart City, and Does This Improve Quality of Life? Proceedings of the 21st International Conference on Urban. Planning, Regional Development and Information Society, Hamburg, Germany, 22–24 June 2016*; CORP—Competence Center of Urban and Regional Planning: Vienna, Austria, 2016.
3. Prudhomme, C.; Homburg, T.; Ponciano, J.-J.; Boochs, F.; Cruz, C.; Roxin, A.-M. Interpretation and automatic integration of geospatial data into the Semantic Web. *Computing* **2019**, 1–27. [CrossRef]
4. Zhang, C.R.; Zhao, T.; Anselin, L.; Li, W.D.; Chen, K. A Map-Reduce based parallel approach for improving query performance in a geospatial semantic web for disaster response. *Earth Sci. Inform.* **2015**, *8*, 499–509. [CrossRef]
5. Luiten, B.; Böhms, M.; O'Keeffe, A.; van Nederveen, S.; Bakker, J.; Wikström, L. A hybrid linked data approach to support asset management. In *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure: Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016)*; CRC Press, Taylor & Francis Group: London, UK, 2016.
6. Luiten, B.; O'Keeffe, A.; Stolk, S.; Wikström, L.; Weise, M. *Interlink D4. Principles for a European Road OTL*; CEDR: Brussels, Belgium, 2017.
7. Karan, E.P.; Irizarry, J. Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services. *Autom. Constr.* **2015**, *53*, 1–12. [CrossRef]
8. Hitzler, P.; Janowicz, K.; Krisnadhi, A.A. Ontology Modeling with Domain Experts: The GeoVocamp Experience. In Proceedings of the Diversity++@ ISWC 2015, Bethlehem, PA, USA, 12 October 2015.

9. Jetlund, K.; Onstein, E.; Huang, L. Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model. *Isprs Int. J. Geo Inf.* **2019**, *8*, 141. [CrossRef]
10. Coetzee, S.; Plews, R.; Brodeur, J.; Hjelmager, J.; Jones, A.; Jetlund, K.; Grillmayer, R.; Wasström, C. Standards Making Geographic Information Discoverable, Accessible and Usable for Modern Cartography. In *Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management*; Döllner, J., Jobst, M., Schmitz, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 325–344.
11. Object Management Group. *Unified Modelling Language Specification Version 2.5.1*; Object Management Group: Needham, MA, USA, 2017.
12. ISO/TC 211. *ISO 19109:2015 Geographic Information–Rules for Application Schema*; ISO: Geneva, Switzerland, 2015.
13. ISO/TC 211. *ISO 19103:2015 Geographic Information–Conceptual Schema Language*; ISO: Geneva, Switzerland, 2015.
14. Object Management Group. *Model Driven Architecture (MDA) Guide Rev. 2.0*; Object Management Group: Needham, MA, USA, 2014.
15. ISO/TC 211. *ISO 19136:2007 Geographic Information–Geography Markup Language (GML)*; ISO: Geneva, Switzerland, 2007.
16. Jeansoulin, R. Review of Forty Years of Technological Changes in Geomatics toward the Big Data Paradigm. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 155. [CrossRef]
17. ISO/TC 211. *ISO 19128:2005 Geographic Information–Web Map Server Interface*; ISO: Geneva, Switzerland, 2005.
18. ISO/TC 211. *ISO 19142:2010 Geographic Information–Web Feature Service*; ISO: Geneva, Switzerland, 2010.
19. Zhang, C.R.; Zhao, T.; Li, W.D.; Osleeb, J.P. Towards logic-based geospatial feature discovery and integration using web feature service and geospatial semantic web. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 903–923. [CrossRef]
20. Iwaniak, A.; Kaczmarek, I.; Strzelecki, M.; Lukowicz, J.; Jankowski, P. Enriching and improving the quality of linked data with GIS. *Open Geosci.* **2016**, *8*, 323–336. [CrossRef]
21. Manola, F.; Miller, E.; McBride, B. RDF 1.1 Primer. W3C Working Group Note 2014. Available online: https://www.w3.org/TR/rdf11-primer/ (accessed on 21 August 2019).
22. Hitzler, P.; Krötzsch, M.; Parsia, B.; Patel-Schneider, P.F.; Rudolph, S. OWL 2 Web Ontology Language Primer (Second Edition). *W3C Recomm.* **2012**, *27*, 123.
23. Hart, G.; Dolbear, C. *Linked Data: A Geographic Perspective*; CRC Press: Boca Raton, FL, USA, 2016.
24. Aditya, T.; Kraak, M.-J. A Search Interface for an SDI: Implementation and Evaluation of Metadata Visualization Strategies. *Trans. GIS* **2007**, *11*, 413–435. [CrossRef]
25. Klien, E. A Rule-Based Strategy for the Semantic Annotation of Geodata. *Trans. GIS* **2007**, *11*, 437–452. [CrossRef]
26. Lutz, M.; Kolas, D. Rule-Based Discovery in Spatial Data Infrastructure. *Trans. GIS* **2007**, *11*, 317–336. [CrossRef]
27. Wiemann, S.; Bernard, L. Spatial data fusion in Spatial Data Infrastructures using Linked Data. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 613–636. [CrossRef]
28. Hietanen, E.; Lehto, L.; Latvala, P. Providing Geographic Datasets as Linked Data in Sdi. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 583–586. [CrossRef]
29. Patroumpas, K.; Georgomanolis, N.; Stratiotis, T.; Alexakis, M.; Athanasiou, S. Exposing INSPIRE on the Semantic Web. *J. Web Semant.* **2015**, *35*, 53–62. [CrossRef]
30. Ulutaş Karakol, D.; Kara, G.; Yılmaz, C.; Cömert, Ç. Semantic Linking Spatial Rdf Data to the Web Data Sources. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII–4*, 639–645. [CrossRef]
31. Aydinoğlu, A.Ç.; Kara, A. Modelling and publishing geographic data with model-driven and linked data approaches: Case study of administrative units in Turkey. *J. Spat. Sci.* **2019**, *64*, 11–31. [CrossRef]
32. Wang, Y.D.; Qiao, M.L.; Liu, H.; Ye, X.Y. Qualitative spatial reasoning on topological relations by combining the semantic web and constraint satisfaction. *Geo Spat. Inf. Sci.* **2018**, *21*, 80–92. [CrossRef]
33. Mobasheri, A. A Rule-Based Spatial Reasoning Approach for OpenStreetMap Data Quality Enrichment; Case Study of Routing and Navigation. *Sensors* **2017**, *17*, 2498. [CrossRef]
34. Tand, J.; van den Brink, L.; Barnaghi, P. Spatial Data on the Web Best Practices. W3C Working Group Note 2017. Available online: https://www.w3.org/TR/2017/NOTE-sdw-bp-20170928/ (accessed on 21 August 2019).
35. ISO/TC 211. *ISO 19150-2:2015 Geographic Information–Ontology–Part 2: Rules for Developing Ontologies in the Web Ontology Language (OWL)*; ISO: Geneva, Switzerland, 2015.

36. ISO/TC 211. ISO/TC 211 Group for Ontology Management. 2018. Available online: https://github.com/ISO-TC211/GOM (accessed on 15 June 2019).
37. ARE3NA Project. Guidelines for the RDF Encoding of Spatial Data. 2017. Available online: http://inspire-eu-rdf.github.io/inspire-rdf-guidelines/ (accessed on 15 June 2019).
38. Echterhoff, J.; Portele, C.; Birkel, P.; Nichols, D.L.; Badgley, E.D. *OGC Testbed-12: ShapeChange Engineering Report*; Open Geospatial Consoritum: Wayland, MA, USA, 2017.
39. Echterhoff, J.; Birkel, P.; Nichols, D.L. *OGC Testbed-14: Application Schema-Based Ontology Development Engineering Report*; Open Geospatial Consoritum: Wayland, MA, USA, 2018.
40. McGlinn, K.; Wagner, A.; Pauwels, P.; Bonsma, P.; Kelly, P.; O'Sullivan, D. Interlinking geospatial and building geometry with existing and developing standards on the web. *Autom. Constr.* **2019**, *103*, 235–250. [CrossRef]
41. Debruyne, C.; Meehan, A.; Clinton, É.; McNerney, L.; Nautiyal, A.; Lavin, P.; O'Sullivan, D. Ireland's Authoritative Geospatial Linked Data. In *International Semantic Web Conference*; Springer: Berlin/Heidelberg, Germany, 2017.
42. Debruyne, C.; Clinton, É.; McNerney, L.; Nautiyal, A.; O'Sullivan, D. Serving Ireland's Geospatial Information as Linked Data. In Proceedings of the International Semantic Web Conference (Posters & Demos), Kobe, Japan, 17–21 October 2016.
43. Ordnance Survey. Ordnance Survey Ontologies. Available online: http://data.ordnancesurvey.co.uk/ontology (accessed on 6 March 2019).
44. Van den Brink, L.; Janssen, P.; Quak, W.; Stoter, J. Linking spatial data: Automated conversion of geo-information models and GML data to RDF. *Int. J. Spat. Data Infrastruct. Res.* **2014**, *9*, 59–85. [CrossRef]
45. Folmer, E.; Beek, W.; Rietveld, L. Linked Data Viewing as Part of the Spatial Data Platform of the Future. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-4/W8*, 49–52. [CrossRef]
46. Vilches-Blázquez, L.M.; Villazón-Terrazas, B.; Corcho, O.; Gómez-Pérez, A. Integrating geographical information in the Linked Digital Earth. *Int. J. Digit. Earth* **2014**, *7*, 554–575. [CrossRef]
47. Greek Linked Open Data. Available online: http://linkedopendata.gr/ (accessed on 6 March 2019).
48. Allemang, D.; Hendler, J.A. *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*, 2nd ed.; Elsevier: Burlington, MA, USA, 2011.
49. Bārzdiņš, J.; Bārzdiņš, G.; Čerāns, K.; Liepiņš, R.; Sproģis, A. UML style graphical notation and editor for OWL 2. In *International Conference on Business Informatics Research*; Springer: Berlin/Heidelberg, Germany, 2010.
50. Brockmans, S.; Haase, P.; Hitzler, P.; Studer, R. A metamodel and UML profile for rule-extended OWL DL ontologies. In *Semantic Web: Research and Applications, Proceedings*; Sure, Y., Domingue, J., Eds.; Springer: Berlin, Germany, 2006; pp. 303–316.
51. Brockmans, S.; Volz, R.; Eberhart, A.; Loffler, P. Visual modeling of OWL DL ontologies using UML. In *Semantic Web Iswc 2004 Proceedings*; McIlraith, S.A., Plexousakis, D., VanHarmelen, F., Eds.; Springer: Berlin, Germany, 2004; pp. 198–213.
52. Cranefield, S.; Purvis, M. A UML profile and mapping for the generation of ontology-specific content languages. *Knowl. Eng. Rev.* **2002**, *17*, 21–39. [CrossRef]
53. Djuric, D.; Gasevic, D.; Devedzic, V.; Damjanovic, V. UML profile for OWL. In *Web Engineering, Proceedings*; Koch, N., Fraternali, P., Wirsing, M., Eds.; Springer: Berlin, Germany, 2004; pp. 607–608.
54. Djuric, D.; Gasevic, D.; Devedzic, V.; Damjanovic, V. A UML profile for OWL ontologies. In *Model Driven Architecture*; Assmann, U., Aksit, M., Rensink, A., Eds.; Springer: Berlin, Germany, 2005; pp. 204–219.
55. Evermann, J.; Wand, Y. Ontological modeling rules for UML: An empirical assessment. *J. Comput. Inf. Syst.* **2006**, *46*, 14–29. [CrossRef]
56. Guizzardi, G.; Wagner, G.; Guarino, N.; van Sinderen, M. An ontologically well-founded profile for UML conceptual models. In *Advanced Information Systems Engineering, Proceedings*; Persson, A., Stirna, J., Eds.; Springer: Berlin, Germany, 2004; pp. 112–126.
57. Guizzardi, G.; Wagner, G.; Herre, H. On the foundations of UML as an ontology representation language. In *Engineering Knowledge in the Age of the Semantic Web, Proceedings*; Motta, E., Shadbolt, N., Stutt, A., Gibbins, N., Eds.; Springer: Berlin, Germany, 2004; pp. 47–62.
58. Keet, C.M.; Fillottrani, P.R. An ontology-driven unifying metamodel of UML Class Diagrams, EER, and ORM2. *Data Knowl. Eng.* **2015**, *98*, 30–53. [CrossRef]

59. Kogut, P.; Cranefield, S.; Hart, L.; Dutra, M.; Baclawski, K.; Kokar, M.; Smith, J. UML for ontology development. *Knowl. Eng. Rev.* **2002**, *17*, 61–64. [CrossRef]
60. Object Management Group. *Ontology Definition Metamodel, Version 1.0.*; Object Management Group: Needham, MA, USA, 2009.
61. Parreiras, F.S.; Staab, S. Using ontologies with UML class-based modeling: The TwoUse approach. *Data Knowl. Eng.* **2010**, *69*, 1194–1207. [CrossRef]
62. Interactive Instruments GmbH, ShapeChange. Available online: http://shapechange.net (accessed on 15 June 2019).
63. Bourahla, M.; Belghiat, A. UML Class Diagrams to OWL Ontologies: A Graph Transformation based Approach. *Int. J. Comput. Appl.* **2012**, *41*, 41–46. [CrossRef]
64. Bourahla, M.; Belghiat, A. Transformation of UML models towards OWL ontologies. In Proceedings of the 2012 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, SETIT, Sousse, Tunisia, 21–24 March 2012.
65. Xu, Z.M.; Ni, Y.Y.; He, W.J.; Lin, L.L.; Yan, Q. Automatic extraction of OWL ontologies from UML class diagrams: A semantics-preserving approach. *World Wide Web Internet Web Inf. Syst.* **2012**, *15*, 517–545. [CrossRef]
66. Bahaj, M.; Bakkas, J. Automatic Conversion Method of Class Diagrams to Ontologies Maintaining Their Semantic Features. *Int. J. Soft Comput. Eng. (IJSCE)* **2013**, *2*, 65–69.
67. Hajjamy, O.E.; Alaoui, K.; Alaoui, L.; Bahaj, M. Mapping UML to OWL2 ontology. *J. Theor. Appl. Inf. Technol.* **2016**, *90*, 126–143.
68. Gasevic, D.; Djuric, D.; Devedzic, V.; Damjanovi, V. Converting UML to OWL ontologies. In Proceedings of the 13th international World Wide Web conference on Alternate Track Papers & Posters WWW Alt. '04, New York, NY, USA, 19–21 May 2004.
69. Gherabi, N.; Bahaj, M. A New Method for Mapping UML Class into OWL Ontology. *Int. J. Comput. Appl.* **2012**, *1*, 5–9.
70. Probst, F.; Bibotti, F.; Pazos, A. Connecting ISO and OGC Models to the Semantic Web | OGC Network. In Proceedings of the 3rd International Conference on Geographic Information Science: Extended Abstracts and Poster Summaries, Adelphi, MD, USA, 20–23 October 2004; pp. 181–184.
71. Buccella, A.; Cechich, A.; Gendarmi, D.; Lanubile, F.; Semeraro, G.; Colagrossi, A. Building a global normalized ontology for integrating geographic data sources. *Comput. Geosci.* **2011**, *37*, 893–916. [CrossRef]
72. Zedlitz, J.; Luttenberger, N. Transforming Between UML Conceptual Models And OWL 2 Ontologies. In Proceedings of the Terra Cognita Workshop on Foundations Technologies and Applications of the Geospatial Web at the 11th International Semantic Web Conference ISWC 2012, Boston, MA, USA, 12 November 2012.
73. ISO/TC 211. *ISO 19107:2003 Geographic Information–Spatial Schema*; ISO: Geneva, Switzerland, 2003.
74. Perry, M.; Herring, J. *OGC GeoSPARQL-A Geographic Query Language for RDF Data*; Open Geospatial Concortium: Wayland, MA, USA, 2012.
75. Cox, S. An explicit OWL representation of ISO/OGC Observations and Measurements. *SSN@ ISWC* **2013**, *1063*, 1–18.
76. ISO/TC 211. *ISO 19125-1:2004 Geographic Information–Simple Feature Access–Part 1: Common Architecture*; ISO: Geneva, Switzerland, 2004.
77. Inspire. *Inspire Generic Conceptual Model*; INSPIRE Drafting Team "Data Specification": Brussels, Belgium, 2013.
78. ISO/TC 204. *ISO 14825:2011 Intelligent Transport Systems–Geographic Data Files (GDF)–GDF5.0*; ISO: Geneva, Switzerland, 2011.
79. Link, V.; Lohmann, S.; Marbach, E.; Negru, S.; Wiens, V. WebVOWL. 2019. Available online: http://www.visualdataweb.de/webvowl. (accessed on 21 August 2019).
80. Inspire. *Inspire Data Specification on Geographical Names–Guidelines*; INSPIRE Thematic Working Group Geogrphical Names: Brussels, Belgium, 2010.
81. Echterhoff, J.; De Paepe, D. Encoding of Geographical Names. 2017. Available online: https://github.com/inspire-eu-rdf/inspire-rdf-guidelines/issues/28 (accessed on 17 June 2019).
82. ISO/TC 37/SC 2. *ISO 639-2:1998 Codes for the Representation of Names of Languages–Part 2: Alpha-3 Code*; ISO: Geneva, Switzerland, 1998.

83. The Library of Congress, ISO639-2 Languages. 2019. Available online: http://id.loc.gov/vocabulary/iso639-2 (accessed on 21 August 2019).
84. ISO/TC 204. *ISO 14823:2017 Intelligent Transport Systems–Graphic Data Dictionary*; ISO: Geneva, Switzerland, 2017.
85. Jetlund, K. ISO 14823 SKOS Concept Schemes. 2019. Available online: https://github.com/jetgeo/GIS2OWL/tree/master/iso14823 (accessed on 21 August 2019).
86. Noy, N.F.; McGuinness, D.L. Ontology Development 101: A Guide to Creating Your First Ontology. 2001. Available online: https://protegewiki.stanford.edu/wiki/Ontology101 (accessed on 21 August 2019).

Article 4

IFC Schemas in ISO/TC 211 compliant UML for improved interoperability between BIM and GIS

International Journal of
Geo-Information

MDPI

Article

IFC Schemas in ISO/TC 211 Compliant UML for Improved Interoperability between BIM and GIS

Knut Jetlund [1,2,*], Erling Onstein [2] and Lizhen Huang [2]

1 Norwegian Public Roads Administration, 2318 Hamar, Norway
2 Department of Manufacturing and Civil Engineering, Norwegian University of Science and Technology, 2821 Gjøvik, Norway; erling.onstein@ntnu.no (E.O.); lizhen.huang@ntnu.no (L.H.)
* Correspondence: knut.jetlund@vegvesen.no; Tel.: +47-9596-2863

Received: 3 February 2020; Accepted: 22 April 2020; Published: 23 April 2020

check for updates

Abstract: This study aims to improve the interoperability between the application domains of Building Information Modelling (BIM) and Geographic Information Systems (GIS) by linking and harmonizing core information concepts. Many studies have investigated the integration of application schemas and data instances according to the BIM model IFC and the GIS model CityGML. This study investigates integration between core abstract concepts from IFC and ISO/TC 211 standards for GIS—independent of specific application schemas. A pattern was developed for conversion from IFC EXPRESS schemas to Unified Modelling Language (UML) models according to ISO/TC 211 standards. Core concepts from the two application domains were linked in the UML model, and conversions to implementation schemas for the Geography Markup Language (GML) and EXPRESS were tested. The results showed that the IFC model could be described as an ISO/TC 211 compliant UML model and that abstract concepts from ISO/TC 211 standards could be linked to core IFC concepts. Implementation schemas for BIM and GIS formats could be derived from the UML model, enabling implementation in applications from both domains without conversion of concepts. Future work should include refined linking and harmonization of core abstract concepts from the two application domains.

Keywords: information models; building information modelling; industry foundation classes; geographic information systems; unified modelling language

1. Introduction

1.1. Building Information Modelling (BIM) and Geographic Information Systems (GIS)

Information models for digital representation of real-world features in a geospatial context are fundamental for understanding, using, maintaining and developing the natural as well as the built environment. The application domains of Building Information Modelling (BIM) and Geographic Information Systems (GIS) have existed and emerged in parallel for decades, with different scopes and distinct models of geospatial information. While applications and information models for BIM mainly have handled the built environment, GIS has handled the natural and built outdoor environment.

Stakeholders from the two application domains have developed domain-specific information models based on distinct conceptual modelling languages. The core information model for BIM is defined in the Industry Foundation Classes (IFC) [1], described with the EXPRESS modelling language [2]. In the GIS application domain, ISO Technical Committee 211 (ISO/TC 211) have defined concepts for modelling of geospatial information founded on a formalized use of the Unified Modelling Language (UML) [3–5] and Model-Driven Architecture (MDA) [6]. The ISO/TC 211 concepts are widely used in implementable information models in the GIS application domain, developed by stakeholders such as INSPIRE and the Open Geospatial Consortium (OGC) [7].

The scopes of the two application domains have intersected each other increasingly over the last years. Planning, constructing, using, and maintaining the built environment in BIM depends on knowledge of the surroundings—the natural, as well as the built, environment. Therefore, BIM applications need to integrate information from geospatial datasets into BIM projects and relate the new constructions to the existing environment. Likewise, the development of the built environment influences the natural environment. GIS applications need information from BIM projects in order to update and combine the existing environment with the new constructions. Therefore, the integration of BIM and GIS has been considered a promising topic for cross-domain knowledge exchange [8–10]. Integration has become especially relevant with the extensive use of BIM for infrastructure, as infrastructure projects extend over large geographic areas and relate more to other features from the natural and built environment than what has typically been handled in BIM [11].

A shared understanding of the digital representation of real-world features in a geospatial context is considered an essential fundament for a successful integration between BIM and GIS. The heterogeneity between application-specific information models described with different modelling languages is a challenge for such common understanding [12–14].

1.2. Contribution and Research Questions

This study concerns the integration of technologies and approaches used for information modelling in the application domains of BIM and GIS. Many studies have investigated the integration of information models at the application schema and instance levels—typically IFC and OGC CityGML [15] models and data. However, CityGML is only one of several GIS application schemas that may be integrated with IFC. Therefore, we aim to enable the integration of core concepts independent of specific application schemas. We are investigating whether heterogenous information models from the two application domains can be integrated into a common, harmonized UML and MDA approach, using shared core concepts.

The focus of the study is on concepts from the IFC information model in the BIM domain and core standards and information models from ISO/TC 211 in the GIS domain. As the main concern is information modelling, complex issues related to solid geometries in BIM and GIS—as discussed in, e.g., [16,17]—are not included in the study.

The study investigates three research questions:

1. What are the potential solutions and challenges for the conversion of information models for IFC to UML models according to ISO/TC 211 standards?
2. What relationships can be defined between core concepts in information models for IFC and concepts from core ISO/TC 211 models?
3. What are the potential solutions and challenges for deriving implementation schemas for EXPRESS and the Geography Markup Language (GML) from IFC UML models?

2. Literature Review

2.1. Information Models for BIM and GIS

BIM was introduced as an application domain in the late 1980s and early 1990s, while the first version of IFC was released in 1996. IFC is developed by buildingSMART International and standardized as ISO 16739-1 [18] and is currently at version 4.1—with further extensions under development. Schemas described with the EXPRESS modelling language [2] are the complete official specifications of the IFC information model. Besides, representations of the model are also available in XML and the Web Ontology Language (OWL). Other specifications supplement IFC for information handling—in particular, Model View Definitions (MVD); Information Delivery Manuals (IDM); Property Set Definition (PSD); and International Framework Dictionary (IFD).

The application domain of GIS emerged in the late 1960s. Since 1994, the major stakeholders in standardization in the GIS domain has been ISO/TC 211 and OGC [7]. Standards developed by

ISO/TC 211 define general abstract concepts for models of geospatial information, where three core standards are fundament for the use of UML and MDA: ISO 19103 [19] formalizes the use of UML and MDA and defines core datatypes. ISO 19109 [20] specifies further rules and semantics for modelling of application schemas. Finally, ISO 19136 [21] defines the standardized GML exchange format; semantics needed in UML for implementation as GML schemas; and rules for conversion from UML to GML. Besides, the abstract concepts for geometry and spatial referencing defined in ISO 19107 [22] and ISO 19111 [23] are vital in the scope of this article.

The general concepts defined by ISO/TC 211 lay the foundation for specific application schemas for direct implementation in GIS applications. Authorities and organizations have defined standardized application schemas at national and international levels for numerous themes, e.g., base maps, geology, infrastructure, land cover, and land administration. Among these are the European INSPIRE Data Specifications [24], the TN-ITS specification for road traffic regulations [25] and specifications from OGC. OGC has developed a wide range of applications schemas, where CityGML [15] for three-dimensional city models and LandInfra/InfraGML [26] for infrastructure are often referred to as the equivalent to IFC in the GIS domain [9,16].

The last years have seen a move towards a closer collaboration between buildingSMART and OGC. The LandInfra conceptual model was developed in collaboration between buildingSMART and OGC [27] and lay the foundation for infrastructure models in both application domains. However, more specific information models for infrastructure are still heterogeneously developed for the two application domains: The buildingSMART Infrastructure Room develops extensions to IFC [28], while OGC has developed the InfraGML models [26]. IFC version 4.1 introduced alignments as a core concept for infrastructure, based on the LandInfra conceptual model. Further extensions for bridge information are introduced in the IFC version 4.2 candidate standard [29], while extensions for road, rail, and other infrastructure domains are under development. Parallel to the development of IFC for Infrastructure, OGC has developed the InfraGML Encoding Standards [26], where the Alignments model in part 3 [30] is based on the LandInfra conceptual model for alignment.

2.2. Integration of BIM and GIS

Integration of BIM and GIS has become a comprehensive research activity in both application domains. Recent review studies found a worldwide increase in research on the integration between BIM and GIS: Zhu et al. [9] noted an evolution from three citations concerning BIM and GIS in 2009 to 37 in 2014 and 313 in 2017; while Song et al. [10] observed the same evolution of citations and a similar trend for publications. They found an average of six published articles over the years from 2008 to 2014, 19 articles in 2015 and 27 in 2016. Liu et al. [8], as well as Zhu et al. [9], considered information loss to be a challenge for the integration and found the conversion of semantics to be an even more challenging issue than the conversion of geometry. Liu et al. [8] pointed at openness and collaboration in the development of information models as keys to successful integration. They categorized most of the current solutions for the exchange of semantics as project-specific.

Liu et al. [8] found that the OGC specifications CityGML, IndoorGML and LandInfra/InfraGML intersect with the BIM domain and may reduce the barrier between BIM and GIS. However, they noted that the OGC specifications provide solutions from only particular views. Donkers et al. [16]; Deng et al. [17]; Hor et al. [31]; Kang and Hong [32]; Knoth et al. [33]; Ohori et al. [34]; and Vilgertshofer et al. [35] are some examples of articles that describe integration between BIM and GIS through schema mapping between IFC and CityGML. Niestroj et al. [36] discussed information models for road asset information exchange and noted that InfraGML might be easier to implement than IFC as GIS applications widely support the GML format. Kavisha et al. [11] discussed the LandInfra/InfraGML standards and the IFC Infrastructure extensions and noted that LandInfra/InfraGML could help bridge the gap between BIM and GIS.

The OGC and buildingSMART Integrated Digital Built Environment (IDBE) joint working group discussed challenges and possibilities for integration between IFC, CityGML and InfraGML [14].

They pointed at four predominant challenges: Different purposes, different conceptualisations of real-world objects, different modelling languages and differences in spatial representation. Among the proposed actions for improved interoperability and integration were a shared vocabulary and harmonization of concepts.

At the more abstract level, Onstein and Tognoni [12] and Roxin et al. [13] have compared core concepts from IFC to equivalent concepts from ISO/TC 211 standards. Onstein and Tognoni [12] compared the IFC Kernel schema to the ISO 19109 General Feature Model (GFM) in a case study on building permits. Roxin et al. [13] suggested to align IFC schemas and ISO/TC 211 models at the different levels of abstraction described in ISO 19103—an approach that is discussed in the joint work between ISO/TC 59/SC 13 and ISO/TC 211 on ISO Technical Report (TR) 23262 [37] as well. ISO TR 23262 investigates barriers and proposes measures to improve interoperability between GIS and BIM standards.

Conversions between information models have mostly been implemented with scripting and ETL (Extract, Transform, Load) tools [36]. Besides, Semantic web technologies such as OWL ontologies have been considered a promising approach for integration. A reference ontology—also known as a LinkSet—can describe relations between concepts in different models and then be used to link information based on heterogeneous information models. This approach is described in, e.g., Liu et al. [8]; Hor et al. [31]; and Roxin et al. [13], and implemented in the European INTERLINK Project for infrastructure [38,39].

2.3. EXPRESS and UML

The EXPRESS language is standardized in part 11 of the ISO 10303 STEP (Standard for the Exchange of Product model data) series [2]. EXPRESS has a formal and lexical syntax, while a graphical representation of a subset of the language can be described with EXPRESS-G. UML is developed and maintained by the Object Management Group (OMG) and standardized as ISO/IEC 19505 part 1 and 2 [4,5]. UML is mainly a graphical language but includes semantics for documenting and implementing the artefacts defined in the models. Extensions and restrictions for a specific use of UML can be formalized in profiles such as the ISO/TC 211 UML profile defined in ISO 19103 [19].

A few publications have discussed conversions between EXPRESS and UML: Arnold and Podehl [40] described a mapping between EXPRESS-G and UML class diagrams. Krause and Kaufmann [41] described mapping of concepts at the metamodel level, while the "exff" Project (exPRESS for free) [42] defined mapping rules from EXPRESS to the UML Interchange format XMI. OMG defines a complete representation of EXPRESS in UML with the UML metamodel for EXPRESS [43].

An essential work for conversion from EXPRESS to other modelling languages was done by Pauwels and Terkaj [44]. They described a conversion from EXPRESS to OWL for the IFC model. Based on their work, OWL has become an official format for IFC schemas [1]. The work has been supplemented with further refinements for modularization [45] and geometry handling [46]. Several conversion principles described by Pauwels and Terkaj may be applied for conversion from EXPRESS to UML as well.

The IFC Infrastructure projects for rail and road has used UML as their primary tool for information modelling, and buildingSMART has announced an upcoming move from EXPRESS to UML as the original modelling language for IFC [47]. The intention is to maintain the conceptual model in UML and generate implementation schemas for EXPRESS, XML and OWL. A draft version of the IFC-UML model and the conversion scripts was recently made available [48]. The IFC-UML model is developed independently of the UML profiles and rules defined by ISO/TC 211. Therefore—although information models from both application domains are available in UML—they are still heterogeneous due to different usage of the modelling language.

3. Materials and Method

Figure 1 illustrates the materials and method used for investigating the research questions in a process with five steps. A base knowledge was established through detailed studies of the EXPRESS and UML language specifications and their use for IFC and ISO/TC 211 information models. The source materials for the studies are listed in Table 1.

Based on the knowledge from the first step, we were able to identify potential solutions and challenges for conversion, and develop a pattern for conversion from IFC concepts represented in EXPRESS to representation in UML according to ISO/TC 211 standards. The conversion pattern is described in Section 4.1, with example results in UML package and class diagrams.

The ISO/TC 211 compliant UML representation of IFC was used for comparing and relating core IFC concepts to core concepts from ISO/TC 211 standards, as described in Section 4.2. Based on the results and knowledge from the previous steps, we developed conversion patterns from UML to implementations schemas for EXPRESS and GML, which is described in Section 4.3. Finally, the results from the bidirectional conversion are evaluated and discussed in Section 5.

The data sources and tools used for the bi-directional conversions are listed in Table 2, while the complete UML models, implementation schemas and conversion scripts are available online at [49]. The notation for UML package and class diagrams as presented in Section 4 is specified in the UML specification [3] and illustrated for use in ISO/TC 211 UML models in the ISO/TC 211 Wiki on UML Best Practices [50]. Further descriptions of UML can, e.g., be found in [51].

Table 1. Sources for comparisons of IFC EXPRESS and ISO/TC 211 UML.

Language	Source
EXPRESS	ISO 10303-11: The EXPRESS language reference manual [2]
	Information modelling: The EXPRESS way (Schenk and Wilson) [52]
EXPRESS for IFC	IFC 4.3 (IFC Road) first draft EXPRESS schemas [53]
	IFC 4.2 candidate standard EXPRESS schemas [29]
UML	The UML Specification version 2.5.1 [3]
UML for IFC	The draft IFC-UML model from buildingSMART [48]
UML for geospatial information	ISO 19103 [19]: Conceptual schema language
	ISO 19109 [20]: Rules for application schema
	ISO 19136 [21]: Geography Markup Language (GML)
	UML models for ISO/TC 211 standards [54]
	UML models from OGC: Land and Infrastructure Conceptual Model (LandInfra) version 1.0, and InfraGML Encoding Standards version 1.0

Table 2. Data sources and tools.

Source or Tool	Description
IFC Road EXPRESS schemas [53]	Converted to UML through the conversion pattern.
UML models for ISO/TC 211 standards [54]	Imported from the ISO/TC 211 Harmonized UML Model.
Enterprise Architect version 15 [55]	The application used for UML modelling and scripting of the bi-directional conversion between UML and EXPRESS.
ShapeChange version 2.8 [56]	The application used for conversion from UML to GML.
Oxygen XML Editor version 22 [57]	The application used for validating and inspecting XML schemas, and for developing example instances.

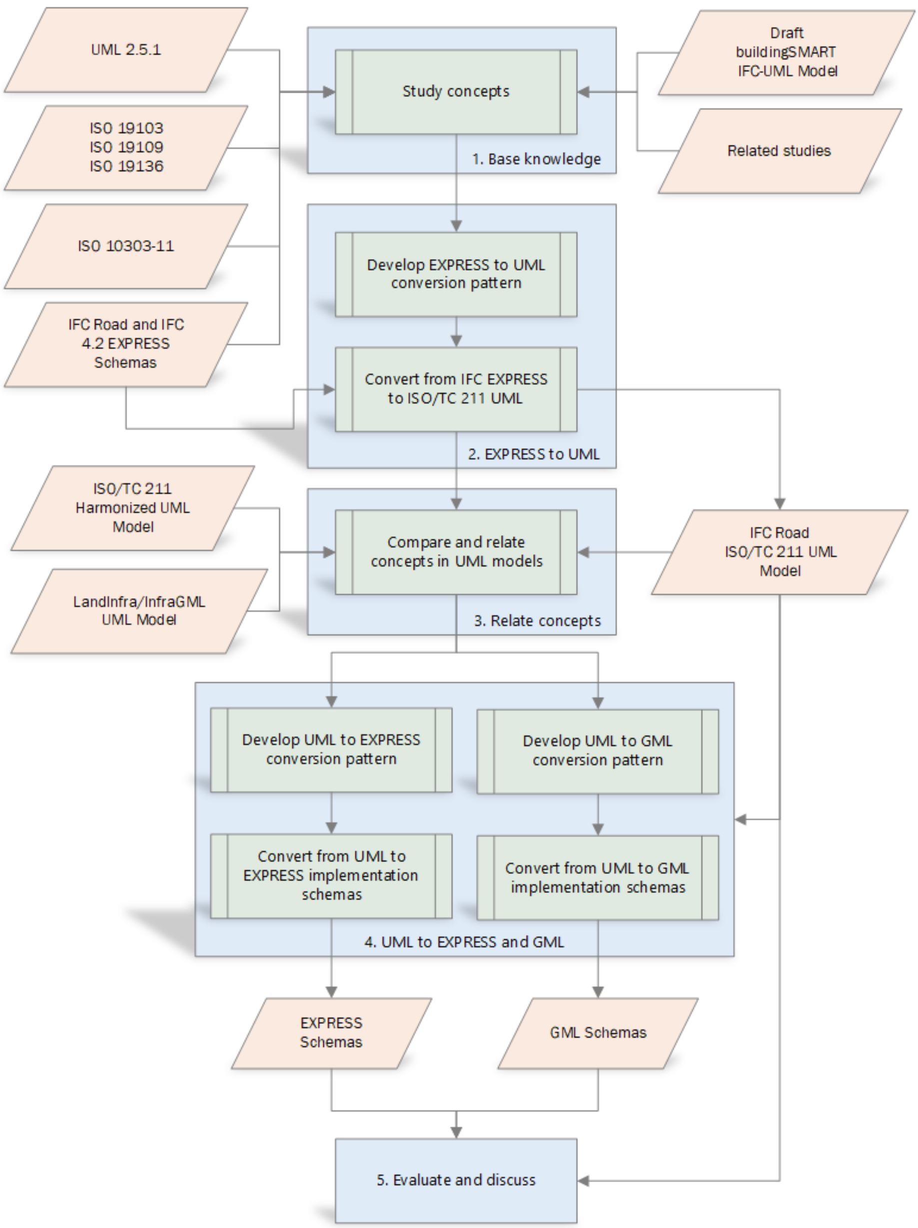


Figure 1. The research process and materials for the study.

4. Results

4.1. Conversion of IFC Models to ISO/TC 211 Compliant UML Models

4.1.1. Conversion Pattern Overview

We developed a pattern for conversion from IFC information models into a UML model according to profiles and rules defined by ISO/TC 211: The UML profile defined in ISO 19103, with extensions defined in ISO 19109 and ISO 19136; the ISO 19109 GFM; and modelling rules defined in ISO 19103, ISO 19109, and ISO 19136. The conversion pattern describes conversion from EXPRESS schemas, as the official IFC-UML is still in development. Table 3 presents an overview of the conversion pattern, while subsequent clauses describe further details of the conversion pattern.

The conversion pattern considers differences between the EXPRESS and UML languages as well as the specific use of EXPRESS for IFC and UML for standards from ISO/TC 211 and OGC. Concepts from IFC EXPRESS schemas are implemented as concepts described in the extended ISO/TC 211 UML profile and the ISO 19109 GFM.

The pattern aims to reduce the heterogeneity between BIM and GIS by moving IFC schemas into the same UML environment as GIS information models and enable the future maintenance of IFC models harmonized with standards from ISO/TC 211 and OGC. Our approach is broader than the one described in [44], where the scope of the converted ontology is conversions of existing IFC files to RDF. Our approach is also different from the work on the official IFC-UML [48], as we are relating the IFC model to core ISO/TC 211 concepts.

Table 3. Overview of the conversion pattern.

IFC EXPRESS Concept		**ISO/TC 211 UML Concept**
Schema		Package
	Reference	Package dependency
Entity		Class with stereotype "FeatureType"
	Abstract	Abstract class
	Supertype/Subtype	Generalization/Specialization
Type	Simple EXPRESS datatype	Simple ISO 19103 datatype
	Simple defined datatype	Datatype, realization of ISO 19103 datatype
	Type of IFC datatype	Datatype, subtype of IFC datatype
	Aggregation datatype	Datatype with attribute or association and tagged values
	Enumeration datatype	Enumeration
	Select datatype	Union
Attribute		Attribute or association of FeatureType
	Datatype	Datatype or associated FeatureType
	Derived	Derived attribute or association
	Optional	Multiplicity with lower bounds = 0
	Aggregation	Multiplicity and tagged values
	List or array aggregation	Ordered attribute or association
Constraint	Unique	FeatureType constraint
	Inverse	FeatureType constraint, role name and multiplicity
	Derive	FeatureType constraint
	Where	FeatureType constraint
Function		Operation grouped in Interface
Rule		Operation grouped in Interface

As our conversion pattern follows profiles and rules defined in ISO/TC 211 standards, it has some differences from the draft buildingSMART IFC-UML model [48]. The differences are listed in Table 4. Most important; the IFC-UML model has more of the EXPRESS semantics specifically defined. Contrary, our pattern converts EXPRESS primitive types to ISO 19103 types; select types to unions; and group functions and rules as operations in interfaces.

Table 4. Comparison of UML models.

IFC EXPRESS Concept		ISO/TC 211 UML Concept	buildingSMART IFC-UML Concept
Schema	Reference	Package dependency	Package to element dependency
Entity		Class with stereotype "FeatureType"	Class
Type	Simple EXPRESS datatype	Simple ISO 19103 datatype	Datatype stereotyped "primitive" in the model
	Simple defined datatype	Datatype, realization of ISO 19103 datatype	Datatype, subtype of primitive EXPRESS datatype
	Aggregation datatype	Datatype with attribute or association	Datatype with an additional EXPRESS definition
	Select datatype	Union	Class stereotyped "EXPRESS SELECT", with substitution relations
Attribute	Aggregation	Multiplicity and tagged values	Multiplicity and additional EXPRESS definition
Constraint	Derive	FeatureType constraint	Attribute constraint
Function		Operation grouped in Interface	Class stereotyped "EXPRESS FUNCTION" with an operation
Rule		Operation grouped in Interface	Class stereotyped "EXPRESS RULE" with an operation

4.1.2. Schema

Grouping of models in smaller portions is particularly useful for large models with many elements. The core structural concept of EXPRESS is the schema, which defines the scope for a collection of items forming part or all of a model [2]. The IFC specification defines four conceptual layers with a total of 38 schemas, as presented in [29]: The Resource layer (21 schemas), the Core layer (four schemas), the Interoperability layer (five schemas) and the Domain layer (eight schemas). Despite the layer and schema architecture, the official IFC specifications are described in one large EXPRESS schema—e.g., IFC version 4.2 [29]. However, the draft IFC 4.3 Roads specification [53] that was used in the conversion was described in separate schemas according to the schema architecture in [29].

The package concept in UML is used for structuring and organizing classes and data types in UML models, equivalent to schemas in EXPRESS. Therefore, each EXPRESS schema can be converted to a UML package. The conceptual layers from IFC are maintained by grouping the packages in four main packages representing the layers, as illustrated in Figure 2. Besides, ISO 19109 defines the stereotype "ApplicationSchema" for UML packages that contain a conceptual schema for a specific domain model. According to rules in ISO 19109, an application schema shall not contain another application schema. The stereotype is therefore added only to the root IFC package in the UML model.

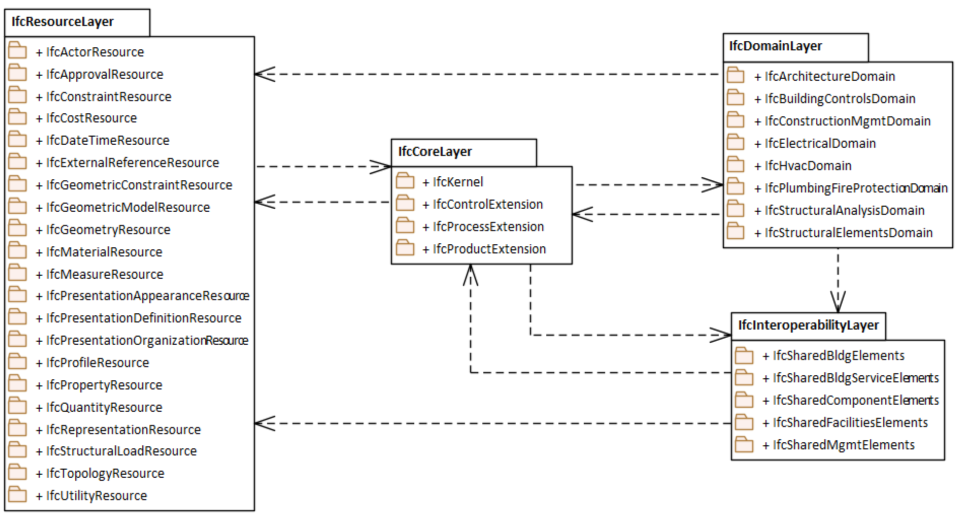


Figure 2. UML package diagram showing the core structure of the IFC-UML Model.

The Reference construct in EXPRESS specifies entities and types from other schemas that are reused in the current schema. A typical example is the reuse of elements from the IFC Resource layer schemas in Interoperability and Domain layer schemas. Our conversion pattern converts the REFERENCE FROM statements from EXPRESS to package-to-package dependencies in UML, and aggregates the dependencies further to the IFC conceptual layer packages, as illustrated in Figure 2. These package dependencies are mainly used for visualizing dependencies. The formal dependencies in the UML model are derived from elements in each package and their relations and property value domains from other packages.

4.1.3. Entity

The Entity concept in EXPRESS and the Class concept in UML has been considered equivalent in several studies, for example [40,41]. The ISO 19109 GFM specializes the UML metaclass "Class" to a UML metaclass "FeatureType", implemented with the stereotype "FeatureType". Therefore, we convert EXPRESS entities to UML classes with stereotype "FeatureType". Furthermore, EXPRESS, as well as UML, have concepts for abstract classes and class hierarchies. Abstract EXPRESS entities are converted to abstract UML classes, while the EXPRESS statement SUPERTYPE OF is converted to UML generalizations. Figure 3 illustrates a part of the IFC class hierarchy, emphasizing the inheritance from the core abstract superclass IfcRoot to the implementable class IfcFacility.

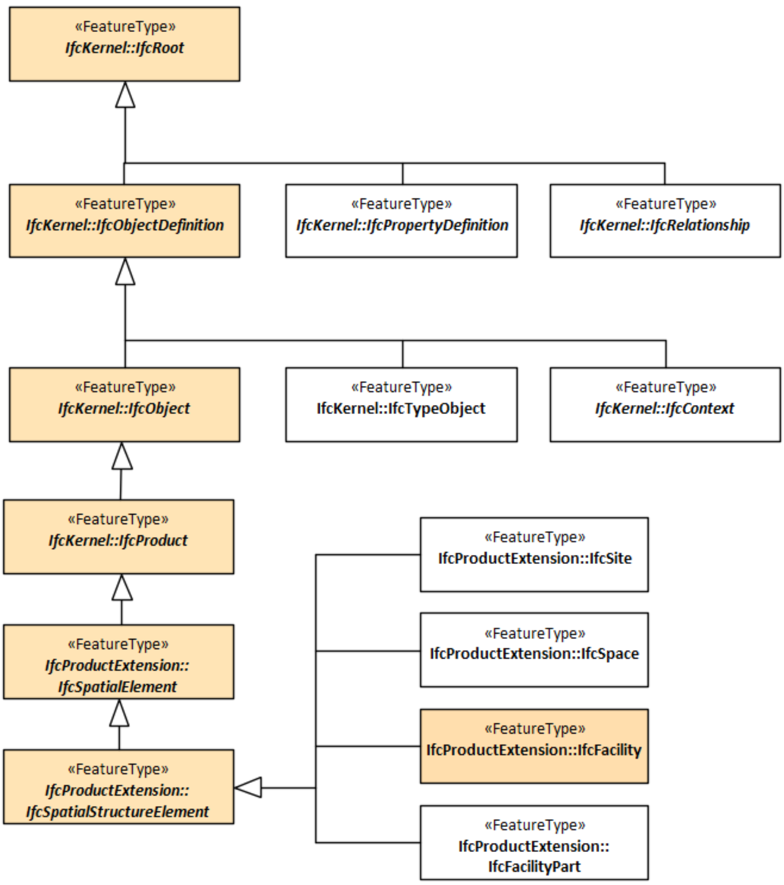


Figure 3. UML class diagram showing a part of the IFC class hierarchy from IfcRoot to IfcFacility.

4.1.4. Default Datatypes and Defined Datatypes

Type declarations in EXPRESS are considered equivalent to datatypes in UML [40,41]. The EXPRESS language defines a set of default simple datatypes: BINARY, BOOLEAN, INTEGER, LOGICAL, NUMBER, REAL, and STRING. ISO 19103 defines equivalent datatypes for use in UML models of geospatial information. Our conversion pattern applies a mapping between the default EXPRESS datatypes and equivalent datatypes from ISO 19103, for use in the conversion of defined datatypes.

The IFC EXPRESS schemas define a range of simple core datatypes for IFC, based on the default EXPRESS datatypes. For example, IfcInteger (INTEGER), IfcReal (REAL) and IfcBoolean (BOOLEAN). Besides, more specific simple datatypes, such as IfcAreaMeasure (REAL) and IfcCountMeasure (NUMBER), are defined as well. The simple (core and specific) IFC datatypes are converted to UML datatypes with realization associations from equivalent ISO 19103 datatypes, as illustrated in Figure 4.

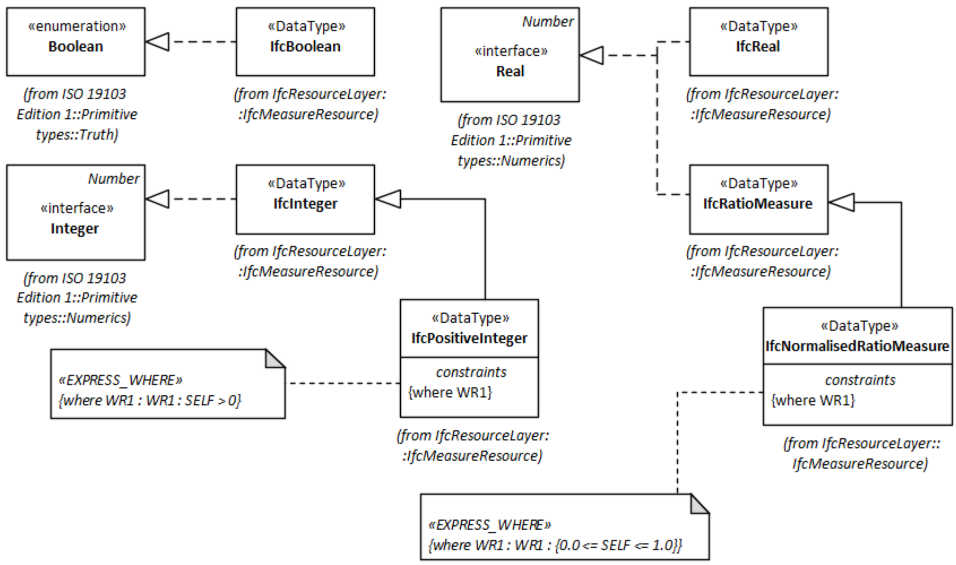


Figure 4. UML class diagram showing examples of IFC simple datatypes as realizations of ISO 19103 datatypes.

Constrained specializations of simple IFC datatypes, e.g., IfcPositiveInteger (IfcInteger) and IfcNormalisedRatioMeasure (IfcRatioMeasure) are converted to subtypes of the datatype they are specializing, as illustrated in Figure 4. Conversions of EXPRESS constraints on types and entities are described in Section 4.1.9.

4.1.5. Aggregation Datatypes

EXPRESS aggregation datatypes represent ordered or unordered collections. Three different EXPRESS aggregation datatypes are used in the IFC EXPRESS schemas: ARRAY, LIST, and SET. An array is an indexed and ordered collection with a fixed number of members; a list is an ordered collection with a flexible number of members; while a set is an unordered collection. The aggregation datatypes are extensively used for attribute domains in the IFC EXPRESS schema but only for a few type declarations: One ARRAY, three LISTs and one SET, as listed in Table 5.

Table 5. IFC Aggregation datatypes.

Type of Collection	IFC Type	Declaration
ARRAY	IfcComplexNumber	ARRAY [1:2] OF REAL
LIST	IfcArcIndex	LIST [3:3] OF IfcPositiveInteger
LIST	IfcCompoundPlaneAngleMeasure	LIST [3:4] OF INTEGER
LIST	IfcLineIndex	LIST [2:?] OF IfcPositiveInteger
SET	IfcPropertySetDefinitionSet	SET [1:?] OF IfcPropertySetDefinition

Our conversion pattern for aggregation datatypes follows a similar approach as the conversion from EXPRESS to OWL described in [44]: the aggregation datatypes are converted to UML datatypes with properties (attributes or associations) that holds the collections. The multiplicity of the properties defines the number of possible instances, and multiple instances form a collection. Details of the conversion pattern are described in Table 6, while Figure 5 shows the five IFC aggregation datatypes as UML datatypes.

Table 6. Conversion pattern for aggregation datatypes.

IFC EXPRESS	ISO/TC 211 UML	Description
All aggregation datatypes (LIST, ARRAY, SET)	Datatype with property "component".	Conversion from IFC EXPRESS schemas to UML properties are described in Section 4.1.8.
Type of aggregation	Tagged value "aggregationType".	The type of aggregation (LIST, ARRAY or SET) is maintained for implementation in EXPRESS.
Simple EXPRESS datatypes	Simple IFC datatypes.	Simple EXPRESS datatypes are replaced with equivalent simple IFC datatypes in order to avoid specific EXPRESS datatypes in the UML model.
ARRAY and LIST	The property "component" is defined as ordered.	The UML tag "IsOrdered" is set to true, to specify that the collection shall be ordered.
LIST and SET lower and upper bounds	Multiplicity of the property "component".	The lower and upper bounds of lists and sets define the number of members in the aggregation.
ARRAY bounds	The multiplicity of the property "component" is derived from the fixed number of ARRAY members which equals the difference between the upper and lower index plus one.	The bounds for arrays represent the lower and upper index numbers in the array—not the number of members [27].
ARRAY lower bounds	Tagged value "minIndex".	The lower index is maintained for implementation in EXPRESS. The upper index in EXPRESS can be calculated from the UML multiplicity.

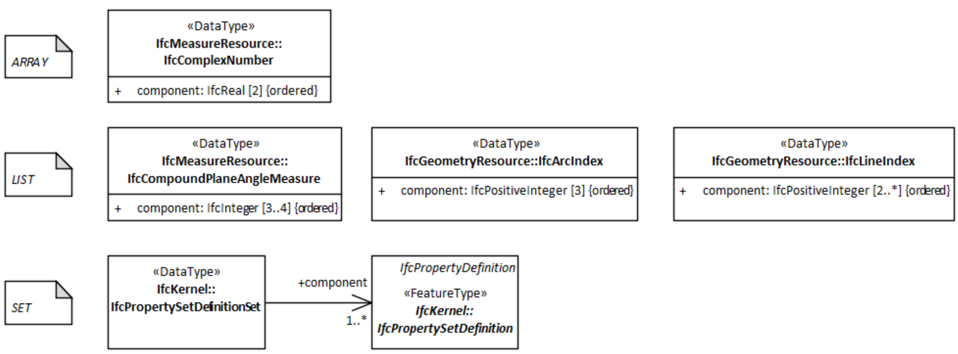


Figure 5. UML class diagram showing IFC aggregation datatypes in UML.

4.1.6. Enumerations

The Enumeration concept is equivalent in EXPRESS and UML and defines a set of names that are valid values for an attribute. Extensible enumerations in EXPRESS are defined with the keyword EXTENSIBLE and are equivalent to the CodeList concept defined in ISO 19103. ISO 19103 requires that the CodeList concept shall be used unless all possible values are known. However, extensible enumerations are not used in IFC. Therefore, enumerations from the IFC EXPRESS schemas are converted to UML Enumerations, as illustrated for some examples in Figure 6.

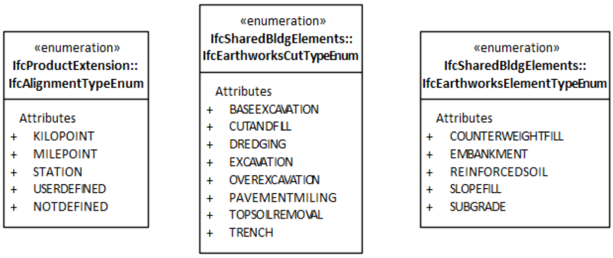


Figure 6. UML class diagram showing examples of IFC enumerations.

4.1.7. Select Datatypes

Select datatypes in EXPRESS defines a set of datatypes that may be used as the value domain for an attribute. Only one of the datatypes in the selection can be applied for an individual attribute instance. The UML Profile in ISO 19103 defines the equivalent concept Union. Therefore, IFC EXPRESS select types are converted to UML unions, as defined in ISO 19103. Figure 7 shows some examples of IFC select datatypes converted to UML unions.

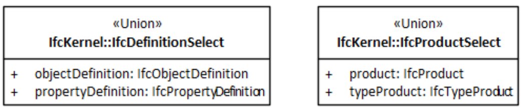


Figure 7. UML class diagram showing examples of IFC select datatypes converted to UML unions.

4.1.8. Attributes

EXPRESS attributes are used for describing characteristics of entities, similar to UML properties (attributes and associations), which are used for describing characteristics of classes. Therefore, attributes from IFC EXPRESS schemas are converted to either UML attributes or associations. According to the accepted best practices for UML [58], properties should be modelled as attributes if the value domain is a datatype and associations if the value domain is a class. IFC EXPRESS attributes with a datatype as value domain are therefore converted to UML attributes, while IFC EXPRESS attributes with an entity as value domain are converted to associations pointing to UML FeatureTypes.

Some concerns must be taken in order to be compliant to ISO 19103 and ISO 19109 and at the same time maintain full declarations from EXPRESS. The concerns are described in the attribute conversion pattern in Table 7.

Figure 8 shows examples of IFC entities and their attributes that have been converted to UML classes and properties. The association from IfcAligment2DHorizontal to IfcAligment2DHorizontalSegment is converted from the EXPRESS attribute statement "`Segments:LIST [1:?] OF IfcAlignment2DHorizontalSegment`". The attribute dim and the accompanying constraint in the class IfcCurve is converted from the EXPRESS attribute statement "`DERIVE Dim:IfcDimensionCount:= IfcCurveDim(SELF)`". The class IfcAlignment shows attributes inherited from the classes IfcRoot and IfcObject.

Table 7. Conversions pattern for attributes.

IFC EXPRESS	ISO/TC 211 UML	Description
■ Attribute name in UpperCamelCase style	Property name in lowerCamelCase style	ISO 19103 requires that property names shall be in lowerCamelCase style.
■ Simple EXPRESS datatypes	Simple IFC datatypes	All occurrences of simple EXPRESS datatypes are replaced with equivalent simple IFC datatypes in order to avoid specific EXPRESS datatypes in the UML model.
■ Aggregation datatypes (ARRAY, LIST, and SET)	Multiplicity from the aggregate declaration, and tagged value "aggregationType"	Aggregation datatypes as value domain are handled as described in Section 4.1.5. Complex aggregations of aggregations are not handled yet but could be handled with additional collection datatypes.
■ UNIQUE Statements	Constraint	Attributes that shall be unique for all instances are defined in EXPRESS with the UNIQUE statement. The equivalent concept in UML is to define properties as unique with the tag "IsUnique". However, UNIQUE statements in EXPRESS can define unique combinations of several attributes. Such requirements can only be described as constraints in UML.
■ DERIVE Statements	Constraint and derived property	Derived attributes are defined in EXPRESS with the DERIVE statement. The equivalent concept in UML is to define properties as derived with the tag "IsDerived". Besides, the derive rule from EXPRESS is converted to a UML constraint, as described in Section 4.1.9.
■ OPTIONAL Statements	Multiplicity with lower bound 0	Optional attributes are defined in EXPRESS with the OPTIONAL statement. The equivalent concept in UML is to set the lower bound of the multiplicity to 0. In some cases, there is a conflict between the OPTIONAL statement and the lower bound for aggregate datatypes. One example is the Roles attribute in IfcOrganization: `Roles:OPTIONAL LIST [1:?]  OF IfcActorRole` We suggest that OPTIONAL shall define the lower bound in UML instead of the aggregate lower bound in such cases.

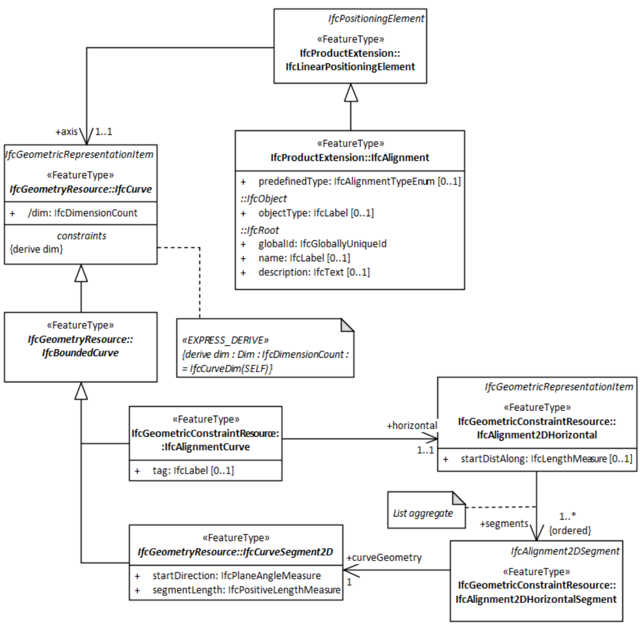


Figure 8. UML class diagram showing examples of IFC attributes converted to UML.

4.1.9. Constraints

The IFC EXPRESS schemas utilize constraints to specify restrictions on types, entities and attributes, based on the EXPRESS keywords UNIQUE, INVERSE, DERIVE and WHERE, while constraints on UML elements are specified in the Object Constraint Language (OCL) [59]. The syntax from EXPRESS constraints may be translated to OCL, but the syntaxes are significantly distinct. The translation is not easy to automate and would most likely require manual work.

Furthermore, WHERE constraints on datatypes and attributes may be split into parts and stored in tagged values that can be used for deriving requirements in implementation schemas. e.g., the constraint "`ValidRange:1 <= SELF <= 31`" can be stored as one tagged value for the lower valid value and one for the upper valid value. The tagged values can then be implemented as "minInclusive" and "maxInclusive" restrictions in XML schemas and as a WHERE declaration in EXPRESS schemas. Similar to translations to OCL, splitting the declaration and storing parts as tagged values is a complex operation that would most likely require manual work.

Our conversion pattern converts EXPRESS constraints to UML constraints with the EXPRESS syntax maintained, and with properties of UML constraints derived, as described in Table 8. A future extension may include conversion to OCL and splitting into tagged values. Figure 9 shows some examples of IFC constraints in the UML model.

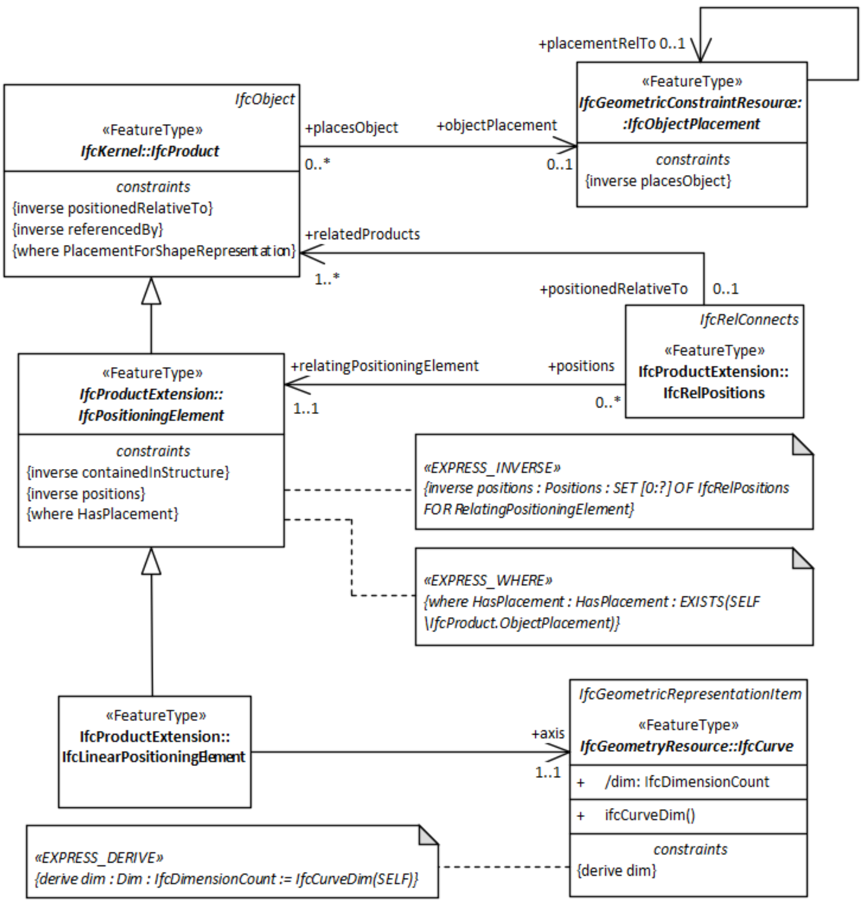


Figure 9. UML class diagram showing examples of IFC constraints converted to UML.

Table 8. Conversion pattern for IFC EXPRESS constraints.

UML Constraint Property	Conversion Pattern
Constraint type	UNIQUE declarations: EXPRESS_UNIQUE INVERSE declarations: EXPRESS_INVERSE DERIVE declarations: EXPRESS_DERIVE WHERE declarations: EXPRESS_WHERE
Constraint name	UNIQUE declarations: "unique" & label WHERE declarations: "where" & label INVERSE declarations: "inverse" & label DERIVE declarations: "derive" & attribute name
Constraint notes	EXPRESS constraint declaration

Besides maintaining the constraints, the conversion pattern extracts the association role name and multiplicity on the side of the related entity from INVERSE declarations. The INVERSE declaration "`Positions:SET [0:?] OF IfcRelPositions FOR RelatingPositioningElement`" for IfcPositioningElement is extracted to the role name "positions" and multiplicity 0..* on the IfcRelPositions association end, as shown in Figure 9.

4.1.10. Functions and Rules

EXPRESS functions are used for calculating values of derived attributes and for validating the content of an IFC dataset, while EXPRESS rules are used for restricting the content of a dataset. Similar to constraints, functions and rules are written in the specific EXPRESS syntax. The equivalent concept to EXPRESS functions and rules in UML is operations.

We propose to create one UML interface for EXPRESS functions and one for rules in each UML package. Every function and rule from the individual EXPRESS schema are defined as operations within these interfaces, and the original EXPRESS syntax is maintained as code. With the interface approach, relevant operations can be realized in the classes where they shall be used. For example, the class IfcCurve implements the operation ifcCurveDim in order to calculate the derived attribute "dim", as illustrated in Figure 10.

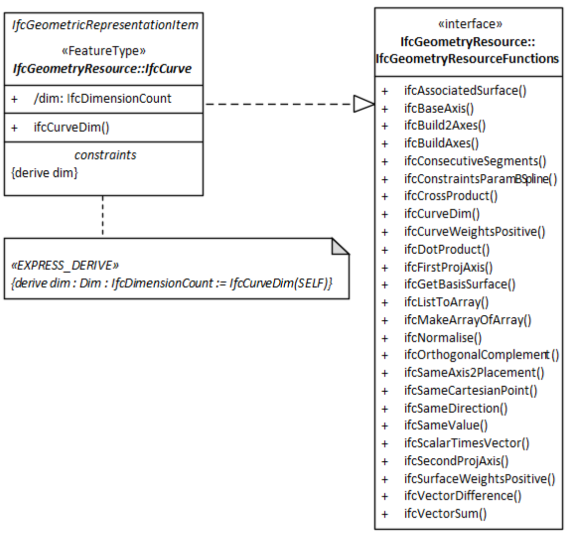


Figure 10. UML class diagram showing realization of the operation IfcCurveDim in the class IfcCurve.

4.2. Relating the IFC Model to Abstract Concepts from ISO/TC 211 Standards

4.2.1. Core Concepts

ISO 19103 defines the MDA approach for models of geospatial information with four layers of abstraction, as illustrated in Figure 11. Metamodels in the first level define how information models shall be specified; abstract schemas in the second level define basic information concepts; application schemas in level 2 reuse the abstract concepts and define specific domain models; while the fourth level contains implementation schemas for specific implementation technologies. The conceptual schemas in level two and three are independent of implementation technologies, while the implementation schemas in level four are derived from conceptual schemas based on rules for conversions.

The IFC model is a domain-specific model for the built environment, similar to OGC CityGML and LandInfra/InfraGML. The conversion pattern described in Section 4.1 converts the IFC information model to a conceptual application schema (level three in Figure 11) based on the metamodels in the top level of Figure 11—the ISO 19103 UML Profile and the ISO 19109 GFM. Besides, the IFC model contains several core concepts that may be compared and related to equivalent concepts from ISO/TC 211 abstract schemas on level 2. In particular, the schemas in the Core and Resource layers from the original IFC schema structure contain abstract concepts where equivalent concepts may be found in ISO/TC 211 schemas.

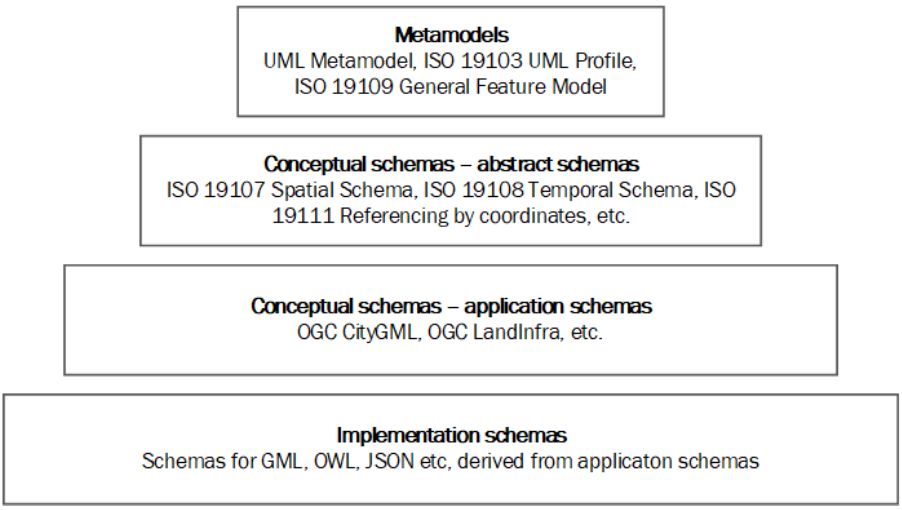


Figure 11. ISO 19103 MDA Levels of abstraction. Adapted from [60].

The most fundamental concepts for objects, properties and relationships in IFC are defined in the IFC Kernel schema. The concepts are similar to those defined in the ISO 19109 GFM (Figure 12), which defines concepts for feature types, attribute types, and feature associations types for use in UML models of geospatial information. However, the schemas for IFC Kernel and ISO 19109 GFM are at different levels of abstraction according to MDA: the ISO 19109 GFM is a metamodel that is implemented in UML models through stereotypes. In contrast, the IFC Kernel schema defines abstract classes with attributes that are inherited by implementable subclasses. Our conversion pattern defines all IFC EXPRESS entities as UML classes that implement the GFM metaclass FeatureType through the stereotype "FeatureType". Entity attributes from IFC EXPRESS are defined as GFM PropertyTypes (AttributeType or FeatureAssociationRole) in UML.

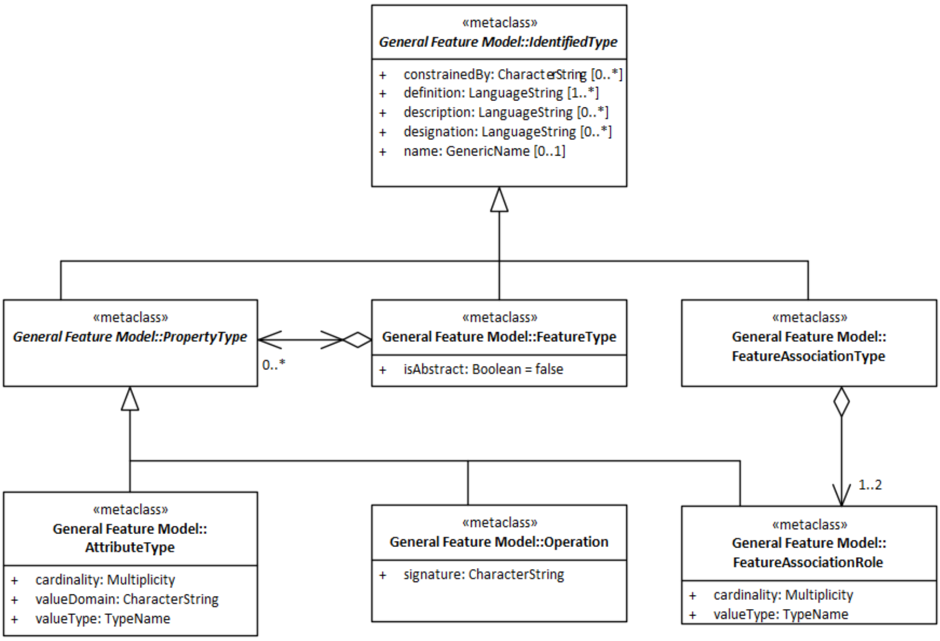


Figure 12. UML class diagram showing the ISO 19109 General Feature Model (GFM).

Figure 13 shows some of the core classes from the IFC Kernel schema. The ultimate superclass in IFC is IfcRoot, which defines common properties that are inherited to the majority of implementable IFC classes through a hierarchy of subclasses. IfcRoot can be considered the core realization of a GFM FeatureType and its superclass IdentifiedType. The class IfcObject is the abstract superclass of implementable classes that typically represent real-world features in GIS standards. For example, the instantiable class IfcAlignment is a subclass of IfcObject and represents the same real-world feature as the InfraGML class Alignment.

The class IfcTypeObject represents a concept that is less used in GIS standards. An instance of IfcTypeObject defines characteristics that are common for several instances of IfcObject subclasses. The complete information for one specific real-world feature—e.g., a railing—may be described with two instances in an IFC data set: one instance of an IfcObject subclass (IfcRailing) with individual information such as location, and one instance of a related IfcTypeObject subclass (IfcRailingType) that contains standard information for all instances of this specific type of railing.

IfcPropertyDefinition defines the concept for flexible assignment of properties to individual objects and type objects through property sets (IfcPropertySet). While GIS standards often define fixed properties for each feature type, IFC has an extensive use of property sets that are defined outside of the IFC schemas, based on the Property Set Definition [61] structure. An equivalent concept exists in InfraGML, with the classes Property and PropertySet that can be assigned to any feature instance. The property set concept represents a flexible approach, but it is more challenging for implementation and information exchange, as the description of the content is defined outside of the schema.

Subclasses of IfcRelationship are used for relating instances of IFC classes, where the relationship class defines the type of relation: assign, associate, connect, declare, decompose or define. Unlike GFM feature associations, IFC relationships are classes that may have properties. Furthermore, the IFC schemas do not define specific relations between specific classes; only the concepts for relations between instances are defined. Similar as to property sets, this approach is more flexible but may also be more challenging for implementation and information exchange.

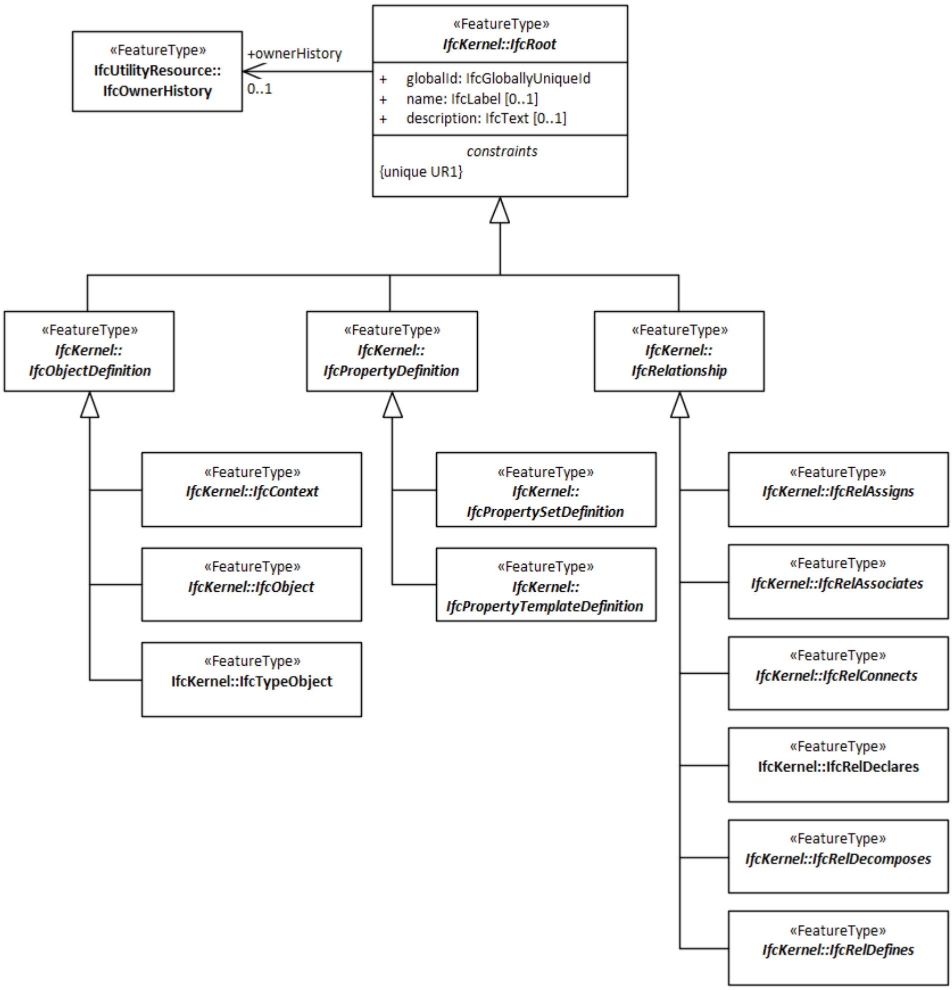


Figure 13. UML class diagram showing the IFC Kernel core classes.

4.2.2. IFC Resource Schemas and ISO/TC 211 Abstract Conceptual Schemas

The relation between simple IFC EXPRESS datatypes and ISO 19103 datatypes is defined in the conversion pattern in Section 4.1, where IFC simple datatypes are automatically defined as realizations of ISO 19103 datatypes. Besides, other abstract concepts defined in IFC resource schemas and ISO/TC 211 abstract conceptual schemas can be considered equivalent as well. Improved interoperability between IFC and GIS schemas could be achieved by utilizing ISO/TC 211 datatypes—and their mapping to XML types—in IFC schemas. Therefore, we investigated relations between three IFC resource schemas and possibly equivalent ISO/TC 211 schemas: IfcDateTimeResource, IfcMeasureResource and IfcGeometryResource. Other IFC resource schemas are candidates for relations to ISO/TC 211 schemas as well and could be investigated through further studies. Table 9 lists the three selected IFC resource schemas and their ISO/TC 211 counterpart.

Table 9. The studied IFC resource schemas and counterpart ISO/TC 211 abstract conceptual schemas.

IFC Schema	ISO/TC 211 Schema
IfcDateTimeResource	ISO 19103 (Date and Time types) ISO 19108 Temporal schema
IfcMeasureResource	ISO 19103 (Measure types)
IfcGeometryResource	ISO 19107 Spatial Schema

The IFC schema IfcDateTimeResource defines datatypes for time-related information. The datatypes are defined as simple datatypes in the EXPRESS schemas, supported by textual descriptions of usage and format in the IFC documentation. For example, IfcDate is defined as a string that shall be formatted as YYYY-MM-DD, while IfcTime shall be formatted as hh:mm:ss. The format requirements are only defined in the textual documentation; they are not defined in the schema. ISO 19103 and ISO 19108 defines similar concepts for date and time types, and ISO 19136 defines default mappings to standard XML schema types. The XML schema sets structure restrictions on the date and time strings. We have defined IFC date and time datatypes as realizations of equivalent ISO 19103 and ISO 19108 datatypes, as illustrated in Figure 14.

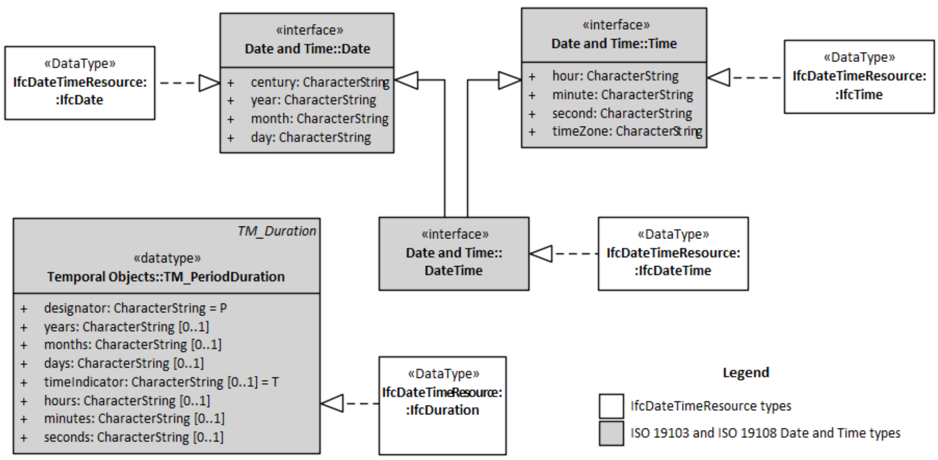


Figure 14. UML class diagram showing IFC Date and time types as realizations of ISO 19103 and ISO 19108 datatypes.

The IFC schema IfcMeasureResource defines datatypes for measure information. Similar to the IFC date and time datatypes, IFC measure datatypes are defined as simple datatypes with textual descriptions of the expected content. The unit of measure is not a part of each measured value—except for in the generic class IfcMeasureWithUnit—but default units for an IFC dataset can be specified in the attribute unitsInContext in the IfcProject class. ISO 19103 defines measure types for use in geospatial information models, where the unit of measure is given for each measured value. In order to utilize the ISO 19103 datatypes for IFC, we have defined IFC measure datatypes as realizations of equivalent ISO 19103 measure datatypes, as shown for some examples in Figure 15. Several IFC measure datatypes do not have equivalent datatypes in ISO 19103. These are defined as realizations of the core class Measure.

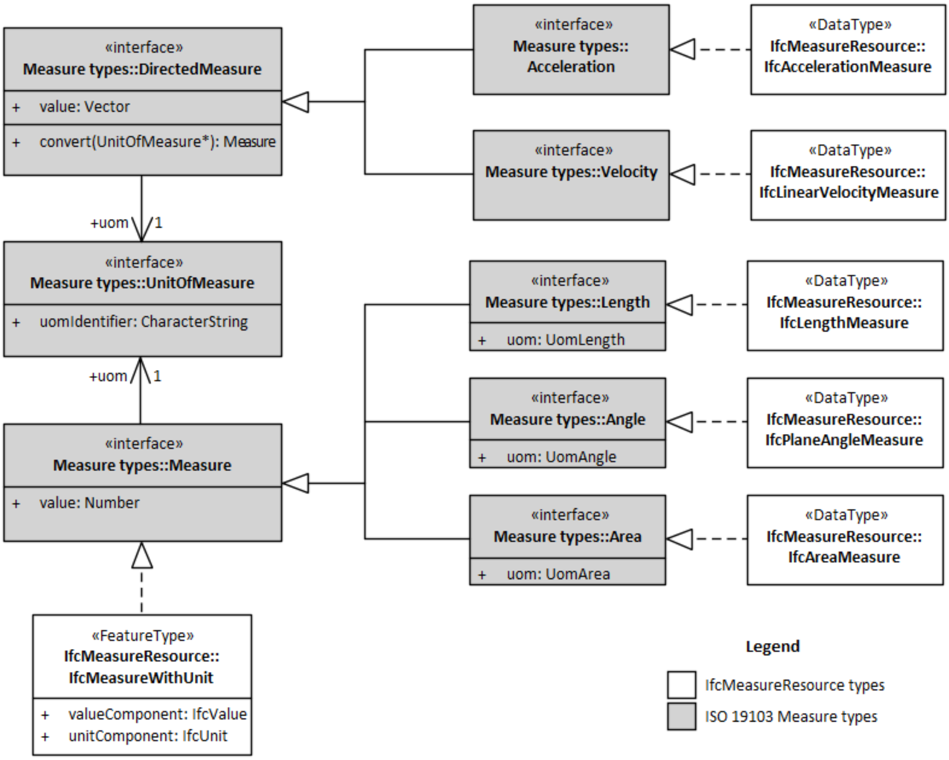


Figure 15. UML class diagram showing examples of IFC measure types as realizations of ISO 19103 measure types.

4.2.3. Concepts for Geometry

Transformation of geometry characteristics between IFC and information models based on ISO/TC 211 concepts is a complex issue that has been discussed in many studies (e.g., Donkers et al. [16]; Deng et al. [17]; Kang and Hong [32] and Ohori et al. [16,34]). The main challenge is the use of swept solids and constructive solids for the geometry of products in IFC, while GIS information models use Boundary Representation (BREP) to represent the geometry of features [17]. The geometry concepts in IFC are defined in three IFC resource schemas: The basic concepts for geometry are defined in IfcGeometryResource. More complex concepts for solid geometry representation of products are defined in IfcGeometricModelResource and IfcGeometryConstraintResource. We have limited our studies on geometry to relations between basic concepts for geometry in IfcGeometryResource and the core ISO/TC 211 geometry model in ISO 19107 [22]. We suggest that several IFC geometry types can be defined as realizations of equivalent ISO 19107 geometry types, as shown for some examples in Figure 16. However, not all IFC geometry types are defined in ISO 19107 and vice versa.

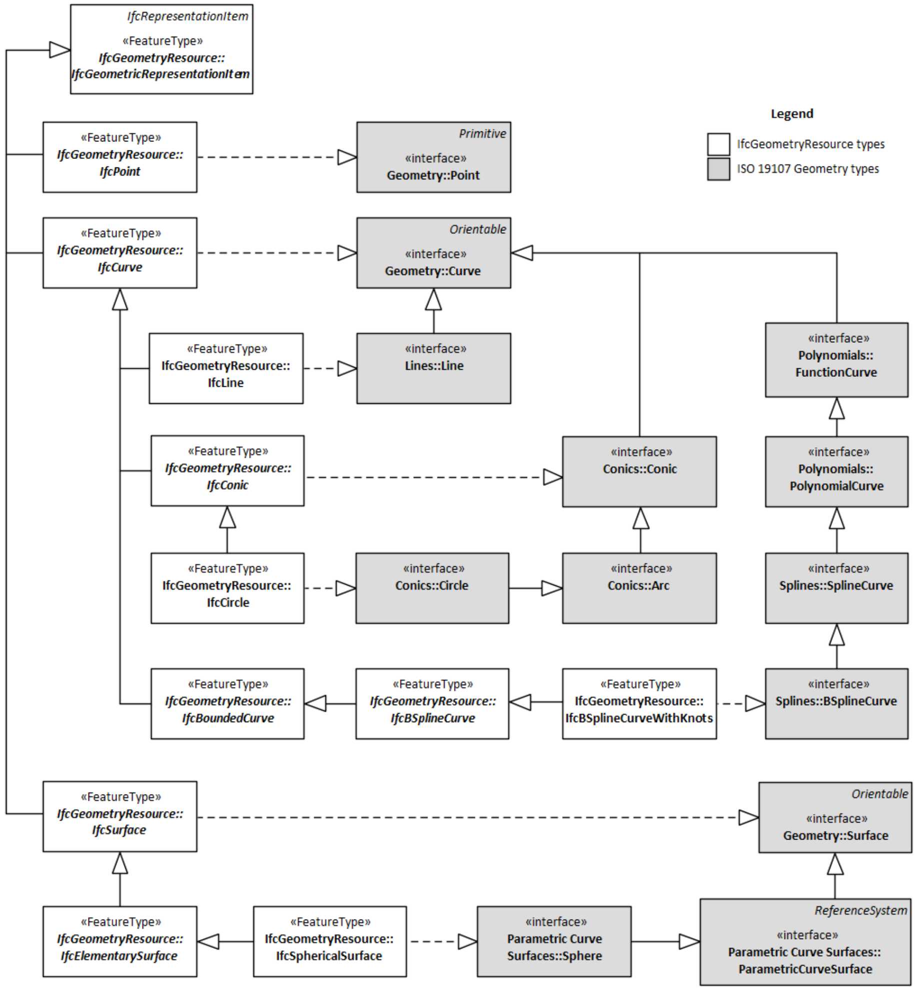


Figure 16. UML class diagram showing examples of IFC geometry types as realizations of ISO 19107 geometry types.

4.3. Model-Driven Derivation of Implementation Schemas

4.3.1. Conversion from UML to GML

ISO 19136 defines the GML exchange format for geospatial information; and rules for conversion from UML to XML implementation schemas for GML. The open-source software ShapeChange [56] has implemented the rules from ISO 19136 as well as additional rules for the automated derivation of GML implementation schemas. Our converted UML model of IFC was designed according to ISO 19109, and GML implementation schemas could then be derived directly with ShapeChange. We used the same modularized structure for the GML schemas as for the UML model, with one main application schema and individual GML schema files for each UML package.

The realization associations described in Sections 4.1 and 4.2 were used for automated schema mappings between IFC and ISO/TC 211 datatypes, in order to improve interoperability with GIS applications and databases. IFC classes (feature types and datatypes) were analyzed for possible mapping to ISO/TC 211 datatypes if they were (1) referenced as the value domain of a property, and (2) were connected to an ISO/TC 211 datatype through a realization. Two approaches for schema mapping were applied, depending on the complexity of the referenced class and its hierarchy of specializations: Replacement with an ISO/TC 211 datatype; or modification of the IFC class.

For simplicity, we made a presumption for the analyzed classes: specializations of an analyzed class describe equivalent concepts as the connected ISO/TC 211 datatype if the subtype hierarchy is entirely within the same package as the class itself or do not have any own properties. For example, specializations of a measure type within IfcMeasureResource all describe equivalent concepts to specializations of the related ISO 19103 measure datatype. Based on the presumption, the value domains of 797 of 827 analyzed properties could be replaced with ISO/TC 211 datatypes. The remaining 30 properties all referenced the classes IfcSurface or IfcCurve. IfcSurface has the complex specialization IfcSectionedSurface in the IfcGeometricModelResource package, while IfcCurve has the complex specialization IfcAlignmentCurve in the IfcGeometricConstraintResource package. We considered that the complex specializations needed to be maintained, and the two classes IfcCurve and IfcSurface could therefore not be replaced with the ISO 19107 datatypes Curve and Surface. The link to ISO 19107 for these property types was instead established by adding geometry attributes with datatype Curve for IfcCurve and Surface for IfcSurface.

Figures 17 and 18 show XML schema views of the classes IfcAligment and IfcAlignmentCurve from the derived GML implementation schemas. IfcAlignment is a subtype of IfcLinearPositioningElement and inherits the "axis" association to IfcCurve. IfcAlignmentCurve is a subtype of IfcCurve, which was connected to the ISO 19107 datatype "Curve". The new attribute "curve" enables IfcCurve and subtypes to have curve geometries according to ISO 19107, while at the same time, the complexity of the classes and the hierarchy is maintained. The property type for the attribute "dim" in IfcCurve is mapped from "IfcDimensionCount" to the XML type "integer", while the property type for the attribute "tag" in IfcAlignmentCurve is mapped from "IfcLabel" to the XML type "string". Figure 19 shows a fragment of an example GML instance of IfcAlignment and IfcAlignmentCurve.

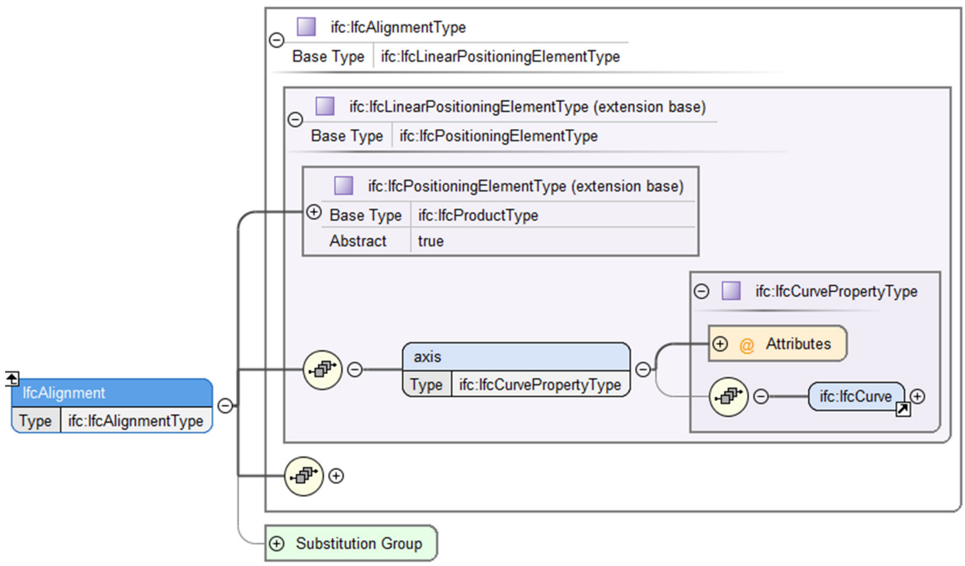


Figure 17. Extract from the IfcProductExtension GML schema, illustrated in Oxygen XML Editor [57].

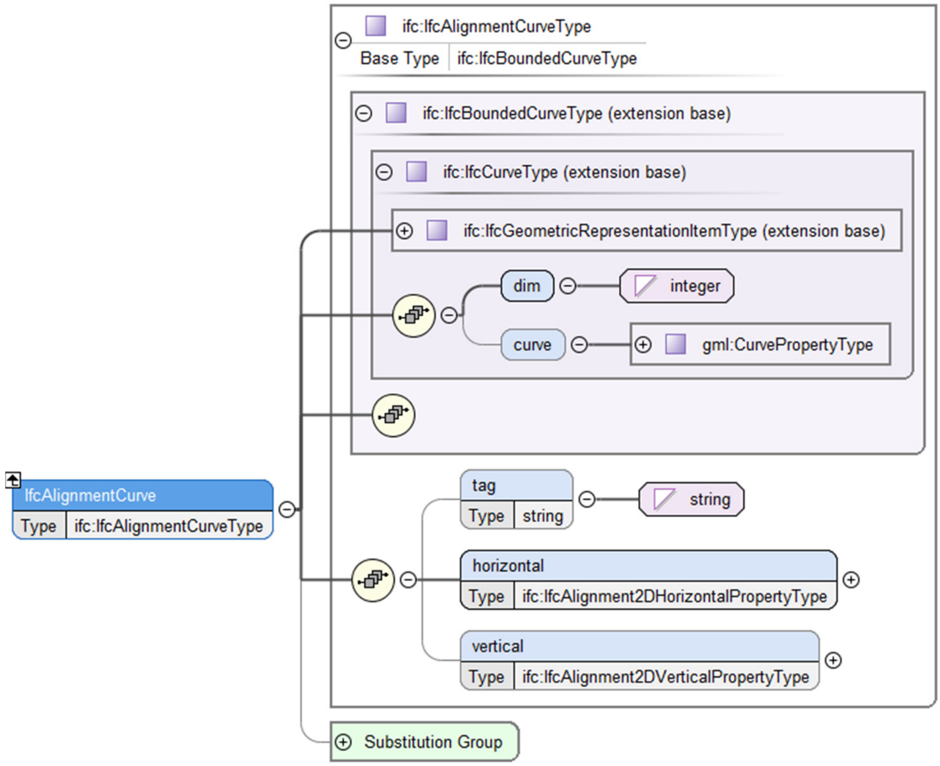


Figure 18. Extract from the IfcGeometricConstraintResource GML schema, illustrated in Oxygen XML Editor [57].

```
<ifc:IfcSite>                                                                    [36 lines]
<ifc:IfcAlignment gml:id="vegvesen.no.nvdb.88263528">
 <ifc:globalId>30ce2eac-48f8-494f-b556-e35fd683687d</ifc:globalId>
 <ifc:name>Alignment_88263528</ifc:name>
 <ifc:description>Example alignment feature</ifc:description>
 <ifc:axis>
  <ifc:IfcAlignmentCurve>
   <ifc:dim>3</ifc:dim>
   <ifc:curve>
    <gml:LineString srsName="EPSG:6150" srsDimension="3">
     <gml:posList>133189.363 1313776.627 135.218 133289.351 1313747.742 142.1
                  133364.851 1313764.775 146.8 133543.175 1313888.363 154.8
                  133679.496 1314048.563 157.3 133839.121 1314127.771 163
                  134048.109 1314279.143 179.9 135091.153 1314562.761 205.954</gml:posList>
    </gml:LineString>
   </ifc:curve>
   <ifc:horizontal xlink:href="IfcAlignmentExample.gml#vegvesen.no.nvdb.88263528_0_H"/>
  </ifc:IfcAlignmentCurve>
 </ifc:axis>
</ifc:IfcAlignment>
<ifc:IfcAlignment2DHorizontal gml:id="vegvesen.no.nvdb.88263528_0_H">             [64 lines]
```

Figure 19. Extract of an example GML file for IFC.

4.3.2. Conversion from UML to IFC EXPRESS

While rules for conversion from UML to GML are defined in ISO 19136, no such rules are yet defined for conversion to EXPRESS implementation schemas. We suggest a set of conversion rules that reverse the EXPRESS to UML conversion described in Tables 3 and 6–8. Table 10 summarizes our suggested rules for conversion from UML to EXPRESS.

Table 10. Rules for conversion from UML to EXPRESS.

UML Concept	Rules
Package	Each UML Package is implemented as an EXPRESS schema.
	Package dependencies are derived from the value domains of UML properties and implemented as EXPRESS schema references.
Datatype	UML Datatypes are implemented as EXPRESS types.
	Classes with realization associations to ISO 19103 simple datatypes are implemented as types of EXPRESS simple datatypes.
	Specializations of UML datatypes are implemented as EXPRESS types of the type they are specializing.
	UML datatypes with tagged value "aggregationType" = "LIST" or "SET" are implemented as EXPRESS aggregated types specified in the tagged value and multiplicity defined from the UML property "component".
	UML datatypes with tagged value "aggregationType" = "ARRAY" are implemented as EXPRESS arrays. The lower index range is extracted from the tagged value "minIndex", and the upper index range is calculated from "minIndex" + difference between the upper and lower multiplicity of the UML property "component".
Enumeration	UML enumerations are implemented as EXPRESS enumerations with identical content as the UML enumeration.
Union	UML unions are implemented as EXPRESS select types with content defined from the value domain of each union member.
FeatureType	UML FeatureTypes are implemented as EXPRESS entities.
	Generalizations and specializations of UML FeatureTypes are implemented as EXPRESS supertype and subtype statements.
	Abstract UML FeatureTypes are implemented as abstract EXPRESS entities
FeatureType property	UML properties (attributes and associations)—except derived properties—are implemented as EXPRESS attributes with name in UpperCamelCase style.
	Derived UML properties are not implemented as EXPRESS attributes.
	The value domains of UML property types are implemented as EXPRESS datatype of attributes.
	UML properties with lower bounds 0 are implemented as optional EXPRESS attributes.
	UML properties with tagged value "aggregationType" = "LIST" or "SET" are implemented as EXPRESS attributes of aggregated types specified in the tagged value and multiplicity defined in the UML property.
	UML properties with tagged value "aggregationType" = "ARRAY" are implemented as EXPRESS attributes of arrays. The lower index range is extracted from the tagged value "minIndex", and the upper index range is calculated from "minIndex" + difference between the upper and lower multiplicity of the UML property.
Constraint	UML Class constraints of type "EXPRESS_UNIQUE" are implemented as EXPRESS entity "UNIQUE" statements. The statement is implemented from the UML constraint notes.
	UML Class constraints of type "EXPRESS_INVERSE" are implemented as EXPRESS entity "INVERSE" statements. The statement is implemented from the UML constraint notes.
	UML Class constraints of type "EXPRESS_DERIVE" are implemented as EXPRESS entity "DERIVE" statements. The statement is implemented from the UML constraint notes.
	UML Class constraints of type "EXPRESS_WHERE" are implemented as EXPRESS entity or type "WHERE" statements. The statement is implemented from the UML constraint notes.
Interface operation	The code for each operation in UML interfaces is implemented as an EXPRESS function or rule statement.

In order to automate the conversion, we developed a script in Enterprise Architect, where we implemented the conversion rules from Table 10. As for the GML implementation schemas, we used the same modularized structure as in the UML model, with one EXPRESS schema file for each UML package. Figure 20 shows a selection of declarations from the exported EXPRESS schema files, representing the same UML classes as the extracts from GML schemas in Figures 17 and 18.

```
ENTITY IfcCurve
 ABSTRACT SUPERTYPE OF
   (ONEOF(IfcBoundedCurve, IfcPcurve, IfcConic, IfcSurfaceCurve, IfcOffsetCurve, IfcLine));
 SUBTYPE OF (IfcGeometricRepresentationItem);
 DERIVE
   Dim : IfcDimensionCount := IfcCurveDim(SELF);
END_ENTITY;

ENTITY IfcBoundedCurve
 ABSTRACT SUPERTYPE OF
   (ONEOF(IfcIndexedPolyCurve, IfcCurveSegment2D, IfcPolyline, IfcCompositeCurve, IfcBSplineCurve,
   IfcTrimmedCurve, IfcAlignmentCurve));
 SUBTYPE OF (IfcCurve);
 INVERSE
   PositioningElement : IfcLinearPositioningElement FOR Axis;
END_ENTITY;

ENTITY IfcAlignmentCurve
 SUBTYPE OF (IfcBoundedCurve);
   Tag : OPTIONAL IfcLabel;
   Vertical : OPTIONAL IfcAlignment2DVertical;
   Horizontal : IfcAlignment2DHorizontal;
END_ENTITY;

ENTITY IfcAlignment2DHorizontal
 SUBTYPE OF (IfcGeometricRepresentationItem);
   StartDistAlong : OPTIONAL IfcLengthMeasure;
   Segments : LIST [1:?] OF IfcAlignment2DHorizontalSegment;
 INVERSE
   ToAlignmentCurve : SET [1:?] OF IfcAlignmentCurve FOR Horizontal;
END_ENTITY;
```

Figure 20. Selected declarations from the derived EXPRESS schema files.

5. Evaluation and Discussion

Literature studies have shown that there has been an increasing interest and amount of research on the integration of BIM and GIS over the last years. Information exchange and conversion of semantics have been identified as a significant challenge for such integration. Still, most studies have investigated the integration of application schemas and data instances and not integration at the abstract conceptual level where the core semantics are defined. We aimed to improve the integration of core concepts from information models for BIM and GIS—independent of application schema—by moving the core concepts into a shared environment. The approach for reaching this goal was to describe concepts from both application domains in UML according to modelling rules from ISO/TC 211.

Our first research question asked for potential solutions and challenges for the conversion of the IFC information model to a UML model according to GIS standards. We developed a pattern that covers core conversion as well as the more specific use of EXPRESS for IFC and UML for GIS information models, and applied the pattern for conversion of the draft IFC Road EXPRESS schemas. The conversion pattern (described in Section 4.1) and the results from the conversion show that most EXPRESS concepts used for IFC have equivalent concepts in UML according to ISO/TC 211 standards and profiles and can be converted through simple conversions. However, we identified challenges with some concepts that need a more complex and controlled conversion—specifically aggregation datatypes (Table 6) and aggregation properties (Table 7); constraints (Table 8); and functions and rules.

We defined a UML representation of EXPRESS aggregation data types that maintained the collections, but that may need further improvement for complex aggregations of other aggregations. For constraints, functions and rules, the distinct syntax for different implementation technologies is a challenge. We considered a translation from EXPRESS syntax to OCL and tagged values, for implementation in other technologies than EXPRESS. All codes would then need to be maintained in the distinct syntaxes, or complex conversion rules from OCL to EXPRESS must be developed. Whether or not such translation is needed will depend on the use of the implementation technologies. We believe a translation is relevant in order to enable the use of GML implementation schemas for controlling data according to constraints, functions and rules.

Our second research question asked for relationships that could be defined between concepts in the IFC information model and core ISO/TC 211 standards. The results in Section 4.2 show that the ISO/TC 211 compliant UML representation of the IFC model was useful for comparing and linking concepts defined in different application domains. Core IFC classes in the IFC Kernel schema can be considered core realizations of ISO 19109 GFM metaclasses. Furthermore, basic IFC datatypes for date, time, and measures can be linked to equivalent datatypes from ISO/TC 211 standards. Therefore, a shared understanding of these datatypes may be achieved. However, we will question whether it is optimal that any of the two application domains shall define these and other universal datatypes. Instead, referring to datatypes independent of application domain—defined in core external vocabularies—would improve interoperability between the two application domains as well as with other domains.

Unlike datatypes for date, time and measures, datatypes and concepts for geometry are more naturally owned by one of the two application domains in question. We defined links between several core geometry concepts in the IFC model and ISO/TC 211 models, but we also observed that not all geometry types could be linked one-to-one. Further harmonization of geometry types between the two application domains would give significant improvement of interoperability. Basic geometry types should only be defined in one of the application domains, and then used in both, as well as in any other application domain concerning geospatial information.

The third research question asked for potential solutions and challenges for deriving implementation schemas for EXPRESS and GML from UML models of IFC. The results in Section 4.3 show that conversions were possible to both GML and EXPRESS directly from the UML model, supported by some implementation-specific semantics in tagged values. We applied a mapping to ISO/TC 211 datatypes for the GML conversion, in order to achieve better interoperability with GIS applications and databases. The mapping of geometry datatypes was essential for enabling interoperability that included geometry characteristics and was also the most challenging part, due to the distinct geometry models in IFC and ISO 19107. The conversion would be less complicated if the two application domains were using a shared model of geometry datatypes.

The IFC-UML model developed by buildingSMART is expected to replace EXPRESS schemas as the original conceptual IFC model in the future. The comparison in Table 4 shows that there are some minor differences between the draft IFC-UML model and our ISO/TC 211 compliant UML model. If the final IFC-UML model has the same structure as the draft, our conversion pattern must be adapted for conversions from the IFC-UML model to an ISO/TC 211 compliant UML model. However, we consider that a better solution would be to model the original IFC-UML model according to ISO/TC 211 profiles and rules. Our bi-directional conversion has shown that all semantics from the IFC model could be maintained in an ISO/TC 211 compliant model and that implementation schemas could be derived from the model.

6. Conclusions and Further Work

This study has investigated whether harmonized core UML concepts for information models in BIM and GIS can be a way forward for improved interoperability between the two application domains. We developed a conversion pattern from IFC EXPRESS schemas from the BIM application domain to UML models according to ISO/TC 211 standards from the GIS application domain. The model was further refined with links between core concepts from IFC and ISO/TC 211 models and tested for conversions to implementation schemas in the GML and EXPRESS formats.

The research showed that all semantics from the IFC model could be converted to an ISO/TC 211 compliant UML model and that implementation schemas for both application domains could be derived from the UML model. Some implementation-specific semantics for GML and EXPRESS need to be maintained as tagged values. Enhanced technology-independent implementation can be achieved by translating constraints, rules and functions specified in the EXPRESS syntax to OCL and tagged values for implementation in GML schemas. Most important; the representation of information models from both application domains in a shared environment was useful for understanding, linking, mapping and reusing concepts between IFC and core GIS standards.

Improved interoperability at application schema and data instance level may be achieved by using shared core concepts as a fundament for application schemas in both domains. Exchange of information for use in planning, constructing, using, and maintaining the built environment would be a less complicated issue with a shared understanding of the semantics for digital representation of real-world phenomena. We consider such common understanding more important than the choice of modelling language.

Further work following this study should include an improved linking and harmonization of core concepts. Domain-specific concepts should be replaced with shared concepts, and externally defined vocabularies should be reused in both application domains when possible. A shared model for geometry types is particularly important for improved interoperability.

Author Contributions: Conceptualization, Knut Jetlund; methodology, Knut Jetlund, Erling Onstein and Lizhen Huang; validation, Knut Jetlund; formal analysis, Knut Jetlund; investigation, Knut Jetlund; data curation, Knut Jetlund; writing—original draft preparation, Knut Jetlund; writing—review and editing, Erling Onstein and Lizhen Huang; visualization, Knut Jetlund; supervision, Erling Onstein and Lizhen Huang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Norwegian Public Roads Administration and The Research Council of Norway. NTNU Open Access publishing funds cover the article processing charges.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. buildingSmart International. IFC Specifications Database. 2019. Available online: https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/ (accessed on 10 January 2020).
2. ISO. *Industrial Automation Systems and Integration—Product Data Representation and Exchange—Part 11: Description Methods: The EXPRESS Language Reference Manual*; ISO/TC 184 SC 4, ISO 10303-11:2004; ISO: Geneva, Switzerland, 2004.
3. Object Management Group. *Unified Modelling Language Specification Version 2.5.1*; Object Management Group: Needham, MA, USA, 2017.
4. ISO. *Information technology—Object Management Group Unified Modeling Language (OMG UML)—Part 2: Superstructure*; ISO/IEC JTC 1/SC 7, ISO/IEC 19505-2:2012; ISO: Geneva, Switzerland, 2012.
5. ISO. *Information technology—Object Management Group Unified Modeling Language (OMG UML)—Part 1: Infrastructure*; ISO/IEC JTC 1/SC 7, ISO/IEC 19505-1:2012; ISO: Geneva, Switzerland, 2012.
6. Object Management Group. *Model Driven Architecture (MDA) Guide Rev. 2.0*; Object Management Group: Needham, MA, USA, 2014.
7. Coetzee, S.; Plews, R.; Brodeur, J.; Hjelmager, J.; Jones, A.; Jetlund, K.; Grillmayer, R.; Wasström, C. Standards—Making Geographic Information Discoverable, Accessible and Usable for Modern Cartography. In *Service-Oriented Mapping: Changing Paradigm in Map Production and Geoinformation Management*; Döllner, J., Jobst, M., Schmitz, P., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 325–344.
8. Liu, X.; Wang, X.; Wright, G.; Cheng, J.C.; Li, X.; Liu, R. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 53. [CrossRef]
9. Zhu, J.; Wright, G.; Wang, J.; Wang, X. A Critical Review of the Integration of Geographic Information System and Building Information Modelling at the Data Level. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 66. [CrossRef]

10. Song, Y.; Wang, X.; Tan, Y.; Wu, P.; Sutrisna, M.; Cheng, J.; Hampson, K. Trends and Opportunities of BIM-GIS Integration in the Architecture, Engineering and Construction Industry: A Review from a Spatio-Temporal Statistical Perspective. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 397. [CrossRef]
11. Kumar, K.; Labetski, A.; Ohori, K.A.; Ledoux, H.; Stoter, J. The LandInfra standard and its role in solving the BIM-GIS quagmire. *Open Geospat. Data Softw. Stand.* **2019**, *4*. [CrossRef]
12. Onstein, E.; Tognoni, M.G. "Building permits" as proof of concepts in merging GIS and BIM information: A case study. *WIT Trans. Built Environ.* **2017**, *169*, 155–166.
13. Roxin, A.; Hbeich, E. Semantic interoperability between BIM and GIS–review of existing standards and depiction of a novel approach. In Proceedings of the 36th CIB W78–Information Technology for Construction, Northumbria, UK, 18–20 September 2019.
14. Gilbert, T.; Rönsdorf, C.; Plume, J.; Simmons, S.; Nisbet, N.; Gruler, H.-C.; Kolbe, T.; van Berlo, L.; Mercer, A. *Built Environment Data Standards and Their Integration: An Analysis of IFC, CityGML and LandInfra*; Open Geospatial Consortium: Wayland, MA, USA; buildingSMART International: London, UK, 2020.
15. Open Geospatial Consortium. OGC CityGML. 2019. Available online: https://www.opengeospatial.org/standards/citygml (accessed on 18 February 2020).
16. Donkers, S.; Ledoux, H.; Zhao, J.; Stoter, J. Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Trans. GIS* **2016**, *20*. [CrossRef]
17. Deng, Y.; Cheng, J.C.; Anumba, C. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Autom. Constr.* **2016**, *67*, 1–21. [CrossRef]
18. ISO. *Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management industries—Part 1: Data Schema*; ISO/TC 59/SC 13, ISO 16739-1:2018; ISO: Geneva, Switzerland, 2018.
19. ISO. *Geographic Information—Conceptual Schema Language*; ISO/TC 211, ISO 19103:2015; ISO: Geneva, Switzerland, 2015.
20. ISO. *Geographic information—Rules for Application Schema*; ISO/TC 211, ISO 19109:2015; ISO: Geneva, Switzerland, 2015.
21. ISO. *Geographic information—Geography Markup Language (GML)*; ISO/TC 211, ISO 19136:2007; ISO: Geneva, Switzerland, 2007.
22. ISO. *Geographic Information—Spatial Schema*; ISO/TC 211, ISO 19107:2019; ISO: Geneva, Switzerland, 2019.
23. ISO. *Geographic Information—Referencing by Coordinates*; ISO/TC 211, ISO 19111:2019; ISO: Geneva, Switzerland, 2019.
24. INSPIRE. INSPIRE Data Specifications. 2020. Available online: https://inspire.ec.europa.eu/data-specifications/2892 (accessed on 31 March 2020).
25. CEN/TC 278. *Intelligent Transport Systems—ITS Spatial Data—Data Exchange on Changes in Road Attributes*; CEN/TS 17268:2018 I; CEN: Brussels, Belgium, 2018.
26. Open Geospatial Consortium. OGC LandInfra/InfraGML. 2017. Available online: https://www.opengeospatial.org/standards/infragml (accessed on 18 February 2020).
27. Open Geospatial Consortium. *OGC®Land and Infrastructure Conceptual Model Standard (LandInfra)*; Open Geospatial Consortium: Wayland, MA, USA, 2016.
28. buildingSmart International. Infrastructure Room. 2019. Available online: https://www.buildingsmart.org/standards/rooms/infrastructure (accessed on 30 October 2019).
29. buildingSmart International. Industry Foundation Classes Version 4.2 bSI Draft Standard. 2019. Available online: https://standards.buildingsmart.org/IFC/DEV/IFC4_2/FINAL/HTML/ (accessed on 10 January 2020).
30. Open Geospatial Consortium. *OGC InfraGML 1.0—Part 3—Alignments—Encoding Standard*; Open Geospatial Consortium: Wayland, MA, USA, 2017.
31. Hor, A.-H.; Jadidi, A.; Sohn, G. BIM-GIS INTEGRATED GEOSPATIAL INFORMATION MODEL USING SEMANTIC WEB AND RDF GRAPHS. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 73–79. [CrossRef]
32. Kang, T.W.; Hong, C.H. IFC-CityGML LOD mapping automation using multiprocessing-based screen-buffer scanning including mapping rule. *KSCE J. Civil Eng.* **2018**, *22*, 373–383. [CrossRef]
33. Knoth, L.; Scholz, J.; Strobl, J.; Mittlböck, M.; Vockner, B.; Atzl, C.; Rajabifard, A.; Atazadeh, B. Cross-Domain Building Models—A Step towards Interoperability. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 363. [CrossRef]

34. Arroyo Ohori, G.A.K.; Diakite, A.A.; Krijnen, T.; Ledoux, H.; Stoter, J.E. Processing BIM and GIS Models in Practice. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 311. [CrossRef]
35. Vilgertshofer, S.; Amann, J.; Willenborg, B.; Borrmann, A.; Kolbe, T. Linking BIM and GIS Models in Infrastructure by Example of IFC and CityGML. *Comput. Civil Eng.* **2017**, *2017*, 133–140.
36. Niestroj, M.G.; McMeekin, D.A.; Helmholz, P. Overview of standards towards road asset information exchange. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-4*, 443–450. [CrossRef]
37. ISO. *GIS (Geospatial)/BIM Interoperability*; ISO/TC 59/SC 13, ISO/WD TR 23262; ISO: Geneva, Switzerland, 2019.
38. INTERLINK Project. EUROTL Framework Overview. 2018. Available online: https://www.roadotl.eu/static/eurotl-ontologies/index.html (accessed on 10 January 2020).
39. Luiten, B.; Böhms, M.; O'Keeffe, A.; van Nederveen, S.; Bakker, J.; Wikström, L. A hybrid linked data approach to support asset management. In *Life-Cycle of Engineering Systems: Emphasis on Sustainable Civil Infrastructure, Proceedings of the Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016) Delft, The Netherlands, 16–19 October 2016*; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: London, UK, 2016.
40. Arnold, F.; Podehl, G. *Best of Both Worlds—A Mapping from EXPRESS-G to UML*; Springer: Berlin/Heidelberg, Germany, 1999.
41. Krause, F.L.; Kaufmann, U. Meta-Modelling for Interoperability in Product Design. *CIRP Ann.* **2007**, *56*, 159–162. [CrossRef]
42. Exff Project. 2003. Available online: http://exff.sourceforge.net/docs/e2u_mapping.html (accessed on 15 January 2020).
43. Object Management Group. *Reference Metamodel for the EXPRESS Information Modeling Language*; Object Management Group: Needham, MA, USA, 2015.
44. Pauwels, P.; Terkaj, W. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Autom. Constr.* **2016**, *63*, 100–133. [CrossRef]
45. Terkaj, W.; Pauwels, P. A Method to generate a Modular ifcOWL Ontology. In Proceedings of the 8th International Workshop on Formal Ontologies Meet Industry, Bolzano, Italy, 21–23 September 2017.
46. McGlinn, K.; Wagner, A.; Pauwels, P.; Bonsma, P.; Kelly, P.; O'sullivan, D. Interlinking geospatial and building geometry with existing and developing standards on the web. *Autom. Constr.* **2019**, *103*, 235–250. [CrossRef]
47. Van Berlo, L. *Future Technology Presentation buildingsmart Summit Beijing*; Linkedin: Sunnyvale, CA, USA, 2019.
48. buildingSmart International. NextGen-IFC. 2019. Available online: https://github.com/buildingSMART/NextGen-IFC (accessed on 24 January 2020).
49. Jetlund, K. IFC2GML. 2020. Available online: https://github.com/jetgeo/IFC2GML (accessed on 23 April 2020).
50. ISO/TC 211. Introduction to UML. 2018. Available online: https://github.com/ISO-TC211/UML-Best-Practices/wiki/Introduction-to-UML (accessed on 20 March 2020).
51. Miles, R.R.; Hamilton, K. *Learning UML 2.0*, 1st ed.; O'Reilly: Sebastopol, CA, USA, 2006.
52. Schenck, D.; Wilson, P. *Information Modelling: The EXPRESS Way*; Oxford University Press: New York, NY, USA, 1994.
53. buildingSMART International. IFC4x3—FIRST_DRAFT 2019. Unpublished.
54. ISO/TC 211. Harmonized Model Maintenance Group, The ISO/TC 211 Harmonized Model. 2019. Available online: http://iso.sparxcloud.com (accessed on 30 October 2019).
55. Sparx Systems Pty Ltd. Enterprise Architect Version 15. 2020. Available online: https://sparxsystems.com/products/ea/index.html (accessed on 10 January 2020).
56. Interactive instruments GmbH. ShapeChange. Available online: http://shapechange.net (accessed on 15 January 2020).
57. Syncro Soft SRL. Oxygen XML Editor 22. 2020. Available online: https://www.oxygenxml.com/xml_editor.html (accessed on 20 March 2020).
58. Bellekens, G. UML Best Practice: Attribute or Association. 2011. Available online: https://bellekens.com/2011/08/10/uml-best-practice-attribute-or-association/ (accessed on 10 January 2020).
59. Object Management Group. *The Object Constraint Language Version 2.4*; Object Management Group: Needham, MA, USA, 2014.

60. Jetlund, K.; Onstein, E.; Huang, L. Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 141. [CrossRef]
61. buildingSmart International. PSD: Property Set Definition XSD. 2013. Available online: http://standards.buildingsmart.org/IFC/RELEASE/IFC4/FINAL/PSD/PSD_IFC4.xsd (accessed on 10 January 2020).

Article 5

A structure of UML profiles for modelling of geospatial information in GIS, ITS and BIM

A structure of UML profiles for modelling of geospatial information in GIS, ITS and BIM

Knut Jetlund [1,2]

[1]Norwegian University of Science and Technology, Gjøvik, Norway
[2]The Norwegian Public Roads Administration, Hamar, Norway - knut.jetlund@vegvesen.no

KEY WORDS: information modelling, unified modelling language, model-driven architecture, geographic information systems, intelligent transport systems, building information modelling

ABSTRACT:

This study aims to improve the interoperability between models of geospatial information from the applications domains of Geographic Information Systems (GIS), Intelligent Transport Systems (ITS) and Building Information Models (BIM). A state-of-the-art analysis showed that the Unified Modelling Language (UML) and Model-Driven Architecture (MDA) are used for modelling information in a geospatial context in all three domains, but with different approaches and levels of formality. A structure of formal UML profiles for modelling of geospatial information in GIS, ITS and BIM is suggested and tested for implementation. The Core Geospatial Profile (GCP) and general encoding profiles for the Geography Markup Language (GML) and the Web Ontology Language (OWL) are based on adapted concepts from ISO/TC 211 standards. Community specific profiles for conceptual models and encodings are based on UML profiles and the use of UML for specific information models in the three application domains. The studies and related research showed that the structure of UML profiles could be implemented and used for information modelling in the UML software Enterprise Architect and that existing profiles and information models could be adapted into the framework. Integration of information models in a common approach based on MDA and UML establishes a fundament for improved interoperability through a shared understanding of the digital representation of the real world.

1. INTRODUCTION

1.1 The Digital Geospatial Environment

The digital representation of the natural and built environment in a geospatial context is fundamental for several application domains. Among these are the application domains of Geographic Information Systems (GIS), Intelligent Transport Systems (ITS), and Building Information Modelling (BIM). As illustrated in Figure 1, the three domains have distinct but related roles in the digital geospatial environment. Applications for GIS are mostly used for handling and analyzing the existing natural and built environment, while applications for BIM are used for planning, developing, constructing and maintaining the built environment. Finally, applications and systems for ITS use information about the built environment for transportation purposes.

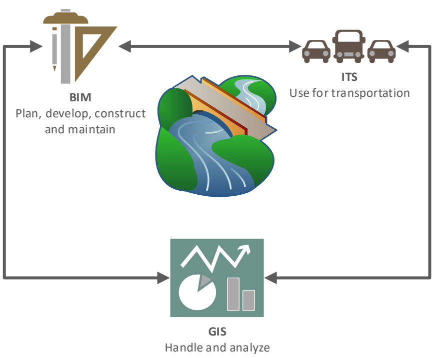


Figure 1. The roles of GIS, ITS and BIM in the digital geospatial environment.

While the roles of GIS, ITS and BIM are distinct and the real world is modelled in different perspectives, many of the real-world features and concepts they handle are the same. Therefore, reuse of information across application domain borders should be possible. For example, the digital representation of a railing along a new road will first come into existence in a BIM project in the planning stage for the road. The feature representation of the railing could later be reused in GIS datasets and High Definition (HD) maps for ITS when the road has been built. Likewise, the existing environment represented in a GIS dataset lays the foundation for new BIM projects. Dynamic data from ITS sensors could be an essential fundament for updating feature information and performing environmental analysis in GIS, and for maintenance planning in BIM.

Reuse of information across application domain borders requires a common understanding of how the real world is described in information models. Stakeholders from GIS, ITS and BIM have developed application-specific information models that describe features and concepts from the natural and built environment in a geospatial context. Information models from all three domains are based on Model-Driven Architecture (MDA) (Object Management Group, 2014) and the Unified Modelling Language (UML) (Object Management Group, 2017). A harmonized use of MDA and UML could play a significant role in a shared understanding of information models across application domain borders.

1.2 Contribution and Research Questions

This study concerns the approaches and technologies used for modelling of geospatial information in the three application domains of GIS, ITS and BIM. We aim to establish a harmonized approach for the use of MDA and UML by investigating two research questions:

1. How can practices for information modelling and semantics for implementation technologies for GIS, ITS and BIM be combined into one common MDA approach with a structure of UML profiles?
2. How can information models based on existing domain-specific technologies be implemented in the common approach?

2. MATERIALS AND METHOD

The fundament for answering the research questions was established through a state-of-the-art analysis on the use of MDA and UML for modelling of geospatial information in the three application domains. The analysis included UML profiles and modelling rules from standardized information models; and relevant research on the topic.

The knowledge gained from the state-of-the-art analysis was the foundation for defining a common structure of UML profiles for the three application domains. Finally, the usability of the structure was tested through implementation and adaption of existing information models.

The UML modelling software Enterprise Architect (EA) (Sparx Systems Pty Ltd, 2020) has been used for developing standardized information models in all three domains. Therefore, we found it relevant to use EA in this study as well, for the development and implementation of UML profiles, and transformation of existing models.

3. STATE OF THE ART

3.1 Model-Driven Architecture

The MDA approach for information modelling provides a methodology for describing conceptual models independent of implementation technology and for deriving implementable models by applying transformations. The conceptual models are defined as Platform-Independent Models (PIM) and are described in a conceptual modelling language – typically UML. Implementable models (e.g., prepared for implementation in XML) are defined as Platform-Specific Models (PSM).

The core concepts for UML are defined as metaclasses in the UML metamodel. Specialized concepts, semantics and restrictions for the use of UML in a specific domain can be formalized in UML profiles through the stereotype mechanism, which defines extensions of UML metaclasses. Stereotypes can have properties for additional semantics – represented as tagged values – and constraints that restrict the concept.

3.2 GIS

Standards developed by ISO/TC 211 define the concepts for using MDA and UML for modelling of geospatial information, as illustrated in Figure 2. ISO 19103 Conceptual Schema Language (CSL) specifies the use of UML for geospatial information, including a formal UML Profile and rules for UML modelling (ISO/TC 211, 2015a). ISO 19109 defines additional rules for application schemas (RAS), and an extension of the UML profile from ISO 19103. Besides, ISO 19109 defines the General Feature Model (GFM) as a metamodel for geospatial information (ISO/TC 211, 2015b). Finally, rules for conversion from UML to implementation schemas are defined for the Geography Markup Language (GML) in ISO 19136 (ISO/TC 211, 2020a), and the Web Ontology Language (OWL) in ISO 19150-2 (ISO/TC 211, 2015c). ISO 19136 also defines additional semantics that is needed for conversion to GML. The semantics are implemented as tagged values but are not specified in a formal UML profile.

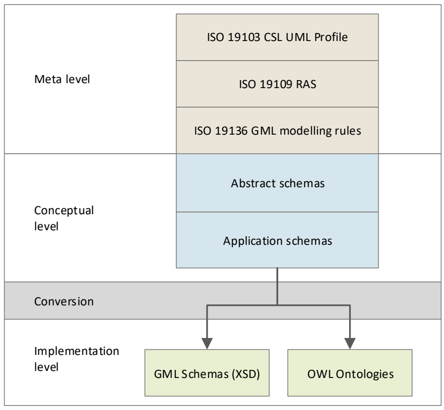


Figure 2. The use of Model-Driven Architecture in GIS.

Authorities and organizations working in the GIS domain, as well as other domains, have developed a wide range of UML models for geospatial information based on the concepts from ISO/TC 211. For example, the European INSPIRE directive defines standard European specifications for 34 different spatial data themes (INSPIRE, 2020). The Open Geospatial Consortium (OGC) has specified a number of standards, among them the CityGML (Open Geospatial Consortium, 2019) and InfraGML (Open Geospatial Consortium, 2017) specifications that are closely related to BIM models. One example from the ITS domain is the ISO 19109 compliant TN-ITS specification (CEN/TC 278, 2018a) for exchange of road-related geospatial information.

3.3 ITS

ITS is an extensive application domain with a wide range of activities and technologies, where geospatial information is vital for many purposes. Standardized information models for ITS in a geospatial context have been developed by ISO/TC 204 and CEN/TC 278, and by consortiums of equipment manufacturers and other stakeholders.

ISO 20524 Geographic Data Files (GDF) defines the primary model for geospatial road-related information used in ITS applications and services (ISO/TC 204, 2019a, b). The GDF information model is described in UML and applies a set of specific stereotypes on model elements. There is no formalized UML profile for the GDF model, but the model is partly based on ISO/TC 211 UML profiles. Joint work between the ISO technical committees for GIS (TC 211) and ITS (TC 204) has studied gaps between GDF and ISO/TC 211 conceptual models (ISO/TC 211, 2020b). A general recommendation from their work is to develop a future version of GDF compliant to ISO/TC 211 UML profiles.

While GDF focuses mainly on static information, two other series of ITS standards define models for dynamic information in a geospatial context: The ISO series ISO 21219 TPEG2 and the CEN European series 16157 DATEX II. Both series define the use of MDA and rules for UML modelling, with specific stereotypes and tagged values and rules for conversion to implementation schemas, as illustrated in Figure 3. Part 2 of

TPEG2 (ISO/TC 204, 2014) defines UML modelling rules – but no formal UML profile, while parts 3 (ISO/TC 204, 2015a) and 4 (ISO/TC 204, 2015b) define conversion rules to binary format and XML schemas. Part 1 of DATEX II defines a formal UML profile and conversion rules to XML schemas (CEN/TC 278, 2018b).

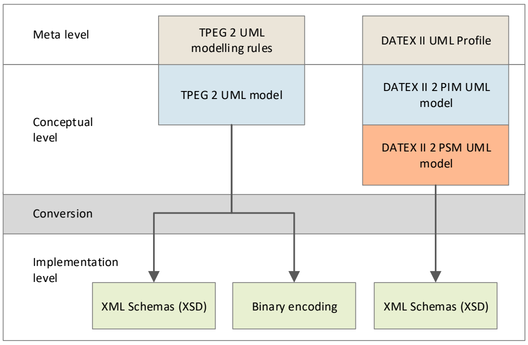


Figure 3. The Model-Driven approaches for TPEG2 and DATEX II.

The European Standard EN 12896 (CEN/TC 278, 2006) defines a reference data model for Public Transport Information (Transmodel). A conceptual model is described in UML, but no specific UML profile is defined (CEN TC278/WG3/SG4 PT0302, 2017). The Transmodel reference model is reused in several more specific models. One example is the Public Transport Network Timetable Exchange model (NeTEx) (CEN/TC 278, 2014). NeTEx applies a model-driven design with a conceptual model (PIM), physical models (PSMs) and implementation schemas – as illustrated in Figure 4.

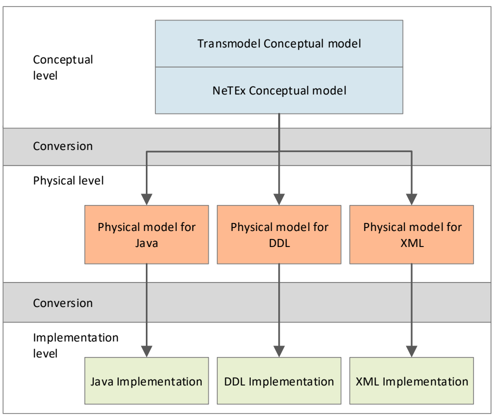


Figure 4. The Model-Driven design in NeTEx. Adapted from (CEN/TC 278, 2014).

3.4 BIM

The core concepts for describing the real world in a geospatial context for use in BIM are defined in the Industry Foundation Classes (IFC) (buildingSmart International, 2019a). IFC defines real-world features, their relations to other features, and their properties – including shapes and positions. The geospatial context and knowledge of the surroundings is vital information for BIM projects – in particular infrastructure projects, which extend over large geographic areas.

The IFC information model is initially described in the EXPRESS modelling language. A representation in UML is under development, and UML is planned to replace EXPRESS as the official modelling language for future versions of IFC (van Berlo, 2019). Implementation schemas for EXPRESS, XML, and OWL will then be derived from the UML model, as illustrated in Figure 5.

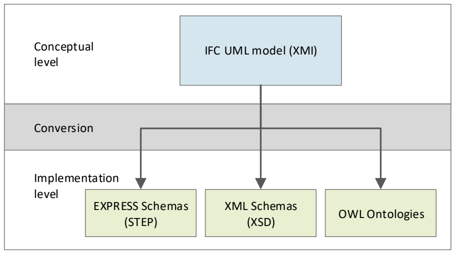


Figure 5. The Model-Driven approach for IFC. Adapted from (van Berlo, 2019).

A draft IFC-UML model has recently been made available (buildingSmart International, 2019b). The model has implemented a set of UML stereotypes and tagged values for the derivation of EXPRESS schemas. However, no official UML profile for IFC is available.

Interoperability between IFC and information models for GIS has been studied by research projects as well as standardization stakeholders over the last years (Zhu et al., 2018, Liu et al., 2017). The ISO technical committees for GIS (TC 211) and BIM (TC 59) have analyzed gaps and the possibilities for harmonization of BIM and GIS standards (ISO/TC 59/SC 13, 2019). One of the recommendations from their work is to link core concepts for IFC with concepts for GIS information models.

3.5 Related research

Kutzner et al. (Kutzner, 2016, Kutzner et al., 2018) presented a significant contribution to the research on UML profiles and model transformation for geospatial information. The studies evaluated the ISO/TC 211 UML profiles and found several deficiencies, and presented a framework with a modular structure of UML profiles. The framework included base and community profiles for platform-independent conceptual models and platform-specific profiles for encoding. Besides, information integration and model-driven transformation were described at distinct levels of abstraction according to the ISO/TC 211 MDA approach.

Jetlund et al. (Jetlund et al., 2019b) suggested that the GDF information model for ITS could be modified to follow ISO/TC 211 UML profiles and then implemented as GML schemas. Only minor modifications were needed for the GDF model. Likewise, Jetlund et al. (Jetlund et al., 2020) demonstrated that the IFC information model for BIM could be transformed from EXPRESS to a UML model compliant with ISO/TC 211 UML profiles. Implementation schemas for the GIS format GML as well as EXPRESS schemas for use in BIM could be derived from the UML model. Some extra semantics for implementation in

EXPRESS were needed in the UML model. Besides, EXPRESS concepts for complex aggregations, constaints and functions needed a more complex transformation.

Jetlund et al. (Jetlund et al., 2019a) also described how transformations from UML models based on ISO/TC 211 profiles to OWL ontologies could be improved by applying extensions to the ISO/TC 211 UML profiles.

Sampaio et al., and Ferreira et al., described the UML profile GeoProfile for conceptual models of geospatial databases (Sampaio et al., 2010, Ferreira et al., 2016). The profile has a high degree of intersection with the ISO/TC 211 profiles, but neither the work by ISO/TC 211 nor OGC is mentioned in the articles. Besides, Ferreira et al. (Ferreira et al., 2016) described transformation at different levels of abstraction, similar to the work by Kutzner et al. (Kutzner, 2016, Kutzner et al., 2018).

4. A STRUCTURE OF UML PROFILES

We propose to establish a structure of formalized UML profiles for modelling of geospatial information in GIS, ITS and BIM, following the framework presented by Kutzner et al. (Kutzner, 2016, Kutzner et al., 2018). The structure is illustrated in Figure 6 for the base and general encoding profiles, and example community-specific profiles for IFC, DATEX II and GDF. The Core Geospatial Profile (CGP) is the root of all profiles.

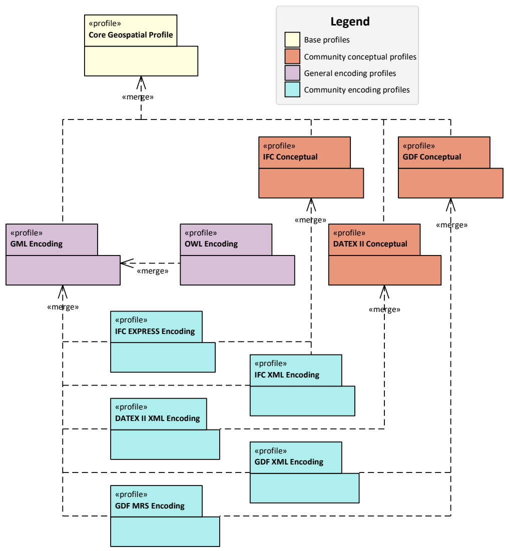


Figure 6. Structure of UML Profiles.

The UML profiles in Figure 6 are related through package merge relations, which merge all concepts from a supplier package to a client package. Concepts that are only defined in the supplier package are added to the client package as-is, while concepts with identical names in the two packages are combined into extended concepts in the client package. For example, all concepts defined in the CGP are merged into the GML Encoding Profile, while all concepts from the GML Encoding Profile are merged into the IFC EXPRESS Encoding Profile. This approach simplifies the modelling and maintenance of profiles: Each profile needs only to define its unique concepts, while more general concepts are merged from supplier profiles.

The CGP contains the core concepts for conceptual models of geospatial information. The profile combines concepts from the profiles in ISO 19103 and ISO 19109, as suggested by Kutzner et al. (Kutzner et al., 2018, Kutzner, 2016). Using concepts only from the ISO 19103 UML profile is relevant for abstract conceptual schemas such as the core ISO/TC 211 standards for geometry (ISO 19107), time (ISO 19108) and reference systems (ISO 19111). However, for modelling of application schemas, concepts from ISO 19103 and ISO 19109 are used in combination. Therefore, a combined profile is more useful as the building-block for all models of geospatial information.

The content of the CGP is shown in Figure 7. We have modified some concepts from ISO 19103 for use in the CGP, according to suggestions by Kutzner et al. (Kutzner, 2016, Kutzner et al., 2018): The CodeList stereotype extends the Enumeration metaclass instead of the DataType metaclass, while the Union stereotype extends the DataType metaclass instead of the Classifier metaclass. Furthermore, the DATEX II UML profile, as well as Jetlund et al. (Jetlund et al., 2019a), describes semantics for defining external concepts and global properties. Jetlund et al. (Jetlund et al., 2019a) suggested these extensions for improved implementation in OWL, but they are also relevant at a PIM level, as well as in other implementation technologies. In particular, reuse of external vocabularies is a good practice that should be considered at an early stage of information modelling (Noy and McGuinness, 2001). Therefore, semantics for unique identification of internal and external concepts are included in the profile through the stereotype ExternalNamespace and the properties URI and vocabulary. Semantics for global properties are included in the profile through the property isGlobal.

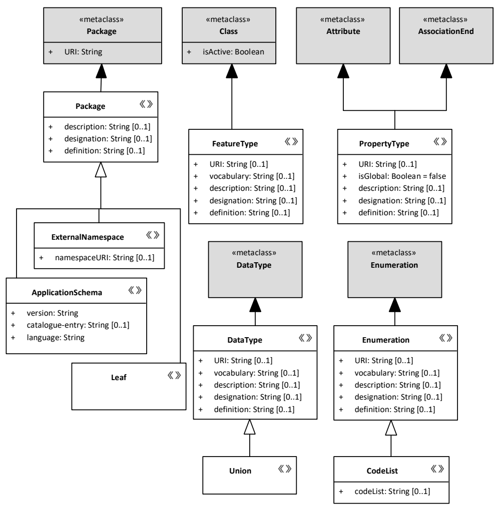


Figure 7. The Core Geospatial Profile – CGP.

The encoding profiles define the semantics needed for conversion from conceptual schemas to implementation formats. We have defined a GML Encoding Profile based on the modelling and conversion rules defined in ISO 19136 as a core encoding profile for geospatial information. GML is the standardized exchange format for geospatial information, and all information models based on the CGP should support implementation in GML – besides implementation in the community-specific technologies.

Furthermore, semantics defined in the GML Encoding profile are also relevant for other implementation technologies. For example, namespace information for packages and sequence number for properties are semantics in the GML Encoding Profile that are relevant for several encodings. Therefore, we have related all other encoding profiles to the GML Encoding Profile through package merge relationships.

Besides GML, we have defined the OWL Encoding Profile to be a general encoding profile, as OWL is the standard implementation technology for the Semantic Web. The OWL encoding from UML models of geospatial information is based on conversion rules defined in ISO 19150-2 with extended rules defined by OGC (Echterhoff et al., 2018, Echterhoff et al., 2017). The conversion rules use existing tagged values defined in the CGP and the GML Encoding Profile. Besides, Jetlund et al. suggested extensions to ISO 19109 for improved OWL encodings (Jetlund et al., 2019a). The suggested semantics for global properties and external vocabularies are added to the CGP, while the semantics for defining ontology name and RDF statements are defined in the OWL Encoding Profile, as shown in Figure 8.

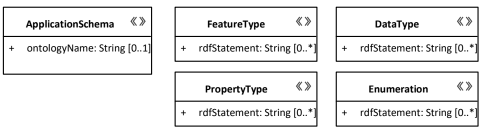


Figure 8. Extension of stereotypes in the OWL Encoding Profile.

The community conceptual UML profiles define concepts and semantics that are relevant only within a specific application domain or for a specific series of models. Likewise, the community encoding profiles define concepts for specific implementation technologies, defined for specific communities. From the findings in the state-of-the-art analysis, possible community-specific profiles for conceptual models and encodings may be needed for IFC, GDF, DATEX II, TPEG2, and Transmodel with possible extensions for NeTEx. The approach for developing community profiles is discussed in Section 7.

5. PROFILE IMPLEMENTATION

Kutzner et al. (Kutzner et al., 2018) pointed out that the concept with profiles related through merge relationships needs to be tested for implementation in UML tools. Therefore, we developed and tested the UML profiles for implementation in EA. The package merge relationship is defined in the UML specification (Object Management Group, 2017) and implemented for use in the design of UML profiles in EA. The profiles can be exported as XML files and then be imported into an EA project where they are applied to UML models.

However, we were not able to maintain the merge relationships when the profiles were exchanged and imported. Only the stereotypes and tagged values defined in each profile were available in an imported profile. Therefore, we developed a script in EA for performing the merge into individual and complete profiles before export to XML. Each complete merged profile could then be imported and applied to models in EA.

Figure 9 shows the extension of stereotypes in the original IFC EXPRESS Encoding Profile and the same stereotypes after being merged with stereotypes from the GML Encoding Profile and the CGP. Figure 10 shows an example of a datatype with semantics both from the CGP and the IFC EXPRESS Encoding Profile.

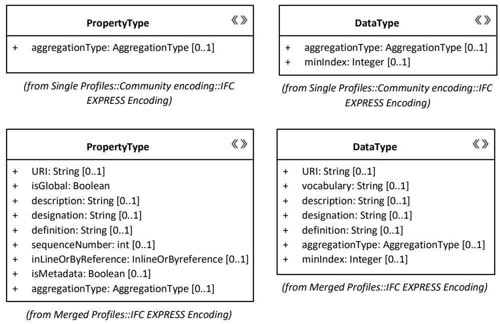


Figure 9. Stereotypes in the original and merged IFC EXPRESS Encoding Profile.

According to the principles in MDA, conceptual models shall be developed as platform-independent (PIM). Semantics for encodings – defined in general and community encoding profiles – are added to PSMs for deriving the specific implementation schemas. This expansion from PIM to PSM can be done by creating individual PSMs – as shown for NeTEx in Figure 4 – or by adding all needed semantics to one PSM. In the latter approach, several encoding profiles must be merged.

Independent of approach, the semantics initially defined in the PIM must be maintained in the PSM. For example, semantics described according to the CGP must be maintained when moving to an IFC EXPRESS Encoding Profile, as illustrated in Figure 10.

Name	IfcComplexNumber
General	
Type	DataType
Stereotype	IFC EXPRESS Encoding::DataType
Alias	
Keywords	
Status	Proposed
Version	4.3
«DataType» (from IFC EXPRESS Encoding)	
URI	IFC4_3.IfcComplexNumber
vocabulary	http://standards.buildingsmart.org/IFC/...
description	The first element (index 1) denotes the r...
designation	
definition	Representation of a complex number ex...
aggregationType	ARRAY
minIndex	1

Figure 10. Semantics for a datatype according to the IFC EXPRESS Encoding Profile with merged semantics from the CGP.

We tested how EA handled semantics when changing from one profile to another, e.g., how the semantics for a FeatureType were handled when extending from the CGP to the GML Encoding Profile. As far as we were able to identify, EA does not maintain the semantics. Therefore, we developed a script for changing from one profile to another, making sure that any tagged values defined for stereotypes in both profiles were maintained.

6. MODEL ADAPTION

Existing information models must be adapted to be compliant with UML profiles in the proposed structure in order to achieve the full potential of the solution. Horizontal adaption can be applied between models at the same level of abstraction, e.g., metamodel to metamodel or conceptual model to conceptual model. Contrary, vertical adaption concerns models at different levels of abstraction.

Horizontal adaption of the GDF and IFC information models to be compliant with ISO/TC 211 UML profiles was demonstrated by Jetlund et al. (Jetlund et al., 2020, 2019b). For our work, we found the adaption of DATEX II information models particularly relevant, as DATEX II has the most formalized community UML profile. If DATEX II models could be made compliant with the framework, they could be implemented in the GML and OWL formats, which would increase the interoperability with other application domains.

Kutzner et al. (Kutzner et al., 2018, Kutzner, 2016) successfully tested the Atlas Transformation Langauge (ATL) for horizontal transformation between UML profiles. ATL is available as an open-source implementation where the transformation is performed on XMI files – the exchange format for UML models. However, ATL is not available in EA, which was our selected tool for implementation. Therefore, we used the scripting facilities in EA for model adaption.

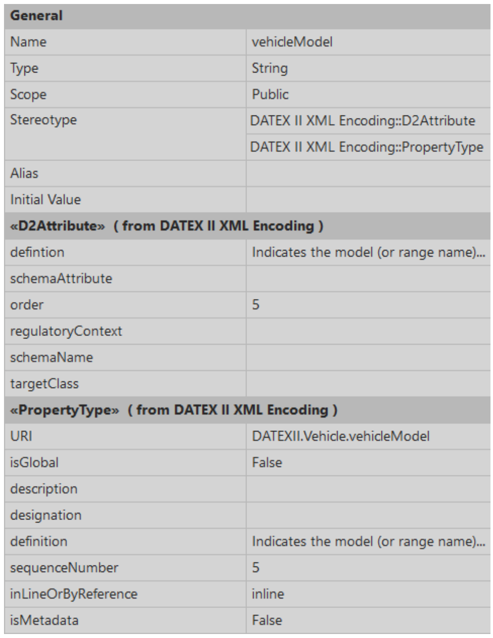
The DATEX II UML Profile is more detailed than the CGP. For example, while CGP extends the metaclass Class with the stereotype FeatureType only, DATEX II has five stereotypes for classes: D2Class, D2Identifiable, D2VersionedIdentifiable, ExternalClass, and D2ModelRoot. Each stereotype has its specific rules for conversion to DATEX II XML implementation schemas. A DATEX II PSM that shall be implemented according to the DATEX II XML conversion rules need to have the DATEX II stereotypes. Therefore, rather than to change the DATEX II stereotypes, stereotypes from the CGP must be added to the DATEX II information model.

Table 1 shows examples of rules for adding stereotypes and semantics from the GCP and the GML Encoding Profile to UML concepts – based on their existing DATEX II stereotype. Semantics that are defined in both the source profile (DATEX II) and the target profile (The CGP or the GML Encoding Profile) are duplicated and stored as semantics according to both profiles.

Figure 11 shows an example attribute from DATEX II with two stereotypes: D2Attribute and PropertType. With the semantics from both stereotypes – and specified rules for conversion to implementation schemas – the model can be implemented in both the DATEX II XML Format and GML.

Source profile	Target profile
Stereotypes: - D2Class, D2Identifiable, D2VersionedIdentifiable, ExternalClass, D2ModelRoot	Stereotype: FeatureType Copy semantics: - definition
Stereotypes: - D2Attribute, D2Literal	Stereotype: - PropertyType Copy semantics: - definition Derive semantics: - sequenceNumber = order
Stereotypes: - D2Datatype, ExternalType	Stereotype: - DataType Copy semantics: - definition, description

Table 1. Examples of rules for mapping from DATEX II to the CGP and the GML Encoding Profile.



General	
Name	vehicleModel
Type	String
Scope	Public
Stereotype	DATEX II XML Encoding::D2Attribute
	DATEX II XML Encoding::PropertyType
Alias	
Initial Value	
«D2Attribute» (from DATEX II XML Encoding)	
defintion	Indicates the model (or range name)...
schemaAttribute	
order	5
regulatoryContext	
schemaName	
targetClass	
«PropertyType» (from DATEX II XML Encoding)	
URI	DATEXII.Vehicle.vehicleModel
isGlobal	False
description	
designation	
definition	Indicates the model (or range name)...
sequenceNumber	5
inLineOrByReference	inline
isMetadata	False

Figure 11. Example of an attribute with stereotypes from the DATEX II profiles and the GCP.

7. DISCUSSION

The state-of-the-art analysis in section 3 showed that models of geospatial information from all three application domains of GIS, ITS and BIM are developed based on UML and model-driven approaches. However, the approaches are specialized for individual application domains and specific series of standards. Furthermore, only a few approaches are based on a formalized use of UML profiles.

Our first research question asked for possibilities for combining the model-driven approaches in a common structure of formalized UML profiles. We defined a structure of UML profiles for GIS, ITS and BIM, based on a framework developed by Kutzner et al. (Kutzner, 2016, Kutzner et al., 2018). The structure includes a Core Geospatial UML Profile (CGP) and general encoding profiles; and more specific community profiles for conceptual models and implementation models. Package merge relations connect the profiles. Table 2 summarises the profiles in the structure and recommended further actions for formalization.

UML Profile	Recommended actions
CGP	Revise the UML profiles in ISO 19103 and ISO 19109.
GML Encoding Profile	Define a formal UML profile in ISO 19136.
OWL Encoding Profile	Define a formal UML profile in ISO 19150-2.
IFC Conceptual Profile	No actions are needed; the CGP can be used.
IFC EXPRESS Encoding Profile	Define a formal profile for encoding in EXPRESS.
GDF Conceptual Profile	No actions are needed; the CGP can be used.
GDF Encoding Profiles	Define formal profiles for encoding in GDF XML and MRS.
DATEX II Conceptual and Encoding Profiles	Define a two-way mapping between concepts in the CGP and the GML Encoding Profile.
TPEG 2 Conceptual and Encoding Profiles	Define formal profiles for the conceptual model and the encodings, based on rules in ISO 21219.
Transmodel and NeTEx	Define a formal profile for specific concepts from the use of UML in existing models.

Table 2. Suggested UML Profiles and recommended actions for formalization.

The CGP presented in Section 4 combines concepts from the core ISO/TC 211 standards ISO 19103 and ISO 19109, with improvements based on related research. We recommend that these improvements are considered for formal ISO/TC 211 UML profiles in revisions of the two standards. Likewise, formalized GML and OWL encoding profiles should be defined in ISO 19136 and ISO 19150-2, as general encoding profiles for all models of geospatial information.

The approach for defining community profiles depends on the maturity and degree of formality – for existing information models as well as rules for modelling and conversion. The formal UML profile and rules defined in DATEX II may be adapted and mapped into the suggested structure. TPEG 2 has a structured set of rules that may be used for defining a profile within the framework. Related research has shown that the conceptual models for IFC and GDF can be modelled according to core ISO/TC 211 profiles, supported by specific encoding profiles for conversion to EXPRESS for IFC; and XML and MRS for GDF. Finally, potential UML Profiles for Transmodel and NeTEx may be defined from the use of UML in the models and representations in implementation schemas.

Our second research question asked how information models based on existing domain-specific technologies could be implemented into a common structure of UML profiles. The results in Section 5 showed that our selected UML application EA could not implement the framework of related profiles directly. However, we were able to perform a merge of the profiles with an internal script in EA and then implement the merged profiles.

The approach for implementing existing models into the common structure will depend on the structure of the original model. Transformations are always concerned with the risk of losing information or expressiveness. Therefore, model adaption by adding more semantics to existing models may be preferred over model transformation. The results in Section 6 showed that model adaption by scripting is possible if the original model is modelled according to a described structure, as was the situation for the DATEX II model. Besides, related research has described how the existing IFC model could be made compliant with ISO/TC 211 profiles through transformation scripts. On the other hand, the GDF model needed more manual modification.

8. CONCLUSIONS

Information models in the application domains of GIS, ITS and BIM describe many of the same real-world features and concepts, but from different views. A common understanding of how the real world is described in the information models is needed to enable reuse of information across application domain borders.

Formalized UML profiles and modelling rules is the fundament for a structured representation of the real world in UML. We developed and tested a structure of UML profiles for modelling of geospatial information in the three application domains and described actions for establishing formal profiles. The results showed that the profiles could be implemented in UML software as complete individual profiles for use in information models. Existing UML profiles and information models from the three application domains could be adapted into the structure.

This study has focused on the core and abstract concepts for information modelling in UML. However, the main advantages of the suggested structure can be achieved at the application schema and data instance level. Transformation and linking between instances that represent real world-features in different views can be defined easier and more accurate when the distinct models are based on a common fundament for information modelling.

ACKNOWLEDGEMENTS

The authors would like to thank Ordnance Survey GB (https://www.ordnancesurvey.co.uk) and 1Spatial (https://1spatial.com/) for sponsoring the publication of this paper.

REFERENCES

buildingSmart International, 2019a. IFC Specifications Database. buildingSMART International, London, UK. https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/ (10. January 2020).

buildingSmart International, 2019b. NextGen-IFC. buildingSMART International, London, UK. https://github.com/buildingSMART/NextGen-IFC (24. January 2020).

CEN TC278/WG3/SG4 PT0302, 2017. Use of UML in Tranmodel. CEN: Brussels, Belgium, http://www.transmodel-cen.eu/.

CEN/TC 278, 2006. EN 12896:2006 Road transport and traffic telematics — Public transport — Reference data model. CEN: Brussels, Belgium.

CEN/TC 278, 2014. CEN/TS 16614-1:2014 Public transport — Network and Timetable Exchange (NeTEx) — Part 1: Public transport network topology exchange format. CEN: Brussels, Belgium.

CEN/TC 278, 2018a. CEN/TS 17268:2018 Intelligent transport systems — ITS spatial data — Data exchange on changes in road attributes. CEN: Brussels, Belgium.

CEN/TC 278, 2018b. EN 16157-1:2018 Intelligent transport systems — DATEX II data exchange specifications for traffic management and information — Part 1: Context and framework. CEN: Brussels, Belgium.

Echterhoff, J, Birkel, P, and Nichols, DL, 2018. OGC Testbed-14: Application Schema-based Ontology Development Engineering Report. Open Geospatial Consortium, Wayland, MA, USA, http://docs.opengeospatial.org/per/18-032r2.html.

Echterhoff, J, Portele, C, Birkel, P, et al., 2017. OGC Testbed-12: ShapeChange Engineering Report. Open Geospatial Consortium, Wayland, MA, USA, http://docs.opengeospatial.org/per/16-020.html.

Ferreira, TB, Stempliuc, SM, and Lisboa Filho, J, 2016. Interoperability of Conceptual Geographical Data Schemas using UML Infrastructure. *SBBD*.

INSPIRE, 2020. INSPIRE Data Specifications. https://inspire.ec.europa.eu/data-specifications/2892 (31 March 2020).

ISO/TC 59/SC 13, 2019. ISO/WD TR 23262 GIS (Geospatial) / BIM interoperability. ISO: Geneva, Switzerland.

ISO/TC 204, 2014. ISO/TS 21219-2:2014 Intelligent transport systems — Traffic and travel information (TTI) via transport protocol experts group, generation 2 (TPEG2) — Part 2: UML modelling rules. ISO: Geneva, Switzerland.

ISO/TC 204, 2015a. ISO/TS 21219-3:2015 Intelligent transport systems — Traffic and travel information (TTI) via transport protocol experts group, generation 2 (TPEG2) — Part 3: UML to binary conversion rules. ISO: Geneva, Switzerland.

ISO/TC 204, 2015b. ISO/TS 21219-4:2015 Intelligent transport systems — Traffic and travel information (TTI) via transport protocol experts group, generation 2 (TPEG2) — Part 4: UML to XML conversion rules. ISO: Geneva, Switzerland.

ISO/TC 204, 2019a. ISO/DIS 20524-1 Intelligent Transport Systems — Geographic Data Files (GDF) GDF5.1 — Part 1 : Application independent map data shared between multiple sources. ISO: Geneva, Switzerland.

ISO/TC 204, 2019b. ISO/DIS 20524-2 Intelligent transport systems — Geographic Data Files (GDF) — GDF5.1 — Part 2 : Map data used in automated driving systems , Cooperative ITS , and multi-modal transport. ISO: Geneva, Switzerland.

ISO/TC 211, 2015a. ISO 19103:2015 Geographic Information — Conceptual schema language. ISO: Geneva, Switzerland.

ISO/TC 211, 2015b. ISO 19109:2015 Geographic information — Rules for application schema. ISO: Geneva, Switzerland.

ISO/TC 211, 2015c. ISO 19150-2:2015 Geographic information — Ontology — Part 2: Rules for developing ontologies in the Web Ontology Language (OWL). ISO: Geneva, Switzerland.

ISO/TC 211, 2020a. ISO 19136-1:2020 Geographic information — Geography Markup Language (GML) — Part 1: Fundamentals. ISO: Geneva, Switzerland.

ISO/TC 211, 2020b. ISO/PDTR 19169 Geographic Information — Gap-analysis: To map and describe the differences between the current GDF and ISO/TC211 conceptual models to suggest ways harmonize and resolve conflicting issues. ISO: Geneva, Switzerland.

Jetlund, K, Onstein, E, and Huang, L, 2019a. Adapted Rules for UML Modelling of Geospatial Information for Model-Driven Implementation as OWL Ontologies. *ISPRS International Journal of Geo-Information*, 8 (9):365.

Jetlund, K, Onstein, E, and Huang, L, 2019b. Information Exchange between GIS and Geospatial ITS Databases Based on a Generic Model. *Isprs International Journal of Geo-Information*, 8 (3):141.

Jetlund, K, Onstein, E, and Huang, L, 2020. IFC Schemas in ISO/TC 211 compliant UML for improved interoperability between BIM and GIS. *ISPRS International Journal of Geo-Information*, 9 (4).

Kutzner, T, 2016. Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles. PhD, Technische Universität München.

Kutzner, T, Donaubauer, A, and Kolbe, TH, 2018. UML profiles and their use in data modelling and model-driven transformation of geospatial data. *gis.Science* (4).

Liu, X, Wang, X, Wright, G, et al., 2017. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 6 (2):53.

Noy, NF, and McGuinness, DL, 2001. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory, Stanford University, CA, USA. https://protegewiki.stanford.edu/wiki/Ontology101 (21 August 2019).

Object Management Group, 2014. Model Driven Architecture (MDA) Guide rev. 2.0. Object Management Group: Needham, MA, USA, https://www.omg.org/mda/.

Object Management Group, 2017. Unified Modelling Language Specification Version 2.5.1. Object Management Group: Needham, MA, USA, https://www.omg.org/spec/UML.

Open Geospatial Consortium, 2017. OGC LandInfra / InfraGML. Open Geospatial Consortium, Wayland, MA, USA. https://www.opengeospatial.org/standards/infragml (18. February 2020).

Open Geospatial Consortium, 2019. OGC CityGML. Open Geospatial Consortium, Wayland, MA, USA. https://www.opengeospatial.org/standards/citygml (18. February 2020).

Sampaio, GB, Nalon, FR, and Lisboa-Filho, J, 2010. UML Profile for Conceptual Modeling of Geographic Databases. *Proceedings of the 12th International Conference on Enterprise Information Systems - Information Systems Analysis and Specification*.

Sparx Systems Pty Ltd, 2020. Enterprise Architect version 15. https://sparxsystems.com/products/ea/index.html (10. January 2020).

van Berlo, L. 2019. "Future technology Presentation buildingSMART Summit Beijing." 8. November 2019. https://www.linkedin.com/pulse/future-technology-presentation-buildingsmart-summit-l%C3%A9on-van-berlo/ (18. February 2020).

Zhu, J, Wright, G, Wang, J, et al., 2018. A Critical Review of the Integration of Geographic Information System and Building Information Modelling at the Data Level. *ISPRS International Journal of Geo-Information*, 7:66.

ISBN 978-82-326-6601-0 (printed ver.)
ISBN 978-82-326-5703-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology