

A structure of UML profiles for modelling of geospatial information in GIS, ITS and BIM

Knut Jetlund^{1,2}

¹Norwegian University of Science and Technology, Gjøvik, Norway

²The 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.

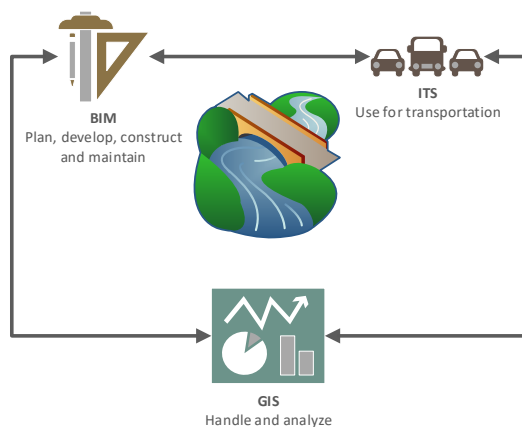


Figure 1. The roles of GIS, ITS and BIM in the digital geospatial environment.

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.

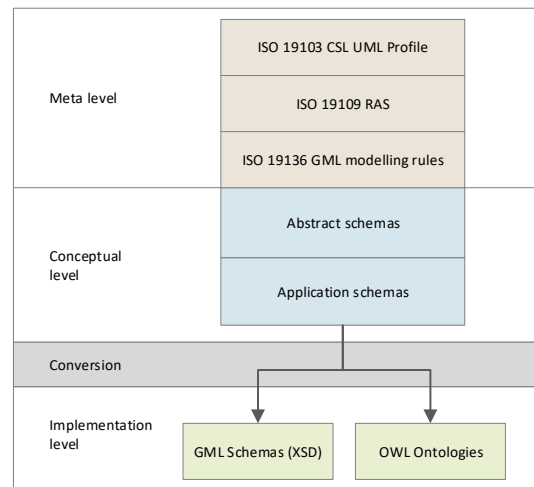


Figure 2. The use of Model-Driven Architecture in GIS.

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).

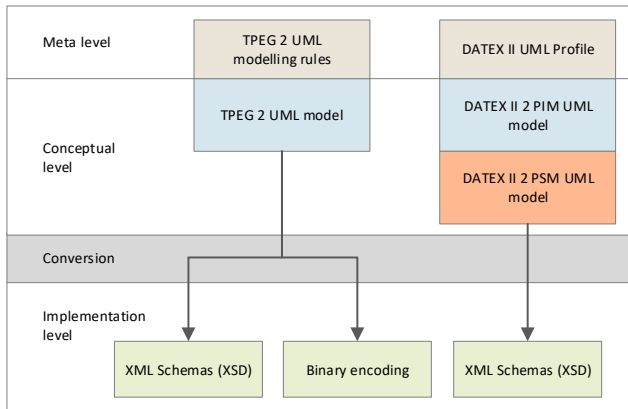


Figure 3. The Model-Driven approaches for TPEG2 and DATEX II.

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.

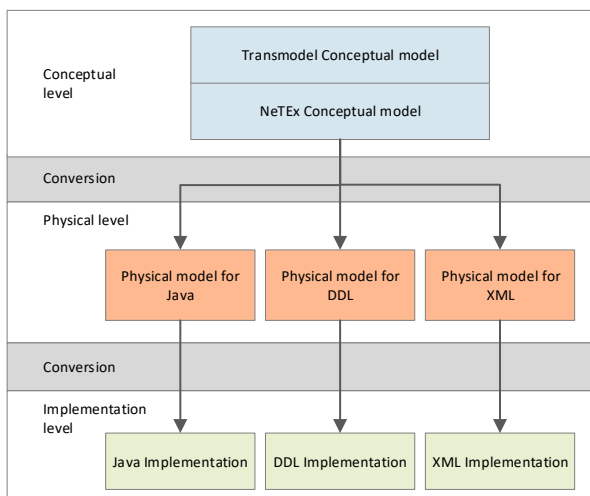


Figure 4. The Model-Driven design in NeTEx. Adapted from (CEN/TC 278, 2014).

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.

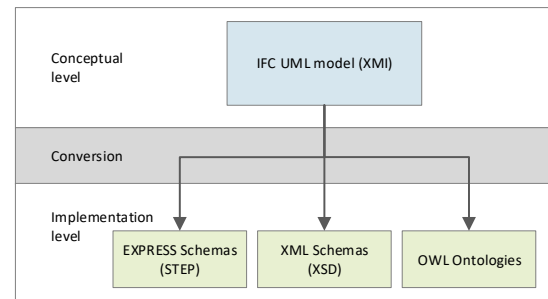


Figure 5. The Model-Driven approach for IFC. Adapted from (van Berlo, 2019).

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, constraints 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.

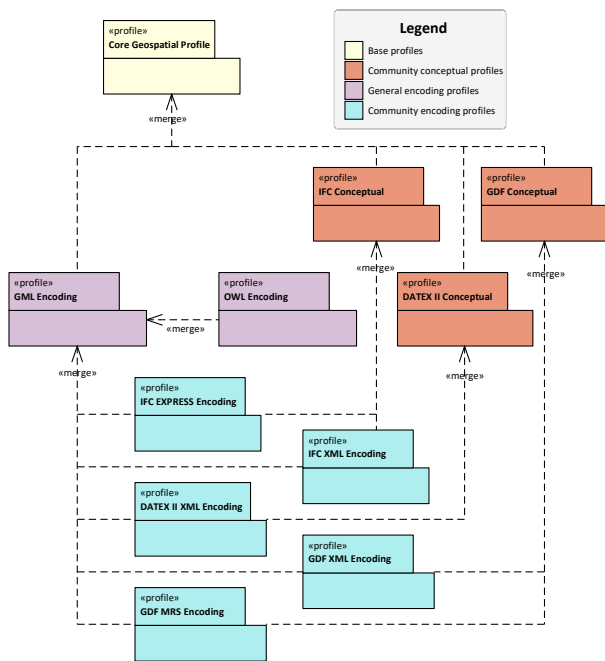


Figure 6. Structure of UML Profiles.

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.

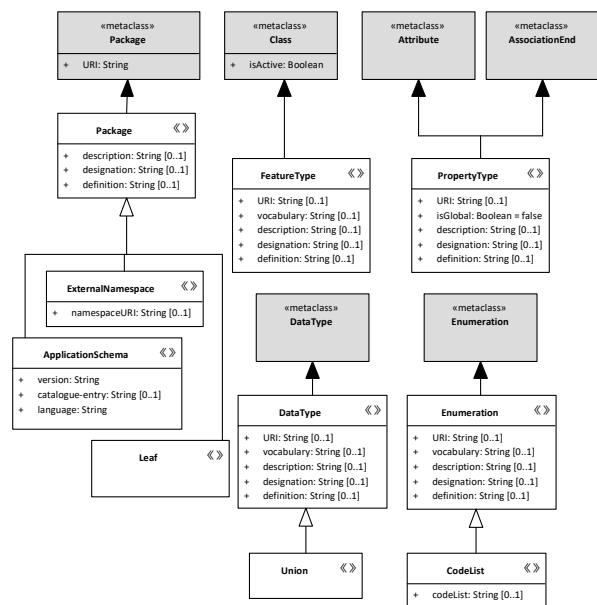


Figure 7. The Core Geospatial Profile – CGP.

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.

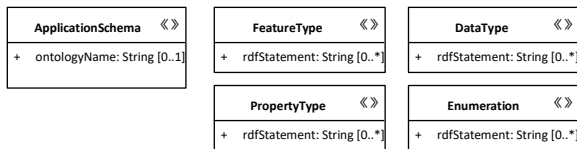


Figure 8. Extension of stereotypes in the OWL Encoding Profile.

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.

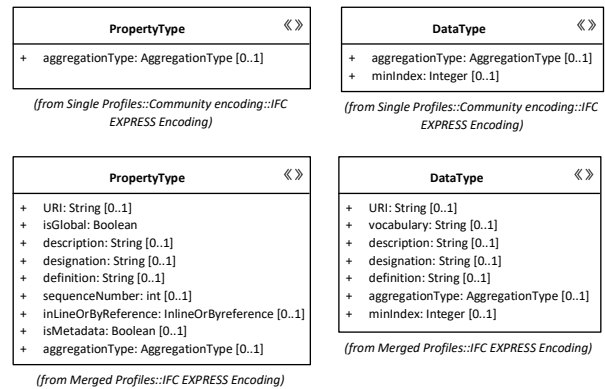


Figure 9. Stereotypes in the original and merged IFC EXPRESS Encoding Profile.

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.IfComplexNumber
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 Language (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.

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.

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.

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 ISpatial (<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, Stempluc, 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, Donaubaauer, 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-1%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.