Interoperability of Traffic Infrastructure Planning and Geospatial Information Systems

Interoperabilität zwischen Verkehrsinfrastrukturplanung und Geoinformationssystemen

Nazereh Nejatbakhsh Esfahani

Schriftenreihe des Instituts für Bauinformatik
Herausgeber: Prof. Dr.-Ing. Raimar J. Scherer

ISBN: 978-3-86780-574-2

Institut für Bauinformatik, TU Dresden

Postanschrift:
Technische Universität Dresden
01062 Dresden

Besucheranschrift:
Nürnberger Str. 31a
2. OG, Raum Nr. 204
01187 Dresden

Tel.: +49 351 / 463-32966
Fax: +49 351 / 463-33975
E-Mail: Raimar.Scherer@tu-dresden.de
WWW: http://cib.bau.tu-dresden.de

Interoperability of Traffic Infrastructure Planning and Geospatial Information Systems

Interoperabilität zwischen Verkehrsinfrastrukturplanung und Geoinformationssystemen

An der Fakultät Bauingenieurwesen der Technischen Universität Dresden als

DISSERTATION

von

Nazereh Nejatbakhsh Esfahani

geboren am 08. September 1981 im Iran

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigt.

Gutachter:

Prof. Dr.-Ing. Raimar J. Scherer, Technische Universität Dresden
Prof. Dr.-Ing. André Borrmann, Technische Universität München
Univ.-Prof. Dr.-Ing. Jörg Blankenbach, Technische Universität Aachen

Tag der Einreichung: 11.07.2016
Tag der Verteidigung: 24.11.2017

Preface

To God, for His great and unwavering love and guidance, without which, I would have given up on life's challenges without any struggle.

Then, my deepest senses of gratitude and extreme thanks are due to my supervisor Prof. Dr. Scherer for his continuous encouragement, valuable ideas and lengthy discussions.

It goes without saying that this thesis would have not reached this level if it were not for the immense efforts and time spent throughout the researches at OBERMEYER Consulting company, direct support of the chief executive of the company Mr. Dipl. –Chem., Dipl. –Ing. Maximilian Grauvogl and the company sponsorship.

Many thanks are due to Mr. Dipl.-Ing. Rolf Uhrig for his valuable time and support throughout the concept phases.

I am indebted to my advisors Dr. Sebastian Fuchs, Mr. Dipl. -Ing. Matthias Frei and Dr. Robert Balder, not only for their generous and insightful advises, help and supervision during this thesis, but also for providing data and information necessary for this research's case study.

I take this opportunity to express my profound gratitude to my beloved family. I cannot help but feel immense gratitude towards them for encouraging and supporting me in all means possible throughout the steps of my life.

Lastly yet importantly, I owe a lot of thanks to the love of my life, my husband, for his extra patience and continuous encouragements.

Dresden, Summer 2016

Nazereh Nejatbakhsh Esfahani

Abstract

Building Information Modelling (BIM) as a *Model-based design* facilitates to investigate multiple solutions in the infrastructure planning process. The most important reason for implementing model-based design is to help designers and to increase communication between different design parties. It decentralizes and coordinates team collaboration and facilitates faster and lossless project data exchange and management across extended teams and external partners in project lifecycle.

Infrastructure are fundamental facilities, services, and installations needed for the functioning of a community or society, such as transportation, roads, communication systems, water and power networks, as well as power plants. **Geospatial Information Systems (GIS)** as the digital representation of the world are systems for maintaining, managing, modelling, analyzing, and visualizing of the world data including infrastructure. High level infrastructure suits mostly facilitate to analyze the infrastructure design based on the international or user defined standards. Called *regulation[1]-based design*, this minimizes errors, reduces costly design conflicts, increases time savings and provides consistent project quality, yet mostly in standalone solutions.

Tasks of infrastructure usually require both model based and regulation based design packages. Infrastructure tasks deal with cross-domain information. However, the corresponding data is split in several domain models. Besides infrastructure projects demand a lot of decision makings on governmental as well as on private level considering different data models. Therefore lossless flow of project data as well as documents like regulations across project team, stakeholders, governmental and private level is highly important. Yet infrastructure projects have largely been absent from product modelling discourses for a long time. Thus, as will be explained in chapter 2 interoperability is needed in infrastructure processes.

Multimodel (MM) is one of the interoperability methods which enable heterogeneous data models from various domains get bundled together into a container keeping their original format. Existing interoperability methods including existing MM solutions can't satisfactorily fulfill the typical demands of infrastructure information processes like dynamic data resources and a huge amount of inter model relations. Therefore chapter 3 concept of infrastructure information modelling investigates a method for loose and rule based coupling of exchangeable heterogeneous information spaces. This hypothesis is an extension for the existing MM to a rule-based Multimodel named extended Multimodel (eMM) with semantic rules – instead of static links. The semantic rules will be used to describe relations between data elements of various models dynamically in a link-database.

Most of the confusion about geospatial data models arises from their diversity. In some of these data models spatial IDs are the basic identities of entities and in some other data models there are no IDs. That is why in the geospatial data, data structure is more important than data models. There are always spatial indexes that enable accessing to the geodata. The most important unification of data models involved in infrastructure projects is the spatiality. Explained in chapter 4 the method of infrastructure information modelling for interoperation in spatial domains generate interlinks through spatial identity of entities. Match finding through spatial links enables any kind of data models sharing spatial property get interlinked. Through such spatial links each entity receives the spatial information from other data models which is related to the target entity due to sharing equivalent spatial index. This information will be the virtual properties for the object. The thesis uses Nearest Neighborhood algorithm for spatial match finding and performs filtering and refining approaches. For the abstraction of the spatial matching results hierarchical filtering techniques are used for refining the virtual properties. These approaches focus on two main application areas which are product model and Level of Detail (LoD).

For the eMM suggested in this thesis a rule based interoperability method between arbitrary data models of spatial domain has been developed. The implementation of this method enables transaction of data in spatial domains run loss less. The system architecture and the implementation which has been applied on the case study of this thesis namely infrastructure and geospatial data models are described in chapter 5.

Achieving afore mentioned aims results in reducing the whole project lifecycle costs, increasing reliability of the comprehensive fundamental information, and consequently in independent, cost-effective, aesthetically pleasing, and environmentally sensitive infrastructure design.

---

[1] Rule-based design in infrastructure, to avoid confusion with rule as semantic, rule is replaced with regulation

Kurzfassung

Building Information Modelling (BIM) als *Model-Based Design* ermöglicht vielfältige Lösungen im Infrastrukturplanungsprozess zu untersuchen. Der wichtigste Grund für die Implementierung des *Model-Based Design* ist Planern zu helfen, und die Kommunikation zwischen verschiedenen Planungspartnern zu erhöhen. Es dezentralisiert und koordiniert die Kollaboration und ermöglicht einen schnelleren und verlustfreien Projektdatenaustausch und Datenmanagement zwischen erweiterten Teams und externen Partnern im Projektlebenszyklus.

Infrastruktur sind grundlegende Einrichtungen, Dienste und Installationen, die für das Funktionieren einer Gesellschaft erforderlich sind, wie Verkehr, Straßen, Kommunikationssysteme , Wasser- und Stromnetze sowie Kraftwerke. **Geospatial Information Systems (GIS)** als die digitale Repräsentation der Welt sind Systeme für die Wartung, Verwaltung, Modellierung, Analyse und Visualisierung der Weltdaten, einschließlich der Infrastruktur. Hochrangige Infrastrukturprogramme ermöglichen vor allem die Infrastrukturplanung, auf Grundlage der internationalen oder benutzerdefinierten Standards zu analysieren. *Regulation-Based Design*[2], minimiert Fehler, reduziert teure Planungskonflikte, erhöht die Zeitersparnis und sorgt für konsistente Projektqualität, dennoch meist als Einzellösungen.

Infrastrukturaufgaben erfordern in der Regel sowohl die *Model-Based* als auch die *Regulation-Based* Design-Pakete. Infrastrukturaufgaben befassen sich mit domänenübergreifenden Informationen, die in mehrere Domänenmodelle aufgeteilt sind. Darüberhinaus verlangen die Infrastrukturprojekte unter Berücksichtigung diverser Datenmodelle viele Entscheidungsfindungen auf staatlicher sowie privater Ebene. Daher ist ein verlustfreier Fluss von Daten und Dokumenten wie z.B. Vorschriften quer durch das Projektteam, den Beteiligten und den staatlichen und privaten Sektoren sehr wichtig. Dennoch war Produktmodellierung bei Infrastrukturprojekten für eine lange Zeit kein Kernthema. Deshalb wird wie in Kapitel 2 erläutert, Interoperabilität in Infrastrukturprozesse benötigt.

Multimodell (MM) ist eine Interoperabilitätsmethode, die die Bündelung heterogener Datenmodelle aus verschiedenen Domänen unter Beibehaltung ihrer ursprünglichen Formate in einem Container ermöglicht. Bestehende Interoperabilitätsmethoden einschließlich der bestehenden MM Lösung können die typischen Anforderungen der Infrastrukturinformationsprozesse wie dynamische Datenressourcen und große Menge an modelübergreifenden Beziehungen nicht zufriedenstellend erfüllen. Deshalb wird im Kapitel 3 Konzept der Infrastrukturinformationsmodellierung ein Verfahren zur losen und regelbasierten Kopplung von austauschbaren heterogenen Informationsräumen untersucht. Diese Hypothese ist eine Erweiterung für das bestehende MM zu einem regelbasierten Multimodell namens erweitertes Multimodell (eMM) mit semantischen Regeln anstelle von statischen Links. Die semantischen Regeln werden verwendet, um die Beziehungen zwischen Datenelemente von diversen Modellen in einer Link-Datenbank zu beschreiben.

Die meiste Verwirrung über Geospatial Datenmodelle ergibt sich aus ihrer Vielfalt. In einigen dieser Datenmodelle sind räumliche IDs die grundlegenden Identitäten von Entitäten und in einigen anderen Datenmodellen gibt es keine IDs. Deshalb ist in Geodaten die Datenstruktur wichtiger als das Datenmodel. Es gibt immer räumliche Indizes, die den Zugriff auf die Geodaten ermöglichen. Die wichtigste Vereinheitlichung der in Infrastrukturprojekten involvierten Datenmodelle ist die Räumlichkeit. Wie in Kapitel 4 beschrieben generiert die Infrastrukturinformationsmodellierung für die Interoperation der räumlichen Domänen die Links durch räumliche Identität von Entitäten. Die Matchfindung durch räumliche Links ermöglicht die Verknüpfung aller Datenmodelle mit räumlicher Eigenschaft. Durch einen räumlichen Link erhält jede Entität der Zieldomäne die räumlichen Informationen aus anderen Datenmodellen, die einen äquivalenten räumlichen Index beinhalten. Diese Informationen sind die virtuellen Eigenschaften des Objektes. Diese Dissertation verwendet den Nächste-Nachbar-Algorithmus für die räumliche Matchfindung, und führt Filterungs- und Verfeinerungsansätze durch. Für die Abstraktion der Ergebnisse der räumlichen Matchfindung wird zur Verfeinerung der virtuellen Eigenschaften hierarchische Filtertechniken verwendet. Diese Ansätze beziehen sich auf zwei Hauptanwendungsgebiete, diese sind das Produktmodell und der Detaillierungsgrad (LoD).

---

[2] Rule-Based Design in Infrastruktur, um die Verwechslung mit der Regel als Semantik zu vermeiden, wurde "rule" durch „regulation" ersetzt.

Für das in dieser Dissertation vorgestellte eMM wurde ein regelbasiertes Interoperabilitätsverfahren zwischen beliebigen Datenmodellen von räumlichen Domänen entwickelt. Die Implementierung dieses Verfahrens ermöglicht einen verlustfreien Informationsfluss in räumlichen Domänen. Die Systemarchitektur und die Implementierung, die in der Fallstudie dieser Arbeit, nämlich Infrastruktur und GIS-Datenmodelle, angewendet wurden, sind in Kapitel 5 beschrieben.

Das Erreichen oben genannter Ziele ergibt die Reduzierung der Kosten während des gesamten Projektlebenszykluses, die Erhöhung der Nachhaltigkeit der umfassenden Grundlageinformationen und somit eine eigenständige, kosteneffektive, ästhetisch und ökologisch sensible Infrastrukturplanung.

Keywords

Model Based Design, Rule Based Design, Regulation Based Design, Infrastructure Design, Intelligent Infrastructure Design, Product Model, Collaborative Model, Dynamic Data Models, Building Information Modelling, BIM, IFC©, BuildingSMART©, OKSTRA©, LandXML©, ProVI©, Geospatial Information Systems, ArcGIS©, CityGML©, OGC©, OpenINFRA, InfraFinBim, Nearest Neighborhood, Level of Details, Interoperability, Multimodel, Rule Based Multimodel, Extended Multimodel, Rule Based Link Model, Semantic, Spatial Links, Spatial Link Model

Table of Content

List of Figures

List of Tables

Chapter 1

Introduction

Since the introduction of Computer-Aided Design (CAD) in the planning and design world some decades ago object-based modeling has been used by the computer products. Most CAD vendors have launched more powerful object-based design software in recent years. This category of software is now commonly known as model-based design or BIM.

Using object based CAD programs Intelligent Infrastructure Design (IID) attempts to establish a connection between two Points of Interest (PoI) through an economical path. For this purpose IID considers different design models, among them the cost models and the alignment optimization models. Jha et al. have discussed in their book (2006) the impact of different optimization methods, especially the revolutionary role of GIS Solutions in infrastructure design models. As construction of roads or railways is great interference into the environment, geospatial data are very important in separate phases of the process.

GIS is developed in feature-based method. The term "feature" encompasses the geographical entity as well as its object representation (NCDCDS, 1988). In GIS Objects are a direct representation of the geographic features rather than of geometric elements such as point, line, and area.

This chapter is an introduction to the overlaps of BIM, Infrastructure and GIS in the context of Infrastructure Projects.

1.1 A GENERAL VIEW

In the Infrastructure design the alignment problem is sufficiently complex to pose substantial research questions, but simple enough to be tractable (Frank & Wallace, 1995). This is a permanent inevitable procedure of data transaction which is performed unnoticeably manually and ends to a long feedback operation which can be speeded up if modeled and embedded in the alignment planning package. Smart technologies, smart heating systems, smart cars and smart highways™© (Studio Roosegaarde, 2011) with the goal to make infrastructure more sustainable and interactive by using light, energy and road signs that automatically adapt to the traffic situation and weather condition are today's and future challenge of infrastructure projects. Figure 1 demonstrates how such smart infrastructure should communicate with the world around it.

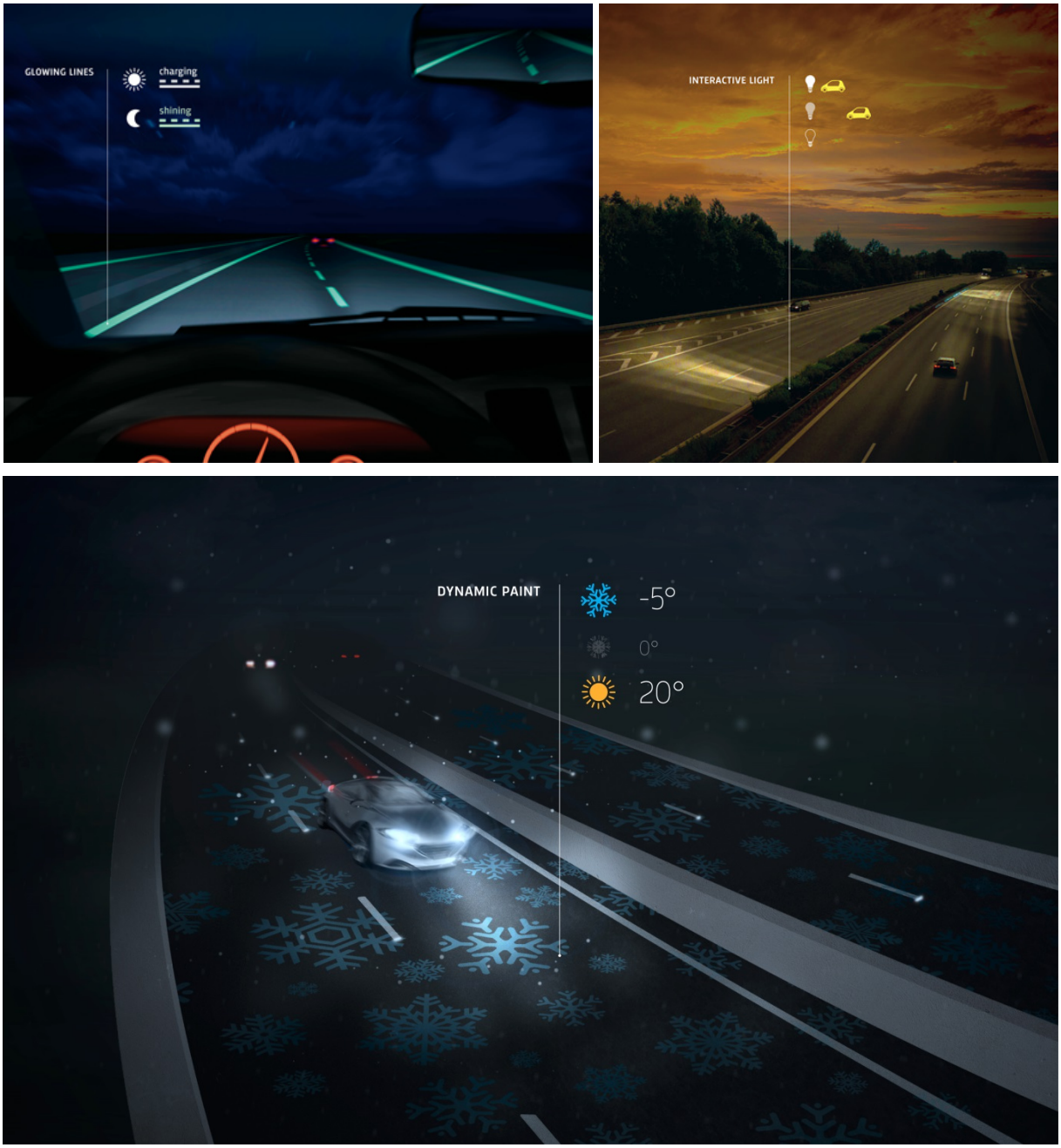


Figure 1 Smart Highway and the World, (Studio Roosegaarde, 2011)

Besides, shown in figure 2 there are many attempts since 2009 to use solar panels™ instead of asphalt. The solar panels *"would generate electricity, which would in turn be fed into the grid. This makes oil conserve twice: Electric cars could be charged with the energy produced by the panels, and the panels*

would replace the use of asphalt, the production of which requires petroleum. Moreover, Solar Roadways™ are heated and equipped with integrated LED screens, which act not only as street markings, but also show warnings directly on the road[3]*"* (Brusaws, 2009).

As the construction management as well as maintenance of the infrastructure projects involves many stakeholders, proper information sharing over the entire lifecycle of the project among the stakeholders is very important. Therefore Open Geospatial Consortium (OGC) has developed geospatial information systems and standards that promote efficient maintenance, management, analysis and visualization of data models of the project and information of the surrounding world.

Traditionally high level infrastructure design suites are regulation based. These regulations are introduced to the high level software packages in form of standard catalogues which restrict the design process. In addition to the entire design these regulations affect the geometry of the project as well. Such design packages use also object-based models, yet the methods are regulation based. This is mentionable that BIM has also regulations or rules; however the methods and processes are object based.

Yet, engineering design is not the only application field for GIS in the infrastructure domain. Using Location Based Services (LBS), smart systems integrate heterogeneous location based information from different service providers, run task based analysis and issue results.

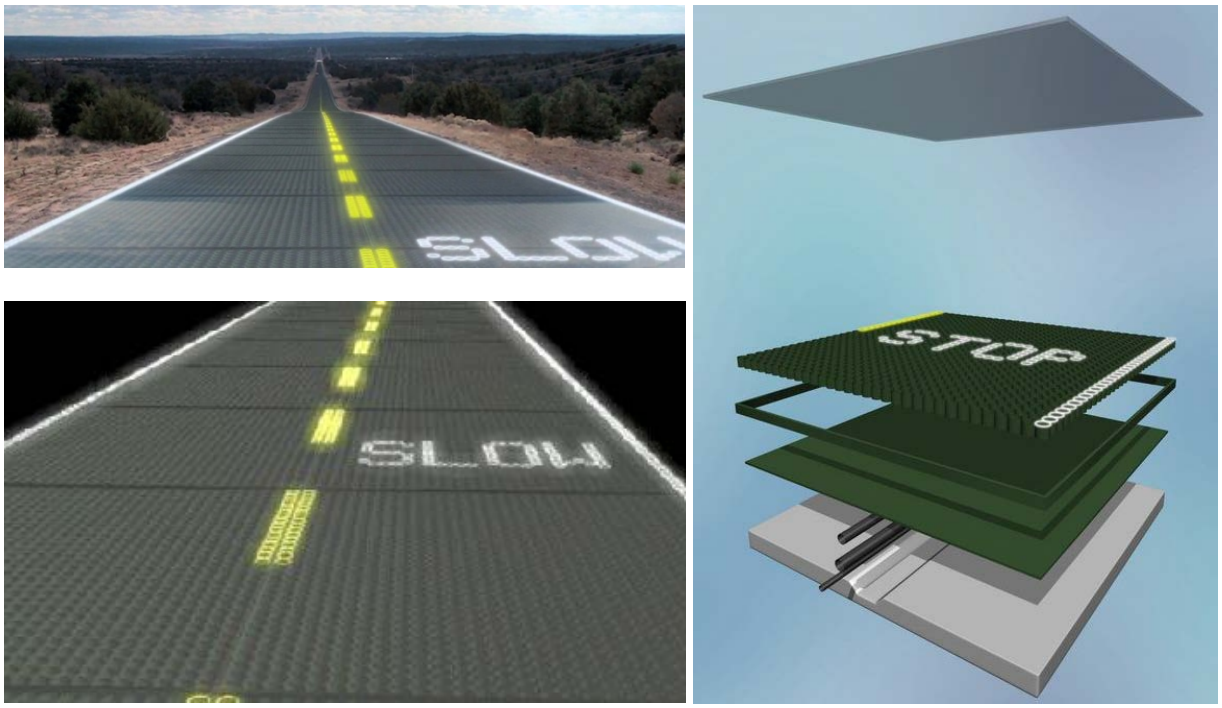


Figure 2 Solar Roadways, (Brusaws, 2009)

1.2 PROBLEM STATEMENT

"Infrastructure development projects are critical in the advancement of local communities' economic development agenda. They reenergize the local and regional economy and create opportunities for new jobs and sustainable economic growth" (Abukhater, 2013). Both geospatial technologies and BIM solutions "*are used to support these developments on various scales and in all stages, including planning, designing, building and management. Geospatial is used in analysis and planning tasks while BIM is used to support project design and operations"* (Jha et al., 2007).

"However, planning and implementation of infrastructure projects can take many years which depletes valuable resources and can result in significant delays and costly deadlocks. This is especially true as communities struggle to engage and involve all stakeholders in the development

[3] https://natgrp.wordpress.com/2013/09/23/solar-road-panels-offer-asphalt-alternative/?share=stumbleupon

process due to the long and sometimes tedious permitting and approval process and the incompatibility of data and technology stacks used in various stages of the aforementioned project lifecycle" (Abukhater, 2013).

As shown in figure 3 "*when planning a large development, community planners tend to rely on two different software applications – GIS technology to analyze and evaluate development scenarios and BIM technology to visualize outcomes and evaluate design scenarios. Although the designers need to collaborate to ensure the success of the planned project, in reality these two workflows are for the most part separate. This causes them to be unable to communicate since they lack interoperability, bringing the development to a crawl"* (Abukhater, 2013).

"Database connectivity, features system, navigation, query, render options, hierarchical and multi-resolution database representation of GIS with its 3D graphics and efficient new algorithms for terrain analysis gives the techniques which can be combined with fast handling methods for large amounts of data to make integrated GIS visualization systems" (Depriton et al., 2008). GIS can be used almost in any field of planning process. Planners for new infrastructure facilities can benefit from a full 3D view from any prospective or from nearby existing landmarks with the planned facility in place. Another *"dimension of GIS "spatially enabling business and analytic intelligence", extends existing concepts of enterprise, integration and interoperability, and creates new value for planers by leveraging the power of place and analytics in support of fact-based decision making, planning, and operation"* (Holland, 2005; Pick, 2008). Notwithstanding its universality of purpose and function, there has been little deep integration of GIS with infrastructure data models used by planers even though many design programs rely on the same or very similar data models.

The OGC defines interoperability as *the capability to communicate, execute programs, or transfer data among various functional units in a way that requires the user to have little or no knowledge of the unique characteristics of those units* (ISO 2382-1). Key interoperability issues include how a given user-defined schema relates to standardized geospatial concepts (e.g. feature, reference system, geometry, etc.), how schemas are represented, and how different user-defined schemas relate to one another. "*Lacking in interoperability is a major problem when dealing with massive infrastructure development projects that cost the community a significant amount of resources and lost opportunities"* (Abukhater, 2013).

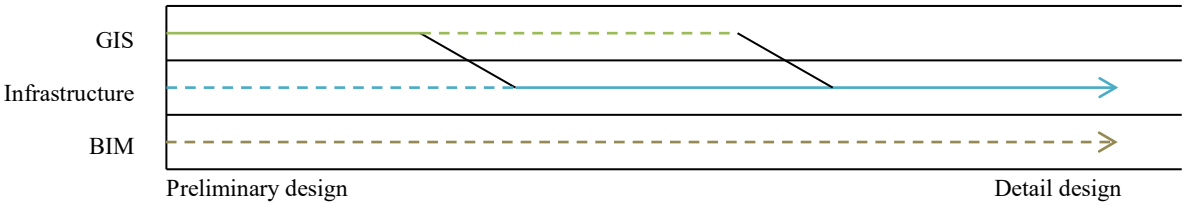


Figure 3 GIS, CAD, and BIM interaction in alignment planning

"Expediting the process of development means a higher degree of data interoperability between the two sides of the project lifecycle by stitching GIS and BIM workflows. This is where GIS meets BIM, where analysis meets design and closes the last ring of the project lifecycle chain as shown in figure 4. The result is a robust set of capabilities that are interoperable and complimentary to each other to support a more informed and streamlined planning, decision making and development process" (Abukhater, 2013).

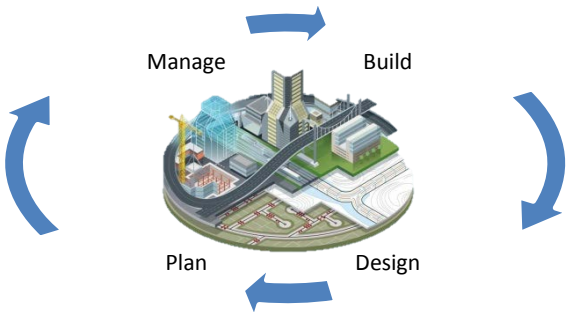


Figure 4 GIS meets BIM, (Abukhater, 2013)

Infrastructure domain is exactly where BIM and GIS overlap strongly. *"BIM allows digital exploration of project's physical and functional characteristics, before the construction. BIM is not just about buildings; it is information modeling for the built environment. As such, BIM encompasses CAD and GIS disciplines by combining model-based design with information and analysis"* (Niemiec, 2010).

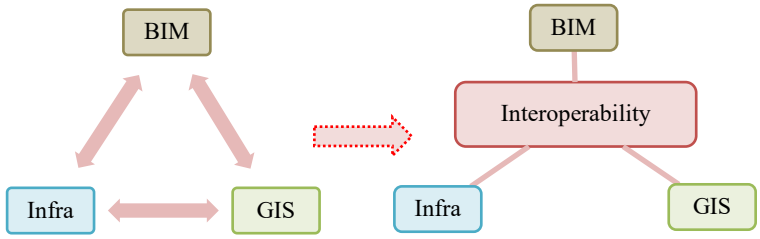


Figure 5 BIM-Infra-GIS triangle

Thus as shown in figure 5 in the entire life cycle of an infrastructure project there is huge overlap between BIM, GIS and Infrastructure processes in different phases. This overlapping calls for interoperability between various data domains.

In the late 80's, when the idea of product models has become more widely accepted, the efforts in standardization of product model data structures in construction were focused on buildings, hence infra-structure projects have been absent from product modelling. Therefore infrastructure design software developers continued to use in-house or national standards for data structuring (Kim et al., 2004).

1.3 OBJECTIVES

As the demand of the infrastructure projects in terms of value engineering increases continuously, integrated project development rises largely. This change alters also the labor market of the infrastructure processes in the future. *Geometric design also affects an emerging objective called livability, which includes providing access to employment, schools, businesses and residences, accommodate a range of travel modes such as walking bicycling transit and automobiles, minimizing fuel use, emissions and environmental damage* (Federal Highway Administration, 2012).

These facts have called for applying model based methods in infrastructure which creates value for such processes. Thus one of the important objectives of this work is to provide an interoperability approach for infrastructure processes in the sense of considering geospatial semantic information already from early phases of design.

On the other hand the type of the information demand in infrastructure projects is widely different from typical BIM task queries. Infrastructure information processes are distinguishable from typical BIM projects due to their cross country nature of most of them, longer project life time, legacy procedure, dynamicity of pre-constraints, having no wide spread product model, exchanging not only data but also regulations, and organizational hierarchies. Mentioned by Esfahani et al. (2014) following are some examples of typical information demand from cross model data.

- Which environmental impacts to be mitigated in designing a road or railway? What do the mitigation activities for the protection areas cost?
- How many trees are along the alignment, how many to be felled, how many to be transplanted?
- Which traffic or city growth to be considered for the next 100 years of the object life time?
- Which parts of the city are under noise emission by a railroad?
- Which compensations are to be considered for socio-economical and land use aspects?
- Who will define such compensation?
- Which utilities to be diverted for construction of a station in a metro line project?
- How to optimize the alignment planning in consideration of all "necessary" factors? Who is in charge of such optimization?

- What is the maximum possible length for a tunnel (Alternative 1) with the available budget or rather choose for the same money a longer bridge (Alternative 2)? (Esfahani et al., 2014)

Obviously most of these issues if not all are spatial in nature. The infrastructure design concerns strongly with the positioning of the physical elements of the infrastructure project according to standards and constraints. Hence infrastructure projects interface strongly with the geospatial data models. The basic objectives in geometric design are to optimize efficiency and safety while minimizing cost and environmental damages.

Thus there is a need for semantic linking of arbitrary data models. Therefore a method is needed which allows for definition of semantic links or an adequate rule based filtering through topological queries.

The integration of GIS in infrastructure design results in avoiding duplications, parallel work or unnecessary work recurring. Interoperability in infrastructure design ends to more effective management and control of the project among the multidisciplinary teams. Integration of all information will only succeed if it can be treated in a coherent way that allows a seamless translation of the databases in such a way that the data loads and proceeds completely lossless to ensure maximum applicability. Interoperability increases reliability, cost and time saving, significantly resulting in a benefit for the designers, managers, and real estate professionals. It also minimizes the changes in the scope of the project.

1.4 APPROACH

Semantic interoperability approaches depend strongly on the data structure. Majority of the operations in infrastructure processes are based on location, and thus spatial in nature (Kayondo et al., 2011). The infrastructure design is much more influenced by geospatial data than the design of most other construction objects (Rebolj, 1993). Infrastructure design concerns strongly with the positioning of the physical elements of the design according to standards and constraints. Thus the most fundamental unification of data models involved in infrastructure domain is the spatial characteristic of them.

The semantic used for the matching purpose depends strongly to the data structure. This thesis has targeted the spatial domain and has used the spatial identity of the data models as the matching approach. The spatial matching illustrates how rule based links are created using spatial indexes. Links can be saved explicitly for example for documenting a specific status of the project for future reviews or the virtual information can be obtained for analysis and processes. Future data models sharing the same semantic are equally manageable in this approach.

Spatiality of the data models involved in information processes in spatial domain can be used for the purpose of semantic matching. In addition, such a rule based linking approach can handle and establish links between dynamic data models to conduct virtual properties. Called spatial matching this approach is independent from and not owned by any of data models. Therefore this rule based matching approach is applicable not only on infrastructure schema but also on any other data model belonging to the spatial domain like geo-mechanical or energy (Esfahani et al., 2014).

Infrastructure and geospatial data models are the most proper application field for spatial matching. In this thesis a high-level spatial matching approach is proposed for infrastructure processes. The semantic and consequently the matching approach create rule based links for the data models using their original format. Therefore as the links are created via rules sharing the same semantic the changes in the data are manageable. I.e. there is no need to up-date the input data model in MM and the links.

1.5 STRUCTURE OF THESIS

To achieve these aims, this thesis reviews in chapter 2 different data models of infrastructure and GIS domain and their standards. The problems of existing data models, their insufficiency and incompleteness in compare to the advance data models will be discussed in chapter 2 as well. Chapter 2 distinguishes in addition the current level of interoperability in infrastructure and GIS domain. The need to a functional interoperability method and the effectiveness of cross domain approaches for

infrastructure and GIS is also discussed in this chapter. Multimodel (MM) introduced by Scherer et al. (2011) will be reviewed as a cross domain interoperability approach to support information analysis and collaborative work over multiple application domains. Besides, chapter 2 investigates the limitations of the existing MM approach for the case of infrastructure and geospatial domain. In addition the role of IDs and data structure in the applicability and effectiveness of the MM approach is discussed also in chapter 2.

According to the discussion in chapter 2 regarding the spatial nature of the infrastructure and GIS domain this thesis suggests in chapter 3 that the existing MM should be extended in a way that semantic links such as spatial links are possible to get created. In chapter 3 the concept of a rule based MM as an extension to the existing MM will be illustrated.

The methodology of this concept has been explained in chapter 4. Chapter 4 describes the match finding solution in different data domains and linking discrete information from different software packages to inspect the coherency of related application models and create complementary analysis models in a rule based link approach.

Implementation demonstrates in chapter 5 linking heterogeneous data models with a rule based link model for rule based filtering and query purpose. As a case study on spatial data, these all implements and customizes a spatial link model between CityGML and LandXML data models so that virtual properties are obtained in different Levels of Detail (LoD).

"This streamlined workflow allows rapid reiteration and collaboration and can significantly shorten the overall development time saving serious time and resources. This makes it possible for state and local government agencies and community development professionals to support all phases of the project lifecycle, including planning, designing, building and management" (Abukhater, 2013).

"With enhanced interoperability, developers will have a common end-to-end solution that supports the entire project lifecycle. Working in unison, this comprehensive technology solution leverages data management and interoperability, powerful spatial analysis, advanced 2D and 3D visualization and rendering of the project in its intended location" (Abukhater, 2013).

Chapter 2

Interoperability in Infrastructure Engineering

By definition, interoperability is the ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE, 1990). Syntactic, semantic and cross domain interoperability are the common interoperability methods in construction engineering. Interoperability is a very vivid topic between cross domain data models in recent years.

Interoperability in CAD based infrastructure design can be considered as in extending parametric models, yet the promising way of interoperability is the one in context of Standards / Data Models / Specifications. In the GIS semantic interoperability is considered in spatial and non-spatial aspects (Laurini et al., 2008).

In the context of this thesis the cross domain interoperability method is the best matching methodology ensuring the purposes of the integration between cross domain data models of BIM. MM is a cross domain interoperability method which enables linking and exchange of data from different data models.

This chapter is a review about the historic development of the data models of GIS and Infrastructure and their data standards. Besides, this chapter gives a review about the existing interoperability approaches between these data models. The application of MM for interoperability purpose between Infrastructure and GIS data models is also investigated in this chapter.

2.1 STATE OF INTEROPERABILITY

Born around 30 years ago as shown in figure 6, computer aided design (CAD) provides powerful and rapid evaluation of alternatives, and optimization techniques for infrastructure design projects (Jha, 2000). Therefore most of the design software if not all, are CAD based. Besides the ability to analyze, recognize and link local, national, or international design standards, CAD processes suite well to the rule based modeling task of infrastructure design. The model aims at design with minimum total cost, conforming to design specifications, environmental demands, and safety parameters.

Industry Foundation Classes (IFC) which is a data model developed for interoperability issues in the building industry. The IFC system is a data representation standard and file format, in fact the only truly universal computer data format, *"used to define architectural and construction-related CAD graphic data as 3D real-world objects. Its main purpose is to provide architects and engineers with the ability to exchange data between CAD tools, cost estimation systems and other construction-related applications"* (Solibri Inc., 2014). Modern BIM systems are able to create rich internal representations on building components. IFC adds a common information transferring language between different BIM applications while maintaining the meaning of different pieces of information. This reduces the need of remodeling one building in different application (Solibri Inc., 2014).

According to the Tekla potential group BIM is known as the "*process of modeling and communicating the structure of a building in detail to benefit the entire building lifecycle. Different terminologies express BIM to facilitate the exchange and visualization of building information in digital 3D format between all building disciplines to ensure integrated project delivery"* [4]. A BIM model is a placeholder for the information about a building, a facility, or a site. A BIM starts with data collection which afterwards can be fed into a BIM ready application to enrich the model with geometry. Physical and functional characteristics of facilities can be digitally stored in BIM.

BIM as a model-based design facilitates to investigate multiple solutions in most of the construction projects and process early enough to help better decision making. The most important reason for implementing model-based design is to help designers and to increase communication between different design parties. It improves team collaboration and it's much easier to manage project data across extended teams and external partners in project lifecycle (Romandi & Vallas, 2013).

The American Heritage Dictionary (2000) defines the term "Infrastructure" as fundamental facilities, services, and installations needed for the functioning of a community or society, such as transportation, roads, communication systems, water and power networks, as well as power plants. High level infrastructure suits mostly facilitate to analyze the infrastructure design based on the international or user defined standards. Called regulation-based design, this minimizes errors, reduces costly design conflicts, increases time savings and provides consistent project quality, yet mostly in standalone solutions (Esfahani, 2013b).

According to Esfahani et al. (2014) infrastructure information processes are distinguishable from typical BIM projects due to their cross country nature of most of them, longer project life time, legacy procedure, dynamicity of pre-constraints, having no wide spread product model, exchanging not only data but also regulations, and organizational hierarchies (Esfahani et al., 2014).

GIS is a platform for managing and presenting spatially referenced information. As the digital representation of the world are systems for maintaining, managing, modelling, analyzing, and visualizing of the world data including infrastructure. GIS with its spatio-temporal analysis tools is a very helpful solution in all steps of alignment planning, as infrastructure design starts with topographic maps and map analysis methods. The general design task is to find the technically feasible alternative and to select the best for further assessment. The design parameters of the alignment and information about the terrain are predefined (Road Design Manual, 2014).

When used together effectively, BIM and GIS have a solid track record for streamlining decision making and reducing inefficiencies in the design, planning, and execution of critical operations and projects. A growing number of engineering tasks in numerous fields including design, architecture,

[4] www.tekla.com

construction, and asset management now require the knowledge of many interrelated yet disconnected BIM / GIS task-specific software.

Both BIM and GIS are decisive in organizing the data tangled in facility operations. Beside 3D model based design BIM can manage huge amount of construction related data and is capable of managing building data itself. A GIS has analysis tools needed for simulating outside of the buildings and is more applicable for projects in multi-site environment and is flexibly customizable (Southwood, 2011). Figure 6 illustrates this development of computer aided design to today's BIM and GIS technologies and their overlapping.

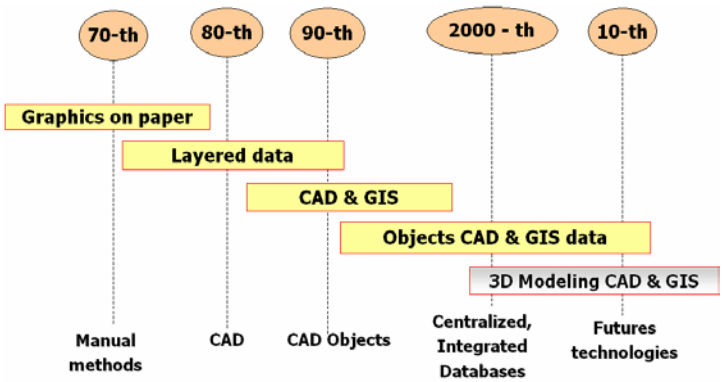


Figure 6 Development schema of computer aided design technologies, (Vitkiene et al., 2008)

Thus interoperability which is the *"data interaction between data models and information sharing between project team members and across the software applications that they use commonly for design, construction, procurement, maintenance and operations is one of the basic challenges in interlinking the data models*[5]*"* (BuildingSMART).

Data interoperability is an important factor to achieve intelligent design process. By coupling the different data models it is possible to activate optimization potential via tempo-spatial interdependencies of the building tasks. During progress in the project, dynamic update will be applied to the models so that different users get technical and commercial information relevant to their project phase. One of the interoperability approaches is to create one exchange format which models all information involved in a project.

"Software applications store the building information in a native and proprietary format. In order to make this valuable information available to other project participants, their software applications either all have to understand the native formats of the other applications, or preferably they support the exchange format as the open format for the data and information models[6]*"* (BuildingSMART).

Therefore due to the lack of or poor interoperability in infrastructure domain and in the absence of an internationally accepted alliance for undertaking the interoperability tasks, each software vendor defines its own domains catalog or objects modelling in accordance to their vendor and customer demands. Thus arbitrary data models might be integrated in future tasks. These facts have called for interoperability and systematic information modelling in infrastructure domain, to create value for such projects in their lifecycle through more transparency and faster and more precise way of exchanging data and documents (Esfahani et al., 2014). Thus there is a need for semantic linking of arbitrary data models. Therefore a method is needed which allows for definition of semantic links or an adequate rule based filtering through topological queries.

When properly executed, it will identify cost savings and minimize delays during construction. After planning, the planners suffer always in operations on the site, due to the matter that most of the times realizing the in-office drawn plan is not the optimum solution in construction site or sometimes not even applicable.

---

[5] https://www.buildingsmart.org/bim, State 2010 Oct

[6] https://www.buildingsmart.org/bim, State 2010 Oct

2.1.1 Interoperability of GIS and BIM

"GIS, is a platform for managing and presenting spatially referenced information. Within this domain, the exchange of geospatial data and the interoperability between systems are established using the Geographic Markup Language (GML). GML is an Open Geospatial Consortium (OGC) standard data model for defining the data types and constructs for describing the geographic features" (Amirebrahimi et al., 2015, p 79).

Geospatial objects are described in two types: a) spatial data introducing the location and b) attribute data defining characteristics at that location. Such vector or raster data models are represented digitally via layer combination based on similar identifications (e.g hydrography, elevation, water lines, sewer lines, grocery sales) and by selecting appropriate layer appearance with respect to projection, scale, accuracy, and resolution (Briggs, 2015). These data are incorporated by Relational Data Base Management System (RDBMS).

With a more specific focus, the *"heterogeneous geospatial information about urban data (e.g. buildings, transport, vegetation and water bodies) at different levels of details is integrated within the framework of the 3D virtual city models such as CityGML"* (Döllner et al., 2006). *"CityGML is the most comprehensive urban information model within the geospatial domain to digitally represent a city in 3D. The building information in CityGML (and other existing GIS formats) however, is not as complete and mature as BIM and for this reason multiple extensions, e.g. utility network in Hijazi et al. 2010, have been created separately over the past years to improve the model"* (Amirebrahimi et al., 2015, p 80).

"GIS and BIM originate from different domains and were developed for the specific needs of that field. Their integration creates a seamless and scale-independent view of the world across both domains. It can benefit a variety of applications that meeting their requirements would not be possible by independent use of BIM or GIS" (Amirebrahimi et al., 2015; Karimi & Akinci, 2010). *"This integration however is not simple due to the differences between the two. Such dissimilarities are discussed in terms of spatial scale, level of granularity, geometry representation methods, storage and access methods as well as semantic mismatches between them"* (Karimi & Akinci, 2010; El-Mekawy & Östman, 2010).

Various attempts have been made for integrating BIM with GIS that can generally be classified into three groups: at application, process, and data levels.

"Initially, the overall scope is to store, manipulate and maintain data that are directly related to the mission of system users. But increasingly the need to share information with other institutions and to reduce data cost has allowed the emergence of tools for the support of not only data sharing, but also interoperability between data and services" (Laurini et al., 2008, p 2). GIS interoperability is a necessity for applications that involve combining data from several data models. Typical example of applications in the context of infrastructure includes street repair, environmental monitoring and studies, international transportation, navigation and location based services, etc.

Interoperability happens at the application level via rearrangement (Karimi & Akinci, 2010) in which software patches modify existing GIS or BIM tools or the tools become developed new to integrate the functionalities of the other. *"This method is generally costly and inflexible. On the other hand, process level integration methods like OWS-4 project by OGC (2007) use Service Oriented Architecture (SOA) to allow the participation of BIM and GIS systems in those tasks that require the capabilities of both while they simultaneously remain live and distinct. This method provides more flexibility than the first group. However, in this method, the challenges of integration are still to be resolved at the underlying data level to provide interoperability between these systems"* (Amirebrahimi et al., 2015, p 80).

The integration of BIM and GIS at the data level happens through linking methods developed via Application Programming Interface (API) at either side. FME (Safe Software, 2010), IfcExplorer (IFCExplorer, 2010) and the work by Nagel et al. (2009) belong to such conversion or translation methods which directly convert between GIS and BIM formats.

As discussed in detail by El-Mekawy (2010) such translation systems between CityGML and IFC data to each other are still being developed. Challenges are to keep semantics, to convert geometry as it is and to concentrate on infrastructure domain elements.

The geometry transformation concern is being addressed by Li et al. (2006) and Wu et al. (2010) and results in partial integration. Advanced flexible integration mechanisms either extend one of data models to incorporate the data from the other or address a new data model as a "meta model" to arbitrate between BIM and GIS at Meta level.

One of such efforts is the work of Nagel (2007) for automatic conversion of IFC building models into CityGML. Another example is using BIM for automatic generation of buildings in CityGML according to building components and their semantics as in the research of Isikdag and Zlatanova (2009). Examples in the side of CityGML contain conceptual requirements discussion for converting CityGML to IFC models (Nagel et al., 2009), Application Domain Extensions (ADE) for integrating IFC into CityGML (Van Berlo, 2009), and the GeoBIM as well as utility network extensions by Van Berlo and Laat (2010) and Hijazi et al. (2010).

On the other hand, IFC-for-GIS project or IFG (BuildingSMART, 2005) intended to extend the IFC model to include information from geospatial world. IFG was initiated to facilitate the interaction of geographic information created using GIS systems within the framework of the IFC schema. Even though providing a complete model for geographic information and supporting complex geodetic coordinate systems has been out of the scope of this work still most of the interfaces for visualization and modeling purpose has been considered and developed in IFG framework. A prominent example of meta models is the "Unified Building Model" (El-Mekawy et al., 2011).

The aim of interoperability between BIM and GIS *"is to meet the increased demand for construction analysis, urban planning applications, disaster management, cadaster and homeland security and other applications"* (Johansson & Roupé, 2010). *"These applications require not only 3D geometry and appearance but also complex semantic information. Conceptual models have been developed in BIM and GIS, in forms of geometric and semantic models. For only visualization purposes, geometric along with texture information are sufficient to represent 3D spatial objects. However, semantic models are needed for different engineering and planning applications that require complex queries and analyses"* (El-Mekawy & Östman, 2010, p 1).

Predicating two GIS and BIM worlds do not really integrate Léon van Berlo and Ruben de Laat (2010) have developed a new so called "GeoBIM" CityGML extension for IFC data which is capable of transforming the IFC geometry as well as semantics and properties into GIS. They solved the previous problem of GIS/IFC integration namely focus on converting just geometry. El- Mekawy and Östman (2010) have also created a unified model oriented approach for semantic mapping and bi-directional integrating the IFC and CityGML. In the case of overlapping objects, concepts were merged; otherwise new concepts are created to capture all indoor and outdoor objects. Finally, relationships among all objects have been built with consideration of semantic and geometric properties.

All these show that the focus of the integration was to develop an intermediate data model for building, building plan submission process and its facility management. *"In general, these models are application-focused and the integration is made for a particular use case with specific requirements. Therefore, the included concepts and relationships within these models may not suit other applications with different functional requirements"* (Amirebrahimi et al., 2015, p 80).

Therefore the interoperability between GIS and BIM does not exist for infrastructure data models. In regard to infrastructure domain, what infrastructure design really needs is something to manage information on a large scale.

This emphasizes the need of integrating infrastructure data model into a coherent community of standards which can simulate the construction site for a proper decision making as virtual model is used for testing decisive processes to omit unnecessary delays. This digital model of the construction site can be used in all phases of the project. In such digital model data from various disciplines can be integrated into a central standardized data platform. The model should be updated dynamically and therefore, the model is used by various team members in all stages of the project to obtain technical and economic information.

2.1.2 Interoperability of GIS and Infrastructure

According to Autodesk in the field of infrastructure design, road corridors are arguably the most powerful and sophisticated objects of design. To create a road corridor, an alignment, profile, and assembly are combined to form a 3-dimensional representation of a linear feature such as a road or channel. The corridor model serves as the "backbone" of the design and from it many useful forms of information mostly GIS models can be derived.

Thus, the infrastructure design world has historically been concerned with geospatial information and hence typically in the area of GIS. GIS is much stronger in the case of infrastructure design due to its ability to analyze tempo-spatial information and integrate data from different systems. *"The built environment general information map can be divided in two parts:*

- *Source and reference data (metadata, land survey registries, maps, city plans ...), which represents information that is rather static in nature, managed by e.g. public authorities. As discussed above this part is covered by OGC and ISO 19100 series of standards.*
- *Design objects and structures (infrastructure, networks, pipelines ...), which represents the more dynamic project information, created and used by those involved in design and construction activities"* (BuildingSMART Nordic, LandXML for IFC, 2012).

Although there are still many other ways to follow the design process, yet the key factor in infrastructure design is not the geometry but the design parameters which will be visualized in the geometry. Therefore, design process is extremely influenced by a huge number of design parameters, which vary with for instance the country of construction, the present road and safety level, the details necessary, regulations for design as well as for construction and the legacy procedure that apply.

GIS being able to handle both geometric and semantic data models has supported infrastructure planning over the years. In the early stages before design procedure is started, GIS plays an influential role in evaluating and choosing sites and data integration during the construction process. Referred to as GIS based design or geo-design this has received increased attention in recent years (Steinitz, 2012). Sustainability in geo-design has comparable objectives to support decision making, maximize infrastructure lifecycle and minimize costs.

With typical length of around 500 kilometers for railway projects and around 100 kilometers for highway projects, it is very important to reduce as much as possible the repetitions in processes and costs in manpower. OBERMEYER consultants applied a GIS model in the 700km Algerian high plateau railway project in 2006 and introduced a low-level integration of GIS / BIM which improved the efficiency of the design task and reduced the planning costs (Esfahani et al., 2011).

Therefore, the standards are already defined either in national or international categories but as infrastructure is not separated from civil projects it is important to merge its standards with civil projects standards. Up to know the main focus of IFC, the international standards catalog for civil projects, has been the building schema, IFC does not yet cover the infrastructure data models completely and includes neither infra domain nor any extension dealing with infrastructure data. Even though IFC-Bridge data model is tightly linked with road data types, still roads, highway and railway design is not included in its scope (BuildingSMART). *"The first attempts to extend the IFC model into the field of infrastructure projects are currently in progress. Recent research and development in this field is marked by the development of the international exchange formats IFC-Bridge and IFC- Road and numerous national standards"* (Ji et al., 2009; Yabuki et al., 2003, 2006).

"On September 22-23, 2005, CSTB Sophia-Antipolis hold an IFC-Bridge and Roads workshop for the final presentation of IFC-Bridge data model before official integration into IFC international standard and discussion on opportunity for launching an "IFC-Road" project" (Lebégue, 2005).

This caused an international cooperation within the Industry Alliance for Interoperability (IAI) society now buildingSMART focusing on roads, which means that roads are supposed to get embodied into the IFC standard. A standard developed within the construction sector and widely adopted as international exchange standard for object oriented design.

Mentioning that an international standard can focus on roads network administration and maintenance in the left loop of figure 7 or on roads detailed design and building operation in the right loop in figure 7, his discussion focuses on following points:

- It must be very helpful to have an integrated product data model for roads and bridges design because of the interaction needs between this two type of structures and dedicated methodology including CAD and analysis software are based on the same geometric concepts
- It must be interesting to derivate a road design product model from a building and very complete one like proposed by IFC
- This building and civil engineering integrated product model shall have interaction with GIS data for eSubmission and detailed urban and territorial roads network planning management purposes
- Such integrated data models shall speed up the creation of faithful virtual reality environments with large capabilities of simulations

To help answering the above mentioned questions for buildingSMART IFC-Bridge has been accepted to be a very good base and background for a so called "IFC-Road" data model as in the figure 7.

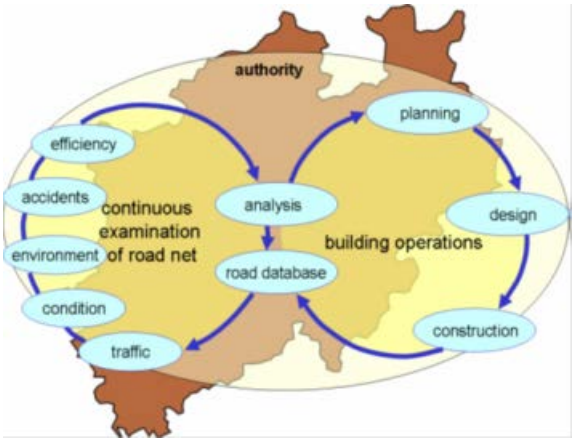


Figure 7 Scope of German OKSTRA solution for roads design management, (BuildingSMART, 2005)

Still the fact is, since both the IFC-Bridge and the IFC-Road, backbone standards for infrastructure part of the package, are still in the development phase, there is currently no integrated IFC product model for alignment, bridge, building, grounds and site installation. Another key matter in this regard is that, IFC-Road still will not cover the railway domain. Much worse is that each railway has its own standards from country to country and yet there is no accepted international design norm for railroads.

According to Obergriesser and Borrmann (2012) *"an analysis of the organization of the entire geotechnical infrastructure process reveals a further problem. Currently the workflow is not structured and is based on fixed timeslots or the experience of several project managers. A detailed description of an ideal infrastructure process is not yet defined"* (Obergriesser & Borrmann, 2012, p 582).

As a part of the IfcAlignment research Prof. Borrmann (2013) and his team has created a new alignment data structure and an alignment meta-model for the comparison of alignment product models. The Alignment Meta Model allows easy transformation between different alignment models and can be used to compare different alignment models. This is a help to find the best model for a specific use case and to evaluate the different design alternatives for alignment models (Amann et. al., 2014). Yet this is a use case dependent and a new alignment data model which does not solve the problem of interoperability between existing data models. This new data model is designed for IFC, from one side it adds to the complexity of IFC and from the other side serves only for specified infrastructure uses of IFC.

Therefor the interoperability between GIS and Infrastructure is either limited to those application based interaction possibilities developed for a specific part where GIS and Infrastructure overlap or the conversion or translation of the data domain of each data model in the other one such as traffic data model of GIS or topology / surveying data model in infrastructure domain.

2.2 MAIN CHALLENGES AND RELATED WORK

According to the discussions above, in most of studies targeted application areas explicitly include urban and landscape planning; architectural design; tourist and leisure activities; training simulators; and mobile robotics. 3D cadasters; environmental simulations; mobile telecommunications; disaster management; homeland security; vehicle and pedestrian navigation with different data model architectures have been not considered very much in detail. As shown in figure 8, alike building projects, infrastructure projects have a lifecycle in which project development, design, construction, operation and maintenance as well as retirement are very common topics.

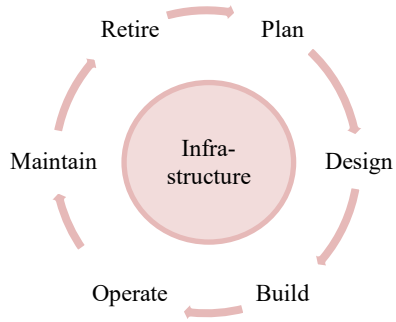


Figure 8 BIM Lifecycle for Infrastructure Projects

In the field of infrastructure domain there are many data formats, which are mainly due to the diversity of labor and the organization of the planning process as well as the construction industry, such as exclusive identity of the structures, project teams, construction, high proportion of the design in the overall services, no mass production, no prefabrication, high regulation, etc. Therefore unlike the wide application of business informatics in the automotive industry process and the production chain, it is not possible, to apply the knowledge of economic computer sciences like SAP for the information management of the infrastructure domain. Existing specialized software applications for design management and construction will continue to produce and process the data in these formats. These specialized applications have spread widely, are expensive and highly specialized, so that it is not possible to simply introduce a new even multipurpose software. Besides these existing data formats are derived from different domains such as construction site, structure, engineering services, operations, etc. The constantly demand increase for cross-domain management of the project tasks, indicates how necessary are the cross-domain data like geo-information and infrastructure.

Therefore motivated by strong market needs, organizations around the world are working together to extend the open standards foundation for also road design information integration. This promises to benefit government agencies and also companies in the value chains associated with road and railway projects planning, construction, sales and management. It also promises to play an important role in time saving in planning procedures. Planning a road is limited to the standard geometric data models namely vector and raster and *"Spatial Multiple Criteria Evaluation (SMCE) which is based on multiple attribute decision analysis techniques and combines multi-criteria evaluation methods and spatio-temporal analysis performed in a GIS environment"* (Farkas, 2009, p 722). While the design is going on, gain insight into construct-ability with high-quality, model-driven visualizations, simulations, 4D modeling, and clash detection capabilities through GIS modeling tools, one can enjoy:

- faster planning
- lower software and maintenance costs
- create more sustainable designs by using precise model data to facilitate environmental impact analysis
- minimize disruption to land and water
- coordinate construction

In its report of 23 May 2012 buildingSMART indicates that *"there is an increasing need to look at the buildings as integral parts of their environment, enabling integrated design and construction of smart built environments and communities. Often it will be difficult to draw a sharp line between a building*

and civil structures, therefore the information should be readily available to actors on both sides of the line — meaning interoperable standards, if not one for all. Both building and infra design and construction are usually done within the same companies, also many times the software for the both are provided by the same vendors. Since many of these are already members of the buildingSMART this would be a natural way to reach the relevant stakeholders to develop and implement infra sector information standard, interoperable with IFC5 (or integrated in IFC5 in the long run. IFC specification rather well covers buildings and to some extent their relationship to the surroundings world. Civil engineering structures are not specifically covered by IFC4 (even the IFC-Bridge extension is still not done), even though many basic capabilities already could be used (and have been used in some cases). Other open standards in this area are scarce, LandXML being the one to cover infra projects' domain" (BuildingSMART, 2012).

Besides there are viewpoints such as of Rangnes (2012) that LandXML as the internationally accepted standard for infrastructure design is not really a model, but a format supporting some semantics. In addition this viewpoint indicates LandXML is not really sustained as the last revision/update of it dates back to 2008. Therefore it suggests buildingSMART take ownership of LandXML and ensure the interoperability with IFC and compatibility with LandXML. Such viewpoints and the need of integrating BIM with Infrastructure have called for a new so called "openINFRA" initiative in many countries especially in Europe. According to buildingSMART in their report of May 2012 "*the proposed strategy for supporting infra project information exchange is two-fold:*

- *In short term, acknowledge LandXML as an interoperable "affiliated schema". This requires IDM/MVD development and mapping to IFC. Software certification should be considered.*
- *In long term (to be specified in the openINFRA roadmap) develop extensions to IFC schemata to move towards one integrated standard. The experiences in short term implementation and use will guide this step"* (BuildingSMART, 2012).

BuildingSMART Norway suggests an overall interoperability standards package as shown in the figure 9.

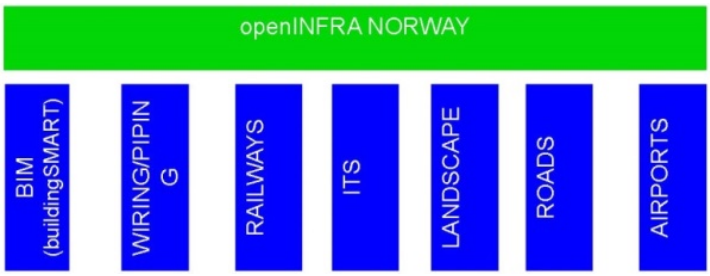


Figure 9 Overall Interoperability Standards Package, (BuildingSMART Norway, 2011)

This shall include involved organizations, processed, terminologies and data formats (IFC, LandXML, GML …). This will be an internationally coordinated Norwegian project, based on open standards, which allow different sectors sharing and using mapping, environmental, and socio-economic data as well as easy access to updated data regardless domain/sector.

Therefore with the goal of focusing on constructability and to serve as a generic guide for integration of BIM and INFRA openINFRA is now creating a catalogue of definitions level with INSPIRE connection to facilitate the effective use of GIS.

Having said it all, one should know that, due to the fact that unfortunately, infrastructure is designed as differently as there are countries in the world, and even worse, each county or municipality tends to do the process according to their own solutions, the IAI work of developing infrastructure domain started in the fall of 2006 still doesn't exist (KJEMS, 2007).

Such package is not only a container for the features and design regulations and parameters but also contributes in maintenance the planning framework and updating it. Moreover it provides an exchange format between different applications dealing with infrastructure design tasks.

Moreover, according to Liebich et al. (2013) and as indicated by Katranuschkov et al. (2014) *"continuously extending the use of BIM-based working and the related needs for BIM-based*

interoperability of more and more specialized AEC tools in various building construction subdomains showed:

- *global all-encompassing models for all data in a construction project are neither realistic nor practical target*
- *BIM data typically have to be combined with other kinds of construction related data in order to be efficiently applied in real AEC tasks"* (Katranuschkov et al., 2014).

On the GIS side there are also main concerns regarding the geospatial data models. Most of the confusion about data models arises from their diversity. The data models as illustrated in figure 10 go from most general at the top level (vector, raster, TIN) to most specific at the bottom level (shapefile, coverage, geodatabase). It is important to note that a geodatabase can handle all three general models, not just the vector model. Geographic data models have evolved under the influences of technology (e.g., increasing storage space and processing power, networking, or software evolution) and even history.

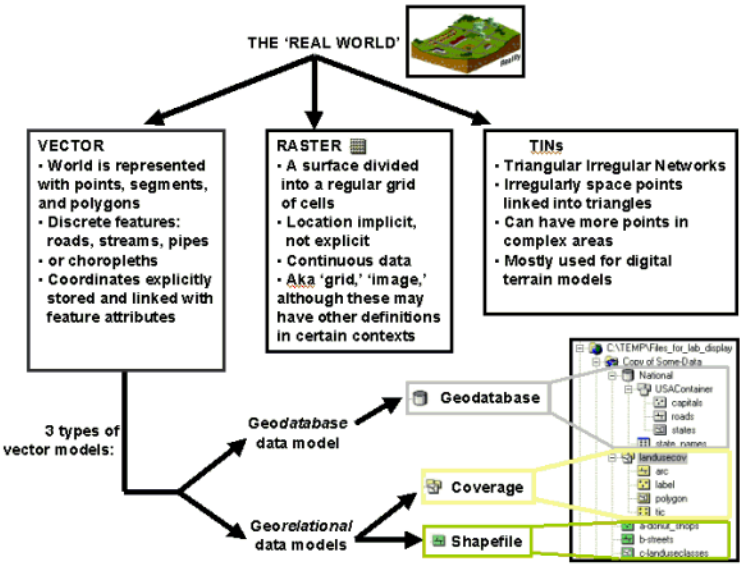


Figure 10 Hierarchy of geospatial data models, (USCB, 2010)

Every GIS software package will be capable of supporting a number of data models, but will also have it's own proprietary format (that none of the others might read). The capabilities of the data models may change with new versions of the software, and compatibility issues may arise between different GIS software, and even between different versions of the same software. Certain functions might be accessible using data in the form of one data model but not another. This is also another challenge which makes the data structure of geospatial data and their spatiality more important and even helpful in the context of interoperability than the data models.

2.3 INFRASTRUCTURE MODELING IN GEOSPATIAL CONTEXT

There is a semantically rich information model behind IFC, which is in fact more detailed than CityGML. However, lacks in other city furniture and has totally limited georeferencing possibilities. *"A lot of data is involved in the planning and management of road maintenance operations. Since majority of these operations are based on location, a great deal of this data is spatial in nature"* (Kayondo et al., 2011). *"Roads have been designed digitally since the early 60's, when IBM introduced the program package HIDES. Since then the most radical changes in the road design software were mostly influenced by the development of computer graphics, which brought road design in line with other CAD software. The design of roads is, on the other hand, much more influenced by geospatial data than the design of most other construction objects; therefore road design software has often been linked with GIS"* (Rebolj, 1993, 2000, p 2). In the late 80's, when the idea of product models has become more widely accepted, the efforts in standardization of product model data structures in construction

were focused on buildings. Therefore road design software developers continued to use in-house standards for data structuring (Kim et al., 2004). *"On the other hand GIS standards for data exchange have been under development in the 90's, which has added to the unclear situation about infrastructure modeling standards"* (Rebolj et. al., 2000, p 2).

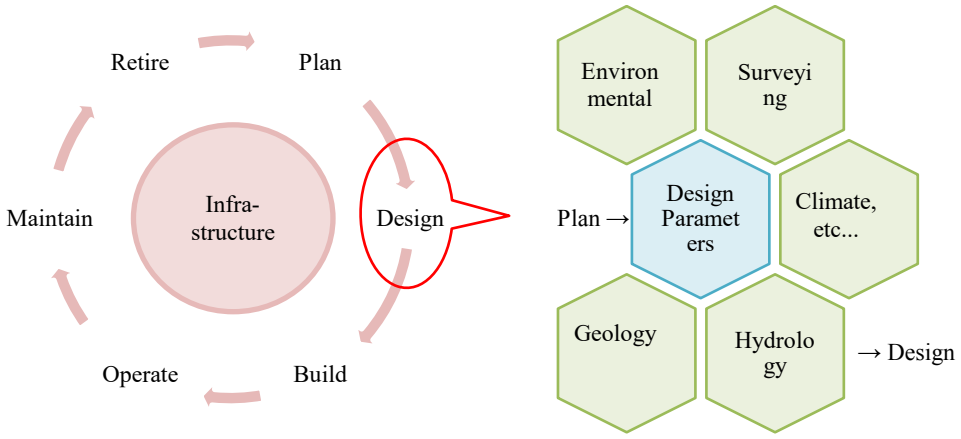


Figure 11 Data Models involved in Infrastructure Design, (Przybylo et al., 2015)

Geometry is the core in the infrastructure software package, which furthermore represents the fundamental characteristic and functionality of the product model. To achieve the best geometry design of the infrastructure with the less risk factor, the BIM concepts have forced to integrate as much as data models which affect an infrastructure project in an interactive platform. Shown in figure 11 most of the data models affecting the infrastructure design process have geospatial nature which differ in share in various design stages. The proportion of different data domains in infrastructure projects are shown according to the experiences from various national and international works in the figure 12.

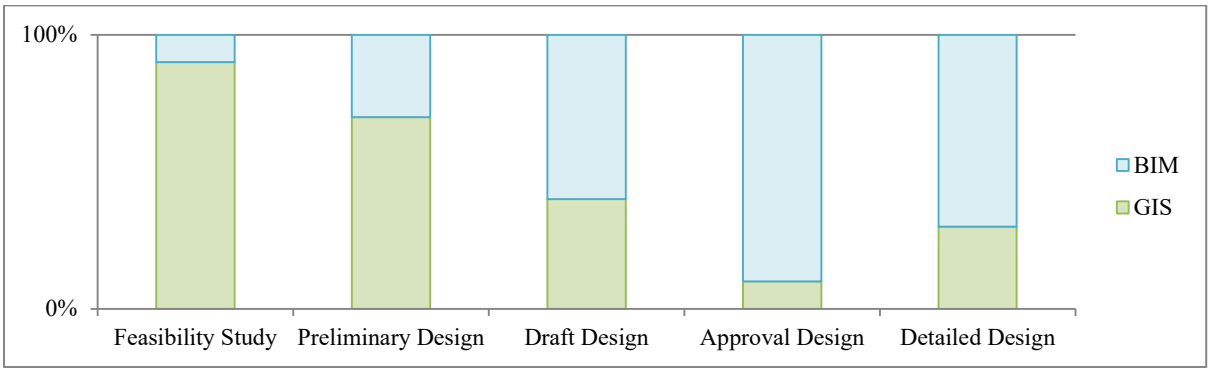


Figure 12 Proportion of GIS and BIM data types in different infrastructure design stages

In infrastructure design an integrated view over the planning data and a comprehensive data exchange are the basis for efficient planning. Besides, for a coherent recording, editing and maintaining the characteristics of design consumptions, in addition to an instruction of design data base, a standardized design object catalog is necessary. That is why there are several national and international STEP and XML based standards which allow a high level data exchange for infrastructure data. Common national standards are German OKSTRA, Swedish EuroSTEP, and American TransXML. IFC-Bridge and LandXML are common international standards, whereas IFC-Bridge is developed more for bridge design than for infrastructure design.

"LandXML is the format of the land-development world, which is a nonproprietary data format that allows the exchange of land-development design elements between various products. As the developer consortiums are different American companies with international offices, they have been able to create a better translation mechanism for the land-development community at large. Different versions of this common language are called "Schema" and contain important information about the format of the data that is included" (LandXML org.).

Another model related to infrastructure was initiated within the object catalog for road and traffic data project (OKSTRA), coordinated by the German federal institute of roads. *"The main objective of the project was to create the common platform for describing, storing and exchanging all data related to road design, road maintenance, and traffic. The authors considered the international standards, German particularities, regulations and the needs of their construction industry and traffic professionals. OKSTRA is a complex network of objects, described using NIAM, and EXPRESS, and is being accepted by the German industry recently"* (Rebolj et al., 2000, p 2). OKSTRA is developed and standardized by the Federal Ministry of Transport, Building and Urban Development. In compare to LandXML which is defined to be based on planning requirements, OKSTRA looks more into construction of infrastructure objects.

Therefore in this object catalog not only all objects of roads and traffic data model, but also the current regulations and principals of road and highway planning in Germany as well as design and draft procedure of road projects are consistently defined in the form of EXPRESS schema. The implementation of the OKSTRA standard in the practice requires great effort for the CAD software companies (Esfahani et al., 2012). The Standard exchange format OKSTRA-CTE has been enforced on a large scale since just the last few years (Ji et al., 2008). OKSTRA describes the object structure of various technical parts of road and traffic data systems which can be used as input and/or output of the design processes involved, particularly in applications and databases. It is a basis for unified technical implementation of the deployed and developing digital procedure of road construction management and assures an almost disruption-less information flow between different courses of the procedures.

"The Building and Construction Research group in the Netherlands (TNO Institute) has developed a model for the description of road geometry (The Road Shape Model Kernel – RSMK), which serves as the basis for more profound description of a road. The model is developed in accordance to the STEP standard. The initial anchorage and the horizontal axis elements, which are longitudinally stretched along the axis and parametrically and functionally described, define the horizontal axis. All the intermittent coordinates of points can be indirectly calculated. In this manner the quantity of data and consequently the complexity of the model are significantly reduced. The problem, however, arises when dealing with longer sections of a road where the summation of arithmetical errors can cause larger deviations. The authors were aware of this problem. Nevertheless, the tests have revealed that if the data is sufficiently accurately stored, the deviations are not significant. Yet, the model doesn't follow the way roads are being designed and it has not been reported about further development or implementation" (Rebolj et al., 2000, p 2).

Swedish National Road Administration has afforded a road product model in which the geometry model is defined through STEP standard and is its integrated sources. As the most complete data model this model includes the biggest quantity of road data. In addition to geometry and the geographic description of an object, this data model contains road elements functional classification, physical objects properties, project execution tasks. Besides, this data model is merged with IFC (Rebolj et al., 2000).

American National Cooperative Highway Research Program (NCHRP) has also developed TransXML which is an XML-based schema for transportation applications. As a broadly accepted public domain TransXML is a platform for exchanging, development, validation, dissemination, and extension of transportation data schemas. The key four domains of this product are "Survey/Roadway Design", "Transportation Construction/Materials", "Highway Bridge Structures", and "Transportation Safety".

TransXML provides a common vocabulary and an information system for transportation tasks as well as a valuable framework for exchanging data between transportation applications. According to TransXML organization "*the products of this project are a family of XML schemas for transportation applications, along with a recommended institutional structure and process for implementing and sustaining its use"* (TransXML org).

In respect to the aims and purposes of this thesis following criteria are defined for evaluation of different standards:

- Handling geometry as well as semantics
- Which regulations are considered?

- Is standard internationally accepted?
- Is standard extendible?
- Which infrastructure domains are dealt with?

Using different data formats; one can reduce the human errors occurring due to misunderstandings or communicational errors that come from different design programs for planning. Table 1 shows a comparison of these different standards in regard of demands and requirements of this thesis:

Table 1 Comparison of different standards in accordance to selected design criteria

Standard / Criteria	OKSTRA	EuroSTEP	TransXML	IFC-Road	LandXML
Geometry, Semantic	x	x	x	x	x
Regulations	german	swedish	Aashtow (USA)	Ifc (Int.)	landXML (Int.)
International	-	/	x	x	x
Extendable	x	x	x	x	x
Road, Rail, Pipes	x	/	/	-	/

As shown in the table 1 the *"German Standard OKSTRA is of particular interest. The entire range of objects that feature in road construction and management are defined in an EXPRESS schema, in particular the rules and regulations in force in Germany with respect to the process of planning, modeling and design of roadways. However, the main shortcomings of the German standard are its inadequate implementation in existing construction software programs as well as being implemented specially for German regulations"* (Ji et al., 2009, p 3). As mentioned before huge standards organizations such as buildingSMART are trying to embed and integrate the LandXML standard.

Although this thesis according to its requirements has chosen LandXML to be the infrastructure design standard for its modeling purposes, however as will be shown later, the proposed approach is independent from the data format and standard.

2.3.1 LamdXML: Infrastructure Data Standards

LandXML org is an international consortium of software firms, equipment manufacturers, consultants and other land-development professionals creating a lingua franca for the industry, and simplifying the sharing of design data between design and analysis packages into a simple import / export operation. One of the most important aspects of this product is to accommodate international requirements. LandXML is made up of 70% non-US participants; making LandXML an international standard.

As indicated in the report of 20.05.2012 of buildingSMART about InfraFinBIM, *"LandXML is an open specification for the exchange of civil engineering and survey measurement data between members of a design team, between designer and surveyor, and from design into construction, including into machine controlled site equipment. The specification covers civil design information, including surveying data, surface data, parcel data and 3D road, street, railway and waterway models, as well as pipe networks. In addition to geometry of roadway alignments, cross sections and surfaces, standard properties (functional or characteristic) can be exchanged, and the schema also allows for user defined features (analogous to property sets in IFC). In principle, LandXML schema covers the fundamental information requirements of land development projects and it is implemented in major industry software applications (70 registered according to LandXML.org); also, it has been used in various projects, but there is no good overview of its usage or the quality of implementations (largely due to the lacking organizational support for LandXML)"* (BuildingSMART, 2012).

The data schema facilitates the exchange of data created during the Land Planning, Civil Engineering and Land Survey process. Land development professionals can use LandXML to make the data they create more readily accessible and available to anyone involved with a project. With LandXML, project data is independent of the authoring software, thus overcoming the interoperability problems that have plagued the land development industry. LandXML provides interoperability between different land application software, as well as varying versions of software. As a result, data can be archived and accessed more readily on future projects. Additionally, other web-based tools can be

used to view, edit, and report LandXML data. Extensible Stylesheet Language (XSL) style sheets can be easily created and applied to LandXML data and then run from a project web page. Examples include XSL style sheets that format raw point data into point tables, or format data to match an organization's internal standards, such as legal descriptions for parcel reports.

"LandXML is not really a model, but a format supporting some semantics, besides landXML is not sustained as the last revision / update has occurred in 2008" (Rangnes, 2012). Yet *"landXML has been widely implemented in today's CAD programs for civil engineering projects internationally. The standard is based on the data exchange format XML (Extensible Markup Language) and structured hierarchically. The top element, the road alignment, is divided into different subparts including the main road design views in 2D, such as the horizontal alignment ("CoordGeom"), the vertical alignment ("Profile") and the road cross-sections ("CrossSects"). These parts are further divided into subparts that are related to one another, making it possible to describe and depict the geometric and semantic attributes of the road alignment. The advantages of the LandXML standard are the wide support provided by computer-aided systems for infrastructure design (Autodesk Civil3D, RIB STRATIS, and Bentley Inroad) and the flexible extensibility of the LandXML schema"* (Ji et al., 2009, p 3).

2.3.2 CityGML: Geospatial Data Standards

"The Federal Geographic Data Committee (FGDC) is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis. This nationwide data publishing effort is known as the National Spatial Data Infrastructure (NSDI). The NSDI is a physical, organizational, and virtual network designed to enable the development and sharing of this nation's digital geographic information resources. FGDC activities are administered through the FGDC Secretariat, hosted by the U.S. geological survey"[7].

International Cartographic Association (ICA) has set a spatial data standards commission with the task to verify principles for identifying the procedure of transferring spatial data, and any norm aligned with that process. This commission studied Spatial Data Infrastructure (SDI) development at all regional levels, and its implications on the world of the ICA. The commission currently develops an SDI conceptual model and defines its technical characteristics[8].

With the task of developing publicly available interface standards 477 companies, government agencies and universities have formed an international industry consortium called the OGC. This standard is supposed to promote interoperability solutions to "geo-enable" the web, wireless and location- based services and mainstream IT (Arctur, 2006). The standards facilitate accessibility for complex spatial information and services and applications and allow spatial assets and queries. A spatial query is a particular database inquiry type created for geospatial and temporal databases. There are important differences between normal SQL queries and this type of queries. These differences are for example that they enable the use of geometry data types and consider the spatial relationship between such geometries.

According to Briggs (2015) GIS spatial data types are:

- continuous: elevation, rainfall, ocean salinity
- areas:
    - unbounded: landuse, market areas, soils, rock type
    - bounded: city/county/state boundaries, ownership parcels, zoning
    - moving: air masses, animal herds, schools of fish
- networks: roads, transmission lines, streams
- points:
    - fixed: wells, street lamps, addresses
    - moving: cars, fish, deer

[7] http://www.fgdc.gov/

[8] http://icaci.org/

GIS attribute data types are:

- Categorical (name)
  - Nominal (no inherent ordering such as land use types, county names)
  - Ordinal (inherent order such as road class, stream class)
- Numerical
  - Interval (no natural zero such as temperature (Celsius or Fahrenheit))
  - Ratio (natural zero such as income, age, rainfall)

According to University of California, in GIS Labs *"most of the confusion about geospatial data models arises from their diversity. Some data models are more abstract / theoretical while others are made with specific database types in mind. For example, the vector data model and the raster data model are very general, whereas the georelational data model and geodatabase data model are made to fit specific categories of database software. Furthermore, a given data model may belong to more than one category: a coverage is both a vector data model (general) and a georelational data model (database specific). That is why in the geospatial data, data structure is more important than data models. The specific format with which the data are stored on the computer is known as the data structure. To illustrate, consider a basic vector data model. The vector model represents features as consisting of lines which individually link together a start node, vertices in between, and an end node. To draw and analyze features represented this way, the computer needs information on the locations of each node and vertex of the lines. This could be provided in the form of a table listing the coordinates of these points, and indicating which line(s) go through them. This table would be the basic data structure"*[9].

"CityGML is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for the Geography Markup Language version 3.1.1 (GML3), the extendible international standard for spatial data exchange issued by the OGC and the ISO TC211. The aim of the development of CityGML is to reach a common definition of the basic entities, semantics, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields"[10].

In addition to spatial properties, CityGML features can be assigned appearances. In the context of representation of observed 3D topography, CityGML introduces explicit 3D shapes such as surfaces and volumes. Besides, CityGML identifies of most relevant feature types usable in a wide variety of applications.

Yet, the geometry and its appearance are only one aspect of the entities. The key issue in CityGML is semantic modeling, such as type of buildings, categories of streets, type of traffic information. Therefore CityGML facilitates external references in external datasets to be joined to corresponding objects (Kolbe, 2011). The greatest benefit of CityGML is modeling representation of well-defined objects with their semantics (among them spatial and graphical ones), structures, and interrelationships from their geometry / graphics oriented models. This modeling results in homogeneous data quality in the same scale or LoD.

With the aim of establishing high degree of semantic and syntactic interoperability between different city data models, CityGML enables multifunctional usage of 3D city models via defining comprehensive thesauri for a common information model (ontology). Therefore, CityGML ensures suitability for spatial data infrastructures via well-defined semantics for geometries.

2.3.2.1 Geospatial Data Standards and Level of Detail

As shown in figure 13 CityGML differentiates five consecutive Levels of Detail (LoD), where objects become more detailed with increasing LoD regarding their geometry and thematic differentiation[11].

---

[9] https://dusk.geo.orst.edu/buffgis/Arc9Labs/Lab2/lab2.html, State 2009 Apr

[10] http://www.opengeospatial.org/standards/citygml, State 2009 Dec

[11] Complementary discussion in chapter 4, part 4.2

- LoD 0 – regional, landscape (2.5D Digital Terrain Model)
- LoD 1 – city, region ("block model" w/o roof structures)
- LoD 2 – city districts, projects(textured, differentiated roof structures)
- LoD 3 – architectural models (outside), landmarks (detailed architecture model)
- LoD 4 – architectural models (interior) ("walkable" architecture models)

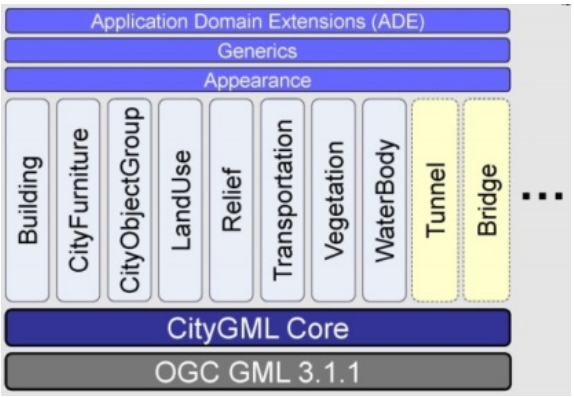


Figure 13 Modular Structure of CityGML, (Kolbe, 2009)

Thus CityGML considers 5 levels of details for its representation. CityGML files can - but do not have to - contain multiple representations (and geometries) for each object in different LoD simultaneously. *"An alternative approach to LoD creation and management is presented in Döllner and Buchholz (2005) and Ohori et al. (2013), who propose the dynamic generation of coarse representations from finer ones and replace finite LoDs by continuous ones"* (Borrmann et al., 2014b, p 2).

However in CityGML the coarsest level LoD0 is essentially a two and a half dimensional Digital Terrain Model over which an aerial image or a map may be draped. For buildings and building parts LoD0 represents footprint and roof edge. Bridges and tunnels can be represented in LoD 1 – 4 and the underlying data models have a coherent structure with the Building model (OGC, 2011).

"In order for various applications to correctly interpret the multiple geometries of a feature, GML geometries are enclosed in an element that indicates the geometry role. E.g. a Building feature is represented through a solid geometry at LoD2" (FME, 2011). The interchange of data models best happens if data models from different domains get interrelated in a specified LoD which pretty suits the demands of interchange.

Like the CityGML model, the IFC data model applies strict separation between the semantic and geometric descriptions of objects. Thus, a semantic object can be associated with multiple geometric representations (Zhang et al., 2014). *"Although this facilitates multi-scale modeling providing different LoDs in the geometric part, yet, the integration of the multi-scale concept in the semantic part and the explicit definition of refinement relationships is lacking so far"* (Borrmann et al., 2014b, p 7).

From the interoperability point of view, LoD indicates the minimum modelling requirement, the level of certainty required for a particular object and its belonging or related elements to make and fulfill a predefined meaningful interpretation in accordance to the documentation. An example of this exists in Minimum Modelling Matrix of USACE11: in LoD 100 and 200 "… elements shall be depicted with necessary intelligence to produce plans, sections, elevations and schedules ...". LoD 300 includes "Accurate to the degree dimensioned or indicated on contract documents".

A project can be modelled as in an indoor aspect and as in an outdoor aspect throughout its lifecycle. With the main focus of modelling the digital world as in an outdoor point of view OGC has considered 5 levels of detail. Nevertheless in the last LoD namely LoD 4, OGC has modelled the buildings and streets up to their furniture object level. This makes it the interface to interior design of the project.

On the other hand most of the CAD as well as BIM programs especially IFC focus mainly on design of the project interiorly. Therefore it is obvious that the designed concept of LoD by OGC might not be sufficient for representing the entire IFC / BIM model. Thus, an enhancement is needed so that there is one general LoD concept which can be used for the purpose of interoperability. As shown in figure 14 Eric Lebègue (2005) introduces the following LoD model for the infrastructure objects:

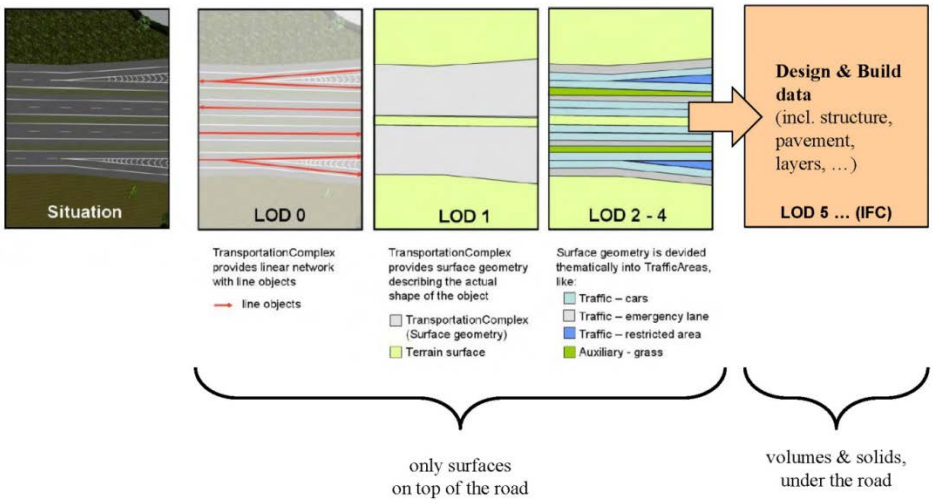


Figure 14 CityGML – Transportation Objects – LODs → IFC, (Eric Lebègue 2005)

"There are important differences in the LoD approaches taken by GIS applications and infrastructure design. In cartography, usually fine-grained data is captured and generalized to create coarser levels of detail (bottom-up approach), whereas in infrastructure design, planners start with a coarse representation and add more and more details (top-down approach)" (Borrmann et al., 2014a, p 18). Borrmann et al. (2014a) have introduced LoD concept for tunnels as well which is shown in figure 15.

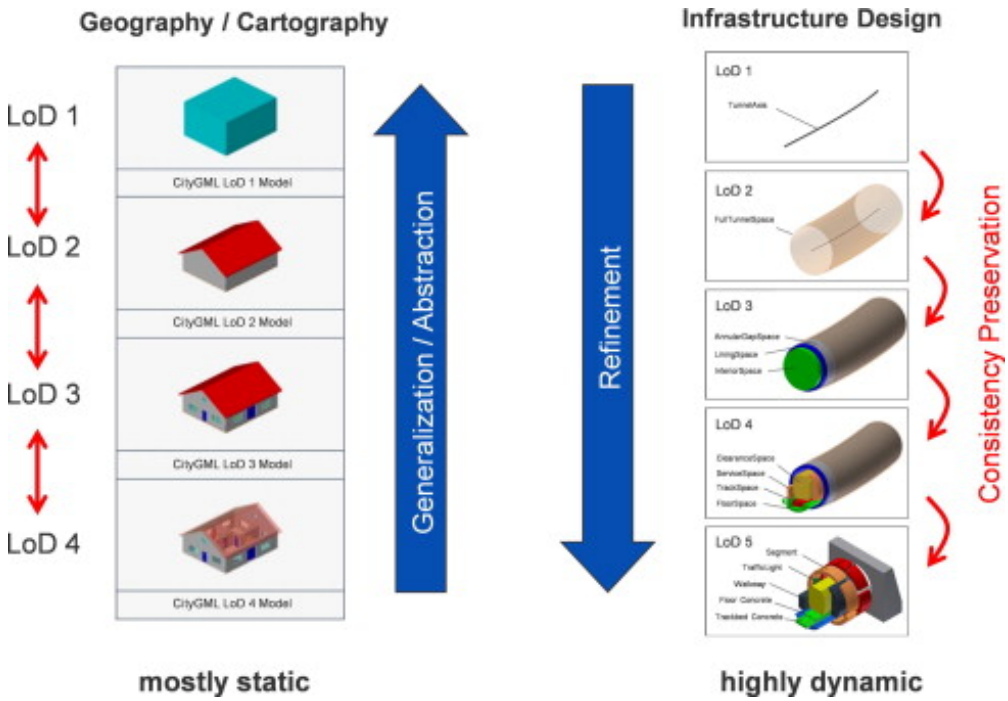


Figure 15 Top-down approach for LoD concept of infrastructure design, (Borrmann et al., 2014a)

As shown in figure 15 in this concept Borrmann et al. (2014) introduce 5 different LoDs. LoD 1 to LoD 4 consist like all of GIS applications spatial elemets. LoD 5 contains physical objects which are presented in the finest level. At the same time, the concept maintains the aggregation relationships across the different LoDs in order to explicitly model a refinement hierarchy.

2.3.3 LandXML and CityGML

"On May 2004, OGC has announced its first interoperability experiment involving the Geography Markup Language (GML) Version 3.1 and LandXML Version 1.0. OGC Interoperability experiments are "brief, inexpensive, low-overhead initiatives led and executed by OGC members to achieve specific technical objectives that further the OGC Technical Baseline; three or more OGC members

launch and run an initiative without the more substantial sponsorship that supports OGC's traditional test beds and pilot projects. These initiatives can be for specification development, refinement, or testing or for other purposes". LandXML Version 1.0 is "an industry- driven, open XML data exchange standard that provides interoperability in more than 40 software applications serving the civil engineering, survey and transportation industries. The LandXML.org Industry Consortium, developing the standard. LandXML is becoming broadly supported in online cadastral applications, GIS applications, Survey field instruments, Civil Engineering desktop and CAD-based applications, instant 3D viewers and high end 3D visualization rendering applications" for American standard based planning as initiated by US Army Corps of Engineers. LandXML XML Schema root nodes include Alignments, Application, CgPoints, Coordinate -System, Grade- Model, Monuments, Parcels, Pipe-Networks, Plan Features, Project, Roadways, Surfaces, Survey, and Units. OGC's Geography Markup Language (GML) is a "widely supported open specification for representation of geographic (spatial and location) information. It defines XML encoding for the transport and storage of geographic information, including both the geometry and properties of geographic features" (Technology Report of OGC).

LandXML has been developed over the last decade or so by volunteer organizations and individuals, outside any standards authority. The objective is to bring about comprehensive integration of survey, civil engineering and other land-based data in CAD and GIS databases. This would promote non-proprietary data transfer within and between these industries and could serve as a common standard for recording topographical features. Organizations wishing to be involved in this process need to be members of OGC[12]. Therefore as in figure 16, openINFRA considers and works on CityGML and LandXML to be interoperable standards.

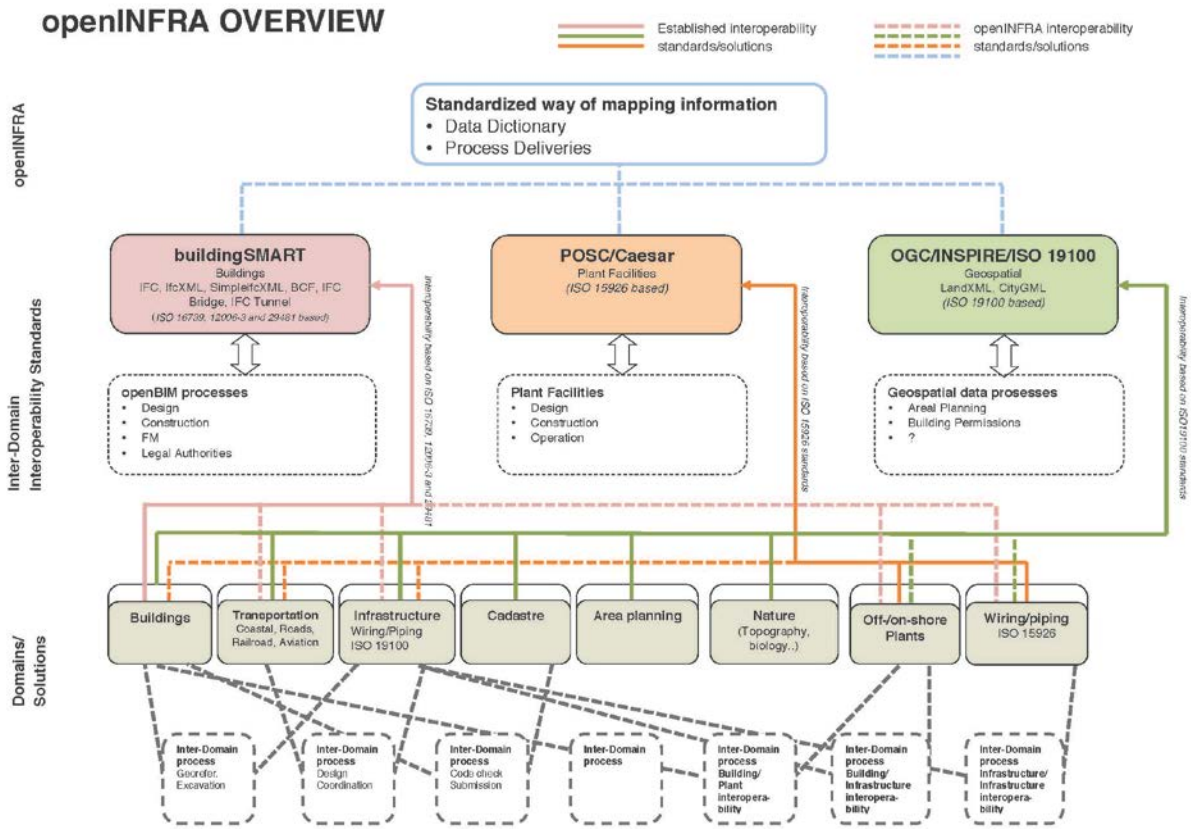


Figure 16 OpenINFRA Overview, (BuildingSMARTnordic[13], 2013)

OpenINFRA will in the first stage support and include the domains listed in table 2.

[12] www.pvpubs.com/archives/files/pdf/GW2012NovDec_06.pdf

[13] http://www.rym.fi/en/programs/builtenvironmentprocessreengineeringpre

Table 2 Domains of OpenINFRA (BuildingSMARTnordic[14], 2013)

1	Traffic
2	Geodetic and topography (and DTM)
3	Geotechnics
4	Hydraulic
5	Utilities: power supply
6	Bridges
7	viaduct
8	Geometry (road, train, channel)
9	Drainage
10	Interchanges
11	Train equipment
12	Rest & services area
13	Pavement
14	Traffic management
15	Utilities: water supply
16	Irrigation
17	Fixed operating equipment
18	FOE Civil Works and Services
19	Street Lighting
20	O&M buildings
21	Landscaping
22	Environmental issues
23	Meteorological
24	Safety equipment
25	Tolling equipment
27	Geotechnical studies
28	Hydraulic studies

In the case of infrastructure projects, design is not the only potential field for the interoperability. Safety, mobility, environment, energy assessment are other fields of application. Data ownership, privacies, liability issues, capture process and handling the data are also nested topics. In addition, the collaboration with the OGC has produced terrific and useable results which enable comprehensive CAD-GIS-CAD data exchange using LandXML and GML. In addition to international requirements and collaboration with adjacent standards bodies, the intent is to make adjustments and add support for additional data based on real world data exchange by LandXML supported applications.

LandXML has been chosen for standardization of information exchange in road; railroad and water planning in most of the countries. To facilitate the exchange of data across multiple GIS systems and indirectly across multiple views of the built environment the OGC has promoted the CityGML standards. De Nier R. (2013) and Hyvärinen (2012a and 2012b) comprise the LandXML and OGC as follows:

LandXML:

- LandXML available as open specification and implemented in major software applications (+)
- Is well suitable for exchange of surfaces, alignments and profiles (+)
- Can pretty handle geometric / parametric road and railway design (+)
- Can support several business cases b/n design and construction (+)
- The schema rather flexible (open for interpretation) and its documentation not very informative (+)
- Preferably international (+)
- Model view definitions with implementation guidelines needed (-)
- Is not appropriate for exchange of the 3D models (-)
- Fails in dealing with semantic information (-)

A lot of current limitations as well as the data loss in export / import between applications are software related.

CityGML:

- Is strong in converging heterogenic descriptive models (+)
- Uses generic modules for transportation networks (+)
- Does not cover the detailed functional aspects of transportation network models (-)

[14] http://www.rym.fi/en/programs/builtenvironmentprocessreengineeringpre

- Has a very abstract definition of the transportation network and does not contain explicit description of transportation objects (-)

2.4 INTEROPERABILITY AND MULTIMODEL TECHNOLOGY

"The GIS interoperability strategies are not mutually exclusive. Data managers and GIS practitioners may use them in complementary ways, as appropriate for specific environments or data holdings. Initially vendors are likely to develop interoperability among their own products. As pressure from specific communities grows and as the underlying technology matures, both vendors and third parties can be expected to offer true, multi-vendor interoperability solutions. GIS interoperability developments are significant for another reason as well: they are contributing to the migration away from the monolithic systems which have dominated the GIS market for so long. Databases, browsers, smart translators, and geoprocessing tools are now being coupled through adherence to open standards and specifications. The backbone of much GIS activity is likely to be network-accessible databases, with connectivity requirements met by common interfaces" (Sondheim et al., 1999, p 357).

Major interoperability aspects include the ability for a given user-defined schema to relate to standardized geospatial concepts such as entities and regulations, representation, and multiple user-defined schemas relations.

Geospatial standards deliver a high degree of semantic reliability insofar as essential geospatial conceptions are concerned. Further to these concepts, however, semantic consistency is an issue. Often a more practical alternative is to build a series of model-to-model conversion, so that the recipient gets a certain schema in a most appropriate way for him. The second approach is more flexible but merely possible with a detailed knowledge of the respective schemas and schema transformation software.

Fuchs believes (2013), according to the transition from building product centric to process centric way of working, cross-domain information supply gains importance, e.g. in the fields of construction management, holistic design, simulation or energy efficiency. However current domain models cannot perform this challenge on their own. Multi modeling is a method to set up arbitrary exchangeable BIM using heterogeneous domain models and explicit external links between their content elements.

Scherer et al. (2011) have introduced a distributed multi-model-based management information system for simulation and decision making on construction projects. In a multimodel (MM) the heterogeneous data models are bundled together keeping their original format in to a container a so called "link model" where the elements of data model will be linked together. The interlinking of the elements to their metadata will remain untouched. The original data, their metadata and the link model will be saved in and exchanged through a container illustrated in figure 17 (eg. ZIP file). MM is an approach to support information analysis and collaborative work over multiple application domains.

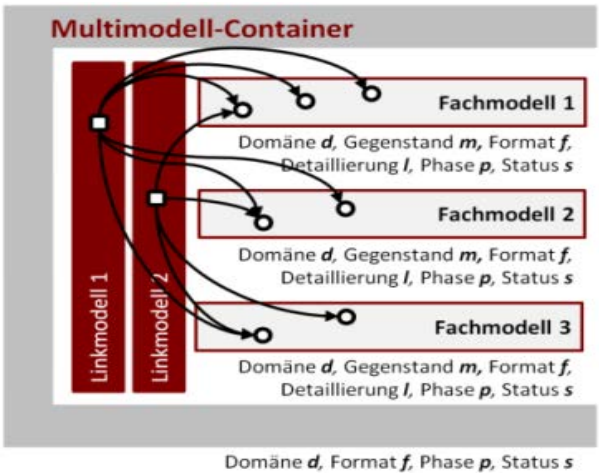


Figure 17 Multimodel Container of Mefisto Project, (Scherer and Schapke, 2011)

"MM allow for joining discrete information from different software applications to inspect the coherency of related application models and create complementary analysis models. Moreover, they

can also support the coordination and documentation of the processes for MM creation and utilization" (Scherer & Schapke, 2011). Where model links were used to link between specified data models such as BIM and climate data, the traditional link models of MM before 2000 were used to serve for metadata integration and create links between BIM model objects, metadata sets and documents. This is illustrated in figure 18.

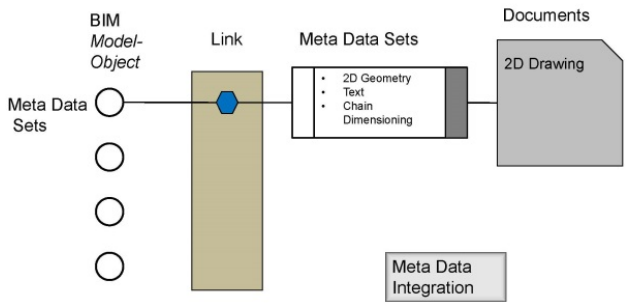


Figure 18 Documents Link Model, (Scherer and Kadolsky, 2014)

According to Schapke et al. (2012) *"the combination of application models in MM provides a basis for a general methodology to inter-relate, compare, analyze and re-use any kind of information throughout construction projects"* (Scherer & Schapke, 2011). Thus, interchangeable cross-domain information sharing environment results, which allows:

- Further use of existing software
- No need to create new data formats
- Obtain cross-domain information

As a part of eeBIM for the HESMOS Project Scherer et al. have introduced a Link Model where the key data of heterogeneous data models such as climate or material data models get transformed into property sets. BIM model-objects get interlinked with property sets. This link model as shown in figure 19 got promoted in the Mefisto Project and used for schema integration of data models such as BIM and HVAC. Here the original models remain unchanged.

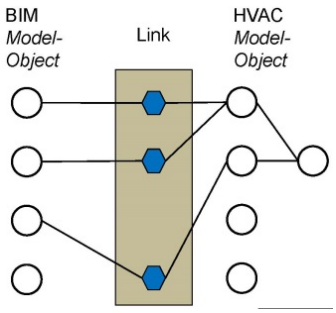


Figure 19 Mefisto-Link Model, (Scherer and Kadolsky, 2014)

The core in a link model is how to use the technology to interlink the data and how to analyze the linked data models to obtain cross-domain information. Scherer et al. have developed a multi model (Scherer & Schapke, 2011) which is equipped with an interlinking language, a query language (Fuchs, 2012) and a graphic which shows the interlinking, exchange of data, and consequently cross domain filter. Figure 20 gives a better understanding how such model will work.

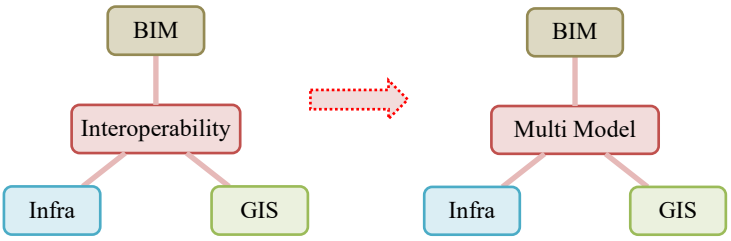


Figure 20 Interoperability of BIM – Infra – GIS in the context of Multimodel

All these interoperability approaches demand a big number of organizations involved, huge range of work done and will be a long term approach for the IFC developing instates. Therefore it would not be in the framework of this thesis, though the key purpose of this thesis is to create an integration process for the GIS and infrastructure models according to the available resources. This thesis therefore is investigating a direct interoperability approach not through BIM domain between GIS and Infrastructure domains.

Multimodel is a method where heterogeneous data models from various domains are bundled together into a container keeping their original format. Therefor this thesis will provide its interoperability approach between GIS and Infrastructure in Multimodel context. For this reason this thesis will study either the existing Multimodels are appropriate for the interoperability purpose for GIS and Infrastructure data models.

With the aim of interoperability Multi Model (MM) is a method where heterogeneous data models from various domains are bundled together into a container keeping their original format. In separate Link Models the elements of the data models will be linked together" (Fuchs, et al., 2011). "MM uses deterministic record linkage which generates links based on the individual identifiers that match among the available datasets. This may create a huge number of links containing a set of identifiers representing concrete entities of a certain data model. Yet, a small decrease in data quality or small increase in data complexity can result in a very large increase in the number of commands necessary to link records (Esfahani, et al., 2014). MM handles just pre specified geometries which results in very limited visualization environments. More importantly the approach lacks in handling the real time changes on elements and dynamic data structure due to being based on predefined ready to use data models. Thus, according to Liebich et al. (2013) using a separate Link Model as a bridge between BIM and non-BIM data can provide generality and modularity for implementation scope.

The OGC has established a Land Information Domain Working Group with the aim of bringing the existing LandXML model and schema under OGC's wing. LandXML / LandGML is a standard for land management, surveying and cadaster, providing a semantic model for parcels, land use, transportation and pipe networks (LandXML, 2001). But Kolbe et al. (2005) believe *"although LandXML supports 3D coordinates; it does not comprise volumetric geometries. Buildings are only represented by their footprints. Further concepts for 3D man-made objects are missing"* (Kolbe et al., 2005).

2.5 LIMITATIONS OF EXISTING APPROACHES

As mentioned previously, various attempts have been made for integrating BIM with GIS in three main categories: at application, process, and data levels.

Interoperability happens at the application level via reconfiguring or rebuilding where software patches are used to either adjust an existing GIS or BIM tool or embed the functions of one in the other. In general this approach is expensive and robust (Amirebrahimi et al., 2015).

Integration at the process level use SOA to enable both BIM and GIS systems participate in tasks that need both systems competences while keeping them separate. Rigidness of this approach is less than the first one. In this approach to be able to work out, the underlying data levels need to understand interoperability between these systems (Amirebrahimi et al., 2015).

There are lots of approaches developed to enable interoperability between BIM and GIS at the data level. In high level and more flexible comprehensive integration approaches, either a new "meta data model" is developed to arbitrate between BIM and GIS systems at higher level; or the data models at GIS or BIM systems side gets extended to manage and handle the data from the other side (Amirebrahimi et al., 2015).

Borrmann (2012) indicates that *"there are many different standards developed for the structural sector. However, in order to realize BIM within the civil engineering domain civil specific standards are necessary"* (Borrmann, 2012, p 3). According to Esfahani et al. (2014) the main focus of IFC up to now has been the building schemata; IFC does not yet cover the infrastructure data models completely and includes neither infra domain nor any extension dealing with infrastructure data.

According to the experiences, it is obvious that various algorithms are under development for different application fields such as spatial information, controlling, energy efficiency, etc. *"First infrastructural product models and first proposals how to define a standardized infrastructural DBB process already exist"* (Obergriesser & Borrmann, 2012, p 584). Borrmann believes that *"implementation of a BIM method within the infrastructural designing, bidding and building industry needs different standards in order to be accepted by engineers, software developers and executive users. Only after this process infrastructural projects can be successfully completed"* (Obergriesser & Borrmann, 2012, p 586).

A simple prototype implementation of infrastructural DBB has been done using commercially available ORDBMS Oracle 10g (Borrmann & Rank, 2009). In this a 3D SQL project has been developed which enables the spatial analysis for BIM and the extraction of partial models that fulfil certain spatial constraints (Borrmann & Rank, 2009). According to Borrmann (2009) and Adachi (2003) *"in existing query languages for BIM, such as product model query language of the EuroStep model server and partial model query language of the IFC model server* (Adachi, 2003)*, the utilization of spatial relations within a query is limited to simple containment relationships predefined in the product model, mainly due to the structure of the underlying BIM which doesn't incorporate the explicit geometry of the building components"* (Borrmann et al., 2007, p 105).

Expecting that spatial modelling and processing play an increasing role in future engineering systems, Borrmann and Rank (2008) have developed a simple spatial query language for BIMs as a first step towards higher spatial concepts. In their opinion *"the partial model resulting from a spatial query may serve as input for a numerical simulation or analysis, or might be made exclusively accessible to certain participants in a collaborative scenario"* (Borrmann & Rank, 2009, p 1).

"The most important advantage of using an object-relational approach is the strong type safety provided by the declaration of user-defined types. E.g. the declaration of the touch member function forces the passed parameter to be of type SpatialObject or one of its sub-types. Thus, type errors may already be detected by the query engine during the interpretation of the SQL statement and more specific error reports can be generated" (Borrmann, 2010, p 25).

In most of these studies targeted application areas explicitly include urban and landscape planning; architectural design; tourist and leisure activities; 3D cadasters; environmental simulations; mobile telecommunications; disaster management; homeland security; vehicle and pedestrian navigation; training simulators; and mobile robotics with different data model architectures. The focus of the integration was to develop an intermediate data model for building, building plan submission process and its facility management. In general, these models are application-focused and the integration is made for a particular use case with specific requirements. Therefore, the included concepts and relationships within these models may not suit other infrastructure applications.

On the other hand the recent attempts to add infrastructure domain to IFC through IFC-Bridge and IFC-Road, backbone standards for infrastructure, are still in the development phase, thus there is currently no integrated IFC product model for alignment, bridge, building, grounds and site installation. Therefore the interoperability between GIS and Infrastructure cannot be covered by interoperability between GIS and BIM.

In the field of infrastructure domain there are many data formats, which are mainly due to the diversity of labor and the organization of the planning process as well as the construction industry. The existing methods which enable direct interoperability between GIS and infrastructure domain are either limited to those application based interaction possibilities developed for a specific part where GIS and Infrastructure overlap or the conversion or translation of the data domain of each data model in the other one such as traffic data model of GIS or topology / surveying data model in infrastructure domain.

Therefore due to the lack of or poor interoperability in infrastructure domain and in the absence of an internationally accepted alliance for undertaking the interoperability tasks, each software vendor defines its own domains catalog or objects modelling in accordance to their vendor and customer demands.

Thus this thesis chooses the MM context the most suitable method to enable interoperability between these two domains. MM is one of cross domain interoperability approaches which enables

heterogeneous data models from various domains get bundled together into a container keeping their original format.

MM uses deterministic record linkage which generates links based on the individual identifiers that match according the join criteria defined using MM query language. This may create a huge number of links containing a set of identifiers representing specific entities of a certain data model. Yet, a small decrease in data quality or small increase in data complexity can result in a very large increase in the number of commands necessary to link records. Thus existing MM gets to its limits in handling dynamic data resources, ID-less data models, or dealing with a huge amount of inter model relations (Esfahani et al., 2014).

Thus, MM and the generated links have no inherent domain semantic. The links created this way regardless of their explicit or implicit types contain no semantic and are static. This is due to the fact that the target data models of MM are static in nature and the links are created after the data models are finalized. As soon as any data is re-modified, the links should get updated. Besides MM is limited in handling data models which is changing constantly (Esfahani et al., 2014).

One of the other limits of existing MM is managing a huge number of links with identical interpretations.

On the other hand the key enabler in a MM is the ID of the data which is the main criteria to store, access and link to the data from data models. Yet the GIS data models lack in having such systematically structured data models. In GIS data models the data structure and the spatial index is the more useful identity of the objects.

Thus this thesis introduces a new approach to overcome the above mentioned limitations and consequently facilitate interoperability between GIS and Infrastructure in a cross domain approach.

Chapter 3

Infrastructure Information Modelling

Information demand in infrastructure information processes is not fulfilled by the existing data models, because the data models involved in these processes are separated. There are specific approaches for joining heterogeneous information spaces such as Integrated Systems (IS) or Multi Models (MM). The existing approaches are not applicable on some information processes of infrastructure domain. Data of IS are not exchangeable and existing MM can't cope with dynamic data sources. Therefore an interoperability method is required for coupling heterogeneous information spaces loosely and dynamically.

Infrastructure Information Model (IIM) is an extension of MM in two ways: a) in semantic way and b) in rule based way. IIM is an instantiation of a rule based MM.

This chapter describes how the hypothesis of this thesis gives an extension for the existing MM to a rule-based Multimodel named extended Multimodel (eMM) with semantic rules – instead of static links. The semantic rules will be used to describe relations between data elements of various models dynamically in a link-database

Matching rules instead of static links are semantic with respect to the aforementioned information demand. Considering the infrastructure domain, it can be stated that the spatiality is a common aspect in most information processes. Therefore spatiality is chosen to be the matching semantic.

LoD is used for the abstraction purposes in the rule based match finding.

3.1 MULTI MODEL FOR GEOSPATIAL AND INFRASTRUCTURE DATA MODELS

For infrastructure domain the next generation of MM functions as to link an element from infrastructure design model such as road by the semantic "spatial" with an element e.g. terrain from the geospatial information model sharing the same spatial index as illustrated in figure 21.

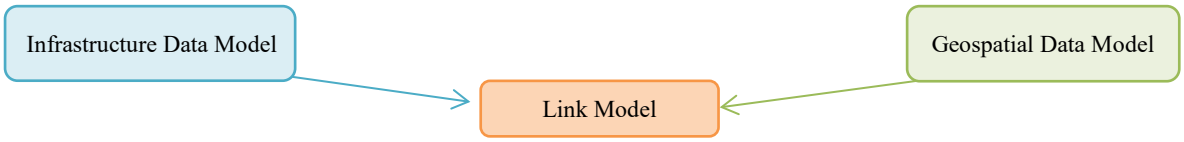


Figure 21 Concept of Link Model for Infrastructure and Geospatial Data Models

The characteristic of the multi model is to understand the correlation between Infrastructure data and the external information in separated data spaces, the Link Model. Model management tools can be then used to decode the correspondences. A Link Model:

- does not require changes in the BIM schema and the external models used (+) (Katranuschkov et. al, 2014)
- guarantees maintenance of each model within its own domain (+) (Katranuschkov et. al, 2014)
- provides greater semantic depth (+) (Katranuschkov et. al, 2014)
- handles almost arbitrary data structures (+) (Madrazo et. al, 2013)
- enables a clear interoperability strategy (+) (Madrazo et. al, 2013)
- has difficulties regarding maintenance of Link Model (-)
- needs additional Link Model management services (-) (Madrazo et. al, 2013)
- some run-time performance deficits due to the increased data complexity (-) (Madrazo et. al, 2013)
- possible consistency problems in the rare case of overlapping Multi Model data (-) (Madrazo et. al, 2013)

Therefore, whilst possible for any multi-model problem, the Link Model approach is most useful where:

- a large amount of external information resources is needed and these resources have non- AEC origin, (Madrazo et. al, 2013)
- a flexible platform for a set of (exchangeable) software tools is sought. (Madrazo et. al, 2013)

Multimodel takes the advantage of being cross domain and handling queries and filters and interrelates different (any type) data models together. MM has been applied up to now on tabular data with IDs as attributes and can join the data through their IDs which is a unique key for accessing one and just one entity. Yet MM is not developed to enable establishing semantic links. In the case of infrastructure and geospatial data models, the link model may operate on the IDs, yet the generated link will not have logical value for the interoperability purpose.

Contrarily to deterministic record linkage used by MM, link methods based on probabilistic or fuzzy matching use another approach for linking the entities. They consider a wider group of potential identifiers, then according to estimated capability of finding a match or a non-match weights will be calculated for each of them. These weights result in computing the plausibility that two given entities relate.

MM and the generated links have no inherent domain semantic. The links created this way regardless of their explicit or implicit types contain no semantic and are static. This is due to the fact that the target data models of MM are static in nature and the links are created after the data models are finalized. As soon as any data is re-modified, the links should get updated. Besides once any of the data models is on constant change, MM hits its limits. While the key issues in MM is to facilitate interoperability for heterogeneous data models, enable query and filtering the data in their original

formats, and to create links independent from the input data models, in the next MM generation the links must be created through matching rules. (Esfahani et al., 2014).

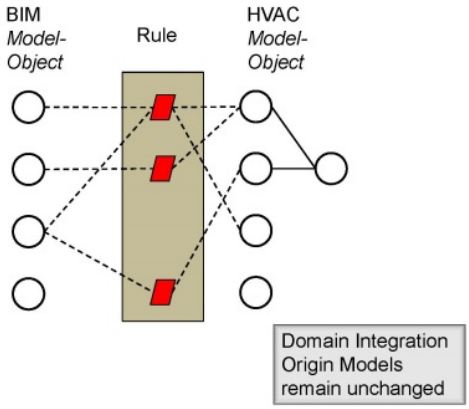


Figure 22 Rule based Link Model, (Scherer and Kadolsky, 2014)

As shown in figure 22, the original models remain unchanged. Such a rule based link model is used to facilitate domain integration. This rule based link model will be generated by the rules. The rules are established by the user and therefore the rule based link model is scalable only if the operator who sets the rules which are in charge of creating rule based links facilitates scalability. Otherwise the scalability is facilitated via creating new rules.

Figure 23 illustrates the next generation of MM where the explicit links will be replaced by one or more linking rules for the matching purpose. As explained and emphasized in detail in previous chapters the most important identification of infrastructure and geospatial data models is the spatial property of both of them. The spatial identification is the basic element that both infrastructure and geospatial data models share in common. Thus, spatial identification is the semantic which links the data in spatial domain appropriately. Through spatial properties, Spatial Links (SL) will be created which are the fundaments in infrastructure processes.

A spatial link model should operate on spatial property and generate interlinks through spatial identity of entities. Thus, for a link model to be able to fulfill the requirements of match finding in a geospatial context, the link engine should understand a spatial query language. This spatial ready link model will be called “Spatial Link Model (SLM)” and the engine which generated the spatial links through spatial identities will be called “Spatial Link Engine (SLE)”.

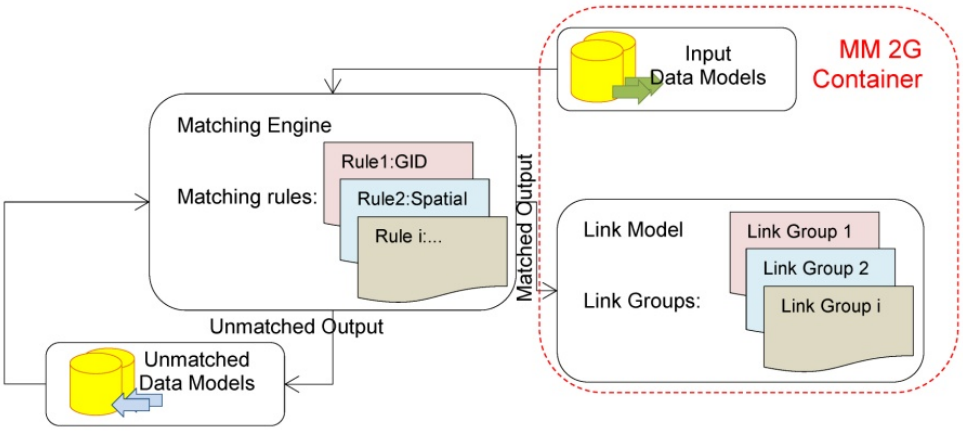


Figure 23 Linking rules replace links in the eMM, (Esfahani et al., 2014)

As shown in figure 23 in the eMM the data models will be input into a container. A matching engine which contains matching rules starts the process of match finding. These matching rules include sharing same ID (as developed in the MM of Fuchs et al. (2013)), or same spatiality – spatial indexes (which is developed in the scope of this thesis), or same temporality – temporal or time based indexes (as suggested for the future work of this thesis), or those who share common security, or being in the related disciplines of the related LoD.

In the process of match finding, according to the rules and inserted data models, some data sets will get matched to some other, those which are matched to each other will get linked together. These rule based linked data will transferred into the rule based link model. The unmatched data will get inserted back to the rule based link engine. The match finding process will change the match tolerance if needed and to the process runs another time. After each match finding loop, the matched results will get inserted into the link model as a new link group. The match finding process keep repeating till the match tolerance reaches its minimum. Therefore as shown in the As shown in figure 24 this spatial link model is an extension to the existing MM and can act as a bridge to transfer the spatial links generated through the spatial link engine to the eMM which handles the normal tabular links.

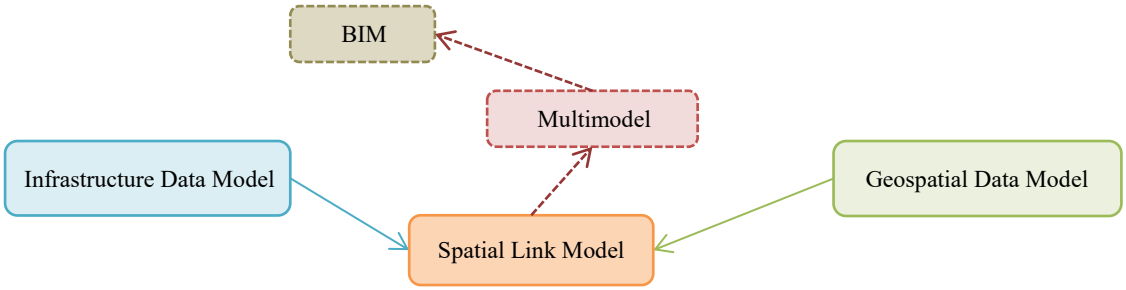


Figure 24 Spatial Link Model in the eMM

The Spatial Link Model is a container for the joint result of data which are implicitly matched. These links can get stored in an MM for the exchange purposes. Once the links are saved in a multi model explicitly, they will be capable of making dialogue with any BIM domain which can be interlinked in an MM (Fuchs, 2015).

Infrastructure and geospatial data models are very well-known examples and fields of application for spatial linking, yet spatial links can be created between any data domain with spatial nature such as geology, hydrology, climate and environment, etc.

Spatial links integrate all spatial data associated with planning, designing, construction and management into a single comprehensive platform. Each data model remains in its original format and interacts with the other domains through spatial properties. Such approach facilitates to efficiently maintain the entire infrastructure project throughout its lifecycle. This approach is called Infrastructure Information Modeling (IIM) which is an instantiation of the generalized MM.

IIM integrates all the elements involved in a project in 3D facilitating clash check at design stage itself thus according to Sutech Solutions saving up to 20% on construction cost, 45% on design cycle time and 50% on interdepartmental site meeting to clear clashes. Integration of the different models in a cross-domain, neutral from infrastructure model type gives the possibility of dynamic specification of the analysis model, runtime-related networking, as well as cross-domain analysis.

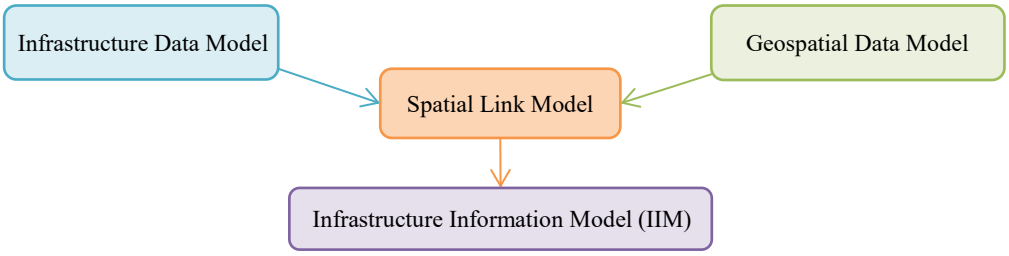


Figure 25 Concept of Infrastructure Information Modeling

IIM is a future ready multi-dimensional cross domain interoperability approach to plan, visualize, analyze, simulate, execute, manage and maintain an infrastructure project more efficiently and more effectively throughout its lifecycle. As the spatiality is the property that the infrastructure data models share in common, it is used for the match finding process in the IIM as demonstrated in figure 25.

3.2 LINKING APPROACH, QUERYING AND FILTERING

Link Models as a part of the Multimodel container are generated to distribute project information among semantically and structurally inhomogeneous domain expert models from one side and on the other side to externalize the implicit relationship between the data elements. The existing Link Model of MM, introduced by Fuchs et al. (2013) is capable of creating structural links. Yet functional links which are created based on fulfilling one or more predefined functions have not been considered.

In the context of interoperability between data models where the content constantly changes or in regard of data domains where the data model itself changes steady, it becomes inevitable to define functions or regulations which create links among the data.

Especially talking about GIS domain mainly with the purpose of modelling the digital representation of the world, there are many subdomains such as climate, energy, pipelines free water bodies and etc. which are not yet entirely included. Therefore one can say that GIS data model is a model on constant change both from content point of view as well as from data model point of view. Such data models are defined in this thesis as dynamic. This dynamicity can either be in the content or in the data model itself.

In such dynamic data models some of the missing subdomains might be found in other data models. For example pipelines are included in some of infrastructure data models like OKSTRA. Yet the problem might still exist that this data model (OKSTRA) containing the missing subdomain (Pipeline) might not be the one internationally accepted data model for the main target domain (Infrastructure).

Therefore it seems that the applicational semantic, in this case spatial, queries cannot be answered using existing Multi Models. In the next generation of MM, namely rule based Multi Models, the functional links will be capable of answering such queries.

In such an approach the functional link 1 (L_1f) queries the objective (Water Pipe(s), Gas Pipe(s), Wastewater Pipe(s), etc.) from the subdomain (Pipelines), regardless of in which data model (OKSTRA or LandXML) to find it. There will be an answer as long as at least one subdomain (OKSTRA) contains the queried item. Then as shown in eq.(1) L_1f links the objective from the subdomain with the related objects from the interface carrier data model (Building). The interface carrier is the data set which exists in both first and target data models (IFC, CityGML).

$$L_1f\ (Pipeline, Building) \xrightarrow{Functional} GIS, Infrastructure \qquad \textbf{Eq.(1)}$$

From here the functional or structural link 2 (L_2f or L_2S) queries and links the interface carrier (Building) with the missing subdomain in the first domain which yet exists in the target domain (Pipe installations inside the building). This is shown in both eq.(2) and figure 26.

$$L_2S\ (Building, Pipes\ ) \xrightarrow{Structural} IFC \qquad \textbf{Eq.(2)}$$

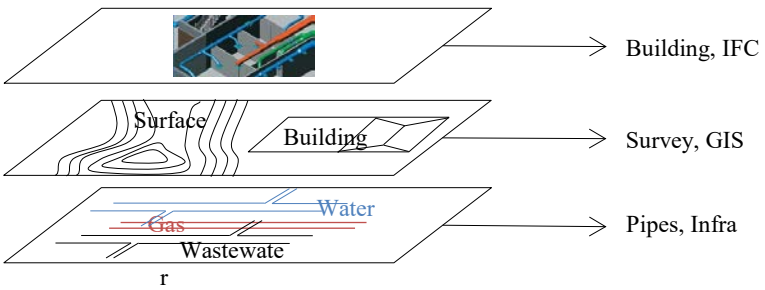


Figure 26 Relationship between functional and structural links

The rule which creates the functional link is in the case of Infrastructure and Geospatial data models the spatiality of them as explained in previous chapters. Therefore as described via eq.(3) in the next step these two links can be set in relationship to each other via their spatial indexes:

$$R(L_1, L_2) = R((L_1, LandXML), (L_2, IFC)) \quad \textbf{Eq.(3)}$$

With this approach questions which were not able to be answered in the existing MM can be answered. For example as shown in the figure 24 consider a building implemented in a BIM model such as IFC. The question is to which utilities of the city with which capacity the building connects. The existence and the type of utilities is provided and stored in the GIS data model. The capacity of the relevant existing utilities is provided or planed in infrastructure data models such as OKSTRA or LandXML. The suggested approach spatially links these two data models to provide capacity information which are stored in the landXML data model for the GIS data model as virtual property. Virtual properties will be explained in the next chapters.

As GIS is the place holder for all real objects in digital format, the discussed building above either exists in the GIS data model or can be integrated through its spatial identity. Therefore the existence of this object is not the matter of discussion but rather its spatial index. Therefore the building of IFC will be abstracted into the target LoD in which GIS data sets are provided. Thus the building will be abstracted to its outline (LoD 0 of buildings in GIS). The spatial index of this object will be defined in this step. From here the building in IFC data model will be linked with the utilities in the GIS data model and capacity information will be obtained through virtual properties from linked objects of LandXML.

Normally the IDs are used to establish links between elements. In the existing MM in the first step the data models are imported into the container, then the IDs are built for the internal use of linking procedure. Thus the main difficulty is that, if there are data models with no IDs or if there is a change or update in the elements or data models, the entire procedure from import to linking needs to be repeated and the links need to get reconstructed. As in a functional linking procedure, there is a rule which is in charge of creating links and the links are not created based on a single ID but in accordance to a certain function which relies on the properties of data models, the changes in the elements of the data model does not require to reimport the data and rebuild the IDs. Only the links need to be updated.

3.2.1 Virtual Properties via Link Model

As mentioned in previous chapter a spatial link model is the place holder for or the data model to represent the generated spatial links between matched data.

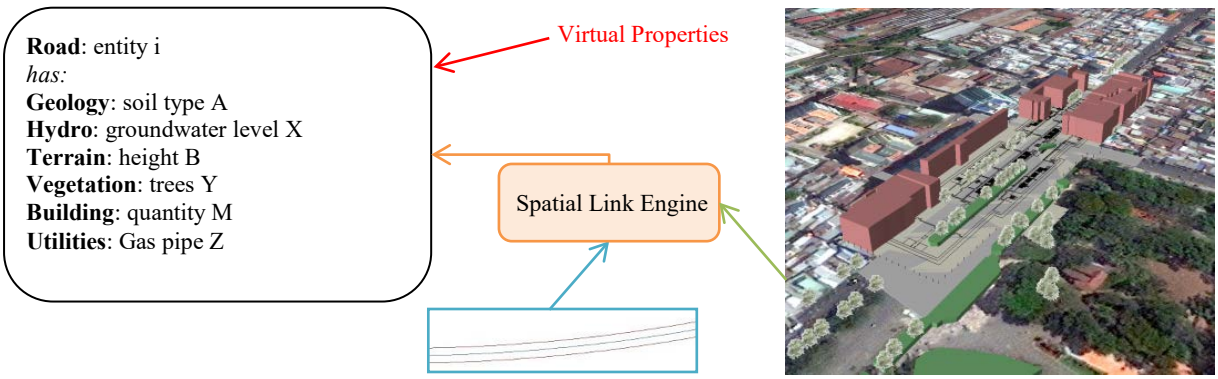


Figure 27 Virtual Properties via Link Model

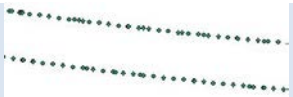
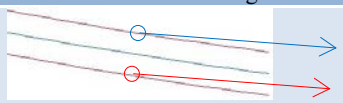
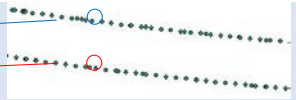
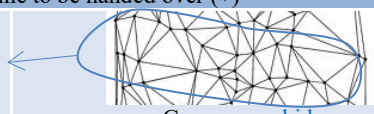
GIS is on the other hand a system where over 90% of information needed for the infrastructure design is stored and maintained. GIS contains different data types from terrain information, hydrological, geological, or climate data, to vegetation and land acquisition. While as in the practice GIS provides the most updated information for the infrastructure design task, BIM software provides the best geometry design for the infrastructure project.

Once SLE links the geometry of planning data with the geospatial information in accordance to the location of elements, each infrastructure entity receives the spatial information which is stored at the location of the entity or is related to the targeted entity due to sharing the same location (Esfahani, 2013b).

The geometrical entity which is devoid of any spatial intelligence gets through this approach all information related to the entity. This information as demonstrated in figure 27 will be the virtual properties for the object (Esfahani, 2013b).

As shown in table 3 the existing data is stored in GIS such as CityGML and the planning data is stored in infrastructure data model such as LandXML, in the first step these data models should be inserted into the eMM. In GIS spatial information, such as location addresses or coordinates are stored in attribute data tables. Yet before they can be processed in the eMM, true spatial data such as geocodes must be conducted.

Table 3 Steps and methods for creating spatial links

	Infra, LandXML, (ex. Road)	Link Model	Geospatial, CityGML (ex. Terrain)
1	 Planning data	 Empty	 Existing data
Spatial interpolation + create virtual objects			
2	 Planning data	 Empty	 Virtual objects, real attribute
Choose Algorithm (ex. Nearest Neighborhood), define rules, criteria, vocabulary (ex. equality)			
3	 Planning data	Spatial properties, Parameters, Ranges, LODs	 Link to virtual objects, real attributes
3a	MM without links, implicit matching according to rules, Less memory required(+), rules to be handed over (-)		
		 Empty	spatial attributes of geodata get allocated to infrastructure design data
3b	Save links in MM, explicit matching, High memory required (-), no need to hand over rules (+)		
	 Infra obj a, elem i Infra obj a, elem j	 Link m Link n	 Geo obj b, elem k Geo obj c, elem z
3c	Group links in MM, explicit matching, Less memory required (+), match finding engine to be handed over (+)		
	 Infra obj a, group elem i	 Link g	 Geo group obj b

In step 2 of table 3, for the spatial linking, a spatial interpolation between both data models should take place so that both data models get similar spatial network grid. This process will be explained in detail in chapter 4. As explained previously digital representation of GIS data occurs by selecting appropriate data properties for each layer with respect to projection, scale, accuracy, and resolution. Therefore creating a spatial network grid tightly relates on the accuracy and LoD of the GIS model as well as on the planning phase of infrastructure data model. After creating spatial network grid the objects of GIS and Infrastructure will be restructured in this network so that virtual objects with real attributes are created.

In the third step rules, criteria, sequences and vocabulary should be defined. In a system of orders, the rule engine searches among the existing rules for the most proper one of rules and fixes the sequence of those. The sequence of rules is strongly based on the application. The rule engine can select all of the rules but it increases the process time. Thus, this is one of the most critical challenges of the rule engine to select the right rules and the proper sequence of them. Besides, the rules can be very

complicated. This deduces the process time too. It is very important for the operator to establish simple rules, but the art in the rule engine is to find the best sequence for these simple rules. Moreover it is very important for the system to choose a specific quantity of rules. This is most of the time a compromise between process time and quality of the results.

In the proposed eMM in this thesis the engine contains only the rules themselves and the sequencing of them will be determined by the operator.

Wrong rules, wrong sequence for rules and uncomplete system of rules may produce wrong results and end to wrong statements. Therefore it is very important that the rule system be validated and tested according to plausibility.

A sequencing mechanism including not only the rules but also the structuring of the rules meaning defining proper sequence algorithm of rules and plausibility check of them should be developed in the next versions.

The procedure of creating links according to match finding process and setting proper tolerances is performed in this step. The exact explanation on match finding procedure and detail information about match tolerance is provided in chapter 4. Creating match finding vocabulary such as limiting the fuzziness is also executed in this step. Detail description about this step is also provided in next chapters.

"Both attribute and spatial data is stored in the same RDBMS (such as Oracle, which supports SDE). This allows mass data capabilities, security and data integrity mechanisms of the RDBMS to be applied to the spatial data" (Briggs, 2015).

After match finding is performed in accordance to the spatial properties, parameters, search ranges and LoDs, virtual properties will be linked to the objects. Up to this step eMM does not contain any links. Implicit matching is performed according to the rules. Such implicit matching requires less memory but to create same linking results, rules and their sequence must be handed over.

The links can also be saved explicitly in eMM. Explicit matching will require higher memory, yet the links can be exchanges and handed over. In this case there is no need to exchange and hand over the rules and their sequences.

A third option is to group the links in eMM. In this case even though less memory is required, the entire match finding engine (rule engine) should be handed over to create same results.

Thus, the technical solution of link models gives the possibility to obtain virtual properties for non-semantical data. Once matches are found, there are three different possibilities to handle the links:

- generate links but don't save them in the MM, which means to create implicit matching according to rules
- generate links and save them in the MM, which facilitates explicit matching, and accessing to the linked objects
- generate group links which means an object from the infrastructure domain gets linked to a group of objects in the geospatial domain which shares the common spatial search criteria

In the case of implicit links, the rules through which the links are generated should be exchanged, if an exchange of data is required. The querying rules need to get applied every time there is a need to access to the information related to a target object. As the linked data will not be stored in the MM, the data size will be small.

Contrary to implicit links, explicit links might need a huge size of memory depending on the project size, because the generated links will be saved in MM in a tabular txt or csv format. In such case, the rules of match finding do not need to be exchanged and accessing to the linked data is fast.

A combination of both of these methods facilitates to link a group of objects from the target data model to a specific entity from the source data model. The virtual properties obtained through this method might not be precise enough but yet the data size is small. The querying rules can also be delivered to make the match finding finer if more accurate results are needed.

3.3 MULTI MODEL AS AN INTERDISCIPLINARY METHOD

As mentioned by Fuchs (2013) *"MM is an interdisciplinary approach which facilitates to overcome the limit of applicability of a single model's domain schema. The task specific combination of established data models allows their optimized usage in relation to semantic practicability and user acceptance. The achievable multi model information is limited to the application domain and the implemented linking instructions of those products. Therefore a neutral method will be developed to create n-ary multi-model links between arbitrary domain models. This allows the realization of the full potential of the multi model life cycle for a wide range of planning information processes: the creation, exchange and retrieval of task specific information spaces, based on established data models"* (Fuchs, 2013).

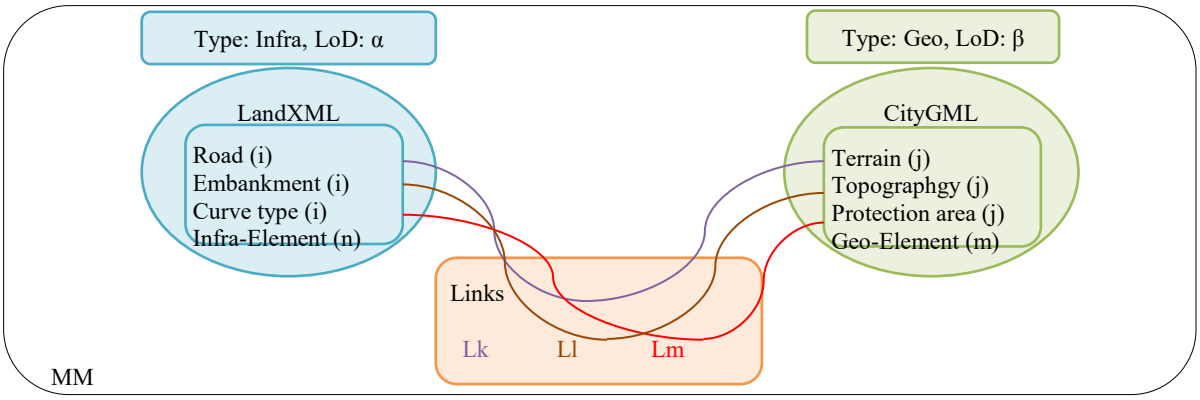


Figure 28 Linking of elements in an explicit infrastructure multi model

As shown in Table 3, part 3 subjacent construction information process result into systematic creation of links of the multi-model developed in the scope of Mefisto project. The process defines the relevant domain dependent rules. The rules will be then generalized and stored in a rule container formally, for example in an ontology system or inside a technical software system. This multi-disciplinary approach defines the rules which describe the potential links between element types and their domain models as well as the special circumstances like attributes value ranges (min, max, deviation …) required to stablish the link between entities. Illustrated in Table 3, part 3a, a common vocabulary containing the level of detail, project phase or link types amongst others is needed to be used for metadata in multi-models. As shown in Table 3 part 3b a generic access to the heterogeneous data models in a multi-model environment where the rules are run must be provided. The element combination mechanism developed in the scope of this thesis is based on setting up the linking space according to potential matches fulfilling the search range. The elements of interest will be then filtered based on match refining approaches and undesired combinations will be omitted. As each link is covering elements meeting a range for a query, the links will be potentially n-ary which build group of links. The original data models together with the resulted link model establish the rule based multi-model for the task based construction information process (Fuchs, 2013).

The concept of "Infrastructure Information Modeling" in the context of "Geospatial Modeling" requires linking the infrastructure data model (example alignment) in a specified stage of design (e.g. preliminary design stage) with geospatial information which is filtered through filtering methods in Spatial Link Engine. Shown in figure 28 these filtering rules are defined in analogy to the design demands in a rules repository in e.g. client's software system. Data manipulation occurs afterward using spatial linking approaches such as Nearest Neighborhood Algorithm. The benefits of such cross domain approach are:

- Neutral Data Exchange
- Reuse of existing applications
- Integration into MM approach

In the eMM links will be replaced by the linking rules and yet the rules must be exchangeable to be interoperable. An abstract interoperable representation is demonstrated in eq.(4):

$$Matching\ Rule = F\ (Matching\ Approach, Match\ Tolerance) \quad \textbf{Eq.(4)}$$

In this thesis one of such matching rules, namely Spatial Matching which is the key enabler of the interoperability between geospatial and infrastructure data models is discussed in detail in chapter 4.

3.4 USING LEVEL OF DETAIL (LOD) FOR FILTERING

"*As planning large infrastructure facilities requires intense collaboration among the numerous experts involved, a periodic handover of planning information has to be performed as soon as well-defined planning stages are reached"* (Borrmann et al., 2014b, p 3). This information is provided in different planning stages with different scale of preciseness. However IFC does not explicitly define different information scales and accuracy nor provide means for a procedural description of geometry.

Scaling is a new context in modeling which scales the data in accordance to the information needed to be obtained from data. This concept the first time got developed in CityGML systematically in the form of Level of Detail (LoD). The concept of LoD in the data modeling controls the complexity of the data model. In the lower levels of detail the representation of data is much simpler while in the higher levels of detail the data model is more and more complex.

This facilitates to interoperate the data models in the proper level of detail which best suits to the demands of integration. Interoperating the data models in a lower level of detail as shown in figure 29 avoids transferring unrequired entities which might slow down the process time as well as make the interoperability complex. Besides, simple representation of the data models is very useful in the context of visualization.

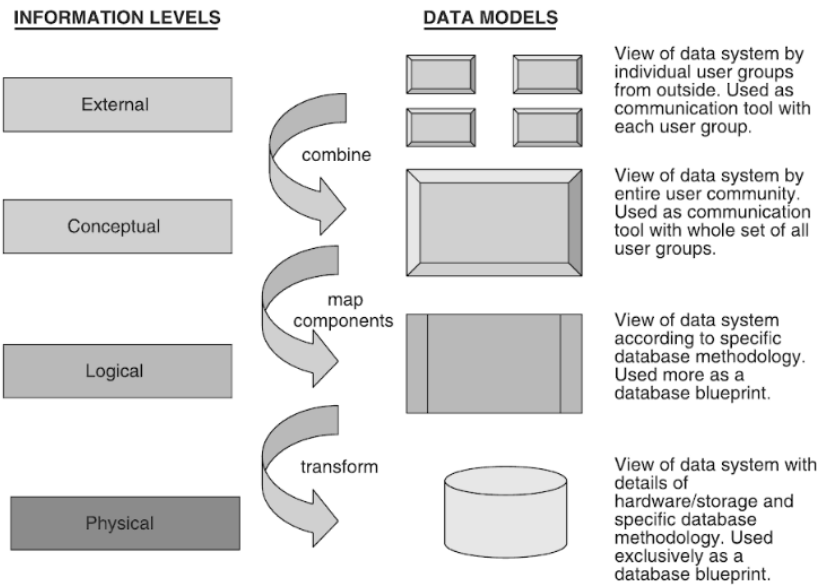


Figure 29 Data models at different information levels, (Ponniah, 2007)

"In GIS mostly, a bottom-up approach is followed, which means coarser representations are abstracted from detailed data by a process called 'generalization'" ((Bormann et al., 2014b, p 2; Forberg, 2007; Meng & Forberg, 2007). Generalization is one of the fundaments of the data models from 1972. According to Bormann et al. (2014b) often, though, *"models on different LoDs are independently generated due to decoupled acquisition processes. In consequence, the available data models for representing and exchanging 3D city models rely on an independent representation of the geometry on the individual LoDs"* (Bormann et al., 2014b, p 2).

Derived from OGC, author suggests following extension for the LoD concept of BIM data models which is briefly demonstrated in Table 4.

- LoD 0:
  - BIM: no 3D BIM modelling, a CAD 2D ground plan
  - Infra: 2D-alignment

    - GIS: GIS-LoD 0 + BIM-LoD 0 + Infra-LoD 0
- LoD 1:
    - BIM: cubature model (LxBxH), block building, virtual texture
    - Infra: 3D-alignment
    - GIS: GIS-LoD 1 + BIM-LoD 1 + Infra-LoD 1
- LoD 2:
    - BIM: cubature model, simple roof model, floor segmentation, photorealistic texture
    - Infra: 3D-alignment, junction tunnel and bridge structure segmentation
    - GIS: GIS-LoD 2 + BIM-LoD 2 + Infra-LoD 2
- LoD 3:
    - BIM: cubature parametric model, exterior façade door and window model, photorealistic texture
    - Infra: parametric design of the infra object and tunnel and bridge structures separately in accordance to BIM-LoD 3
    - GIS: GIS-LoD 3 + BIM-LoD 3 + Infra-LoD 3

Here the LoDs 0, 1, 2, and 3 are focusing on modeling the surrounding or current situation of the outdoor. LoDs 4 is the interface between outdoor and indoor.

- LoD 4:
    - BIM: parametric model, building envelope, properties of building envelope
    - Infra: infra object envelope, geometric design of the infra object and structures in accordance to BIM-LoD 4, street furniture
    - GIS: GIS-LoD 4 + BIM-LoD 4 + Infra-LoD 4

That is where OGC stops modelling of the environment. However modelling continues for the infrastructure and building processes with at least two more levels of detail as illustrated below:

- LoD 5:
    - BIM: parametric model, building envelope, building components, structural design, all properties of building like material
    - Infra: infra object components, structural design of infra object in accordance to BIM-LoD 5
- LoD 6:
    - BIM: building furniture, MEP model, irregular objects
    - Infra: MEP model for tunnel and bridge structures, irregular objects

The table of LoD is one of those rules which will be inserted into rule based link model as an input to define some rules of relationships. In such table one can define which objects of the infrastructure are defined in which specified LoD, which LoD from different data domains are related to each other and which LoDs of which data domains are to be considered in which project phase.

Such table includes definitions such as:

- in project phase “xxxx” GIS objects of domain “xxxx” and “yyyy” from LoD “x” and “y”, GIS objects of domain “zzzz” with an Accuracy of “xx” or “yy”, as well as infrastructure objects of domain “kkkk” as well as BIM objects of subdomain “qqqq” should be integrated.

Another type of rules are conflict detection rules. It is essential to define there the circumstances and cases which are in conflict with each other. One of such conflict cases can be for example to have an axis of road alignment and not bridge alignment crossing a lake or river. Therefor defining type of objects and attributes which are in conflict with each other such as water bodies and road alignments and their type of conflict and allocating them proper attribute for conflict detection is highly important

More examples of such determinations are mentioned in chapter 5, implementation and 5.2 case study of this dissertation. Sample codes of such implementation are to be found in the chapter 5 “implementation”. Such definitions are to be settled and harmonized using ontology based semantic matching approaches but until then they should be one of the fundamental rules to be fed into the match finding procedure of this thesis. A schematic definition for such objects, domains and LoD are described in the table 4.

Table 4 Extension of LoD Concept for BIM / Infra Models

Models / LoD	BIM	Infra	GIS
LoD 0	2D ground plan	2D Alignment	GIS-LoD0 +BIM-LoD0 +Infra-LoD0
LoD 1	Cubature model (l*b*h) Block building Virtual texture	3D-alignment Horizontal and vertical alignment	GIS-LoD1 +BIM-LoD1 +Infra-LoD1
LoD 2	Cubature model Simple roof model Floor segmentation Photorealistic texture	3D-alignment Structure segmentation Junction, tunnel, bridge	GIS-LoD2 +BIM-LoD2 +Infra-LoD2
LoD 3	Cubature model Parametric model Exterior model Façade door and window Photorealistic texture	Parametric design Tunnel and bridge structures separately in accordance to BIM-LoD3	GIS-LoD3 +BIM-LoD3 +Infra-LoD3
LoD 4	Parametric model Building envelope Properties of building envelope	Infra object envelope Geometric design Structures in accordance to BIM-LoD4 Street furniture	GIS-LoD4 +BIM-LoD4 +Infra-LoD4
LoD 5	Parametric model Building envelope Building components Structural design All properties of building	Infra object components Physical elements Structural design of structures in accordance to BIM-LoD5	-BIM-LoD5/IFC
LoD 6	Building furniture MEP model Irregular objects	MEP model for tunnel and bridge structures Irregular objects	-

Another set of rules to be implemented is the table of relations between team members. Organizational hierarchies and their influences on LoDs as well as data domains and their range of impact as well as field of interferences is to be implemented and fed into matching system as well.

A direct interpretation of this item is that for example the high level managers or *Team Members of Level 0* do not seek each and every detail in the planning procedure. They are interested to know about the overall managerial issues such as a very rough time table, very rough costs and benefits as well as possible risks and threads.

Project level managers or *Team Members of Level 1* are more interested and at the same time responsible for the technical issues. This causes that they need more detailed overview about the project. Yet such managers might be from different divisions and therefore will have the highest interface to the other domains or sub domains. Such team members are not only responsible for their own area of impact but also as a part of the whole for the entire realization process. For example in a 200 KM railway project there might be a need for tunnels and bridges as well as adjacent roads. Therefore the team members responsible for the sub domains need to interoperate with the other domains as a whole and with their own domain with a higher activity individually. In such case the technical managers of each domain are to be integrated with their own domain in more detail in compare to the other domains. Yet they should be informed about the interfacing domains.

Project experts or planers or engineers or technicians build the *Team Members of Level 2*. These members have little interface to the other domains. Their area of impact will be their own domain for example in the above mentioned example the bridge domain or the tunnel domain. This level of team members is supposed to know each and every detail about the task according to their area of activity.

For the definition of the LoD this thesis refers to the organization as the leading model for the specification of different LoDs as shown in figure 30. As demonstrated in table 5 organization can be divided into 3 main levels; the company management level, the project management level and the planning team. Each of these main groups has similar interests regarding a project.

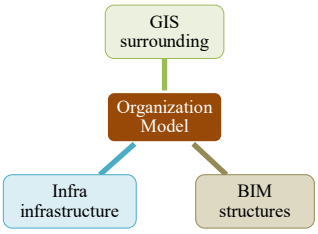


Figure 30 Organization the leading model for defining the LoDs of product model

The company management cares more about the monetary and scheduling, yet in a very brief level and technical properties of the project are less interesting for them. Whereas the project managers are involved essentially with monetary and timing issues, still they need to know some technical coordinates of the project.

Table 5 Table of Levels of Detail of Modeling in regard to Organization from the sight of Building Planning

Organization		Surrounding / GIS		Structures / BIM		Infrastructure
LoD	Organizational Hierarchies	Constructed	Unconstructed	Building	Tunnel & Bridge	
LoD 0	Managing Director	LoD 0 (LoD 0)	LoD 0 (LoD 0)	LoD 0 (LoD 0)	LoD 0[1]	LoD 0 (LoD 0)
	Office Manager					
	Head of Division					
LoD 1	Project Coordinator	unrelated domains[2]: LoD 1 (LoD 1)	unrelated domains: LoD 0 (LoD 0)	unrelated domains LoD 1 (LoD 1)	LoD 1 (LoD 1)	unrelated domains LoD 1 (LoD 1)
		related domains: LoD 2 (LoD 2)	related domains: LoD 1 (LoD 1)	related domains: LoD 2 (LoD 2)		related domains: LoD 2 (LoD 2)
	Project Manager	unrelated objects: LoD 1 (LoD 1)	unrelated objects: LoD 0 (LoD 0)	unrelated objects: LoD 1 (LoD 1)	LoD 1 (LoD 1)	unrelated objects: LoD 1 (LoD 1)
		related objects: LoD 2 (LoD 2)	related objects: LoD 1 (LoD 1)	related objects: LoD 2 (LoD 2)		related objects: LoD 2 (LoD 2)
LoD 2	Senior Planer	LoD 3, 4, 5, 6[3, 4] (LoD 3, LoD 4)	LoD 2, 3, 4 (LoD 2, LoD 3, LoD 4)	LoD 3, 4, 5, 6 (LoD 3, LoD 4)	LoD 2, 3, 4 (LoD 2, LoD 3, LoD 4)	LoD 3, 4, 5, 6 (LoD 3, LoD 4)
	Junior Planer					

- 1: There is no LoD 0 for tunnel and bridge module in CityGML, this can be added to CityGML as a 2D alignment representation for Tunnel & Bridge
- 2: Domains are engineering domains such as electrical, structural...
- 3: Depends on project phase and services such as feasibility, preliminary...
- 4: LoD 5 can be added to CityGML as building or tunnel or bridge components

The planners are not interested in monetary and scheduling issues; however the technical identities of the project are to be carried on by them. Therefore depending on which services and which project phases are offered, they need to know each and every technical detail of the project. From the organization point of view IIM can be divided into three main elements or domains, namely surrounding, structures, as well as existing and to be planned infrastructure. Table 6 shows the level of detail in which the link model can interchange the data most efficiently.

To be able to adopt this table for the case of infrastructure planning data model, one should define the levels of detail for infrastructure data models. This definition specifies a least LoD for different product models. In analogy to CityGML this dissertation introduces a the following draft for data model integration in regard to organization from the building planer point of view as shown in table 6.

Table 6 Levels of Detail of Modeling in regard to Organization from the sight of Infrastructure Planning

Organization		Surrounding	Structures		Infrastructure
LoD	Hierarchies	Constructed or Unconstructed	Building	Tunnel & Bridge	
LoD 0	Managing Director	LoD 0 (LoD 0)	LoD 0 (LoD 0)	LoD 0	LoD 0 (LoD 0)
	Office Manager				
	Head of Division				
LoD 1	Project Coordinator	unrelated domains: LoD 1 (LoD 1)	LoD 1 (LoD 1)	LoD1, 2 (LoD 1, LoD 2)	unrelated domains LoD 1 (LoD 1)
		related domains: LoD 2 (LoD 2)			related domains: LoD 2 (LoD 2)
	Project Manager	unrelated objects: LoD 1 (LoD 1)	unrelated objects: LoD 1 (LoD 1)	unrelated objects: LoD 2 (LoD 2)	unrelated objects: LoD 1 (LoD 1)
		related objects: LoD 2 (LoD 2)	related objects: LoD 2 (LoD 2)		related objects: LoD 2 (LoD 2)
LoD 2	Senior Planer	LoD 3, 4, 5, 6 (LoD 3, LoD 4)	LoD 3, 4 (LoD 3, LoD 4)	LoD 3, 4, 5, 6 (LoD 3, LoD 4)	LoD 3, 4, 5, 6 (LoD 3, LoD 4)
	Junior Planer				

However there is no level of detail classification for infrastructure design data model in standards like LandXML or OKSTRA. But the project phases in alignment planning reflect a similar concept. The German fee catalogue for architects and engineers (HOAI) classifies the alignment design project phases in five main categories; from feasibility study to preliminary design, conceptual design, approval design, and detailed design. As the project stages move forward, the detailing and complexity of design increases. As demonstrated in figure 31, for Infrastructure and Geospatial Data Models to get interlinked in the spatial link model efficiently the fuzziness should be reduced as far as it is possible. To make the results as relevant as possible one method is to define proper LoDs. The result of modeling process gets an LoD consequently.

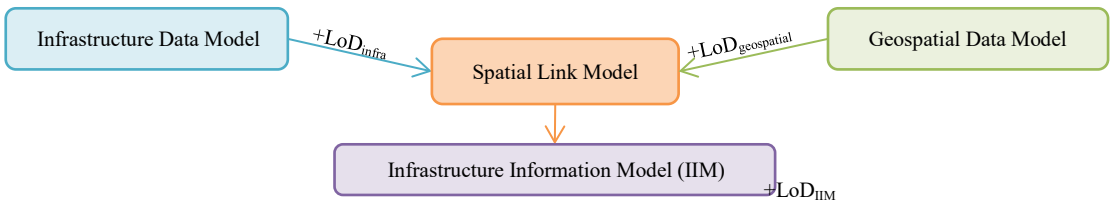


Figure 31 Concept of IIM and Levels of Detail

In a rule based link model various input data models don't need to have similar LoDs. Therefore the link model facilitates mixed LoDs. As shown in the figure 32 in the IIM the LoD of the linked model is a function of the LoD of each input data models. This creates a dynamic LoD definition. Dynamic LoD concept means that depending on the task or process defining the rules of the interlinking, the link result might have a different level of complexity.

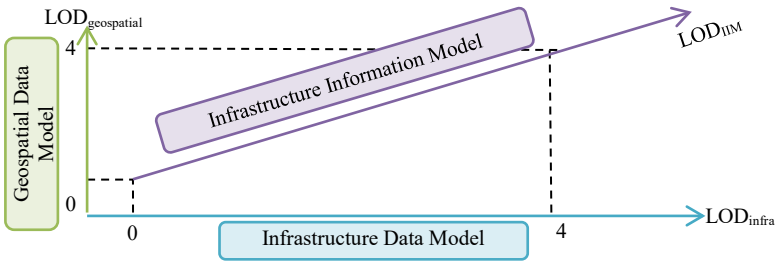


Figure 32 Level of Detail diagram in IIM

Chapter 4

Spatial Modelling and Processing

Since the 1980s, with the emergence of new database applications, such as geographic information systems (GIS), image and multimedia databases and so on, spatial database technology has been a great asset to the infrastructure applications. This thesis has implemented spatial query methods for the applications of infrastructure projects. Chapter 4 describes the match finding solution in different data domains and linking discrete information from different software packages to inspect the coherency of related application models and create complementary analysis models in a rule based link approach.

Explained in this chapter the method of infrastructure information modelling for interoperation in spatial domains generate interlinks through spatial identity of entities. Match finding through spatial links enables any kind of data models sharing spatial property get interlinked. Through such spatial links each entity receives the spatial information from other data models which is related to the target entity due to sharing equivalent spatial index. This information will be the virtual properties for the object.

This chapter illustrates the application of Nearest Neighborhood algorithm for spatial match finding and explains abstraction concept via filtering and refining approaches. For the abstraction of the spatial matching results hierarchical filtering techniques are used for refining the virtual properties. These approaches focus on two main application areas which are product model and Level of Detail (LoD).

4.1 SPATIAL IDENTIFIERS

"*The digital representation might be for some past, present or future time period (or contain some combination of several time periods in an organized fashion)*" (NCGIA, Unit 10). *"Experience has shown that the different branches of civil engineering have developed their own concepts of BIM. As a result, terminologies, data formats and work processes differ. To solve the non-uniform thread of BIM approach the buildingSMART organization introduced three technologies due to standardizes the three core components – terminology, digital storage and work processes"* (Obergriesser & Bormann 2012, p 584; BISM, 2012).

Geographic information systems (GIS) is the most common platform for using spatial data (Stolze, 2003) and a very good example of spatial databases is to be found in the field of transportation and infrastructure design like highway data and their accessing roads, characterization, the level of usage by the public or detailed information about condition and maintenance.

The spatial databases are designed to consider that for each organization only certain phenomena are important enough to be collected and represented in a database. Thus, a spatial database involves a sampling of geographic reality, to determine the status of that reality. Therefore such databases are also called Geospatial Information Systems (GIS). All different types of such databases are common in having the following characteristics according to NCGIA, Unit 10:

- *"Contemporaneous - contain information of the same vintage for all its measured variables*
- *Are as detailed as necessary for the intended applications*
- *Are positionally accurate*
- *Are exactly compatible with other information that may be overlain with it*
- *Are internally accurate, portraying the nature of phenomena without error - requires clear definitions of phenomena that are included*
- *Are readily updated on a regular schedule*
- *Are accessible to whoever needs it"* (NCGIA, Unit 10)

Due to some characteristics of spatial data models, like huge size, complex structure and high operational costs, the optimization of spatial query becomes a difficult yet focal point of handling such data models.

As described by Sondheim et al. (1999) *"there has been a great deal of interest within the database world in managing atypical, non-tabular data, which together are considered as multimedia. The SQL language, long the backbone of the relational database market, was originally designed to handle only tabular data. However, the latest version, SQL3, is comparable to a full programming language and includes the ability to define Abstract Data Types (ADT). An ADT may include functions; a polygon may have one function to return its boundary and another to return the number of holes within the polygon. An ADT is roughly equivalent to an object type in object oriented technology. instead of thinking about data in a database as consisting of rows of attributes with primitive domains (integers, real numbers, character strings, and the like), we can now include higher level data types such as lines and polygons, each of which may have various functions as an inherent part of its definition. The creation of a standardized set of data types, based on the ADT capability in SQL3, is the basis for the multi model extensions known as SQL/MM and defined in the SQL3"* (Sondheim et al., 1999, p 352).

"One particular application area included in SQL/MM is Spatial (ISO/IEC JTC 1/SC 21 1996). These developments are significant because they enable databases to store and process a wide variety of data types, including spatial data types, all in the same environment. The enhancement of extended-relational and object relational databases to geospatial applications parallels the migration of conventional APIs to more object-oriented designs encapsulating both data and methods. SQL/MM Spatial and the OpenGIS specification are compatible standardization efforts which will make it relatively easy for database vendors intending to comply with the former to also comply with the latter". Consequently, connectivity between different databases housing geospatial data and between such databases and other software products will become practical through common interface defined by the OpenGIS specification" (Sondheim et al., 1999, p 352).

Thus, location specification of real or modelled objects is demonstrated in spatial data types as long as those objects belong to the real or modelled space in which they allocate. Spatial identities are basic elements not only in infrastructure and geo-data models but also in many other engineering domains such as energy, traffic, geological and geomechanical data models, climate and weather, seismic data such as earthquake or tsunami and etc. As shown in figure 33, using such identities, it is possible to unify various data models in one unique spatial system for further spatial queries and linking.

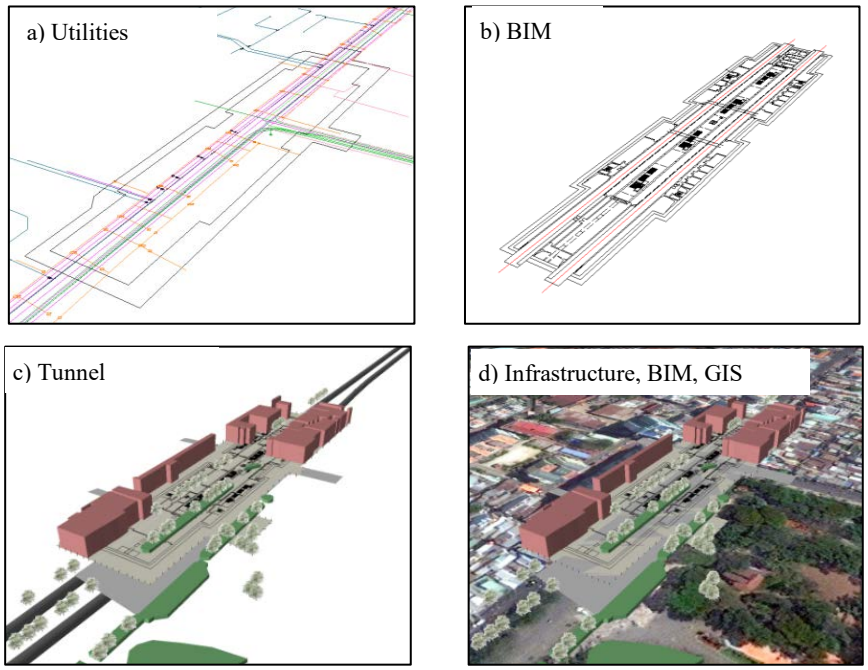


Figure 33 Different data models in infrastructure projects spatially linked to each other, (Przybylo et al., 2015)

4.1.1 Spatial Indexes

The fundamental question in this thesis is to query and relate objects which are positionally close to each other at a specific location. The ideal aim is to find the position of one nearest point to the queried point between lots of others. This question can be answered by an individual through a rapid look at the queried point in a two dimensional environment. The computer needs spatial indexing to solve this problem, unless a sequential scan is sufficient.

Spatial index is one of the three key features of a spatial data model, which is one of the target data models of this work. Indexes are what make using a spatial database for large data sets possible. Spatial index is like an ordinary index with this difference that spatial objects are not 1D data points rather are in higher dimension space (e.g. 2D or 3D) and thus ordinary indices are not appropriate for indexing such data. Spatial indices are used to quickly locate features that match a spatial query. Therefore, an appropriate spatial index is highly important, especially when working with large amounts of data, such as geo data in large infrastructure projects.

Spatial indices work differently depending on the data source. The spatial databases in SQL Server and Oracle which use binary geometry storage or ST_Geometry storage use a system of up to three grids as the spatial index (Oracle Spatial Users's Guide and Reference, 2008). "*Personal geodatabases use a single grid. A grid is defined by a size, referred to as the grid size. This is the size of each cell in the grid, specified in the units of the spatial reference system. Oracle Spatial, Informix, and PostgreSQL and SQL Server do not use grid sizes—they use an R-tree index*" (Esri Inc.).

Without indexing, any search query for a specified feature in infrastructure data model or geo data model would require a "sequential scan" of every record in the database. In large data models of infrastructure projects this will be very time consuming and not functional. Once a spatial index created, one can specify an access method and an operator class. The access method speeds up the access to the entities. The operator class decides for a query to apply an index or not.

4.1.2 Tree-Based Spatial Indexes

"Indexing speeds up searching by organizing the data into a search tree which can be quickly traversed to find a particular record" (Workshop of Boundless Geo, part 15 spatial indexing, 2010). Measuring has shown there is an improvement in the query performance from 55 milliseconds to 9 milliseconds. Figure 34 illustrates the concept of spatial indexing using dynamic grids by Google Earth in different LoDs.

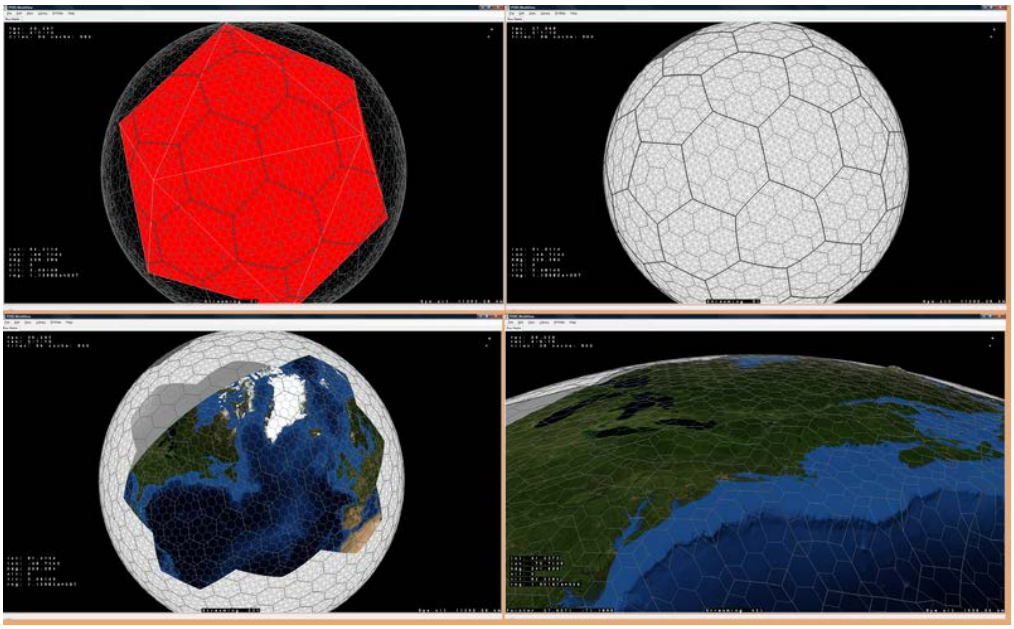


Figure 34 Spatial indexing using grids in Google Earth, (Peterson, 2007)

According to Pu (2004) *"a query using a tree based spatial index can quickly descend the tree to find objects in the general area of interest and then perform more exact tests on the objects themselves. A tree based index can improve performance because it eliminates the need to examine objects outside the area of interest. Without a spatial index, a query would need to evaluate every object to find those that match the query criteria"* (Pu, 2004).

In object relational spatial databases like ORDBMS, spatial data types and spatial operators can be made directly available to the end-users, enabling them to formulate spatial queries.

According to Borrmann and Rank (2008) *"the spatial operators available for the spatial types are the most important part of the algebra. They have discussed in detail operators comprising metric (distance, closerThan, fartherThan etc.), directional (above, below, northOf etc.) and topological (touch, within, contains etc.)"* (Borrmann & Rank, 2008, p 2).

The focus of this thesis is to create a high-level link model for infrastructure process and provide strategies to extend it to further related domains with spatial identities like the geomorphological and energy processes. Besides this dissertation demonstrates how to break down the domain entities into their fundamental units to detect identifiers which enable a semantic linking. An example of this is the climate data model and its temporal identifiers or the traffic data model with its tempo-spatial identities.

Spatial Index is a property of spatial data models and will be delivered along with the data. Spatial index has a direct impact on the performance. Therefore the concept of LoD should be strongly considered in definition of such indices. Yet, ordinary indices such as B-tree are not appropriate for indexing such data. The well-known spatial index technique is R-tree which fits very effective for large datasets (Marik et al., 2003).

According to PostGIS (2010) *"standard database indexes create a hierarchical tree based on the values of the column being indexed. Spatial indexes are a little different – they are unable to index the geometric features themselves and instead index the bounding boxes or extent of the features. Both PostGIS and Oracle Spatial share the same "R-tree" spatial index structure"* (Workshop of Boundless

Geo, part 15 spatial indexing, 2010). As illustrated in figure 35, R-trees break up data into rectangles (cubes), and sub-rectangles (cubes), and sub-sub rectangles (cubes), etc according to their geometry and search loops based on generic index structures (GiST). "*It is a self-tuning index structure that automatically handles variable data density and object size. Whereas the PostgreSQL query planner intelligently chooses when to use or not to use indexes to evaluate a query"* (Workshop of Boundless Geo, part 15 spatial indexing, 2010).

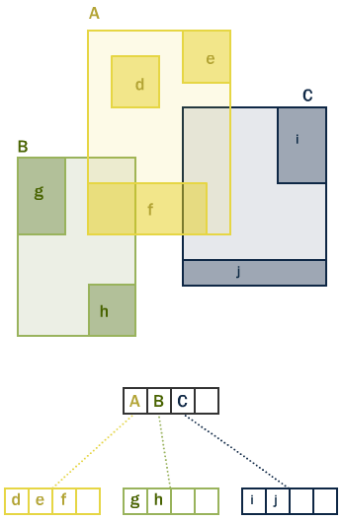


Figure 35 R-tree based spatial index structure and hierarchy for geometries, (PostGIS, 2016)

Shown in figure 35 this approach runs in two phases, the approximate index will be evaluated in phase one, in phase two an exact test will be carried out which can significantly reduce the number of calculations necessary to answer a query. Such reduction minimizes the time needed for a specific query on a large data model. For example consider the query to be finding out the quantity, time and cost of transporters for transporting the proper soil (illustrated with letters A, B, C, … in the figure 36) from the available mines (the points in the figure 36) around the project location for different kilometer station sections and portions of the targeted road (shown as a line in the figure 36).

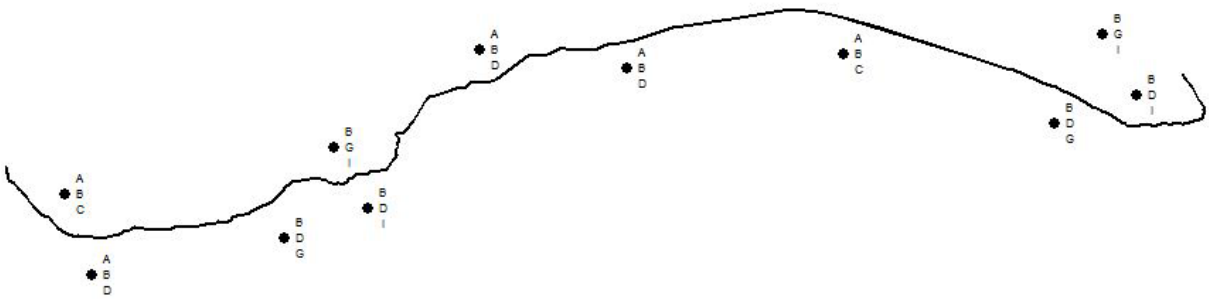


Figure 36 An example of querying mines with different soil types around the project location

For this query it is very important to be able to define a spatial structure and indexing for the targeted and queried road in a flexible way.

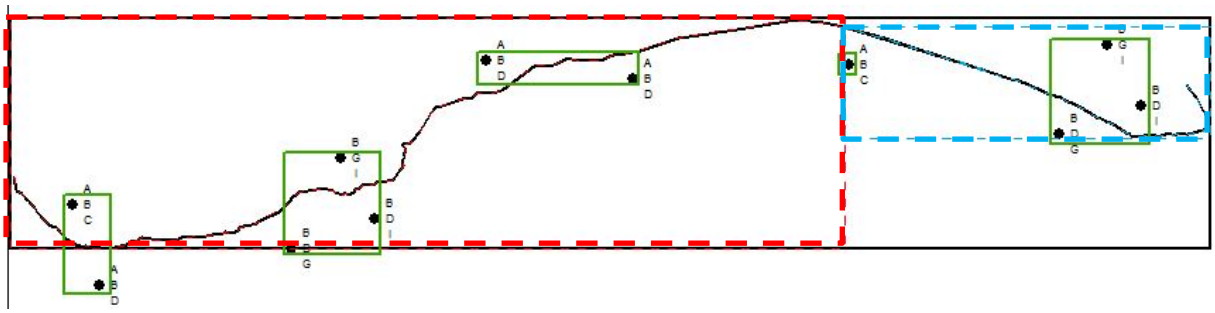


Figure 37 An example of bonding boxes around different spatially structured infrastructure and geodata models

It means it should be possible to consider the road as a whole object (black continuous bounding box in figure 37) and at the same time in different sections (red and blue dashed bounding box in figure 37) for each section of the road in which the mines will be eligible for consideration which fulfill the querying requirements (green bounding box in figure 37).

There is a standard deviation associated with all spatial data components, both absolute and relative. This defines the precision of the spatial index in the spatial data. If the standard deviation of any component is zero, the quantity is either known very precisely and/or the value (e.g. a control point) is being used as a "fixed" quantity (Burkholder, 2004).

Infrastructure and geospatial problems are the typical fields of application where Spatial Indexing and spatial query can be of great asset. With the goal of spatial matching for such applications where there is not yet reference libraries.

Unlike MM approach which allows static links with no semantic meaning, interlinking through spatial identities enables to create links through matching rules. This means to create potential dynamic links which gets updated if any change happens in the data model which might change the location of the objects. This method allows interlinking of other data domains with spatial identity like noise emission as well as unknown future domains, if they contain spatial properties (Trepanier & Chapleau, 2001).

4.2 NEAREST NEIGHBORHOOD AS A BASIC LINK METHOD

With the goal of spatial matching, this thesis uses Nearest Neighborhood (NN) algorithm for match finding. According to NN algorithm and by definition explained in eq.(5) and obtained from Altman (1992) the nearest to P: Position is found through Euclidean Distance:

$$\left|\left|P_{geo} - P_{infra,i}\right|\right| = \sqrt{\sum_j \left(P_{geo} - P_{infra,i}\right)^2} \rightarrow \left|\left|P_{geo} - P_{infra,i}\right|\right|^2 = \sum_j (P_{geo} - P_{infra,i})^2 \qquad \textbf{Eq.(5)}$$

NN depends critically on the distance, thus the input data types should be normalized, i.e. should have the same range of values. In the case of geospatial and infrastructure data model, the best solution is to transform the spatial unit of both of them into a global spatial reference system as presented by the eq.(6) (O´Neil, 2010).

$$P_{geo}, P_{infra} \xrightarrow{transform} global\ reference\ system \qquad \textbf{Eq.(6)}$$

For NN to be able to avoid random perturbations for the calculations, irrelevant (noisy) input should be eliminated. For the geospatial and infrastructure data models it is done if different domains get interlinked separately. That means not to link for example vegetation and building domains from geospatial data model with infrastructure data model in one stage. But rather to consider data domains in separate steps, because each data domain might act as noise for the other domain.

If there are still noises in match finding results, smoothing the NN results can reduce them and increase the precision of the results. This can be done in the form of:

- weighting features by mutual information according to eq.(7)

$$if: w_{infra} = l\left(P_{infra}, y\right)$$
$$then: d\left(P_{geo}, P_{infra}\right) = \sum_{j=1}^{n} w_{infra}(P_{geo} - P_{infra})^2 \qquad \textbf{Eq.(7)}$$

- using the Mahalanobis distance according to eq.(8)

$$D_M\left(P_{geo}, P_{infra}\right) = \left(P_{geo} - P_{infra}\right)^T \sum^{-1} \left(P_{geo} - P_{infra}\right) \qquad \textbf{Eq.(8)}$$

The nearest-neighbor problem and its variants are used in a variety of applications domains. Key applications of this algorithm are Spatial Network (ex. GIS) and Artificial Intelligence (AI) (ex. Pattern Recognition), whereas queries of spatial network, like GIS, are not similar to classical Euclidean space problems, as spatial network distance might be more concerned with time than space.

That is why the temporal data and applications are the next generation which is of value for infrastructure data models such as high speed trains or maglevs. The proposed approach of this thesis handles in the first step the spatial data models. Although temporal data is not the main stream of this thesis, yet some solutions will be suggested in the next chapter for the temporal data models. NN uses different algorithms for finding the matches which are based on two main computational models, namely *Computational Geometry* as in figure 38-a, or *Spatial Databases* as in figure 38-b.

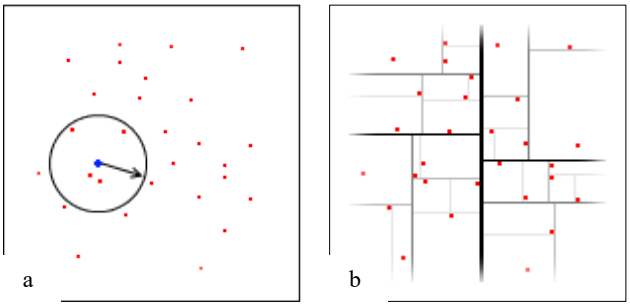


Figure 38 Computational models for NN, left a. computational geometry and right b. spatial database, (ALGIB[15])

Table 7 comprises different data structures designed with Nearest Neighborhood in mind in accordance to criterion regarded in the methodology of this thesis.

Table 7 Nearest Neighbor Algorithm for different data structures, (Dietterich, 2014)

Criterion \ Data Structures	Perc	Logistic	LDA	Trees	Nets
Mixed data	-	-	-	x	-
Missing Values	-	-	x	x	-
Outliers	-	x	-	x	x
Monotone transformations	-	-	-	x	/
Scalability	x	x	x	x	x
Irrelevant inputs	-	-	-	/	-
Linear combinations	x	x	x	-	x
Interpretable	x	x	x	x	-
Accurate	x	x	x	-	x

In their research Borrmann et. al. (2009) have performed an Octree-based implementation of directional operators in a 3D Spatial Query Language in Building Information Models for the extraction of partial models that fulfil certain spatial constraints.

Octree implementation technique cannot realize the dimension operator (Borrmann & Rank, 2008). An Octree is developed to enable spatial queries but does not facilitate k-Nearest Neighborhood searches. Octree cells are guaranteed to be cubical which makes it impossible to use volume bounds to control the number of cells that one has to examine when solving approximate nearest neighbor queries. The detection of the nearest neighbor object to a given location in the reference space (NN query) is a common problem in Geospatial Information System. Data structure supporting range queries are not always adequate to support NN queries.

Search mechanism through computational geometries using for example quad-trees based on sequential scan is very time consuming procedures especially when the geometry is complicated or dynamic. This means query time increases exponentially once the dimensions rise.

"Typical algorithms often assume all data fits in memory. Of course this assumption fails in many real world systems. Particularly in the case of geospatial data very large data sets exist that are orders of magnitude larger than a single computer's memory, and the majority of data in these cases must reside on slower storage media. The performance disparities between main-memory and secondary media, and in some cases, even physical characteristics (e.g. platter rotation and geometry in a hard-disk; sequential nature of tape) motivate the need for specialized algorithms and data structures that best suit each medium" (Davies, 2011).

[15] http://www.alglib.net/other/nearestneighbors.php

As storage spaces and power of processing increase very fast, size and complexity of spatial databases grow rapidly. GIS applications, like Google's Google Earth, contain huge datasets such as detailed images and maps, extensive address database and information on enterprises and their profiles. Due to such huge data volume which is supposed to be processed in nearest neighbor queries, innovative approaches for nearest neighbors are developed. These approaches call for novel data structures (such as KD-trees or R-trees) designed for nearest-neighbors and use approximation techniques that decrease search space through reasonable assumptions. Table 8 shows how NN algorithm fulfills the required criteria for different data structures.

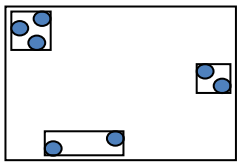


Figure 39 R-tree cardinality

That is what makes disk based algorithms such as R-tree suit to the NN problems even better than KD-trees, because they answer queries in logarithmic time. In a simpler word R-trees are Bottom-up approaches which start with partitioning a set of points/rectangles into groups of small cardinality and find minimum rectangle containing objects for each group as shown in figure 39. R-tree is a dynamic structure, so the goodness of the insertion algorithm is very important.

R-tree family is the only data structure which can practically handle not only points but also rectangles. Therefore in geospatial data model, to be able to build an R-tree based data structure, a rectangular boundary will be considered around the geometries. This rectangular boundary is called "Extent" or "Minimal Bounding Rectangle" (MBR) which is the minimal bounding n-dimensional rectangle bounds its corresponding objects is used to build a good R-tree. For example as represented by the eq.(9) in 2D with x and y axis:

$$MBR: I = (Ix, Iy), where \begin{cases} Ix = [\min\_x, \max\_x] \\ Iy = [\min\_y, \max\_y] \end{cases} \quad \textbf{Eq.(9)}$$

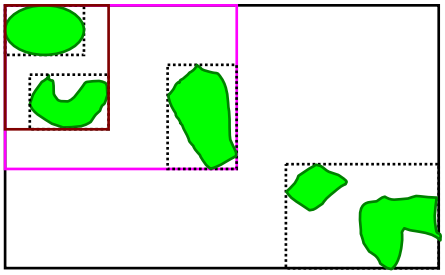


Figure 40 Minimal Bounding Rectangle in R-trees

Thus illustrated in the figure 40, a spatial database consists of a collection of tuples representing spatial objects, known as Entries. In R-tree based data structures, each Entry has a unique identifier that points to one spatial object, and it's MBR.

$$Entry = f(MBR, pointer) \quad \textbf{Eq.(10)}$$

Yet, the R-tree makes no guarantees about the efficacy of the bounding. *"The only guarantee is that each node in the tree will contain no more than B rectangles, where B is the block-size, (number of rectangles stored in each leaf on disk). This overflow constraint is again similar to the KD-tree, where each node can contain no more than B values, but an important difference between the two data structures also emerges. With a KD-tree, the remedy for an overfull node is evident: split the node into two nodes through the median value. While with the R-tree there is no obvious remedy. However there is no exact definition, how one best partitions (B+1) smaller rectangles into two larger bounding rectangles such that the partition minimizes future disk accesses!*

That is why Priority R-trees (PR-trees) are developed. The PR-tree, by considering each rectangle as a 4-dimensional point (xmin, ymin, xmax, ymax) in a KD-tree, is able to achieve a splitting optimum that guarantees good asymptotic performance (similar to the KD-tree) for subsequent accesses.

But unfortunately this KD-tree technique only works during the bulk-loading phase, and is unsuitable for maintaining a dynamic index. Thus there is no support for adhoc additions and deletions after the initial construction of the PR-tree" (Davies, 2011).

The dynamicity is one of the key factors in the case of construction engineering procedures in their lifecycle. Therefore, despite this shortcoming, this thesis has decided to continue the work with R-trees.

Table 8 shows the advantages and disadvantages of each data structure which can be used with nearest neighborhood search mechanism. As illustrated Trees are the best data structures fulfilling most of the criterion observed for the purpose of this thesis.

Among the trees, as shown in table 8 the R-tree provides the best data tree structure for the goals of this work.

Table 8 Key aspects observed for choosing the most adequate Tree Structure for data models, (Esfahani et al., 2014)

Criterion \ Data Structure	B-tree	Octree	KD-tree	R-tree	PR-tree
Adjacency in tree sense	-	-	x	x	x
Huge Datasets	-	-	-	x	x
Dynamicity	-	-	-	x	-
Scalability	x	x	x	x	x
Interpretable	x	x	x	x	x
LoD, Accuracy	-	-	-	x	-
Points AND Rectangles	-	-	-	x	x
Multidimensional	-	-	-	x	x

Developed and introduced by Guttman (1984), the main idea of R-trees is to allow parents overlap. This results in a guarantee for 50% utilization and easier insertion and split algorithms.

According to Burdescu et al. (2005) *"an R-tree is a height-balanced tree with index records in its leaf nodes containing pointers to data objects. Nodes correspond to disk pages if the index is disk-resident and the structure is designed so that a spatial search requires visiting only a small number of nodes"* (Burdescu et al., 2005, p 2). R-tree manages the data in a tree structure, where nodes contain bounding boxes. Bounding cubes represent the maximum extent of the data connected to the sub tree. Indices are completely dynamic, addition and removing of data can be inter-indexed with queries and does not need periodic restructuring. R-tree index enables faster querying and thus a significant performance improvement. Access to multidimensional data happens quicker in R-tree access method. R-tree enables split or extending MBRs if an insertion does not allocate an object to a specified MBR. This makes the structure of MBRs and indexing dynamic.

Structurally an R-tree is a multi-way external memory tree with index nodes and data (leaf) nodes, where all leaf nodes appear on the same level. Every node contains between m and M[16] entries and the root node has at least 2 entries (children).

Therefore in NN algorithm:

- variable sized hypothesis space (+)
- very flexible decision boundaries (+)
- growing a good R-tree can be expensive (-)
- distance function must be carefully chosen (-)
- irrelevant or correlated features must be eliminated (-)
- computational costs are memory, classification and time computation (-)

In an R-tree search all "children" of a node are completely covered by their parent node. More than one parent may cover a child bounding box but the child belongs to ONLY ONE of the covering parents to eliminate duplications, but yet a point search may follow more than one leaf.

[16] M is the maximum number of entries that will fit in one node. And $m \leq M/2$ is the minimum number of entries in one node.

Therefore the most critical part of applying NN on spatial data models is how good the data structure is understandable for the NN. Depending how well spatial data is structured, there will be an index of approximation and error in the results. This is explained by the eq.(11).

$$\hat{I}_v = \hat{I}(Ps, r, R, \varepsilon) where \begin{cases} Ps: point\ set \\ r, R: search\ range[r, R] \\ \varepsilon: approximation\ error \end{cases} \quad \textbf{Eq.(11)}$$

This error is either calculable if the NN is supposed to find kNN (k-th NN) or at least one result, or settable if the results are supposed to have a minimum accuracy.

If ε-approximate nearest neighbor performs a point query q, and the actual nearest neighbor is at distance d, the procedure typically ends up with a search that examines every data structure cell that reaches from the inside to the outside of an annulus or annular shell with inner radius d and outer radius $(1+\varepsilon) * d$. If the cells have bounded aspect ratio[17], as they are in an octree, then there can be at most $\frac{1}{\varepsilon^{(d-1)}}$ such cells, and that proves good bounds on the time for the query. If the aspect ratio is not bounded, as in a KD-tree, these bounds do not apply.

The cells in a KD-tree can have high aspect ratio, meaning they are guaranteed to have at most logarithmic depth, which also contributes to the time for a nearest neighbor query. KD-trees are balanced binary trees and rebalancing can be expensive, but balancing handles heterogeneity better because it is adaptive and therefore better applicable to dynamic approximate nearest neighborhood problems. In the case of the NN application for infrastructure and geospatial data models this approximation error is a function of LoD of CityGML data, the project phase of infrastructure data model and the grid size (GS).

$$\varepsilon = f(LoD_{geo}, LoD_{infra}, GS) \quad \textbf{Eq.(12)}$$

As shown in eq.(13) this approximation ε can be equal or different in various dimensional directions, this means for example as shown in figure 41 in the case of 3-dimensional directions of x, y, z:

$$\begin{cases} \varepsilon_x \neq \varepsilon_y \neq \varepsilon_z \rightarrow Tri - Axial\ Ellipsoid \\ \varepsilon_x = \varepsilon_y \neq \varepsilon_z \rightarrow Spheroid, equal\ semi\ axis \\ \varepsilon_x = \varepsilon_y = \varepsilon_z \rightarrow Normal\ Sphere \end{cases} \quad \textbf{Eq.(13)}$$

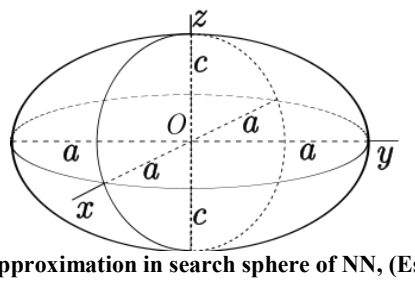


Figure 41 Different approximation in search sphere of NN, (Esfahani et al., 2014)

This thesis considers a search cube with equal approximation in all directions as illustrated in figure 42. According to this approximation one can define the equality for match findings procedure based on the eq.(14).

$$P_{geo}\ \underbrace{is\ equal\ to}\ P_{infra}\ \ if\ \begin{cases} \hat{x}_{geo} = \hat{x}_{infra} \pm \varepsilon \\ \hat{y}_{geo} = \hat{y}_{infra} \pm \varepsilon \\ \vdots \end{cases} \quad \textbf{Eq.(14)}$$

[17] The aspect ratio of a rectangular box is the ratio of its longest edge length to its shortest edge length.

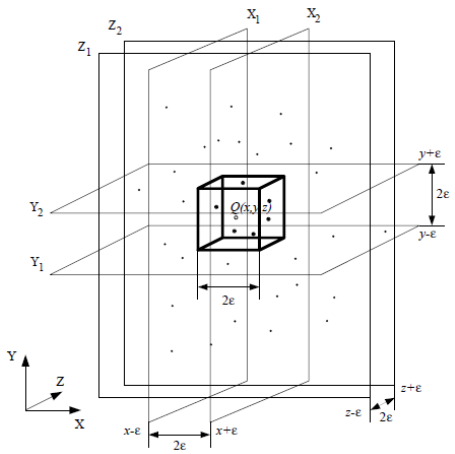


Figure 42 Search cube for NN with equal $\mathcal{E}$ in all directions

In the case of temporal (time dependent) data, there is a need to consider time interval for the match finding results. This can create a unique ID for the equivalents as in the following combination:

$$P_{infra} = f\big((coord), (time), (theme), (project), (alignment), (Stand.Deviation)\big) \qquad \textbf{Eq.(15)}$$

This approximation increases if the data structure is not prepared for NN queries. Such unprepared structure is shown in figure 43. In this figure infrastructure data such as road model and GIS data such as environmental model represented by tree and water body are fed into k-NN. In the first step the entities will be replaced by their bonding rectangles (MBR). In the next step the search or query cube will be constructed. The touching MBRs with the query cube will be selected as the possible results for the spatial querying step. In the practice the output of the spatial matching step through NN might include a huge range of multiple related and unrelated results if the approximation increases. This means the function of "is equal to" returns a set of points instead of one desired point.

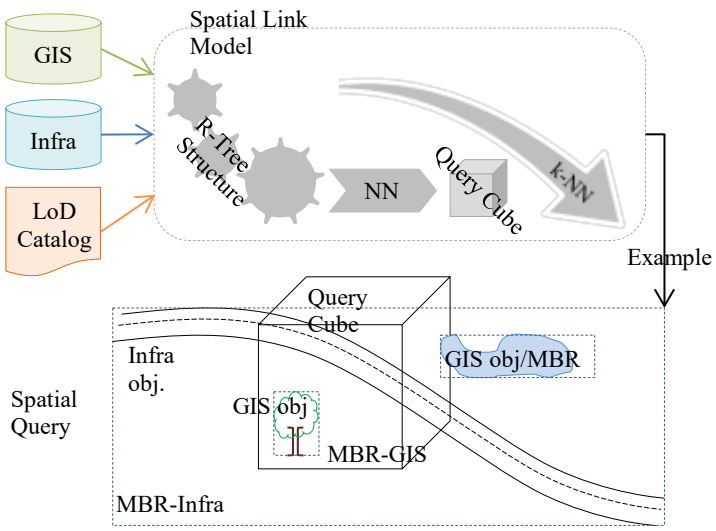


Figure 43 Example of search cube for NN

In addition in heavily populated geospatial or infrastructure data models composed of complex curves, an R-tree based flat organization of spatial objects could result in an occupation of large memory and awkward performance. One solution has been to use a hierarchical index system that organizes both the object and space. But since space is not hierarchically organized, there is no spatial connection among space at different levels of hierarchy. This has motivated to set up dynamic indexing model by combining R-tree with a hierarchical organization of the objects. One approach has been discussed in detail in Zhao et. al. work (1999), to combine R-tree and Voronoi structure.

For solving this problem, this thesis has used hierarchical filtering approaches for optimizing the results of the spatial matching.

Filtering approaches refine the matching results through two main criteria of hierarchy, a) considering the concept of Levels of Detail and b) applying the sub model pattern. I.e. the result of spatial query will be filtered in the next stage in accordance to

- the LoD in which the interoperability is being established
- the trades and products which are being observed

Therefore a tabular filter (TF) can be defined which is a function of LoD and product model (PM) as following:

$$TF = f(LoD_{geo}, LoD_{infra}, PM) \qquad \textbf{Eq.(16)}$$

As the objects available in the data model strongly depend on the product model an the LoD in which the project phase exists, therefore feeding this TF as a parameter into the spatial filter (SF) can adjust the result of whole filtering.

$$if \begin{cases} SF: Spatial\ Filter \\ G: Geometry \\ TF: Tabular\ Filter \\ SRel: Spatial\ Operator \end{cases} \overset{then}{\Longrightarrow} SF = f(G_{Geo}, G_{Infra}, SRel, TF) \qquad \textbf{Eq.(17)}$$

The effect of such filtering is shown schematically for scenario in figure 44. In the abstraction step the bonding boxes will be intersected with the boundary of the query cube. Then it will be decided by the system if the result is relevant to the query or not. The spatial querying and abstraction steps may occur several times to achieve the best related results. The maximum of iterations is also to be set in the system.

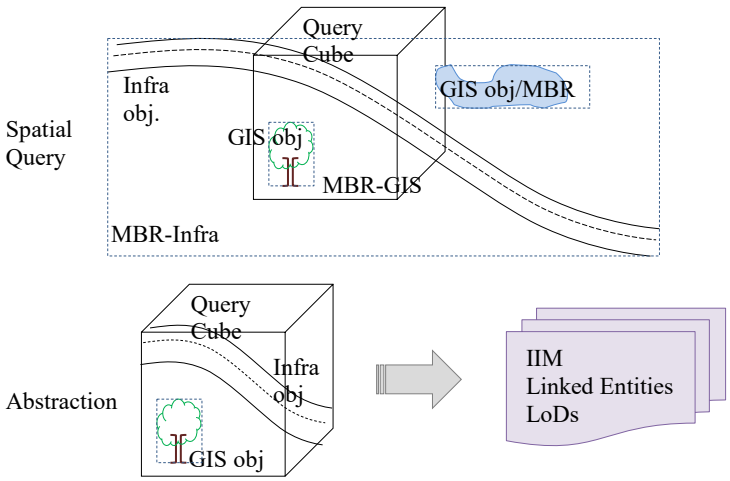


Figure 44 Spatial Filtering and Refining in IIM, (Esfahani et al., 2014)

In the reality, there will be always situations, where such automatism may not deliver proper results. Therefore in the real implementations should be considered that the last stage of refinement of spatial matching, should facilitate a manual selection for the operator in the case the previously provided results are not satisfactory. This will be discussed in detail in next part.

4.3 HIERARCHICAL FILTERING

According to Timpf (2008) *"a hierarchy is an ordered structure. Order can be established between individuals or between classes of individuals. The ordering function may be any function defining a partial order"* (Timpf, 2008):

$$\vartheta_1 < \vartheta_2 \quad iff \quad \exists f: \vartheta_2 = f\{\vartheta_i\}; \quad \vartheta_1 \in \{\vartheta_i\} \qquad \textbf{Eq.(18)}$$

As shown in figure 45 a vertex ϑ_1 in a graph is on a lower level than ϑ_2 <u>if and only if</u> (iff) there exists a function f, such that ϑ_2 can be calculated by applying f to a set of ϑ_i of which ϑ_1 is a member. The function f must be reflexive, antisymmetric and transitive (Timpf, 2008).

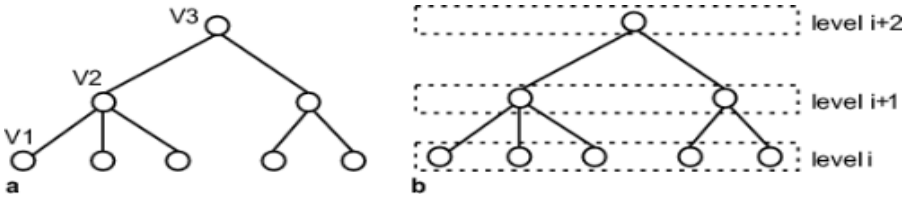


Figure 45 A partial ordering: tree; b levels in a hierarchy, (Timpf, 2008)

"The notion of levels is introduced by the idea that vertices at the same depth of the tree belong to the same level of the hierarchy. Thus, there are as many levels in the hierarchy as the tree is deep. Individuals on the same level share a common property" (Timpf, 2008). *"In the case of GIS data models, hierarchies appear in many different spatial contexts, for example road networks, political subdivisions, land use classes or watersheds. Hierarchies appear also in non-spatial situations such as organizational hierarchies as discussed previously"* (Shekhar & Xiong, 2008, p 431).

According to Pattee (1973) *"hierarchies have many advantages for computational sciences. They provably reduce processing time and increase the stability of any system"* (Pattee, 1973; Shekhar & Xiong, 2008, p 431). *"Besides according to Timpf (1992) they break down the task into manageable portions enhancing the potential for parallel processing, while inferences can be refined depending on the type and level of query"* (Shekhar & Xiong, 2008, p 431).

Development of the GIS enables data storage, data management and integration, and improvement of public participation. To serve different groups of end users, as well as the decision makers with a variety of key data and to assist them with providing up-to-date comprehensive information about the current situation of the project area, GIS has inevitably grown very huge. As GIS is getting currently used in lots of totally varying fields such as environment, health, planning and marketing, the heredity problems of analyzing and management of data are becoming more and more important.

In "global GIS", the geospatial data is very large and changes continuously in local areas. Volume of such geospatial data can be hierarchically organized into different levels of objective. Figure 46 illustrates the concept of spatial hierarchy developed by Coffey (1981).

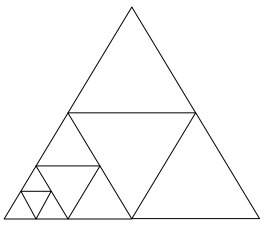


Figure 46 Spatial hierarchy through the use of triangles as a key spatial unit, (Coffey, 1981)

These hierarchies are spatial for example like local, regional and national levels or postal areas or model based such as levels of detail. *"One of the many roles of a Spatial Data Infrastructure (SDI) is to facilitate the integration of spatial data. It is well recognized that within a SDI a number of administrative boundary systems exist"* (Eagleson et al., 2002, p 186). As indicated by Timpf (2008) *"hierarchies are formed through abstraction mechanism, i.e. through factoring out commonalities in the description of several concepts into the description of a more general concept. Four abstraction mechanisms have been identified:*

- *Classification / instantiation*
- *Aggregation / decomposition*
- *Generalization / specialization*
- *Filtering"* (Timpf, 2008, p 432)

Normally generalization approach is used once entities from different classes are to be queried or compared in regard to each other. E.g. if a railroad from the infrastructure domain is supposed to get

linked to a main station from the building domain the generalization approach will be used as demonstrated in figure 47.

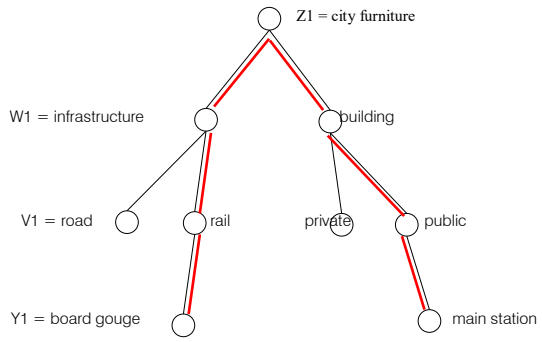


Figure 47 Visualization of generalization, (Timpf, 2008)

The generalization hierarchy relates a class to a super class. Y1, V1, W1, and Z1 denote classes in figure 45. Generalization which is a non-nested hierarchy is defined as forming a new concept by leaving out the properties to an existing concept.

$$V1 < W1 \quad iff \quad \exists f_{gen}: W1 = f_{gen}(V1) \qquad \textbf{Eq.(19)}$$

Timpf (2008) indicates the rules, determining what class belongs to what generic class, should be predetermined. The generalization function states explicitly what classes generalize to what generic classes. Therefore generalization mechanism will be used for refining the query result in accordance to product model.

Filtering is the mechanism used in this thesis for abstraction of the nearest neighborhood results in accordance to the concept of Level of Detail. According to Timpf (2008) *"the filter hierarchy applies a filter function to a set of individuals on one level and generates a subset of these individuals in a higher level"* (Timpf, 2008, p 433).

$$S1 < S2 \quad iff \quad \exists f_{fil}: S2 = f_{fil}(S1) \qquad \textbf{Eq.(20)}$$

Such filtering mechanism is used in this thesis to exclude the entity of infrastructure domain which belongs to an LoD other than the required one for the purpose of spatial matching. E.g. this approach can pass out the cross sections of the alignment at a specific location from being linked to the terrain data.

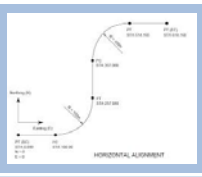
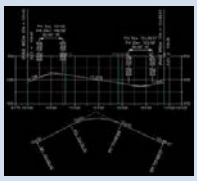
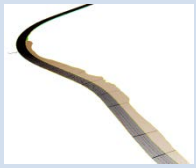
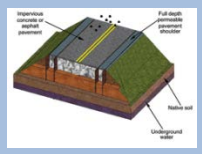
The method used in this thesis is to imply the concept of LoD for hierarchical organization and filtering the data. Each GIS commonly has a number of LoDs which corresponds to a map at a given scale or resolution or detailing of the digital representation of the modelled objects getting loaded. It is quite beneficial for interoperability to organize spatial objects into hierarchical-index structures so that it is possible to develop a mechanism for integrating different data models in homogeneous detailing degree.

Such homogeneity should be applied to the space as well as the object. Generally, in GIS raster and vector approaches are commonly used for representing hierarchical spatial structure. In R-tree based spatial structures hierarchy in space and object are handled separately (Zhao et al., 1999). Therefore there is a need for catalogs which provide description and explanation about LoD of objects in R-tree structures. Such documentation as explained in previous chapters (2.3.2.), divides the project life cycle in exterior and interior design points of view. The interface of both of them appears in OGC LoD 4. Breunig et al. (2011, 2012) have introduced a backbone for LoD definition for track alignment of tunnels. Inspirited from that, the author suggests the following LoD table for infrastructure data models.

Therefore according to LandXML, the class representation of alignment and its belonging objects will differ from LoD 0 to higher ones. This means for example the definition of profile and cross sections will not exist in LoD 0 to LoD 2 and will appear in LoD 3. In LoD 3 this will also be possible to perform first earth work calculations. The design of the structures such as tunnels and bridges shall run parallel in the same LoD as the infrastructure project design. In LoD 4 which is the interface of interior

and exterior aspects of the project pavement design will be performed as well as the drainage as network. The exact definition of other infrastructure components such as shoulders, islands, ditches, curbs and etc. will be done in the LoD 5.

Table 9 Table of LoD for infrastructure data models in coordination with GIS models, (Esfahani et al., 2014)

Models / LoD	Infra	Product Model
LoD 0	Horizontal 2D Alignment	
LoD 1	Horizontal and Vertical Alignment	
LoD 2	3D Alignment Structure Segmentation Junction, Tunnel, Bridge	
LoD 3	Parametric / Geometric design Cross Sections Profile Cut and Fill Structures such as Tunnels and Bridges according to BIM-LoD3	
LoD 4	Infra object envelope Geometric design Pavement design Drainage Network design Structures in accordance to BIM-LoD4 Street furniture	
LoD 5	Infrastructure object components Hydrological design Structural design in accordance to BIM-LoD5 Travel Lanes, Turn Lanes, Shoulders, Islands, Curb, Guard Rail, Side Walk and Ditch	
LoD 6	MEP model for Infrastructure Design as well as for the Tunnel and Bridge structures Irregular objects	

As mentioned before, different objects of the alignment class in the LandXML schema can be modelled in various levels of detail. This thesis introduces the LoD allocation to the objects of the class diagram alignment in accordance to figure 48. The hierarchical filtering of this thesis has used the schema of LandXML 1.0 for the abstraction of spatial matching results based on the interlinking task.

Moreover, as discussed previously the LoD 4 is the interfacing environment between interior and exterior design, or GIS and BIM world. Therefore colored in purple most of the street furniture, structural properties and geometrical design model objects belong to the LoD 4.

Besides, because the geometry is one of the most fundamental elements in geometrical infrastructure design, therefore the main focus has been to deal with the geometry descriptively and attributes are not pointed out. In this thesis, author differentiates between level of approximation, level of accuracy and the LoD. Level of approximation is used to classify the accuracy of the results and outcomes of

applications of different methods and algorithms. Accuracy describes the resolution of implying or changing the parameters of each algorithm and the search range. LoD will be defined in accordance to the provided document in table 9.

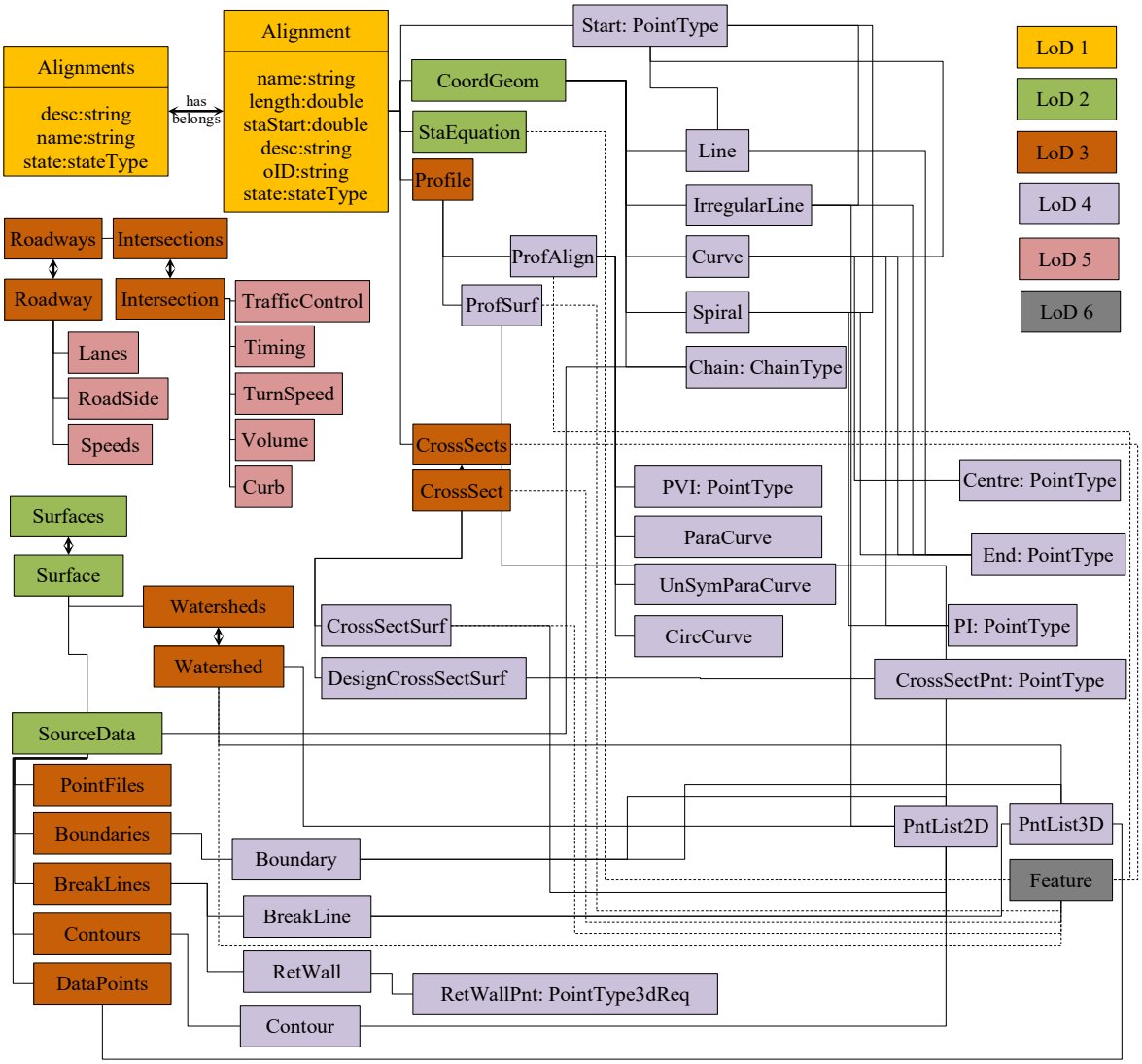


Figure 48 Suggested LoD classification schema for alignment class diagram shown in the public LandXML

For the purpose of interoperability as explained in previous chapters it is very important to design a task / products model based catalog which defines the requirements for selecting a minimum LoD from each product model. For example if according to this catalog the infrastructure data model such as road model shown in figure 49-a has a level of detail α, and geospatial data model such as terrain model shown in figure 49-b is in level of detail β, LoD of infrastructure information model arises from / is a function of LoD of infrastructure and geospatial data models described in eq.(21).

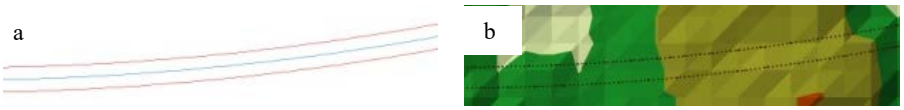
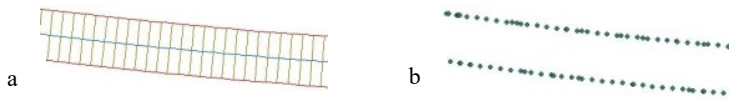


Figure 49 Model of road (a) and terrain (b) in infrastructure catalog in their own LoD

$$LoD_{IIM} = f(\alpha, \beta) = \gamma \qquad \textbf{Eq.(21)}$$

Therefore in a reverse approach, to be able to make the two data models interoperable, the first step will be the segmentation of the infrastructure and geospatial data model as illustrated in figure 50 according to the LoD of infrastructure information model, for example to transfer terrain and road model to the LoD γ.



$LoD'_{Infra} = \gamma$ $LoD'_{Geospatial} = \gamma$

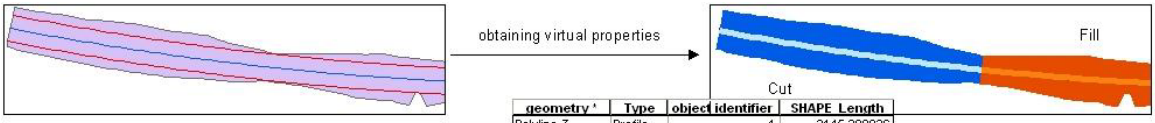
Figure 50 Model of road (a) and terrain (b) in infrastructure catalog in the LoD of IIM meaning γ

This reverse approach prepares the data models to build a homogenous R-tree data structure in spatial database for interlinking, interpolating, or match finding, through nearest neighborhood algorithm. This can result in obtaining virtual properties of geospatial data for infrastructure data. The example below shows how to obtain Z value for infrastructure design data:

$$P(x,y)^{local}_{Infra} = P(x,y)^{global}_{Geospatial}.Z \qquad \textbf{Eq.(22)}$$

This could be any other spatial or non-spatial attribute stored in the database of geospatial data model, such as environmental data (trees located in the design site), protection areas (flood area or animal protection area), or climate data (tornado risk), or hazard risks. This facilitates to plan in respect to risk management parameters which increases the planning efficiency.

Example above shows how engineering structures which belong to the LoD 2 in accordance to the table 9 can be calculated through integrating the infrastructure data model with LoD 1 with GIS data in a higher LoD, here LoD 2. Volume of cut and fill, length of bridges, tunnels, intersections and junctions are also other results of such interoperation.



geometry *	Type	object identifier	SHAPE_Length
Polyline Z	Profile	1	2145,209836
Polyline Z	Street	2	2145,209835
Polyline Z	Cut	3	199,352798
Polyline Z	Fill	4	416,040713
Polyline Z	Cut	5	260,025462
Polyline Z	Fill	6	853,750176
Polyline Z	Cut	7	21,668836
Polyline Z	Fill	8	4,333775
Polyline Z	Cut	9	52,00501

Total Cut and Fill Volumes

Code	Sum_Volume	Sum_Area
-1	100570,538207	35713,680088
1	91673,01682	28190,50655

Figure 51 Example of interoperability and obtaining virtual properties

The Spatial Link Model as the result of interoperability including the links and virtual properties obtained from geospatial data will have more objects which belong to higher LoDs. This is demonstrated in the figure 51. Thus this object can be organized under higher LoDs virtually.

$$LoD^{Link\ Model}_{Infra} = \alpha + \varepsilon \qquad \textbf{Eq.(23)}$$

In a revers approach γ or $LoD^{Linked\ Model}_{Infra}$ to be set as the minimum of (α,β) with some manual adjustments ε if the operator believes γ is too small and might make performance conflicts. This means for instance if the GIS data model is provided in LoD 1 and infrastructure data model is in LoD 5, the interoperability should have LoD 1 unless one of the following scenarios happens:

- Either transferring both data models into the coarser LoD here LoD 1 causes loss of information for the data model with higher LoD, which is in this example the infrastructure data model in LoD 5.
- Or spatial linking in the coarser LoD here LoD 1 causes computational insufficiency for example being hugely time consuming for any of data models in the performance of nearest neighborhood.

In such cases the operator should apply an adjustment of ε to the LoD of Link Model. For the example above this means to transfer both data models into the LoD 2 and perform the interoperability in this level.

4.4 OTHER FUNCTIONAL LINK METHODS

There are other functional link models and researches running about them. Such functional link models comprise different linking approaches at various levels. Some of them perform this modelling

on the language level on Meta Data or Meta Models using Meta classification approaches. Some of them apply semantic linking approaches on schema level.

High level functional linking approaches are algorithmic and use reasoning functions on both schema and instance levels. The conceptual linking methods create equivalent classes of semantic index terms by using description logic reasoner. The reasoner can work on the ontology level or rules level.

One of such conceptual models is the ontology based link model which in being developed in the institute of construction informatics of technical university of Dresden. Kadolsky et al. have created an ontology framework for rule-based inspection of eeBIM-systems and applied it on Energy Models in building domain (Kadolsky et al. 2014).

In the framework of the ISES Project the team has developed an ontology based concept and dedicated the algorithms and functions to link external data models to BIM models.

Kadolsky et al. (2014) have proposed applying "*inference rules to pre-check the input data and to pre-analyze the energy performance"* on a lower detail level, prior to simulation in an eeBIM-ontology-based (Energy Enhanced BIM) framework (Kadolsky et al., 2014, p 293).

The ontology specification "*provides the concepts and relations to describe the building, the external data like the climate data as well as the linking between the BIM-concepts and the external data"* (Kadolsky et al., 2014, p 293). Furthermore, "*constraints and calculation methods are transferred as far as possible into logical rules. The surrounding ontology platform enables the integrating of the input data and manages the process of the calculation methods. It also defines certain external connection points for calculating methods, which cannot appropriate represented in rules"* (Kadolsky et al., 2014, p 293).

Kadolsky et al.'s ontology framework helps to identify building designs which have problems to fulfil energy performance requirements even at a very early stage of the design phase. So, the ontology follows the idea of typical engineering regulations including simplified methods for easy and practical use.

Ontology based data extraction has been practiced in the world of GIS since many years. Web Map Services (WMS) or Web Feature Service (WFS) have been OGC's and GML's standards for sharing geospatial information among different end users. However, "*ontology reasoners only can apply to ontology instances such as Resource Description Framework (RDF) instances or Web Ontology Language (OWL) individuals, which correspond to the database records and WFS feature instances"* (Zhao et al., 2008, p 2).

4.5 ADVANCES AND LIMITATIONS OF FUNCTIONAL LINK METHODS

Generally, the link methods specify linking techniques *"according to which data models are linked and the associated linking operations / linking functions that are permitted. It may also be seen as a technique for the formal description of link approaches, usage constraints and operations. The facilities available vary from one link model to another"* (Scerbakov, 2010, p 2).

In contrast, the functional link model provides a unified approach to manipulation both data models and link procedures. Main idea of the functional data model is a definition of all components of an information system in the form of functions. Thus, for example, the functional link model defines linking rules, sequences and abstraction approaches as so-called link model functions. Moreover, a Functional Rule Manipulation Language is a number of rule manipulation functions which can be applied to rule model functions. Finally, users are provided with a special mechanism to define their own functions which can be seamlessly combined with data models and data manipulation functions mentioned above.

Still the functional linking approaches have their limitations. In rule based linking methods when selecting related information or properties, the type of relation cannot be observed. In the case of a self-sequencing method of rule based linking the sequence of the rules which has resulted into the related information cannot be controlled. Some objects, even if they rationally and reasonably relate,

might remain unlinked due to the fuzziness. The ruling mechanism can quickly lead to overloaded user interfaces and distract the operator. On the other hand the ruling mechanism may remain incomplete.

"Adding semantic or ontology might be one of the solutions to overcome some of the limitations. Single ontology approaches use one global ontology providing a shared vocabulary for the specification of the semantics. However all information sources and domains must be related to this one global ontology" (Arens et al., 1996).

Multiple Ontologies or adding a hierarchical terminological knowledge base sequencing approach is one of the solutions. This will cause the global ontology become a combination of several specialized ontologies. *"The advantage of such approaches is that each source ontology can be developed without respect to other sources or their ontologies. However, the lack of a common vocabulary makes it difficult to compare different source ontologies"* (Laallam, 2014). Identifying semantically corresponding terms of different source ontologies which is called inter-ontology mapping, e.g. definition of equality or similarity can solve this problem. But the mapping has also to consider different views on a domain e.g. different aggregation and granularity of the ontology concepts which is unpractical.

In the spatial domain as also illustrated completely in chapter 2 and according to Cova et al. (2008) *"to efficiently support ontology-based reasoning on geospatial data, it is necessary to keep legacy data stored in geodatabases and other data files while providing an ontology-enabled interface to translate user requests into queries to legacy data stores"* (Cova et al., 2008, p 372). But, it is unclear how the extracted ontology information can help translate user requests into queries to legacy data sources.

Transforming all legacy geospatial data to ontology form is time consuming, error prone, and inefficient. *"In the case of geospatial data ontology tools cannot efficiently manipulate ontology instances in large quantity due to memory consumption. Moreover, it is not cost-effective for ontology tools to include the functionalities provided by geodatabases and WFS services such as transaction management and spatial query"* (Zhao et al., 2008, p 2).

Chapter 5

Implementation of the Proposed IIM Method

For the eMM suggested in this thesis a rule based interoperability method between arbitrary data models of spatial domain has been developed. Following chapter demonstrates implementation of linking heterogeneous data models with a rule based link model for rule based filtering and query purpose. The implementation of this method enables transaction of data in spatial domains run loss less.

As a case study on spatial data, these all implements and customizes a spatial link model between CityGML and LandXML data models so that virtual properties are obtained in different Levels of Detail (LoD).

The system architecture and the implementation which has been applied on the case study of this thesis namely infrastructure and geospatial data models are described in this chapter.

With enhanced interoperability, this implementation will have a common end-to-end solution that supports the entire project lifecycle. Working in unison, this comprehensive technology solution leverages data management and interoperability, powerful spatial analysis, advanced 2D and 3D visualization and rendering of the project in its intended location.

The suggestion of grid size out of the data models, interlinking of the entities, creating new entities and allocating virtual properties to the entities are some of innovations of this implementation. After creating new entities, the grid size will be recalculated for the new geometries and bounding box dimensions and adjusted again in a new interval for new tasks and queries.

Keeping the R-tree in a dynamic flexible format which means the ability of resizing the data structure for adding new objects to the data base and data models or removing objects has also been one of the challenges in the implantation. Removing objects happens rarely but still possible under specific circumstances.

5.1 IMPLEMENTATION

Daum and Borrmann (2014) believe *"there are gaps in the analysis and handling functionality available for BIM, concerning the formal spatial exploration of models and the extraction of sub-models. They have therefore introduced new BRep-based methods for the verification of topological predicates as part of the SQL for BIM which describes a mapping between the 9-Intersection Model and the boundary-based examination of topological predicates"* (Daum & Borrmann, 2014, p 14). There are two main approaches for spatial query and filter, one is oracle based and the other is Structured Query Language (SQL) based.

Shown in figure 52 Oracle Spatial (OS) as a full function program set for quickly and efficiently storing, accessing and analysis spatial data which provides SQL model and many functions, making it quite easier to store, retrieve, update and query the spatial data sets in the oracle database has been the leading spatial data base of the implementation process of this work.

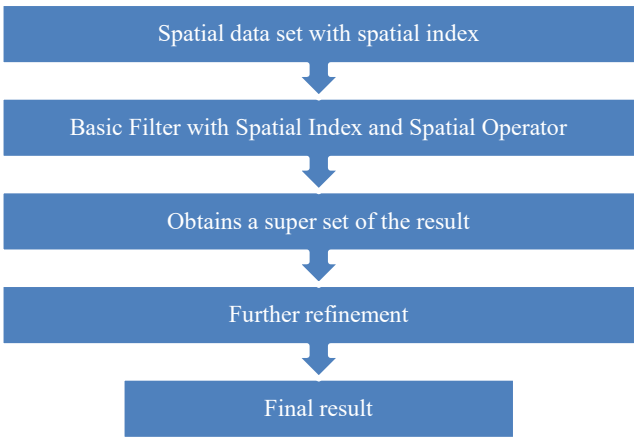


Figure 52 Query model of Oracle Spatial, (Chunhui, 2009)

"Spatial Data Engine is the tool that is used in this dissertation to store and manage spatial data in their chosen relational database management system (RDBMS). The method to access data is in accordance to SQL specification. On this basis, this thesis has implemented the realization of spatial linking" (Chunhui, 2009).

"As defined in ISO99, SQL is extended to manage data like texts, as well as images, spatial data, or to perform data mining"[18]. ISO/TC 211 is in the field of digital geographic information. The standard is grouped into several parts. *"ISO/IEC 13249 or SQL/MM standardizes several major database systems provide extensions to support the management and analysis of spatial data in a relational database system"* (Stolze, 2003).

Queries to databases are posed in high level declarative manner. SQL is the "lingua-franca" in the commercial database world. Standard SQL operates on relatively simple data types. SQL3 / OGIS (GML) support several spatial data types and operations. Additional spatial data types and operations can be defined. In oracle query language, *"oracle spatial is an integrated set of functions and procedures that enables spatial data to be stored, accessed, and analyzed quickly and efficiently in an Oracle9i database. Oracle spatial supports the object-relational model for representing geometries. The object-relational model corresponds to a SQL with Geometry Types implementation of spatial feature tables in the OpenGIS ODBC/SQL specification for geospatial features"*[19].

[18] http://www.isotc211.org

[19] www.oracle.com

Table 10 Comparison between different standards in accordance to selected query criteria

Criteria \ Standard	ISO 19107	SQL 2008	Oracle	OGIS
Spatial Queries	simple	basics	x	strong
Spatial Filtering	-	simple	x	strong
Geometries	x	x	x	strong

According to Shekhar among the mentioned query languages SQL and OGIS (GML) standards are being adopted by many vendors and are similar in regard of spatial data types and operations. Shown in table 10 OGIS has the strongest query language for spatial linking. But the syntax might differ from vendor to vendor and the users may need to alter the queries given in these two standards to make them run on specific commercial products.

Data access is built around the standard cursor model defined in SQL. The data access model builds on the traditional model by supporting spatial and raster constraints in a query. The query model of ArcSDE is shown in figure 53.

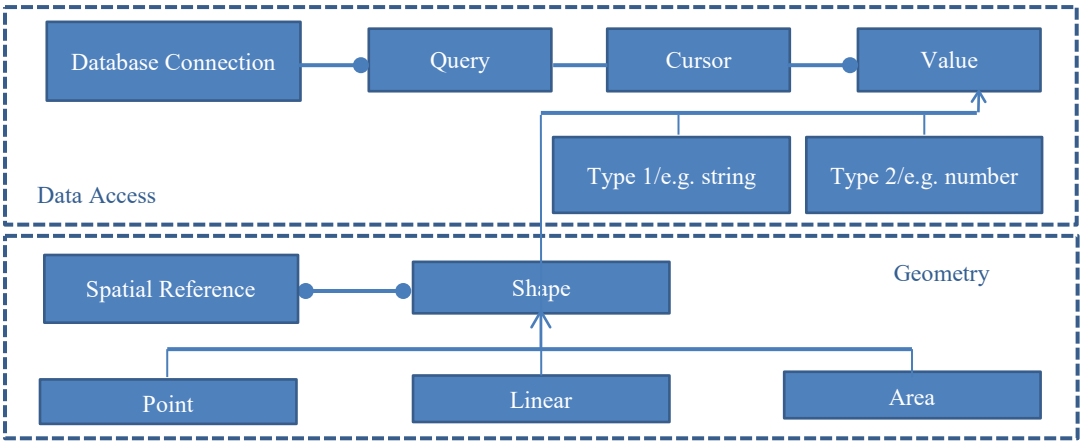


Figure 53 Query model of ArcSDE, (ESRI, 2001)

"Geospatial Data Abstraction Library (GDAL) is a translator library for raster geospatial data formats. The OGR Simple Features Library is a part of the GDAL library. OGR is a C++ open source library providing read (and sometimes write) access to a variety of vector file formats including ESRI Shapefiles, S-57, SDTS, PostGIS, Oracle Spatial, and Mapinfo mid/mif and TAB formats etc.[20]*. The GEOS library provides the spatial predicates originally used in PostGIS, now OGR, and MapServer. It is a C++ topology suite. The complete functionality includes all the OpenGIS Simple Features for SQL spatial predicate functions and spatial operators"* (Chunhui, 2009).

"In OGR's abstraction layer, the space predicate functions are implemented as virtual interface, and OGR does not provide specific implementation. GEOS library implements these functions, and provides external interfaces for better bonding with OGR. The OGR/GEOS libraries implement the spatial predication functions. Though such functions are just spatial operators of geometries, which can help us implement the spatial query based on two or more spatial data models" (Chunhui, 2009).

"In other words, actually, OGR/GEOS libraries have no query model. The implementation of spatial queries is left for the developers" (Chunhui, 2009). According to Childs (2009)

- Improved versatility and usability
- Optimized performance
- Few size limitations
- Easy data migration
- Improved editing model
- Storing raster in the geodatabase
- Customizable storage configuration

[20] http://www.gdal.org/ogr/

- Allows updates to spatial index settings
- Allows the use of data compression
- Registered by INSPIRE (Infrastructure for Spatial Information in the European Community)

The link model developed in this thesis is adapted to ESRI query model, with the difference that the geometry data does not need necessarily to get converted into shape format. The data models will remain in their original format. To remedy the memory and performance problem as discussed in previous chapters, R-tree based multidimensional indices will be devised to the infrastructure data. This prepares the data to be embedded in a RDBS which can directly handle spatial data and spatial join algorithms.

Figure 54 shows the system architecture developed in order to establish interoperability between infrastructure and geospatial data model through spatial linking.

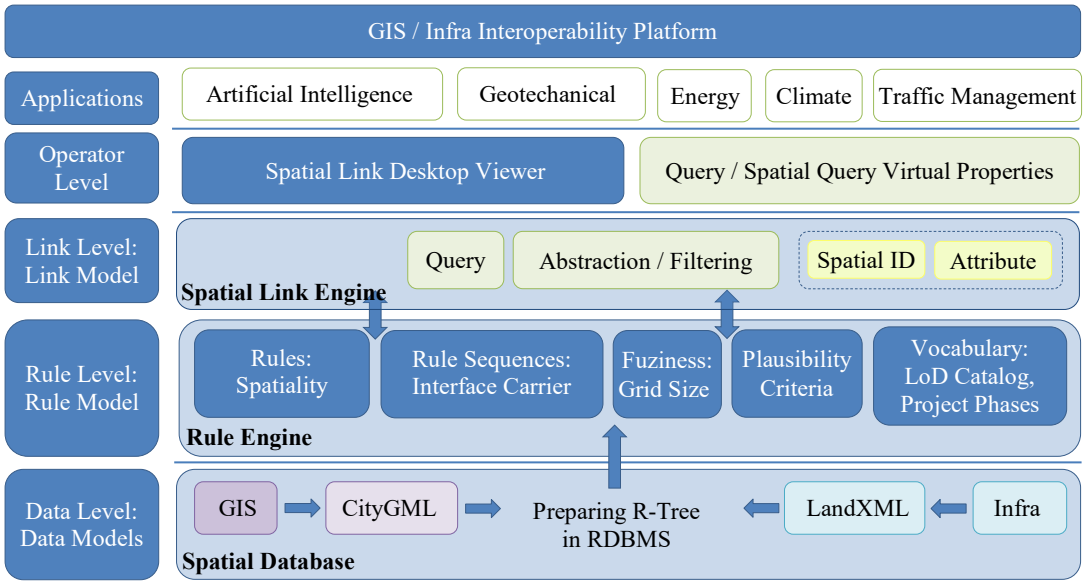


Figure 54 GIS / Infra interoperability system architecture

As shown in figure 54 system architecture for GIS/Infra Interoperability Platform is a bottom-up system which considers in the lowest level the data models involved in infrastructure tasks. Such data models are normally based on a standardized format such as LandXML or ProVI. GIS data is also based on GIS standards like CityGML. This level is a container where data which is being interoperated will be inserted.

This data will be checked and prepared for RDBMS data base according to their structure. If a spatial structure is not available, spatial IDs will be allocated according to the geometries based on R-tree concept. For R-tree using balancing algorithm of quadratic complexity, quadratic methods should be considered. An example of program codes of structuring such r-tree is presented below:

```
using namespace boost::geometry;
typedef std::pair<Box, int> Value;
index::rtree< Segment, index::quadratic<16> > rtree(index::dynamic_quadratic(16));
```

For the purpose of building R-tree structure it is important to virtually create a grid out of the objects of infrastructure and geospatial domain. For example alignment elements of the infrastructure domain should be broke down into the grid size. In this context, the interface carrier will be defined according to which the first recommendation for the grid size will be proposed by the system.

The suggestion of grid size comes in the first step out of the data models and their bounding box dimensions. After interlinking of the entities, new entities might be created subsequently. The next step is allocating virtual properties to the entities. Therefore it is very important that the R-tree is built dynamically to have the ability of being re-sized if new objects get created.

After this step it is important to decide either the proposed grid size according to the interface carrier is the proper grid size or not. As explained in chapter 3 and 4 the gird size should match to the planning

phase of infrastructure domain and LoD of geospatial domain. After creating new entities, the grid size will be recalculated for the new geometries and bounding box dimensions and adjusted again in a new interval for new tasks and queries. This means that grid size should be a dynamic value which changes according to the project progress and increase of the details and corresponding LoD. Therefore it should be possible to define the accuracy of the R-tree dynamically and quadratic for queries in multiple directions.

Keeping the R-tree in a dynamic flexible format which means the ability of resizing the data structure for adding new objects to the data base and data models or removing objects is very essential for the implantation. Removing objects happens rarely but still possible under specific circumstances.

After having structured the spatial IDs for the data models, the rule model and rule engine must be established. In the scope of this thesis, the rule engine has considered dealing with spatial domain and therefore this thesis has focused on spatial link engine as the rule based linking approach. Thus all of settings for the rule engine have been developed for the spatial domain. Furthermore the vocabulary has been developed for the spatial link engine which. The vocabulary which this thesis considers to be functional and thus advisable for the case of spatial domain is the LoD catalog. Therefore LoD catalog should be defined to determine linking rules and sequences of linking for the spatial link engine. Inserting the LoD catalog and the spatial data base into the spatial link model the IIM concept steps into the link level.

In the link level after having established an R-tree for non-spatial data, for the spatial linking approach it is pretty important to determine the tuples who's MBRs overlaps with a spatial query region MBR. This step is not computationally time consuming since R-tree structures support nearest neighborhood algorithms. Besides even if due to any reason, the Tree structure does not support nearest neighborhood function, at most four computations are required to determine whether two rectangles intersect. Spatial linking according to criteria defined in the rule engine will be performed in this step.

In the link level the linked data according to linking rules should be checked and linking results must be controlled according to plausibility criteria. The plausibility criteria are one of the critical items which are considered in this thesis very briefly for the spatial domain yet should be discussed in detail. After the plausibility check has run, the reasonable results must be abstracted and filtered according to set filters such as hierarchical or tabular filters discussed in chapter 3 and 4.

In this level, the tuples that passed the spatial filtering are fed to the refinement step, where the exact spatial representation is used and a spatial predicate is checked on these spatial representations. The refinement step is the challenge of the programmers and developers, since it is computationally expensive, but the number of tuples it processes in this step is less, due to the initial filter step. Figure 44 illustrates an example of the filter-refine policy and the performance of these steps on the case study of this thesis.

In the next level after achieving proper results the user or operator benefits from a viewer where the linked data will be shown in the viewer and virtual properties will be listed in a table of joined related attributes. Most of spatial data formats are supported by this viewer, for those data formats which are not supported the segmented elements of R-tree will be shown in the closest geometry (point, polyline or polygon).

As mentioned previously in chapter 2 for the purpose of data and information exchange standardization in infrastructure models in road; rail-road and water planning LandXML has been chosen in most of the countries. In LandXML an alignment is an object of class *Alignment* which has all class attributes among them a position in the form of Start and *CoordGeom* which defines the alignment type shown in figure 53. Through these attributes and some interpolations the spatial identity of the object alignment will be set in the spatial database (Esfahani et al., 2014).

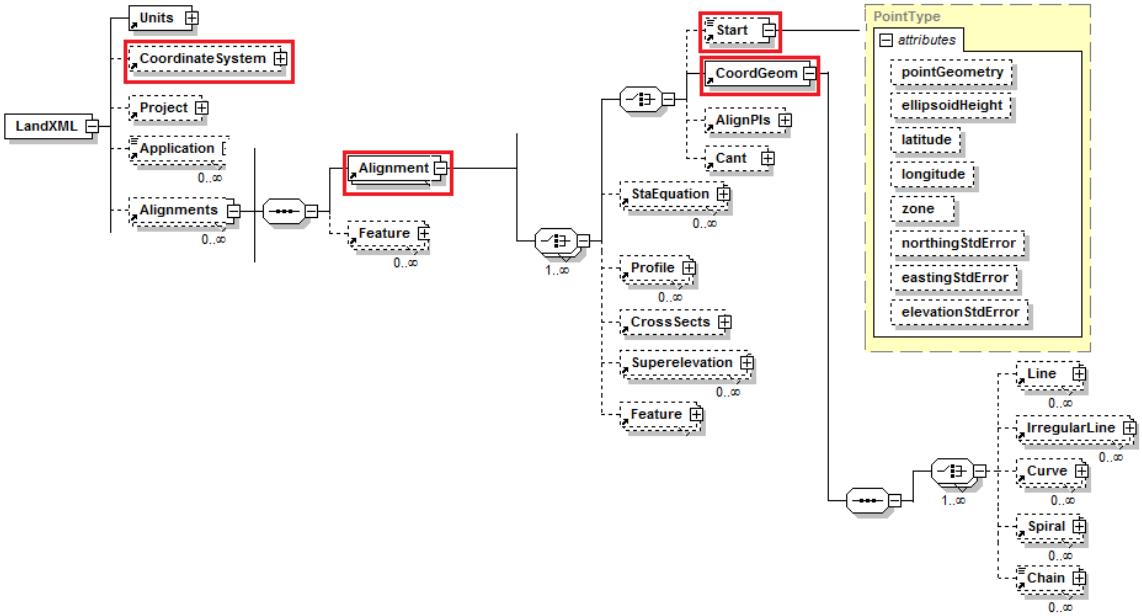


Figure 55 Alignment partial class representation (LandXML[21], last visited 2016)

From the GIS side each object modelled in CityGML has a bounding box which is the envelope containing the geometry of the object and spatially relating it to a reference system. E.g. the objects positioned in the coordinate reference system DHDN/Gauss-Krueger 3° (2nd zone), and normal heights above sea level (DHHN92) are referred to (Esfahani et al., 2014) and bounded by bounding box represented by XML codes below:

```
<gml:boundedBy>
        <gml:Envelope srsName="urn:ogc:def:crs,crs:EPSG:6.12:31466,crs:EPSG:6.12:5783">
                <gml:pos srsDimension="3">5616000.0 2540097.5 54.5</gml:pos>
                <gml:pos srsDimension="3">5673522.3 2576495.6 172.9</gml:pos>
        </gml:Envelope>
</gml:boundedBy>
```

Such reference system should be normalized for all dataset involved in the integration process. Either the datasets have a reference system which should be transformed into a project reference system, or they have no definition. In this case they need to be given a reference system. The codes below show how such reference system allocation can be programmed.

```
prjSR = createSpatialReference(prjHorizontalSR, prjVerticalSR)
for dataSet in dataSets
        dataSet->getSpatialReference(&currentSR)
        if currentSR is Nothing
                dataSet->putSpatialReference(prjSR)
        else
                accuracy = convertSpatialReference(dataSet, prjSR)
        dataset->Accuracy = accuracy
```
[22]

The accuracy involved in this normalization step should be considered for the proximity of the NN calculations such as defining the equality. This is conceptually explained in the following codes:

```
bool Equal (Object obj1, Object obj2)
        double x1, y1, z1, x2, y2, z2
        obj1->QueryCoords(&x1, &y1, &z1)
        obj2->QueryCoords(&x2, &y2, &z2);
        return (|x1 - x2| <= accuracy && |y1 - y2| <= accuracy && |z1 - z2| <= accuracy);
```

The source codes below show one of the NN processes which link the alignment from infrastructure data domain with the surface model from geospatial data domain as illustrated in figure 51.

[21] http://www.landxml.org/schema/LandXML-2.0/documentation/LandXML-2.0.html

[22] Not primary source codes

```
for alignmentPart in alignmentParts
        relationship = createRelationship(alignmentPart, tinSurface, cardinality)
        relationships->Add(relationship)[23]

for Alignment in Alignments
        Alignment->getType(&alignType)
        If alignType is Cothoid
                clothoid = Alignment;
                clothid->InQueryCube(&object)
```

"Applying procedural technologies for multi-scale modeling provides the possibility for a stringent definition of dependencies between individual geometric elements on different LoDs. Thus, LoDs of models are not isolated from each other, but inter-related by means of the construction history" (Borrmann et al., 2014b, p 267).

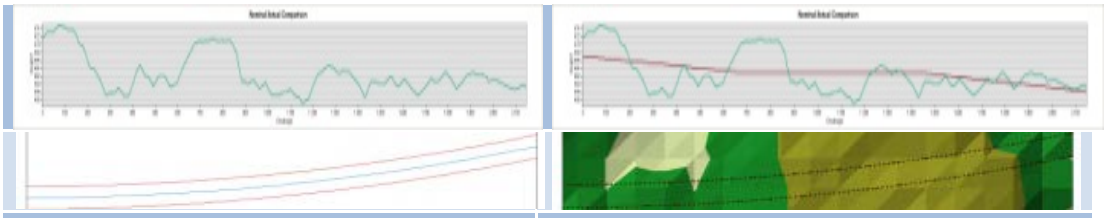
The LoD catalog contains the definition of objects in object model and their preciseness in respect to the 7 LoDs developed in this thesis in table 12. Therefore in accordance to figure 51, in the implementation phase after the linkage to the data in their original format, the input data should get normalized so that it is possible to handle the spatial indices in a common reference system. The source code below shows a sample implementation for such a step.

This thesis uses a multi model approach for the integration of product models in cross-domain analysis processes. This enables the integration of relevant information to the infrastructure data model from more than one product model. Besides as in the infrastructure projects this approach is very powerful in the case of variety in model types and product-system models. This approach simplifies the increasing complexity of the underlying product data models. Export and import through data models cause loss of data because of limited modeling expressivity of domains and formats and is very difficult to preserve object identities. Therefore for the data models transfer between the processing steps it is enough to exchange the spatial linking rules.

5.2 CASE STUDY

With the application of any data domain which can be modelled spatially like Energy, Environment, Climate, Infrastructure, etc. this approach enables linking and querying related objects using their spatial property.

Table 11 Results of Interoperability between Geospatial and Infrastructure Data Models

	Alternative 1	Alternative 2
Engineering Point of View		
Monetary Point of View	?? € no reference to surface model to obtain monetary information	Link to the data base of unit prices for different structure types, here tunnel

For example this methodology monitors and evaluates energetically the changes caused on a railway station due to updates on a railway alternative. As another example this facilitates the monetary comparison between 2 alternatives of a road far before the detail planning stage in accordance to the history recorded in the pricing catalogue of similar projects. Table 11 is a demonstration of such comparison.

"However, there are limitations to the degree of accuracy on coarse levels which can be propagated to finer ones. These limitations are mainly driven by operations in the construction history which only produce results if certain conditions are fulfilled by their operands" (Borrmann et al., 2013, p 80). Therefore the author bases the design in large scale and according re-pricing list of different tunnel

[23] Secondary codes

types and according to the tunnel data base of OBERMEYER Company. Result of the linking approach in shown in table 12:

Table 12 Results of pricing query through interlinking approach to the tunnel database of OBERMEYER company

Structures_Result

object i	Type	Cnt	Last_Layer	Last_Sta	Cost	Sum_SHAP
6	Tunnel	4	Var_0001	State1	39824788.990147	1443.7765

Besides, such spatial matching enables obtaining virtual properties for data involved in matching process. Figure 56 shows how the case study alignment of this thesis is divided into its structural parts such as tunnels, bridges, cut and fill sections in respect to the surface model of GIS.

In figure 56, parts in blue color are defined after matching process to be embankments, red parts are cuts and yellow green parts are to be earth. Different alternatives are evaluated in this figure according to fixed criteria defined in the rule engine. The profiles of these alternatives are shown in the profile graph of the figure 56. The codes which are described below are references of such criteria check.

```
getStructurePartType(double refElv, double checkFrmElv, double checkToElv, double accuracy)
        if (checkFrmElv == refElv && checkToElv == refElv)
                return L"Earth"
        if (bridgeCriteria)
                return L"Bridge"
        if (tunnelCriteria)
                return L"Tunnel"
        if ((checkFrmElv > refElv - accuracy && checkFrmElv < refElv) || (checkToElv > refElv - accuracy &&
checkToElv < refElv))
                return L"Fill"
        if ((checkFrmElv <refElv + accuracy && checkFrmElv > refElv) || (checkToElv < refElv + accuracy &&
checkToElv > refElv))
                return L"Cut"
```

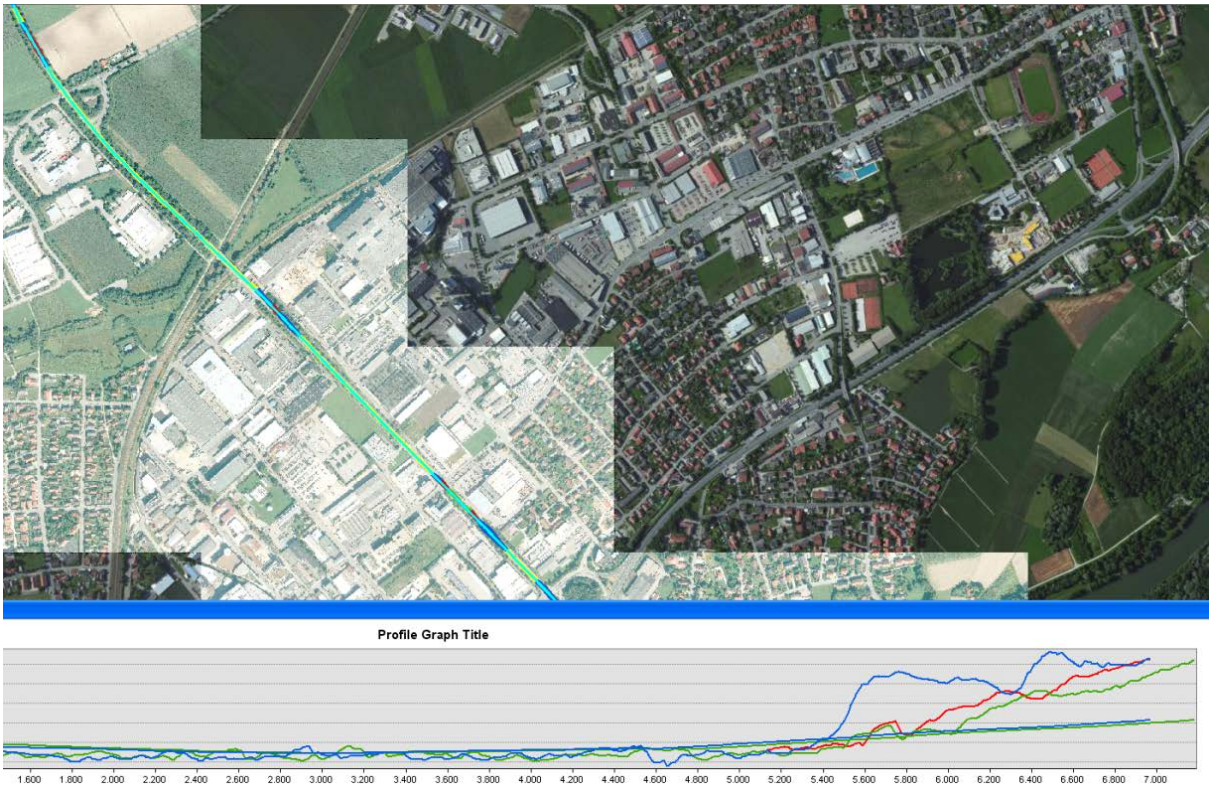


Figure 56 Spatial Matching, obtaining virtual properties for infrastructure data out of geospatial data

Figure 57 shows the process explained above in detail. In this procedure as illustrated in figure 57, bottom left the area of influence of alignment from the infrastructure domain will be defined using the

allowed slopes of gradient defined in rule engine. Grid size will be defined by the rule engine and according to that an R-tree structure will be constructed.

As demonstrated in figure 57, bottom right the alignment will be interpolated into this grid size so that the alignment can get linked partially to the related objects from other domains. As shown in figure 57, top left, from the geospatial domain, terrain data will be queried for the grid in which the alignment data is interpolated.

From there the alignment data will be linked to terrain data and the linked data will be checked according to the criteria which have been predefined in the link engine. The alignment will be then subdivided into its structure parts. The figure 57, top right shows the result of this process and definition of cut in red and fill in blue after the linking step.

As shown in figure 57 interlinking of the alignment model from the infrastructure data domain with the surface model from the geospatial data domain results in defining the geometry and area of influence of the related structures such as bridges, tunnels, reinforcements, cuts and fills, etc. The link model is also exchangeable limited to the geometrical links.

The grid size will be then recalculated according to the geometry and the size of the bounding boxes of the new created objects.

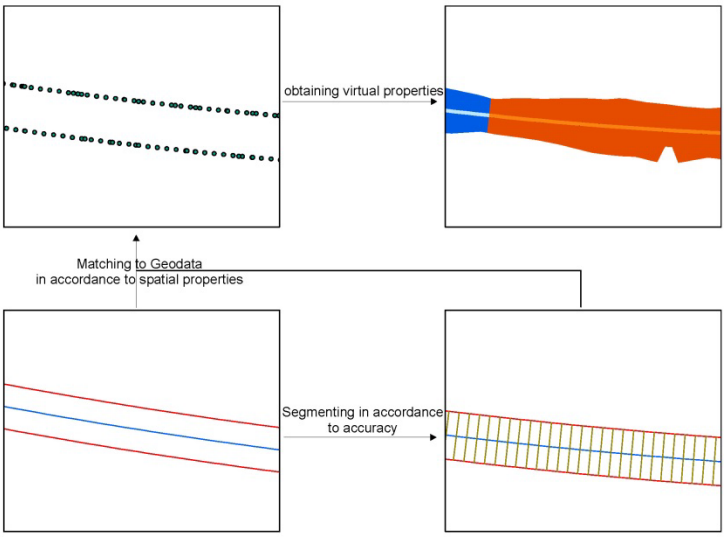


Figure 57 Application of the interlinking of infrastructure and GIS objects

The other part of the rule engine is to define conflicts in regard to the planning procedure. For example the rule engine defines the alignment to be from the bridge type where ever it is located over a river. Therefore if the axis crossing a river is not a bridge axis the system detects a conflict at this position. Another type of spatial conflict detection is for example to be too close to a certain object. I.e. not to enter the protection buffer zone of a nuclear power plant. Such zones are defined as no-go area in the rule engine and are controlled respectively.

Parametric modeling concepts of CAD systems enable the step-by-step advancement of infrastructure models which progress from a rough LoD to sequentially higher LoDs. The very important advantage of the proposed approach in particular in the highly dynamic planning procedures in a real-time real-world mechanism, is that data models from different data domains don't need all to be in the same level of preciseness. In above proposed and discussed approach, planning team can flexibly define the LoDs based on the infrastructure project requirements. Of course for the state of the information now some of the queries get limited to early stages of design and coarser LoDs. For finer LoDs as Borrmann et al. (2014a) believe a bi-directional modeling of dependencies is required and will be the subject of future researches.

Therefore, figure 58 demonstrates how the alignment can be further transferred from the LoD 0 which is the 2D axis into the LoD 1 which is a 3D axis or even LoD 2 which includes the definition of part-alignments and structures after spatial link procedure based on rule definitions.

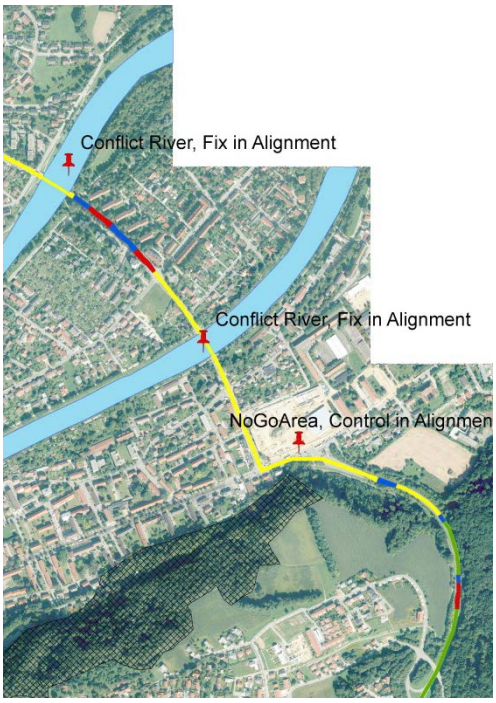


Figure 58 Application of conflict detection according to criteria defined in the rule engine

In this content as the figure 59 shows in a three dimensional viewer the alignment will be divided into a three dimensional model of proposed sections of tunnels (in green color), bridges (in brown color), embankments (in blue color) and cuts (in red color) and earth (in yellow color). Figure 59 shows the proposed tunnel under the 3D terrain based aerial photo.

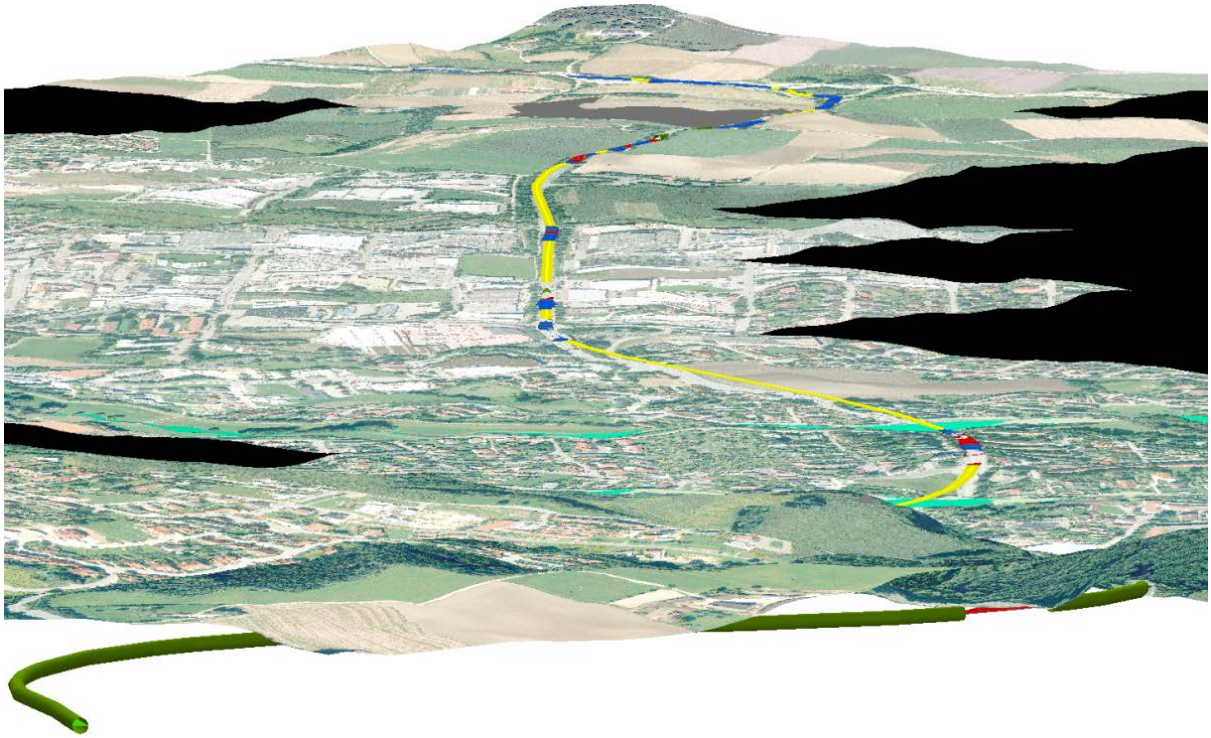


Figure 59 Result of spatial interlinking approach for the Tunnels in LoD 2 in a 3d viewer

During the modeling process, the changes between LoD of one data model and another are explicitly decided by the designing engineer who in this way decides which data model in which preciseness should be fed into the task for the best results.

The proposed multi-scale approach enables different engineers to adjust and share different data models involved in infrastructure task on various scales parallel while using the different design tools.

This is highly vital because a significant identity of infrastructure tasks and processes is the varying scales– ranging from the kilometer scale for the general alignment definition task down to the centimeter scale for the detailed planning of individual track details.

Figure 60 shows how the result of interlinking of different objects from different domains in different LODs and obtaining virtual properties for the object models through relationship container.

Other application areas are for example environmental simulations like noise emission mapping which needs information about noise absorption of surfaces. (Trees located in the design site), protection areas (flood area or animal protection area), and etc. Another application is cultural heritage, where the data model needs augment of objects by their heritage and history, and has to consider the development along time. Utility networks as another area of application need to represent pipes, pipe tunnels, connectors, transforming devices. Climate data can also use the approach for defining for example the tornado risks.

City and urban planning, real estate management, environmental simulation, crisis and disaster management, telecommunication, facility management hazard management and etc. are other application fields.

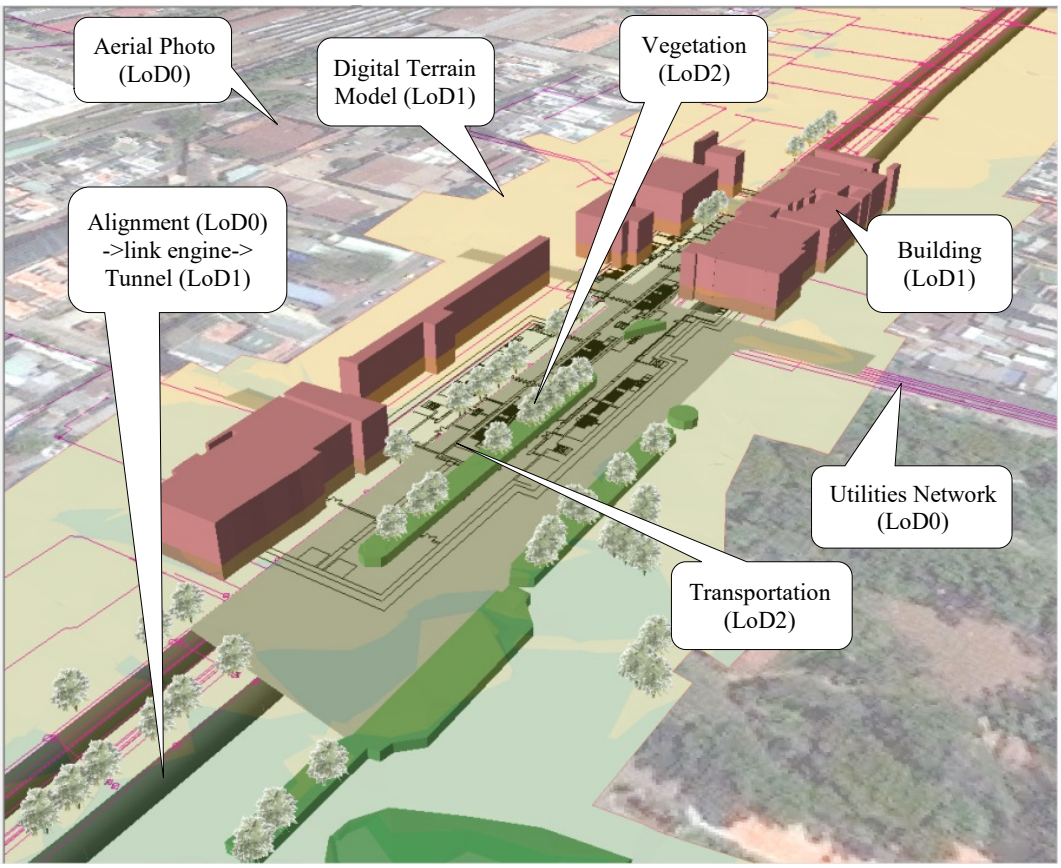


Figure 60 Example for an application MM, merging domain models with different LoD for a specific view of a certain design task (Przybylo et al., 2015)

Chapter 6

Conclusion

This thesis has introduced an approach for reducing the whole project lifecycle costs, increasing reliability of the comprehensive fundamental information, and consequently in independent, cost-effective, aesthetically pleasing, and environmentally sensitive infrastructure design.

Special attention has been paid to the integration of heterogeneous data and regulations involved in the design phase of infrastructure projects. In this phase the current status and level of interoperability among infrastructure data model and Geospatial Information Systems has been discussed in detail.

The dissertation has focused on the available standards in regard of infrastructure design and has discussed the advantages and disadvantages of each of them in detail.

Concept of infrastructure information modelling investigated in this work is a method for loose and rule based coupling of exchangeable heterogeneous information spaces. The MM extension introduced in this work and implemented in the framework of this thesis is a rule-based Multimodel which enables linkage of steady changing data resources via semantic rules. Huge amount of inter model relations will be avoided using hierarchically constructed linking regulations as well as using filtering and abstraction approaches.

Focusing on data structure instead of data model, this thesis has been able to solve the problem of dealing with data models having no IDs has been solved using the eMM implemented in this thesis.

6.1 SUMMERY

One of the most important aspects which has been the road map of this thesis has been the integration of GIS, BIM and Infrastructure data models in the case of infrastructure processes is an effective context of optimization of collaboration across the involved team members, organizations, decision makers, etc. throughout project lifecycle. This aims at minimization of the objective function of different project costs resulted after an assessment of different data domains integrated in the project.

As both BIM and GIS approaches enable a fully integrated project process, the first step which has been done in this thesis was to detect the overlapping and shortcoming of each of them in the case of infrastructure tasks. Because regarding infrastructure tasks and processes there is no defined internationally accepted product or process modelling and therefore infrastructure tasks can partially be covered by BIM approaches and partially fall in GIS approaches. Thus infrastructure projects are exactly where GIS and BIM meet each other.

For integration of infrastructure data models with BIM or GIS data models, in the next step different approaches were investigated in this work. In infrastructure information processes, there is a need for semantic linking of different data models, because it is not known which domain models might be integrated in future tasks. Therefore a method was needed which allow definition of semantic links or an adequate rule based query (Esfahani, 2013, p 17).

The multimodel (MM) method as a general approach which takes the heterogeneous data from their original format, interlinks them together and filters them in accordance to task based queries was investigated and examined for infrastructure data models. Even though in an MM heterogeneous data models from various data domains can get integrated through a unique key ID, yet dynamic data models are where the existing MM gets to the end of its capacity. In this thesis diversity of geospatial data models has been appointed to be their dynamicity.

Besides, in existing MM fundamental integration key is the unique IDs of the entities. As discussed in this work some of geospatial and infrastructure data models have spatial IDs and some other data models there are no IDs. Therefore this dissertation has proposed a next generation for MM where the linking approach does not necessarily require unique IDs but rather general rules. In the proposed approach of this thesis linking rules replace links. The linking rules are semantic matching approaches which are capable of creating links automatically intelligently.

The proposed extended multimodel (eMM) of this thesis uses semantics and consequently the matching approaches to create rule based links for diverse data models independent from their IDs using their original format so that the changes in the data are manageable. I.e. there will be no need to update the input data model in eMM and the links. This is therefore the solution for aforementioned problems.

For the purpose of creating rule based links this dissertation has targeted the spatial domain and has used the spatial identity of the data models as the ruling approach for match finding purpose. The focus in this research has been to support the infrastructure project processes from a spatial planning perspective. Subsequently, a BIM based solution has been suggested. According to Niemiec (2008) originated in the 1980s, BIM has following strengths:

- Precision in planning
- Object oriented
- In most of applications drawing based
- Managing large product model datasets

In compare, GIS with its origin in 1990s benefits from following advantages:

- Strong modelling and visualization machine
- Managing large spatial datasets
- Database oriented

Therefore as discussed in this thesis, geospatial and infrastructure data models are common in dealing with spatial data. Spatial data is the data with spatial properties. In relational databases, spatial data cause spatial relationships. Spatial relationship means to handle spatial linking and spatial querying.

Such querying will answer typical questions which are important for the infrastructure design tasks and is well defined in SQL 2008 or SQL 3.0. Thus the spatial matching proposed in this thesis illustrated how semantic links will be created using spatial indexes. Links can be saved explicitly for documenting a specific status of the project for future reviews.

"Use of topological queries can be very powerful and is supported in BIM. A suggestion in that regard is done in (Borrmann & Rank, 2009). However, albeit foreseen in IFC, topological data are hardly supported by current BIM-CAD. Borrmann and Rank (2008, 2009) show how this can be overcome by additional reasoning but the suggested algorithms based on Octree geometry are complex and hence of limited applicability for non CAD tools. Thus, whilst definitely useful, topological queries" (Katranuschkov et al., 2010) and therefore spatial linking has been excluded in the scope of MM.

To solve this complexity problems mentioned above, for the case that data models do not have spatial IDs, this thesis has used R-tree data structure for spatial data models which can be extended or shortened dynamically. The bounding boxes which are common search boundaries in geospatial applications have been successfully used for infrastructure data models in this work. This means structuring of data happens according to the requirements which are determined in the rule engine for a specific task.

A project can be modelled as in an indoor aspect and as in an outdoor aspect throughout its lifecycle. Most of the CAD as well as BIM programs especially IFC focus mainly on design of the project interiorly, whereas OGC models the digital world as in an outdoor point of view. OGC considers 5 levels of detail for modelling the objects, in which the last LoD namely LoD 4, for having building and street furniture acts as an interface to the interior design. That is why OGC might not cover the desired LoDs required for the entire IFC / BIM model. Thus, an enhancement has been suggested in the form of extending the current LoD template of OGC so that there is one general LoD concept which can be used for the purpose of interoperability.

As discussed in previous chapters the suggested LoD catalog of this thesis is one of the fundamental inputs of rule engine. This catalog contains constraints and relationship definitions for match finding approach. LoD catalog and its definitions has a strong impact on the search cube in the query procedure. The approximation of searching and allowed deviation from expected results are all depending on the LoD catalog. The more exact definition of LoDs and related objects belonging to a specific LoD from a specific related domain, the higher will be the accuracy of the results. Therefore the most critical part is to determine which domain with which LoDs are to be considered for tasks and processes in the queries.

For the alignment class diagram of LandXML, as an example, this thesis has implemented a definition schema for various class objects and their levels of detail, in which they can be modelled. This thesis has used this introduced LoD concept for the hierarchical filtering for the abstraction purpose of spatial matching results based on the interlinking task.

Furthermore a pretty new issue which raised in this thesis is the idea of relationships between team members and data models. This idea deals with the question how team members from various domains such as building, infrastructure or geospatial can be influenced by or impact on each other. To deal with this topic this thesis has suggested an LoD concept for the organizational hierarchies for the matrix of responsibilities and communication purposes. A matrix of the planning team and their roles and subsequently responsibilities has been suggested and an LoD concept has been implemented to show which relationships between organization and product models exist.

Another part of the rule engine proposed in this thesis consists of conflict detection approach part which is responsible for the self-fine tuning and self-control. Conflict detection determinations has been developed in this thesis according to the definitions of attributes and inserted into the rule engine.

This work has used Nearest Neighbor (NN) algorithm for spatial matching. As mentioned above to solve the problem of complexity mentioned by Borrmann and Rank (2009), this work suggests an enhancement in search mechanism through R-tree indexing the spatial data in spatial databases. Thus R-trees are chosen, as they consider NN matching in their structure and therefore facilitate a search mechanism which runs fast, accurate and at the same time dynamic. R-tree manages spatial structure for huge datasets the best, and yet is scalable. This property indicates that even if the dataset gets

larger, still the R-tree will manage the new size of the data, without being enforced to structure the data again. R-tree deals with new geometries and not just points, which makes it multidimensional and interpretable. Besides NN depends critically on the distance, thus the input data types should be normalized or have the same values range.

To avoid unrelated data different abstraction methods have been investigated in this dissertation for the infrastructure and geospatial domain. The filtering concept implemented in this thesis is based on LoD concept. This concept refers to the organization to be the leading model for the definition of minimum LoD required for each product model, GIS, BIM and Infrastructure. Exact specification of LoDs is done instantly for some products of the product model as an implementation part of this work.

6.2 DISCUSSION OF RESULTS

As mentioned in previous chapters high level infrastructure suits mostly facilitate to analyze the infrastructure design based on the international or user defined standards. Called regulation-based design, this minimizes errors, reduces costly design conflicts, increases time savings and provides consistent project quality. Yet design packages either don't consider GIS domains such as energy and environmental impacts or consider their own data domains like materials and land which might not meet the requirement of the other project members (Esfahani, 2013).

Besides infrastructure projects demand a lot of decision makings in governmental as well as public private partnership (PPP) level considering different data models. Therefore lossless flow of the project data as well as regulations across project teams, stakeholders, and governmental and PPP is highly important (Esfahani, 2013).

Multimodel (MM) is a method where heterogeneous data models from various domains are bundled together into a container keeping their original format. In separate Link Models (LM) the elements of the data models will be linked together (Fuchs, et al., 2011). Links contain a set of identifiers (IDs) representing specific entities of a certain data model. Thus, MM and the generated links have no inherent domain semantic. It is quite appropriate, if there are many different stakeholders, labors, types of tasks, questions and queries involved with the heterogeneous data. The data is re-usable in its generic software.

The multimodel method aims at cross-domain information interchange while reusing existing data formats. Therefore links contain a set of identifiers (IDs) which represent concrete entities of a certain data model. Thus, the multi-model and the links themselves have no inherent domain semantic. Semantic is only defined implicitly by applying a multimodel in a certain context, e.g. a task, information process or design procedure.

The method used in this thesis is independent from IDs. It creates an R-tree for the data models involved in the process which is proper for spatial query procedure using NN approach. The advantage of this method is that the query planner does not need to know the IDs and the attributes to build a query. Using the method in this work the best option is evaluated even if instead of the entire table a small part of the table is read. Statistics about data distribution which is recorded in indexed columns of the search table shows that the querying runtime is accelerated.

The method investigated in this thesis has been applied on geospatial and infrastructure domains to establish rule based relationships between them. The linking approach in spatial domain using rule based links will cover even data domains which are under permanent change in content for example climate domain. Rule based linking such as semantic linking between infrastructure data models and GIS, gives computer systems the possibility to transmit design and geospatial data with unambiguous, shared meaning.

As there is always a match tolerance associated with the matching mechanism, the results of the spatial matching might contain multiple tuples. Thus filtering approaches investigated in this work has enhanced the results of the spatial matching. Such optimization helps the system avoid seduces or irrelevant information. Thus retrieved from CityGML this thesis has proposed a concept of LoD and has specified the relationship of different LoDs between GIS, BIM and Infrastructure data models.

This has demonstrated how to rationally refine the results of spatial matching which helps the system avoid irrelevant information and seduces

Applying the developed LoD concept and generating hierarchical LoDs devote in better visualization of objects as well as effective memory management and improving the render speed while loading data from the data streaming point of view. Yet it results into simplification and generalization of objects, which in the terms of modelling cause inaccuracy or fuzziness but not inconsistency.

Overall, rule based linking in spatial domain is an aid for communication between disciplines and organization through unified automatic rule-based intelligent linking method which is created automatically and independent from data models. This becomes even more important in infrastructure projects with very long life time. This facilitates a high degree of coordination, transparency, and cost enhancement while maintaining a high level of optimization in the interdisciplinary cooperation between various disciplines. Arising benefits for the involved parties as well as the project process are added value principally through holistic evaluation of projects across all disciplines and throughout the entire life cycle of a project.

This helps project team members, authorities, stakeholders, and decision makers benefit from a homogenized data and to access the same data pool and decision base. Besides, this minimizes errors and costly conflicts, improves team collaboration and management, wins work and simplifies approval.

6.3 FUTURE WORK

The multimodel method is an approach to interlink heterogeneous data models with each other. Yet to establish the semantic-free links in a way which makes sense, one should manage the linkage procedure. Although this work has introduced an automatic approach of semantic free interlinking which interoperates the data models via dynamic linking and dynamic indexing; yet, the main focus of the work has been the planning phase, and the alignment data model.

Yet, a number of open problems must be solved to allow the development of a fully integrated system for infrastructure projects and geospatial domains. These problems suggest a variety of research directions that need to be pursued to make such a system feasible.

One such direction would be to investigate expanding the proposed rule engine for covering the entire infrastructure project life cycle considering further project phases such as construction, maintenance etc. This idea was pointed out in detail in chapters 3 and 4 and led to the development of IIM. Expectation propagation is to investigate the rule mechanism in the rule engine and complete different parts of the rule engine such as Rules, Sequences, Plausibility Criteria and Vocabulary Definitions as demonstrated in the implementation chapter, system architecture to cover an infrastructure project entirely.

The focus of this work has been to create a high-level spatial link model for infrastructure process and to provide strategies to extend it to further relative domains with spatial identities like the geomechanical and energy process. Another future work is to expand the extended multimodel of this thesis for further rule based linking mechanism. An example of this is the time based (temporal) domain with temporal identifiers. Climate and energy data models are proper data models of time based (temporal) domain. Temporal (time dependent) matching approaches have been introduced in SQL 3.0. Based on this interlinking of temporal data models using their identifiers and tree based data structures can be one of the research areas in the future.

4D-information systems are also another research field for the future. Developing interoperability approaches for the construction processes on site such as digital tools for management of “virtual construction site” can be another future field of investigation.

Ontology based linking approaches are also other research topics which are currently running in the Institute of Construction Informatics in the TU Dresden for the building and energy domain.

Chapter 7

Bibliography

‘To see a world in a grain of sand
And a heaven in a wild flower,
Hold infinity in the palm of your hand
And eternity in an hour.’

William Blake

7.1 BOOKS AND PAPERS

A

Abukhater A. (2013): *GIS Meets BIM: Delivering End-to-End Solutions for Government and Infrastructure*. Pitney Bowes Software Solutions, article date May 31, 2013 in press: http://blogs.pb.com/pbsoftware/2013/05/31/gis-meets-bim-delivering-end-to-end-solutions-for-government-and-infrastructure/

Altman N. S. (1992): *An introduction to kernel and nearest-neighbor nonparametric regression*. The American Statistician 46 (3), p 175–185.

Amann J., Borrmann A., Hegemann F., Jubierre R.J., Flurl M., Koch C. And König M. (2013): *A Refined Product Model for Shield Tunnels based on a Generalized Approach for Alignment Representation*. In Proc. Int. Conf. On Civil And Building Engineering Informatics (ICCBEI), ISBN 978-4-9907371-0-8, Tokyo, Japan, November 2013.

Amann J., Flurl M., Jubierre R.J. and Borrmann A. (2014): *An alignment meta-model for the comparison of alignment product models*. ECPPM 2014, Vienna, Austria.

Amirebrahimi S., Rajabifard A., Mendis P., and Ngo T., (2015): *A Data Model for Integrating GIS and BIM for Assessment and 3D Visualisation of Flood Damage to Building*. Locate Journal, Volume 2015, p 78-89.

Arctur D. (2006): *GIS Standards and Interoperability: Understanding and Using GML*. Esri Federal User Conference (FedUC), 31 January–2 February 2006, Washington D.C., USA.

Axmann Geoinformation GmbH: *ArcGIS Data Interoperability Extension*. Gaenserndorf, Austria.

B

Borrmann A. (2010): *From GIS to BIM and back again – A Spatial Query Language for 3D building models and 3D city models*. 5th International 3D GeoInfo Conference, November 3–4, 2010, Berlin, Germany.

Borrmann, A., Flurl, M., Jubierre, J.R., Mundani, R.-P., Rank, E. (2014a): *Synchronous collaborative tunnel design based on consistency-preserving multi-scale models*. Advanced Engineering Informatics.

Borrmann A., Kolbe T. H., Donaubauer A., Steuer H. and Jubierre J. R. (2013): *Transfering Multi-Scale Approaches from 3D City Modeling to IFC-Based Tunnel Modeling*. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-2/W1, ISPRS 8th 3DGeoInfo Conference & Commission II, WG II/2 Workshop, 27 – 29 November 2013, Istanbul, Turkey, p 75-85.

Borrmann A., Kolbe T. H., Donaubauer A., Steuer H., Jubierre J. R. and Flurl M. (2014b): *Multi-scale geometric-semantic modeling of shield tunnels for GIS and BIM applications*. Computer- Aided Civil and Infrastructure Engineering, Wiley, DOI: 10.1111/mice.12090.

Borrmann A. and Rank E. (2009): *Specification and implementation of directional operators in a 3D spatial query language for building information models*. Advanced Engineering Informatics Journal, Volume. 23, Number 1, p 32–44.

Borrmann A. and Rank E. (2008): *Topological Operators in a 3D Spatial Query Language for Building Information Models*. 12th International Conference on Computing in Civil and Building Engineering (ICCCBE-XII) 2008, Beijing, China.

Borrmann A. and Rank E. (2009): *Topological analysis of 3D building models using a spatial query language*. Advanced Engineering Informatics, Volume 23, Issue 4, October 2009, p 370-385.

Borrmann A., Schraufstetter S., van Treeck C. and Rank E. (2007): *An Octree-Based Implementation of Directional Operators in a 3D Spatial Query Language for Building Information Models.* Proceedings of the 24th Conference on Information Technology in Construction p 105 – 110.

Breunig M., Borrmann A., Rank E., Schilcher M., Hinz S., Donaubauer A. and Mundani R.-P. (2012): *Towards 3D Geoinformatics and Computational Civil Engineering Support for Cooperative Tracks Planning*. Knowing to manage the territory, protect the environment. FIG Working Week 2012, 6–10 May 2012, Rome, Italy.

Breunig M., Rank E., Schilcher M., Borrmann A., Hinz S., Mundani R.-P., Ji Y., Menninghaus M., Donaubauer A., Steuer H. and Vögtle T. (2011): *Towards computer-aided collaborative subway track planning in multi-scale 3D city and building models*. In Proceeding. of the 6th International ISPRS Conference on 3D Geoinformation, 3D GeoInfo, 17, Wuhan, China, 2011.

Briggs R. (2015): *GIS Data Structures, From the 2-D Map to 1-D Computer Files*. POEC 5319, Introduction to GIS, UT Dallas, 27 March 2015.

Burdescu D. D., Ruxanda M. M., Stanescu L., Badica C., Brazovan M. and Mocanu M. (2005): *New Algorithms for Content Based Retrieval in Multimedia Databases (The study of a content based retrieval system for similar images, using R and M trees for multi-dimensional data indexing)*. Contract GAR 153/2005, University of Craiova, Faculty of Automatics, Computers and Electronics.

Burkholder E. F. (2004): *Fundamentals of Spatial Data Accuracy and The Global Spatial Data Model (GSDM)*. Global COGO Inc. - Las Cruces, NM 88003, April 2004. Article in press.

Burrough P. A. and Frank A. U. (1996): *Geographic Objects with Indeterminate Boundaries, Volume 2 of GISDATA Series*. Taylor & Francis, March 1996. ISBN: 9780748403875.

C

Child C. (2009): *The Top Nine Reasons to Use a File Geodatabase.* ArcUser, 2009, pp. 12 – 15

Chunhui Z. (2009): *Models and Optimization Methods of Spatial Query on Multi-platforms*. First International Workshop on Database Technology and Applications, 04/2009

Cova T. J., Miller H. J., Beard K., Frank A. U. and Goodchild M. F. (2008): Geographic Information Science. 5th International Conference, GIScience 2008, Park City, UT, USA, September 23-26, 2008.

D

Daum S. and Borrmann A. (2014): *Processing of Topological BIM Queries using Boundary Representation Based Methods*. Advanced Engineering Informatics, http://dx.doi.org/10.1016/j.aei.2014.06.001

Deprizon A., Atiq R. and Karim O. A. (2008): *Expert System for Geometric Highway Design Using GIS*. EASTS International Symposium on Sustainable Transportation incorporating Malaysian Universities Transport Research Forum Conference 2008 (MUTRFC08). University Technology Malaysia, 12–13 August 2008.

Dietterich T. G. (2014): *The Nearest Neighbor Algorithm*. Lecture notes of class 534, part 3, Institute for Collaborative Robotics and Intelligence Systems (CoRIS), School of Electrical Engineering and Computer Science, 1148 Kelley Engineering Center, Oregon State University

Döllner J. and Hagedorn B. (2007): *Integration von GIS-, CAD- und BIM-Daten mit dienstbasierten virtuellen 3D-Stadtmodellen*. GIS – Zeitschrift für Geoinformatik, 11/2007.

Döllner J., Baumann K. and Buchholz H. (2006): *Virtual 3D City Models as Foundation of Complex Urban Information Spaces*. CORP 2006 & Geomultimedia06, Vienna, Austria, 13–16 February 2006.

Döllner, J. and Buchholz, H. (2005): *Continuous level-of-detail modeling of buildings in 3D city models*. Proc. of the 13th annual ACM Int. Workshop on Geographic information systems.

E

Eagleson S., Escobar F. and Williamson I. (2002): *Hierarchical spatial reasoning theory and GIS technology applied to the automated delineation of administrative boundaries*. Computers, Environment and Urban Systems, 26 (2002) p 185–200.

El-Mekawy M. and Östman A. (2010): *Semantic Mapping: an Ontology Engineering Method for Integrating Building Models in IFC and CITYGML*. In Proceeding of the 3rd ISDE Digital Earth Summit, 12–14 June 2010, Nessebar, Bulgaria.

El-Mekawy M. (2010): *Integrating BIM and GIS for 3D city modelling; The case of IFC and CityGML*. Licentiate Thesis at Geoinformatics Division, Department of Urban Planning and Environment, Royal Institute of Technology (KTH), November 2010, Stockholm, Sweden.

Esfahani N.N. (2013a): *Interoperability methods for Infrastructure Design and Information Systems*. Forum Bauinformatik 2013, p 161–171, Munich, Germany

Esfahani N.N. (2013b): *Interoperability of Infrastructure Planning and Geo-Information Systems*. In: Scherer R. J. Research and Lecture activities in 2014 Technische Universität Dresden, Institut für Bauinformatik, Dresden, Germany, December 2013

Esfahani N.N., Franken W. and Uhrig R. (2011): *Space Technology and GIS in ProVI based Alignment Planning*. The 4th Geodesy Architecture & Construction Conference (GAC 2011), Lviv, Ukraine.

Esfahani N.N., Fuchs S. and Scherer R.-J. (2014): *Infrastructure and geospatial data models; spatial links*. eWork and eBusiness in Architecture Engineering and Construction, Volume 10, p 335–341.

Esfahani N.N., Scherer R.-J. and Balder R. (2012): *Traffic Infrastructure Design and Geo-Information Systems: a case of Interoperability*. eWork and eBisoness in Architecture, Engineering and Construction, Volume 9, p 597–602.

F

Farkas A. (2009): *An Intelligent GIS-Based Route/Site Selection Plan of a Metro-Rail Network*. In I.J. Rudas et al. (Eds.) (2009): *Towards Intelligent Engineering & Information Technology Studies in Computational Intelligent*, Volume 243/2009, p 719–734, Springer, Berlin- Heidelberg, Germany.

Federal Highway Administration (2012): *The Role of FHWA Programs in Livability: State of the Practice Summary*. Article in press on: https://pediaview.com/openpedia/Geometric_design_of_roads

Federal Ministry of Transport, Building and Urban Development (1979): *Regelungen für die elektronische Bauabrechnung Verfahrensbeschreibung 21.003 Massenberechnung aus Querprofilen*. Elling, 1979.

Flurl M. (2012): *Kollaborative Zusammenarbeit in Multi-Skalen-Geometriemodellen: Ein erster Entwurf eines Kollaborationsservers*. Forum Bauinformatik 2012, p 87–94, article in German, Bochum, Germany.

Frank A. U. and Wallace M. (1995): *Construction Based Modeling in a GIS: Road Design as a Case Study*. International Symposium on Computer-Assisted Cartography, Auto-Carto XII.

Fuchs S. (2012): *Eine Abfragesprache für Multimodelle*. Forum Bauinformatik 2012, p 95–102, article in German, Bochum, Germany.

Fuchs S. (2013): *Rule based creation of Multi-Model Links*. In: Scherer R. J. Research and Lecture activities in 2013 Technische Universität Dresden, Institut für Bauinformatik, Dresden, Germany, December 2012.

H

Hyvärinen J. (2012a): *OpenINFRA: buildingSMART MVD for LandXML v1.2*. BuildingSMART ITM / IUG, 17 October 2012, Tokyo, Japan.

Hyvärinen J. (2012b): *OpenINFRA LandXML Proposal*. BuildingSMART ITM / IUG, 18–22 March 2012, Oslo, Norway.

I

IEEE (1990): *IEEE standard computer dictionary: a compilation of IEEE standard computer glossaries*. New York: Institute of Electrical and Electronics Engineers, 1990.

Isikdag, S. and Zlatanova S. (2009): *Towards defining a framework for automatic generation of buildings in CityGML using BIM*. In Lee and Zlatanova (Eds.), 3D geo-information sciences, Springer LNG&C, pp. 79-96.

J

Jha M. K. (2000): *A Geographic Information Systems-Based Model for Highway Design Optimization*. PhD Thesis, University of Maryland.

Jha M. K., Davis C. and Kang M. –W. (2007): *State-of-the-art intelligent road design model with generic algorithms, Geographic Information Systems, and CADD.* Advanced in Transportation Studies an international Journal, Section A&B 13, p 41.

Jha M. K., Schonfeld P., Jong J. –C. and Kim E. (2006): *Intelligent Road Design, Series: Advances in Transport*, Vol. 19. Southampton, WIT Press.

Ji Y., Borrmann A., Rank E., Wimmer J. and Günthner W. A. (2009): *An Integrated 3D Simulation Framework for Earthwork Process*. 26th CIB-W78 Conference on Managing IT in Construction, 1–3 October 2009, Istanbul, Turkey.

Ji Y., Lukas K., Obergriesser M. and Borrmann A. (2008): *Entwicklung integrierter 3D-Trassenproduktmodelle für die Bauablaufsimulation*. Tagungsband des 20 Forum Bauinformatik, Dresden, Germany, September 2008, article in German, article in press.

Johansson M. and Roupé M. (2010): *How can GIS and BIM be integrated?* New Frontiers CAADRIA Conference, 7–10 April 2010, Hong Kong, poster in press.

K

Kadolsky M., Baumgärtel K. and Scherer R.J. (2014): An ontology framework for rule-based inspection of eeBIM-systems. Creative Construction Conference CCC2014, 21–24 June 2014, Prague, Czech Republic.

Karimi H. A. and Akinci B. (2010): *CAD and GIS Integration*. Edited by Karimi, H. A. and Akinci, B. 1st ed: Taylor and Francis Group (CRC Press).

Katranuschkov P., Baumgärtel K., Scherer R., vanWoudenberg W., and Geißler M. (2014): *Using IFC for energy-extended building data modelling – challenges, achievements, first experiences and evaluation*. eWork and eBusiness in Architecture Engineering and Construction, 2014, p 791–799.

Katranuschkov P., Weise M., Windisch R., Fuchs S. and Scherer R. J. (2010): *BIM Based Generation of Multi-Model Views*. Proceedings of the CIB W78 2010: 27th International Conference – Cairo, Egypt, 16-18 November

Kayondo M. L., Bax G. and Togboa T. (2011): *Geographic Information Systems for Transportation - Data Model for Road Infrastructure Maintenance in Uganda*. 1st Conference on Advances in Geomatics Research, 3–4 August 2011, Kampala, Uganda.

Kim E., Jha M. K. and Son B. (2004): *Improving the computational Efficiency of highway alignment optimization models through a stepwise generic algorithms approach*. ELSEVIER, Journal of Transportation Research, Part B 39, p 339–360, article in press.

Kolbe T. H. (2011): *Current Status and Roadmap of CityGML*. Joint ISPRS 3DCMA & 3D GeoInfo Conference, 26–28 June 2011, Wuhan, China.

Kolbe T. H. (2009): *Representing and Exchanging 3D City Models with CityGML*. In Lee J. and Zlatanova S. (Eds.): *3D Geo-Information Sciences*. XVII, p 15-31, Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-87394-5.

Kolbe T. H. and Gröger G. (2005): *CityGML –Status and Roadmap*. Institute for Cartography and Geoinformation, University of Bonn, November 2005, Bonn, Germany.

Kolbe T. H., Gröger G. and Plümer L. (2005): *CityGML – Interoperable Access to 3D City Models*. Proceedings of the Int. Symposium on Geo-information for Disaster Management, 21– 23 March 2005, Delft, Netherlands, Springer.

Kymmell W. (2008): *Building Information Modeling: Planning and managing construction projects with 4D CAD and simulations*. USA, McGraw-Hill.

L

Laallam F. Z., Kherfi M. L. and Benslimane S. M. (2014): *A survey on the complementarity between database and ontologies: principles and research areas.* International Journal of Computer Applications in Technology (IJCAT), Vol. 49, No. 2, 2014.

Laurini R., Yétongnon K., and Benslimane D. (2008): *GIS Interoperability, from Problems to Solutions*. Advanced Geographic Information Systems, Vol. II, Encyclopedia of Life Support Systems EOLSS, ISBN: 978-1-84826-992-7.

Lebégue E. (2005): *IFC: The beginning of the road*. IFC Bridge & Road Workshop Report, CSTB, 22–23 September 2005, Sophia-Antipolis, France.

Liebich T., Katranuschkov P., Weise M., Guruz R. and Scherer R.-J. (2013): *BIM to Energy: Extending BIM for Multi-Model Domain Tasks*. 4th international conference of ICT for sustainable places, 9–11 September 2013, Nice, France.

M

Madrazo L., Nemirovski G., Sicilia Á. (2013): *Shared Vocabularies to Support the Creation of Energy Urban Systems Models*. Proceedings of the 4th Workshop organised by the EEB Data Models Community ICT for Sustainable Places, Nice, France, 9-11 September, 2013. Retrieved July 31, 2013 in Web site: http://sustainable-places.eu/wpcontent/uploads/2014/02/4thW-eeb_dm_booklet.pdf

Marik V., Retschitzegger W. and Stepankova O. (2003): *Database and Expert System Applications*. In Processing of 14th International Conference DEXA 2003, September 2003, Prague, Czech Republic.

Murray C. (2010): *Oracle Spatial Developer's Guide*. 11g Release 2 (11.2), E11830-06, March 2010.

N

Nagel C. (2007): *Conversion of IFC to CityGML*. Meeting of the OGC 3DIM Working Group at OGC TC/PC Meeting, July 2007, Paris, France.

Nagel C., Stadler A. and Kolbe T. (2009): *Conceptual Requirements for the Automatic Reconstruction of Building Information Models from Uninterpreted 3D Models*. Academic Track of Geoweb 2009 Conference, Vancouver.

NCDCDS (1988): *A Proposed Standard for Digital Cartographic Data*. National Committee on Digital Cartographic Data Standards, The American Cartographer, 15(1):1-140.

Niemiec N., Southwood P. and Schlosser M. (2010): *A Cultural Shift in the GIS Industry*. CAD and GIS Integration Workshop at AU2010 Autodesk University, 30 November – 2 December 2010, Las Vegas, USA.

O

Obergriesser M. and Borrmann A. (2012): *Infrastructural BIM standards – Development of an Information Delivery Manual for the geotechnical infrastructural design and analysis process*. eWork and eBusiness in Architecture, Engineering and Construction – Gudnason & Scherer (Eds)© 2012 , p 581 – 587, Taylor & Francis Group, London, ISBN 978-0-415-62128-1.

Ohori A., Biljecki K. F., Stoter J. and Ledoux H. (2013): *Manipulating higher dimensional spatial information*. In: Proc. of the 16th AGILE Conference on Geographic Information Science, Leuven, Belgium

O´Neil D.J. (2010): *Nearest Neighbors Problem*. Department of Computer Science and Engineering, University of Minnesota, Minneapolis, United States. Oracle Spatial User's Guide and Reference (2008): Release 8.1.7.

P

Pattee H. H. (1973): *Hierarchy Theory – The Challenge of Complex Systems.* Braziller, New York, USA, p 156.

Pick J.B. (2007): *Geo-Business, GIS in the Digital Organization*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.

Ponniah P. (2007): *Data Modeling Fundamentals: A Practical Guide for IT Professionals*. ISBN: 978-0-470-14101-4, Jun 2007, 464 pages, published by John Wiley & Sons.

Przybylo J., Esfahani N. N. and Egger M. (2015): *BIM bei OBERMEYER Planen + Beraten.* In Borrmann, A., König, M., Koch, C. und Beetz, J. *Building Information Modeling, Technologische Grundlagen und industrielle Praxis*. Springer Vieweg, ISBN: 978-3-658-05605-6, p 445-461.

Pu S. (2004): *Database Storage of 3D Freeform Curves and Surfaces*. Research Project report, TU Delft, pp. 55, 2004. Article on Geo-Database Management Center (GDMC: http://www.gdmc.nl).

R

Rangnes J. (2012): *Challenges for interoperability in openINFRA*. Nordic OpenINFRA Workshop – LandXML, BuildingSMART Norway, 7 June 2012, Oslo, Norway.

Rebolj D. (1993): *Computerunterstützter integrierter Straßenentwurf in einer objektorientierten Umgebung*. dbv - Verlag für die Technische Universität Graz.

Rebolj D., Cus-Babic N., Tibaut A. and Magdic A. (2000): *Evolution of a Road Product Model*. International Council for Research and Innovation in Building and Construction, CIB w78 conference 2002, Aarhus School of Architecture, 12 – 14 June 2002

Romandi I. and Vallas I. (2013): *Implementing Model-Based Design in Infrastructure Projects*. In the XXVIII International Baltic Road Conference. Vilnius, Lithuania.

S

Scerbakov N. (2010): *Functional Data Model*. Institute of Interactive Systems and Data Sciences, 2010, Technical University of Graz, Austria

Schapke S.-E, Pflug C. and Bögl M. (2012): *Multi-models: New potentials for the combined use of planning and controlling information*. Mefisto-project, research.

Scherer R.-J. and Schapke S.-E. (2011): *A distributed multi-model-based Management Information System for simulation and decision-making on construction projects*. Advanced Engineering Informatics, Volume 25(4), p 582–599.

Scherer R.-J. and Kadolsky M. (2014): *Design4Energy, Models and Link-Models*. technical workshop, presentation notes, 17-18 Feb 2014, Salford.

Shekhar S. and Xiong H. (2008): *Encyclopedia of GIS*. Print ISBN 978-0-387-30858-6, Springer, Boston, MA, US 2008.

Sondheim M., Gardels K., and Buehler K. (1999): *GIS Interoperability*, in Longley P., Goodchild M., Maguire D., Rhind D., (Eds.) Geographical Information Systems 1 Principles and Technical Issues, John Wiley & Sons, New York, 1999, p347–358.

Stolze K. (2003): *SQL/MM Spatial: The Standard to Manage Spatial Data in Relational Database Systems*. 10th Conference on Database Systems for Business, Technology and Web. 26–28 February 2003. Leipzig, Germany.

T

The American Heritage Dictionary of the English Language, Fourth Edition (2000): Houghton Mifflin Company, Boston, MA.

Timpf S. (2008): *Hierarchies and Level of Detail*. In Shekhar S. and Xiong H. (Eds.) (2008): Encyclopedia of GIS, p. 431-434, Springer, Berlin, Germany.

Trepanier M. and Chapleau R. (2001): *Linking Transit Operational Data to Road Network with a Transportation Object-Oriented GIS*. URISA Journal, Volume 13, No. 2, p 23–30.

V

Van Berlo L. (2009): *CityGML extension for Building Information Modelling (BIM) and IFC*. Free and Open Source Software for Geospatial (FOSS4G), Sydney, Australia.

Van Berlo L. and De Laat R. (2010): *Integration of BIM and GIS: The development of the CityGML GeoBIM extension*. The 5th International 3D GeoInfo Conference, 3–4 November 2010, Berlin, Germany.

Vitkiene J., Puodziukas V., Zilioniene D. and Gintalas V. (2008): *Analysis of the Computer Aided Design Models for Roads*. The 7th International Environmental Engineering Conference, 22–23 May 2008, p 1235–1241, Vilnius, Lithuania.

Z

Zhao T., Zhang C., Wei M. and Peng ZR. (2008): *Ontology-Based Geospatial Data Query and Integration*. In: Cova T.J., Miller H.J., Beard K., Frank A.U., Goodchild M.F. (eds) *Geographic Information Science*. GIScience 2008. Lecture Notes in Computer Science, vol 5266. Springer, Berlin, Heidelberg

Zhao X., Chen J. and Zhao R. (1999): *Dynamic Spatial Indexing Model Based on Voronoi*. Proceeding of the International Symposium on Digital Earth – Towards Digital Earth. Science Press, Beijing, China.

7.2 WEBSITES

BISM. (2012): buildingSmart standards (28-03-2012)

http://buildingsmart.com/standards

BuildingSMART, International home of openBIM

http://www.buildingsmart-tech.org/

BuildingSMART Nordic, Report of LandXML for buildingSMART, 23 May 2012

http://www.buildingsmartnordic.org/resources/bSI%20LandXML%2020120520.pdf/at_download/file

CAD and GIS, Integration and Beyond. Available at:

http://cadngis.blogspot.com

Construction Sites Inc. Construction management

http://www.csites.com/

De Nier R. (2013), LandXML-OGC. Available at:

http://prezi.com/vn6n7o4-oqz8/landxml-ogc

ESRI Inc, ArcGIS.

http://www.esri.com

ESRI Development Network

http://edn.esri.com

GDAL, Geospatial Data Abstraction Library

http://www.gdal.org/

GEOS, Geometry Engine Open Source

http://geos.refractions.net

IFCExplorer CityGML Export

http://www.ifcwiki.org/index.php/IfcExplorer_CityGML_Export

International Cartographic Association

http://icaci.org/

ISO 2382-1, revision 2015, section 02, part 01, 01.01.47

https://www.iso.org/obp/ui/#iso:std:iso-iec:2382:-1:ed-3:v1:en

ISO/TC 211

http://www.isotc211.org/

LandXML

http://www.landxml.org/

Marine Metadata Interoperability (2013), Metadata Interoperability-What Is It, Why Is It Important? available at:

https://marinemetadata.org/guides/mdataintro/mdatainteroperability

NCGIA: National Center for Geographic Information Analysis, Unit 10: Spatial Databases as Models of Reality, available at:

http://ibis.geog.ubc.ca/courses/klink/gis.notes/ncgia/u10.html

Network Centric Operations Industry Consortium (2008)

https://www.ncoic.org/home

OGR, OGR Simple Feature Library

http://www.gdal.org/ogr/

OGC Technical Report

http://xml.coverpages.org/ni2004-05-12-a.html

Open Geospatial Organization

http://www.opengeospatial.org

Road Design Manual (2014)

http://onlinemanuals.txdot.gov/txdotmanuals/rdw/design_general.htm

Safe Software, FME Desktop Translator/Converter Software

http://www.safe.com/products/desktop

Studio Roosegaarde (2011)

http://www.studioroosegaarde.net/project/smart-highway/

Solibri Inc. (About BIM and IFC, 2014)

http://www.solibri.com/support/bim-ifc/

GIS Data Models, 2010

http://www.geog.ucsb.edu/~jeff/gis/data_models1.html

Workshop of Boundless Geo, Part 15, Spatial Indexing, 10 Jan 2013

http://workshops.boundlessgeo.com/postgis-intro/indexing.html

List of Abbreviation

ADT	ABSTRACT DATA TYPES
AEC	ARCHITECTURAL, ENGINEERING, CONSTRUCTION
AECOO	ARCHITECTURAL, ENGINEERING, CONSTRUCTION, BUILDING OWNERSHIP & OPERATION
API	APPLICATION PROGRAMMING INTERFACE
BIM	BUILDING INFORMATION MODEL
CAD	COMPUTER AIDED DESIGN
DBB	DESIGN, BID AND BUILDING
DEM	DIGITAL ELEVATION MODEL
FGDC	FEDERAL GEOGRAPHIC DATA COMMITTEE
GIS	GEOSPATIAL INFORMATION SYSTEM
GDM	GEOSPATIAL DATA MODEL
GML	GEOGRAPHY MARKUP LANGUAGE
IAI	INTERNATIONAL ALLIANCE FOR INTEROPERABILITY
IDM	INFRASTRUCTURE DATA MODEL
IFC	INDUSTRY FOUNDATION CLASSES
IID	INTELLIGENT INFRASTRUCTURE DESIGN
IS	INTEGRATED SYSTEMS
INSPIRE	INFRASTRUCTURE FOR SPATIAL INFORMATION IN THE EUROPEAN COMMUNITY
LBS	LOCATION BASED SERVICES
LM	LINK MODEL
LoD	LEVEL OF DETAILS
MBR	MINIMAL BOUNDING RECTANGLE
MM	MULTI MODEL
NSDI	NATIONAL SPATIAL DATA INFRASTRUCTURE
OGC	OPEN GEOSPATIAL CONSORTIUM
OKSTRA	OBJEKT KATALOG FÜR STRAßEN- UND VERKEHRSWESEN
ORDBMS	OBJECT RELATIONAL DATABASE MANAGEMENT SYSTEM
OS	ORACLE SPATIAL
OWL	WEB ONTOLOGY LANGUAGE
PDMS	PRODUCT DATA MANAGEMENT SYSTEM
PLM	PROJECT LIFECYCLE MANAGEMENT
POI	POINT OF INTEREST
PROVI	PROGRAMMSYSTEM FÜR VERKEHRS- UND INFRASTRUKTURPLANUNG
RDF	RESOURCE DISCRIPTION FRAMEWORK
SDI	SPATIAL DATA INFRASTRUCTURE
SF	SPATIAL FILTER
SL	SPATIAL LINK
SLE	SPATIAL LINK ENGINE
SLM	SPATIAL LINK MODEL
SMCE	SPATIAL MULTIPLE CRITERIA EVALUATION
SOA	SERVICE ORIENTED ARCHITECTURE
SQ	SPATIAL QUERY
SQL	STRUCTURED QUERY LANGUAGE
TF	TABULAR FILTER
WFS	WEB FEATURE SERVICES
WMP	WEB MAP SERVICES
XML	EXTENSIBLE MARKUP LANGUAGE

Published Dissertations, Habilitations and Booklets of the Institute of Construction Informatics[24]

MARKUS HAUSER	EINE KOGNITIVE ARCHITEKTUR FÜR DIE WISSENSBASIERTE UNTERSTÜTZUNG DER FRÜHEN PHASEN DES ENTWURFS VON TRAGWERKEN	LOGOS VERLAG BERLIN, 06/1998 DEUTSCH
MARTIN ZSOHAR	STOCHASTISCHE GRÖßEN DER RESONANZFREQUENZEN UND DER VERSTÄRKUNG SEISMISCHER WELLEN IM HORIZONTAL GESCHICHTETEN ZUFÄLLIGEM MEDIUM	SHAKER VERLAG AACHEN, 1998 DEUTSCH
CHRISTIAN STEURER	MODELLIERUNG UND BERECHNUNG DES ERMÜDUNGSFORTSCHRITTS MIT STOCHASTISCHEN DIFFERENTIALGLEICHUNGEN	SELBSTVERLAG, 1999 DEUTSCH
PETER KATRANUSCHKOV	A MAPPING LANGUAGE FOR CONCURRENT ENGINEERING PROCESSES	HEFT 1, 2001 ENGLISH
KARSTEN MENZEL	METHODIK ZUR NACHHALTIGEN, RECHNERGESTÜTZTEN RESSOURCENVERWALTUNG IM BAUWESEN	HEFT 2, 2003 DEUTSCH
MICHAEL EISFELD	ASSISTANCE IN CONCEPTUAL DESIGN OF CONCRETE STRUCTURES BY A DESCRIPTION LOGIC PLANNER	HEFT 3, 2004 ENGLISH
MATTHIAS WEISE	EIN ANSATZ ZUR ABBILDUNG VON ÄNDERUNGEN IN DER MODELLBASIERTEN OBJEKTPLANUNG	HEFT 4, 2006 DEUTSCH
JÖRG BRETSCHNEIDER	EIN WELLENBASIERTES STOCHASTISCHES MODELL ZUR VORHERSAGE DER ERDBEBENLAST	HEFT 5, 2006 DEUTSCH
MARTIN KELLER	INFORMATIONSTECHNISCH UNTERSTÜTZTE KOOPERATION BEI BAUPROJEKTEN	HEFT 6, 2007 DEUTSCH
KAMIL UMUT GÖKCE	IT SUPPORTED CONSTRUCTION PROJECT MANAGEMENT METHODOLOGY BASED ON PRODUCT MODEL AND QUALITY MANAGEMENT EINE METHODOLOGIE FÜR DAS IT-UNTERSTÜTZTE BAUPROJEKTMANAGEMENT AUF DER GRUNDLAGE VON PROZESS- UND PRODUKTMODELLEN UND QUALITÄTSMANAGEMENT	HEFT 7, 2008 ENGLISH
GERALD FASCHINGBAUER	SIMULATIONSBASIERTE SYSTEMIDENTIFIKATION IM RAHMEN DER BAUBEGLEITENDEN GEOTECHNISCHEN ÜBERWACHUNG SIMULATION-BASED SYSTEM IDENTIFICATION IN CONSTRUCTION-SIMULTANEOUS GEOTECHNICAL MONITORING	HEFT 8, 2011 DEUTSCH
WAEL SHARMAK	DYNAMIC NETWORK PLANNING IN CONSTRUCTION PROJECTS USING CONGURABLE REFERENCE PROCESS MODELS DYNAMISCHE NETZPLANUNG IN BAUPROJEKTEN UNTER VERWENDUNG VON KONFIGURIERBAREN REFERENZPROZESSMODELLEN	HEFT 9, 2011 ENGLISH

[24] July 2018 Faculty of Civil Engineering, Institute of Construction Informatics

AMIN ZAHEDI KHAMENEH	WAVE-TYPE BASED REAL-TIME PREDICTION OF STRONG GROUND MOTION ECHTZEIT-WELLENBASIERTE MODELLE FÜR DIE VORHERSAGE DER ERDBEBENBESCHLEUNIGUNG	HEFT 10, 2012 ENGLISH
SEBASTIAN FUCHS	ERSCHLIEßUNG DOMÄNENÜBERGREIFENDER INFORMATIONSRÄUME MIT MULTIMODELLEN ACCESS OF CROSS-DOMAIN INFORMATION SPACES USING MULTI-MODELS	HEFT 11, 2015 DEUTSCH
ALEXANDER BENEVOLENSKIY	ONTOLOGY-BASED MODELING AND CONFIGURATION OF CONSTRUCTION PROCESSES USING PROCESS PATTERNS ONTOLOGIE-BASIERTE MODELLIERUNG UND KONFIGURATION DER BAUPROZESSE MIT HILFE VON PROZESSVORLAGEN	HEFT 12, 2016 ENGLISH
FRANK HILBERT	KONTEXTADAPTIVE INFORMATIONSRÄUME– UNTERSTÜTZUNG INTERDISZIPLINÄRER BAUINFORMATIONSPROZESSE DURCH EINE KONTEXTBEWUSSTE INFORMATIONSLOGISTIK	HEFT 13, 2016 DEUTSCH
HELGA TAUSCHERT	CONFIGURABLE ND-VISUALIZATION FOR COMPLEX BUILDING INFORMATION MODELS KONFIGURIERBARE ND-VISUALISIERUNG FÜR KOMPLEXE BUILDING INFORMATION MODELS	HEFT 14, 2016 ENGLISH