

# The need for a differentiation between heterogeneous information integration approaches in the field of “BIM-GIS Integration”: a literature review

F. Beck<sup>1,\*</sup>, A. Borrmann<sup>1</sup>, T. H. Kolbe<sup>2</sup>

<sup>1</sup> TUM Department of Civil, Geo and Environmental Engineering, Technical University Munich, Germany - (fritz.beck, andre.borrmann@tum.de)

<sup>2</sup> TUM Department of Aerospace and Geodesy, Technical University Munich, Germany - thomas.kolbe@tum.de

**KEY WORDS:** BIM-GIS Integration, GeoBIM, Information Integration, Data Integration, BIM, GIS

## ABSTRACT:

The heterogeneous character of information models results in communication barriers between subsystems in railway organizations dealing with Building Information Modeling (BIM) and Geographic information systems (GIS). The integration of information is a promising way to bridge the heterogeneity of information models and satisfy the need for a more efficient communication. Integration efforts exploited in expert literature are often referenced using umbrella terms like “BIM-GIS Integration” or “GeoBIM”, although dealing with different challenges and addressing different purposes. This paper highlights the need for a differentiation between integration efforts covered by the umbrella term “BIM-GIS integration”. For this, a new approach for the categorization of information integration efforts was developed based on a literature research. Afterwards, challenges concerning information integration efforts in the field of “BIM-GIS Integration” were exploited and assigned to the respective categories to illustrate the importance of differentiation between heterogeneous information integration efforts.

## 1. INTRODUCTION

### 1.1 The need for efficient, lateral communication

Organizations are highly specialized due to the division of labor (Smith 2010) and the increasing complexity of tasks an individual has to carry out (Bar-Yam 2002). Due to the independent development of highly specialized subsystems of organizations and the related heterogeneity of information requirements, organizations deal with heterogeneous information models developed for different purposes. The heterogeneous character of these information models causes communication barriers. However, the increasing complexity of the demands of the environment requires an increasingly efficient communication across the specialized subsystems, also known as *lateral communication* (Bar-Yam 2002). In particular, the increasing complexity of the demands of the environment lead to more complex task specifications accompanied by a higher specialization, whereby the demand for intradisciplinary solutions remains present. Simultaneously, this need for more efficient, lateral communication runs counter to the communication barriers caused by the heterogeneous character of information models. Consequently, this kind of communication barriers need to be bridged to unfold synergetic effects between subsystems and satisfy the increasing demands of the environment of organizations.

Among other domains, the life cycle of railway infrastructure projects requires a high degree of specialization and encounters a multitude of heterogeneous information models (Wunsch et al. 2016). The increasing complexity of the demands of the environment with respect to the railway industry are driven by demographic, environmental, technological changes and the evolving needs and expectation of passengers and stakeholders (Steward et al. 2014; Powrie 2014). The resulting need for a more efficient, lateral communication might be supplied by the digitalization of the railway industry which provides the potential to break down communication barriers between specialized subsystems in railway infrastructure projects. A promising way to achieve this kind of communication is the *integration* of heterogeneous information models. For instance, information models describing asset

information used in the whole lifecycle of a physical asset in terms of Building Information Modeling (BIM) and information models describing geographic information provided in Geographic Information Systems (GIS) are major subjects of information integration efforts in the railway domain.

### 1.2 Research value

Integration efforts exploited in expert literature are often referenced using umbrella terms (e.g. “BIM-GIS Integration”, “GeoBIM”), although dealing with different challenges and addressing different purposes. The resulting vagueness leads to difficulties in communication and understanding in and across the respective domains. As categorizations play a major role for understanding, communication and logical inference (Waldmann 2008) this paper develops a categorization of integration efforts. Additionally, challenges regarding integration efforts identified in the literature research are assigned to the developed categories to show that heterogeneous information integration efforts deal with different challenges.

### 1.3 Research Method

This paper is part of a doctoral research investigation following the Design Science Research (DSR) method. The DSR method aims to create and evaluate Information Technology (IT) artifacts intended to solve the identified problems. According to Peffers et al. (2007) the DSR process generally includes six steps: 1) Problem identification 2) Definition of objectives 3) Design and development of artifacts (e.g. frameworks, methods, models) 4) Demonstration by using the artifact to solve the problem 5) Evaluation of the solution, comparing the objectives and the observed results 6) Communication of the problem, the artifact, its utility and effectiveness. This paper addresses solely the step *problem identification* and its *communication* to provide an understanding of integration efforts concerning the keyword “BIM-GIS Integration” as basis for the further research investigation.

The paper is structured into four major parts (Figure 1). First, a literature research concerning “BIM-GIS Integration” was conducted. Second, categories concerning information integration efforts were developed. Third, challenges of information integration efforts were identified. Last, the challenges were assigned to the respective categories. The literature research was mainly conducted using Scopus, a database of peer-reviewed literature. The scope of the paper is limited to integration efforts addressed by the term “BIM-GIS Integration”.

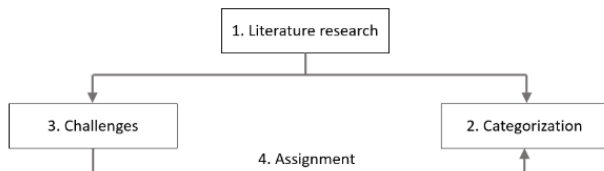


Figure 1: Structure of the research investigation

#### 1.4 Categorization Method

The developed categorizations are following three different approaches based on the concept of enumerative definitions. The application of enumerative definitions means that the information integration efforts are identified and assigned to categories in order to provide an abstract meaning of the respective category (Waldmann 2008). The concept of enumerative definition is applied due to the difficult creation of formal definitions. The first approach makes use of existing categorizations developed in similar or same context. The second approach adapts existing categorization developed in the same or similar context. The third approach creates new categories from scratch. The three methods use the concept of enumerative definition to assign meaning to the adapted or created categories. The creation and adaption of categories is an iterative process and the validity of the developed categorization is related to the context in it is created. To allow an easy understanding, the categories are described, whereas the descriptions are not seen as formal definition.

#### 1.5 Conceptualization

In this paper, the usage of the keyword “BIM-GIS Integration” is seen as problematic due the vagueness of the comparison *BIM-GIS* and the term *integration*. While BIM is in general associated with a *method*, the term GIS refers to *systems*. The comparisons CAD and GIS on system level, BIM and Urban Information Modeling (UIM) on method level, and e.g. Industry Foundation Classes (IFC) and CityGML on data level would be rather convenient (Hijazi and Donaubaer 2017). Similarly, Herle et al. (2020) compare BIM and Geospatial Information Modelling (GIM) and Amirebrahimi et al. (2015) distinguish between integration efforts at *data-*, *process-*, and *application level*. Thus, the vagueness regarding the comparison “BIM-GIS” leave scope for interpretation, e.g. if the integration occurs at *system-*, *method-* or *data level*. Furthermore, the *integration subject* at data level is not specified, since the integration effort may address e.g. databases, information models or ontologies. This paper deals with integration efforts concerning information models, whereas the understanding of information models follows the specification of the Meta-Object Facility (MOF) from Object Management Group (OMG) (2019).

On the other side, there is no common definition regarding the term *integration*. Thus, the term *integration* does not further specify if the information integration effort addresses e.g. the combination, the inclusion or the conversion of information models. Furthermore, integration subjects may refer to information models from different domains, heterogeneous infor-

mation models or different instantiations of the same information model, e.g. different versions. A conceptualization of the term integration is challenging, since the conceptualization need to fit to the context of the research investigation including its limitations. For example, information integration in terms of combination of information models excludes conversion of an information model into another. Therefore, this paper follows and adapts the approach from Scherer and Schapke (2011) and refers to interdependent information models as integration subjects. For instance, three major interdependencies are vertical- (different detailedness), horizontal- (*here*: heterogeneity), and longitudinal interdependencies (changes in time). The differences between the information models occurring along these interdependencies must be “bridged” (Kolbe and Plümer 2004) to create a whole, integrated artifact. Additionally, particular interdependencies belong to different levels of the MOF specification, e.g. instance- and schema level. Consequently, the goal of the information integration process across the domains BIM and GIS is understood as the creation of a whole by bridging differences of interdependent information models at instance- or/and schema-level, whereas combining and interlinking information models are seen as different approaches to achieve this goal. In the scope of this paper, information models provided as input for the information integration effort are called source information model.

#### 1.6 Related literature

Four different categorization subjects were identified in the literature concerning “BIM-GIS Integration” at data level. The first categorization subject is related to the heterogeneity of information models and can be further subdivided into two different categorization approaches: First, the accumulation of information model differences (Brüggemann and von Both 2015; Kolbe and Plümer 2004; Liu et al. 2017; Herle et al. 2020). Here, differences of information models are listed without following a specific approach, e.g. different reference systems, geometric representations, granularity. The second categorization approach is based on rather general approaches following related research subjects like ontology integration (Brodeur 2012) or interoperability models (Herle et al. 2020). For example, the distinction between syntactic, structural and semantic heterogeneities.

The second categorization subject covers integration approaches or -methods. The categorization from Amirebrahimi et al. (2015) is complemented by Zhu et al. (2018) and Liu et al. (2017) by providing examples, whereas Zhu et al. (2018) additionally differentiate between semantic and geometric differences at data level and Liu et al. (2017) enhance the categorization at data level by the subcategories meta-model, extension, conversion/ translation. Furthermore, Hijazi and Donaubaer (2017, p. 44) distinguish between the integration methods conversion of IFC to CityGML, conversion of GIS/CityGML to IFC, Unified Modeling and Linking BIM and UIM. This categorization approach is picked up by Herle et al. (2020). The categorization approaches from Amirebrahimi et al. (2015) and Hijazi and Donaubaer (2017) served as orientation for the developed categories of integration approaches in this paper. Further categorization approaches of information integration methods are provided by e.g. Juan et al. (2006), Kang and Hong (2015).

The third categorization subject addresses the categorization of use cases related to “BIM-GIS Integration” (Liu et al. 2017; Fosu et al. 2015; Song et al. 2017; Noardo et al. 2019). However, the vagueness of the terminologies complicates comparison and evaluation of those categorizations. The use cases were relevant for categorization development concerning the purpose of “BIM-GIS Integration”.

The fourth categorization approach refers to superordinated categorization of information integration efforts. Wang et al. (2019) distinguish between the following information integration efforts: BIM leads and GIS supports, GIS leads and BIM supports, and BIM and GIS equally involved. Additionally, Wang et al. (2019) has assigned expert literature to these superordinated categories with respect to integration methods and use cases. This superordinated categorization approach is considered in the developed categorization concerning the purpose of information integration efforts.

## 2. CATEGORIZATION

### 2.1 Overview

The categorization of information integration efforts is based on three main categories: *information characteristics*, *solution characteristics* and *purpose*. The information characteristics refer to the categorization of information provided by the source information models. The solution characteristics address the categorization of integration- and communication methods. Finally, the categorization of the purpose refers to heterogeneous intentions of the respective information integration efforts. An overview of the categories is provided in Figure 2.

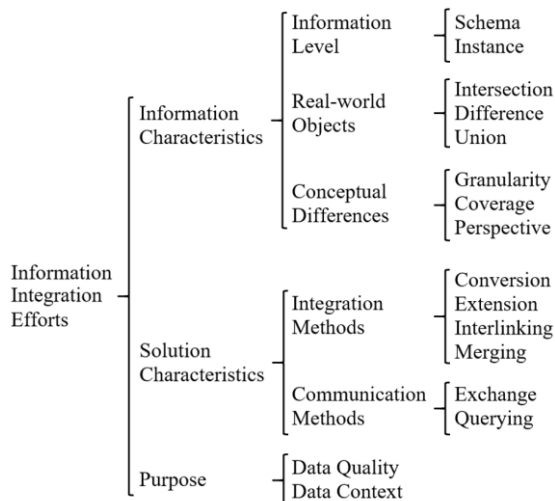


Figure 2: Overview of categories concerning information integration efforts

### 2.2 Purpose

Following West (2011), the “business value of data comes from its use in contributing to sound decisions”. This contribution can occur in different steps of the decision-making process, e.g. problem analysis, simulation, or alternative evaluation. Consequently, the purpose of information integration efforts is based on the use of the integrated information in the respective steps of the decision-making process. Additionally, there is the assumption that the decision-making process refers to a specific subject. Following Wang et al. (2019), there are three different points of views: the decision refers to subjects related to BIM and is supported by information related to GIS, the decision refers to subjects related to GIS and is supported by information related to BIM, or the decision subject refers to BIM and GIS equally. According to the understanding in this paper, the requirement for sound decisions is the sufficient quality of the data embedded in the right context.

- *Data Quality*: The information integration effort aims to improve the data quality of the provided information

set by obtaining relevant information from interdependent information models to satisfy the respective information requirements. Among others, key data requirements are timeliness, reliability, consistency and completeness. For instance, Karan and Irizarry (2016) write that planners require information from different domains for decision-making processes. Moreover, information required for decisions in supply chain management is stored in heterogeneous information models (Karan and Irizarry 2014). Also energy analysis and simulations may need information from heterogeneous information models (Sicilia and Costa 2017).

- *Data Context*: The information of the provided information models is created and embedded in a specific context. Transferring the information to a related context is in some cases purposeful for the decision-making process, e.g. to apply GIS-functionalities on BIM-data or vice versa. For instance, Benner et al. (2005) write that the transformation of data is necessary to make tools from CA(A)D accessible to city modeling. And Li et al. (2020) transfer city- and building information to the context of Precinct Information Modeling (PIM).

In expert literature, the creation of software product interoperability is often highlighted as a major purpose of information integration efforts concerning “BIM-GIS Integration” (Li et al. 2020; Sicilia and Costa 2017). In this paper, software product interoperability is understood as potential capacity to fulfill the purpose of the integration process. Software product interoperability refers to the ability of software products to work with each other, whereas information integration means bridging differences between interdependent information models. Thus, information integration efforts are understood as necessary for the creation of interoperability between software products based on heterogeneous information models. However, not all information integration efforts are intended to achieve software product interoperability, e.g. linking information models to evaluate information consistency.

### 2.3 Information characteristics

**2.3.1 Schema- and Instance-level:** Information integration efforts can be applied at schema- and/or instance-level. The schema level refers to the level M1 of the MOF, whereas the instance level refers to level M0. A clear differentiation concerning schema- and instance-level is in some cases difficult, since information integration efforts at instance-level often implicitly deal with differences at schema level, e.g. when converting an instance-model from IFC to CityGML. In the following, the referenced information integration efforts in the category schema-level address solely differences at schema level, whereas integration efforts referring to the category instance-level may cover differences at both levels.

- *Schema-level*: The information provided in the schema-level are specified using textual or visual information modelling languages like Unified Modelling Language (UML) or EXPRESS. For example, integration efforts mainly focussing on schema level are provided by Kumar et al. (2019), Li et al. (2020) and El-Mekawy et al. (2012).
- *Instance-level*: The instance-level covers instantiations referring to natural or abstract real-world objects. The instance-level information is stored in textual specifications like STEP Physical Model or XML. For example, Donkers et al. (2016) converted IFC datasets to CityGML. Akob et

al. (2019) transferred BIM data to ArcGIS. Moreover, Vilgertshofer et al. (2017) linked instance information of IFC- and CityGML models.

**2.3.2 Conceptual Differences:** Among others, heterogeneous information models cover differences at conceptual level. Following Benerecetti et al. (2001), conceptual heterogeneity of ontologies can be grouped in differences in *granularity*, *coverage*, and *perspective*. The transfer of conceptual differences from the field of ontologies to information models follows the assumption that schemas and ontologies are similar since both provide a vocabulary of terms that describes a domain of interest and both constrain the meaning of terms used in the vocabulary (Euzenat and Shvaiko 2013). However, the clear assignment of conceptual differences to source information models is seen as problematic due to the vagueness of the terms *granularity*, *coverage*, and *perspective*. Furthermore, interdependencies between heterogeneous information models often refer to combined conceptual differences. In the following, three conceptual differences are described and examples concerning combined conceptual differences are provided subsequently.

- *Difference in granularity:* Information models may differ regarding the granularity of the entities. However, a comparison of the granularity level of information models from different domains is often problematic due to the missing consistency.
- *Difference in coverage:* Integration efforts dealing with similar coverage of source information models mean that the information models represent similar real-world entities.
- *Difference in perspective:* Differences regarding the perspective of two source information models result from different views on information.

In the integration effort of Beetz and Borrmann (2018) source information models represent the street network in Netherlands (RWS-OTL and CB-NL) and Germany (OKSTRA). Here, the information models are based on similar coverage and similar perspective. Furthermore, major amount of information integration efforts address the interdependencies of the information models IFC and CityGML concerning building information (see categories *intersection* and *conversion*). The coverage overlaps since both source information models refer (partly) to buildings. The granularity differs because e.g. IFC models provide more detailed entities regarding interior characteristics of physical building elements. The source information models cover differences in perspective, because IFC models are generally created in the design-view (prescriptive) and CityGML models are based on a topographic view (descriptive).

**2.3.3 Real-world Objects:** The objects provided by information models at instance-level refer to real-world objects. However, conceptual and semantic differences impede a clear description whether the respective instances of heterogeneous information models represent the same or different real-world object. The following categories represent a simplistic view using the principles of quantity operators and should be referenced with caution. Nevertheless, this categorization is deemed to be necessary and adequate for the purpose of this paper.

- *Intersection:* The intersection of source information models addresses objects referring to the same real-world object. For example, the source and the target information model of a conversion process refer to the same real-world objects, e.g. IFC model to CityGML model (El-Mekawy et al. 2012; Tauscher 2019; Kang and Hong 2018)

- *Difference:* Information models based on instance-level difference describe different real-world objects. For example, the integration of information models representing road-networks from different countries refer to different real-world objects (Beetz and Borrmann 2018)
- *Union:* Source information models based on both instance-level intersection and -difference refer to the same *and* different objects. For instance, converting an IFC model into CityGML model and embedding that building information in an information model representing the surrounding environment. The source IFC model and the resulting CityGML model refer to the same real-world object, whereas the surrounding environment refers to different real-world objects.

## 2.4 Solution characteristics

**2.4.1 Integration Methods** address the process of bridging the heterogeneity of the source information models. As mentioned previously, a variety of categorizations regarding information integration methods across the domains BIM and GIS were published during the past years. The categorizations approaches from Amirebrahimi et al. (2015) and Hijazi and Donaubaue (2017) served as basis for the developed categorization, which is illustrated in Figure 3.

- *Conversion:* The conversion means the translation of a source information model into a target information model. For instance, Jusuf et al. (2017) converted an IFC model to CityGML model using FME. Tauscher and Stouffs (2018) have applied a triple graph grammar as formal framework for semantic and geometric conversion to CityGML. Rafiee et al. (2014) converted IFC geometry and semantics to vector geographic format using an Extract Transform Load Process. Moreover, Esser and Borrmann (2019) converted a PlanPro-instance model to an IFC model.
- *Extension:* In addition to the conversion method, the extension method enhances the schema of the target model to prevent information loss. Laa and van Berlo (2011) have developed the GeoBIM extension for CityGML and implemented the conversion of IFC to CityGML in the open source Building Information Model server. Hijazi et al. (2011) investigated the CityGML ADE utility network extensions to map interior building utilities in city models by semantic mapping. Borrmann et al. (2015) extended the IFC model for shield tunnels with multi-scale representation and converted the extension to CityGML.
- *Interlinking:* The interlinking of information models means the establishment of explicit links between different source information models. Beetz et al. (2014) have developed a method defining RDF dictionaries for wall structure of IFC models. Hor et al. (Hor et al. 2016; Hor et al. 2018) propose a method to integrate IFC and CityGML using RDF-graphs. Vilgertshofer et al. (2017) connected IFC tunnel proposal with CityGML instance models using Semantic Web technologies. Esfahani (2013) propose the multi-model method according to Scherer and Schapke (2011) link the elements of heterogeneous information models in the field of infrastructure. Beetz and Borrmann (2018) discuss linking RWS-OTL and CB-NL with okstraOWL data. Similarly, Zheng (2017) uses Semantic Web technologies for integrating the OKSTRA road information model and CityGML transportation model.
- *Merging:* The merging method address the creation of a new information model from two existing information models. El-Mekawy et al. call the resulting model *meta-model*. Here, the new information model is called *shared model* since the

name meta-model may be confused with the meta-level of information models according to MOF specifications. The naming *shared model* follows Wache (2002) who used the term *shared vocabulary* for similar approaches in the field of ontology integration. El-Mekawy et al. (2012) have developed the Unified Building Model, in which the schema of CityGML and IFC are merged. Amirebrahimi et al. (2015) have developed an Urban Flood Model as meta-model. Teo and Cho (2016) propose a multi-purpose geometric network model (MGNM) to connect indoor and outdoor network connections. Choi et al. (2008) have developed an Ubiquitous Space Information Model for Indoor GIS. Moreover, Aien et al. (2013) have proposed a 3D cadastral data model (3DCDM). Benner et al. (2005) have developed the QUASY model as meta model concerning IFC and CityGML.

Additionally, combinations of information integration methods at schema and instance level were identified during the literature research. For example, the respective information models are merged at schema-level, and the resulting shared model may be used as intermediate model (El-Mekawy et al. 2012) or as target model for the conversion process (Benner et al. 2005). Another example for combined methods is the interlinking between the source information model and a shared model. Among others, the three methods extension, interlinking and merging are intended to prevent information loss which may occur during plain conversion.

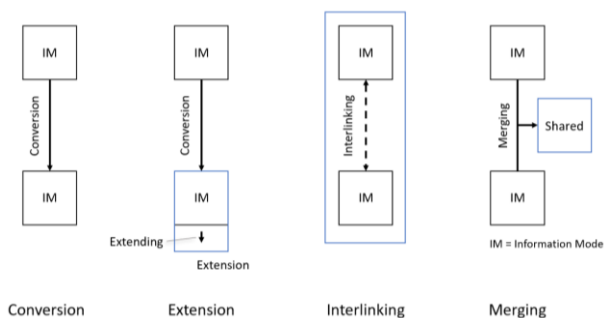


Figure 3: Categorization of integration methods.

**2.4.2 Communication Methods** refer to the way of transmission of the information from the sender to the receiver. Two different ways communicating the information with respect to integration efforts were identified.

- **Exchange:** The exchange of information means the import and export of the integrated information model in software products using network solutions or external storage devices. For instance, Akob et al. (2019) imported BIM data into the software product ArcGIS.
- **Querying:** Communication of information is achieved through querying the integrated information model. For instance, several integration efforts concerning interlinking methods use SPARQL to access the integrated information (Karan and Irizarry 2014; Hor et al. 2016; Zhao et al. 2019; Zhang and Beetz 2016; Sicilia and Costa 2017)

Here, modifications of software products are either part of integration efforts at data or application level. For instance, modifications of software products at data level may be required when interpreting an extended information model, querying, or implementing web services. In distinction to that, modifications of software products implying changes of software product functionalities refer to integration efforts at application level.

### 3. CHALLENGES

#### 3.1 Identified challenges

**Information overload:** The scope of the integrated information needs to fit to the purpose of the integration efforts. The availability of data irrelevant for the respective decision-making process step impedes its target-oriented conduction (Laat and van Berlo 2011; Hijazi et al. 2019).

**File size:** The file size of the integrated information model may be significantly bigger than the file size of the source model. For instance, after conversion of IFC to CityGML (Laat and van Berlo 2011) or after converting the IFC model based on EXPRESS to an ifcOWL model (Pauwels and Roxin 2016). However, the trend in information systems is moving away from file-based information exchange. Thus, file size as challenge may become less relevant or altered.

**Semiotic heterogeneity:** The meaning of symbols depends on the expectations of the information sender or -receiver associated with these symbols (Ogden and Richards 1923). Shared expectations are essential for a common understanding of the relevant concepts and depend on the previous knowledge provided by the communication participants. In general, this previous knowledge differs between communication participants coming from different domains or disciplines. The resulting heterogeneous interpretation of concepts in formal schema specification of source information models (Beetz 2009; Brodeur 2012) or other messages hinders the successful integration of heterogeneous information models.

**Selectivity:** Several information integration efforts are designed for specific, artificial application scenarios and are intended to bridge selected heterogeneities (Hijazi and Donaubaer 2017). Additionally, the developers come from a specific discipline, such that their solutions are driven by a specific perspective. These circumstances often result in selective characteristics concerning the information integration solutions.

**Lack of expert knowledge:** The development of information integration solutions and its usage often requires comprehensive knowledge, e.g. knowledge about the integration subjects from both BIM and GIS and technical knowledge about data modelling and integration (Karan et al. 2016; Liu et al. 2017). However, individual developers and users only seldomly possess this comprehensive knowledge, a fact that results in time intensive training or insufficient solutions.

**Automation of matching processes:** The information integration procedure needs to be automated to prevent time-expensive and error prone matching tasks carried out by the user. Additionally, the automation is required due to the lack of expert knowledge of the user necessary for successful matching. However, the full automation of the matching process is difficult to achieve (Schneider 2019).

**Legal issues:** Legal protection concerning the usage of the information is essential for the acceptance of the respective information integration solution for industrial application. In the field of information integration, the source data comes from different stakeholders maintaining the respective usage rights of the data. Especially in distributed environments with a multitude of stakeholders the protection of intellectual property rights is challenging, e.g. moving BIM data to public city databases (Hijazi and Donaubaer 2017).

*Object identification:* The information referring to the same real-world object need to be identified to create an alignment between the information models at instance-level. This task is called object identification in the field of database integration (Batini and Scannapieco 2016). The identification of information referring to the same real-world object is complicated, since a clear 1:1 mapping is often not possible (Kolbe and Plümer 2004).

*Conflict handling:* Instance information referring to the same object may conflict with each other. The conflict handling strategies is a relevant research topic in the field of database integration (Batini and Scannapieco 2016). The conflicts between source information models are non-trivial in some circumstances. For instance, conflict handling strategies may need extensive expert knowledge concerning the integration subject and depend on the intended context. Moreover, the conflicting information may need to be processed first to identify an existent conflict.

*Validity of information:* The integrated information needs to be valid in the respective context in which it is used. For instance, the functionalities of software applications may be limited to data related to real-world objects at a specific time or specific reference system. Information deviating from these requirements may be invalid in the context wherein the software products functionalities are embedded.

*Assessment of data quality:* The value of a decision strongly depends on the quality of the data on which the decision is carried out. The assessment of the data quality is relevant for information integration processes intended to attain a better data quality. Among others, this challenge addresses the question which data requirements exists for the specific use cases.

#### 4. ASSIGNMENT

The assignment of the challenges to the respective categories means that the challenges rather occur when the information integration effort refers to that category. Here, the aim is to illustrate the importance of a differentiation between heterogeneous information integration efforts. For example, information integration efforts at instance-level may face challenges regarding the model size and information overload, whereas integration efforts at schema level rather deal with semantic and semiotic heterogeneity issues. Object identification and conflict handling primarily occur in integration efforts dealing with several intersecting information models. Consequently, object identification and conflict handling are primarily relevant for the integration methods interlinking and merging. Difference in perspective of information models result in major challenges in terms of bridging semiotic heterogeneities, whereas difference in coverage may emerge in difficulties concerning the deployment of the interface between the information models. Challenges regarding software product interoperability mainly occur in information integration efforts dealing with extension, interlinking, or merging as integration method. Furthermore, interlinking information using Semantic Web Technologies in distributed environments result in challenges regarding data security and access rights. Information integration efforts intended to bring the information into another context may face challenges regarding the validity of the information. And the assessment of the data quality is especially relevant for information integration efforts dealing with multiple information models and intended to improve the respective data quality. In addition, some challenges cannot be assigned to specific categories, but are relevant for all information integration efforts. For instance, the selectivity of the developed solution, the lack of expert knowledge and the need for automation.

#### 5. DISCUSSION

In this paper, the integration of information models is understood as the creation of a whole by combining or linking interdependent information models. In the field of “BIM-GIS Integration” the major interdependency refers to the heterogeneity of information models. Additional interdependencies, like differences caused by time-dependent changes, are generally not considered in the developed information integration efforts. However, this consideration might be relevant since heterogeneous information models are often not intended to represent the real-world objects at the same time.

The concept of enumerative definitions was followed to provide a meaningful assignment of the conducted literature review. However, this process has turned out difficult in some cases due to the vagueness of the respective subcategories. For instance, the assignment of expert literature to categories referring to the real-world objects is complicated, since a clear understanding of the situation when two objects refer to each other is missing. Similarly, the categorization regarding conceptual differences is too vague to assign expert literature in a clear manner. For instance, a comparison of the granularity level between heterogeneous information models is hardly achievable without further specifications. Also, the differences concerning the coverage are hard to identify, due to the ambiguity of the concepts used in the respective schemata.

The developed categorization of integration methods could be adapted with respect to the integration level since some information integration efforts refer to different categories at instance- and schema-level. For example, the usage of a shared model as intermediate model in the conversion process refer to both categories merging and conversion. An adaption of the categorization approach may assign these combined integration approaches in a clearer manner. Additionally, the advantages and disadvantages of the respective integration methods and information exchange methods need to be analysed in more detail.

The dichotomous categorization of the purpose in terms of data quality and data context could be further subcategorized. For this, the different information requirements addressed by information integration approaches need to be investigated and the different data contexts need to be specified. Furthermore, an investigation of the correlation between the subcategories may be valuable, e.g. the correlation between information and solution characteristics categories may indicate to patterns useful for the understanding of information integration efforts in the field of “BIM-GIS Integration”.

#### 6. CONCLUSION

In summary, there are two major findings: First, there is the need for a common understanding of the concept *information integration* referenced by the keyword “BIM-GIS Integration”. This need is underlined by the cumbersome character of the deduction process of an adequate conceptualization concerning these terms for the context of this paper. Second, there is a need for the differentiation between heterogeneous information integration efforts. This need is illustrated by the developed categories regarding information integration efforts and their assignment to corresponding challenges. In conclusion, a scientific discourse which supplies these needs (e.g. by more precise keywords) would be accompanied by positive effects regarding both communication and understanding of information integration efforts currently referenced by the umbrella term “BIM-GIS Integration”.

## 7. ACKNOWLEDGEMENT

This research work was enabled by the support of the Leonhard-Obermeyer-Center ([www.loc.tum.de/en/home/](http://www.loc.tum.de/en/home/)) through the funding of the Gertrude-Obermeyer-Scholarship. Furthermore, the authors would like to thank Ordnance Survey GB ([www.ordnancesurvey.co.uk](http://www.ordnancesurvey.co.uk)) and 1Spatial ([www.1spatial.com/](http://www.1spatial.com/)) for sponsoring the publication of this paper.

## REFERENCES

- Aien, A.; Bennett, R.; Kalantari, M.; Rajabifard, A.; Williamson, I. 2013. Utilising data modelling to understand the structure of 3D cadastres. *Journal of Spatial Science*, 215–234. doi.org/10.1080/14498596.2013.801330.
- Akob, Z.; Abang Hipni, M.; Abd Razak, A. 2019. Deployment of GIS + BIM in the construction of Pan Borneo Highway Sarawak, Malaysia. *IOP Conf. Ser.: Mater. Sci. Eng.* doi.org/10.1088/1757-899X/512/1/012037.
- Amirebrahimi, S.; Rajabifard, A.; Mendis, P.; Ngo, T. 2015. A data model for integrating GIS and BIM for assessment and 3D visualisation of flood damage to building. *CEUR Workshop Proceedings* 1323, 78–89.
- Bar-Yam, Y. 2002. Complexity Rising: From Human Beings to Human Civilization, a Complexity Profile. *UNESCO Encyclopedia of Life Support Systems*.
- Batini, C.; Scannapieco, M. 2016. Data Quality Issues in Data Integration Systems. *Data and Information Quality*. Springer International Publishing, Cham, 279–307.
- Beetz, J. 2009. Facilitating distributed collaboration in the AEC/FM sector using Semantic Web Technologies. With assistance of B. de Vries, JP. van Leeuwen.
- Beetz, J.; Borrmann, A. 2018. Benefits and Limitations of Linked Data Approaches for Road Modeling and Data Exchange. *Advanced Computing Strategies for Engineering*, vol. 10864. Springer International Publishing, Cham, 245–261.
- Beetz, J.; van den Braak, W.; Botter, R.; Zlatanova, S.; Laats, R. de 2014. Interoperable data models for infrastructural artefacts – a novel IFC extension method using RDF vocabularies exemplified with quay wall structures for harbors. *ECPPM 2014*. CRC Press, 135–140.
- Benner, J.; Geiger, A.; Leinemann, K. 2005. Flexible generation of semantic 3D building models. *Proceedings of the 1st International Workshop on Next Generation 3D City Models*, Bonn.
- Borrmann, A.; Kolbe, T. H.; Donaubaue, A.; Steuer, H.; Jubierre, J. R.; Flurl, M. 2015. Multi-Scale Geometric-Semantic Modeling of Shield Tunnels for GIS and BIM Applications. *Computer-Aided Civil and Infrastructure Engineering* 30 (4), 263–281. doi.org/10.1111/mice.12090.
- Brodeur, J. 2012. Geosemantic Interoperability and the Geospatial Semantic Web. *Springer Handbook of Geographic Information*, vol. 22. Springer, Berlin, Heidelberg, 589–611.
- Brüggemann, T.; von Both, P. 2015. 3D-Stadtmodellierung: CityGML. *Building Information Modeling*. Technologische Grundlagen und industrielle Praxis. Springer Vieweg, Wiesbaden, 177–192.
- Choi, J. W.; Kim, S. A.; Lertlakhanakul, J.; Yeom, J. H. 2008. Developing Ubiquitous Space Information Model for Indoor GIS Service in Ubicomp Environment 2. doi.org/10.1109/NCM.2008.97.
- Donkers, S.; Ledoux, H.; Zhao, J.; Stoter, J. 2016. Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. *Trans. in GIS* 20 (4), 547–569. doi.org/10.1111/tgis.12162.
- El-Mekawy, M.; Östman, A.; Hijazi, I. 2012. A Unified Building Model for 3D Urban GIS. *IJGI* 1 (2), 120–145. doi.org/10.3390/ijgi1020120.
- Esfahani, N. N. 2013. Interoperability methods for Infrastructure Design and Information Systems. *Forum Bauinformatik 2013*, 161–171.
- Esser, S.; Borrmann, A. 2019. Integrating Railway Subdomain-Specific Data Standards into a common IFC-based Data Model. *26th International Workshop on Intelligent Computing in Engineering*.
- Euzenat, J.; Shvaiko, P. 2013: *Ontology Matching*. Berlin, Heidelberg. Springer Berlin Heidelberg.
- Fosu, R.; Suprabhas, K.; Rathore, Z.; Cory, C. 2015. Integration of Building Information Modeling (BIM) and Geographic Information Systems (GIS). *Proceedings of the 32nd CIB W78 Conference*. Eindhoven.
- Herle, S.; Becker, R.; Wollenberg, R.; Blankenbach, J. 2020. GIM and BIM. *PFG*. doi.org/10.1007/s41064-020-00090-4.
- Hijazi, I.; Bolduan, T.; Donaubaue, A. 2019. A new approach for integrating BIM and GIS based on Model View Definitions and Profiling Geospatial Application Schemas. *3D GeoInfo Conference*, Singapore.
- Hijazi, I.; Donaubaue, A. 2017. Integration of Building and Urban Information Modeling – Opportunities and Integration Approaches. *Geoinformationssysteme 2017*. Beiträge zur 4. Münchner GI-Runde. Wichmann, Berlin, Offenbach, 42–53.
- Hijazi, I.; Ehlers, M.; Zlatanova, S.; Becker, T.; van Berlo, L. 2011. Initial Investigations for Modeling Interior Utilities Within 3D Geo Context: Transforming IFC-Interior Utility to CityGML/UtilityNetworkADE. *Advances in 3D Geo-Information Sciences*, vol. 13. Springer, Berlin, Heidelberg.
- Hor, A.-H.; Jadidi, A.; Sohn, G. 2016. BIM-GIS Integrated Geospatial Information Model using Semantic Web and RDF Graphs. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* III-4, 73–79. doi.org/10.5194/isprsannals-III-4-73-2016.
- Hor, A.-H.; Sohn, G.; Claudio, P.; Jadidi, M.; Afnan, A. 2018. A Semantic Graph Database for BIM-GIS Integrated Information Model for an Intelligent Urban Mobility Web Application. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-4. doi.org/10.5194/isprs-annals-IV-4-89-2018.
- Kang, T. W.; Hong, C. H. 2015. A study on software architecture for effective BIM/GIS-based facility management data integration. *Automation in Construction* 54, 25–38. doi.org/10.1016/j.autcon.2015.03.019.
- Kang, T. W.; Hong, C. H. 2018. IFC-CityGML LOD mapping automation using multiprocessing-based screen-buffer scanning including mapping rule. *KSCE J Civ Eng* 22 (2), 373–383. doi.org/10.1007/s12205-017-0595-9.
- Karan, E. P.; Irizarry, J. 2014. Developing a Spatial Data Framework for Facility Management Supply Chains. *Construction Research Congress*. Atlanta, Georgia. ASCE, Reston, 2355–2364.
- Karan, Ebrahim P.; Irizarry, Javier; Haymaker, John 2016. BIM and GIS Integration and Interoperability Based on Semantic Web

- Technology. *J. Comput. Civ. Eng.* 30 (3), p. 4015043. doi.org/10.1061/(ASCE)CP.1943-5487.0000519.
- Kardinal Jusuf, S.; Mousseau, B.; Godfroid, G.; Soh Jin Hui, V. 2017. Integrated modeling of CityGML and IFC for city/neighborhood development for urban microclimates analysis. *Energy Procedia* 122, 145–150.
- Kolbe, T. H.; Plümer, L. 2004. Bridging the gap between GIS and CAAD: Geometry, referencing, representations, standards and semantic modelling. *GIM International*, 12–15.
- Kumar, K.; Labetski, A.; Otori, K. A.; Ledoux, H.; Stoter, J. 2019. The LandInfra standard and its role in solving the BIM-GIS quagmire. *Open Geospatial Data, Software and Standards*. doi.org/10.1186/s40965-019-0065-z.
- Laat, Ruben de; van Berlo, Leon 2011. Integration of BIM and GIS: The development of the CityGML GeoBIM extension. *Advances in 3D geo-information sciences*. Springer, 211–225.
- Li, Wei; Zlatanova, Sisi; Diakite, Abdoulaye A.; Aleksandrov, Mitko; Yan, Jinjin 2020. Towards Integrating Heterogeneous Data: A Spatial DBMS Solution from a CRC-LCL Project in Australia. *IJGI* 9 (2), p. 63. doi.org/10.3390/ijgi9020063.
- Liu, Xin; Wang, Xiangyu; Wright, Graeme; Cheng, Jack; Li, Xiao; Liu, Rui 2017. A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS). *IJGI* 6 (2), p. 53.
- Noardo, Francesca; Ellul, C.; Harrie, L.; Overland, I.; Shariat, M.; Arroyo Otori, Ken; Stoter, Jantien 2019. Opportunities and challenges for GeoBIM in Europe: developing a building permits use-case to raise awareness and examine technical interoperability challenges. *Journal of Spatial Science* 42 (9).
- Object Management Group 2019. OMG Meta Object Facility (MOF) Core Specification. Available online at www.omg.org, checked on 3/25/2020.
- Ogden, Charles Kay; Richards, Ivor Armstrong 1923: *The Meaning of Meaning*. Harcourt, Brace and World.
- Pauwels, Pieter; Roxin, Ana 2016. SimpleBIM: From full ifcOWL graphs to simplified building graphs. *ECPPM*. Limasol, Cyprus, 2016.
- Peffer, Ken; Tuunanen, Tuure; Rothenberger, Marcus; Chatterjee, S. 2007. A design science research methodology for information systems research. *Journal of Management Information Systems* 24, 45–77.
- Powrie, William 2014. On track: the future for rail infrastructure systems. *Proceedings of the Institution of Civil Engineers - Civil Engineering* 167 (4), 177–185.
- Rafiee, Azarakhsh; Dias, Eduardo; Fruijtier, Steven; Scholten, Henk 2014. From BIM to Geo-analysis: View Coverage and Shadow Analysis by BIM/GIS Integration. *Procedia Environmental Sciences* 22, 397–402. doi.org/10.1016/j.proenv.2014.11.037.
- Scherer, R. J.; Schapke, S.-E. 2011. A distributed multi-model-based Management Information System for simulation and decision-making on construction projects. *Advanced Engineering Informatics* 25 (4), 582–599.
- Schneider, Georg 2019. Automated Ontology Matching in the Architecture, Engineering and Construction Domain - A Case Study. *7th LDAC Workshop*.
- Sicilia, Álvaro; Costa, Gonçal 2017. Energy-Related Data Integration Using Semantic Data Models for Energy Efficient Retrofitting Projects.
- Smith, Adam 2010: *The Wealth of Nations. An inquiry into the nature and causes of the Wealth of Nations*. Harriman House Limited.
- Song, Yongze; Wang, Xiangyu; Tan, Yi; Wu, Peng; Sutrisna, Monty; Cheng, Jack; Hampson, Keith 2017. Trends and Opportunities of BIM-GIS Integration in the Architecture, Engineering and Construction Industry. *A Review from a Spatio-Temporal Statistical Perspective*. *IJGI* 6 (12), p. 397.
- Steward, Colin; Luebke, Chris; Morrell, Marcus; Goulding, Lynne 2014. Future of Rail 2050. ARUP. London.
- Tauscher, H. 2019. Creating and Maintaining IFC-CityGML Conversion Rules. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 115–122. doi.org/10.5194/isprs-annals-IV-4-W8-115-2019.
- Tauscher, Helga; Stouffs, Rudi 2018. An IFC-to-CityGML Triple Graph Grammar. *Computing for a better tomorrow*. Lodz, Poland.
- Teo, Tee-Ann; Cho, Kuan-Hsun 2016. BIM-oriented indoor network model for indoor and outdoor combined route planning. *Advanced Engineering Informatics* 30 (3), 268–282. doi.org/10.1016/j.aei.2016.04.007.
- Vilgertshofer, S.; Amann, J.; Willenborg, B.; Borrmann, A.; Kolbe, T. H. 2017. Linking BIM and GIS Models in Infrastructure by Example of IFC and CityGML. *International Workshop on Computing in Civil Engineering*. Seattle. ASCE, 133–140.
- Wache, Holger; Voegelé, Thomas; Visser, Ubbo; Stuckenschmidt, Heiner; Schuster, Gerhard; Neumann, Holger; Hübner, Sebastian 2002. Ontology-Based Integration of Information. *Proceedings of the IJCAI*.
- Waldmann, Michael 2008. Kategorisierung und Wissenserwerb. *Allgemeine Psychologie*. 2<sup>nd</sup> ed. Spektrum Akademischer Verlag, Heidelberg, 377–427.
- Wang, Hao; Pan, Yisha; Luo, Xiaochun 2019. Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis. *Automation in Construction* 103, 41–52.
- West, Matthew 2011: *Developing high quality data models*. Amsterdam, Boston. Elsevier/Morgan Kaufmann.
- Wunsch, Susanne; Lehnert, Martin; Krimmling, Jürgen; Easton, John 2016. Datenformate, -modelle und -konzepte für den Eisenbahnbetrieb. *Der Eisenbahningenieur*, 8–14.
- Zhang, C.; Beetz, J. 2016. Querying linked building data using SPARQL with functional extensions. *ECPPM*. Limasol, Cyprus, 2016.
- Zhao, Linlin; Liu, Zhansheng; Mbachu, Jasper 2019. Highway Alignment Optimization: An Integrated BIM and GIS Approach. *IJGI* 8 (4), p. 172. doi.org/10.3390/ijgi8040172.
- Zheng, Y. 2017. Improving the interoperability between city and road semantics an integration of cityGML and OKSTRA data based on semantic web technologies. Master Thesis. Eindhoven University of Technology, Eindhoven.
- Zhu, Junxiang; Wright, Graeme; Wang, Jun; Wang, Xiangyu 2018. A Critical Review of the Integration of Geographic Information System and Building Information Modelling at the Data Level. *IJGI* 7 (2), p. 66. doi.org/10.3390/ijgi7020066.