Conception, Validation and Publication of ETL Processes for the Conversion of Geospatial Data for openBIM Projects with FME

Christoph FREY, Christian DAHMEN, Christian CLEMEN, Germany

SUMMARY

Building Information Modelling (BIM) is a methodology that facilitates collaboration through the creation and shared use of digital models. These models contain comprehensive information for the entire life cycle of a building. This methodology improves interoperability, significantly enhances planning efficiency, and reduces the need for redundant data entry that often burdens conventional workflows. Since buildings exist in an environment, geospatial data plays a crucial role in various phases of BIM projects and needs to be seamlessly integrated into these processes. To achieve this integration, openBIM projects rely on open standards such as the Industry Foundation Classes (IFC) developed by buildingSMART International (bSI). However, transforming geospatial data into IFC is a complex task due to the different modeling intents of BIM and GIS. They utilize different paradigms for semantics, geometric representation, and georeferencing. Given this complexity, many software providers have developed specialized solutions to efficiently and accurately transform geospatial data into the IFC format. One such solution is the Feature Manipulation Engine (FME) by Safe Software Inc., an Extract-Transform-Load (ETL) software application widely used in industry, public authorities / agencies and academia.

The research findings are implemented workflows tailored to transform geospatial data into valid IFC models using FME. These workflows will be published as a framework that can be expanded and adapted in the future. In preparation for creating the workflows, decision trees were designed to outline the transformation steps for each data source and create a clear roadmap for implementation. The ETL process is not a simple format conversion or 1:1 schema mapping. It is crucial to effectively communicate the IFC concepts so that non-geospatial BIM modeling and coordination software can interpret the geospatial content generated in the IFC models effectively, considering that BIM and GIS serve different purposes.

Overall, this research highlights the importance of integrating geospatial data into BIM processes and presents FME as a powerful tool for achieving this integration. The goal of this research was to develop automated workflows using FME to transform 2D cadastral data and 2.5D terrain models in different formats into IFC.

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

Conception, Validation and Publication of ETL Processes for the Conversion of Geospatial Data for openBIM Projects with FME

Christoph FREY, Christian DAHMEN and Christian CLEMEN Germany

1. INTRODUCTION

Building Information Modelling (BIM) is a collaborative working method that uses digital models to record, manage, and exchange information and data relevant to the life cycle of a building. This process involves all parties participating in the project and aims to promote interoperability. The use of BIM leads to improved data exchange and planning efficiency by eliminating time-consuming and error-prone re-entry of information that is common in conventional work processes. Geospatial data, which serves as the basis for planning in various service phases of BIM projects, must also be integrated into the processes. Therefore, open standards are necessary to enable the sharing of all collected data.

To facilitate this, buildingSMART International (bSI) has developed the Industry Foundation Classes (IFC), an open standard for the vender-neutral exchange of data. IFC is the most well-known and widely used format for openBIM since many years.

Geospatial data plays a critical role in BIM projects, providing information such as topography, terrain, and legal rights of the building environment. It ensures accurate visualization and analysis of the site, enabling designers and other stakeholders to make informed decisions. Geospatial data is particularly useful in urban planning, infrastructure projects, and site analysis. Overall, geospatial integration enables better decision-making, reduces errors and conflicts, and improves project coordination and collaboration. Geospatial integration also supports the optimization of building performance and the implementation of sustainable design principles.

However, most software systems used in BIM projects do not support geospatial data. Therefore, the geospatial data drops need to be prepared in such a way that they can be used for specific aspects of the design, construction, and operation of buildings. It should be noted that the IFC data schema is not suitable for maintaining geospatial data in a Geospatial Data Infrastructure (GDI).

To achieve cross-domain information continuity, geospatial and BIM models can be converted (BIM-2-GIS, GIS-2-BIM), merged into a unified model (BIM-and-GIS), or linked (BIM-vs-GIS). The solution presented in this paper belongs to the GIS-2-BIM category.

The primary focus of this study is to investigate the effectiveness of FME in converting 2D parcel data (vector) and 2.5D terrain models (raster, TIN) into the IFC format while preserving geometry, georeferencing, topology, and semantics. To achieve this, four key research questions are addressed:

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

- *How can 2D parcel data and 2.5D terrain models (DTM) be converted to IFC while retaining their classification and attributes?*

- *What criteria can be used to evaluate the success of these conversions and how can independent validation be achieved?*

- *How can different georeferencing methods within IFC be implemented programmatically using FME?*

- *How can geospatial data be aggregated according to BIM conventions to make it usable for collaboration in BIM software?*

Overall, the incorporation of geospatial data enables decision making, reduces errors and conflicts, and improves project and collaboration. Geospatial integration also supports optimization of building performance and the implementation of sustainable design principles.

2. Related Work

The integration of GIS and BIM has made significant progress in recent years, both in research and practice. Efforts have been made to improve interoperability between GIS and BIM technologies.

Several software solutions are available to facilitate interoperability between BIM and GIS. Revit, Allplan, ArchiCAD, and Blender-BIM are well-known BIM authoring software products that support comprehensive building and infrastructure modeling and coordination of design and construction processes. ArcGIS and QGIS are widely used desktop GIS that specialize in capturing, managing, analyzing, and publishing geospatial data. They allow you to create maps and perform geographic analysis. FME is a powerful Extract-Transform-Load (ETL) tool that originated in the GIS domain. FME enables the integration of BIM and GIS, facilitating the seamless transfer of data between different applications and formats.

The textbook "Building Information Modeling: Technology Foundations and Industry Practice" by Borrmann et al. (2018) provide a comprehensive overview of the technological foundations of BIM and its application in the construction industry, including the integration of geospatial information. Part I discusses the technological foundations of BIM, exploring computational methods for both geometric and semantic building modeling, as well as process modeling techniques. In Part II, the focus shifts to the critical aspect of BIM software interoperability, with a detailed examination of the standardized IFC data format. The section also discusses various classification systems, introduces the CityGML data format for 3D city models, and provides insights into BIM programming tools and interfaces.

The 2019 GeoBIM benchmark study conducted by Noardo et al. (2021) provides many valuable insights into the performance of software solutions in processing geospatial and BIM data,

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

leading to further exploration of their integration in different application domains. The study highlights the underutilization of Industry Foundation Classes due to problems in their implementation and use. Although practitioners and researchers recognize its potential, there has been a lack of systematic research into these issues. The GeoBIM benchmark project aims to fill this gap by engaging external volunteers to report on aspects such as geometry, semantics or georeferencing. The results reveal inconsistencies and significant problems in various IFC software programs, likely due to incomplete implementation of the IFC data schema. With regard to FME, the GeoBIM study highlights the software's commendable ability to facilitate the seamless integration of geospatial and BIM data sets. FME excelles at data transformation tasks and demonstrating its ability to handle multiple data formats and complex spatial relationships. However, the study also identifies potential areas for improvement and highlights the need for improved performance, particularly when dealing with large and complex data sets.

Correct georeferencing is highlighted as an important aspect. The concept of Level of Georeferencing (LoGeoRef), developed by Clemen and Görne (2019), breaks down the depth of georeferencing using existing IFC entities. Donaubauer et al. (2022) emphasize that particularly large coordinate pairs and the resulting large distances from the coordinate origin can cause problems in some software applications, leading to loss of accuracy or other undesirable effects. It is therefore essential to perform the transformation into project coordinates.

The converter software IfcTerrain, as described by Clemen et al. (2021), focuses on the integration of terrain models in BIM projects, especially in construction projects where terrain is critical. IfcTerrain allows a comprehensive conversion of 2.5D terrain models to IFC and analysis of terrain information in the BIM context. The primary objective of this study is to generate different types of geometric representations within IFC for Digital Terrain Models (DTMs).

The paper "Integration of BIM and GIS: Geometry from IFC to shapefile using open-source technology" by Zhu et al. (2019) addresses the challenge of transforming geometry from IFC to shapefile using open source technology. The researchers propose an open source approach (OSA) that exploits the spatial structure of IFC, specifically the IFC tree, and uses an automatic multipatch generation algorithm (AMG) for the conversion process. The research finds that OSA is more efficient than the commonly used Data Interoperability Extension for ArcGIS (DIA), comparable to Feature Manipulation Engine (FME), and results in a transformed model that is easier to manage and applicable in different scenarios. Despite these advantages, the study acknowledges the need to further improve the efficiency of OSA compared to FME.

Ongoing research and development in the integration of GIS and BIM, particularly using the IFC format, is progressing. These approaches are contributing to the seamless collaboration of these two technologies in the construction industry, increasing the efficiency of construction projects worldwide.

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

3. FME AND IFC: REQUIREMENTS AND IMPLEMENTATION

The process of transforming geospatial data into the IFC format is complex and requires an appropriate software solution such as the Feature Manipulation Engine (FME). Despite the benefits and versatility of FME, preparing geospatial data for openBIM requires both knowledge of the FME toolsets and a thorough understanding of the object-oriented conceptual IFC schema. A particular challenge is the implementation of objectified relationships in IFC (concepts starting with *IfcRel...*) and the integration of geospatial data into a spatial aggregation hierarchy (*IfcSpatialStructureElement*), as spatial decomposition is not widely used in GIS.

3.1 FME functionalities

The FME platform includes FME Form, FME Flow, FME Flow (Hosted), FME Mobile, and FME AR (Augmented Reality). FME Form allows users to read, inspect, transform and write data within the Workspace, a virtual workspace where the process is stored. It uses spatial queries, geometry transformations, and data manipulation to transform data into a new format. The main components of FME Flow are the Reader, Transformer, and Writer.

- **Readers.** In FME terminology, a reader is a tool that imports data from various sources into the workspace and acts as an interface to make geospatial data available for further processing. FME supports the import of over 450 different source data types, including geospatial data, databases, and web services.

- **Transformers.** Transformers are at the heart of FME, enabling the transformation, manipulation, and analysis of source data. With 485 officially integrated transformers and over 900 user-created transformers available on FME Hub, they offer versatility. Transformers can modify geometry, calculate or change attributes, filter, validate, sort, group, and perform various other transformations. They are modular and can be used individually or in combination to perform complex operations.

- **Writer.** A writer is an FME tool that writes processed data to a target dataset, making the data accessible to other applications or systems. Writers require the selection of a supported file format, path, and name after being added to the workspace. The choice of writer name can be critical, especially for formats such as IFC, as it affects the contents of the written file. Not all formats supported by FME support both reading and writing, some support only one or the other.

3.2 Methodology

It is not possible to map a geospatial data schema 1:1 to the IFC schema. The process steps between the extraction of the geospatial data (reader) and the writing of the IFC file (writer) therefore include transformation steps (transformer) that take place on several conceptual levels. In order to create a valid IFC dataset, certain requirements must be met:

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

- **Attribute and value generation.** The designed FME workflow must be able to generate specific attributes and values for the corresponding building components, as IFC includes different entities such as walls, ceilings, windows and doors with different properties.
- **Generate the spatial aggregation hierarchy of the IFC data model.** The designed FME workflow needs to represent the spatial structure in the IFC dataset. This includes mapping rooms to floors, buildings, and parcels.
- **Correctly process geometry.** The designed FME workflow must process geometry objects without data loss to ensure a valid and complete geometric representation in the IFC model.
- **Correct processing of semantic information.** The designed FME workflow must transfer semantic information of the geospatial data such as classification, properties and functions of building components to provide a comprehensive description of the geographic elements in the IFC file.
- **Georeferencing.** The designed FME workflow must locate georeferenced GIS features in the project coordinate system to ensure proper coordination of federated models.

The high-level steps of the implemented FME workflow are described by the following diagram (see figure 1).
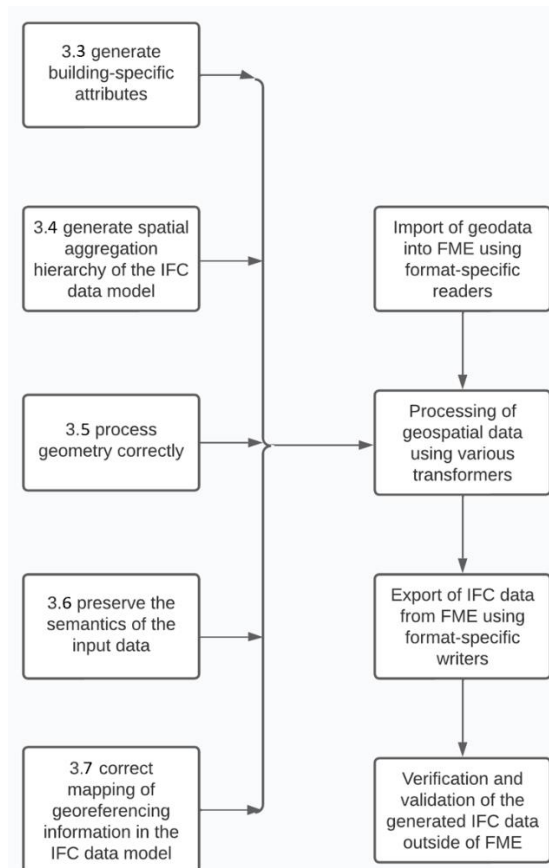

Figure 1 General process for transforming geospatial data to IFC

3.3 Generation and assignment of IFC attributes

When transforming geospatial data into the IFC format, the creation of new attributes and the assignment of values to the appropriate IFC classes is critical. FME offers a range of transformers that provide different methods for attributes, depending on the requirements of the IFC data structure. The following FME transformers are particularly useful: The *AttributeRenamer*, which flexibly renames selected attributes, the *AttributeKeeper*, which restricts the attributes of the object to a selected subset, and the *AttributeCreator* and *AttributeManager*. The latter is used to create, manage, and edit attributes and their values within a dataset. The *AttributeCreator* allows the creation of new attributes and values, while the *AttributeManager* allows new attributes and values to be accessed, edited or values to be accessed, edited or added.

The assignment of attribute values can vary depending on both the type of underlying geospatial data and the resulting IFC class. In the designed FME workflow, semantic information is conceptualized using constants, direct values from the source data, functions, or values assigned by user parameters. User parameters as shown in figure 2 are special settings that allow certain aspects of the workflow to be controlled by user input. They are particularly useful for parameterizing a set of FME workflows for automated processing. Before each execution of the FME workspace, these parameters are automatically queried and can be customized by the user.

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

Figure 2 Graphical user interface (GUI) for the parameterization of the desigend FME workflow

User parameters are mainly used to control data flow, handle input and output data, and configure coordinate systems. In the designed FME workflow, all other attributes are organized according to IFC classes and written to an external JSON file. The data is represented as nested key-value pairs enclosed in parentheses. This JSON file is added to the workspace at an appropriate location, allowing the user to adjust the attribute values before each run of the workspace. This simplifies the process of inserting attribute values at the appropriate locations in the IFC dataset.

3.4 The spatial aggregation hierarchy in the IFC data model

The spatial aggregation hierarchy facilitates the description of spatial elements and their relationships within the IFC model. The spatial aggregation hierarchy consists of several IFC

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

classes that are linked in a hierarchical structure. The top level is represented by the *IfcProject* class, which represents the entire building project. This is followed by the *IfcProject*, *IfcSite*, *IfcBuilding*, *IfcBuildingStorey* and *IfcSpace* classes, which refine the spatial hierarchy of the building into non-overlapping spatial decompositions. In order to model the spatial aggregation hierarchy in the IFC data model with FME, the *ifc_ids* of individual classes must be linked.
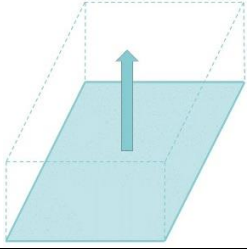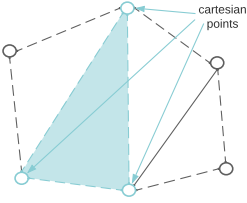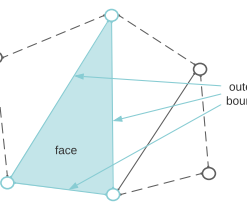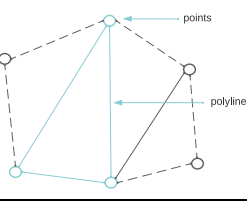
Assuming that *ifc_ids* of features such as parcels can be inherited from one IFC class to another, the hierarchy can be modeled using the *AttributeCreator* transformer. A new attribute is assigned the corresponding value of the parent class from the aggregation hierarchy. For example, the *IfcBuilding* class contains one or more objects of the *IfcBuildingStorey* class. The *IfcBuilding* class has an attribute *ifc_id* that must be passed to *IfcBuildingStorey* to establish the hierarchical relationship between the two classes. To do this, a new attribute called *ifc_parent_id* is created in the *IfcBuildingStorey* class again using the *AttributeCreator* transformer and populated with the value of *ifc_id* from *IfcBuilding*. In addition, a new unique *ifc_id* must be generated for IfcBuildingStorey, which can be accomplished in FME using the UUIDGenerator transformer or by calling the *UUID()* function within the *AttributeCreators*. These steps create a hierarchical relationship between the *IfcBuildingStorey* and *IfcBuilding* classes, and the spatial aggregation hierarchy can be modeled for each class under the conditions described above.

It is important to note that neither parcels nor DTMs show up with any buidlings. Therefore, careful consideration must be given to whether to create "empty" *IfcBuilding* objects to accurately represent the aggregation hierarchy, or to omit them for these cases. While most validation software may throw an error if the hierarchy doesn't have an *IfcBuilding*, this is not essential for parcel or terrain data.

3.5 Geometry processing of input data

The cadastre in Germany is two-dimensional. However, BIM projects require 3D geometries or even solids. Therefore, the polygons of the parcels must either be extruded with the FME Transformer "Extruder" (see SweptAreaSolid in table 1). Or the triangulated terrain (TIN) is represented as a 2.5D coverage on which the parcels are draped. The designed FME process can be parameterized according to the desired variant.

Table 1 The desigend FME workflow alows for four variants how to covert the 2D/2.5D-data to IFC

geometry representation	Description (buildingSMART, 2023)	example
SweptAreaSolid (IFC2x3, IFC4)	Swept Area Solid includes several specific types of geometry created by moving a cross-sectional area along a path curve. It allows the representation of various geometry operations such as extrusion, rotation and translation of a circular area.	
Triangulated-FaceSet (IFC4)	Triangulated Face Set is a geometry representation that represents the surface of a three-dimensional object by triangulation. The surface is a collection of triangles described by their vertices.	
ShellBased-SurfaceModel (IFC2x3)	Shell-Based Surface Model is a geometry representation that represents the outer surface of a three-dimensional object using a shell. The shell is a collection of closed faces that together define the surface of the object.	
Geometric-CurveSet (IFC2x3)	Geometric Curve Set is a geometry representation that describes a three-dimensional object using a collection of curves. These curves can be lines, arcs or complex geometric shapes.	

The way the Digital Terrain Model (DTM) is parameterized is particularly important for BIM practice. For DTMs, TriangulatedFaceSet, ShellBasedSurfaceModel, and GeometricCurveSet are possible geometric representations. However, the practical problem of data exchange is often that the BIM software used in the project does not fully support all geometric representations of the IFC standard. This is especially true for DTMs. For these, the geometric representation must be specified in the information exchange requirements so that the software used in the BIM project can actually use the DTMs in IFC.

Some IFC classes, such as *IfcProject*, have no geometric representation, or geometry is not required at certain points in the transformation. Therefore, before implementing the FME Workbench, it is advisable to determine when geometry is needed and when it is not, in order to make the process as efficient as possible and to avoid unnecessarily increasing the size of the

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

result. Processing geometry from geospatial data follows a workflow that rarely changes. The following steps, shown in figure 3, have proven to be effective.
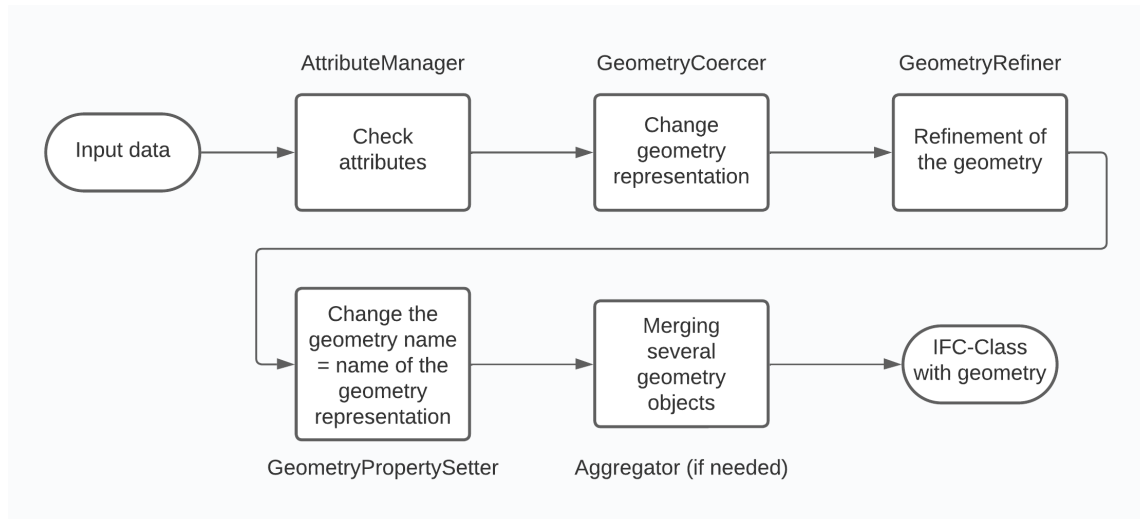

Figure 3 Processing of geometry with FME

3.6 Processing of semantic information

The creation of an IFC model relies on the incorporation of semantic information, namely classification and attribution, where classification is realized with 1) IFC entities, 2) entitytypes, and 3) IFCClassificationReference. Attribution is realized with 1) schema attributes of IFC entities, 2) predefined property sets, and 3) user-defined property sets. In particular, user defined property sets (*IfcPropertySet*) play a critical role in GIS-2-BIM conversions. An *IfcPropertySet* is a dedicated collection of properties associated with a particular IFC element. The definition and use of property sets follows the buildingSMART IFC specifications and guidelines. FME allows the creation of custom property sets, which must be specified in FME as *ifc_property_set_name*. Overall, property sets increase the semantic depth and richness of an IFC dataset, allowing for a nondestructive conversion of semantics from GIS to openBIM.

As with any other IFC class, a dedicated Writer or FeatureWriter transformer is required in FME to store property set definitions. Writers are named according to the IFC element being generated; for example, the writer for the walls of a structure is named *IfcWall*, while the writer for the entire project is named *IfcProject*. However, it is important to note that a new writer is not created for each element. Instead, either a single writer per project is used, duplicated for each element in the structure and renamed based on the IFC element, or a *FeatureWriter* transformer is used that is equipped with all the necessary IFC classes. The spelling of the IFC classes must be exactly the same as specified by (buildingSMART, 2023) in the relevant IFC documentation. The correct spelling ensures consistent and standards-compliant representation of IFC elements in the model.

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

3.7 Options for georeferencing in the IFC data model

The representation of georeferencing in the IFC data model facilitates the connection between the project coordinate system used by IFC and the corresponding (geo) coordinate reference system (CRS). IFC provides several options for storing georeferencing information, with different and sometimes redundant concepts. These options range from simple address information to specifying the geolocation of a reference point and defining an offset between the project coordinate system and the CRS, including a corresponding rotation of the XY plane. Clemen and Görne (2019) have categorized these georeferencing options into different "levels of georeferencing" (LoGeoRef), ranging from LoGeoRef 0 to LoGeoRef 50 (see table 2). Although not formally defined in the IFC standard, this classification is specified in the ISO technical report 23262:2021 (ISO, 2021) and helps to assess the georeferencing information available in the IFC file.

Table 2 level of georeferencing (LoGeoRef) provied by the IFC schema

LoGeoRef	supported CRS	Description
LoGeoRef 0	No CRS	No georeferencing
LoGeoRef 10	No CRS	Approximate location based on the address
LoGeoRef 20	WGS84	geographical coordinates in the attributes RefLatitude, RefLongitude, RefElevation attributes within IfcSite
LoGeoRef 30	Cartesian CRS, Projected coordinates	Definition of the coordinate system for IfcSite or IfcBuilding in IfcCartesianPoint without relative reference to higher-ranking coordinate system / IfcDirection stores Rotations in relation to global north
LoGeoRef 40	Cartesian CRS, Projected coordinates	IfcGeometricRepresentationContext stores the coordinates of the project coordinate system in the WorldCoordinateSystem attribute and the TrueNorth direction (local direction vector from true north)
LoGeoRef 50	Specific projected CRS Specified by EPSG code	Reference point coordinates saved in IfcMapConversion using the attributes Easting, Northing and OrthogonalHeight, rotation of the XY plane saved in XAxisAbscissa and XAxisOrdinate / CRS saved in IfcProjectedCRS using the "Name" attribute with corresponding EPSG code

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

In this research, georeferencing is implemented according to the requirements of LoGeoRef 10, 20 and 50. For LoGeoRef 10, various address data are taken from input data or configuration files and stored in *IfcSite*. For LoGeoRef 20, geographic coordinates are stored in the *IfcSite* attributes *RefLatitude*, *RefLongitude* and *RefElevation*. However, implementing LoGeoRef 50 was challenging because the required IFC entities, *IfcMapConversion* and *IfcProjectedCRS*, could not be created in FME using standard FME writer or creator functions. To achieve better LoGeoRef 50 georeferencing, a custom transformer called *IfcMapConversionWriter*, developed by Stijn Goedertier (2022), is used. The transformer uses a Python script to create the *IfcMapConversion* and *IfcProjectedCRS* attributes, successfully implementing LoGeoRef 50. It requires an existing IFC dataset with an *IfcProject* entity where the two attributes are stored and filled with the configured values. The custom transformer enhances FME's georeferencing capabilities and is an important addition for modeling IFC data.

4. CREATION AND IMPLEMENTATION OF SAMPLE WORKFLOWS IN FME

First, the information exchange requirements were specified and an FME-independent, i.e. purely technical, decision tree for the GIS-2-BIM conversion was created. Two FME workbenches were then developed as examples: One for a parcel map (Parcel-2-IFC), a second one for the DTM (DTM-2-IFC). The workflow is as follows:

1. Input data management
2. JSON configuration file management
3. Attribute management and generation
4. Georeferencing preparation
5. Conversion to BIM project coordinates
6. Classification into IFC classes
7. Geometry processing of input data
8. Create Property and QuantitySets
9. Output data management (writing the entire IFC dataset)

The flow of data within the workspaces is controlled by predefined user parameters and guided by various choices within the workspaces. This provides a high degree of customisation and flexibility in the processing of geospatial data and its conversion to IFC format.

4.1 2D geospatial data (parcels) to IFC

The process of converting 2D parcels GML-geospatial data to IFC4 has been carried out through the nine steps mentioned above. This approach allows for the generation of different geometric representations and IFC classes, with users having the ability to control the depth of information in the IFC file to a certain extent through different user parameters and the JSON configuration file, adapting it to the specific needs of their projects. The IFC data is georeferenced at three levels (10, 20, and 50). The result of the transformation is illustrated by a parcel in the cadastral district of Bad Kleinen, as shown in Figure 4. The spatial aggregation

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

hierarchy of the IFC data model is shown in the browser toolbar on the left. The parcels have been draped over the terrain, with shading visible in the 3D view. The properties of each IFC class can be found on the right hand side in the property toolbar.

However, there is room for improvement. One approach is to create multiple custom IFC property sets, especially for personal parcel owner data, which can be extracted from the input data and stored in attributes. Another improvement would be the possibility to create an *IfcSite* entity for each parcel instead of an *IfcGeographicElement*.
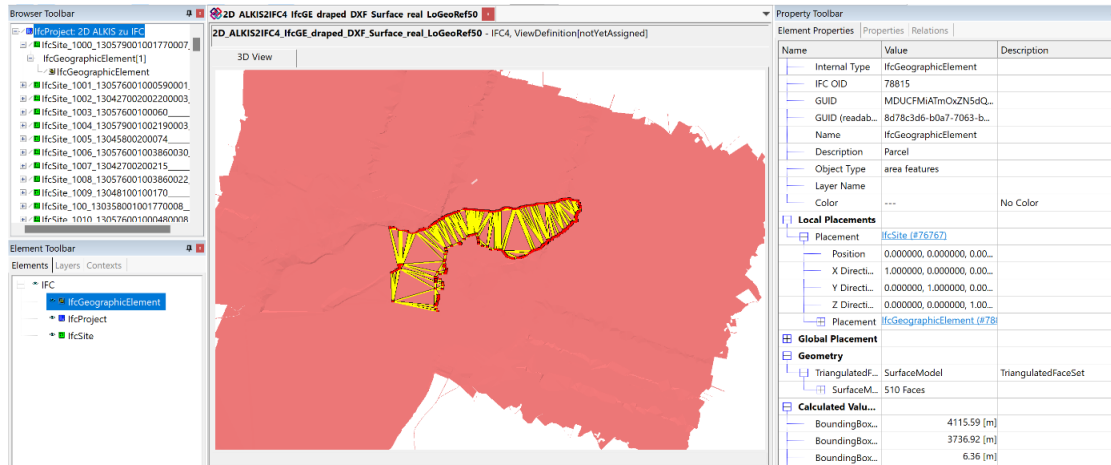

Figure 4 2D parcels of Bad Kleinen shown as IFC in the FZK Viewer

4.2 2.5D geospatial data (DTM) to IFC

The transformation of terrain models into IFC was carried out according to the nine specified steps. In particular, the workflow introduces a user-friendly aspect by incorporating user parameters and integrating a configuration file that allows extensive customization of the IFC data. This allows users to create different combinations of IFC classes and geometric representations. Of note is the workflow's ability to generate both IFC2x3 and IFC4 versions, with the geometric representations in IFC2x3 designed for wider interpretability and display compatibility with most BIM software, as highlighted in Clemen et al. (2021). The workflow ensured successful georeferencing according to LoGeoRef requirements at levels 10, 20 and 50. The result of the transformation is shown in Figure 5 for the Bad Kleinen DTM as IFC4 data.

However, challenges arose during the transformation process regarding the conversion of the georeferencing raster (geoTIFF) to the BIM project coordinates.
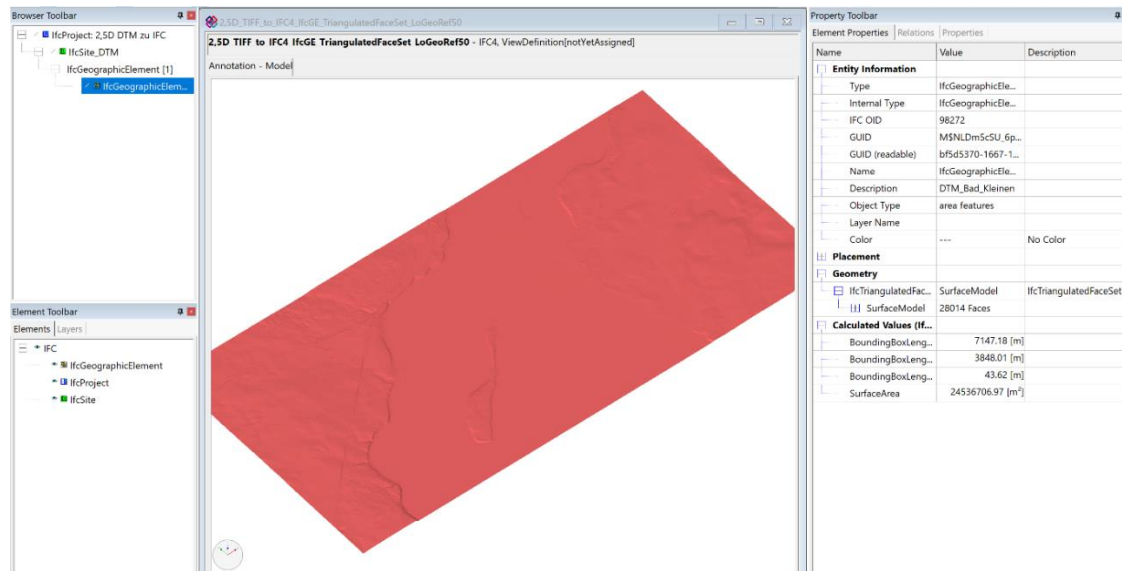
ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

Figure 5 2.5 DTM of Bad Kleinen shown as IFC in the FZK Viewer

5. CONCLUSION AND OUTLOOK

The research aims to develop different concepts for transforming 2D geospatial data and 2.5D terrain models into openBIM. FME software is used as an implementation tool to test different file formats considering geometry, topology and semantics. The concepts are realized in FME Flow through different workspaces that address the formulated FME requirements with a special focus on georeferencing. Automated workspaces generate various IFC data from the underlying dataset, which are then validated and evaluated according to specified criteria. All research questions are successfully answered throughout the study and can be summarized on the basis of these questions.

- *How can 2D parcel data and 2.5D terrain models (DTM) be converted to IFC while retaining their classification and attributes?*

Decision trees were used to guide the transparent transformation process, enabling early problem identification, solution development and process documentation. Clear requirements for FME were defined and successfully implemented using appropriate transformers and parameters, resulting in two automated workspaces generating valid IFC files for a set of test data.

- *What criteria can be used to evaluate the success of these conversions, and how can independent validation be achieved?*

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

Once the various data sets have been successfully transformed into openBIM, the IFC data must be tested and validated against appropriate parameters. Validation of the written IFC records is performed using several parameters:

1. **Schema Conformance.** Verifies that the generated IFC model conforms to the specified schema conventions.
2. **Completeness.** This includes two aspects:
   a. Hierarchy: It is checked whether the IFC inheritance hierarchy is correctly modeled, including all parent and child elements.
   b. Semantics: Checks that the created attributes, property and quantity sets are represented correctly. It also checks that each component is identified according to its semantics.
3. **Pragmatics.** Verifies that the IFC file works outside of the FME platform to ensure that the created IFC records can be easily used in other applications.
4. **Geometry representation.** verifies that the geometry in the IFC dataset is represented correctly and accurately.
5. **Georeferencing according to LoGeoRef.** Verifies that the georeferencing is correctly implemented according to the LoGeoRef concept.

− *How can different georeferencing methods be implemented programmatically within IFC using FME?*

The concepts of LoGeoRef 10, 20 and 50 are effectively implemented. The implementation of LoGeoRef 50 required a workaround using a custom transformer developed by Stijn Goedertier (2022) due to unimplemented FME attributes.

− *How can geospatial data be aggregated according to BIM conventions to make it usable for collaboration in BIM software?*

The geospatial data is aggregated and grouped according to the respective database and certain properties, such as the geometry and how it can be represented in IFC, the resulting IFC classes, and the generation of the corresponding attributes and values and their assignment to the corresponding IFC classes. This allows various other BIM software tools to access, retrieve and display this information correctly. This is demonstrated by importing and validating the data in several other software solutions such as Autodesk's Revit 2023, ESRI's ArcGIS Pro and the BIMcollab collaboration platform. All these aspects are successfully displayed, interpreted and mapped according to the software.

Finally, the developed FME workspaces are highly customizable, with extensible user parameters and configuration files, allowing for project-specific requirements. The ongoing development of of transformers and FME in the context of BIM improves processes and provides to work with future versions of IFC. Overall, the successful implementation demonstrates FME's ability to handle complex 2D geospatial and 2.5D terrain models and

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

seamlessly integrate them into the IFC data model. In this context, our research provides a comprehensive investigation and analysis of the integration of BIM and GIS with FME. However, in order to work with FME, a deep knowledge of both the software and the database is required.

REFERENCES

Borrmann, A., König, M., Koch, C. and Beetz, J. (eds) (2018) *Building Information Modeling : Technology Foundations and Industry Practice* [Online], Cham, Springer International Publishing. Available at https://doi.org/10.1007/978-3-319-92862-3 (Accessed 5 January 2024).

buildingSMART (2023) *buildingSMART Industry Foundation Classes* [Online]. Available at https://standards.buildingsmart.org/IFC/DEV/IFC4_2/FINAL/HTML/ (Accessed 9 September 2023).

Clemen, C. and Görne, H. (2019) 'Level of Georeferencing (LoGeoRef) using IFC for BIM', *Journal of Geodesy, Cartography and Cadastre*, vol. 10, no. 3, pp. 15–20 [Online]. Available at https://jgcc.geoprevi.ro/docs/2019/10/jgcc_2019_no10_3.pdf (Accessed 5 January 2024).

Clemen, C., Schröder, M., Kaiser, T. and Romanschek, E. (2021) 'IFCTERRAIN – A free and open source tool to convert digital terrain models (DTM) to openBIM Industry Foundation Classes (IFC)', *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, VIII-4/W2-2021, pp. 145–151 [Online]. DOI: 10.5194/isprs-annals-VIII-4-W2-2021-145-2021.

Donaubauer, A., Kutzner, T., Gruber, U., Borrmann, A. and Krause, K.-U. (2022) 'BIM und GIS-Integration – standardisierte, offene Datenformate', in DVW e.V. and Runder Tisch GIS e.V. (eds) *Leitfaden Geodäsie und BIM: Version 3.1* [Online], Bühl/München, pp. 115–131. Available at https://dvw.de/BIM-Leitfaden.pdf.

ISO (2021) *ISO/TR 23262: GIS (Geospatial) / BIM-Interoperabilität*.

Noardo, F., Krijnen, T., Arroyo Ohori, K., Biljecki, F., Ellul, C., Harrie, L., Eriksson, H., Polia, L., Salheb, N., Tauscher, H., van Liempt, J., Goerne, H., Hintz, D., Kaiser, T., Leoni, C., Warchol, A. and Stoter, J. (2021) 'Reference study of IFC software support: The GeoBIM benchmark 2019—Part I', *Transactions in GIS*, vol. 25, no. 2, pp. 805–841.

Stijn Goedertier (2022) *How to georeference a BIM model* [Online]. Available at https://github.com/stijngoedertier/georeference-ifc#conclusion (Accessed 5 January 2024).

Zhu, J., Wang, X., Wang, P., Wu, Z. and Kim, M. J. (2019) 'Integration of BIM and GIS: Geometry from IFC to shapefile using open-source technology', *Automation in Construction*, vol. 102, pp. 105–119 [Online]. DOI: 10.1016/j.autcon.2019.02.014.

BIOGRAPHICAL NOTES

Christoph Frey, born 1996. Completed an apprenticeship as a geomatics technician in 2018 and began studying for a bachelor's degree in geomatics at HTW Dresden in the same year.

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024

Graduated in 2022 as B. Eng. in Geoinformatics and in Oktober 2023 as M. Eng. in Geomatics/Management from the HTW Dresden. Since then he is a research assistant at the Faculty of Spatial Information at HTW Dresden.

Christian Dahmen is the team lead of con terra's Data Integration Product Services Team. He has been working for con terra since 2005 and has 18 years of experience in the geospatial industry. Christian loves all challenges around data integration, data management and data harmonization. He is a long-term FME technology enthusiast. As a member of OGC's CityGML Standard Working Group he supports the development and dissemination of open data standards. He studied Geoinformatics at the University of Muenster and holds a master's degree. Christian is married and has two children. He loves to run half marathons.

Christian Clemen, born 1976. Graduated in 2004 as a Dipl.-Ing. in Surveying from the Technical University of Berlin. From 2004 to 2010 he was an Assistant at the Department of Geodesy and Geoinformation, Technical University of Berlin. Since 2013 he has been a full professor at the HTW Dresden, University of Applied Sciences. He teaches adjustment calculation, CAD and BIM. His research topic is BIM from a geomatics perspective.

CONTACTS

Christoph Frey
Research Assistant, Faculty of Spatial Information
HTW Dresden, University of Applied Sciences
Friedrich-List-Platz 1
D-01069 Dresden, Germany
E-Mail: christoph.frey (at) htw-dresden.de

Christian Dahmen
con terra GmbH
Martin-Luther-King-Weg 20
D-48155 Münster, Germany
E-Mail: c.dahmen (at) conterra.de

Prof. Dr.-Ing. Christian Clemen
Surveying and BIM, Faculty of Spatial Information
HTW Dresden, University of Applied Sciences
Friedrich-List-Platz 1
D-01069 Dresden, Germany
E-Mail: christian.clemen (at) htw-dresden.de

ETL processes for the conversion of parcels (2D) and terrain models (2.5D) to openBIM projects using IFC (12626)
Christoph Frey, Christian Dahmen and Christian Clemen (Germany)

FIG Working Week 2024
Your World, Our World: Resilient Environment and Sustainable Resource Management for all
Accra, Ghana, 19–24 May 2024