



## ENHANCING BIM DATA INFORMATION EXCHANGE PROCESSES: A PULL-BASED FRAMEWORK FOR TARGETED DATA EXTRACTION AND VALIDATION

Timur Weilbach-Eyüboğlu, Cornelius Preidel, and Simon Vilgertshofer  
Hochschule München University of Applied Sciences, Munich, Germany

### Abstract

Traditional *push-based* workflows in Building Information Modeling (BIM) often suffer from inefficiencies that hinder precise and effective data exchange. The following study introduces a *pull-based* approach designed to address challenges, induced by inefficient *push-based* workflows, by focusing on targeted information requests and automating validation processes. The proposed method improves data quality, ensures compliance, and streamlines workflows by minimizing redundancies, reducing coordination cycles, and prioritizing relevant data. The *pull-based* approach enhances digital data processes and improves stakeholder communication, providing a practical and efficient alternative to the currently dominant *push-based* workflows while unlocking new opportunities for BIM adoption.

### Introduction

The exchange of information in the construction industry, particularly within the Building Information Modeling (BIM) framework, is notably complex due to the highly fragmented process chain, the large number of stakeholders involved, and the growing volumes of data (Sacks et al., 2018). BIM introduces a model-based approach to streamline these processes, combining disciplines with widely varying information requirements and generating immense data streams – ranging from construction details and technical specifications to construction site reports. Furthermore, these data streams are often maintained and exchanged in diverse formats and structures, adding another layer of complexity to the process (Sacks et al., 2018). Ensuring a smooth and targeted flow of information remains a key challenge (Leite et al., 2020).

Standards such as DIN EN ISO 19650 aim to overcome the difficulties associated with information exchange by specifying clearly defined processes and roles: The information requester (requester) formulates the information requirements, for example, project-specific needs, technical specifications, or compliance-related details, while the information provider (provider) delivers the corresponding information in the required format and structure (DIN, 2019).

In practice, however, defining information specifications in advance can be challenging, particularly in the early

stages of a project (Sacks et al., 2018). Not all details may be finalized yet; specific information depends on various influencing factors (Sacks et al., 2018). For instance, the construction method for walls – whether prefabricated or built on-site – may change throughout a construction project, leading to different information needs at different stages.

Similarly, in the early planning phase, information exchange presents its challenges. The architect might provide a shell model of a building to the structural engineer, containing essential information for the dimensioning process, such as fire resistance classes, materials, and load-bearing capacities of elements. However, issues arise when dealing with *vertical interoperability*, which refers to the ability to exchange and integrate information between different types of models – such as an architectural model created with Computer-Aided Design (CAD) software and a structural engineering model created with Computer-Aided Engineering (CAE) tools (Chiaia et al., 2015). Despite the intended interoperability of Industry Foundation Classes (IFC) as a data model designed to enable information exchange across software platforms, its effectiveness heavily depends on the quality and structure of the data provided to the model. While IFC provides a standardized format for data exchange, compatibility issues still often arise when trying to import or interpret data across different platforms (Gerbino et al., 2021). Therefore, teams working with architectural models as a baseline for structural engineering planning often encounter difficulties. These compatibility issues force teams to revert to traditional methods, such as permit drawings or submission plans, to extract the necessary information, highlighting a key challenge in current model-based approaches (Sacks et al., 2018): Although digital models contain valuable data, their practical application is still hindered by issues in sharing data across platforms (Chong et al., 2020). These challenges underscore a flaw in the traditional information exchange model, as data may be available but remains difficult to access or use in downstream processes, leading to inefficiencies and delays. Accordingly, the provider may deliver excessive or irrelevant data. Insufficient clarity regarding specific requirements or uncertainty about the precise data needed to meet the objectives often contributes to this issue (Aksenova et al.,

2019). As a result, stakeholders tend to share everything they have or consider relevant, aiming to cover different scenarios and potential future requirements, or to compensate for collaboration issues (Matarneh et al., 2019). However, adopting such a practice can have adverse effects, such as important information being misunderstood or overlooked due to the overwhelming amount of data (Piroozfar et al., 2019).

The described *push-based* approach – where information is delivered based on contractually defined specifications in advance (Sacks et al., 2018) – requires subsequent data validation to ensure that the content is correct and the quality is high (Cavka et al., 2017). Issues and inconsistencies are frequently identified only after delivering the required information, mainly when the data is being utilized and its inadequate quality becomes apparent (Sacks et al., 2018). Project timelines are affected when exchanged information must be revisited and adjusted due to insufficient or flawed data (Sacks et al., 2018). The essence of digital processes – streamlining workflows and improving overall efficiency – is hereby undermined. Instead, the reliance on repeated coordination cycles introduces avoidable complexity, thereby significantly impairing the effectiveness of the decision-making process (Dave et al., 2018).

At this point, the principle of the *pull-based* approach, which originates from Lean Construction Management (LCM) (GLCI, 2019), provides a promising alternative to the traditional *push-based* approach. LCM aims to reduce inefficiencies, minimize waste, and make processes more focused and efficient. It also improves the flow of information, ensuring every effort contributes directly to value creation.

Unlike the *push-based* approach, which relies on forecasts and often leads to overproduction or unnecessary inventories – for example, ordering more concrete than required due to outdated quantity data, resulting in excess concrete that must ultimately be discarded – the *pull-based* approach is demand-oriented (Koskela et al., 2010). Information is provided only when needed, ensuring the focus remains on the most relevant and actionable data. While the *pull* principle was initially designed for physical production processes, its application can also effectively optimize model-based information processes (Sacks et al., 2018). The *pull* approach enhances efficiency and keeps processes lean and purpose-driven by tailoring information flow to actual needs. In this research, we aim to explore whether shifting from *push-based* workflows to a *pull-based* approach could offer advantages, like improving the quality of the exchanged data, reducing the amount of data to the essentials, simplifying processes, and minimizing unnecessary iterations. An implementation of the *pull-based* approach has not yet been investigated in depth for information exchange in the early planning stages. The aim is to experimentally develop and test a functional *pull-based* workflow for exchanging information in the BIM environment. The focus is not on proving superiority but critically assessing whether this approach could present a

viable alternative to established workflows in construction projects' planning and design phases.

The main contributions of this work are:

- Defining the requirements and key conditions for *pull-based* information flows to ensure that only relevant information is provided exactly when needed, avoiding unnecessary data exchange.
- Development of a conceptual framework that extends and adapts existing methods for specifying information requirements to enable *pull-based* requests in the BIM environment.
- Demonstration of a prototypical workflow that showcases a *pull-based* approach illustrating its potential to effectively improve information exchange processes.

## Background on BIM Data Exchange

Effective information exchange is fundamental to collaboration (Sacks et al., 2018). Optimizing the processes related to this exchange enhances data flows, making them more efficient and adaptable to diverse project needs. Providing stakeholders with the specific data required for analysis, decision-making, or project execution supports data interoperability, ensures model accuracy (Sacks et al., 2018), and facilitates seamless project coordination. Numerous data exchange and information extraction approaches have been developed and proposed as BIM workflows evolve over the past decades. The ongoing development of new methods highlights that the question of how stakeholders can effectively access the information they need remains an open challenge. (Olawumi and Chan, 2019)

Integrating BIM processes across architectural, structural, and detailing workflows faces key challenges, including data interoperability and communication gaps. These barriers hinder seamless integration and highlight the need for standardized data exchange methods to optimize BIM workflows and facilitate more effective information extraction (Chong et al., 2020).

Cavka et al. (2017) emphasize the importance of standardized approaches for defining Owner Information Requirements (OIR) to facilitate improved data exchange throughout the project lifecycle. Highlighting how inefficiencies in information extraction often arise from a lack of standardization, their work demonstrates how a structured OIR framework enables stakeholders to extract relevant data based on specific project needs, ensuring optimal information flow.

One of the most common data transfer practices is the file-based exchange, which operates as a one-way transfer of information, necessitating repeated file transfers for each design iteration to account for changes. Managing data inconsistencies and redundancy becomes challenging with this method. Object-level data management is also limited, as access control is only possible at the file level, not for specific objects. Furthermore, the increasing number of

Model View Definitions (MVDs) – used to extract subsets of data from IFC files – presents a challenge for software vendors to implement. (Lou et al., 2021)

Recognizing the challenges associated with file-based exchanges, lack of object-level access control and difficulties in managing MVDs, buildingSMART is leading modernization efforts for IFC 5, as discussed by van Berlo et al. (2021), aiming to enhance interoperability and data exchange. The modularization and improved semantic expressiveness of IFC 5 support a shift toward object-based data exchange, which allows for more granular, context-specific updates. These advancements enable more efficient and flexible BIM integration, facilitating incremental data loads and updates, which are crucial for optimizing information extraction. Although these developments are promising, they are still in progress.

Lee et al. (2020) propose an entity-based integration approach to improve BIM data exchange standards. Their work aids in generating MVD concept modules, which enhance interoperability and data exchange accuracy. However, since MVDs are primarily developed for software certification rather than to directly address Information Requirements, their implementation presents additional challenges. Its complex structure allows the definition and constraining of concepts but limits broader adoption due to bespoke implementation requirements, as noted in the review from Tomczak et al. (2022).

Information Delivery Specifications (IDS) have emerged as a tool for automating the specification of Information Requirements, particularly in defining alphanumeric data. IDS facilitates the authoring of Information Requirements by providing predefined templates for required BIM data, focusing on entities and properties defined within the IFC standard (buildingSMART International, 2025). However, IDS is restricted to openBIM-compliant tools and does not support geometric or process mapping, limiting its application (Tomczak et al., 2022). It primarily supports the traditional (*push-based*) process, in which 3D models are created and subsequently checked.

Building on prior efforts to enhance BIM data retrieval, Zhang and El-Gohary (Zhang and El-Gohary, 2015) propose a method that employs JSDAI for structured extraction of IFC-based BIM data. By utilizing late binding, their approach ensures compatibility across multiple IFC schema versions, allowing entities and attributes to be accessed dynamically. The extracted data is then processed according to metadata schema and transformed into structured logic facts using semantic NLP techniques. This transformation facilitates semantic alignment and reasoning, addressing challenges in structuring IFC data for automated interpretation.

Semantic web technologies have introduced diverse approaches to optimizing BIM information extraction. A key strategy involves leveraging formal mechanisms, such as query languages, to define precisely what data should be extracted from a model. By working with targeted subsets of a model instead of its entirety, workflows are stream-

lined, and the usability of the extracted information is improved. Within this context, Oraskari et al. (2021) explore how the Shapes Constraint Language (SHACL) operates within Linked Building Data (LBD) frameworks. By defining and validating constraints, SHACL functions similarly to mvdXML in the IFC schema, ensuring that only relevant and validated data is exchanged. The ability to enforce data quality and compliance significantly enhances information exchange processes. Nevertheless, several prerequisites remain. The data rarely exists in a form directly processable by SHACL and often needs to be converted first, which is a time-consuming process that introduces the potential for errors. Furthermore, SHACL is inherently complex and likely to remain a tool primarily suited for expert users.

Yin et al. (2023) present an innovative model-based ontology population and semantic parsing (MOP-SP) method that advances BIM data retrieval. The approach integrates a tailored IFC Natural Language Expression ontology with automated extraction of project-specific terms from BIM models. By converting natural language queries into SPARQL commands via sophisticated NLP techniques, it enables precise multi-constraint data extraction, streamlining workflows and enhancing the accuracy of targeted information retrieval.

BIMQL (Building Information Model Query Language) is an open, domain-specific query language designed for IFC-based building information models. It enables users to flexibly extract, modify, and interact with model data, supporting ad hoc queries tailored to practical information requirements. BIMQL simplifies information retrieval through natural language-like syntax and domain-specific shortcuts, such as property-based selection and hierarchical filtering. It also supports schema-level and instance-level queries, making it suitable for compliance checks, data transformation, and model simplifications. (Mazairac and Beetz, 2013)

Developments in natural language processing (NLP) and machine learning are transforming information extraction in BIM. Tang et al. (2022) introduce an NLP framework for quantity take-off (QTO) automation, utilizing Named Entity Recognition models to identify key construction elements from work descriptions. Their approach reduces manual effort and improves the accuracy of extracted data.

Generative AI technologies, such as BIM-GPT, as proposed by Zheng and Fischer (2023), further revolutionize information extraction. Using prompt-based frameworks, BIM-GPT interprets natural language queries to generate precise answers from BIM datasets, enhancing accessibility and reducing the need for specialized training.

We aim to empower requesters to play a more active and central role in the information exchange process, allowing them to directly influence the flow of information and ensure the data is both high-quality and relevant to their specific needs.



Figure 1: Comparison Push-based Approach (left) and Pull-based Approach (right)

## Methodology

Our methodology follows a structured *pull-based* process, contrasting with the traditional *push-based* approach, in which information is provided proactively, often without consideration for the recipient’s specific needs, leading to subsequent coordination loops. In comparison, a *pull-based* process allows the requester to actively define and request the required information based on clearly articulated needs and hereby ensures a more targeted and efficient exchange of information. Figure 1 illustrates the conceptual framework for our *pull-based* approach leveraging IDS in contrast to a traditional *push-based* approach.

The IDS schema is designed to streamline BIM workflows by defining specific information requirements for deliverables in a structured and interoperable manner. It specifies how objects, classifications, properties, values, and units should be delivered and exchanged using IFC. IDS employs *Applicability* to define the scope of a specification, such as relevant IFC classes, predefined types, or properties, while *Requirements* outline the actual information constraints, including expected properties, values, and formats. Code 1 showcases a section of an IDS-File, in which *Applicability* targets a specific element identified by its Globally Unique Identifier (GUID). The *Requirements* define that this element must contain a PropertySet named *ARC*, which must include a Property called *IsExternal*.

Code 1: IDS file section

```

<ids ...>
  <specification name="IDS by using GUID">
    <applicability>
      <attribute>
        <name>
          <simpleValue>GlobalId</simpleValue>
        </name>
        <value>
          <simpleValue>FdWA0oMU</simpleValue>
        </value>
      </attribute>
    </applicability>
    <requirements>
      <property>
        <propertySet>
          <simpleValue>ARC</simpleValue>
        </propertySet>
        <baseName>
          <value>IsExternal</value>
        </baseName>
      </property>
    </requirements>
  </specification>
</ids>
  
```

```

    </property>
  </requirements>
</specification>
</ids>
  
```

IDS enables the automated validation of digital models by expressing information requirements in machine-readable formats, such as XML (as shown in Code 1), thereby reducing misinterpretations and rework. BIM validation tools parse the IDS file to cross-check the IFC model, ensuring that the specified elements meet predefined rules and context-dependent criteria for information exchange. Clear expectations among project members are established, enhancing collaboration and ensuring that essential data is accurately captured.

Beyond validation, IDS can also be a key input for the modeling process. Model authors can use IDS as a template to identify key placeholders representing information requirements and populate them with the necessary data, providing clear guidance on which components still lack specific information to meet the requirements. Alternatively, IDS can function as a mapping tool, ensuring that existing data in the model or authoring software correctly aligns with the appropriate locations in the IFC model.

Although IDS was not initially designed for information requests, it can enable a *pull-based* workflow, allowing the requester to articulate their information requirements and act as a digital transport medium for the requested information. Validating the requested data content is a valuable side effect of using IDS, ensuring the quality of the information.

Our approach leverages IDS slightly differently as a dynamic tool to bridge the gap between the requester and the provider, enabling the targeted extraction, transfer, and integration of information. Instead of validating the information delivery, we *pull* specific data from the provider by locating the information with IDS based on predefined criteria. Once located, the requested information is extracted and postprocessed to be implemented in a new IDS file – the *Answer-IDS*, which functions as a transport mechanism for capturing and delivering the information to the requester.

Upon receipt, the *Answer-IDS* is used to implement the pulled information into the requester’s digital model, ensuring that relevant, structured, and validated data is seamlessly incorporated into the requester’s model, maintaining consistency and enhancing interoperability. Differ-

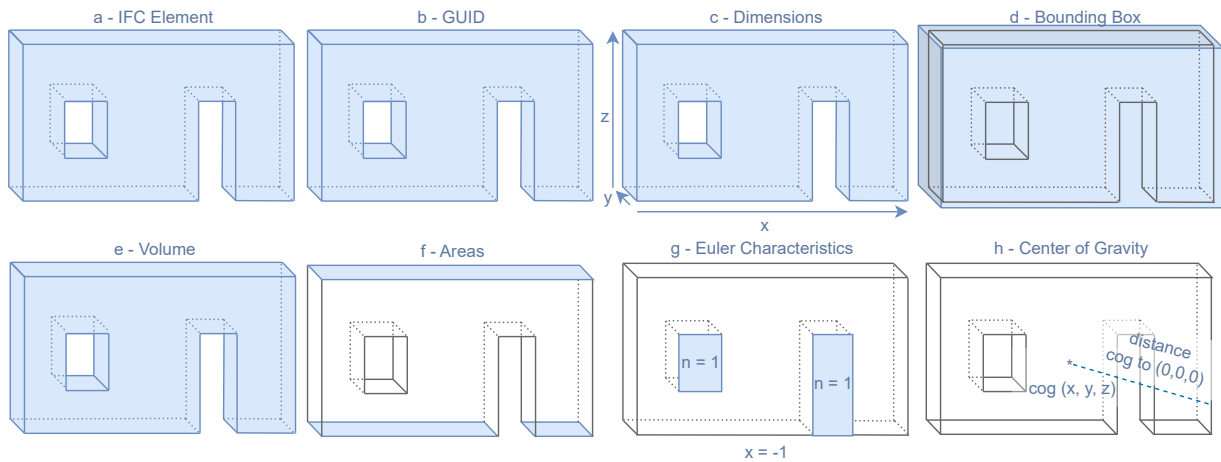


Figure 2: Key Geometric Features

ing GUIDs of corresponding objects for different domain models challenge the request and transport mechanism. Precise alignment between the requester and provider models is essential to enabling information transfer, necessitating a robust mapping mechanism.

Using the IDS schema in our approach offers several advantages:

- The schema is user-friendly and allows precise targeting of specific elements, enabling users to define their information requirements in a machine-readable format. This facilitates seamless integration into automated digital workflows.
- Various IDS editors are available, reflecting widespread adoption and adaptability. The schema's straightforward logic makes it accessible for users across different tools, promoting flexibility and reducing learning curves.
- IDS files can function as validation tools and transport containers, facilitating the extraction and exchange of targeted information between requesters and providers while ensuring compliance with pre-defined criteria.

## Implementation

To implement the methodology, we utilize a command-line-based Python script for geometry feature extraction, element mapping, IDS creation, and validation. The developed software framework processes IFC files to align requester and provider models, enabling targeted information exchange using a *pull-based* approach.

### Geometry Feature Extraction

As outlined in the Methodology section, effectively addressing the challenge of aligning correlated elements with different GUIDs between the requester and provider domain models requires a detailed understanding of the respective elements' geometric and topological characteristics. Geometry feature extraction provides a foundational basis for the mapping mechanism that facilitates the com-

parison of elements across IFC files, thereby supporting the alignment process.

By analyzing IFC models, key geometric features that uniquely define each building element are derived. Focusing on these features can bridge discrepancies between domain models, ensuring accurate matching and reliable information exchange. The workflow is built on advanced parsing and processing techniques, leveraging the *ifcopenshell* and *Open Cascade Technology* libraries to extract the necessary data. Extracted features include key geometric properties and semantic information, as shown in Figure 2. For each element in the requester and provider models, the extracted features are computed to ensure consistency and accuracy for downstream processing. The feature values are stored in a CSV file, forming the basis for subsequent analyses and element mapping. To mitigate potential misclassifications, such as when Element A in Model A is classified as an *IfcWall* but as an *IfcBeam* in Model B due to geometric thresholds, we rely exclusively on geometry for matching, ignoring semantic discrepancies.

### Element Mapping Process

By analyzing geometric features, the element mapping process establishes correspondences between building elements in the requester and provider models. While GUIDs are essential for IDS validation, they are excluded from the feature dataset used for geometry-based mapping to focus purely on spatial characteristics.

To ensure consistency, feature values are normalized using the *MinMaxScaler* from *sklearn.preprocessing*, scaling all features to a standard range. Pairwise Euclidean distances are computed to identify each requester element's closest matching component of the provider dataset. The mapping process captures both datasets' IFC element types and GUIDs, assembling the results into a structured data frame, as exemplified in Table 1. This dataset enables further analysis and integration into workflows requiring element correspondence.

Matching elements based on their geometric features, ensures a reliable mapping mechanism. However, some chal-

Table 1: Exemplary Element Mapping Information Requester - Information Provider

IFC-Element (IR)	GUID (IR)	IFC-Element (IP)	GUID (IP)
IfcBeam	0kn..ONQ	IfcBeam	3V8..AQC
IfcWall	0JN..s6k	IfcWall	20i..Byj
IfcDoor	33f..CNk	IfcDoor	3Sp..y7X

Challenges persist, mainly due to differences in modeling practices. For instance, in one model, a wall may be represented as a single multi-layered *IfcWall*. In contrast, in another model, the same wall could be broken into several separate components, such as individual *IfcWall* and *IfcCovering* elements. Discrepancies in the decomposition or grouping of elements can create difficulties in achieving a one-to-one mapping.

### Generating the Pull-Request

To define the content of a pull request, the user specifies different element types, such as *IfcWall* or *IfcSlab*, which apply to all *IfcBuildingElements*. Additionally, the user provides details about desired properties, such as *Fire Resistance Rating (FRR)*, *Material*, or *Load Bearing*. Multiple properties can be assigned to each element, creating a structured set of specifications.

After the requester defines the specifications, the dataset is further refined, as illustrated in Table 2.

Table 2: Exemplary Element Mapping with Appended Property

IR-Element	IR-GUID	IP-Element	IP-GUID	IR-Property
IfcWall	3p..fP	IfcWall	0\$..LN	FRR
IfcWall	3p..fP	IfcWall	0\$..LN	Material
IfcWall	3p..fP	IfcWall	0\$..LN	Load Bearing

The next step generates the IDS for the *pull* request. Leveraging the requester’s input, it defines the two key components of the IDS: *Applicability* and *Requirements*. *Applicability* specifies the elements for which information is to be extracted (e.g., *IfcWall*), while *Requirements* outline the specific properties of interest (e.g., *FRR*, *Material*, *Load Bearing*).

To implement the *pull-based* request, the requester’s specified element type (*IfcWall*) is used to filter the mapped dataset. From this filtered dataset, the provider’s corresponding GUIDs are identified and used as targets for information extraction. Figure 3 illustrates how the IDS is constructed to target specific elements within the provider model based on the requester’s input.

The provider receives a *pull* request from the requester in the form of the previously created IDS file. The developed software framework analyzes the provider’s model,

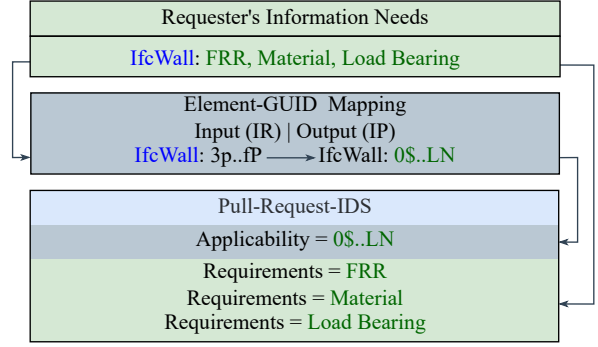


Figure 3: Pull Request IDS for information extraction

validating it against the criteria outlined in the IDS file. Elements that meet these criteria have their associated data—such as *IfcEntity*, *GUID*, *Property*, *Value*, and *Data Type*—extracted. The provider’s approval is required to extract the requested information. Once approved, the extracted data, as shown in Table 3, is stored for further processing.

Table 3: Extracted information from provider model

Element	GUID	Property	Value	Data Type
IfcWall	0\$..LN	FRR	R30	IfcText
IfcWall	0\$..LN	Material	Concrete	IfcText
IfcWall	0\$..LN	Load Bearing	Yes	IfcBoolean

### Information Transfer

A key challenge when pulling information is the transport mechanism. As detailed in the Methodology section, we use IDS as a bridge to transfer the extracted information from the provider back to the requester.

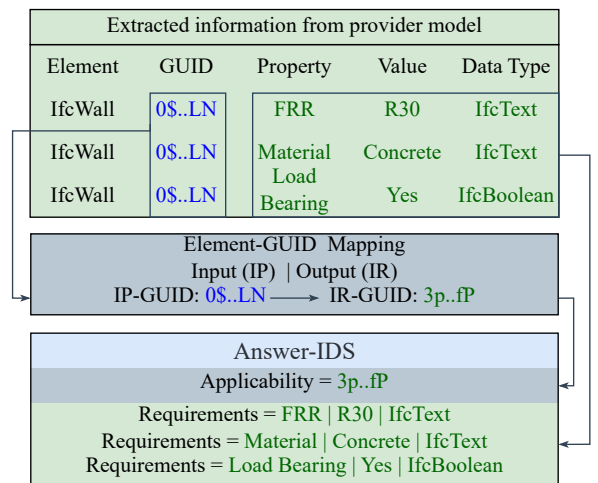


Figure 4: Answer-IDS for information transport

After the provider approves the information extraction, the *pull* request fires, generating a new IDS file—the *Answer-IDS*. Prior to this, the GUID mapping is reversed, con-

verting the provider GUIDs back to the requester GUIDs, which are then used to identify the applicable target elements within the requester model. A concrete value and a data type supplement the requirements. Figure 4 provides a visual representation of the *Answer-IDS* generation process.

The generated *Answer-IDS* file is subsequently processed within the developed software framework, where it is used to validate the requester's model. Any failed validations indicate missing information, which is then integrated into the requester's IFC model using the *ifcopenshell* library. A new property set (PSET) containing the extracted properties, values, and data types is created within the model.

By automating data extraction and validation, the *pull-based* workflow reduces manual coordination, streamlines processing, and ensures precise, project-specific information exchange.

## Discussion and Conclusion

The present study illustrates how a *pull-based* workflow enhances BIM data exchange by reimagining the role of IDS, enabling precise and efficient information extraction and transportation triggered by a *pull request*. While IDS has primarily been utilized for compliance checking, the presented framework extends its functionality to support a demand-driven workflow, reshaping the methods of information request and delivery within BIM processes.

By incorporating principles from LCM, the *pull-based* workflow represents a significant improvement over traditional *push-based* systems, addressing persistent inefficiencies in data exchange. Its implementation has shown notable successes, including enhanced precision and compliance, reduced iterative rework, improved stakeholder collaboration, and increased data usability within typical BIM workflows. Compared to traditional *push-based* workflows, demand-driven data exchange in BIM offers significant advantages, extending the application of IDS beyond compliance checking to support interactive and role-specific data augmentation. Prioritizing targeted data requests over comprehensive submissions minimizes redundant interactions and enhances clarity in communication between data requesters and providers. A particularly impactful outcome is the reduction of coordination loops and iterative revisions, which are prevalent in *push-based* workflows due to unclear data requirements. As a result, stakeholders can allocate more time to value-adding tasks by minimizing repetitive or non-essential data handling.

IDS demonstrates flexibility and the potential to overcome persistent challenges in BIM information exchange. The improvements observed in precision, compliance, and stakeholder collaboration suggest that the *pull-based* workflow holds considerable potential for adoption in projects prioritizing accuracy and efficiency. This study investigates opportunities for more efficient, user-focused workflows in the construction industry. The goal is to contribute to the ongoing development of BIM practices and provide insights that may inform future advancements.

Despite these promising results, several challenges and limitations remain. A notable challenge is harmonizing object modeling standards across different domain models. As mentioned, this is exemplary with multi-layer walls versus separated components.

Adopting standardized modeling guidelines can help mitigate these discrepancies but may not eliminate the risk of errors, particularly in complex projects with diverse stakeholders and software ecosystems.

Another limitation is IDS's inability to specify geometric data, which restricts its application for tasks reliant on spatial information. While the *pull-based* workflow performed effectively for well-defined tasks, its scalability to larger, more complex projects with numerous interdependencies has yet to be fully assessed.

## Outlook

Future developments in *pull-based* BIM data exchange processes hold significant promise for addressing current limitations and advancing the methodology. Key areas for refinement include enhancing scalability, improving IDS usability for exploring ways to integrate geometric data into the *pull-based* framework and integrating multiple different but dependent information sources. Resolving issues related to element and GUID mapping inconsistencies, especially in environments with varying modeling guidelines, will be vital.

Leveraging machine learning and automation advancements may also help scale these workflows to accommodate more intricate BIM processes. Knowledge graphs represent a promising solution, offering the potential to prepare data for AI-driven mapping, establish intelligent connections across providers, and enable predictive, computer-assisted information requests. By leveraging these advancements, the *pull-based* BIM data exchange process can be further optimized to deliver higher accuracy, adaptability, and efficiency in diverse and complex scenarios.

## Data availability statement

The data supporting this study's findings are available at [https://gitlab.lrz.de/tweilbach/ids\\_pull](https://gitlab.lrz.de/tweilbach/ids_pull), ensuring that they are Findable, Accessible, Interoperable, and Reusable (FAIR).

## Acknowledgments

The authors gratefully acknowledge the support and resources provided by the Bavarian Construction Industry Association (Bayerischer Bauindustrieverband), enabling our research and its practical application.

## References

- Aksenova, G., Kiviniemi, A., Kocaturk, T., and Lejeune, A. (2019). From BIM to collaborative working environment: A project management perspective. *Automation in Construction*, 108:102968.
- buildingSMART International (2025). IDS: Informa-

- tion Delivery Specification. <https://github.com/buildingSMART/IDS>. Accessed: 2025-01-07.
- Cavka, H. B., Staub-French, S., and Poirier, E. A. (2017). Developing owner information requirements for BIM-enabled project delivery and asset management. *Automation in Construction*, 83:169–183.
- Chiaia, B., Davardoust, S., Osello, A., et al. (2015). Structural calculations by horizontal and vertical interoperability for the redevelopment of existing buildings. In *Heritage and Technology. Mind Knowledge Experience*, pages 650–658. La Scuola di Pitagora.
- Chong, O. W., Baker, C., Afsari, K., Zhang, J., and Roach, M. (2020). Integration of BIM processes in architectural design, structural analysis, and detailing: current status and limitations. In *Construction Research Congress 2020*, pages 1203–1212. American Society of Civil Engineers Reston, VA.
- Dave, B., Kubler, S., Främling, K., and Koskela, L. (2018). Opportunities for enhanced lean construction management using Internet of Things standards. *Automation in Construction*, 93:35–45.
- DIN (2019). EN ISO 19650-2:2018 – Organization and digitization of information about buildings and civil engineering works, including BIM – Information management using BIM – Part 2: Delivery phase of the assets. DIN Standard. German version.
- Gerbino, S., Cieri, L., Rainieri, C., and Fabbrocino, G. (2021). On BIM Interoperability via the IFC Standard: An Assessment from the Structural Engineering and Design Viewpoint. *Applied Sciences*, 11(23).
- GLCI (2019). *Lean Construction: Begriffe und Methoden*. German Lean Construction Institute – GLCI e.V., Karlsruhe, Germany. 1st Edition, October 2019.
- Koskela, L., Sacks, R., and Dave, B. (2010). Interaction of Lean and Building Information Modeling in Construction. *Journal of Construction Engineering and Management*, 136(9):968–980.
- Lee, Y.-C., Shariatfar, M., Ghannad, P., Zhang, J., and Lee, J.-K. (2020). Generation of entity-based integrated model view definition modules for the development of new BIM data exchange standards. *Journal of Computing in Civil Engineering*, 34(3):04020011.
- Leite, F., Akcamete, A., Akinci, B., Atasoy, G., and Kiziltas, S. (2020). Information Integration Challenges in Building Information Modeling: Updated Perspectives from Project Teams. *Journal of Construction Engineering and Management*, 146(2):100–115.
- Lou, J., Lu, W., and Xue, F. (2021). A review of BIM data exchange method in BIM collaboration. In *Proceedings of the 25th International Symposium on Advancement of Construction Management and Real Estate*, pages 1329–1338. Springer.
- Matarneh, S. T., Danso-Amoako, M., Al-Bizri, S., Gaterell, M., and Matarneh, R. (2019). Building information modeling for facilities management: A literature review and future research directions. *Journal of Building Engineering*, 24:100755.
- Mazairac, W. and Beetz, J. (2013). BIMQL—An open query language for building information models. *Advanced Engineering Informatics*, 27(4):444–456.
- Olawumi, T. O. and Chan, D. W. (2019). Building information modelling and project information management framework for construction projects. *Journal of Civil Engineering and Management*, 25(1):53–75.
- Oraskari, J., Senthilvel, M., and Beetz, J. (2021). SHACL is for LBD what mvdXML is for IFC. In *Proc. of the Conference CIB W*, volume 78, pages 11–15.
- Piroozfar, P., Farr, E. R., Zadeh, A. H., Inacio, S. T., Kilgallon, S., and Jin, R. (2019). Facilitating Building Information Modelling (BIM) using Integrated Project Delivery (IPD): A UK perspective. *Journal of Building Engineering*, 26:100907.
- Sacks, R., Eastman, C., Lee, G., and Teicholz, P. (2018). *BIM handbook: A guide to building information modeling for owners, designers, engineers, contractors, and facility managers*. John Wiley & Sons.
- Tang, S., Liu, H., Almatared, M., Abudayyeh, O., Lei, Z., and Fong, A. (2022). Towards automated construction quantity take-off: An integrated approach to information extraction from work descriptions. *Buildings*, 12(3):354.
- Tomczak, A., Berlo, L. v., Krijnen, T., Borrmann, A., and Bolpagni, M. (2022). A review of methods to specify information requirements in digital construction projects. *IOP Conference Series: Earth and Environmental Science*, 1101(9):092024.
- van Berlo, L., Krijnen, T., Tauscher, H., Liebich, T., Van Kranenburg, A., and Paasiala, P. (2021). Future of the industry foundation classes: towards IFC 5. In *Proc. of the 38th International Conference of CIB W*, volume 78, pages 11–15.
- Yin, M., Tang, L., Webster, C., Xu, S., Li, X., and Ying, H. (2023). An ontology-aided, natural language-based approach for multi-constraint BIM model querying. *Journal of Building Engineering*, 76:107066.
- Zhang, J. and El-Gohary, N. (2015). Automated extraction of information from building information models into a semantic logic-based representation. In *Computing in civil engineering 2015*, pages 173–180.
- Zheng, J. and Fischer, M. (2023). BIM-GPT: a prompt-based virtual Assistant framework for BIM information retrieval. *arXiv preprint arXiv:2304.09333*.