Technical University of Munich
TUM School of Engineering and Design

# BIM-based semantic enrichment for environmental analyses using Large Language Models

**Kasimir Forth**

Vollständiger Abdruck der von der TUM School of Engineering and Design der Technischen Universität München zur Erlangung eines

**Doktors der Ingenieurwissenschaften (Dr.-Ing.)**

genehmigten Dissertation.

Vorsitz: Prof. Dr.-Ing. Frank Petzold

Prüfende der Dissertation:

1. Prof. Dr.-Ing. André Borrmann
2. Prof. Dr. Jakob Beetz
3. Associate Prof. Dr.-Ing. Patricia Schneider-Marin

Die Dissertation wurde am 12.08.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Engineering and Design am 19.11.2024 angenommen.

## Abstract

The AEC industry and the operation of buildings are responsible for approximately a third of global greenhouse gas emissions. There is also a very high demand for resources and, at the same time, a very high volume of waste. In order to meet the current ecological challenges of the construction industry, environmental analyses are an established approach in the early design stages. These include life cycle assessments (LCA), material passports (MP), and building energy performance simulations (BEPS), all of which are applied as different use cases in this dissertation.
Building Information Modeling (BIM) is a digital designing and planning method that can be used for carrying out such environmental analyses using semantically rich geometric models of building designs. These BIM models are used as data sources to derive analysis results without manual remodeling. A distinction is made between open and closed BIM workflows. Open BIM differs from closed BIM by the use of manufacturer-neutral data formats, including Industry Foundation Classes (IFC) for geometric and semantic model information.
The main contribution of this dissertation is the (semi-)automated semantic enrichment of open BIM models for environmental analyses using Natural Language Processing (NLP) and Large Language Models (LLM). The semantic enrichment approach consists of matching the semantically most similar data from a database to the items from the IFC model and adding the missing information for the respective environmental analysis. The degree of automation of this matching differs depending on the use case due to differently structured databases and further domain-specific training of language models, the so-called LLM fine-tuning. Depending on the use case, pre-trained LLMs (LCA), monolingual fine-tuned LLMs (MP), or multilingual fine-tuned LLMs (BEPS) are used, whereby different strategies are combined to increase the matching accuracy. The three environmental analyses were tested using several real-world case studies and models, trained with their semantic information, and evaluated.
The last chapter presents a decision-making approach regarding element and material selection using open BIM data formats, including IFC and the BIM collaboration format (BCF), as well as different visualization strategies. The results of the embodied greenhouse gas emissions are visualized together with the uncertainties of the early design stages. This method was tested and evaluated with experts and stakeholders without LCA expertise using a case study to enable reliable decision-making based on various visualizations.

## Zusammenfassung

Das Bauwesen und der Gebäudebetrieb sind für gut ein Drittel der globalen Treibhausgasemissionen verantwortlich. Außerdem besteht ein sehr hoher Ressourcenbedarf bei gleichzeitig sehr hohem Abfallaufkommen. Um den aktuellen ökologischen Herausforderungen der Bauwirtschaft zu begegnen, sind Nachhaltigkeitsanalysen bereits in frühen Entwurfsphasen ein bewährter Ansatz. Dazu gehören Ökobilanzierungen (LCA), Materialpässe (MP) sowie Gebäudeenergiesimulationen (BEPS), welche in dieser Dissertation als verschiedene Anwendungsfälle angewandt werden.
Building Information Modeling (BIM) ist eine digitale Arbeitsmethode, um mithilfe semantisch reicher Geometriemodellen von Gebäuden solche Nachhaltigkeitsanalysen durchzuführen. Dabei werden diese BIM Modelle als Datengrundlage verwendet, um Nachhaltigkeitsanalysen ohne manuelle Nachmodellierung abzuleiten. Man unterscheidet dabei zwischen offenen und geschlossenen BIM Arbeitsabläufen. "Open BIM" unterscheidet sich von "closed BIM" durch die Nutzung hersteller-neutraler Datenformate, unter anderem Industry Foundation Classes (IFC) für Modellinformationen.
Der wesentliche Beitrag dieser Dissertation liegt auf der (halb-)automatisierten semantischen Anreicherung von open BIM Modellen für Nachhaltigkeitsanalysen mithilfe von Natural Language Processing (NLP) und künstlichen Sprachmodellen, auch als Large Language Models (LLM) bekannt. Die semantische Anreicherung besteht aus der Zuordnung, im Englischen auch matching genannt, der semantisch ähnlichsten Datensätze aus einer Datenbank zu dem Typen aus dem IFC-Modell sowie der Ergänzung der fehlenden Informationen für die jeweilige Nachhaltigkeitsanalyse Der Automatisierungsgrad dieser Zuordnung unterscheidet sich je nach Anwendungsfall aufgrund verschieden strukturierter Datenbanken und weitertrainierten Sprachmodellen, dem sogenannten LLM fine-tuning. Je nach Anwendungsfall werden vortrainierte LLM (LCA), monolingual weitertrainierte LLM (MP) oder multilingual weitertrainierte LLM (BEPS) verwendet, wobei verschiedene Strategien zur Erhöhung der Zuordnungsgenauigkeit kombiniert werden. Die drei Nachhaltigkeitsanalysen wurden anhand von verschiedenen realen Beispielprojekten und -modellen getestet, mit deren semantischen Informationen trainiert und die Methode ausgewertet wurde.
Im letzten Kapitel wird eine Methode zur Entscheidungsfindung von Konstruktions- und Materialauswahl mithilfe von open BIM Datenformaten, darunter IFC und BIM Collaboration Format (BCF), sowie verschiedenen Visualisierungsstrategien vorgestellt. Dabei werden die Ergebnisse der gebundenen oder grauen Treibhausgasemissionen gemeinsam mit den Unsicherheiten der frühen Entwurfsphasen dargestellt. Diese Methode wurde neben Experten auch mit Personen ohne Ökobilanzexpertise anhand einer Fallstudie getestet und ausgewertet, um auf Basis verschiedener Visualisierungen belastbare Entscheidungen zu treffen.

## Acknowledgements

This dissertation is the outcome of my research at the Chair of Computational Modeling and Simulation at the Technical University of Munich. During the preparation of this work the authors used Grammarly in order to check grammar and terminologies. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

I want to express my sincere gratitude to Prof. Dr.-Ing. André Borrmann. Thank you for giving me this opportunity to research freely and open-mindedly and for always giving valuable feedback, ensuring I stay focused and not lose the Ph.D. path. And thank you for the comfortable environment for self-development!

A big thank you also goes to Prof. Dr.-Ing. Alexander Hollberg for the valuable discussions on my scientific path, and to Prof. Jakob Beetz and Prof. Dr.-Ing. Eva Patricia Schneider-Marin, for your interest in my research and for taking over the task of reviewing my dissertation!

Thank you, Prof. Dr.-Ing. Frank Petzold, for supporting my journey in this Ph.D. in many ways, including your valuable advice as a mentor.

Furthermore, I want to thank Siemens Real Estate, especially Christian Schunicht and Sabrina Fellermann, for funding the research project, which gave me the time and scope for major contributions to this dissertation, and for providing case studies for testing.

Additionally, I also want to thank LIST Eco and Forth Grünig Architekten for providing real-world case studies for training. Special thanks go to Sebastian Theißen and Jannick Höper for the exchange and constructive discussions over a couple of years in this emerging field of BIM-based LCA.

Thanks to my colleagues at the Chair of Computational Modeling and Simulation and at the Leonhard Obermeyer Center for the continuous exchange, support, and good times. Special thanks go to Florian, Ina, and Sebastian for always being willing to listen to me and my issues, as well as Martin for his great LaTeXsupport.

Last, I want to thank my family, especially my parents, for their support and trust in me. Biggest thanks to my partner Bilge for your continuous support, motivation, and joy in my life.

# Contributions

This cumulative dissertation is based on four published, peer-reviewed research papers, which are presented in Chapters 3 to 6.

## Paper I

**Forth, K.**; Abualdenien, J.; Borrmann, A.: *Calculation of embodied GHG emissions in early building design stages using BIM and NLP-based semantic model healing*. Energy and Buildings 284, 2023, DOI: 10.1016/j. enbuild.2023.112837

Contributions:
Kasimir Forth developed a methodology for Life Cycle Assessment of buildings in early design phases using Building Information Models (BIM) and Natural Language Processing (NLP). Jimmy Abualdenien contributed to the conceptualization and reviewed the manuscript. André Borrmann supervised this study, contributed to the conceptualization, and reviewed the manuscript.

## Paper II

**Forth, K.**; Berggold, P.; Borrmann, A.: *Domain-specific fine-tuning of LLM for material matching of BIM elements and Material Passports*. Proceedings of the 2024 ASCE International Conference on Computing in Civil Engineering, Carnegie Mellon University, Pittsburgh, PE, USA, 2024.

Contributions: Kasimir Forth developed the methodology for semantic enrichment and automated matching material datasets to IFC materials using Semantic Textual Similarity and evaluating the results of the case studies. Patrick Berggold contributed to the conceptualization and reviewed the manuscript. André Borrmann supervised this study and reviewed the manuscript.

## Paper III

**Forth, K.**; Borrmann, A.: *Semantic enrichment for BIM-based Building Energy Performance Simulations using Semantic Textual Similarity and fine-tuning multilingual LLM*. Journal of Building Engineering, 2024, DOI: 10.1016/j. jobe.2024.110312 )

Contributions: Kasimir Forth developed the methodology for semantic enrichment by automated matching space types and thermal constructions to BIM models using Semantic Textual Similarity, multilingual LLM fine-tuning strategies, and evaluating the results using several case studies. André Borrmann supervised this study and reviewed the manuscript.

## Paper IV

**Forth, K.**; Hollberg, A.; Borrmann, A.: *BIM4EarlyLCA: An interactive visualization approach for early design support based on uncertain LCA results using open BIM*. Developments in the Built Environment 16, 2023, DOI: 10.1016/j. dibe.2023.100263

Contributions:
Kasimir Forth developed a design-decision-making approach for reducing embodied greenhouse gas emissions by interactive, model-based visualizations of uncertain LCA results. He also designed a user study and analyzed its results. Alexander Hollberg supervised this study, contributed to the conceptualization, and reviewed the manuscript. André Borrmann supervised this study and reviewed the manuscript.

# Further Related Scientific Contributions

## Book chapters

- Höper, J.; Theißen, S.; Forth, K. (2024): BIM-basierte Gebäudeökobilanz. In: Wimmer, R.; Bartels, N.; Maile, N. (Ed.): Next Generation BIM. Aus der Praxis für die Lehre. 1. Auflage. Berlin: bSD Verlag - Haus der Bundespressekonferenz / 4103 (BIM Basics). ISBN: 978-3-948742-93-5

## Peer reviewed journal papers

- Selimovic, E.; Noichl, F.; Forth, K.; Borrmann, A. (2022): Retrofitting Potential of Building envelopes Based on Semantic Surface Models Derived From Point Clouds. Journal of Facade Design and Engineering 10 (2). p. 127-139. DOI: 10.47982/jfde.2022.powerskin.8

## Peer reviewed conference papers

- Ogunjinmi, G. J.; Forth, K.; Theißen, S.; Borrmann, A. (2024): Estimating the Circularity of Building Elements using Building Information Modelling. World Sustainable Built Environment Conference 2024, online. IOP Conference Series: Earth and Environmental Science. DOI: 10.1088/1755-1315/1363/1/012043
- Forth, K. (2023): Multilingual semantic enrichment of room-specific load profiles using BIM models for whole building energy simulation, In: Proc. of the 34. Forum Bauinformatik, Ruhr-Universität Bochum. p. 250-258. DOI: 10.13154/294-10093
- Forth, K.; Hollberg, A.; Borrmann, A. (2023): Interactive visualization of uncertain embodied GHG emissions for design decision support in early stages using open BIM. In: Proceedings of Eighth International Symposium on Life-Cycle Civil Engineering (IALCCE), Milan, Italy. p. 3634-3641. DOI: 10.1201/9781003323020
- Forth, K.; Noichl, F.; Borrmann, A. (2023): LCA Calculation of Retrofitting Scenarios using Geometric Model Reconstruction and Semantic Enrichment of Point Clouds and Images. Proc. of the ASCE International Conference on Computing in Civil Engineering 2023. Corvallis, Oregon, USA. DOI: 10.1061/9780784485231.047
- Lammers, B.; Forth, K. (2022): IFC-based variant analysis considering multicriterial sustainability analysis of buildings. In: Proc. of 33. Forum Bauinformatik, München. p. 265-272. DOI: 10.14459/2022 md1686600
- Forth, K.; Abualdenien, J.; Borrmann, A. (2022): NLP-based semantic model healing for calculating LCA in early building design stages. In: Proceedings of the 14th European Conference on Product & Process Modelling (ECPPM 2022). Trondheim, Norway. p. 77–84. DOI: 10.1201/9781003354222
- Forth, K.; Höper, J.; Veselka, J.; Theißen, S.; Borrmann, A. (2022): Towards life cycle assessment of technical building services in early design phases using building information modelling. In: Proceedings of the 2022 European Conference on Computing

in Construction. 2022 European Conference on Computing in Construction (EC3 2022), Rhodes, Greece. DOI: 10.35490/EC3.2022.178

- Forth, K.; Schneider-Marin, P.; Theißen, S.; Höper, J.; Svane, N. D.; Borrmann, A. (2022): Connected design decision networks: Multidisciplinary decision support for early building design LCA. In: APP 38, p. 124–130. DOI: 10.14311/APP.2022.38.0124

- Schumacher, R.; Theißen, S.; Höper, J.; Drzymalla, J.; Lambertz, M.; Hollberg, A.; Forth, K. et al. (2022): Analysis of current practice and future potentials of LCA in a BIM-based design process in Germany. In: E3S Web Conf. 349, p. 10004. DOI: 10.1051/e3 sconf/202234910004

- Kolbeck, L.; Forth, K. (2021): Interoperability of BIM based Life Cycle Energy Analysis in Early Design Stages. Proc. of the Forum Bauinformatik 2021. Online

- Forth, K.; Abualdenien, J.; Borrmann, A.; Fellermann, S.; Schunicht, C. (2021): Design optimization approach comparing multicriterial variants using BIM in early design stages. Proc. of 38th International Symposium on Automation and Robotics in Construction (ISARC 2021). Online. DOI: 10.22260/ISARC2021/0034

- Forth, K.; Braun, A.; Borrmann, A. (2019): BIM-integrated LCA - model analysis and implementation for practice. In: IOP Conf. Ser.: Earth Environ. Sci. 323, p. 12100. DOI: 10.1088/1755-1315/323/1/012100

## Reports

- Bahlau, S.; Schumacher, R.; Lambertz, M.; Theißen, S.; Höper, J.; Borrmann, A.; Forth, K.; von Both, P.; Ebertshäuser, S.; Horn, R. (2024): Digital Twin Footprint - Erarbeitung eines ganzheitlichen Meilensteinplans mit Handlungsempfehlungen und notwendigen Forschungsbausteinen zur zielführenden Verknüpfung der Lebenszyklusanalyse (Gebäudeökobilanzierung) und BIM-Planungsprozesse mit einem Fokus auf den frühen Planungsphasen. Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR). ISSN: 1868-0097

# Contents

# Acronyms

| | |
|---|---|
| **AEC** | Architecture, Engineering, and Construction |
| **AI** | Artificial Intelligence |
| **BCA** | Building Circularity Assessment |
| **BCF** | BIM Collaboration Format |
| **BDL** | Building Development Level |
| **BEM** | Building Energy Modeling |
| **BEPS** | Building Energy Performance Simulation |
| **BERT** | Bidirectional Encoder Representations from Transformers |
| **BIM** | Building Information Modeling |
| **ContrastiveL** | Contrastive Loss |
| **CosL** | Cosine Similarity Loss |
| **CV** | Computer Vision |
| **DL** | Deep Learning |
| **DPP** | Digital Product Passports |
| **DSR** | Design Science Research |
| **gbXML** | Green Building Extensible Markup Language |
| **GHG** | Greenhouse Gas |
| **GIS** | Geographic Information Systems |
| **GWP** | Global Warming Potential |
| **HBJSON** | Honeybee JSON |
| **IDM** | Information Delivery Manual |
| **IDS** | Information Delivery Specification |
| **IFC** | Industry Foundation Classes |
| **JSON** | JavaScript Object Notation |
| **LaBSE** | Language-agnostic BERT Sentence Embedding |
| **LCA** | Life Cycle Assessment |
| **LCI** | Life Cycle Inventory |
| **LCIA** | Life Cycle Impact Assessment |
| **LKdb** | LCA knowledge database |
| **LLM** | Large Language Model |

**LOD** Level of Development
**LOG** Level of Geometry
**LOI** Level of Information
**LOIN** Level of Information Needs
**LSTM** Long Short Term Memory Networks

**MEP** Mechanical Electric Plumbing
**MNRL** Multiple Negatives Ranking Loss
**MP** Material Passport
**MSEL** Mean Squared Error Loss
**MVD** Model View Definition

**NLP** Natural Language Processing
**NREL** National Renewable Energy Laboratory
**NSC** Neuro-Symbolic Computing

**PLM** Pre-trained Language Model
**Pset** Property Set

**RNN** Recurrent Neural Networks

**SNLI** Stanford Natural Language Interference
**STS** Semantic Textual Similarity

**UUID** Universally Unique Identifier

# Chapter 1

# Introduction

## 1.1 Motivation

According to the International Energy Agency, the Architecture, Engineering, and Construction (AEC) sector and building operations are responsible for approximately 37% of the global final energy consumption and around 40% of the global Greenhouse Gas (GHG) emissions (IEA, 2024). The impact of operational energy is approximately 27%, and further energy consumption related to the construction-related manufacturing industry is around 13%.

Furthermore, around 100 billion tons of waste annually originates from construction, renovation, and demolition according to the UN Environment Programme (United Nations Environment Programme, 2022), while about 35% of it is sent to landfill (Chen et al., 2022). Conversely, global material use might increase from 79 Gt in 2011 to 167 Gt in 2060, with the construction industry having the highest material intensity of all sectors (OECD, 2019).

These statistics show the necessity of holistic assessment of environmental impacts for new constructions and renovation or demolitions of buildings, including operational and embodied emissions, as well as the demand to increase the circular economy of the built environment. The European Commission established through the EU taxonomy a legal framework for, among others, real-estate-related financial transactions that address these challenges by measures for climate change mitigation, climate adaptation, and circular economy (European Commission, 2021).

### 1.1.1 Environmental analyses using BIM

Environmental analyses, such as Building Energy Performance Simulation (BEPS), Life Cycle Assessment (LCA), or Building Circularity Assessment (BCA), have been introduced to address the issues of global warming, waste generation, and resource demand in the construction industry. Material Passport (MP), also known as (Building) Ressource Passports or used as synonyms for Digital Product Passports (DPP) on product-level, form the information basis for different environmental analyses, such as LCA, or BCA, as shown in more detail in Section 2.2.2.

Figure 1.1 shows schematically material- and element-based design decisions in early design stages compared to detailed design stages. Material and element decisions are often relevant to the results of environmental assessments of buildings, such as operational or embodied energy and emissions. As some materials and elements are not decided yet

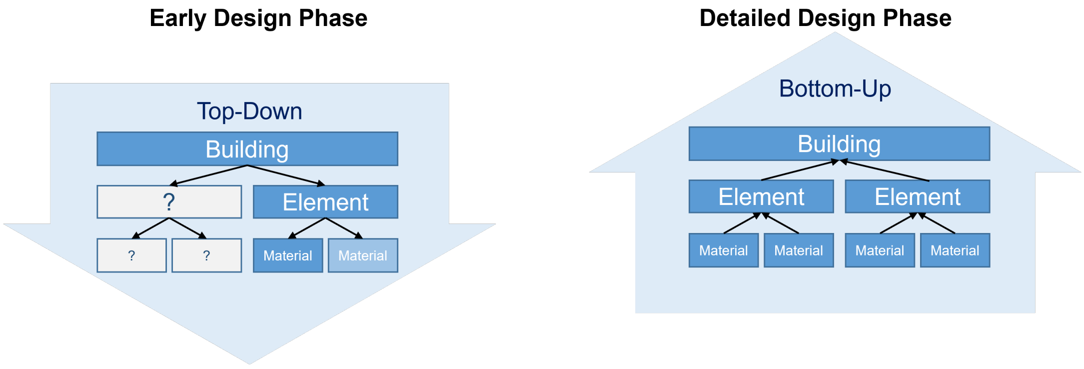



Figure 1.1: Material- and element-based design decisions in early and detailed design stages

in early design stages or only vaguely described in general terms, such as concrete without a specific stiffness class, environmental analyses are often conducted using benchmarks and a top-down approach. When all decisions have been made in detailed design stages, the environmental analyses are calculated from the bottom up.

Building Information Modeling (BIM) is an established method for a data-driven approach for these environmental analyses using its models as a single source of truth (Borrmann et al., 2021). However, these analyses are often conducted in detailed design stages when all necessary information is available (Schumacher et al., 2022). This information is often vague, uncertain, or completely missing in the early design stages.

A distinction is made in BIM workflows between closed BIM, meaning the usage of software-specific, and open BIM, the usage of vendor-neutral data formats. Closed BIM is often preferred over open BIM for LCA (Schumacher, 2021). The advantages of closed BIM for environmental analyses are reduced data loss due to direct implementation in the authoring tools using plugins. Nevertheless, all stakeholders of the BIM-based project need to agree on one authoring tool and have licenses. However, this is not always the case, so open BIM is preferred, especially by public clients. One main challenge here is ensuring the correct data quality by exporting open BIM models from different authoring tools using the data format Industry Foundation Classes (IFC). Especially in the early design stages, not all models are checked and solved in detail when design decisions with high environmental impacts must be made.

### 1.1.2 Problem statement

Relevant data for LCA or BEPS are often not in the same hierarchy as BIM models in the early design stages. LCA databases, such as Ökobaudat (BBSR, 2024), are usually not very detailed and structured but only material- or product-specific, while BIM elements often do not have all layers and materials modeling.

Figure 1.2 shows a simple BIM model as an example of semantic matching relevant information for environmental analyses. One usual naming of the BIM exterior wall with

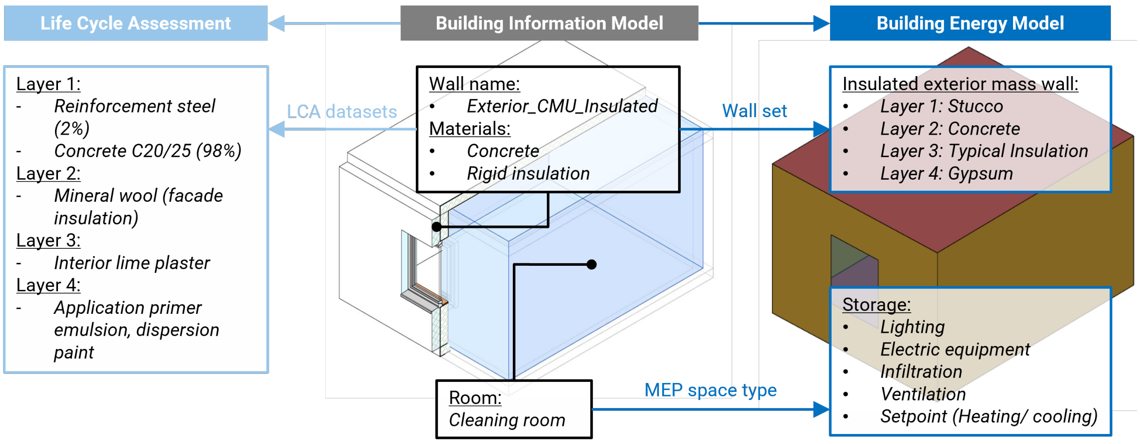



Figure 1.2: Semantically matching relevant information for LCA (left) and Building Energy Modeling (BEM) (right) based on an exemplary BIM model

reinforced concrete in teh authoring software Autodesk Revit is "Exterior_CMU_Insulated", with the BIM material "concrete" and "rigid insulation".

The related LCA datasets would include "reinforcement steel wire" and a generic concrete dataset, which needs a specific stiffness class assigned, e.g. "concrete C20/25". However, wall finishes, plaster, and paint materials are usually not modeled in the BIM model but must be manually added for holistic and reliable LCA results.

Another environmental analysis use case covers the semantic enrichment for BEM models based on BIM in order to ensure reliable BEPS results. In this case, space types for Mechanical Electric Plumbing (MEP), including internal loads and schedules, need to be assigned based on the naming of architectural rooms. For example, garbage or cleaning rooms are both assigned to the MEP space type of "storage". Furthermore, the thermal properties of all element layers must be enriched based on BIM elements, also considering missing layers, similar to the before-mentioned use case of LCA.

Consequently, domain experts need to check and manually enrich the models with missing information to ensure reliable environmental analysis results. However, this enrichment is time-intensive and costly. Therefore, this BIM model enrichment process for environmental analyses is usually done only once and in detailed stages. Furthermore, the results of environmental analyses are often interpreted by domain experts rather than decision-makers, such as clients, although the BIM method promises an integral and more transparent design approach.

## 1.2 Aim and scope of this thesis

To overcome barriers of BIM-based sustainability analyses in early stages, this dissertation proposes a holistic framework for (semi-)automated semantic enrichment of open BIM models for environmental analyses to support design decision-making in early design stages using Natural Language Processing (NLP), especially its subtask Semantic Textual Similarity (STS), and Large Language Model (LLM). This section introduces the challenges addressed by this thesis, based on the identified research gaps that will be presented in Section 2.4:

1. **Automated semantic enrichment of information for decision-support using environmental analyses**: The impact of design decisions in early design stages is significantly lower compared to detailed design stages (MacLeamy, 2004). The aim of automatically enriching semantic information relevant to these environmental analyses is to include the results of these analyses for data-based design decision support in early design stages. Pre-trained and domain-specifically fine-tuned LLM can support this enrichment by matching the semantically most similar data from a use-case-specific database. For the use case of LCA, a knowledge database based on the German Ökobaudat (BBSR, 2024) is proposed, for MP, an unstructured material database by EPEA (EPEA GmbH, 2022), and for BEPS, the American databases by National Renewable Energy Laboratory (NREL) are used (NREL, 2024).
2. **Adding missing, uncertain element or material information for LCA in early design stages**: Early design stages are characterized by uncertain or missing information, for example, about elements and their materials. In the detailed design stages, this information has already been decided upon and is available in the BIM model. To cover these uncertainties for life cycle assessments, the main aim is to calculate ranges of LCA results rather than exact values in these early stages.
3. **Using open BIM workflow and data formats to integrate the environmental analyses results and design decisions**: To establish a vendor-neutral collaboration framework, open BIM workflow and data formats are essential. Therefore, the proposed method exchanges BIM models using IFC. Furthermore, the BIM Collaboration Format (BCF) is introduced to communicate design decisions and LCA results to BIM modelers in their authoring tools. The BCF schema can be extended to save all relevant information about the LCA results and the design decision-making.
4. **Visualization of environmental impacts for design-decision-making of non-LCA-experts**: BIM and LCA both require expert knowledge and training to be able to use the tools and interpret the results. Therefore, clients and building owners often need experts, such as BIM or sustainability consultants, to interpret the LCA results and use the tools for them. The aim is to use a combination of different visualization strategies incorporated in an open BIM workflow to enable a transparent design decision framework that can be also used by non-LCA experts.

## 1.3 Research methods

This dissertation follows the research method Design Science Research (DSR) by Peffers et al. (2008) and was applied in all four publications I-IV. It consists of six operative steps:

(1) problem identification, (2) objective identification, (3) design and development, (4) demonstration, (5) evaluation, and (6) conclusion.

The problems and objectives are identified as research questions, hypotheses, and objectives in the following Subsections 1.3.1 and 1.3.2 and shown in Figure 1.3. The design and development stages result in proposing different frameworks resulting in DSR artifact. Prototypical implementations of these frameworks represent the demonstration phase. Each publication has different evaluations of the demonstrations using different experiments in order to answer the raised research questions. Additonally, Paper IV uses a user study testing the prototype and a survey following qualitative and quantitative evaluations. Finally, the conclusion phase is included in each publication, but also overall for the overall framework and whole dissertation in Chapter 7.

| | Research questions | Hypotheses | Objectives | Publications |
|---|---|---|---|---|
| Matching method | **Research question I:** Which degree of automation is possible to ensure reliable environ. analyses results? | **Hypothesis I:** Automation of matching method depends on database structure and LLM fine-tuning. | **Objective I:** Develop. (semi-)automated matching methodologies for three environmental analyses (LCA, MP, BEPS) | **Paper I - III** |
| Matching method | **Research question II:** How can BIM models be semantically healed such that a reliable LCA can be calculated? | **Hypothesis II:** NLP and pre-trained LLM support automated matching for holistic and reliable LCA results. | **Objective II:** Automated matching methodology using a structured LCA knowledge database with uncertainties | **Paper I:** Forth, K., Abualdenien, J., Borrmann, A.; Energy and Buildings, 2023 |
| LLM fine-tuning | **Research question III:** Which fine-tuning strategy improves the monolingual matching performance for Material Passports? | **Hypothesis III:** A combination of different strategies of fine-tuning LLM improves the MP matching accuracy. | **Objective III:** Monolingual material matching methodology using a combination of fine-tuning strategies | **Paper II:** Forth, K., Berggold, P., Borrmann, A.; Proc. of ASCE i3CE, 2024 |
| LLM fine-tuning | **Research question IV:** Which fine-tuning strategies improve the multilingual matching performance for BEPS? | **Hypothesis IV:** Use-case-specific comb. of LLM fine-tuning strategies improves the BEPS matching accuracy. | **Objective IV:** Multilingual room- & element matching method using a combination of fine-tuning strategies | **Paper III:** Forth, K., Borrmann, A., Building Engineering, 2024 |
| Visualization & decision support | **Research question V:** How can uncertainties and LCA results in early design stages be intuitively visualized also for non-LCA-experts? | **Hypothesis V:** A combination of visualization strategies using 3D models and color-coding support non-LCA experts. | **Objective V:** Develop. visualization strategies incorporating uncertainties and LCA & evaluation by a user study. | **Paper IV:** Forth, K., Hollberg, A., Borrmann, A., Developments in the Built Environment, 2023 |
| Visualization & decision support | **Research question VI:** How can open BIM data formats support the design decision-making process for environmental analyses? | **Hypothesis VI:** IFC models support calculating environmental analyses, and extended BCF schema enables feedback communic. | **Objective VI:** Develop. a methodology for automat. calculating environmental analyses and digitally communi-cating information back | **Paper I - IV** |

Figure 1.3: Scope of the publications presented in the following chapters

### 1.3.1 Research questions

Based on the identified scope from Section 1.2, the following research questions arise:

1. Which degree of automation is possible for the matching and enrichment process to ensure reliable environmental analysis results for LCA, MP, and BEPS?
2. How can BIM models be semantically healed to enrich correct element types and materials to the respective model elements so that a reliable LCA can be calculated?
3. Which monolingual LLM strategies of fine-tuning pre-trained LLM improves the matching performance from IFC materials for MP?
4. Which multilingual LLM fine-tuning strategies improve the accuracy for matching and enriching program types based on architectural rooms and construction with thermal properties based on IFC elements and materials for BEPS?
5. How can uncertainties and LCA results in early design stages be intuitively visualized for design decision support by non-LCA-experts?
6. How can open BIM data formats support the design decision-making process for environmental analyses?

### 1.3.2 Research hypothesis and objectives

The derived hypotheses and objectives are addressed in dedicated chapters as shown in Figure 1.3. These questions are clustered into three main topics: matching approach, LLM fine-tuning, and visualization & decision support.

To approach the raised research questions, the following hypotheses and objectives were derived:

1. **Hypothesis:** Depending on the structure of databases and fine-tuned LLMs with domain knowledge, the matching method is fully or only semi-automated ensuring reliable results for the different use cases of LCA, MP, and BEPS.

   **Objective:** Development of (semi-)automated matching methods for three environmental analyses, such as LCA, MP, and BEPS, and (un-)structured databases and fine-tuned LLM with domain knowledge.

2. **Hypothesis:** Semantic Textual Similarity and pre-trained Large Language Models support the automation of matching BIM elements to LCA datasets for calculating a holistic, whole-building LCA. Missing layer information is added by structuring an LCA-specific knowledge database, which heals the incomplete model.

   **Objective:** Development of a framework for automatically matching BIM to LCA data using a structured LCA knowledge database (LKdb). Uncertainties in early design stages lead to a range of LCA results.

3. **Hypothesis:** A combination of different strategies of fine-tuning pre-trained LLM improves the matching performance of IFC materials for Material Passports based on small datasets, such as structuring and filtering the material database, adding domain knowledge, using an optimal loss function and adding negative pairs.

**Objective:** Development of a framework for domain-specific fine-tuning of LLM for enriching Material Passports by matching IFC materials to unstructured databases.

4. **Hypothesis:** Depending on the use case, differentiating between space type to room matching or matching constructions with thermal properties, different combinations of multilingual LLM fine-tuning strategies, such as different loss functions, adding negative pairs, domain-specific abbreviations, adding context labels, or different multilingual student LLM, improve the matching accuracy of BIM-based building energy performance simulations.

   **Objective:** Development of a framework for multilingual use case-specific matching method ("rooms" & "elements") based on small datasets derived from three case studies using a combination of fine-tuning strategies.

5. **Hypothesis:** A combination of different visualization strategies using 3D models and color-coding supports non-LCA experts in making design decisions with comparable improvements as by LCA experts.

   **Objective:** Development of visualization strategies incorporating uncertainties and LCA results and evaluating them using a user study.

6. **Hypothesis:** IFC models support automated calculation and visualization of environmental analyses' results, and extending BCF schema enables feedback communication of decision and LCA results.

   **Objective:** Development of a framework for automatically calculating reliable results of environmental analyses for decision support and development of an extended BCF schema for communicating relevant information back to BIM modelers.

## 1.4 Overall concepts

This section introduces a generalized matching approach, which represents the common contribution and is applied to different use cases of environmental analyses in the following chapters. Additionally, it shows how the semantic enrichment of all use cases of environmental analyses and the proposed design decision support approach interact in a general framework.

### 1.4.1 General matching approach

Figure 1.4 shows the initial idea of a matching approach for element-specific semantic model enrichment based on Semantic Textual Similarity (STS) as presented in (Forth et al., 2021). Three main steps are proposed to match a product from the IFC schema to one of the LCA knowledge database (LKdb):

1. Filtering
2. Similarity calculation

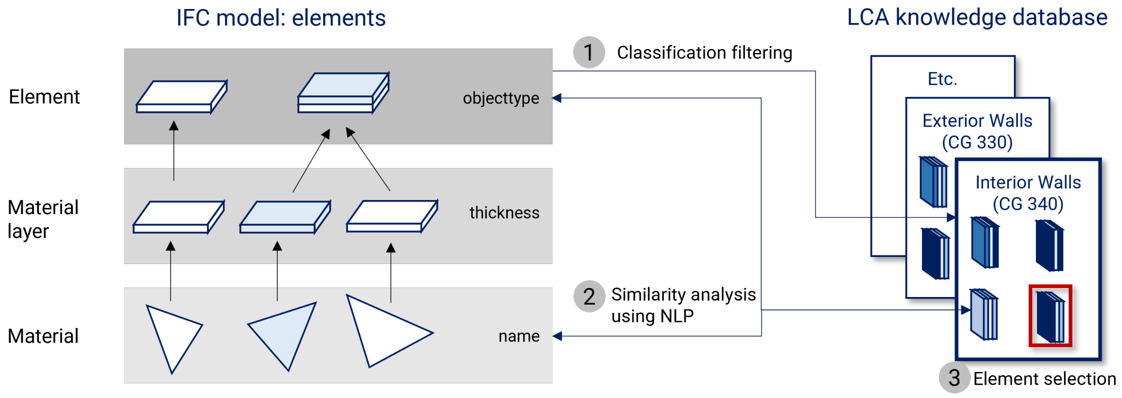



Figure 1.4: Matching approach for element-specific semantic model enrichment by (Forth et al., 2021)

3. Selection and enrichment

The IfcProduct entity "is an abstract representation of any object that relates to a geometric or spatial context" (buildingSMART International Ltd., 2024). Depending on the matching use case, it varies and can be either IfcSpace, IfcElement or the IfcMaterial associated with this IfcElement and is further described in Figure 2.3 in Section 2.1.3. The first filtering step aims to narrow the solution space and ensure meaningful results using a classification system. In this case, the German cost group system for building elements is used (DIN 276, 2018). The second step includes the similarity calculation of every filtered element from the database with the one from the BIM model. Finally, the semantically most similar element is selected, and all information is enriched for the environmental analysis. Depending on each use case, different levels of automation can be achieved, starting from semi-automated design decision support to full automation.

Algorithm 1.1: Generilzed algorithm for matching data from a filtered external database to products from IFC model

```
for ifc_product in all_relevant_products_in_IFCmodel:
    similarity_dictionary = {}
    for data in filtered_database:
        term_similarity = cosine_similarity(vectorize(ifc_product),
            vectorize(data))
        token_similarities = []
        for token_ifc_product in tokenize(ifc_product):
            for token_data in tokenize(data):
                token_similarity = cosine_similarity(vectorize(
                    token_ifc_product), vectorize(token_data))
                token_similarities.append(token_similarity)
        if max(token_similarities) > term_similarity:
            data_similarity_dictionary[data] = max(token_similarities)
        else:
            similarity_dictionary[data] = term_similarity
    selected_data = max(similarity_dictionary)
```

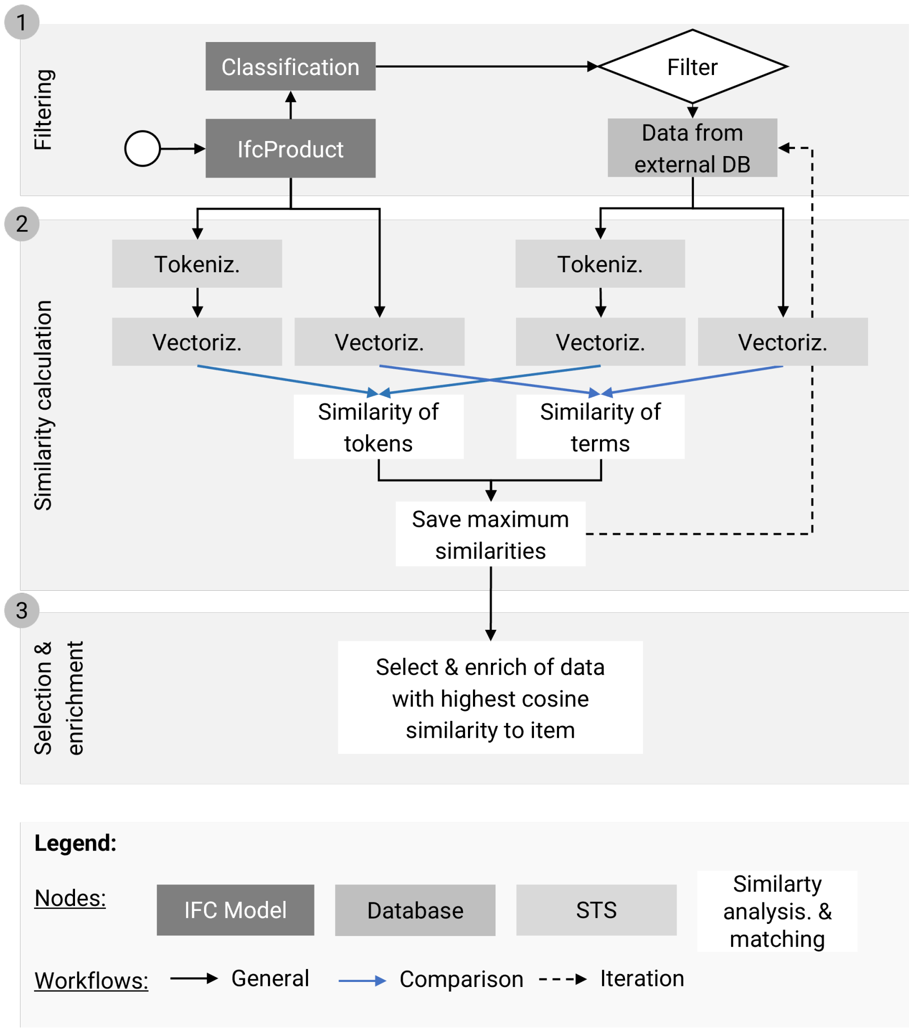



Figure 1.5: Generalized matching concept using Semantic Textual Similarity of IFC products and filtered databases

Figure 1.5 and Algorithm 1.1 generalize the initial matching concept so that it can be applied not only for element or material matching but also space types to rooms or similar products from IFC to a hierarchically structured database. It shows the different matching steps following the iterative workflow in more detail and is also publicly available with an example from Section 2.3.3 on Github[1]. This general concept is applied to the three different use cases of LCA, MP, and BEPS in the following chapters.

In the first step, the respective product from the IFC model, including its classifications, is iterated. Next, the matching database is filtered according to the product's classification. This is necessary in order to narrow down the solution space and thereby significantly increase the overall matching accuracy and ensure reliable and meaningful results. The classification, relevant data in the IFC schema, and the database for matching vary depending on each use case.

After each product, such as elements, materials, or rooms, in the IFC model is selected, its tokens and whole expressions/terms are vectorized by its vector embeddings and com-

[1] https://github.com/kasforth/ifcProductMatching

pared to every filtered data set from the database. In Section 2.3.3, different approaches are discussed to measure the semantic similarity of two terms. As transformer-based, Pre-trained Language Model (PLM) are used and fine-tuned, cosine similarity of the terms embeddings is used as a similarity measure. For every token of the term and the term itself, the cosine similarities are measured.

The maximum similarity of either the token or the term is saved for each iterated dataset from the filtered database. Finally, the maximum similarity of all cosine similarities is found, and the data with the maximum similarity is selected and enriched to the IFC product afterward.

### 1.4.2 General framework

Figure 1.6 shows an overview of how all proposed use cases and the decision-support part are combined into one overall framework. The framework is divided into five different main steps:

1. BIM modeling
2. Geometric processing
3. Semantic enrichment
4. Environmental analyses
5. Design decision support

The first two steps include the BIM modeling and the geometric processing. As only pre-modeled real-world use cases are used in this dissertation, the BIM modeling and the

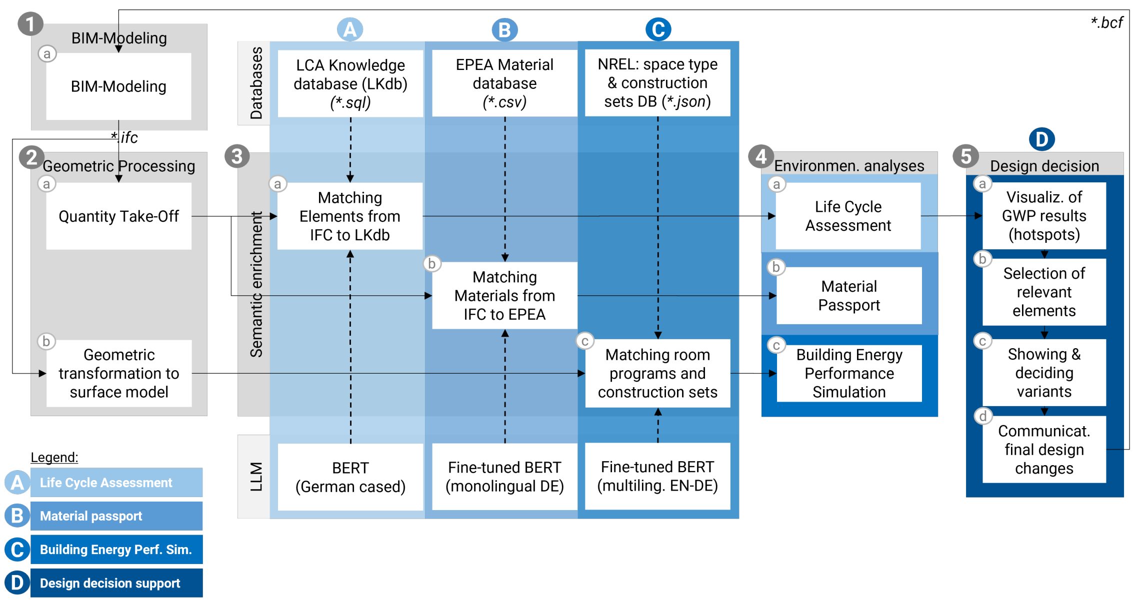



Figure 1.6: Overall framework for Semantic model enrichment for environmental analyses using Semantic Textual Similarity for design decision support in early design stages

model export using the IFC schema are excluded from the scope of this thesis. For the second step, depending on the environmental analyses, either a quantity take-off for life cycle assessment and material passports or geometrically transformed surface models for building energy performance simulation are processed as input for the following semantic enrichment step.

The main contribution of this dissertation is step 3, the semantic enrichment. This step consists of the matching approaches based on different databases for each environmental analysis, and using different Large Language Models. The use-case-specific databases have different degrees of structure, with the EPEA material database being unstructured, the NREL space types database being structured only with one hierarchy level, and the NREL construction database as well as the LKdb having the highest degree of being structured with multiple hierarchal levels. Furthermore, the highlighted colors differentiate between four parts: three different applied environmental analysis use cases of LCA, MP and BEPS, and the design decision part. Each part will be described in more detail in the following chapters and represents the four published peer-reviewed research papers.

Use case A covers the use case of LCA in early design stages and is presented in Chapter 3. For the semantic enrichment of this environmental analysis, a LKdb is proposed. Furthermore, three different pre-trained LLM, such as GermaNet (Henrich & Hinirchs, 2010), SpaCy (Honnibal & Montani, 2017) and the German BERT (Chan et al., 2020), are tested for matching and the German BERT is evaluated as the best performing LLM.

Use case B contains semantic enrichment for MP, described in more detail in Chapter 4. The unstructured EPEA material database is used (EPEA GmbH, 2022), but the main focus is further fine-tuning the German BERT LLM, including domain knowledge, using Reimers and Gurevych's methodology of using siamese BERT networks (Reimers & Gurevych, 2019).

The third use case, C, focuses on BEPS presented in Chapter 5. It included two enrichment steps for enriching space types based on architectural rooms and thermal properties based on elements. Both databases, one for space types and one for constructions with thermal properties, are provided by the NREL of the United States of America in English (NREL, 2024; Wilson et al., 2021), and the available real-world case studies and BIM models are in German. Therefore, the main focus is multilingually LLM fine-tuning of pre-trained BERT models for this use case following Reimers and Gurevych's approach of knowledge distillation for multilingual sentence embeddings (Reimers & Gurevych, 2020).

In step four, the use-case-specific environmental analyses are calculated or simulated after the BIM models are semantically enriched. The fifth and last step contains the design decision support based on the LCA results. This step is divided into the visualization of the LCA results using the embodied GHG emissions for hotspot analysis, followed by selecting the relevant elements, showing their variants and deciding on one, and lastly, communicating the final design changes back all relevant information as issues to the BIM modeler using the BCF.

## 1.5 Structure of the thesis

This cumulative dissertation is divided into seven chapters. This Chapter 1 first introduces the aim and scope of the research conducted, as well as six research questions and hypotheses are defined. Finally, overall concepts are described, focusing on the general framework and general matching approach. Chapter 2 gives an overview of the relevant background, related works, and current State of the Art about Building Information Modeling, BIM-based environmental assessments, and Natural Language Processing in the AEC industry. The following chapters present three journal publications and one conference publication according to the general framework and according to the contributions.

Chapter 3 contains **Paper I** which was published in the Journal "Energy and Buildings" (Forth, Abualdenien, & Borrmann, 2023). It describes the proposed method of semantic enrichment and model healing for life cycle assessments in early design stages using pre-trained LLM and a structured LCA knowledge database (LKdb). The most similar elements are matched from the database based on BIM elements and material names using semantic similarity, and ranges of embodied LCA results are calculated, mainly focusing on Global Warming Potential (GWP) and GHG emissions.

In the following Chapter 4, depicting **Paper II** from the 2024 ASCE International Conference on Computing in Civil Engineering (Forth et al., 2024), the matching method is adapted for the use case of Material Passports. However, the material database used is unstructured, and the focus is on the LLM fine-tuning strategies for the monolingual semantic matching of material datasets to BIM materials.

In Chapter 5, the matching method is further developed multilingually for enriching space types and thermal properties to enable the use case of Building Energy Performance Simulations. This chapter includes **Paper III**, which was published in the "Journal of Building Engineering" (Forth & Borrmann, 2024). Based on architectural room names, the most similar space types and the most similar constructions with thermal properties are matched based on the BIM elements and materials. The LLM fine-tuning has a monolingual and multilingual step, as BIM models in German and NREL databases in English are used.

Chapter 6 contains **Paper IV** which was published in the Journal "Developments in the Built Environment" (Forth, Hollberg, et al., 2023). It introduces a novel approach for design decision-making through interactive, model-based visualizations of uncertain LCA outcomes from Chapter 3. The proposed methodology uses open BIM data formats, including IFC and BCF. It is tailored to provide decision support for non-LCA experts in the early design stages. Finally, Chapter 7 evaluates all the defined objectives of the thesis, concluding with remaining gaps and an outlook for future research.

# Chapter 2

# Background and related works

This chapter about the current State of the Art is divided into three main parts: Building Information Modeling (BIM), environmental analyses using BIM, and Natural Language Processing (NLP) and Large Language Model (LLM).

## 2.1 Building Information Modeling

In this section, the concepts of Building Information Modeling (BIM) are briefly introduced, which are relevant for environmental analyses. First, the challenges of early design stages, including uncertainties, are discussed, followed by the concept of Level of Development (LOD) and Level of Information Needs (LOIN), data exchange formats for environmental analyses, and semantic enrichment of BIM models. Generally, the focus is on the open BIM method, which is vendor-neutral and follows a federated model approach (Borrmann et al., 2021).

### 2.1.1 Early design stages and uncertainties

According to the MacLeamy curve (MacLeamy, 2004), the impact of cost and functional capabilities are higher in schematic design and design development stages, and the costs of design changes are lower compared to detailed design stages. However, many decisions have not been made yet in early design stages, and several uncertainty in design decisions exists (Knotten et al., 2015). Conducting environmental assessments in early design stages helps domain experts in data-based decision-making, as Abualdenien showed in his dissertation (Abualdenien, 2023).

As shown in Figure 1.1 in the Motivation Section 1.1, material- and element-based design decisions are often not decided yet in early design stages or only vaguely available using generic terms, e.g., concrete. However, this detailed information is relevant for a bottom-up calculation of environmental analyses. If uncertainties are not addressed in design decisions in early design stages, the results of environmental analyses might be misleading or wrong. Therefore, uncertainties must be defined along every design stage, especially in early phases (Tian et al., 2018).

Critical decisions significantly influence the building's carbon footprint in the early design stages. However, these stages are marked by substantial uncertainty due to incomplete information and pending decisions, complicating the task of performing a holistic and consistent Life Cycle Assessment (LCA) to support design decisions and optimize performance (Schneider-Marin et al., 2020). Goulouti et al. recommend a probabilistic approach for service lives of building elements to increase the reliability of LCA and results (Goulouti et al., 2020). Harter et al. (2020) and Schneider-Marin et al. (2020) analyzed the effect of geometric and material-related uncertainties of embodied and operational emissions testing.

Warrier et al. classified several sources of uncertainties in building LCA, such as stakeholder decisions, input data and data quality, future stages (service life of components, end-of-life choices, etc.), and uncertainties related to LCA methods (Warrier et al., 2024). Relevant to this dissertation are the uncertainties due to input data, which are differentiated into the quantity of materials and modeling choices of construction materials.

### 2.1.2 Level of Development and Level of Information Need

As previously described, building design progresses iteratively, evolving from initially vague information to more detailed specifications, leading to increased accuracy and reliability of BIM models throughout the modeling process (Abualdenien et al., 2021).

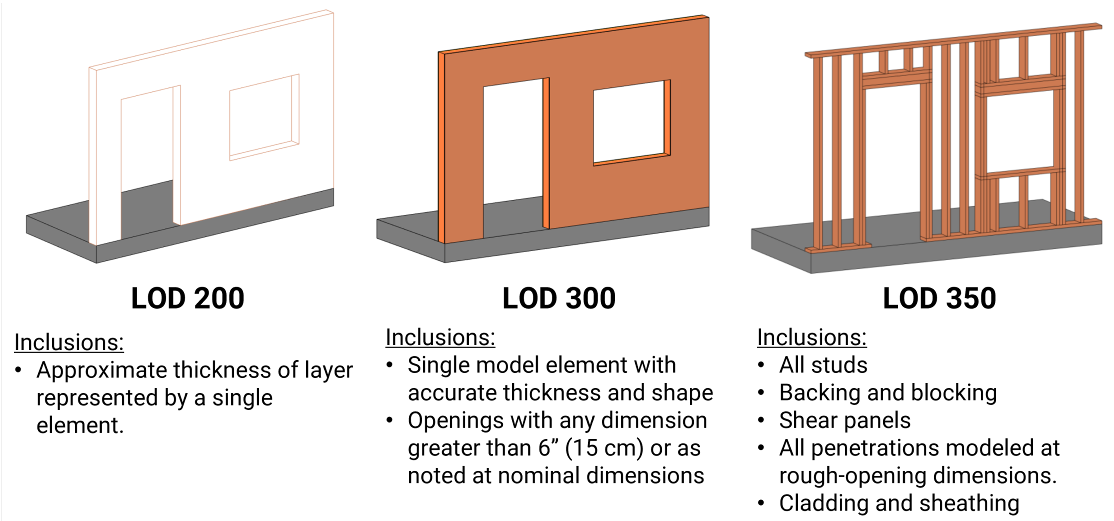



Figure 2.1: Level of Development (LOD) of an exterior wooden wall construction (B2010.20.10) according to (BIM Forum, 2023).

The Level of Development (LOD) defines the degree of completion, maturity, or elaboration. While the BIMforum, the US chapter of buildingSMART International, has delineated specific LODs, they have yet to be universally adopted as an international standard (BIM Forum, 2023). Figure 2.1 shows exemplarily the LOD from 200, 300, and 350 of exterior wooden wall construction and its inclusions.

In the European standardization effort (DIN EN 17412-1, 2021), Level of Information Needs (LOIN) aligns with LOD, encompassing geometric and alphanumerical information, but is tailored to particular use cases and milestones. However, Level of Information Needs (LOIN) is a framework that does not define overall specific levels, as in the LOD concept by BIM Forum. In Germany, LOIN is known as the combination of Level of Geometry (LOG), specifying geometric detailing, and Level of Information (LOI), representing the extent of alphanumerical information. Level of Information (LOI) is contingent on project specifics and client requirements, thus lacking generalizability. In practice, LOI is often articulated through "Type-and-attribute tables" (TAT), detailing object types and attributes (Borrmann et al., 2021). Additionally, buildingSMART International has proposed Information Delivery Specification (IDS) aiming to author and validate non-geometrical information requirements, such as material or classifications (Tomczak et al., 2022).

Abualdenien and Borrmann introduced a meta-model approach where multi-LOD data depict buildings at various design phases, building upon BIMForum's LOD definitions and introducing the Building Development Level (BDL) concept (Abualdenien & Borrmann, 2019). While LOD delineates specific components, the BDL concept includes the overall building maturity with multiple LODs for each component. The BDL concept resembles the Modelldetaillierungsgrad (MDG) by VBI (Verband Beratender Ingenieure VBI, 2016).

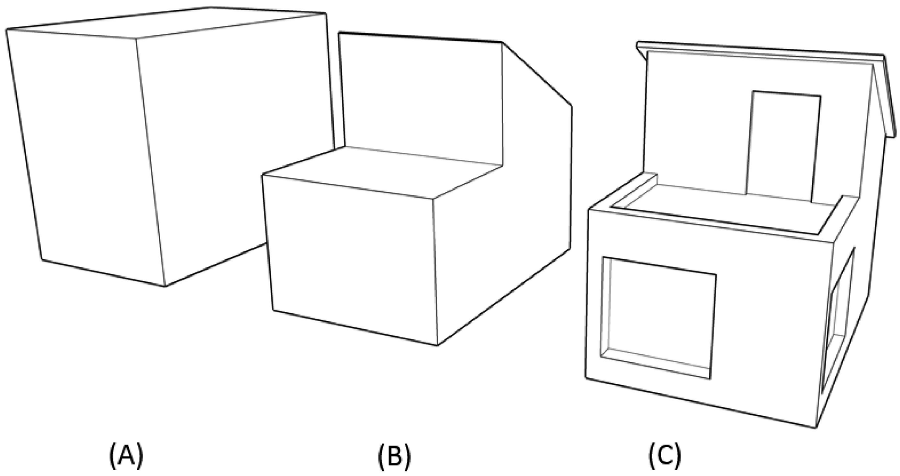



Figure 2.2: Example of symbolic (A), simplified (B) and detailed (C) graphical representation of a building for 3D modeling to support master planning (A), early light analyses (B) and detail light analyses (C) according to (DIN EN 17412-1, 2021).

Figure 2.2 graphically represents a 3D BIM model in the steps of a symbolic (A), simplified (B), and detailed (C) design stage for early (B) and detailed (C) daylight simulations according to (DIN EN 17412-1, 2021). LOIs and LOGs are pivotal in BIM-based environmental analyses by specifying required information or accommodating information uncertainty in early design stages. As less information is available initially, generic datasets are utilized, necessitating assumptions for missing material layers. Product-specific datasets can be incorporated into calculations based on the components utilized during construction.

### 2.1.3 Data exchange formats for environmental analyses

The Industry Foundation Classes (IFC) data model (Liebich, 2013) constitutes an open data exchange format formulated and upheld by buildingSMART, aimed at fostering interoperability across the Architecture, Engineering, and Construction (AEC) sector. It furnishes a unified data model facilitating seamless exchange of both geometric and semantic information without loss. As a vendor-neutral standard, IFC encompasses a broad array of building information representations, encompassing diverse geometry representations and a comprehensive set of semantic objects structured in an object-oriented fashion.

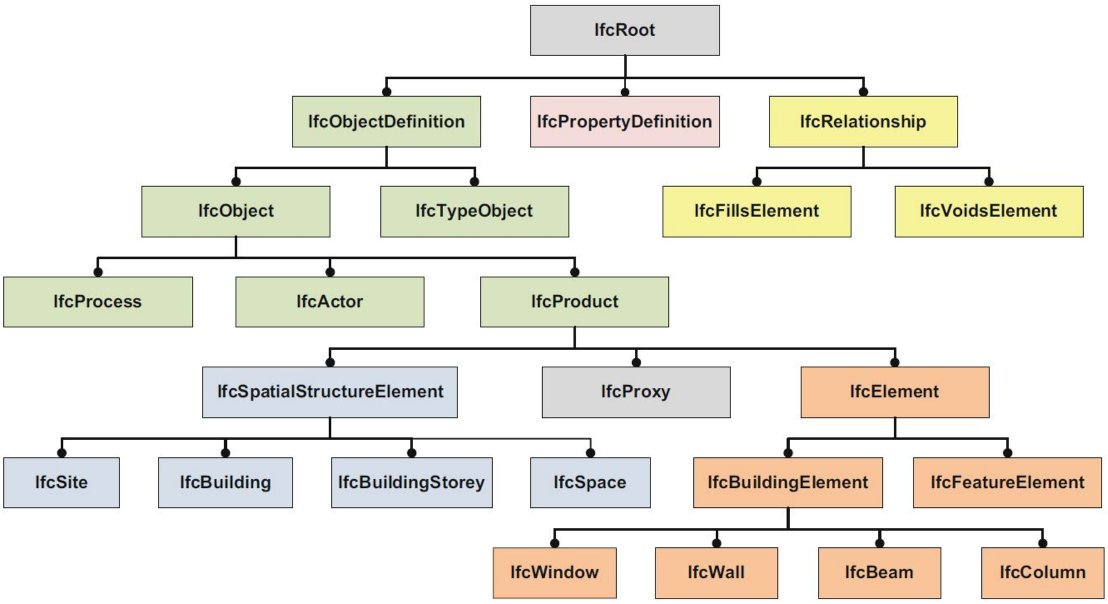



Figure 2.3: Part of the IFC data model showing the most important entities in the upper layers of the inheritance hierarchy according to (Borrmann et al., 2018).

Since its inception in 2009, the Green Building Extensible Markup Language (gbXML) exchange format has emerged as a publicly accessible, non-commercial schema, primarily oriented towards exchanging building information relevant to operational energy simulations (Green Building Foundation, 2021). Initially developed by Green Building Studio and acquired by Autodesk, Green Building Extensible Markup Language (gbXML) lacks official standardization oversight. The gbXML schema does not include a complete BIM model but rather encapsulates relevant environmental and geometric data of buildings. This derived data model based on the BIM model is often called a Building Energy Modeling (BEM). The schema features a "campus" container that houses one or multiple buildings, each delineated by a closed building envelope described by surfaces. These surfaces are characterized by type specifications (e.g., "InteriorWall"), Boundary Representation (B-Rep) geometry, references to adjacent spaces linked to zones, and designated openings.

The gbXML format finds application in Life Cycle Assessment (LCA) during early design phases, for instance, through utilization in CAALA software (CAALA, 2024), or Building Energy Performance Simulation (BEPS), e.g., using Honeybee from Ladybugtools (Sadeghipour Roudsari & Pak, 2013), facilitating consideration of both embodied and operational emissions. However, the schema lacks a detailed representation of specific element layers and materials, rendering it unsuitable for precise alignment with environmental datasets at the material level.

Recently, the initiative opensource.construction and Christian Kongsgaard published an open data schema for LCA, called LCAx (Kongsgaard, 2024). It aims to establish an open-source, machine, and human-readable data format for exchanging LCA results based on JavaScript Object Notation (JSON) format. However, the input data does not necessarily need to be derived from BIM models.

### 2.1.4 Semantic enrichment of Building Information Models

Many advanced techniques are available for automating the augmentation of semantic content for deriving different BIM use cases. Bloch's analysis provided an overview of various strategies, methodologies, and application domains for enriching the semantic content of BIM (Bloch, 2022). Two primary avenues were identified: leveraging IFC to represent building information coupled with inference-driven enhancement and integrating IFC with external data sources.

Additionally, Semantic Web technologies were utilized to process building information, focusing on building design, performance assessment, and notably, energy simulations (Scherer & Schapke, 2011). Exemplary cases included using Natural Language Processing (NLP) to classify spatial elements for Korean school buildings and deriving code compliance regulations through rule extraction (Guo et al., 2021).

Costa and Sicilia employed semantic query languages to automatically convert BIM data, concentrating on harmonized data models to facilitate building-scale energy simulations utilizing EnergyPlus (Costa & Sicilia, 2020). In another approach, Baumgärtel et al. utilized ontologies to dynamically modify and assess thermal energy performance in building contexts (Baumgärtel & Scherer, 2016).

Generally, the majority of semantic methodologies leverage ontologies, Semantic Web tools, and linked data concepts to achieve automated semantic enhancement in the BIM domain. While some incorporate NLP, its application for enhancing detailed insights from BIM models for energy simulations remains relatively limited.

In Sections 2.2.1 to 2.2.3, we focus more on current manual semantic enrichment processes for different BIM-based environmental analyses.

## 2.2 Environmental analyses using BIM

To assess the ecological dimension of sustainability of building designs, different environmental analyses are conducted, such as Building Energy Performance Simulation (BEPS), Life Cycle Assessment (LCA), or Building Circularity Assessment (BCA), for which MP constitute the basis. BuildingSMART International lists several open BIM use cases related to sustainable building design and environmental assessments, such as daylight analysis, thermal comfort and energy simulations, life cycle assessments, and material passports (buildingSMART International, 2024).

In the following, the current State of the Art using BIM for use cases of LCA in Chapter 2.2.1, Material Passport (MP) in Chapter 2.2.2, and BEPS in Chapter 2.2.3 is introduced.

### 2.2.1 BIM-based Life Cycle Assessment

The method of a Life Cycle Assessment (LCA) is generally standardized in ISO norms (DIN EN ISO 14040, 2021; DIN EN ISO 14044, 2021), and is divided into four phases: (1) Goal and Scope, (2) Life Cycle Inventory (LCI), (3) Life Cycle Impact Assessment (LCIA), and (4) life cycle interpretation. The European norms specifically for buildings divide building into three main life cycle phases: (A) Production and erection phase, (B) Use Phase, (C) End of Life Cycle, followed by model (D) including benefits and liabilities outside of the system boundaries (DIN EN 15804, 2022; DIN EN 15978, 2012)

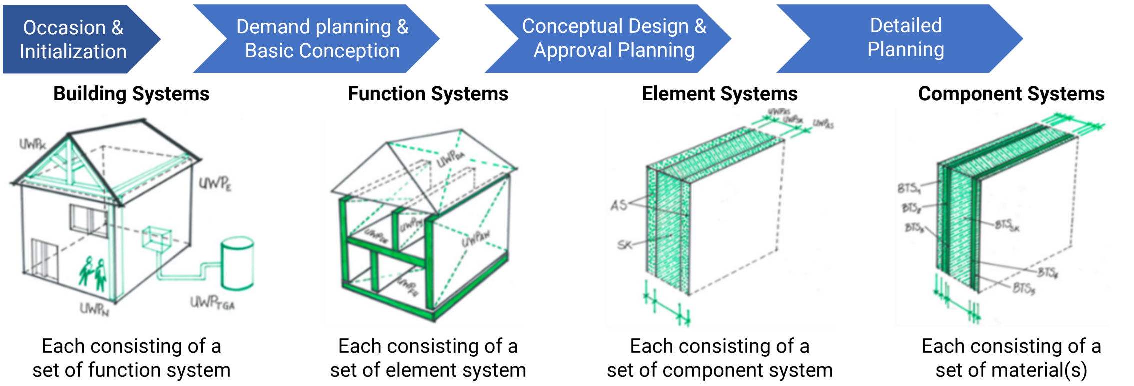



Figure 2.4: Phase-specific specification of the planning object as a reference system for the Life Cycle Assessment (LCA) (Horn et al., 2020).

The field of BIM-based LCA has gained increasing attention in research within the last decade. Horn et al. showed different scopes of decision-making using Building Information Modeling (BIM) for LCA in different design stages, as shown in Figure 2.4 (Horn et al., 2020). They showed the correlation of different system levels, such as building, function, element, and component systems, and the related design stage from very early (occasion & initialization) to detailed planning. The main focus of this dissertation is on early design stages, such as conceptual design stages following LOD 200-300, starting with element systems but also taking component systems and materials into account.

In the following, three main aspects of this research field relevant to this dissertation are highlighted, such as BIM-LCA integration strategies, matching and enriching LCA datasets to BIM models, as well as visualization for LCA results and uncertainty using BIM.

#### BIM-LCA integration startegies

Wastiels and Decuypere introduced a classification of five strategies on how to integrate BIM and LCA, as shown in Figure 2.5 (Wastiels & Decuypere, 2019). As part of the BBSR research project "Digital Twin Footprint," these integration strategies were analyzed towards its data exchange losses, degree of automation, suitability for practical use, and more (Bahlau et al., 2024). Open BIM workflows have bigger advantages for complex building designs with several planers involved using the federated model approach. However, the enrichment of LCA datasets with IFC elements or materials must still be done manually and semi-automated. With current software tools, a semi-automated BIM-LCA process is also relevant to ensure reliable and holistic LCA (Forth et al., 2019).

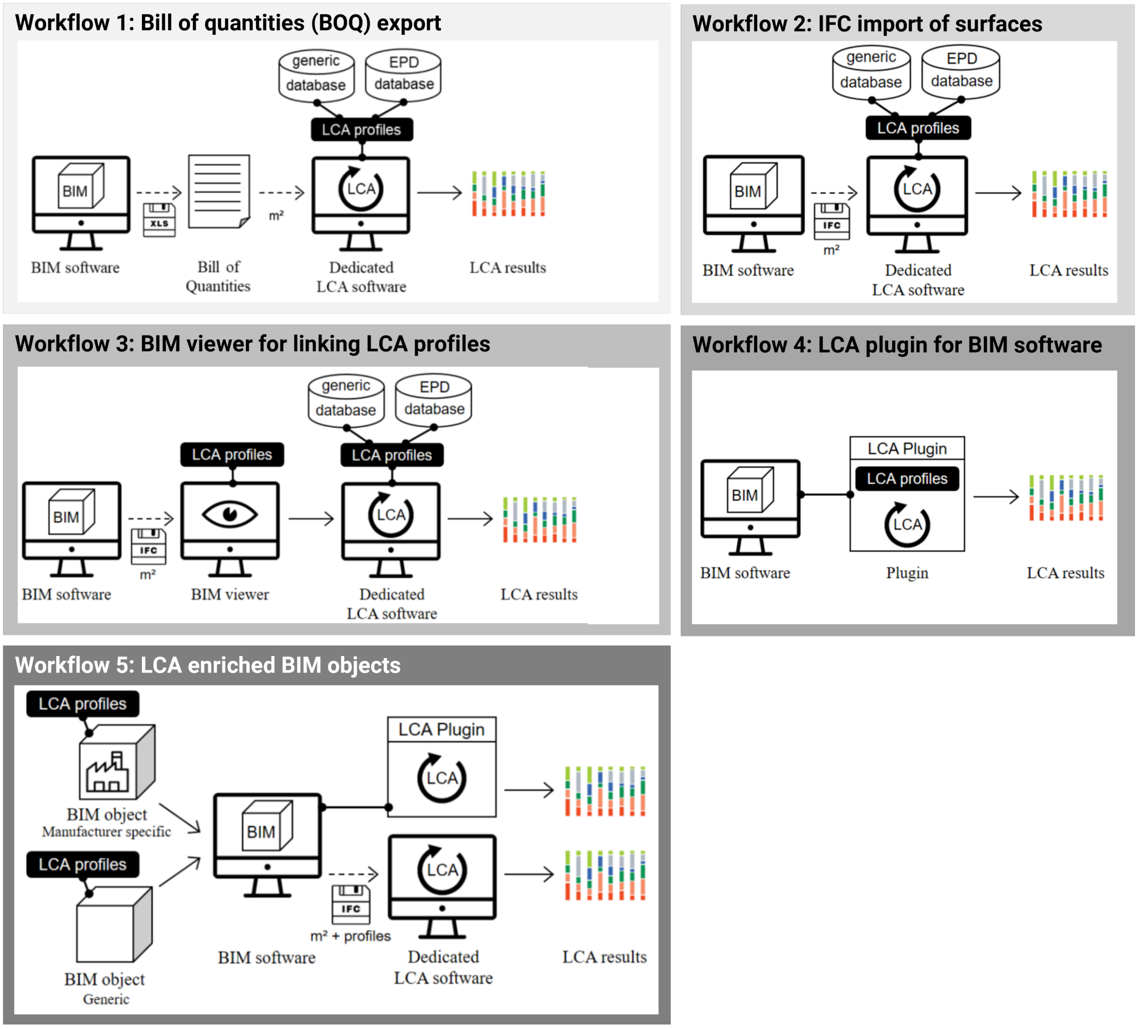



Figure 2.5: BIM-LCA integration strategies according to (Wastiels & Decuypere, 2019).

In Section 3.3, an initial literature review on BIM-based LCA was conducted. This section discusses the most recent literature reviews and research gaps within the last few years. Tam et al. reviewed open research gaps and future perspectives in the field of BIM-LCA integration under the framework of ISO 14040 (Tam et al., 2022). They clustered the eight unaddressed issues and eleven future perspectives in five topics:

1. Data collection at the Life Cycle Inventory (LCI) phase, such as ranges of LCA results in early stages,
2. Data mapping at the LCI phase, such as hierarchical structures of LCA data,
3. Data exchange at the LCI phase,
4. Presenting environmental impacts of buildings at the LCIA phase, such as 3D visualization of LCA results in BIM, and
5. Research topics of BIM-LCA integration at the interpretation phase.

Fonseca Arenas and Shafique reviewed the recent progress on BIM-based LCA and BEM integration and identified several research gaps and future directions (Fonseca Arenas & Shafique, 2023). One direction considered the usage of Artificial Intelligence (AI) for BIM-LCA integration to easier interact with BIM models, e.g., using a voice assistant interface. However, they mainly understand BIM-based LCA in a closed BIM workflow, neglecting open BIM.

#### Matching and enriching LCA datasets to BIM models

Chen et al. discussed three major topics of future prospects of BIM-LCA integration, such as linking BIM to dynamic LCA, automated data linking of BIM to LCA data, and combining further digital methods, such as semantic web or Geographic Information Systems (GIS) (Chen et al., 2024).

Current approaches of matching LCA datasets to IFC models often use manual linkage of Universally Unique Identifier (UUID) from external LCA databases and store these as Property Set (Pset) in the IFC models (Theißen, Drzymalla, et al., 2020). BuildingSMART International provides the Pset attribute "PSetEnvironmentalImpactIndicators" and "PSetEnvironmentalImpactValues" according to (BuildingSMART International Limited, 2020).

Parece et al. proposed an approach for automatically mapping LCA data to BIM objects using a construction classification system SECClasS (Parece et al., 2024). However, the approach follows the closed BIM approach, or the classifications need to be manually assigned for each BIM object.

Another matching approach of matching LCA datasets to IFC materials has been introduced by Reitschmidt using Ökobaudat based on tokenization of material names and a distinct matching or via Levenshtein distance (Reitschmidt, 2015). However, this linking approach and the one by Reitschmidt is only applied for holistic LCA if all materials are modeled in detail in the BIM model, which is the usually case for detailed design stages.

#### Visualization of LCA results and uncertainty using BIM

In their review of LCA result visualization, Hollberg et al. evaluate current practices and offer a thorough overview of various strategies and their potentials (Hollberg et al., 2021). This overview categorizes different visualization strategies based on LCA goals and the amount of information conveyed, as shown in Figure 2.6.

3D model representations of buildings, such as BIM, are used to identify hotspots of environmental impact with little information. Several researchers have implemented this 3D-model-based visualization strategy in recent years (Mousa et al., 2016; Naneva, 2022; Röck et al., 2018a, 2018b; Tsikos & Negendahl, 2017). These methods predominantly use color coding within authoring tools to visually represent the final LCA outcomes. Kiss and Szalay employ a distinct visualization technique for detailed LCA analysis,

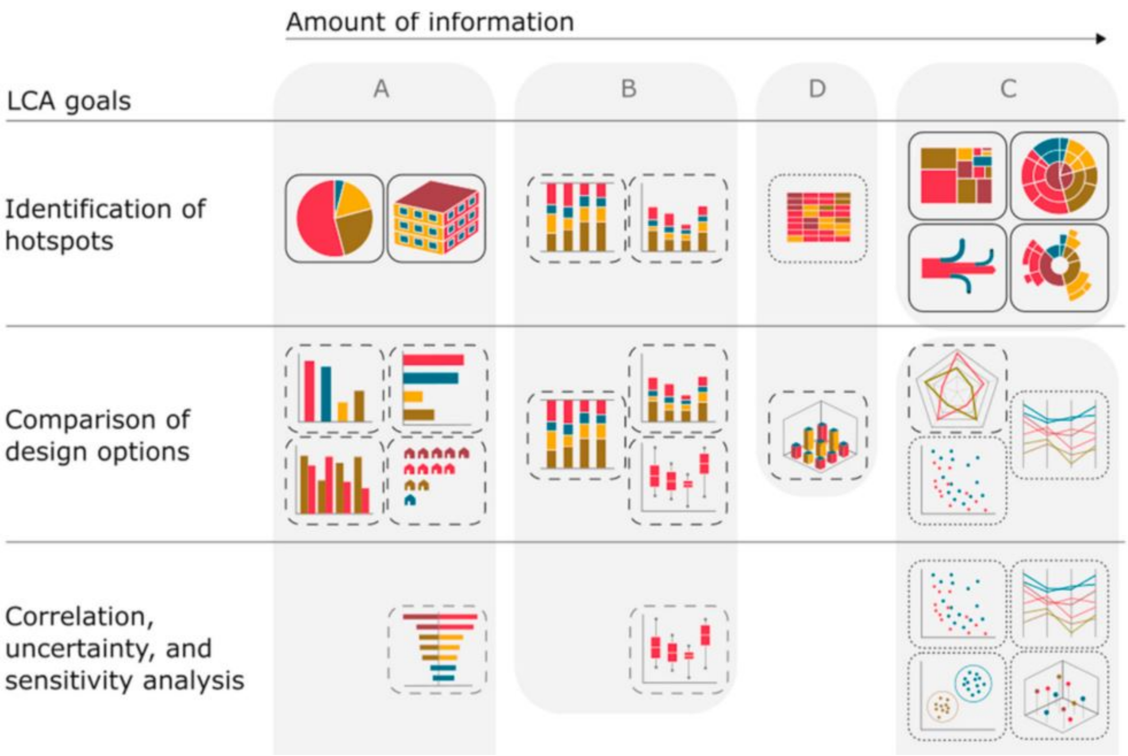



Figure 2.6: Detail of Hollberg et al.'s synthesis of the LCA goals, the group of visualization types, and the amount of information displayed in the visualization according to (Hollberg et al., 2021).

combining model-based color coding with a sunburst diagram to highlight specific aspects of the results. Their implementation leverages Rhino and Grasshopper for enhanced visualization capabilities. However, none of the mentioned approaches has tested its effect and intuitiveness for users, especially non-LCA experts.

Another visualization strategy for identifying hotspots is heat maps, which was also implemented by several researchers (Cer et al., 2017; Eberhardt et al., 2019; Goossens et al., 2018; Kiss et al., 2020; Vuarnoz & Jusselme, 2018). However, Hollberg et al. identified that heat maps are mainly used by LCA experts in detailed design phases, not in early ones (Hollberg et al., 2021). Box plot diagrams are widely used to include the LCA goal of comparing design options and uncertainties of the LCA results. These were found to be also used in early design stages and also by building design professionals and decision-makers, but less in existing LCA tools.

Marsh et al. reviewed uncertainties in LCA for the built environment, identifying several sources, clustered by the LCA phases of Goal & Scope, LCI, and Life Cycle Impact Assessment (LCIA) (Marsh et al., 2023). They also pointed out barriers such as data quality, human error, practitioner expertise, carbon data comparability, data availability, unknown early-stage material specifications, and time requirements.

Ströbele introduced a fuzzy LCA (fLCA) approach that manages vagueness using distribution curves instead of singular outcomes (Ströbele, 2022). Schneider-Marin et al. created the EarlyData knowledge database to guide material choices during design stages with limited details, visualizing semantic uncertainty with box plots representing GWP ranges (Schneider-Marin et al., 2022).

To address uncertainties in BIM models, Abualdenien and Borrmann proposed several methods for visualizing geometric and semantic uncertainties in building elements during early design stages (Abualdenien & Borrmann, 2020). They concluded that using a combination of color value and transparency to quantify semantic reliability yielded a high degree of intuitiveness and acceptance.

However, there are currently no visualization approaches for LCA results, including uncertainties, using BIM models in early design stages, which have been tested and evaluated by non-LCA experts.

### 2.2.2 BIM-based Material Passports

To tackle the current challenges of high waste generation and resource demand of the AEC industry, the concept of circular economy is been used to close resource cycles and material flows. However, material-related information is insufficiently documented to realize the reuse and recycling of building components and materials. Therefore, the concept of BCA has been introduced, such as Design for Disassemble (Elma, 2006), the Material Circularity Index (Ellen MacArthur Foundation, 2015), or the Urban Mining Index (Rosen, 2021). The German Sustainable Building Council (DGNB) agreed on a standard for circularity indices for buildings distinguishing between today's contributions and future contributions (Braune & Wellstein, 2024). They count material origin, construction, and demolition waste as today's contribution, while material compatibility, disassembly capability, detachability, and material utilization are future contributions.

| | Digital Product Passport | Material Passport | Digital Building Logbook |
|---|---|---|---|
| Scale | Product | Area, complex, building, element, product, material, raw material | Building |
| Industry | Cross-industry | (mainly) built environment | Built environment |
| Regulation | EU Ecodesign Directive | - | EU-wide framework for a digital building logbook |

Figure 2.7: Differences and similarities between digital product passports, material passports, and digital building logbooks according to (Çetin, 2023).

Several similar approaches exist for digitizing this BCA, such as Material Passport (MP), Digital Product Passports (DPP), circularity passports, building renovation passports, or (Building) Ressource Passports (Çetin, 2023; Honic et al., 2024). The main difference is that DPPs are used for any products in any industry (European Commission, 2022), while MP (BAMB, 2019) and Digital Building Logbooks (Volt et al., 2020) were specifically developed for buildings, as shown in Figure 2.7. Building Ressource Passports, in German "Gebäuderessourcenpass" (DGNB, 2024), is often used as synonyms for MP or Digital Building Logbooks in the German market for the above-mentioned concepts. The digital Building Logbook can align with MP but extends the approach beyond circularity aspects with other sustainability aspects, such as energy performance or renovation history (Çetin, 2023).

Several researchers investigated the potential of using different digital methods, such as BIM, Internet of Things, or Digital Twins, for circularity assessments and MPs (Çetin et al., 2021; Dervishaj & Gudmundsson, 2024). Wolf et al. introduced an overview of a circular built environment in the digital age, covering business and governance aspects, design and fabrication, as well as data-related aspects, such as GIS, Scan-to-BIM, AI, and BIM for MP (de Wolf et al., 2024).

Using BIM for MP has been introduced to document all relevant information about BCA, LCA, and others (Heinrich & Lang, 2019). Different approaches have been introduced using authoring tools (Atta et al., 2021; Honic et al., 2019a), or following open BIM workflows (Tomczak et al., 2024).

Tomczak et al. analyzed the IFC schema by developing IDS for circularity information (Tomczak et al., 2024). They discovered that information related to disassembly instructions, End-of-Life predictions, and connections between components are difficult to incorporate in IFC. Sanchez et al. developed a BIM-based Semantic Enrichment Engine for Disassembly Planning (SEEDP) using open BIM standards, such as IFC, Information Delivery Manual (IDM), and Model View Definition (MVD) for disassembly planning (Sanchez et al., 2024). The relevant semantic information needs to be manually inputted but afterward automatically enriched.

Currently, a few software providers have implemented tools for open BIM-based circularity assessments, such as Madaster (Madaster, 2024), or circularity consultancies, such as EPEA (Gebetsroither et al., 2024).

However, current approaches have not solved the gap in automated semantic enrichment of relevant information, such as material information from external databases. To assess open BIM models in early design stages, they still need to be manually enriched, which makes iterative optimization and decision-support in early design stages costly and time-expensive.

### 2.2.3 BIM-based Building Energy Simulations

The term Building Energy Performance Simulation (BEPS) is commonly used interchangeably with Building Energy Simulation (BES), Building Performance Simulation (BPS), Building Energy Modeling (BEM), or simply energy simulation (Hong et al., 2018). This overarching designation encompasses simulations primarily focused on energy demand and indoor environmental factors such as thermal comfort, both based on a BEM model.

The geometry representations and semantic information are the main differences between BIM and BEM models. Consequently, one main challenge of transforming BIM models to BEM models is the correct geometric transformation from volumetric to a watertight surface representation without any gaps or holes. Furthermore, semantic information in BIM models are defined use-case specific, while the relevant information for BEM models are specified for the use case of energy-related simulations.

Eckstädt et al. conducted a comparative study examining three distinct methodologies for conducting whole-building energy simulations, often termed building performance simulations (Eckstädt et al., 2022). Their investigation centered on utilizing open BIM models and the IFC data format as the principal input file. However, challenges persist within existing tools, stemming from the need for accurate IFC export settings tailored to individual simulation tools, limitations in the IFC import process, and inherent constraints of the simulation tools themselves.

Van Treeck et al. discussed different data exchange formats for BIM-based BEPS, such as using MVD with IFC or gbXML (van Treeck et al., 2018). Löhr et al., on a different note, proposed a partially automated procedure for generating multi-zone thermal models from IFC models (Löhr et al., 2022).

Ramaji et al. introduced an alternative approach for converting IFC-based BIM into BEM, directly converting IFC models into OpenStudio's native IDD format (Ramaji et al., 2020). Despite encountering challenges during conversion, they addressed these issues using MVD. Similarly, Spielhaupter conducted a comparative analysis of various IFC-based strategies for transforming BIM to BEM (Spielhaupter, 2021).

Yang et al. adopted IFC files as the foundation of their methodology, albeit with a different approach wherein they initially converted IFC data into the gbXML format before further transformation into the IDF format, which is the native file by EnergyPlus (Y. Yang et al., 2022). However, their workflow indicates the frequent necessity of additional adjustments, prompting the proposed strategy to employ the Honeybee JSON (HBJSON) format as the transformation schema due to its open-source nature, adaptability across different file formats, and enhanced geometric export reliability.

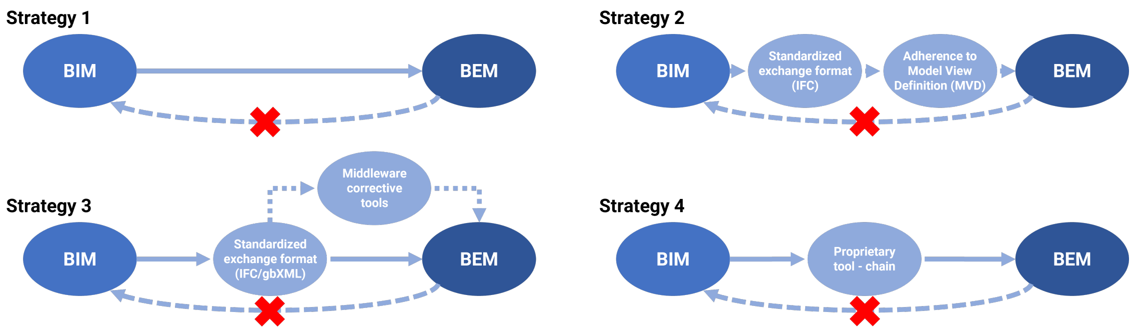



Figure 2.8: Comparison of different BIM-BEM interoperability strategies according to (Ciccozzi et al., 2023).

Ciccozzi et al. reviewed interoperability strategies in BIM to BEM workflows, identifying four key approaches, as shown in Figure 2.8 (Ciccozzi et al., 2023): real-time connection, standardized exchange formats, and middleware tools, MVD, and proprietary toolchains. They also analyzed various methods of automatically mapping energy-related information to the BIM model.

Following over 15 years of research in the intersection of BIM and BEM, this research domain persists in confronting unresolved challenges. Gao et al. highlighted the automated conversion of intricate spatial functions across all rooms as a critical area for future exploration, given the ongoing manual nature of the current process (Gao et al., 2019).

Di Biccari et al. reviewed the state-of-the-art and research trends in BIM and BEPS interoperability (Di Biccari et al., 2022). They emphasized the necessity for research to propose practical solutions for describing occupancy and MEP component schedules in BIM. Additionally, they noted that despite the availability of thermal properties in authoring tools, manual mapping is still required during the IFC export stage as part of post-processing.

Wang et al. proposed a method using PLM for transforming BIM to BEM matching metamodels of different types of BEMs (Z. Wang et al., 2024). They compare the matching accuracy of the PLM LSTM with their fine-tuned LLM T5-small (Colin Raffel et al., 2020) using on collected metamodel pairs for training. However, they recommend integrating constraints and rules for instance-level transformation by filtering the instances. Furthermore, they do not use their approach to automatically semantically enrich BIM models for BEM.

In general, the current state-of-the-art still shows a research gap for robustly matching BEPS-related semantic information to BIM objects, such as space types of thermal properties of elements and materials, particularly in the early design stages when the BIM model contains ambiguous information.

## 2.3 Natural Language Processing and Large Language Models

Natural Language Processing (NLP) is a sub-domain of Artificial Intelligence (AI) and Deep Learning (DL). The field of NLP and the usage of Large Language Model (LLM) has seen notable research progress in recent years, showcasing advancements in efficiency and accessibility. Thereby, NLP is expanding its utility across various sectors, including the construction industry. In the following sections, current approaches of NLP in the AEC industry, the NLP subtask of Semantic Textual Similarity (STS), and, finally, domain adaptation and fine-tuning of LLM are introduced.

### 2.3.1 Natural Language Processing in the AEC industry

Locatelli et al. conducted a scientometric analysis exploring the synergies between Natural Language Processing (NLP) and Building Information Modeling (BIM) (Locatelli et al., 2021). In addition to the domain of Automatic Compliance Checking, they identified Information Retrieval from BIM models and Information Enrichment of BIM objects as further fields of significant application. Wang et al. developed a query-answering (QA) system for BIM information extraction (IE) using NLP techniques, achieving notable accuracy scores in their evaluation (N. Wang et al., 2022). Xie et al. proposed a method for associating real-world facilities with BIM elements utilizing NLP for word segmentation and keyword extraction, employing the LTP word segmentation module (Xie et al., 2019). Their matching method evaluates matching matrices based on HiTree paths, aiming for optimal alignment with natural language feature vectors.

Cornago et al. conducted a SWOT analysis employing Transformers for Life Cycle Assessment (LCA) studies, revealing internal strengths such as automation, integration support, and relatively low marginal costs (Cornago et al., 2023). However, concerns were raised regarding data quality, electricity intensity during model training, and rapid technological evolution. External opportunities encompass community-building and augmented data availability, while threats include regulatory gaps, standardization issues, and a scarcity of interdisciplinary expertise. Transformers promise to aid LCA practitioners by mitigating scalability challenges and enabling data-centric environmental decision-making support.

Zheng et al. investigated the utility of domain-specific corpora in augmenting deep learning and BERT-based models for Information Retrieval (IR) tasks within the AEC domain (Zheng et al., 2022). Their findings indicate that domain-specific corpora enhance traditional word embedding models for select tasks while detrimentally affecting others. Conversely, BERT-based models consistently outperform traditional approaches, culminating in the development of RegulatoryBERT, a highly effective model.

Wu et al. conducted an exhaustive review of NLP utilization in construction management, highlighting advancements in information extraction and document organization (C. Wu et al., 2022). They also deliberated on the potential and challenges of NLP applications in construction management, serving as a valuable reference for project teams interested in harnessing NLP techniques for intelligent construction practices.

One of the recent research trends in AI is combining Deep Learning techniques, like NLP or Computer Vision (CV), with symbolic reasoning (symbolic logic and knowledge representation), also known as Neuro-Symbolic Computing (NSC) or Neuro-Symbolic AI. Luo et al. state that NSC "has the potential to enable more robust, interpretable, and accurate AI systems in construction by harnessing the strengths of Deep Learning (DL) and symbolic reasoning" (Luo et al., 2023).

While the relevance of developing NLP methodologies for various tasks within the AEC domain is escalating, none of the discussed studies proposed an NLP-based enrichment process for BIM-based LCA.

### 2.3.2 Large Language Models

In recent years, language models have been the basis of NLP's research development. Large Language Model (LLM) have proven emergent abilities compared to small-scale models (Wei et al., 2022).

Ding et al. identified in their scientometric analysis three main phases of NLP in the construction industry: the germination stage (2000-2011), the gradual development stage (2012–2018), and the rapid development stage (2019–2020) (Ding et al., 2022). The first stage is characterized by a small number of models and expensive computing power. The most dominant techniques were Recurrent Neural Networks (RNN), as well as Long Short Term Memory Networks (LSTM) (Hochreiter & Schmidhuber, 1997). A more efficient variant of LSTM was realized by bi-directional training, meaning not only memorizing the previous words but also the subsequent ones (Bach et al., 2024).

In the second stage, DL and neural networks became the dominant technology (Krizhevsky et al., 2012). Open-source frameworks, such as Tensor Flow (2015) or Pytorch (2017), as well as improved computing power and graphics cards, supported this technology, enabling training on larger datasets (Wolber et al., 2024).

Furthermore, Vaswani et al. introduced transformer-based NLP models in 2017 (Vaswani et al., 2017). It is based on the self-attention mechanism, which makes it possible to learn contextual relationships between different parts of the input sequence (Bach et al., 2024).

The development and release of Bidirectional Encoder Representations from Transformers (BERT), as shown in Figure 2.9, by Google in 2019 accelerated NLP research and applications in the AEC industry (Devlin et al., 2018). BERT is trained for multiple tasks and on large datasets.

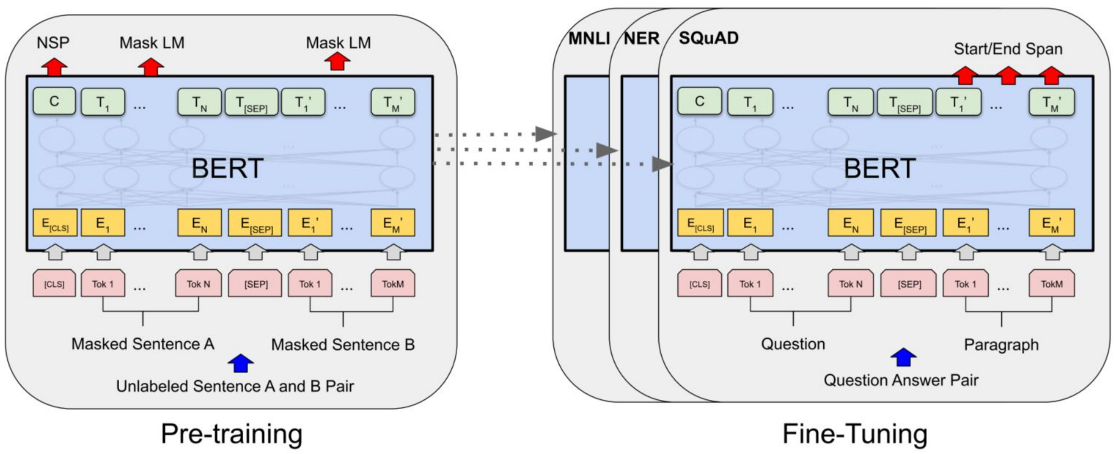



Figure 2.9: Overall pre-training and fine-tuning procedures for BERT (Devlin et al., 2018).

In 2022, the release of openAI's ChatGPT enabled an end-to-end solution, which is easy to use by its Chat interface (OpenAI, 2022). However, there are different LLMs specialized in different NLP tasks, such as 'text-embedding-ada-002' using embeddings for STS tasks, but these models are not free of use.

### 2.3.3 Semantic Textual Similarity

The domain of NLP encompasses various functionalities, including text summarization, text classification, named entity recognition, and sentiment analysis. This section primarily focuses on introducing the foundational principles underlying Semantic Textual Similarity (STS), a key task within NLP, as shown in Figure 2.10.

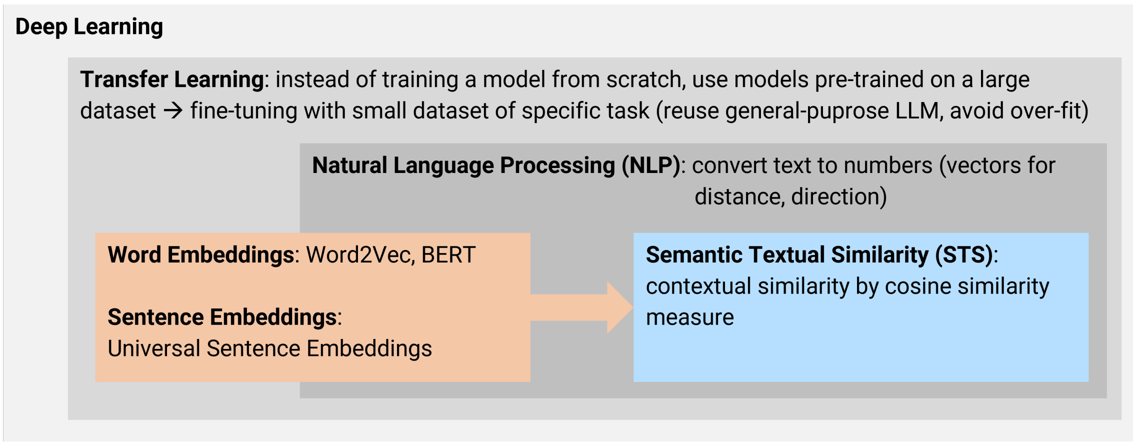



Figure 2.10: Overview of Deep Learning, Natural Language Processing, and Semantic Textual Similarity according to (Park, 2019).

One of the big advantages of NLP and STS is that it uses transfer learning, one feature of DL. This allows training specific tasks on small datasets as the model is not trained ab initio, but pre-trained LLM are used (Park, 2019). The core methodology for this task involves sentence pair modeling, a technique also utilized in Natural Language Inference (NLI) or Recognizing Textual Entailment (RTE) (Lan & Xu, 2018). Achieving STS involves discerning the semantic relatedness between a statement and its associated premise, also known as semantic similarity (Bowman et al., 2015).

Chandrasekaran and Mago comprehensively analyzed semantic similarity's evolution, categorizing approaches into knowledge-based (such as lexical-semantic nets), corpus-based, deep neural network-based (such as transformer-based LLM), and hybrid methods (Chandrasekaran & Mago, 2022). Each approach presents distinct merits and drawbacks, with discernible trends favoring the development of embeddings and transformer models imbued with greater semantic understanding. Corpus-based methods predominantly employ cosine similarity to gauge the disparity between word vectors. However, alternative metrics such as Euclidean or Manhattan distance are also employed in STS (R. Li et al., 2023).

Corpus-based semantic similarity methodologies leverage word or sentence embeddings, which encode vector representations of words, encapsulating linguistic associations (Tobias Schnabel et al., 2015). Word embeddings capture individual word semantics, while sentence embeddings encapsulate entire sentence semantics. Widely adopted pre-trained word embeddings include Word2Vec (Mikolov et al., 2013) and BERT (Devlin et al., 2018). BERT, comprising pre-training and fine-tuning stages, also accommodates sentence embeddings. Reimers and Gurevych introduced Sentence-BERT (SBERT), also known as sentence transformers, a refined iteration of BERT employing Siamese and triplet network architectures to derive semantically meaningful sentence embeddings, facilitating cosine-similarity comparisons and excelling in STS tasks (Reimers & Gurevych, 2019).

STS tasks typically necessitate substantial datasets for training, fine-tuning, and evaluation. The Stanford Natural Language Interference (SNLI) corpus, comprising 570k human-authored English sentence pairs, is a commonly utilized dataset for this task (Bowman et al., 2015). The STS benchmark serves as a standard evaluation metric for validating STS tasks (Cer et al., 2017).

In the following subsections, different approaches of measuring semantic similarities are introduced and exemplarily applied using the word pairs of "masonry" and "brick", as well as "masonry" and "concrete".

#### Synset similarity using Lexical-Semantic Nets

Before the development of LLM using embeddings, language models represented in knowledge graphs were used for STS tasks. Most known Lexical-Semantic Net are WordNet in English (Fellbaum, 1998), or GermaNET in German (Hamp & Feldweg, 1997; Henrich & Hinirchs, 2010). Vossen introduced a multilingual Lexical-Semantic Net, called EuroWordNet, first including Dutch, Italian, Spanish and English language (Vossen, 1998).

These Lexical-Semantic Nets semantically relate nouns, verbs, and adjectives by grouping lexical units that express the same concept into synsets and defining semantic relations between these synsets, which stands for sets of synonyms. This is depicted as a graph, where the nodes are synsets and the edges represent semantic relations (Navigli & Martelli, 2019). Figure 2.11 shows the synset relations between the example words "brick", "masonry" and "concrete".

Semantic similarity is often measured by the shortest path similarity, which is the inverse of the shortest path length between two synsets (Rada et al., 1989). Other path-related similarity measures include Wu-Palmer similarity (Z. Wu & Palmer, 1994), Leacock-Chodorow similarity (Leacock & Chodorow, 1998), Resnik similarity (Resnik, 1995), Lin similarity (Lin, 1998), and Jiang-Conrath similarity (Jiang & Conrath, 1997).

| Similarity measure | Score (masonry-brick) | Score (masonry-concrete) |
|---|---|---|
| Shortest path | 0.200 | 0.200 |
| Wu-Palmer | 0.714 | 0.714 |
| Leacock-Chodorow | 2.028 | 2.028 |
| Resnik | 2.305 | 2.305 |
| Lin | 0.196 | 0.200 |
| Jiang-Conrath | 0.053 | 0.054 |

Table 2.1: Different similarity measures between the words "masonry" and "brick", as well as "masonry" and "concrete" using WordNet

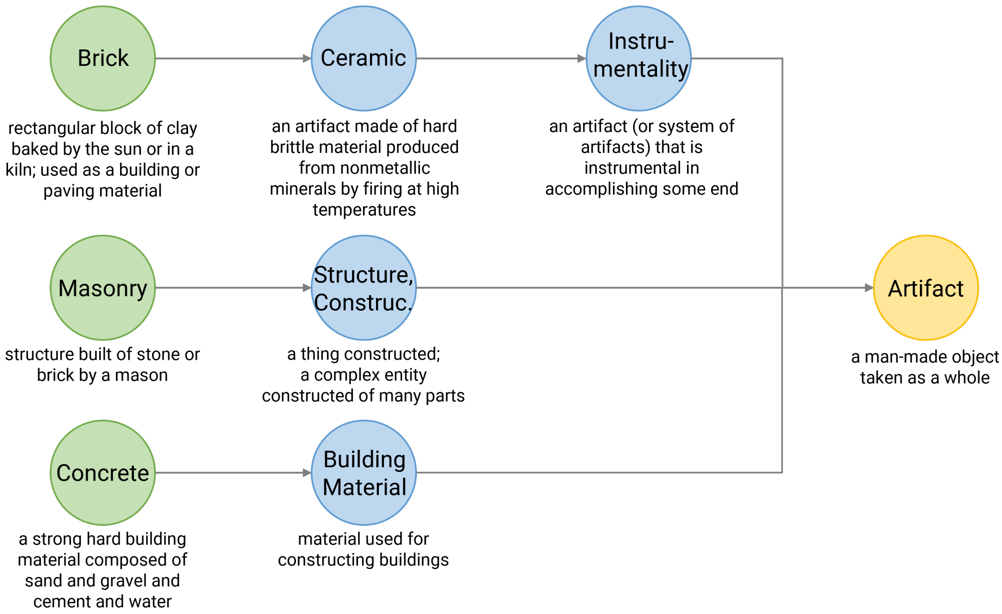



Figure 2.11: Relatedness of synset "masonry" and "brick", as well as "masonry" and "concrete" using WordNet (Fellbaum, 1998).

Table 2.1 shows the different similarity measures of shortest path, Wu-Palmer, Leacock-Chodorow, Resnik, Lin, and Jiang-Conrath similarity between the example word pairs using WordNet, as shown in Figure 2.11. Except for minor deviations in Lin and Jiang-Conrath similarity, there are no differences in the scores of the word pair "masonry"-"brick" and "masonry"-"concrete".

### Semantic similarity using Levenshtein distance

There are several other approaches to measure string-based similarities. Ghomaa and Fahmy differentiate between character-based and term-based similarity measures, also known as token-based similarity measures (Gomaa & Fahmy, 2013). One of the most used character-based similarity measures is the Levenshtein distance, which follows the idea of an edit-based algorithm (Levenshtein, 1965). It defines the distance between two strings $a$ and $b$, as shown in Equation 2.1. The distance is measured by counting the minimum number of single-character allowed operations needed to transform one string into the other, such as insertions, deletions, or substitutions.

$$lev(a,b) = \begin{cases} |a| & \text{if} |b| = 0 \\ |b| & \text{if} |a| = 0 \\ lev(tail(a), tail(b)) & \text{if} |a| = |b| \\ 1 + min \begin{cases} lev(tail(a), b) \\ lev(a, tail(b)) \\ lev(tail(a), tail(b)) \end{cases} & \text{otherwise.} \end{cases} \tag{2.1}$$

| | | b | r | i | c | k |
|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 |
| m | 1 | 1 | 2 | 3 | 4 | 5 |
| a | 2 | 2 | 2 | 3 | 4 | 5 |
| s | 3 | 3 | 3 | 3 | 4 | 5 |
| o | 4 | 4 | 4 | 4 | 4 | 5 |
| n | 5 | 5 | 5 | 5 | 5 | 5 |
| a | 6 | 6 | 6 | 6 | 6 | 6 |
| r | 7 | 7 | 6 | 7 | 7 | 7 |
| y | 8 | 8 | 7 | 7 | 8 | 8 |

| | | c | o | n | c | r | e | t | e |
|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| m | 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| a | 2 | 2 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| s | 3 | 3 | 3 | 3 | 4 | 5 | 6 | 7 | 8 |
| o | 4 | 4 | 3 | 4 | 4 | 5 | 6 | 7 | 8 |
| n | 5 | 5 | 4 | 3 | 4 | 5 | 6 | 7 | 8 |
| a | 6 | 6 | 5 | 4 | 4 | 5 | 6 | 7 | 8 |
| r | 7 | 7 | 6 | 5 | 5 | 4 | 5 | 6 | 7 |
| y | 8 | 8 | 7 | 6 | 6 | 5 | 5 | 6 | 7 |

Table 2.2: Levenshtein distance between the words "masonry" and "brick", as well as "masonry" and "concrete"

As shown in Table 2.2, the Levenshtein distance between the word "masonry" and "concrete" is smaller, scoring 7, compared to the Levenshtein distance between "masonry" and "brick", resulting in a distance of 8 operations.

### Cosine similarity using transformer-based LLM

Cosine distance or cosine similarity is a well-used term-based similarity measure for corpus-based STS, and it calculates the cosine of the angle between two encoded word or expression vectors. Every string is converted from text to a vector representation, also known as embeddings, to measure semantic similarity by encoding the text using the weights of the LLM. In this case, a vector is a list of numerical values, and their combination represents the overall meaning (Wilbur & Sirotkin, 1992). Afterward, the cosine similarity ($cos(\theta)$) between two different vectors, $A$ and $B$, can be calculated using the cosine similarity, as shown in Equation 2.2, while $n$ is the dimension of the vector:

$$cosine - similarity := cos(\theta) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{2.2}$$

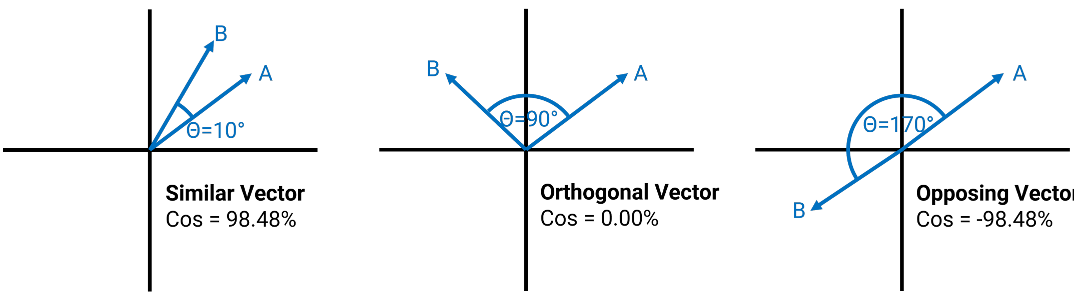



Figure 2.12: Schematic overview of cosine similarities using two dimensional vectors

Figure 2.12 shows schematically the similarity of two-dimensional vectors using cosine similarity. The closer the vectors point in the same direction, the higher the cosine similarity is with a maximum of 100%. The further the vectors point in the opposite directions, the smaller the cosine similarity is with a minimum at -100%. When the vector is orthogonal, the cosine similarity is 0%. However, corpus-based embeddings have more dimensions, BERT has 768 dimensions, but all vectors are positive, so the cosine similarity is always between 0% and 100%. The cosine similarity between the words "concrete" and "cement" is 94.72%, while between "concrete" and "brick" it's lower with 83.33% using BERT ('bert-base-uncased') (Devlin et al., 2018).

### 2.3.4 Domain adaptation and LLM fine-tuning

Given that most LLM are trained on generic text, they may not always align optimally with domain-specific tasks, necessitating domain adaptation. Typically, domain adaptation involves fine-tuning a PLM on a domain-specific dataset (Kohle & Jannidis, 2020). This fine-tuning process entails adjusting the weights of the original model to better accommodate the specific characteristics of the domain data and the targeted task. The Huggingface platform offers numerous PLMs in various languages, as well as multilingual LLMs that are fine-tuned for domain-specific tasks (Wolf et al., 2019).

Selecting the appropriate loss function is crucial for fine-tuning PLM, depending on the training data and the overall task at hand. For instance, to enhance the performance of fine-tuning BERT in multitask domains like sentiment analysis, paraphrase detection, and STS, Jadwin and Huang utilized in-domain pre-training and Multiple Negatives Ranking Loss (MNRL) (Jadwin & Huang, 2023). They found that MNRL fine-tuning had the most significant impact on performance optimization.

Contrastive Loss (ContrastiveL), proposed by Hadsel et al., adjusts the distance between two embeddings based on labels (0 or 1) (Hadsell et al., 2006). Cosine Similarity Loss (CosL) utilizes manual labels indicating the expected cosine similarity between two embeddings for fine-tuning LLMs, usually 0.8 for high similarity and 0.3 for contradicting word pairs. Reimers and Gurevych employed either Softmax classifier (classification objective function) for fine-tuning on SNLI dataset or CosL (regression objective function) to compute similarity scores in their concept of sentence embeddings using Siamese BERT-Networks (Reimers & Gurevych, 2019), as shown in Figure 2.13. For fine-tuning multilingual LLMs, they utilized Mean Squared Error Loss (MSEL) to train the student model (Reimers & Gurevych, 2020).

In addition to different loss functions and fine-tuning frameworks, multilingual PLMs play a crucial role in fine-tuning. Conneau et al. introduced a multilingual masked LLM, "XLM-R," trained with data from 100 languages, including German and English, leveraging the strengths of multilingual XLM models and the monolingual RoBERTa model (Conneau et al., 2019). Feng et al. proposed a Language-agnostic BERT Sentence Embedding (LaBSE), focusing on multilingual sentence embeddings supporting 109 languages, including German and English (Feng et al., 2020). Reimers and Gurevych also introduced a framework using knowledge distillation to create multilingual sentence embeddings from monolingual ones (Reimers & Gurevych, 2020), offering versions of multilingual PLMs, such as "distiluse-base-multilingual-cased-v2," which supports over 50 languages, including German and English, while requiring fewer samples and lower hardware requirements for training.

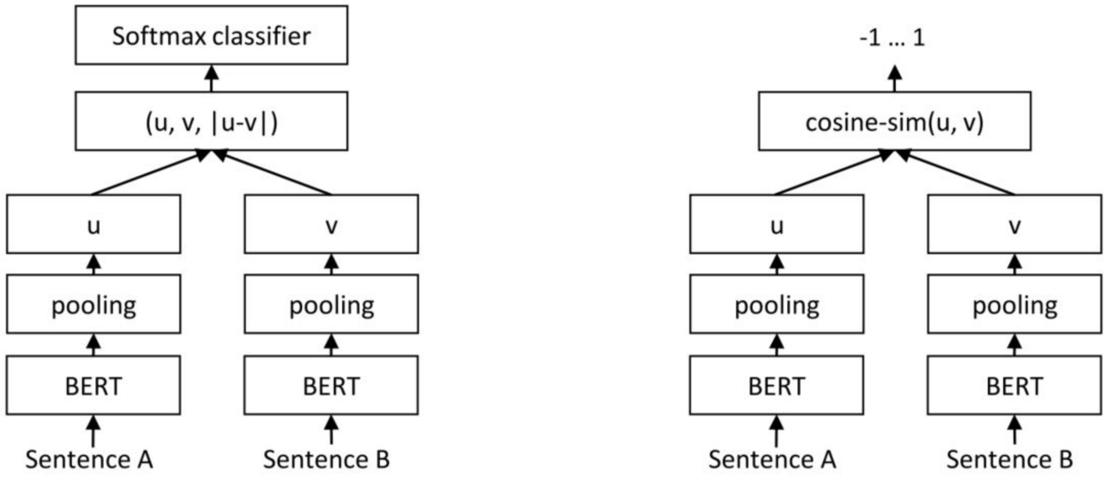



Figure 2.13: SBERT architectures with classification objective function for fine-tuning on SNLI dataset (left) vs. regression objective function to compute similarity scores (right) according to (Reimers & Gurevych, 2019).

Recent research projects showed the limitations of current LLM and the necessity of combining LLM to databases or knowledgebases (Suchanek & Holzenberger, 2024). Similar to the previously mentioned trend of neuro-symbolic AI, it aims to use structured domain knowledge and the advantages of LLM to increase the performance of domain-specific automation tasks.

## 2.4 Overall research gaps

Based on the previous chapters about the research background and related works, the identified research gaps within the scope of this dissertation are highlighted, which were previously introduced in Chapter 1.2:

1. **Automated and robust matching of LCA datasets to BIM data**: BIM model information and LCA datasets follow different structures and hierarchies. The same convention of hierarchy/ structure and naming needs to be addressed to link LCA data to BIM data to guarantee a correct matching (Potrč Obrecht et al., 2020). "Since the data structure and naming convention in LCA databases are fixed and hard to disaggregate, the data structure and naming convention of data from BIM models are often modified for mapping into LCA data during the data mapping process" (Tam et al., 2022).

   Furthermore, the same research gap also applies to matching material datasets to BIM elements for Material Passports. Real-world case studies shall be used to address a robust matching approach. Chapter 3 discusses this research gap for the use case of LCA, and in Chapter 4 for the use case of Material Passports.

2. **Ranges of LCA results in early design stages**: Tam et al. identified the information shortage of low LOD BIM models in early design stages as a challenge for reliable

LCA calculations (Tam et al., 2022). A range of LCA results should be applied instead of calculating specific LCA results for one uncertain element. This range can also be integrated into the design decision-making. This research gap will be addressed in Chapter 3.

3. **Automated and robust matching of space types to architectural rooms and thermal properties to BIM elements for BEPS**: To create Building Energy Models out of BIM, manual enrichment steps are still necessary. "The detailed space type of all spaces/rooms of a building should be automatically transformed from the BIM model, rather than by manual data setting" (Gao et al., 2019). Also, additional data regarding thermal properties must be manually enriched before a holistic energy simulation with reliable results is calculated (Raggi et al., 2021). These two research gaps of automated and robust enrichment of space types and thermal properties to BIM models for holistic BEPS are addressed in Chapter 5.

4. **Combination of Large Language Models and Domain Knowledge**: Current Large Language Models have limitations for different domain-specific tasks. Promising recent research areas are shown by neuro-symbolic AI (Luo et al., 2023) and combining knowledgebase and structured databases with LLM (Suchanek & Holzenberger, 2024) to automate domain-specific tasks, such as the previous matching of LCA, MP and BEPS related information to BIM models. This gap is addressed in the Chapters 3, 4 and 5.

5. **Visualization of environmental impacts for design-decision-making**: Tam et al. mentioned 3D visualization of environmental impacts in BIM models as another open BIM-LCA integration challenge (Tam et al., 2022). They call for more effort in this research direction "to make it more accessible to LCA practitioners to present the environmental assessment results in an intuitive and visualized way" (Tam et al., 2022). However, non-LCA experts shall also use these intuitive visualization strategies discussed in Chapter 6.

# Chapter 3

# Calculation of embodied GHG emissions in early building design stages using BIM and NLP-based semantic model healing

Previously published as: Forth, K.; Abualdenien, J.; Borrmann, A.: *Calculation of embodied GHG emissions in early building design stages using BIM and NLP-based semantic model healing*. Energy and Buildings 284, 2023, DOI: 10.1016/j .enbuild.2023.112837

## Abstract

To reach the goals of limiting global warming, the embodied greenhouse gas (GHG) emissions of new buildings need to be quantified and optimized in the very early design stages, during which design decisions significantly influence the success of projects in achieving their performance goals. Semantically rich building information models (BIM) enable to perform an automated quantity take-off of the relevant elements for calculating a whole building life cycle assessment (LCA). However, imprecise type and property information often found in today's BIM practice hinders a seamless processing for downstream applications. At the same time, the early design stages are characterized by high uncertainty due to the lack of information and knowledge, making a holistic and consistent LCA for supporting design decisions and optimizing performance challenging. In assessing this often vague information, it is essential to consider different levels of element and material information for matching BIM to LCA data. For example, the structural properties of concrete are not yet defined in early design stages and should instead be considered as a range of material options due to different compressive strength classes.

This paper presents a novel methodology for automatically matching the coarse information available in BIM models of the early design stages to the respective entries in LCA databases as a basis for a fully automated calculation process of the embodied GHG emissions of new buildings. This approach solves the existing gap in the automation process of manually enriching BIM models and adding information of LCA data and missing layers of vague models. In more detail, the proposed method is based on Natural Language Processing (NLP), using different strategies to increase performance in matching elements and materials from a BIM model to a knowledge database to enrich environmental indicators of commonly used elements' materials. The knowledge database contains all missing information for LCAs and has different levels of information for a range of several potential design options of elements and materials, including their

dependencies. Accordingly, this paper investigates multiple NLP techniques and evaluates the performance of state-of-the-art deep learning models such as GermaNet, SpaCy, or BERT. Following this, the most performant NLP approach is used to provide an automatic workflow for matching Industry Foundation Classes (IFC) elements to the knowledge database, facilitating a seamless LCA in the early stages of design. For five different case studies, the performances of the proposed matching method are analyzed. Finally, one case study is selected to compare the embodied emissions results to those of the conventional process.

## 3.1 Introduction

According to the United Nations, the construction industry, specifically through the production of materials for building construction, is responsible for 11% of the global energy-related carbon emissions (Abergel et al., 2017). In order to reach the international goals of the Paris Agreement and reduce the environmental impacts, Green House Gas (GHG) emissions of new buildings must be significantly reduced. To assess the Global Warming Potential (GWP) of buildings, life cycle assessment (LCA) is an established method for calculating environmental indicators along the whole life cycle. At its core, it is based on environmental impact datasets for individual materials, typically provided through dedicated databases. During the design phase, a careful LCA of the different design options is required in order to identify the main drivers and optimize the building design accordingly. However, in conventional projects in today's practice, the main focus is still on improving the economic performance of buildings, while environmental qualities are usually not prioritized or even considered.

Until recently, LCA has mainly been calculated manually, which is time-consuming, especially when it comes to quantifying the building elements and matching them to environmental datasets, which have a different classification system and ontology. BIM combines geometry and semantics and thus facilitates deriving consistent and automated quantity take-off of the relevant elements for calculating whole building LCA. Using and enriching the semantic information of e.g., materials has great potential to completely automate the calculation of whole building LCA (Safari & AzariJafari, 2021).

In early design stages, significant decisions are taken that have a major impact on the carbon footprint of the building to be realized. This is a primary reason for conducting a holistic multi-criteria variant analysis in the early design stages. At the same time, the early design stages are characterized by a high degree of uncertainty due to the lack of information and not-yet-taken decisions, making a holistic and consistent LCA for supporting design decisions and optimizing performance challenging (Schneider-Marin et al., 2020). In more detail, in the rough BIM models of early design stages, materials are typically defined by material groups rather than specific types, which allows a wide range of possibilities for each material group. Furthermore, several materials or element layers might not yet be defined, which gives the opportunity to explore and compare different design options. While several approaches for BIM-LCA integration exist, they are limited in implementing a fully automated workflow with open BIM models, in particular when it comes to early design phases (Forth et al., 2019). A major challenge lies in the fact that imprecise type and property information in BIM models hinder a seamless processing for LCA applications.

To overcome this issue of vague model information in early design phases resulting in labor-intensive processes with additional manual input, we introduce the concept of "semantic healing" for automatically calculating embodied greenhouse gas (GHG) emissions. In doing so, we propose a novel automated method of matching LCA and BIM data on the element level by using Natural Language Processing (NLP). This gap of a fully automated matching process has not been filled yet (Safari & AzariJafari, 2021), while research on NLP has recently advanced significantly and has strong potential for solving problems in the AEC industry (Locatelli et al., 2021).

This paper focuses on supporting decision-making in the early design phases. To support the decision-making in these phases, decisions for more detailed phases are also anticipated and analyzed. Based on the current approaches in the literature, the findings are considered to further extend the approach in the sense of a holistic analysis that is adaptable for further sustainability criteria.

The main contribution of this paper to the previously described problem involves a novel approach for semantically healing conceptual BIM models to assist the calculation of a holistic LCA, informing design decisions to detail the design further. The model healing process is conducted by enriching all necessary information to the model by automatically matching elements from BIM models to a knowledge database (discussed in detail in section 4) using Natural Language Processing (NLP).

In summary, this paper aims to answer the following research question: *Is automated semantic healing of BIM models possible in a way that allows assigning correct element types and materials to the respective model elements such that a reliable LCA can be calculated?*

It is structured as follows: Section 2 provides the relevant background in the field of BIM, classification systems, NLP, and its application with BIM. Afterwards, Section 3.3 focuses on the state of the art of BIM-based LCA and discusses existing literature reviews, highlighting their limitations. Section 3.4 presents the methodology for enriching BIM models for LCA and proposes a new methodology for the semantic model healing process. The proposed methodology is then evaluated in Section 3.5 through different real-world case studies, where the potential, as well as limitations, are highlighted. Finally, Section 3.6 presents our conclusions and recommendations for future research.

## 3.2 Background

This Section describes multiple fundamental topics about BIM, level of development, classification systems, and Natural Language Processing (NLP), providing the necessary background for the following Sections.

### 3.2.1 Level of Development (LOD) and Building Development Level (BDL)

As building design is a progressive process in which initially vague information is further detailed, also BIM models gain more accuracy and reliability along the modeling process. Level of Development (LOD) represents the degree of completion, maturity, or elaboration (Abualdenien et al., 2021). While the BIMforum, the US chapter of buildingSMART International, has defined individual LOD (BIM Forum, 2020), they have not been adopted as an international standard, yet. Defined in the European standardization effort EN 17412, Level of Information Needs (LOIN) describes similar content like LOD, such as geometric and alphanumerical information (Abualdenien et al., 2021), but specifies a particular use-case and milestone it is supposed to be applied for.

In Germany, LOD is known as the aggregation of Level of Geometry (LOG), specifying the geometric detailing, and LOI, representing the extent of alphanumerical information. Borrmann et al. discuss that Level of Information (LOI) is highly dependant on the project and client, so they can not be generalized. In BIM practice, LOI is often described with "Type-and-attribute tables" (TAT) specifying object types and attributes (Borrmann et al., 2021). Additionally, buildingSMART International proposed the Information Delivery Specifications (IDS). "The main goal of IDS is to provide a simple yet comprehensive way to author and validate nongeometrical [Information Requirements]", for example specifying material or classifications (Tomczak et al., 2022).

Abualdenien and Borrmann developed a meta-model approach where multi-LOD data represent buildings at different design phases (Abualdenien & Borrmann, 2019). It is based on the BIMForum's LOD definitions and introduces a new concept, Building Development Level (BDL). While LOD defines specific components, the BDL concept defines the maturity of the overall building with multiple LODs for each component.

LOIs and LOGs are of great importance for BIM-based LCA as they provide a means to specify the required information, or in turn, allow to take into account the vagueness and uncertainty of information provided in early design phases. Since less information is available in early design phases, generic datasets are used and missing material layers have to be assumed. During construction, on the other hand, product-specific data sets can be included in the calculation depending on the components used.

### 3.2.2 Open BIM and open formats

The design and construction of a building is a collaborative process that incorporates multiple disciplines. Each expert, such as the architect and structural engineer, uses different authoring tools and requires specific information to be present in the model to support a particular type of simulation and analysis. With the increasing specialization of the stakeholders, the building industry requires a high level of interoperability. The US National Institute of Standards and Technology (NIST) (GCR, 2004), as well as many researchers and case studies (Cemesova et al., 2015; Hernández et al., 2018; Lai & Deng, 2018) have confirmed the difficulties and high annual costs resulting from the lack of interoperability between the AEC industry software systems.

The Industry Foundation Classes (IFC) schema (Liebich, 2013) is an open data exchange format developed and maintained by buildingSMART with the goal of enabling interoperability across the AEC industry. It provides a common data model for lossless geometric as well as semantic data exchange. IFC is a free vendor-neutral standard and includes a large set of building information representations, including a variety of different geometry representations and a large set of semantic objects modeled in a strictly object-oriented manner.

Since 2009, the exchange format Green Building XML (gbXML) has been established as a public, non-profit schema (Green Building Foundation, 2021) focusing on exchanging building information for operational energy simulations. Initially developed by Green Building Studio and later acquired by Autodesk, it is currently not maintained by an official standardization body. The extension markup language (XML) schema does not intend to describe a complete BIM model but represents the relevant building's environmental and geometric information. Often, the reduced BIM model is referred to as building energy model (BEM). The schema provides a container denoted as "campus" for one or several buildings, each of which has a closed building envelope described by surfaces. The surfaces have a type specification (e.g., "InteriorWall"), B-Rep geometry, references to adjacent spaces, which are referenced to zones, and assigned openings.

The gbXML format is used for LCA in early design stages, e.g., using CAALA software, considering both embodied and operational emissions. Nevertheless, the details about specific element layers and materials are not represented and therefore, not suitable for accurately matching environmental datasets on material level.

### 3.2.3 Classification systems

The classification of elements in BIM models enables the project-wide, uniform structuring of information in order to be read and used in an uniform and automatic manner. Applying "a classification system for component types in a digital building information model" enables all stakeholders "to have a common understanding of the information contained in the building model and, in conjunction with a system for model development, enables the realization of a high degree of automation for the processes to be operated by them" (VDI 2552 Blatt 9, 2022).

In the international context, the classification systems Omniclass and Uniclass are among the most widespread. In Germany, due to the lack of a full-scale classification system, the most common classification systems are DIN 276 for cost groups (DIN 276, 2018) and DIN 277 for room usage types(DIN 277, 2021). According to German standards for calculating LCA, e.g., certification systems like DGNB or BNB, the classification system of the cost groups of DIN 276 is used (BMI, 2015; DGNB GmbH, 2020).

For the LCA context, DIN V 18599, focusing on the Energetic evaluation of buildings [1], has been recently established (DIN EN 15643-2, 2021).

For LCA of buildings, a uniform classification of building elements defines the system boundary, especially for the manufacturing phase (A1-A3) as well as the end of life cycle (C3-C4) and module D. Thus, it is part of the "target and investigation framework" according to DIN EN ISO 14040 (DIN EN ISO 14040, 2021). In German certification systems, according to Deutsches Gütesiegel Nachhaltiges Bauen (DGNB) and Bewertungssystem Nachhaltiges Bauen für Bundesgebäude (BNB), the classification of cost groups is carried out according to DIN 276 (DIN 276, 2018), taking into account the building elements for the cost groups KG 300 "Building - Structures" (see A.1). The system boundary for the operational phase, in particular the energy consumption during operation (B6), on the other hand, refers to DIN 18960, which however is not relevant to this paper. For the classification of relevant areas, on the other hand, the net room area (NRF) according to DIN 277 is used (DIN 277, 2021).

### 3.2.4 Natural language processing (NLP)

Natural language processing allows computers to analyze and "understand" text created by human authors. At its core, natural text is transformed into a computer-readable representation through various techniques, including tokenization, lemmatization, and vectorization. Those techniques convert each word to its original/dictionary form and represent each word with a numerical value, describing the semantic similarity through their distance (e.g., the word *window* has a smaller distance to *door* than to a *tree*). Semantic similarity is a key feature of the matching process described in this paper.

As in other domains, artificial intelligence revolutionized its advancement. In this regard, long short-term memory (LSTM) and recurrent neural networks (RNN) dominated NLP as they learn bidirectional links between the vector representations of words and sentences to capture the overall meaning. Recently, those networks were outperformed by transformer-based models. One example of a pretrained deep bidirectional transformers is BERT by Google (Devlin et al., 2018). The structure of transformers consists of an encoder and a decoder, and transformer-based models themselves consist of multiple layers of

[1] Full title: Energetic evaluation of buildings in the context of the energy consumption in the use phase (B6) relevant for the life cycle assessment in accordance with DIN EN 15643-2

transformers (Vaswani et al., 2017). This enables learning the contextual representations of input data.

### 3.2.5 NLP application in AEC

Locatelli et al. investigated in their scientometric analysis the synergies between NLP and BIM (Locatelli et al., 2021). Beside the field of Automatic Compliance Checking, they also identified Information Retrieval from BIM models and Information Enrichment of BIM objects as a further fields of relevant application. Wang et al. developed a query-answering (QA) system for BIM information extraction (IE) by using NLP and achieved high accuracy scores in their evaluation (N. Wang et al., 2022). Xie et al. introduced a method for matching real-world facilities to BIM using NLP for word segmentation and keyword extraction by adopting the LTP word segmentation module (Xie et al., 2019). For the matching method itself, matching matrices based on HiTree paths are evaluated using the highest degree of matching with the natural language feature vector. Reitschmidt proposed an matching method of IFC materials to the LCA database Ökobaudat based on tokenization of material names and a distinct matching or via Levenshtein distance (Reitschmidt, 2015). Nevertheless, automated matching of LCA and IFC data on the element level using NLP has not been developed yet (Safari & AzariJafari, 2021). Finally, Zahedi et al. proposed an NLP approach for documenting design decisions by searching building codes and request for proposal documents (Zahedi et al., 2022).

## 3.3 State of the Art of BIM-based Life Cycle Assessment (LCA)

This Section focuses on a literature review of the current approaches of BIM-based Life Cycle Assessment (LCA). First, existing literature reviews are compared. Based on this, a structured literature analysis is conducted by analyzing each publication according to several topics. Finally, the findings and limits of conventional and current BIM-based LCA methodologies are shown.

### 3.3.1 Existing literature reviews

Before presenting the literature analysis, existing ones are analyzed to prevent repetition. The focus is primarily on embodied emissions and energy rather than operational emissions or energy. Nevertheless, the aspect of multi-criteria approaches will be investigated too, for example, a combination of embodied and operational energy with life-cycle costs (LCC). Analyzing eleven publications from 2013 to 2015, the literature review of the BIM-based LCA method by Soust-Verdaguer et al. differentiates between Data input (BIM-LOD, LCA goal & scope, stages, and inventory), Data analysis (BIM software, Energy Consumption Calculation, LCA tool) and Outputs and communication of results (Environmental impact indicators, sensitivity analysis, embodied and operational CO2 emissions) (Soust-Verdaguer et al., 2017).

In 2019, Wastiels and Decuypere classified existing approaches and identified five different strategies for BIM-LCA integration (Wastiels & Decuypere, 2019). Later literature reviews base their findings on these five strategies, which contain Bill of Quantities (BOW) export, IFC import of surfaces, BIM viewer for linking LCA profiles, LCA plugin for BIM software, and LCA-enriched BIM objects.

Potrč Obrecht et al. classified in their literature review all analyzed methods according to the five strategies by Wastiels and Decuypere (Potrč Obrecht et al., 2020). In the second step, they differentiated between manual, semi-automated, and automated approaches. In 2020, several other literature reviews were published focusing on different aspects. Roberts et al. identified in their literature review about LCA in building design process three different trends: integration of LCA into BIM, combining LCA and LCC, and using parametric approaches (Roberts et al., 2020).

Cavalliere et al. concentrate on the capabilities of the combination of BIM and parametric-based tools, analyzing 25 different publications between 2013 and 2018 (Cavalliere et al., 2020). Most of the analyzed methods focused on BIM and only a few had a parametric approach included. Hollberg and Ruth were the first ones to develop a parametric-based LCA (PLCA) in 2016, using Visual Programming Language (VPL) but no BIM integration (Hollberg, 2016). Llatas et al. focus in their systematic literature review on Life Cycle Sustainability Assessment (LCSA) and add, besides LCA and LCC, also sLCA in their investigation approach (Llatas et al., 2020). In total, they reviewed 36 papers about BIM-LCSA integration, but only six methods included LCA and LCC and none sLCA. Tam et al. analyzed in their critical review on BIM and LCA 61 articles by using content analysis method (Tam et al., 2022). Furthermore, they identified several unaddressed issues, for example, the lack of a standardized structure between BIM and LCA data.

### 3.3.2 Literature analysis

Based on the findings of existing literature reviews in the field of BIM-based LCA, a systematic literature analysis was conducted. After reviewing more than 60 publications in this field, published in 2018-2022, 25 were selected and analyzed. In the following, the main findings are described. The main focus of several approaches is on detailed design stages such as (Eleftheriadis et al., 2018; Santos et al., 2019; Theißen, Drzymalla, et al., 2020). However, optimization of the building design can be achieved in early design stages, when information is still uncertain. Therefore, Rezaei et al. are suggesting a workflow that is based on Autodesk Revit but doesn't include an optimization process (Rezaei et al., 2019). Only a few methodologies implemented uncertainties in their approach (Cavalliere, Hollberg, et al., 2019; Eleftheriadis et al., 2018; Rezaei et al., 2019).

As previously shown, Wastiels and Decuypere classified five different integration strategies (Wastiels & Decuypere, 2019). The two mainly implemented approaches of the analyzed publications are the one which uses authoring tools for getting the Bill of Quantities (BoQ), which was analyzed by (Potrč Obrecht et al., 2020). The second primary strategy is using BIM objects enriched with property sets (Pset) (Eleftheriadis et al., 2018; Llatas et al., 2022; Santos et al., 2020; Theißen, Höper, et al., 2020). A new approach by Lee et al. suggests BIM templates for authoring tools to avoid data loss due to exchange formats (Lee, 2021). Only a few of the analyzed publications use existing LCA Plugins for Autodesk Revit, such as Tally, eToolLC, or One Click LCA (Atik et al., 2020; Carvalho et al., 2020; Nilsen & Bohne, 2019; Veselka et al., 2020; X. Yang et al., 2018). As most of the approaches use the BIM model only for downstreaming LCA-related information, only one includes a computer-readable feedback communication process of the calculated results back to the BIM model (Horn et al., 2020).

Most of the analyzed approaches used the open BIM format, mainly IFC, such as (Figl et al., 2019; Horn et al., 2020; Llatas et al., 2022; Santos et al., 2019; Theißen, Drzymalla, et al., 2020). Nevertheless, another open BIM exchange format specialized in energy simulation is gbXML (Green Building Extensible Markup Language), which was used by (X. Yang et al., 2018). Other approaches use the closed BIM approach with software tools like Autodesk Revit (Abu-Ghaida & Kamari, 2021; Figl et al., 2019; Nizam et al., 2018) or additionally in combination with the VPL tool Autodesk Dynamo (Bueno et al., 2018; Hollberg et al., 2020; Kiamili et al., 2020; Naneva et al., 2020; Röck et al., 2018a). Another used VPL tool is McNeel's Rhino and Grasshopper, which was used by (Cavalliere, Habert, et al., 2019; Hollberg & Ruth, 2016; Lobaccaro et al., 2018), which is not considered as a BIM tool just as little as Trimble's Sketchup, used by (Meex et al., 2018).

Although this publication focuses on LCA, the framework allows it to be extended to multiple criteria for design optimization. Only a few analyzed approaches show a few more criteria, which can be included in their workflows. While Kiamili et al. focus only on embodied energy of HVAC (heating, ventilation, air conditioning) systems (Kiamili et al., 2020), other approaches include both the embodied emissions of building construction and HVAC (Cavalliere, Hollberg, et al., 2019; Theißen, Drzymalla, et al., 2020). In a next step, further publications even include operational energy besides embodied energy (Di Bari et al., 2019; Figl et al., 2019; X. Yang et al., 2018). Besides LCA, Life Cycle Costs (LCC) and social Life Cycle Assessments (sLCA) are further relevant criteria to consider in the field of LCSA. A few approaches include both LCA and LCC (Abu-Ghaida & Kamari, 2021; Eleftheriadis et al., 2018; Figl et al., 2019; Santos et al., 2019). Llatas et al. propose the only approach, which considers all three criteria of LCSA, while the main focus of sLCA is on working hours (Llatas et al., 2022). Nevertheless, there is no methodology that integrates embodied emissions of building construction and HVAC, as well as operational emissions in early design phases.

As a functional unit of the approaches, most of the analyzed publications focus on the whole building. Global Warming Potential (GWP) was considered by all approaches, while other publications also considered further environmental impact categories such as acidification potential (AP), eutrophication potential (EP), ozone depletion potential (ODP), and photochemical creation potential (POCP) (Atik et al., 2020; Meex et al., 2018; Palumbo et al., 2020; Santos et al., 2020; Theißen, Höper, et al., 2020). Depending on the country of the publication, several different international Life Cycle Inventory (LCI) databases were used, such as German Ökobaudat, or ecoinvent and KBOB from Switzerland, and sometimes even product-specific Environmental Product Declarations (EPD). Palumbo et al. investigated the challenge of using EPDs in early design stages to obtain accurate LCA results (Palumbo et al., 2020).

As a result of the literature analysis, there is great potential for including LCA calculations in an optimization process in early design stages using open BIM models. Furthermore, most of the analyzed publications focused only on the criterium of LCA, extending the focus on multiple criteria such as LCC is also becoming more relevant. Nevertheless, the process of matching LCA and IFC data on element and material levels is still manual, and an automated approach is not developed or solved yet.

### 3.3.3 Limits of conventional BIM-based LCA calculation

As the findings of the literature review showed, there are still challenges and opportunities in the field of BIM-LCA integration. In this Section, the limits of conventional BIM-based LCA approaches will be critically investigated using a case study.

Safari and AzariJafari stress in their publication out that a major focus will be in early design stages, considering LODs and uncertainties in future approaches (Safari & AzariJafari, 2021). Zimmermann et al. showed in their investigation of industry practice and needs different challenges, such as manual workflows, matching model data with LCA data, quality in models, and many more (Zimmermann et al., 2021).

Nevertheless, in conventional projects in practice, the main focus is still on the economic performance of buildings, while environmental qualities are not widely spread yet. This is the reason to approach the holistic multi-criteria variant analysis, in the early design stages, based on existing approaches of BIM-integration strategies for LCA. Current approaches still have limits of fully automated workflow with open BIM models (Forth et al., 2019). Scherz et al. propose in their methodology of hierarchical reference-based know-why models design support for several sustainability criteria focusing on building envelopes (Scherz et al., 2022). Nevertheless, BIM integration is only envisioned in their future work.

The main scope of this paper focuses on the early design phases. To support the decision-making at these phases, detailing decisions from more detailed phases are additionally analyzed. Based on the current approaches in the literature analysis, the findings are considered to further extend the approach in the sense of a holistic analysis that is adaptable for further criteria, for example, LCC or similar.

## 3.4 Methodology for semantic model healing for early BIM models for LCA calculation

The aim of this paper is to develop a framework for calculating ranges of embodied emissions of building designs based on element-specific design variants to support decision-making in early design phases. The methodological approach includes open BIM data exchange in early design stages, environmental impacts of construction, operation, and End-of-Life phase of buildings), as well as an automated matching of relevant information from the model. Therefore a robust implementation should take different modeling approaches (model authors & software products) into consideration. Furthermore, the framework provides flexibility to add economic impact or individual cost benchmarks and the calculation of further criteria.

As shown in the previous Section 3.3.3, the BIM-integration of LCA lacks an approach for early design stages, which fully automatically matches all information from BIM models to LCA datasets and considers uncertainties and missing information in early design stages. Therefore, the proposed methodology focuses on the following key features:

- Semantic model healing by using an LCA knowledge database (LKdb)
- Automated matching of IFC elements to the elements of LKdb using pretrained NLP models
- Calculation of LCA result ranges according to the early design uncertainties

The details of the method are described in the following Sections. First, the general framework is introduced, followed by more detailed descriptions of each part, such as semantic model healing, LKdb, and the matching method.

### 3.4.1 Proposed methodology

To perform multi-criteria analyses using BIM, engineers need a set of information to be present in the BIM model. Usually, in early design stages, some of the required information is uncertain or even completely missing, which has a significant influence on analysis or simulation results. For this reason, the concept of a knowledge database is introduced, which provides all relevant information and default values in the case that relevant information is missing.

As this paper focuses on embodied GHG emissions, the database is filled with LCA-relevant information. Nevertheless, the database can be easily extended to cover other criteria as well. In case of missing or uncertain information, such as elements or properties, the LKdb provides a set of possible options or ranges of values. Furthermore, several design variants can be explored in these cases, and their performances can be evaluated according to the influence of the incorporated uncertainties on the environmental qualities.

In the proposed methodology, design decisions are made by selecting one of these variants. To implement the conducted selections in the design, these are communicated back to the BIM authoring software. The proposed methodology follows the open BIM approach to support a wide range of authoring tools. Therefore, it uses Industry Foundation Class (IFC) and BIM Collaboration Format (BCF) as exchange data formats.

### General Framework

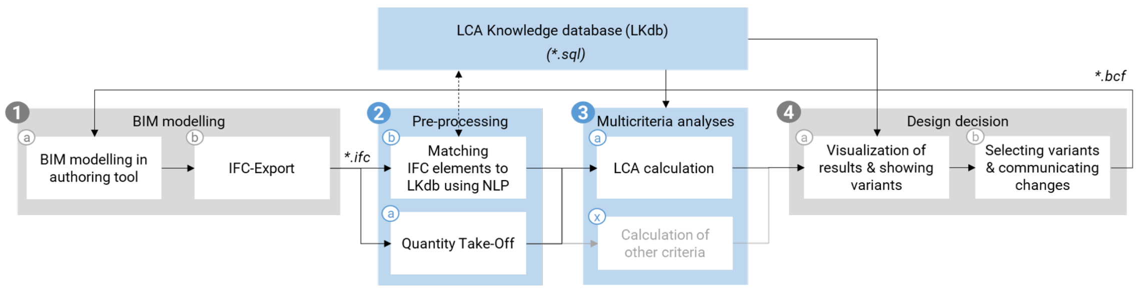



Figure 3.1: General Framework of NLP-based semantic model healing of early BIM models for LCA calculation

Figure 3.1 presents the different steps of the proposed methodology, which was briefly described above. In the first step (1), the BIM model is exported from the authoring tool as an IFC file. In the next step (2), the IFC data are pre-processed for the following analyses. This is split into the Quantity Take-Off (2.a) and the NLP-based matching method (2.b), which is explained in more detail in the upcoming Section 3.4.1. The Quantity Take-Off (QTo) contains information about the element type, the classification, the sum of all type-specific element areas, the area unit, the amount of type-specific elements, the element-specific materials, and the thicknesses of the material layers. In terms of the multi-criteria analyses to be performed in the next step, the focus is on the LCA calculation in this publication (3.a). The final step consists of visualizing the results (4.a), supporting the selection process, and communicating the design decisions and changes back to the BIM author (4.b). In this regard, this publication concentrates on the visualization of the LCA result ranges, including relevant benchmarks.

### Semantic model healing

The semantic model healing process is performed to add all relevant but missing information for the model-based LCA calculation. The first step in this process is to collect all available and relevant information from the IFC model. Based on this information, the second step focuses on how existing techniques of NLP help to match IFC elements to those of a knowledge database. Different strategies are used for the NLP-based healing process to increase the performance of the matching element from an "imperfect" BIM model to this knowledge database. In the last step, all missing element information is added by those of the matched knowledge database. The knowledge database contains all missing information for LCA and has different levels of detail for a range of several potential design variants of elements and materials, including their dependencies. The semantic model healing process is performed when the incomplete IFC element data are matched to the most similar element in the LCA knowledge database (LKdb) and afterwards enriched by all missing element information provided by the LKdb.

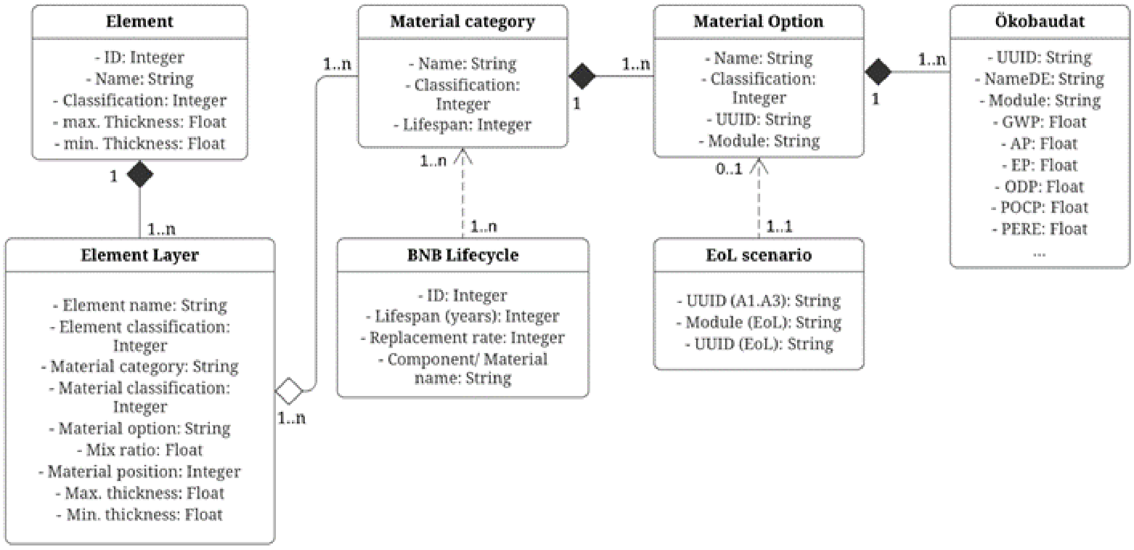



Figure 3.2: UML diagram of the proposed LCA Knowledge database with different hierarchies, such as elements, material categories and material options, and external databases such as BNB life cycle (BBSR, 2017) or Ökobaudat (BBSR, 2021)

### 3.4.2 LCA knowledge database (LKdb)

The LCA knowledge database, based on elements, layers, and materials, will contain all information that is relevant for the holistic calculation of different criteria and is typically not provided in the IFC model. This database is similar to the recently published "EarlyData knowledge base" by Schneider-Marin et al., which has a similar purpose of calculating reliable LCA results in early design stages considering uncertainties (Schneider-Marin et al., 2022). Nevertheless, the focus of their database was not focusing on using it for testing a robust matching approach. This LCA knowledge database is linked with different external databases, for example, databases for environmental criteria, such as Oekobaudat (BBSR, 2022) (Figure 3.2). The main aim of the LKdb is to provide all necessary input information for a holistic and correct LCA calculation and analysis, which is typically missing in early design phases. Another aim is to combine several external databases with different input information on different levels of information. International databases can be added to the database using this methodology. Material and product-specific Environmental Product Declarations (EPD) can be linked, too.

The database provides additional information on different levels, which are needed for a sufficient LCA calculation, such as the lifespan of an element, End-of-Life scenarios if missing in the original external dataset, or densities. Due to the German LCA classification standard according to cost groups, the database itself is structured similarly to the classification system of DIN 276 on the third level but provides a material-specific level of different element layers. Other criteria information like cost values or U-values (if missing in the model) for calculating operational energy can be stored in the database as well but are out of scope in this publication. This ensures that a change in the variants leads to a change in all criteria calculations and shows the complex dependencies of the multi-criteria design decision process. A first extension, including LCC, was tested recently (Lammers & Forth, 2022).

The general structure of the proposed LCA LKdb consists of three different levels: element, material category, and material option. As the LCI database, Ökobaudat was chosen (BBSR, 2021). The main reason for this decision is that the selected case studies are located in Germany. Thus the BIM models use German terminology for components and properties. Furthermore, Ökobaudat is the official LCI database for German certification systems and consists of more than 1400 datasets specifically for building products.

Every single dataset of Ökobaudat has as keys the Universally Unique Identifier (UUID) and the relevant life cycle modules (A1-A3, C3, C4, D). All datasets from Ökobaudat consist of several environmental impact categories, such as Global Warming Potential (GWP), Acidification Potential (AP), Eutrophication Potential (EP), Ozone Depletion Potential (ODP), Photochemical Creation Potential (POCP), Primary Energy Renewable (PERE) and many more. Nevertheless, the quality of some datasets is not sufficient for a holistic LCA, as there are some End-of-Life scenarios missing. Therefore, generic End-of-Life (EoL) datasets from Ökobaudat have to be manually matched to those datasets, which are lacking this information. For this reason, up to two UUIDs are linked to the material option dataset of the LKdb: one for the production phases and, if necessary, one for the End-of-Life scenario. Stenzel conducted in her master thesis this manual mapping as well as a classification of all UUID according to German cost groups using DIN 276 (Stenzel, 2020). This information is used for the prototypical implementation of the LKdb. All material options have a name and classification as their keys, which is derived from the German name in Ökobaudat. Further entries are the classification, UUID, included Modules, and the encoded NLP vectors of the name (spans and tokens), which are stored because of calculation performance reasons.

According to the structure of Ökobaudat, every material is classified according to specific material categories. As there are three different levels of categories, only the most specific one is used for material classification in the LKdb. Every material category is mapped to potential cost groups of the German classification system (DIN 276, 2018). This is necessary to map the service life of building components on this level, according to (BBSR, 2017). This external input is named "BNB life cycle" and contains an ID, the lifespan in years, the replacement rate according to 50-year buildings life, and an element or material name according to its own classification. The key for material categories is the name and the classification. Additional information is the encoded NLP vectors of the material category name (spans and tokens) due to calculation performance reasons.

For setting up element layers, material options and categories are used in the next level. Elements themselves can consist of one or multiple element layers. Both elements and element layers have a default maximum and minimum thickness. The material layer corresponds to the third level of the German cost group system (DIN 276, 2018). As the material layer can consist of composite materials, different mix ratios need to be defined. For monolithic layers, the ratio is 100%. As an example of composite materials in one element layer, reinforced concrete consists of different materials, such as concrete and reinforcement steel. Every element layer has a unique material position, which describes the order of the material in the specific element. For the element levels, every entry gets a unique ID as a key. Due to calculation performance reasons and also for the elements, the encoded NLP vectors of the element names (spans & tokens) are stored in the LKdb.

All entries for elements are inserted due to common domain knowledge. The most typical construction types were considered and modeled using the proposed schema. Due to the versatility of constructions, the database is continuously updated and has no claim to be ever completed.

### 3.4.3 Matching method

In later design stages, conventional methods rely on manually matching each IFC material to a UUID of external databases and store this information as a Pset attribute in "PSetEnvironmentalImpactIndicators" and "PSetEnvironmentalImpactValues" according to (BuildingSMART International Limited, 2020) or self-defined Psets, such as "Plca_Lca" according to (Theißen, Drzymalla, et al., 2020). To avoid the laborious manual work of matching elements and materials of the BIM models to the related ones in the LKdb, an automated matching method is proposed in this paper. Another approach by Reitschmidt also follows automated matching on material level (Reitschmidt, 2015). In contrast, in early design stages, information about the materials is missing or incomplete. For this reason, the proposed method is matching on an element-level, so this vague or missing information about material layers can be added using the LKdb.

The main challenge of this method is to automatically and correctly match IFC and LKdb elements and materials so that calculation and analysis results are also reliable. In early design stages, materials are often defined in a more general way and not as specific as in LCA databases, e.g., "concrete" rather than "concrete C20/25". Sometimes, for some elements, material information is completely missing, while in the element naming, some material information is included, for example "brick wall". Furthermore, the proposed methodology aims to be a robust approach, which also considers poor model quality due to multiple ways of modeling BIM models and exporting them as IFC files. As the structure and nomenclature of elements and materials in IFC and the used LCI database Ökobaudat differ, the goal is to find the semantically most similar pairs on material and element level.

Figure 3.3 shows the proposed matching method, which is divided into three steps:

1. Filtering of element classification
2. Similarity analysis using NLP
3. Element selection

First, IFC elements are filtered according to their classification type. This classification, according to the German cost group schema (DIN 276, 2018), is an exchange requirement

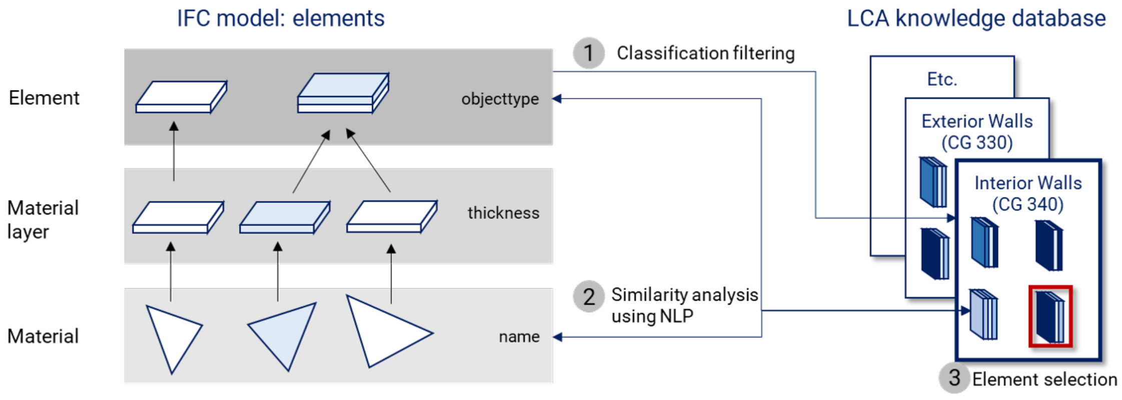



Figure 3.3: General steps for an automated method of matching elements of IFC models to those of the LCA knowledge database (LKdb)

(ER) and is stored as "IfcRelAssociatesClassification". If the element does not comply with the ER and no classification is available, the method can also classify the IFC element using its "IfcProduct" class types (e.g., IfcWall, IfcColumn, IfcSlab, etc.) and properties (e.g., IsExternal, IsLoadBearing, etc.) according to (C. Richter & Liedtke, 2021).

In the second step, every IFC element and its properties are analyzed and semantically compared according to its similarity with the filtered element variants in the LKdb. Not only the element expressions but also the material expression is analyzed according to the NLP technique used. In order to measure semantic similarity, every expression needs to be converted from text to a vector representation. In this case, a vector is a list of numerical values, and the combination of them represents the overall meaning (Wilbur & Sirotkin, 1992). Afterwards, the similarity between two different vectors A and B can be calculated using the cosine similarity, while $n$ is the dimension of the vector:

$$cosine - similarity := cos(\theta) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{3.1}$$

In the following Sections, these three main steps of the matching method are explained in more detail, as shown in Figure 3.4. The choice of NLP technique will be investigated in Section 3.5.2.

### Element filtering

The starting point is iterating through each element type from the IFC model. Each element type consists of an element name, its classification according to DIN 276, and its material name. Based on the classification, a list of LKdb elements is filtered to compare similarities with the IFC element. For performing a robust matching method, the elements are compared on material and on element levels. Therefore, the IFC element name is compared to the filtered list of LKdb elements. And furthermore, the IFC material is compared to the material categories and material options which are contained in the filtered element list. The differentiation between material category and material options is

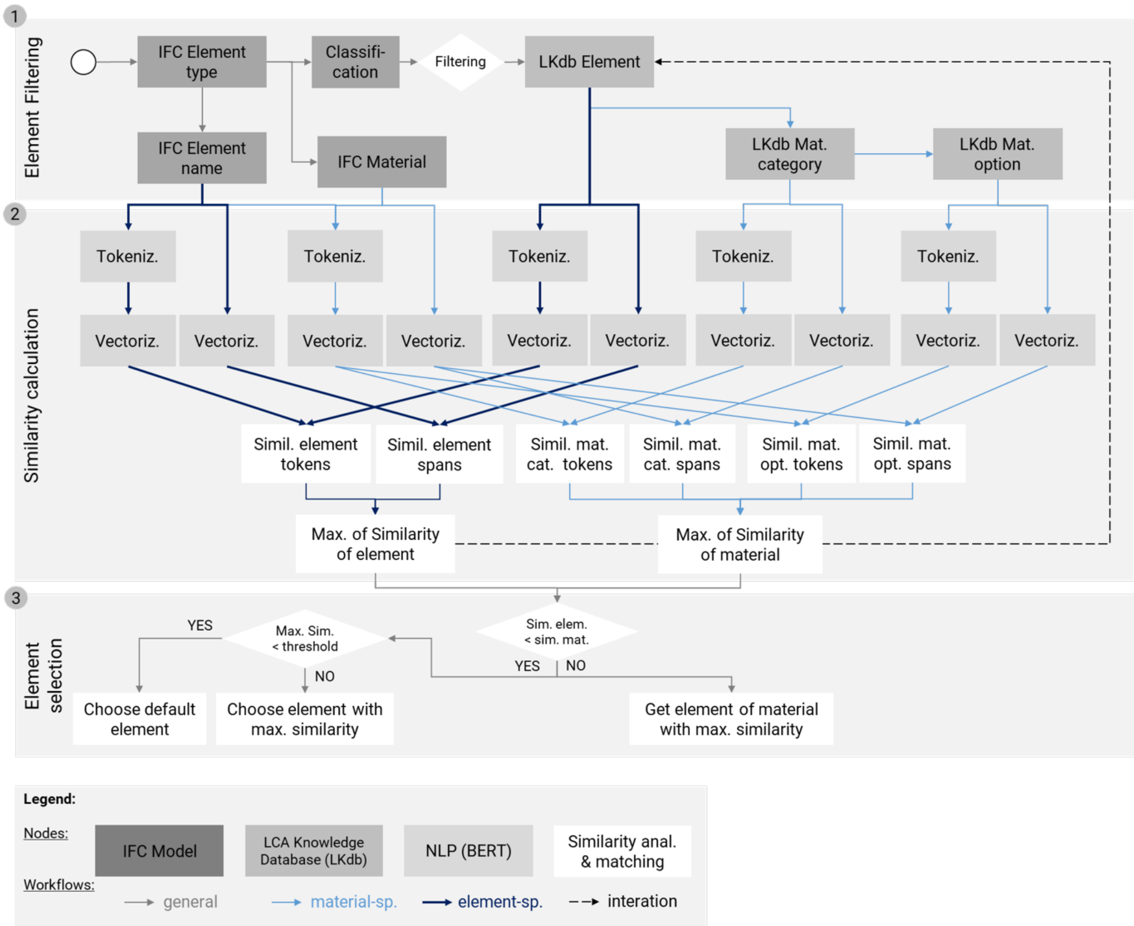



Figure 3.4: Detailed workflow for matching IFC elements to LKdb elements using Natural Language Processing (using BERT language model) and cosine similarity on different levels of information (element, material category, and material option)

required due to the fact that the matching method considers different LOIs for the naming of materials (see Section 3.4.3).

### Similarity calculation

In the calculation of semantic similarities, three couples of IFC and LKdb are considered: on element level, material level comparing with the material category, and comparing with the material option. Each of these three couples is split into calculating the whole span and all tokens. To this end, the word encoding or vectorization is conducted for twelve different words per every iteration step, while the tokens themselves are also iterated. For each token set, only the maximum token is considered in the following selection process. The calculation of the cosine similarity is conducted six times per iteration step and is stored in a list for the following selection process:

- element tokens
- element spans
- material category tokens

- material category spans
- material option tokens
- material option spans

After the calculation of all cosine similarities, the most similar element and material are identified. The maximum similarity of all element tokens and element spans are compared for the most similar element. Accordingly, the maximum similarity of all material category tokens and spans, as well as material option tokens and spans, are derived for the most similar material.

**Element selection**

In the next step, the final element selection is performed based on the previously derived most similar element and material. Therefore, the two cosine similarities of the most similar element and most similar material are compared. If the similarity of the material is higher, the corresponding element of the material is searched and selected. In case the similarity of the element outperforms the one of the material, this element is selected if its cosine similarity is higher than a threshold. As a threshold, 80% was set, according to the material similarity analyses using the BERT model in Section 3.5.2. If this threshold is not reached, the default element of the classification group is chosen, as the identified element similarity is too low to ensure the quality of this matching method. For IFC elements with multiple material layers, the steps of the previously explained workflow are derived for every material layer. Nevertheless, in the end, the different results have to identify only one selected element. For this, the different elements of each layer are counted, and their cosine similarities are summed up. Finally, the element with the highest summed-up cosine similarity is selected as the overall multi-layer matched element.

### 3.4.4 LCA calculation of LKdb elements

This paper focuses exclusively on embodied emissions. For this reason, for the LCA calculation, the operational part B6 is omitted. This study does not focus on different environmental impact potentials but on the reliability of the calculation process. The system boundaries of the LCA include the life cycle phases production (A1-A3), maintenance and replacement (B4), and End-of-Life (C3, C4, D).

Generally, the Environmental Impact Potential ($EIP_{c_o}$) of the construction phase ($c$) for each element ($e$) is the sum of the production phase ($P_e$), recovery and disposal phase ($D_e$), and the maintenance and replacement ($M_e$) in a reference period ($t_D$). As in the LKdb, different material options for one material layer exist. The element-specific environmental impact potential can consist of a range of results rather than a single value. In the following, the different steps are described for calculating the Environmental Impact Potential of one specific option set ($o$). The final LCA result ranges are derived by the different options and can be clustered on element or cost group level or determined for the whole building.

$$EIP_{c_o} = \sum_{e_o}^{n} \frac{P_{e_o} + D_{e_o} + M_{e_o}}{t_D} \tag{3.2}$$

The maintenance and replacement $M_{e_o}$ of each element are calculated by the frequency of replacement ($n_{replacement,e}$) and the sum of the production $P_{e_o}$ and recovery and disposal phase $D_{e_o}$, while the frequency of replacement depends on the ratio of the reference period $t_D$ and the service life of the element ($t_e$).

$$M_{e_o} = n_{replacement,e} * (P_{e_o} + D_{e_o}) \tag{3.3}$$

$$n_{replacement,e} = roundup(\frac{t_D}{t_e}) - 1 \tag{3.4}$$

The production $P_e$ of each element is the sum of the product of each layer-specific dataset for the production phase ($EIP_{e_{o,i}}^{A1-A3}$) and element-specific quantities ($f_{e_{o,i},x}$) over each element layer ($i$) of the element-specific maximum amount of layers ($m_{e_o}$). The recovery and disposal $D_{e_o}$ is, accordingly, just taking the datasets for different life cycle phases into account (C3-C4, D).

$$P_{e_o} = \sum_{i=0}^{m_{e_o}} EIP_{e_{o,i}}^{A1-A3} * f_{e_i,x} \tag{3.5}$$

$$D_{e_o} = \sum_{i=0}^{m_{e_o}} EIP_{e_{o,i}}^{C3-C4,D} * f_{e_i,x} \tag{3.6}$$

The datasets $EIP_{e_i}^{A1-A3}$ or $EIP_{e_i}^{C3-C4,D}$ are stored in the LKdb. Depending on the functional unit ($x$), the quantity of each dataset can either be area $a_e$, length $l_e$, volume depending on the layer-specific thickness $d_{e_{o,i}}$, mass depending on the material-specific density $\rho_{o,i}$, or amount $s_e$.

$$f_{e_{o,i},a} = a_e \tag{3.7}$$

$$f_{e_{o,i},l} = l_e \tag{3.8}$$

$$f_{k_{o,i},v} = a_e * d_{e_{o,i}} \tag{3.9}$$

$$f_{e_{o,i},m} = a_e * d_{e_{o,i}} * \rho_{o,i} \tag{3.10}$$

$$f_{e_{o,i},s} = s_e \tag{3.11}$$

| Case study number | Net floor area (sqm) | Total amount of elements | Total surface area of all elements (sqm) |
|---|---|---|---|
| 1 | ca. 11.870 | 2.110 | 68.949,39 |
| 2 | ca. 1.950 | 307 | 5.823,82 |
| 3 | ca. 35.300 | 13.966 | 85.193,77 |
| 4 | ca. 11.390 | 7.144 | 118.155,97 |
| 5 | ca. 8.710 | 5.822 | 117.562,25 |

Table 3.1: Information about the five case studies considering net floor area, total amount of elements, and total surface area of all elements

Depending on the level of the matching and available attributes of the IFC elements, different quantities can be used for this calculation step. The total area, length, and amount of all IFC elements of one specific object type are always derived by the Quantity Takeoff. If no material information is available in the IFC element and the matching is performed on the element level, the default quantities, such as thicknesses and densities, from the LKdb are used. If the matched element is based on most similar materials, the material layer information of the IFC element is used for the LCA calculation. This is also valid if, for a multi-layer element, only a few materials were identified in the matched element. For these matched materials, the material layer thicknesses of the IFC element are used, while for the missing ones, the default values are used according to LKdb. This selection ensures that all available and relevant information of the IFC model is used for LCA calculation. The LKdb provides all geometric and semantic information of the material layers, which are not modeled in the IFC model but are crucial for a holistic LCA calculation.

## 3.5 Evaluation and results

In this Section, we first briefly introduce five case studies, which are used to evaluate the proposed methodology. In the first evaluation, the best-performing language model is identified by testing three different models (GermaNet, spaCy, BERT) using the manually matched couples (IFC-LKdb) of case study 1. In the following Subsection, the whole element matching workflow is evaluated on all five case studies. Case study 2 is used for evaluating the whole procedure, including the LCA calculation using Global Warming Potential (GWP) as environmental impact category. Finally, we discuss the limitations of the proposed methodology based on the evaluations.

### 3.5.1 Case studies

To validate the proposed matching method, five case studies from real-world projects were selected, as shown in Figure 3.5 and Table 3.1. They are all office buildings, so the performance of the proposed approach is comparable but from different modelers and designers. Nevertheless, the quality of material and element naming, as well as the

Figure 3.5: Selected case studies for validating the proposed matching method (Picture of case study 1: ("Baustart für neues Siemens Technology Center in Garching", 2022), case study 3 (Siemens Deutschland, 2022)

modeled BDL and classification, differ in all five case studies and need to be taken into account in the following analysis.

In Figure 3.6, the element distributions of the 2nd and 3rd levels of the German classification system according to DIN 276 are shown. Case studies 2, 4 and 5 do not have elements in classes 320 (foundations) and 360 (roofs).

### 3.5.2 Evaluation of different NLP techniques for material matching

Following this, this publication investigates multiple NLP techniques and evaluates the performance of state-of-the-art deep learning models such as GermaNet, SpaCy, or BERT. They will be introduced in the following Sections and are the basis for the previously introduced matching method. The best-performing NLP technique is applied for the prototypical implementation and validation.

For comparing the three different NLP techniques and the performance of their workflows as well as calculating the whole building LCA, case study 1 was chosen, which was presented in Section 3.5.1. This real-life project guarantees that the material naming is not optimized but according to current industry standards so that the matching performances are tested under realistic conditions. In total, the IFC model of case study 1 consists of 2110 individual elements, which are summed up to 133 unique elements from the same families. Those consists of 59 unique IFC materials, which were manually matched to LCA material options and categories.

#### GermaNet

GermaNET is a Lexical-Semantic Net for the German language and is also known as the German version of the Princeton WordNet (Hamp & Feldweg, 1997; Henrich & Hinirchs,

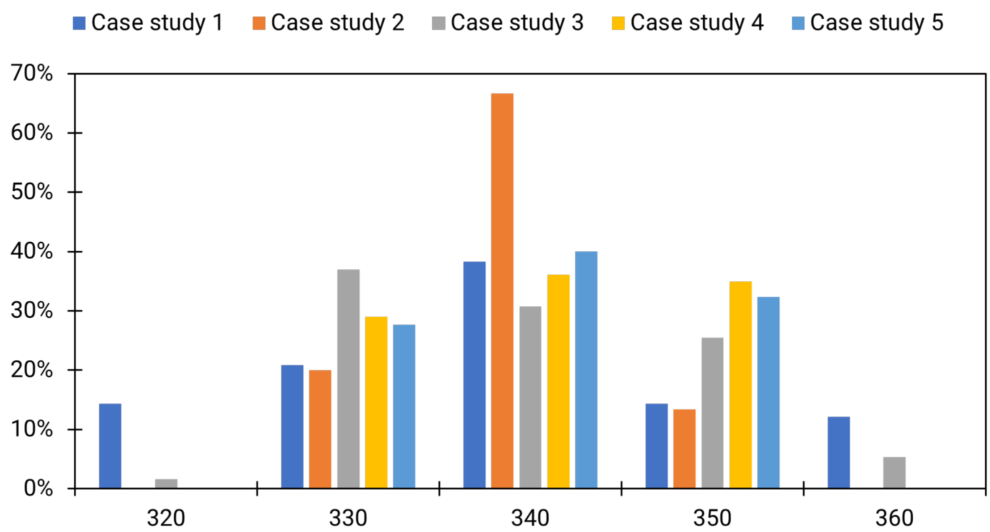



Figure 3.6: Overview of elements' classification distribution of the five case studies

2010). GermaNet relates German nouns, verbs, and adjectives semantically by grouping lexical units that express the same concept into synsets and by defining semantic relations between these synsets (sets of synonyms). It can be represented as a graph whose nodes are synsets and its edges its semantic relations (Navigli & Martelli, 2019). Therefore, the similarity is not measured using cosine similarity but graph-related shortest path similarity, which is equal to the inverse of the shortest path length between two synsets. There are other path-related similarity analyses, such as Wu-Palmer similarity (Z. Wu & Palmer, 1994) or Leacock-Chodorow similarity (Leacock & Chodorow, 1998), which are not considered in this paper.

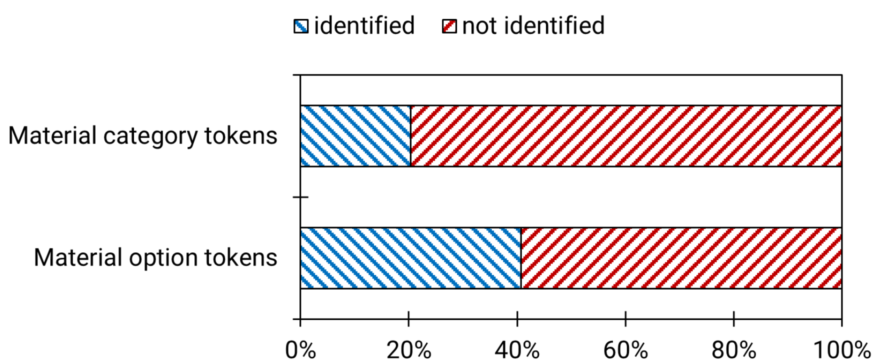



Figure 3.7: Identification rate of material token synsets using GermaNet for case study 1

As the workflow of the GermaNet differs partially from the other two NLP techniques, the identification rate of the material token's synsets needs to be analyzed before analyzing the shortest path similarity. After the tokenization of the IFC material names, material options, and their related material categories of the LKdb, synsets are identified to calculate the shortest path similarity. Nevertheless, not for every token set, synsets could be identified. As shown in Figure 3.7, only for 20.3% of the material category tokens and 40.7% of the material option tokens, a pair of synsets with the IFC material could be identified.

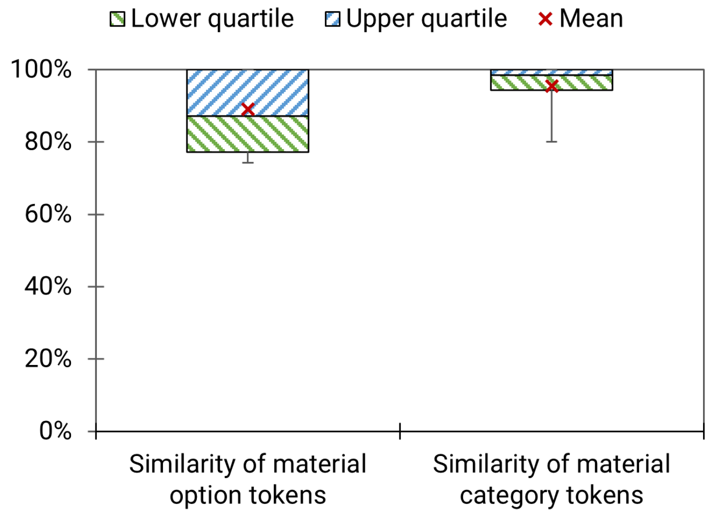



Figure 3.8: Shortest path similarity of identified, pre-matched material couples (IFC-LKdb) using GermaNet for case study 1

Nevertheless, the shortest path similarities of the identified pairs of synsets show promising results (Figure 3.8). The median of the similarity of material option tokens is 87.1%, and of the material category tokens, even 98.6%, both with little deviation. However, including the low synset identification rate of both material options and material categories from the LKdb, the total similarity are very low and not sufficient for being used in the proposed matching methodology.

#### spaCy

SpaCy is a pretrained neural network model and a promising implementation of the state of the art in the field of NLP (Honnibal & Montani, 2017). Its large German model ("de_core_news_lg") includes 500k unique vectors in its corpus and represents every word or expression with a vector of 300 dimensions. As sources for training data, existing corpi were used, such as e.g., TiGer Corpus (Brants et al., 2004).

For the results of spaCy and BERT, the vectorization of both tokens and whole spans of the material options and material categories are compared.

As shown in Figure 3.9, the ranges of the cosine similarity of all different comparisons, according to Section 3.4.3, differ a lot. Generally, the similarities of IFC materials to the material option spans have the worst performance, with the median being 13.6%. The tokenization improves the performance of matching the material performances up to a median of 60.0%. Also, the spans of the material categories are much better (median at 44.4%). The tokenization of the material categories improves the performance results by up to 60.3%. As an additional performance result, the maximum similarity of all comparisons (material option spans and tokens, as well as material category spans and tokens) is calculated. Its median is 74.4%, but also the quartile ranges improved compared to all other ranges. In general, the results are not sufficient for further usage in the proposed

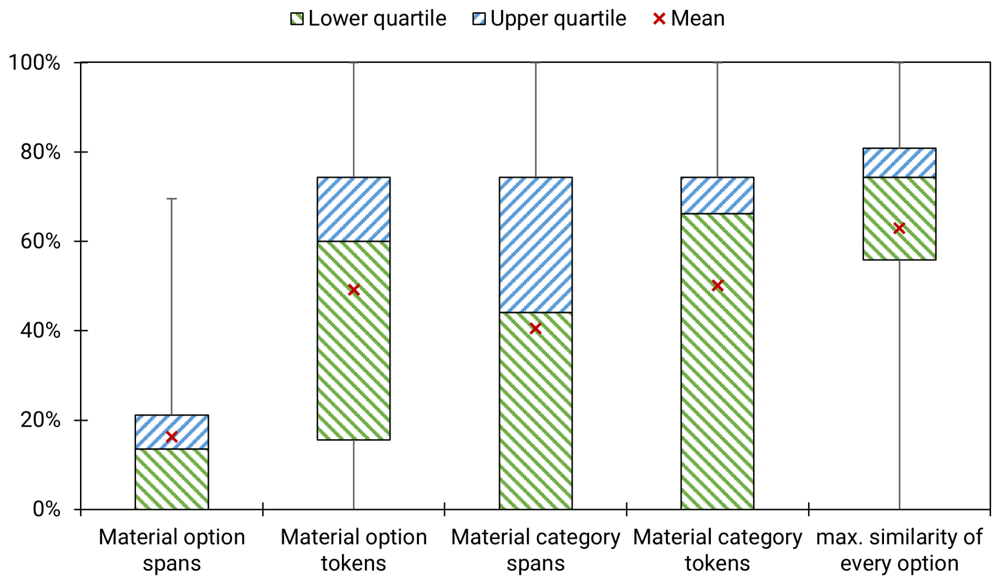



Figure 3.9: Cosine similarity of pre-matched material couples (IFC-LKdb) using spaCy for case study 1

framework but show a promising strategy for getting the maximum similarity of every option.

### BERT

BERT stands for Bidirectional Encoder Representations from Transformers and was released by Google in 2018 (Devlin et al., 2018). Transformers-based pretrained models are currently state of the art and are capable of solving a wide range of tasks as they "can represent the characteristics of word usage such as syntax and how words are used in various contexts" (Locatelli et al., 2021). BERT represents each word or expression with a vector of 768 dimensions, which is significantly higher compared to spaCy and makes the similarity calculation more time-consuming.

For the NLP technique BERT, the same similarity comparisons using cosine similarity are calculated as previously shown with spaCy. Figure 3.10 is showing the results as ranges of the material option spans and tokens and material category spans and tokens according to the workflow described in Section 3.4.3.

Generally, all result ranges differ much less compared to the results using spaCy. Additionally, all medians are between 79.2% (material category spans) and 87.2% (material option tokens). Also, the strategy of getting the maximum similarity of every option is improving the promising general results (median 87.7%). In addition, the minimum values of each result range show that BERT generally performs much better than spaCy.

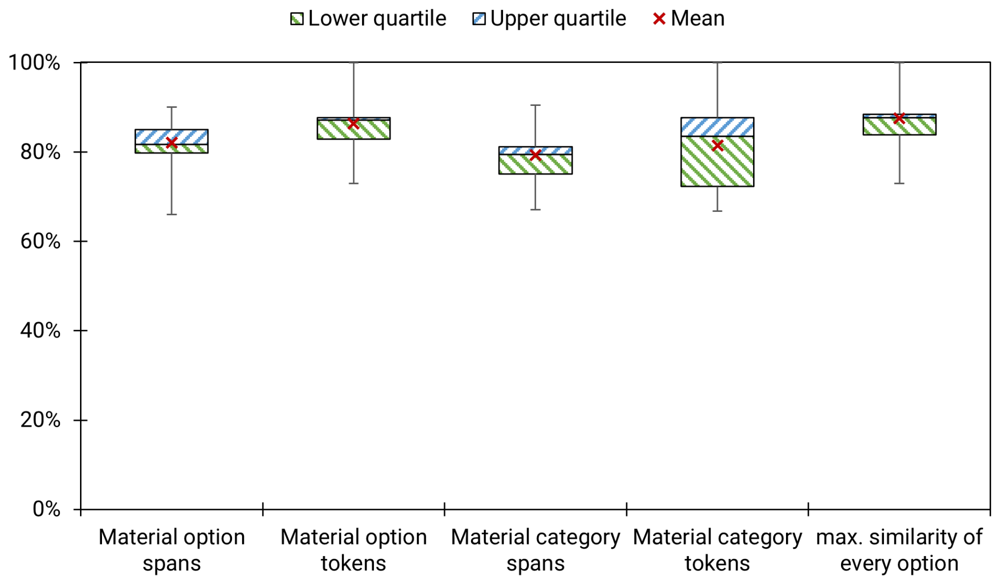



Figure 3.10: Cosine similarity of pre-matched material couples (IFC-LKdb) using BERT for case study 1

#### Conclusions regarding NLP-based matching performance

It was possible to apply all three NLP techniques to the case study, although their language body was not specifically trained for material expressions in the construction industry. While GermaNET shows promising results in the ranges of shortest path similarity, the identification rate of synsets is too low. Therefore, using GermaNET for the proposed matching methodology is not pursued further.

The NLP library spaCy shows that different strategies of calculating the cosine similarity of material option spans and material category spans are improving the results. Furthermore, the tokenization of both material options and material categories, as well as choosing the maximum similarity of every calculated option, improve the result ranges significantly. However, the ranges are deviating too much and are generally too low, so further consideration for implementation is not planned.

The NLP technique BERT showed the most promising results. Low deviations of the result ranges and high cosine similarity of all strategies lead to applying it for the matching approach. Nevertheless, due to its large vectors with 786 dimensions, the calculation time is significantly higher than with spaCy and needs to be considered for further optimization.

### 3.5.3 Evaluation of element matching method

In this Section, the proposed matching method is tested with real-world case studies. In the first step, five office buildings were chosen, consisting of the required model information, such as element classification according to DIN 276 and materials. In the next step, the performance of the previously proposed matching method on element level using the best-performing NLP model, BERT, is analyzed for all case studies. In the last step, the ratio of correctly matched versus complete set is evaluated for each case study depending on their specific model quality.

According to the proposed matching method, as shown in Section 3.4.3, all elements and their materials are filtered and encoded, the similarities are calculated, and finally, the most similar element is selected. To evaluate the performance of the proposed matching method, all matched elements are evaluated according to correctness. If not matched correctly, the reason for wrong matching is recorded. For validation, a manual element matching is set as ground truth, also using the same LKdb.

Besides correct and wrong element matching, there are other reasons why correct matching was not possible. As the LKdb is just taking the most common elements into account, it is not covering all potential element structures. Therefore, one of the reasons for incorrect matching is the insufficient amount of available elements. Another reason for incorrect matching is that there is no valid cost group classification according to the German classification system DIN 276 available for the element to be matched. As a result, the algorithm cannot filter the relevant list of elements in LKdb, and no default element can be selected. Furthermore, also wrong classifications of the model's elements can lead to incorrect matching. This reason will be described in more detail in the following Sections. Finally, incorrect matching can also occur if the element's name and material's name are too generic or not existing. In this case, the default element of the classification group is matched according to the proposed matching method. In total, there are five different error clusters:

a) correctly matched

b) no correct matching element available in LKdb

c) wrong element classification

d) no valid element classification

e) too little information/ details

f) wrong matching

Figure 3.11 shows the matching performance of all case studies summed up, once weighted by the amount of individual elements (left) and, on the other hand, weighted by the element areas (right). The area-weighted result shows the influence of wrong matching according to the LCA relevant quantities, while the element-weighted results show the performance compared to the manual matching step.

The total element-weighted matching performance results show a correct matching of 78.1% for all five case studies. The biggest drivers of incorrect matching are due to too little information/ details (8.62%), no correct matching element available in LKdb (5.65%), and wrong element classification (5.50%). Nevertheless, the different ratios between element-weighted and area-weighted matching performances differ so widely that wrong

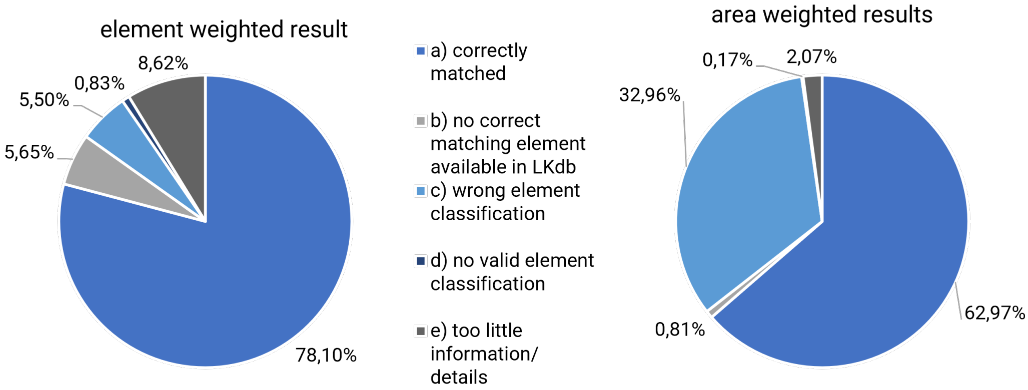
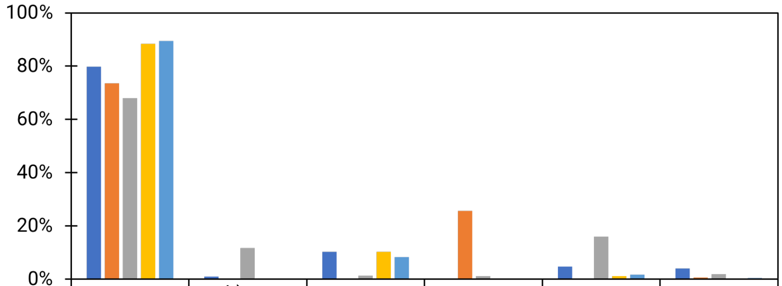



Figure 3.11: Total element matching performance of all case studies according to correct matches or matching error cluster, weighted by the amount of elements (left) and area of elements (right)

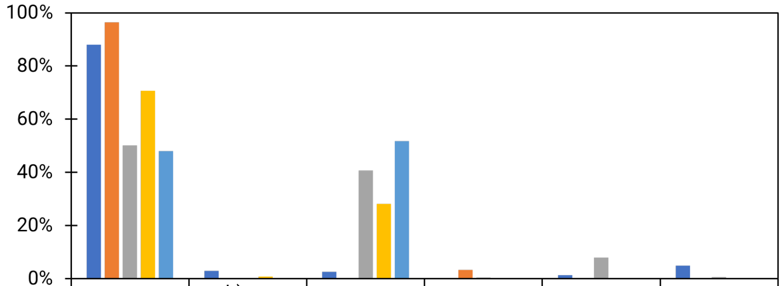



| | a) correctly matched | b) no correct matching element available in LKdb | c) wrong element classification | d) no valid element classification | e) too little information / details | f) wrong matched |
|---|---|---|---|---|---|---|
| Case study 1 | 79,76% | 1,04% | 10,24% | 0,14% | 4,83% | 3,98% |
| Case study 2 | 73,62% | 0,00% | 0,00% | 25,73% | 0,00% | 0,65% |
| Case study 3 | 67,98% | 11,68% | 1,30% | 1,13% | 16,04% | 1,87% |
| Case study 4 | 88,37% | 0,07% | 10,23% | 0,03% | 1,26% | 0,04% |
| Case study 5 | 89,40% | 0,00% | 8,33% | 0,02% | 1,70% | 0,55% |

Figure 3.12: Case-study-specific element matching performance according to correct matches or matching error clusters, weighted by the amount of element

element classification is 32.96%, and only 62.97% of the elements are correctly matched. Therefore, the results need to be analyzed in more detail and case-study-specific in the following.

As shown in Figures 3.12 and 3.13, there are major differences in the error clusters between the different case studies and the weighting scenario. When looking at the element-weighted incorrectly matched elements of case study 2, the main error is no valid element classification with more than 25.0%, which is mainly due to a different classification nomenclature for windows ("B20" instead of "334"). For weighting the scenario using the areas of the elements, the error is only 3.42%, and the correctly matched elements show the best performance of all case studies. Similar differences can be seen for case study 3, where the main error is due to clusters b) (11.68%) and e) (16.04%) in element weighting. In the area-weighted performance, these two clusters seem less significant compared




| | a) correctly matched | b) no correct matching element available in LKdb | c) wrong element classification | d) no valid element classification | e) too little information / details | f) wrong matched |
|---|---|---|---|---|---|---|
| Case study 1 | 88,05% | 2,94% | 2,54% | 0,08% | 1,42% | 4,98% |
| Case study 2 | 96,36% | 0,00% | 0,00% | 3,42% | 0,00% | 0,22% |
| Case study 3 | 50,23% | 0,12% | 40,60% | 0,44% | 7,90% | 0,70% |
| Case study 4 | 70,69% | 0,90% | 28,15% | 0,03% | 0,21% | 0,03% |
| Case study 5 | 48,07% | 0,00% | 51,72% | 0,00% | 0,19% | 0,02% |

Figure 3.13: Case-study-specific element matching performance according to correct matches or matching error clusters, weighted by the area of elements

to cluster c) (40.6%). This is due to the fact that the amount of elements is a different weighting factor. Nevertheless, as in case studies 4 and 5 are more columns modeled, which do not have the quantity of area but only length, the area-weighted performance results become significantly worse, although the element-weighted performance seems satisfying.

Generally, the matching performance shows satisfying results as, in total, 86,72% of the elements were correctly matched, or due to too little information, the default element was matched. 11,15% of the total elements were wrongly matched as there are not sufficient classifications available. For only 0,83% of the total elements, the matching method results in wrong matches. The performance results differ due to model complexity and the quality of correct element classification according to DIN 276 of each real-world case study. The quality of LOD, sufficient amount of elements in LKdb, and wrong matching due to the proposed methodology and chosen NLP model seem to have a minor influence on the matching performance. There can be different matching performances depending if the total amount of matched elements or their areas are considered, which is mainly driven by influences of columns without area quantity sets. Considering the fact that tested IFC models were not optimized for this use case, the performance results prove the proposed matching method for real-world projects. The performance can be further increased by checking the model requirements of the elements' classification.

### 3.5.4 Evaluation of LCA result range calculation

Next, we chose one case study to validate the whole semantic healing process by evaluating the calculation of the embodied GHG emissions. As case study 2 shows in the area-weighted performance the best results, we select it for calculating the LCA results.

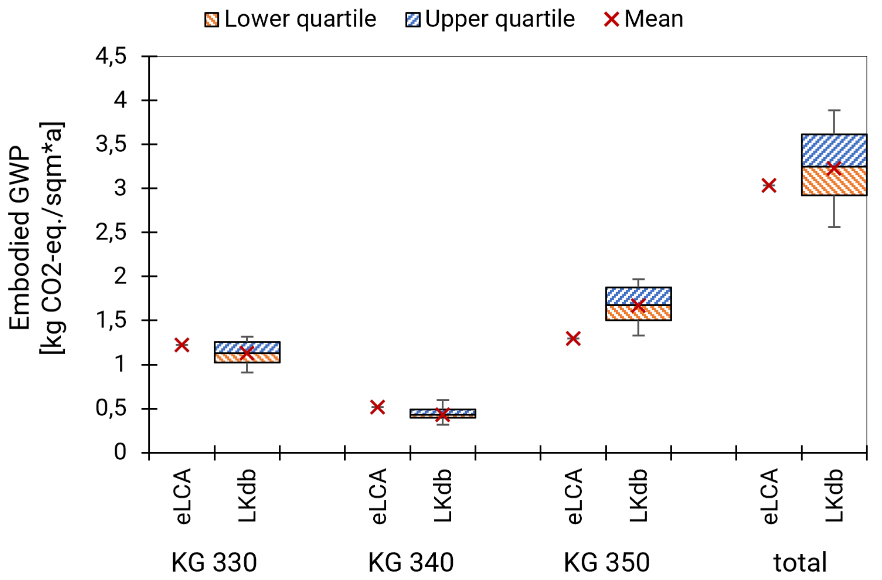



Figure 3.14: Total and cost group-specific results of Global Warming Potential (GWP) of case study 2 in [kg CO2-eq./ sqm*a]

The results will then be compared to a manual calculation, focusing on GWP as the main impact indicator. For the conventional LCA calculation, we chose the German LCA calculation tool eLCA (BBSR, 2022). Furthermore, only the total sum of all life cycle phases (A1-A3, B4, C3-C4, D) is considered to directly compare the final results of the examples. The reference period for this office building is 50 years, according to DGNB and BNB standards. The main goal of this evaluation is to show the results of the entire semantic healing workflow and its advantages compared to conventional processes. The optimization of element-specific LCA results itself is not the focus of this Section.

Figure 3.14 shows the GWP results clustered by cost groups (KG) and the total sum of the case study. Generally, the results show that the specific values of the conventional calculation following the manual, conventional workflow using eLCA are in the same range as the result ranges using the proposed methodology, including the matching method and the LKdb. The total manual result of 3,04 kg CO2-eq./ sqm*a calculated with eLCA is slightly lower than the range calculated by the proposed methodology and LKdb (Minimum 2,56, Median 3,25, Maximum 3,89 kg CO2-eq./sqm*a). To evaluate the difference in more detail, the element-specific results have to be analyzed.

Figure 3.15 shows the GWP results of the most relevant elements for each class according to the total sum of GWP over all life cycle phases. For each of the five chosen elements, on the left side, the results of the manual calculation using eLCA are shown, and on the right side, the automated calculated results using the matching method and LKdb are shown. The shown IFC elements consist of different element types, such as single- and multi-layer solid elements, windows and doors, or elements with composite materials. For the element with the cost group 331 and 351, reinforced concrete was matched, which consists of the materials reinforcement steel and concrete. While for the reinforcement steel, only one material option is available, for the concrete, there are several according to

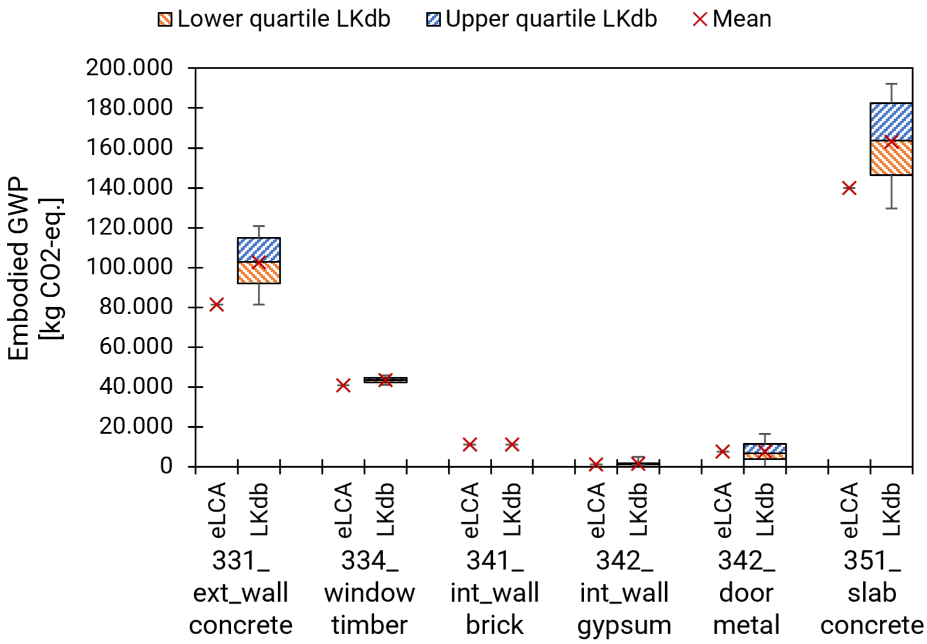



Figure 3.15: Element-specific results of Global Warming Potential (GWP) for selected elements of each classification group and different materials of case study 2 in [kg CO2-eq.]

the specific compressive strength, which hasn't been specified in this early design phase yet. These different material options lead to a range of results for the total GWP.

In comparison, for the element of the cost group 341, the monolithic brick wall was chosen, while only one material option of brick is available in this case. For this reason, both results of eLCA and LKdb are identical and do not differ. For the selected door (KG 344), different EPDs are used in the LKdb, while for the manual selection, only one EPD was chosen. Usually, the LCA calculation of windows needs different quantity inputs as solid elements, as the functional units for the window frame are the length of the perimeter and the area for the transparent glass. The only varying material for the implemented LKdb windows is the frame material, which is, in this matched case, wood. In the LKdb, glass was implemented as only one material option per element, either single, double, or triple pane, and is therefore not varying. The total GWP range is not varying a lot due to a few different wood-based frame options, but also close to the manual calculation results.

Finally, the interior wall (KG 342) consists of a multi-layer element of plasterboard and mineral wool. In the IFC model, the element consists of four different layers of plasterboard, while in the LKdb, there are only two. Therefore, the different thicknesses were summed up so that the total thickness for plasterboard layers is the same. Nevertheless, also in this case, there are 26 different material options for plasterboard, which leads to a range for the total GWP results.

In general, the evaluation of the whole process shows reliable GWP results compared to manual calculation using eLCA. The results depend on the different element types and the level of information, which was decisive for the matching. Another aspect is that with the manual workflow in early design stages, the total GWP results of this case study seem to be lower than the average of the result range derived from the proposed methodology. This underlines the need for a semantic healing process to enable more realistic LCA result ranges based on this uncertain information.

### 3.5.5 Limitations

The authors had to make a couple of assumptions to validate the proposed methodology, which led to certain limitations. For implementing the LKdb and its embodied emissions values, the German database Ökobaudat was used, as all the applied case studies are located in Germany, and German material naming was used. An extension using other databases and mapping them to elements and material options can be easily realized and has been prototypically tested (Lammers & Forth, 2022). Nevertheless, the implemented elements in the LKdb only cover the most common element structures. Specific element structures for special cases need to be included in future work. So far, neither operational energy simulation nor life-cycle cost calculation is included in the database, as the focus of this publication is solely on embodied GHG emissions. Although we only discussed GWP results for evaluating the LCA calculation, other environmental impact metrics have been calculated, too, such as AP, EP, POCP, and ODP, as well as energy-related impact metrics.

The results of the element matching of five case studies presented in Section 3.5.3 show that a correct classification is crucial to match the IFC element to realistic LKdb elements. However, the German classification system DIN 276 was used, which cannot be directly transferred to other countries' classification systems. If IFC models have no or a lot of incorrectly classified cost group elements, the LCA results will differ significantly and are not meaningful.

Furthermore, the NLP model BERT employed here was not specifically trained for the application in the AEC context. Nevertheless, the results from the material and element matching showed that this circumstance does not affect the results due to the robust selection process of the matching method. Nonetheless, the bidirectional trained model leads to a high amount of vector dimensions for each expression and, as a result, a time-intensive computation process. A specific trained model could decrease the computational effort while providing similarly satisfying results as with BERT. For training such a model, a high amount of real-world data from different companies and designers is needed, which is difficult to collect due to privacy issues.

## 3.6 Conclusions and future research

To enable the calculation of embodied emissions of buildings in early design phases, automated workflows based on BIM models can be used to compare different design alternatives and find those solutions that have a minimal environmental impact. However, the uncertainties in these stages are unavoidable and missing information can lead to erroneous LCA results. Therefore, enriching vague models is crucial for calculating meaningful results, which are usually a range of results rather than single values. Among the most challenging boundary conditions is the fact that early-stage BIM models often lack precise specifications of object types and material properties. Instead, a wide range of mixed terminology is used, and some information remains completely unprovided. With this unstructured data, however, finding correct LCA information from the respective databases is almost impossible.

To overcome this issue of manual material matching and vague model information, in this paper, we propose a novel approach for automated semantic healing of BIM models. The proposed method allows assigning correct LCA information of element types and materials to the respective model element such that a reliable and holistic LCA can be calculated in early design stages. For the semantic healing process, an NLP-based method is used to enrich the model by automatically matching elements of an LCA Knowledge database (LKdb) to close the missing gap of the automation process of enriching LCA datasets to IFC materials and elements, and adding missing layer information of imprecise model elements. This LKdb contains all relevant information for the LCA calculation process, including LCA datasets on material level and different design alternatives, such as element variants of the same classification group or different material options of each element layer. Missing element layers are added to ensure reliable and consistent LCA results. The elements are matched by the most similar material or element names using the cosine similarity of the pre-trained NLP model vectors.

In an initial evaluation, different NLP models were compared by aligning the results of pre-matched materials of a case study. BERT was identified as having the best-performing results and proved to be suitable for the element-matching method. In a second evaluation, the proposed matching method was tested using five real-world BIM models, and their performances were analyzed. Generally, the proposed matching method proved to be satisfactory, correctly matching the majority of the IFC elements (86,72% success rate in total) to the corresponding LKdb elements. Nevertheless, the importance of correct classification of the IFC models is a relevant requirement for correct element matching. The success rate depends on the semantic model quality, mostly on correct and valid element classification for the initial filtering step. In a third evaluation, one of the five case studies was selected to calculate the embodied emissions focusing on global warming potential of each element and summing the resulting ranges up for the whole building. These results were compared to a manually calculated LCA using the tool eLCA, showing that the manual results are in the range of the results using the proposed method.

Finally, answering the research question raised, it can be confirmed that the proposed automated semantic healing methodology is sufficient for calculating embodied emissions based on early design BIM models. The main limitations are the processing time of the prototypical implementation using large NLP vector dimensions and the correct element classification, which can be error-prone in a manual workflow.

In our ongoing research, we plan to investigate the visualization of the results and selection process of element variants or specific material options. Using the geometric BIM model as an interactive representation and mapping the LCA results as color ranges has great potential for the visualization and selection process. Furthermore, the developed methodology and the LCA Knowledge database will be extended according to other element groups, such as HVAC, as well as further criteria, such as for operational energy simulation, LCC calculation, or circularity aspects. These criteria will also be included in the visualization and selection process.

## 3.7 Acknowledgments

The authors would like to gratefully acknowledge the support of Siemens Real Estate (SRE) for providing financial support for this research and giving access to the analyzed case studies.

# Chapter 4

# Domain-specific fine-tuning of LLM for material matching of BIM elements and Material Passports

Previously published as: Forth, K.; Berggold, P.; Borrmann, A.: *Domain-specific fine-tuning of LLM for material matching of BIM elements and Material Passports*. Proceedings of ASCE International Conference on Computing in Civil Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.

## Abstract

Material Passports (MP) enable a combined assessment of life cycle assessments and circularity assessment of buildings. Semantically rich 3D models, such as Building Information Models (BIM), facilitate deriving consistent and automated creation of MPs. Nevertheless, a time-consuming effort is still needed to manually match material and element information to automate the BIM-based MP. To improve this step, we propose a method of semi-automatically matching BIM materials to the relevant material datasets using Semantic Textual Similarity and fine-tuning pre-trained Large Language Models (LLM). The method matches the semantically most similar environmental material datasets to every BIM material to enrich further information. We are fine-tuning the LLM by proposing different strategies, such as adding domain knowledge, testing different loss functions, applying different labeling, adding negative pairs or filtering, and using manually matched pairs of datasets from 23 real-world case studies. Combining different strategies for fine-tuning a pre-trained LLM significantly increases the accuracy of the proposed method of matching BIM elements and materials to environmental material datasets.

## 4.1 Introduction

In 2020, buildings and the construction industry were responsible for 36% of the Greenhouse gas (GHG) emissions as well as for ca. 37.5 % of the waste generation within the European Union (European Commission, 2020). To tackle the insufficient documentation for realizing the reuse and recycling of buildings, the concept of material passports (MP) is introduced in different scales, such as material, product, or building (Çetin et al., 2023). Building Information Models (BIM) contain geometric and semantic information about buildings and can facilitate MPs (Honic et al., 2019b). Besides precise quantity take-offs, further semantic information about the elements' layers, materials, and detachability can be included. However, manual steps are still required for enriching materials from circularity databases to those used in BIM (Honic et al., 2019b), as architectural nomenclature differs from the more precise databases. These manual enrichments are expensive in costs and labor. We define this automated enrichment step as the primary technology gap to be addressed by this publication.

To close this gap, we propose a novel method to automatically enrich open BIM models with material information from LCA and circularity databases using Natural Language Processing (NLP) and its subtask Semantic Textual Similarity (STS). Usually, the naming of IFC materials is more generic, e.g., “pre-cast concrete”, while datasets for MPs are more specific, e.g., “reinforced concrete” with specific compressive strengths. In a previous publication, we showed a similar approach using a well-structured database and a pre-trained Large Language Model (LLM) for LCA (Forth, Abualdenien, & Borrmann, 2023). In this publication, though, we propose a domain-specific fine-tuning of pre-trained LLM using different strategies for this task. These include domain-specific abbreviations, loss functions, and additional information from the BIM model. Our method is based on open BIM data formats, such as Industry Foundation Classes (IFC).

## 4.2 Background & Related Works

### 4.2.1 BIM for Material Passports

Recently, different researchers have proposed BIM-based methods for material passports (MP). The findings suggest that LCA-based BIM plugins have significant potential for improving circularity in early design stages but emphasize the importance of data accuracy, effective management, clear guidance for modeling, and increased knowledge in implementing LCA and circular economy concepts. Honic et al. introduced a BIM-based MP approach to optimize the recyclability of buildings. However, they identified the manual material matching by a specialist as a significant obstacle (Honic et al., 2019b). Atta et al. developed a framework for digital MPs using BIM, considering the deconstructability of elements (Atta et al., 2021). However, their approach is based on the BIM authoring tool Autodesk Revit and is limited to its closed BIM workflows. Gebetsroither et al. compared current BIM-based approaches for building Material passports mainly in the German-speaking market (Gebetsroither et al., 2024). They came to the conclusion that the approach by Madaster and from EPEA is currently practicable, and the BIM integration not only saves time but also supports the documentation and archiving of the building. The discussed approaches lack open BIM data exchange and a fully automated process of matching material datasets from external databases to those of the BIM model. The detachability, deconstructability, or connection types of elements for circularity assessments are out of the scope of this publication and part of future research.

### 4.2.2 Fine-tuning Large Language Models

As most large language models (LLM) were trained on generic text, they do not always fit well into domain-specific tasks. Accordingly, domain adaptation needs to be applied for domain-specific use cases. Usually, domain adaptation is fine-tuning a pre-trained language model (PLM) on a domain-specific, new dataset. This fine-tuning process adjusts the original model's weights, aligning them with the specific attributes of the domain data and the targeted task. Reimers and Gurevych present Sentence-BERT (SBERT), “a modification of the pre-trained BERT network that uses Siamese and triplet network structures to derive semantically meaningful sentence embeddings that can be compared using cosine-similarity" (Reimers & Gurevych, 2019). Their approach focuses on semantic textual similarity (STS) and outperforms other sentence embedding methods. For improving the performance of fine-tuning BERT in a multitask domain, such as sentiment analysis, paraphrase detection, and STS, Jadwin and Huang employed an in-domain pre-training and Multiple Negative Ranking Loss Learning (MNRL) (Jadwin & Huang, 2023). They concluded that MNRL fine-tuning leads to the highest performance optimization impact. Sachidananda introduced adaptive tokenization (AT), a method for efficiently adapting PLMs to new domains by expanding the tokenization vocabulary with domain-specific token sequences (Sachidananda et al., 2021). AT achieves significant performance improvements without requiring further language model pre-training, offering a promising approach for domain adaptation in natural language processing tasks. Generally, these methods show different approaches for domain adaptation and fine-tuning of pre-trained language models, which will be further discussed later.

## 4.3 Method

### 4.3.1 General workflow

As shown in Figure 4.1, the general workflow consists of four main steps. The first step includes the BIM modeling in the authoring tool (1.a) and the IFC export (1.b). The detailed requirements for the IFC export are described in the implementation section 4.2. The next step 2.a contains a quantity take-off of all relevant elements and materials using the base quantities of each element and layer. Step 3.a describes the main part of the proposed method, called semantic model healing. The quantity take-off derived in the previous step is used to automatically match the corresponding datasets from the material database (3.b). For this process of semantically healing the IFC model, the highest semantic similarity of the material datasets with each material of each IFC element is used. We fine-tune a monolingual LLM domain specifically for this task based on the German language (3.c). In the final step 4.a, we can upload the semantically healed and enriched IFC model to a Material Passport platform for further analysis, such as life cycle assessments (LCA) or circularity assessments.

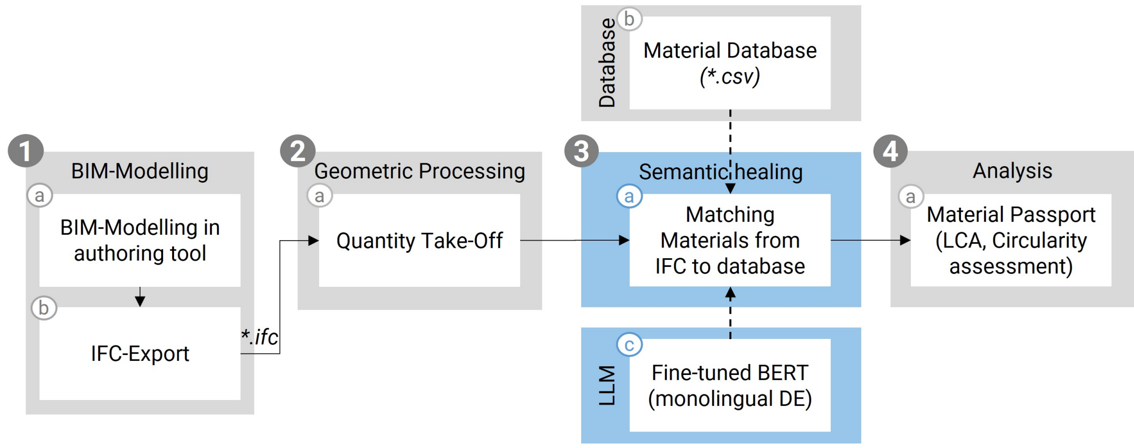



Figure 4.1: The general workflow of semantically healing IFC models for Material passports.

### 4.3.2 Strategies for fine-tuning domain-specific LLM and improving matching performance

We propose five different strategies for the domain-specific LLM fine-tuning to improve the STS and matching performance: Strategy 1 - Adding domain-specific abbreviations, Strategy 2 - Applying different loss functions for fine-tuning, Strategy 3 - Adding different/ multiple labels for further context information (Element name, classification, IfcType, etc.), Strategy 4 - Adding negative/ contradicting word pairs, and Strategy 5 - Filtering word pairs according to IfcType. As shown in previous studies with similar model healing tasks but for different analysis types (Reimers, 2023; Sachidananda et al., 2021), domain-specific abbreviations were a big challenge for the matching approach. Therefore, our first fine-tuning strategy is to train LLM with these AEC- and BIM-specific abbreviations. For the second strategy, the following suitable loss functions are proposed for our application of fine-tuning using manually matched word pairs. In brackets, the typical labels for positive or negative word pairs are shown according to (Reimers, 2023):

a) Cosine Similarity Loss: manual positive matches (0.8), negative matches (0.3)

b) Contrastive Loss: positive (1), negative (0)

c) MNRL – Multiple Negatives Ranking Loss: no labels needed

Another strategy for improving the fine-tuning performance is to add further knowledge of the BIM models using different labels for each type of information in the training process. For every material pair, we also know the IFC element name, the IfcType, and usually the classification. With the Softmax loss function, we can use different labels for fine-tuning, including this additional information. Therefore, we propose the following labels: Abbreviations (0), IFC material – positive material dataset (1), IFC element name – positive material dataset (2), classification – material dataset (3), IFC Type – material dataset (4), IFC material – contradicting material dataset (5). As shown for the last label, the fourth strategy for improving the fine-tuning and matching performance is to include negative

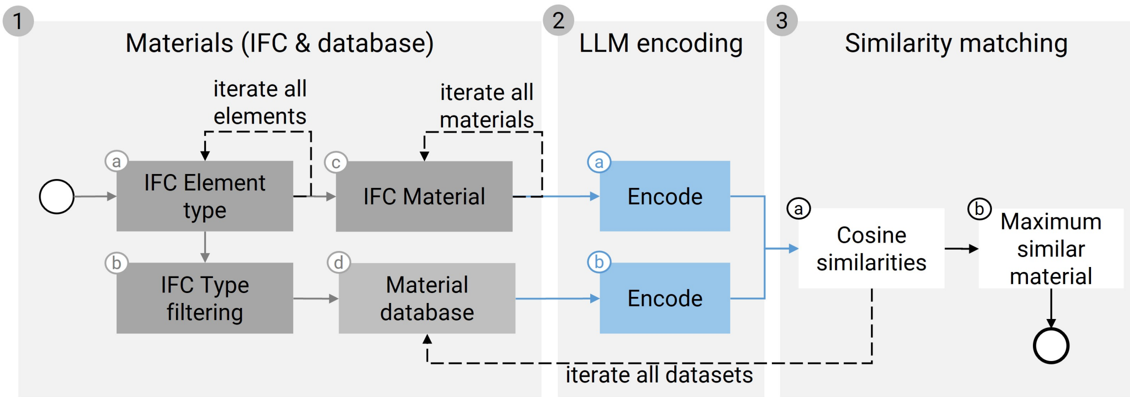



Figure 4.2: Automated material matching process of IFC materials to MP datasets

pairs. The MNRL has only positive word pairs with anchor ai, and pi being positive. But it assumes all other positives pj are the negative pairs, so ai and pj for i!=j are negative pairs. We can manually create negative pairs for all other loss functions according to the same logic but also check that pi!=pj. The negative labels have already been introduced in paragraphs of the previous strategies. This strategy can be realized by the Cosine Similarity Loss (4a) as well as Contrastive Loss (4b). The fifth strategy includes a filtering step of the material database. As the used material database is unstructured, we add a filter structure using different IfcTypes, such as IfcWall, IfcSlab, IfcCovering, IfcColumn, IfcDoor, IfcWindow, and IfcRoof. To enrich only applicable material datasets per IfcType, we check for all positive word pairs for their related IfcType and save the material dataset. Instead of comparing all 387 material datasets for each IfcMaterial, we can limit the material datasets to 82 for IfcSlab, 79 for IfcWall, 51 for IfcCovering, 21 for IfcColumn, etc. This strategy doesn't improve the fine-tuning process but improves the matching performance afterward.

### 4.3.3 Combination of different strategies

We briefly describe how different strategies can be combined with each other to improve the fine-tuning process and match performance further. Adding abbreviations (strategy 1) and the filtering process (strategy 5) can be combined with all different strategies. Multiple labels (strategy 3) can only be realized with the Softmax Loss function, as the other functions don't allow multiple labels. Nevertheless, negative pairs can be realized with the Cosine Similarity Loss and Contrastive Loss functions. MNRL already incorporates the negative pairs, as described in the previous subsection. Different Loss functions could be combined in case more model context was fine-tuned with multiple labels and Softmax, and this LLM is used afterward as the base model for another fine-tuning process with Cosine Similarity, Contrastive, or MNR Loss function or in the opposite order.

### 4.3.4 Matching approach of highest semantic similarity

Figure 4.2 shows the general matching workflow of matching the semantically most similar material of the material database to each IFC element layer's IFC material. To this end, first (1), all IFC elements are iterated (1.a) and, next, filtered according to their IfcType if the filtering strategy (S5) is applied. If not, we go to the following step (1.c) of iterating for each element and its material layers. These materials are then compared with the whole or the filtered material database, so the material datasets are iterated (1.d). Each IfcMaterial (2.a) and each material dataset from the database (2.b) are encoded in the next step using different fine-tuned LLM. Next, the cosine similarity with each material from the database is calculated for STS (3.a). The material from the database with the highest cosine similarity is selected as the matched material for each IFC material (3.b).

## 4.4 Case studies and implementation

### 4.4.1 Case studies and datasets

For the domain-specific abbreviations, we used 571 general AEC abbreviations and their descriptions (Bundesamt für Bauwesen und Raumordnung 2021) and 155 BIM-specific abbreviations (Helmus et al., 2021), such as construction types or material abbreviations. Both abbreviation datasets are in the German language. We employed 23 IFC models as case studies, where the IFC materials had already manually enriched PSets for the Madaster Platform (Frank, 2021). Besides the availability of the provided case studies by LIST Eco, the Madaster platform is one of the few Building Material Passport providers using open BIM data format and the only commercial platform that embeds and manages a portfolio of MPs of several buildings (Gebetsroither et al., 2024). The case study projects are a mix of logistic, residential, and office buildings from different designers and clients. This assures a high diversity in the data and real-world adaptability. The matched materials from the material database mainly include 387 EPEA datasets (EPEA GmbH, 2022), but some were customized and added to the overall database. We extracted the matches of IFC materials and MP dataset following Madaster-specific PropertySet called "MaterialOrProductName". Based on these case studies, we derived 245 unique material matches and split them into 75% training, 184 positive word pairs for training, and 25% test samples (61 test pairs).

### 4.4.2 Prototypical Implementation

To implement the proposed method of fine-tuning domain-specific LLM, we used the cased version of the German BERT model ('bert-base-german-cased') as a base model for training (Chan et al., 2020). All IFC models, including their element and material names, and the EPEA database are provided in German language. For the prototypical implementation of the training pipeline, we used SentenceTransformers packages based

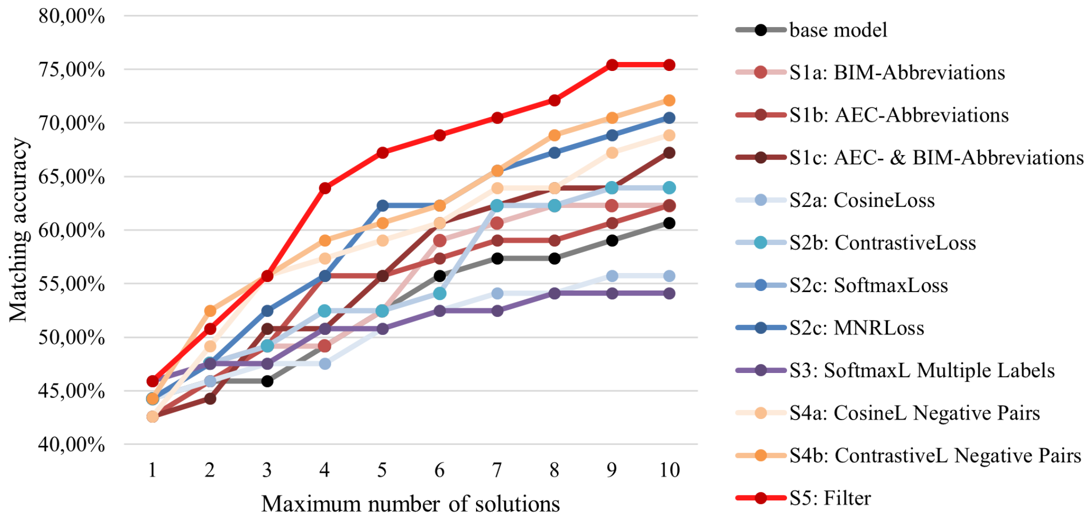



Figure 4.3: Matching accuracy for strategies 1- 5 compared to the base model

on the SBERT method by Reimers and Gurevych (Reimers & Gurevych, 2019). These packages incorporate all mentioned loss functions from Subsection 3.2. The different labeling for the additional domain knowledge from the IFC models was pre-processed accordingly after parsing all quantity take-offs.

## 4.5 Results

### 4.5.1 Results of the overall matching approach using different fine-tuned LLM strategies

Figure 4.3 depicts the achieved matching accuracy, so the ratio of correct and total matches/ predictions, when applying the different strategies. Instead of only showing the correct matches of the most similar solution, we add a continuous solution space of the maximum ten most similar matches. This is because the initial results would not have a significant difference, and a deeper analysis would not be possible. The results show that the base model ('bert-base-german-cased') has only 44,26% correct matches, taking the most similar match into account, but increases up to 60,66% of correct matches considering the top 10 maximum similar matches. Different individual matching accuracies exist for domain-specific information considering AEC-overall and BIM-specific abbreviations. Each abbreviation source slightly increases the matching accuracy. But by combining both abbreviations, the matching accuracy significantly increases to 67,21% for the top 10 matches. Adding multiple labels (strategy 3) increases the matching accuracy for the top four matches, but it even underperforms the base model for the following matches. The highest increase in the matching performance is using the filter strategy. The filtering is applied to the base model and reaches up to 75,41% correct matches. The loss function with the highest matching accuracy is Multiple Negative Ranking Loss (MNRL), which already considers negative pairs, followed by Contrastive Loss. Cosine

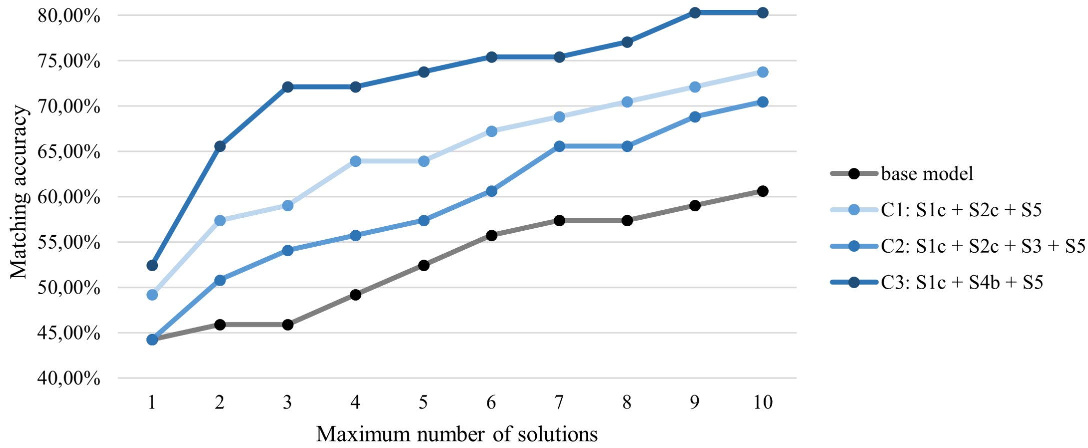



Figure 4.4: Matching accuracy for combined strategies compared to the base model

Similarity Loss even underperforms compared to the base model. Nevertheless, adding negative pairs significantly increases the matching performance. Overall, the loss function with the highest matching accuracy is Contrastive Loss, including negative pairs, although it's computationally more expensive than MNRL by a factor of ca. 100.

### 4.5.2 Results of the overall matching approach using different combinations of the strategies

We defined a combination set of different strategies as follows:

- C1: Training AEC-, BIM-Abbreviations (S1c), Material Datasets with MNRL (S2c) and filtering (S5)
- C2: Training AEC-, BIM-Abbreviations (S1c), Material datasets with MNRL (S2c), multiple labels with SoftmaxL (S3) as base model, and filtering (S5)
- C3: Material Datasets with ContrastiveLoss including negative pairs (S4b), AEC- and BIM abbreviations using MNRL (S1c) as base model, and filtering (S5)

Figure 4.4 shows the results of the matching accuracies of combining different strategies compared to the base model. Generally, the results indicate that combining the individual strategies increases the matching accuracy even more. Nevertheless, adding more context with multiple labels (S3) did not improve the overall performance (see Section 5.1). Adding this strategy to the combination of C1 lowers the accuracy. The best-performing combination of strategies is C3, reaching up to 80,33% matching accuracy for the top ten matches. The following section analyzes the results of the best-performing individual strategies and the best-performing combination.

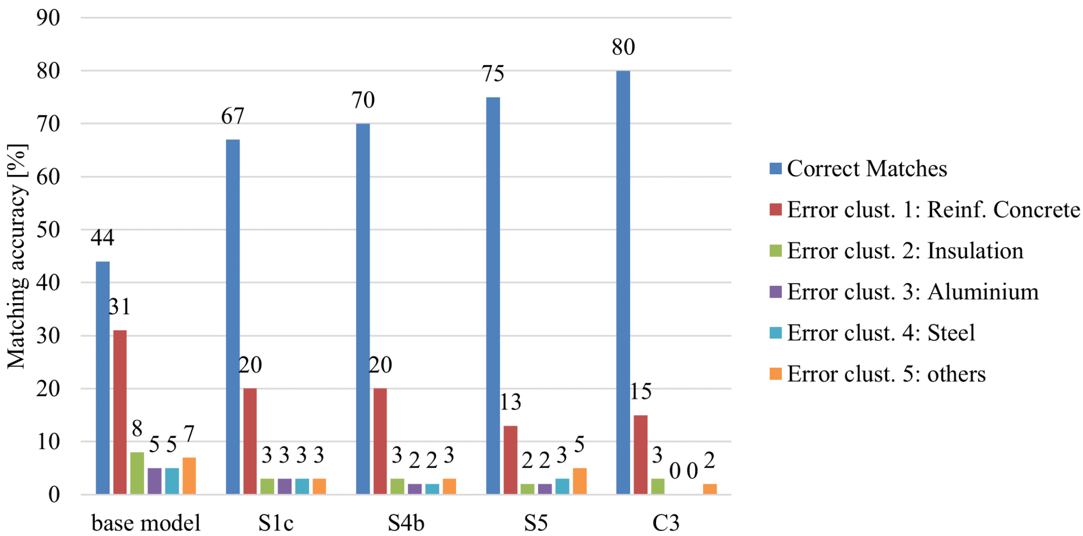



Figure 4.5: Matching analysis of correct and reasons for wrong matches comparing the base model and strategies S1c, S4b, S5, and C3

### 4.5.3 Analysis of correct and wrong matches

As shown on the left side of Fig. 4.5, the base model has approximately 44% correct matches but ca. 66% wrong matches. The reason for false matches was classified according to the main material group of the IfcMaterials. Most of the wrong matches (31%) are related to reinforced concrete. Accurate matching is challenging, as most of the IfcMaterials are named "Stahlbeton" (Engl. "Reinforced Concrete"), but the Material Datasets are more diverse, including specific compression strength classes and reinforcement ratios. The reason for wrong matches with aluminum, steel, and others (mainly asphalt and Larch wood) is primarily that in the IfcMaterial, more than one material is included, while in the ground truth, only one Material Dataset is matched. For the insulation materials, there are in the ground truth matching instead of "XPS" other Material Datasets matched, such as "EPS", making direct matching impossible. This can be avoided by including multiple similar materials.

Compared to the base model, the strategies S1c and S4b decrease the error of wrongly matching reinforced concrete by 11% (error cluster 1) and the error cluster 2-5 to 10-12%. Furthermore, only adding the filter reduces the error cluster of reinforced concrete by 18% and the other error clusters by 13%. Combining these three strategies solves the errors with steel and aluminum, and only the insulation error remains at 3% due to wrong classification. Furthermore, error cluster 5 still includes the error with the specific wood material. The error cluster 1 about reinforced concrete remains and can't be further solved. This is mainly because, for the wrong matches, different reinforcement ratios are added, which can not be predicted by the IfcMaterial alone. There are 23 different material datasets for reinforced concrete with varying ratios of reinforcement, priorities, or user-specific datasets. This issue could be handled by adding this information before automatically matching to increase matching accuracy.

### 4.5.4 Limitations

The most significant limitation of this publication is the limited number of matching samples in the dataset. From 23 real-world BIM models, only 245 unique matching samples were extracted, so the fine-tuning process took place with 186 samples. Also, the test datasets were limited to 61 matches. However, besides the limited number of samples, their quality also limits the accuracy of the matching. As previously analyzed, having multiple IfcMaterials matched to only one input limits the overall performance. Another limitation is that in this study, we only took one LLM network architecture into account. As we have German material expressions, we used the German version of BERT ('bert-base-german-cased') as the base model (Chan et al., 2020). Finally, there is no 100% matching accuracy possible. This means this approach is a support tool than fully automating the process. For this reason, we included the Material Dataset with the highest similarity and extended it to the top 10 most similar samples.

## 4.6 Conclusion & Outlook

In this paper, we introduce a method of (semi-)automatically matching BIM materials to the relevant material datasets using Semantic Textual Similarity (STS) and different strategies of domain-specific fine-tuning pre-trained Large Language Models (LLM). The method matches the semantically most similar material datasets to every BIM material for further analysis. We used the German BERT LLM and sentence embeddings using Siamese BERT-Networks for fine-tuning. The five strategies and their combination increase the matching accuracy from 44,26% to 80,33% by extending the solution space to ten material datasets with the highest semantic similarity. Therefore, the low matching accuracy of the most similar match leads to using this method as a support tool instead of a fully automated approach. Although we had 23 real-world case studies, the 245 material samples with different data quality are still limited. In our future research, we will use more case studies and material samples for training and testing with cross-validation for more robust solutions. Furthermore, a more structured database, rather than only differentiating by IfcTypes, could increase the accuracy of the matching in the filtering step. As we identified too many similar material datasets for reinforced concrete, we suggest an interim step of adding more information about reinforcement ratio and priority. Additionally, these fine-tuning strategies shall be transferred to multilingual training for enriching building energy models for building performance simulations (Forth, 2023a). Last, more information is missing, such as the connection type of different elements to derive the detachability index to enable fully automated circularity assessments in early design stages.

#### Acknowledgements

We thank LIST Eco GmbH & Co. KG for providing real-world case studies.

# Chapter 5

# Semantic enrichment for BIM-based Building Performance Simulations using Semantic Textual Similarity and fine-tuning multi-lingual LLM

Previously published as: Forth, K.; Borrmann, A.: *Semantic enrichment for BIM-based Building Performance Simulations using Semantic Textual Similarity and fine-tuning multi-lingual LLM*. Journal of Building Engineering, 2024, DOI: 10.1016/j. jobe.2024.110312

## Abstract

To achieve the global targets of the Paris Agreement of limiting global warming, it is necessary to reduce the operational energy of buildings, which are responsible for around 30% of the global greenhouse gas emissions. Building Energy Performance Simulation (BEPS) is an established method to estimate the building's energy demand in early design stages. Building Information Models (BIM) provides geometric and semantic information to create precise Building Energy Models (BEM) in early design stages. However, manual enrichment of missing semantic information is still a time-consuming and laborious process. Therefore, we propose a novel methodology to automatically enrich missing information to BIM using Semantic Textual Similarity (STS) and fine-tuned Large Language Models (LLM). For every IfcSpace, we match room-specific space types and constructions with missing thermal properties using the semantic most similar pairs of the BIM model and the according databases. We use three real-world case studies to fine-tune LLMs, and two case studies evaluate the whole methodology. Different fine-tuning strategies, such as using different loss functions, adding opposing word pairs or domain-specific abbreviations, significantly improve the accuracy of the matching. At the same time, however, findings show that semantic matching based on multilingual fine-tuned LLM performs worse than translated, monolingually fine-tuned LLM. Finally, BEPS results from automatically enriched BEM only slightly deviate from manually enriched BEM.

## 5.1 Introduction

According to the United Nations, buildings contribute to 40% of the world's greenhouse gas emissions, mainly in the operational phase (United Nations Environment Programme, 2022). In order to reduce the environmental emissions of building operations, conducting energy simulations throughout the entire design process assists in quantifying and enhancing the building's energy performance. Building Energy Performance Simulations (BEPS) include simulation of the whole building energy demand as well as thermal comfort. Early design stages show a significant influence on decision-making and optimizing the later real-life demand (Gao et al., 2019).

Building Information Modeling (BIM) digitally represents a physical building design geometric, semantic information, and topological relationships (Borrmann et al., 2018). It can support reducing the effort of Building Energy Modeling (BEM) by automating the energy simulation workflow. The research field of *BIM to BEM* has been ongoing for the last two decades. Several challenges on interoperability and data exchange formats or geometric transformation from volumetric BIM models to surface-based BEM (van Treeck et al., 2018) have been solved. Nevertheless, in early design stages, BIM models contain vague information or are even lack relevant information for a holistic and reliable whole-building energy simulation.

Furthermore, most of the current BIM-to-BEM approaches map or match BEPS-related semantic information to BIM objects either manually (Di Biccari et al., 2022; Raggi et al., 2021) or using templates (Müller et al., 2021). Manual matches mean an increased amount of time, effort and required expertise, while template-based approaches usually do not take semantic information from the BIM models into account but rather enrich it based on statistical information. Additionally, in early design, detailed BEPS-related properties, such as room-specific schedules and occupancies or element- and material-specific thermal properties, are not yet defined. In the following, we focus on the Development Design as the early design stage, as defined by Schneider-Marin & Abualdenien (Schneider-Marin & Abualdenien, 2019). Moreover, most of the current approaches propose ideal workflows but do not take the diversity of real-world case studies into account. Therefore, they are not robust in this regard.

To the authors' best knowledge, the gap of semi-automated enriching BEPS-related information robustly based on real-world Building Information Models has not been filled yet.

To close the described research gap, the main contribution of this publication is to propose a novel methodology for semi-automating the enrichment of BIM-based room-specific space types and element-specific thermal properties using Semantic Textual Similarity and multilingual fine-tuning pre-trained Large Language Models (LLM). For the semantic enrichment process, we use existing databases by the National Renewable Energy Laboratory (NREL) and propose adapted matching methods for the room-specific space types and element-specific thermal properties. Based on a number of German real-world case studies, we test different fine-tuning strategies in order to improve the multilingual matching performance.

In summary, we aim to answer the following research questions:

a) How can semantic textual similarity and fine-tuning of pre-trained LLM support the two multilingual matching use cases of space types to architectural rooms and thermal construction to BIM elements in early design stages?

b) Which fine-tuning strategies improve the multilingual matching accuracy for each use case?

c) Is a fully or semi-automated matching possible to ensure reliable BEPS results?

To answer these research questions, we first propose the general workflow, the matching methods, and the multilingual fine-tuning process and strategies. Second, we test the methodology using five different case studies and a prototypical implementation. Finally, we evaluate the matching accuracy of the different strategies of the room and element-specific matching, and the final BEPS results of the overall workflow.

The publication is structured as follows: Section 5.2 introduces the most relevant background and works related to the topics of BIM-based BEPS, the current BIM to BEM matching and semantic enrichment approaches, Natural Langauge Processing (NLP) in the AEC industry, Semantic Textual Similarity (STS), as well as multilingual fine-tuning of pre-trained LLM. Afterward, Section 5.3 presents the methodology of multilingual semantic enrichment of BEPS, including the general workflow, geometric transformation to surface models and BEPS, and describes in more detail the semantic enrichment part, as well as the multilingual fine-tuning part. The proposed methodology is then evaluated in Section 5.4 using different case studies differentiating between the room-specific matching of space types, the element-specific matching of constructions' thermal properties and BEPS results, and finally, highlighting the limitations. Finally, our conclusions and aimed future research are discussed in Section 5.5.

## 5.2 Background and Related works

### 5.2.1 BIM-based building energy performance simulation

Building energy performance simulations (BEPS) is a term mostly used as a synonym for building energy simulation (BES), building performance simulation (BPS), building energy modeling (BEM), or energy simulation (Hong et al., 2018). This umbrella term includes simulation for mainly energy demand but also indoor environmental qualities, such as thermal comfort.

Ciccozzi et al. reviewed interoperability strategies in BIM to BEM workflows (Ciccozzi et al., 2023). They identified four strategies, including real-time connection, standardized exchange formats and middleware tools, MVD, and proprietary toolchains. They also analyzed different approaches of automatically mapping energy-related information to the BIM model.

Eckstädt et al. conducted a comparative analysis encompassing three distinct pathways for calculating whole-building energy simulations, often referred to as building performance simulations. Their investigation was grounded in utilizing open BIM models and Industry Foundation Classes (IFC) data format as the primary input file (Eckstädt et al., 2022). However, the prevailing tools still confront certain challenges arising from accurate IFC export settings tailored to each simulation tool, limitations within the IFC import process, and inherent constraints of the simulation tools themselves. Löhr et al., in a different vein, proposed a partially automated process for constructing multi-zone thermal models through IFC models (Löhr et al., 2022).

Ramaji et al., on their part, introduced an alternative approach for transforming IFC-based BIM into BEM (Ramaji et al., 2020). Their methodology involved a direct conversion of IFC models into OpenStudio's native IDD format. However, this conversion posed challenges that were addressed using Model View Definitions (MVDs). Similarly, Spielhaupter, in his master's thesis, undertook a comparison of diverse IFC-based strategies for BIM to BEM transformations (Spielhaupter, 2021).

In a related study, Yang et al. also employed IFC files as the foundation of their approach. However, their route differed in that they initially transformed the IFC data into the gbXML format before progressing to the IDF (EnergyPlus native file) (Y. Yang et al., 2022). Nonetheless, their workflow overview indicates that additional adjustments are frequently necessary.

Hence, the proposed strategy aims to leverage the HBJSON format as the transformation schema. This choice is informed by its open-source nature, versatility in being converted to various file formats (such as gbXML, IDF, etc.), and its heightened reliability in geometric export. The IFC format will be only used as open BIM input for further geometric and semantic transformation towards a fully enriched HBJSON format which is then used as input for BEPS.

### 5.2.2 Matching or mapping BEPS-related semantic information to BIM objects

After more than 15 years of research in the field of Building Information Modelling (BIM) to Building Energy Modelling (BEM), this research domain continues to grapple with unresolved challenges. Gao et al. identified the automated conversion of intricate spatial functions across all rooms as a pending avenue for future investigation, given that the present process remains manual (Gao et al., 2019). Elnabawi's assessment similarly underscored the persistent manual nature of assigning occupancy operating schedules, emerging as a central obstacle in achieving seamless interoperability (Elnabawi, 2020). In alignment with these findings, Raggi et al. concluded that "additional data (e.g., regarding some HVAC data [...], must be manually added [...] to the models before the energy simulation can run" (Raggi et al., 2021). Notably, occupancy and operating schedules are specifically linked to Heating, Ventilation, and Air-Conditioning (HVAC) data.

Di Biccari et al. currently reviewed the State-of-the-art and research trend in the field of BIM and BEPS interoperability (Di Biccari et al., 2022). They request that “research should propose practical solutions for describing occupancy and MEP component schedules in BIM” (Di Biccari et al., 2022). They also identified that even if thermal properties are available in authoring tools, the IFC exports still need manual mapping as postprocessing.

Besides room-specific information, such as program-specific internal loads, occupancy, etc., thermal properties of constructions and elements need to be mapped or matched from libraries or databases to BIM models. Wimmer et al. presented an approach using mapping rules (Wimmer et al., 2015). Nevertheless, a “continuous adaptation has to be done on the mapping rule side”. Kim et al. proposed an approach of mapping IFC information to building energy analysis models (Kim et al., 2016). Nevertheless, only if the material name from the IFC model matches exactly the one in the ASHRAE library, the thermal properties can be enriched. If not, a new material and its properties need to be manually added.

Müller et al. used a template-based approach to enrich thermal properties based on IFC models in their BIM2SIM research project (Müller et al., 2021). They are using templates from a project called TEASER(Remmen et al., 2018). Richter et al. recently further developed this approach to improve the accessibility of thermal comfort analysis (V. E. Richter et al., 2023). Li et al. recently developed an algorithm to map BIM-based objects, such as spaces, using invariant signatures of AEC objects (H. Li et al., 2023). Different features, such as coordinates, surfaces, or quantities of the surface, are used for this algorithm to calculate the distance between IFC and IDF zones.

Generally, the current State of Art still shows a research gap in the field of robustly matching BEPS-related semantic information to BIM objects, especially in early design stages, such as the Design Development stage (Schneider-Marin & Abualdenien, 2019), when the BIM model contains vague information. Hence, this publication focuses on the (semi-)automated matching and enrichment of space types to architectural rooms, and construction with thermal properties to BIM elements and materials.

### 5.2.3 Semantic enrichment of BIM

Various cutting-edge techniques are available for the automated enhancement of meaning to serve different objectives. In her analysis, Tanya Bloch presented diverse strategies, methodologies, and domains of application for enriching the semantic content of BIM (Bloch, 2022). Within this context, two primary avenues were recognized: the utilization of IFC to represent building information coupled with inference-driven enhancement and the integration of IFC with external data sources. Furthermore, Semantic Web technologies were harnessed to process building information. Among the trio of key application domains, emphasis was placed on building design, performance assessment, and particularly energy simulations (Scherer & Schapke, 2011). Noteworthy instances encompassed the employment of NLP to classify spatial elements for Korean school buildings (Song et al., 2019), and the derivation of regulations for code compliance through rule extraction (Guo et al., 2021).

Bloch and Sacks propose a semantic enrichment approach of BIM models by comparing a machine learning approach and a rule-based approach for room type classification (Bloch & Sacks, 2018). They used their own classification for labeling room types, a large number of 32 case studies for training, and supervised artificial neural networks with different features. However, this approach depends on large datasets for training and does not use transfer learning.

Costa and Sicilia adopted semantic query languages to facilitate the automatic conversion of BIM data, concentrating on harmonized data models to facilitate building-scale energy simulations utilizing EnergyPlus (Costa & Sicilia, 2020). In a distinct approach, Baumgärtel et al. harnessed ontologies to dynamically modify and assess thermal energy performance in building contexts (Baumgärtel & Scherer, 2016).

Generally, the majority of semantic methodologies leverage ontologies, semantic web tools, and linked data concepts to achieve automated semantic enhancement in the realm of BIM. While a few also employ NLP, its application for augmenting intricate insights from BIM models for energy simulations remains limited. Therefore, this study uses NLP for automatically enriching the semantic information of BEM, as it uses pre-trained Large Language Models and transfer learning and, therefore, promises to work with fewer case studies for training.

### 5.2.4 Natural Language Processing for AEC-related tasks

The realm of Natural Language Processing (NLP) has demonstrated substantial advancements in research, highlighting enhancements in both performance and user-friendliness. This progress has extended its relevance beyond academic boundaries, permeating diverse sectors, including the construction industry.

Cornago et al. published a SWOT analysis using Transformers for Life Cycle Assessment (LCA) studies that reveals internal strengths, including automation and integration support and relatively low marginal costs (Cornago et al., 2023). However, they formulate concerns about data quality, electricity intensity during model training, and the rapid evolution of technology. External opportunities include community-building and enhanced data availability, but threats include a lack of regulation, standardization, and interdisciplinary talent. Transformers have the potential to benefit LCA practitioners by addressing scalability issues and enabling data-driven environmental decision-making support.

Zheng et al. explore how domain-specific corpora can enhance deep learning and BERT-based models for Information Retrieval (IR) tasks in the AEC domain (Zheng et al., 2022). They find that domain corpora improve traditional word embedding models for some tasks but have a negative effect on others. In contrast, BERT-based models consistently outperform traditional models. In consequence, they created a high-performing model called RegulatoryBERT. This work provides valuable insights and resources for future investigations and applications in the AEC domain.

Wu et al. conducted a comprehensive review of Natural Language Processing (NLP) usage in construction management, highlighting improved information extraction and document organization (C. Wu et al., 2022). They also discussed the potential and challenges of NLP applications in construction management and serve as a valuable resource for project teams seeking to leverage NLP techniques for smart construction.

Although developing NLP for other tasks in AEC is becoming increasingly relevant, to our best knowledge, none of the reviewed papers uses fine-tuned LLM for automated semantically enriching BIM models for BEPS.

### 5.2.5 Semantic Textual Similarity

Natural Language Processing (NLP) covers different tasks, such as text summarization, text classification, named entity recognition, sentiment analysis, and more. In this subsection, we are mainly focusing on the basic concepts for deriving Semantic Textual Similarity (STS) as another task of NLP. The fundamental technique for this NLP task is sentence pair modeling, which is also used for Natural Language Inference (NLI), also known as Recognizing Textual Entailment (RTE) (Lan & Xu, 2018). It requires an understanding of semantic similarity between a hypothesis and its premise (Bowman et al., 2015).

Chandrasekaran and Mago surveyed the evolution of semantic similarity and distinguished between knowledge-based, corpus-based, deep neural network-based, and hybrid methods (Chandrasekaran & Mago, 2022). Each approach has different advantages and disadvantages. However they identified the trends towards building more semantically aware embeddings and transformer models. There are different types of corpus-based semantic similarity methods, most of them using cosine similarity to measure the distance between word vectors (Chandrasekaran & Mago, 2022). Other distance measures used in the field of STS are Euclidian distance or Manhattan distance (R. Li et al., 2023).

Corpus-based semantic similarity methods use word or sentence embeddings, which are vector representations of words, including linguistic relationships between words (Tobias Schnabel et al., 2015). Word embeddings represent individual words, while sentence embeddings represent whole sentences.

One of the most used pre-trained word embeddings include word2vec (Mikolov et al., 2013) or BERT (Devlin et al., 2018). BERT consists of a pre-training step and a rather inexpensive fine-tuning step and can also incorporate sentence embeddings. Reimers and Gurevych present "Sentence-BERT (SBERT)", a modification of the pre-trained BERT network that uses Siamese and triplet network structures to derive semantically meaningful sentence embeddings that can be compared using cosine-similarity" (Reimers & Gurevych, 2019). Their approach focuses on semantic textual similarity (STS) and outperforms other sentence embedding methods.

Usually, big data sets are needed to train, fine-tune and evaluate semantic textual similarity. One of the most used datasets for this task is the Stanford Natural Language Interference (SNLI) Corpus consisting of 570k human-written English sentence pairs (Bowman et al., 2015). For validating STS tasks, the STS benchmark is considered as a standard benchmark (Cer et al., 2017). For our use case, however, there is only a small number of parallel sentences or word pairs available, which makes the usual benchmarks obsolete.

However, we use STS for semantically matching and enriching BIM information to BEM, and, additionally, the introduced SBERT approach is used for fine-tuning LLM based on a small number of training datasets.

### 5.2.6 Multilingual fine-tuning of pre-trained Large Language Models

As most large language models (LLM) were trained on generic text, they do not always fit well for domain-specific tasks. Therefore, domain adaptation is applied. Usually, domain adaptation is fine-tuning a pre-trained language model (PLM) on a domain-specific, new data set (Kohle & Jannidis, 2020). This fine-tuning process leads to an adjustment of the original model's weights, aligning them with the specific attributes of the domain data and the targeted task. The platform Huggingface provides multiple PLMs in different languages or also multilingual LLMs that can be fine-tuned for further domain-specific tasks (Wolf et al., 2019).

To fine-tune pre-trained LLM, the choice of the correct loss function is significant depending on the training data and the overall task.

For improving the performance of fine-tuning BERT in a multitask domain, such as sentiment analysis, paraphrase detection, and semantic textual similarity (STS), Jadwin and Huang employed an in-domain pre-training and Multiple Negative Ranking Loss Learning (MNRL) (Jadwin & Huang, 2023). They concluded that MNRL fine-tuning has the highest impact on optimizing performance.

Contrastive loss is another loss function proposed by Hadsel et al. where either the distance between two embeddings is increased (label 0) or reduced (label 1) (Hadsell et al., 2006). Cosine Similarity Loss uses the manual label of the expected cosine similarity between two embeddings to fine-tune the LLM. Reimers and Gurevych used the Softmax Loss when introducing their concept for sentence embeddings using Siamese BERT-Networks (Reimers & Gurevych, 2019). In their approach to fine-tune multilingual LLM, Reimers and Gurevych are using the mean square loss for training the student model (Reimers & Gurevych, 2020).

Besides different loss functions and fine-tuning frameworks, pre-trained multilingual LLMs are also important for the fine-tuning process.

Conneau et al. introduce a multilingual masked LLM, "XLM-R", trained with 2.5 TB in 100 languages, including German and English, taking the advantages of multilingual XLM models and the monolingual RoBERTa model (Conneau et al., 2019). They showed that their pre-trained multilingual LLM reduces the amount of parallel training data needed to achieve sufficient performance results.

Feng et al. propose a language-agnostic BERT Sentence Embedding, "LaBSE", approach focusing on multilingual sentence embeddings (Feng et al., 2020). LaBSE supports 109 languages, including German and English, and promises good results for finding translation pairs in multiple languages, as well as for assessing the similarity of sentence pairs without translations.

Reimers and Gurevych proposed another framework using knowledge distillation to make multilingual sentence embeddings out of monolingual ones (Reimers & Gurevych, 2020). They published several versions of a pre-trained multilingual LLM, the latest called "distiluse-base-multilingual-cased-v2". This version supports more than 50 languages, including German and English, and uses multilingual universal sentence encoders (Y. Yang et al., 2019). Their advantage is that it can work with few samples, and the hardware requirements for training are lower.

## 5.3 Methodology of multilingual semantic enrichment for BEPS

In the following section, we first introduce the general workflow for the proposed methodology for semantic enrichment for multilingual Building Energy Performance Simulations (BEPS). Next, we briefly describe the different steps, such as geometric transformation from volumetric BIM models to BEPS surface models. Nevertheless, the main focus of this publication is on semantic enrichment, including relevant databases, the multilingual matching method of space types to architectural rooms and thermal properties of constructions to BIM elements, as well as the workflow for multilingual LLM fine-tuning and different domain-specific strategies for improving the fine-tuning and matching performance. Based on these matches, we describe the further semantic enrichment part for room-specific enrichment of space types and construction sets. Finally, we briefly describe which BEPS metrics we are integrating with this methodology.

We are applying Peffer's Design Science Research (DSR) as our research method (Peffers et al., 2012). Therefore, the proposed methodology serves as the artifact that answers the defined research questions representing the design and development phase of the previously described problem and motivation as well as the objective phase. The prototypical implementation demonstrates the utility and suitability of the artifact using different real-world case studies. The evaluation and results section represents the evaluation phase, while this publication serves as the communication phase.

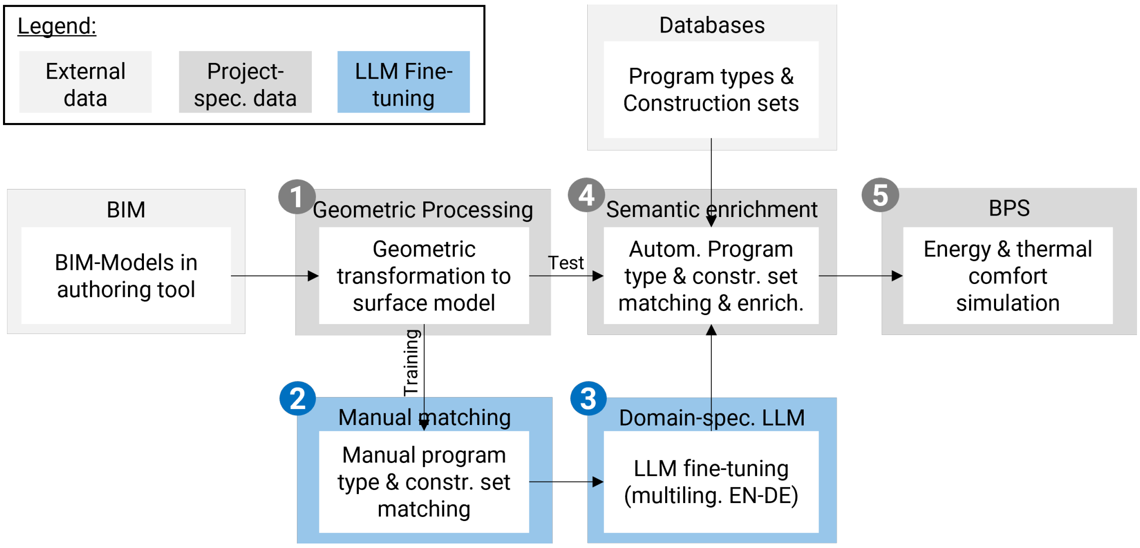



Figure 5.1: General workflow of proposed methodology of multilingual semantic enrichment for building performance simulations

### 5.3.1 General workflow

As shown in Figure 5.1, the general workflow consists of five steps comprising external and project-specific data processing and LLM fine-tuning. As external data, we consider the Building Information Models in the authoring tool, as well as the space types and construction sets databases, which we introduce in more detail in Section 5.3.3. The first processing step has the BIM model as input and geometrically transforms it to a surface model. In the next project-specific processing step (4), the semantic information consisting of space types and construction sets are automatically enriched by the semantically most similar matches using Semantic Textual Similarity (STS) and Large Language Models (LLM). The matching of space types and construction sets are described in more detail in Sections 5.3.3 and 5.3.3. After this matching, we run the building performance simulations (BEPS), such as annual energy demand or thermal comfort simulation.

Previously, we have shown that the existing LLM is not sufficient for matching the semantic information (Forth, Hollberg, et al., 2023). Therefore, we propose in this method two fine-tuning steps. We split the five case studies into training and testing case studies. For the training case studies, we extract the manually matched pairs of architectural rooms and space types, and the element-construction or material pairs for LLM fine-tuning. For the test case studies, we use the whole building BIM and BEM for testing the semantic matching, enrichment and whole BEPS for evaluating the proposed method. The testing data are used to validate the proposed workflow while the training data are only used to fine-tune pre-trained LLM. First (step 2), the space types and construction sets of the training case studies are manually matched. Next, we fine-tune pre-trained LLM using the manually matched space types and construction sets and their BIM-related expressions, as well as further domain knowledge, which is described in more detail in Section 5.3.4.

### 5.3.2 Geometric transformation to surface model

As the main focus of this publication is the automated semantic enrichment of space types and construction sets based on BIM in early design stages, we include existing methods for the geometric transformation to surface models. In the prototypical implementation, we use the Pollination Plugin for Autodesk Revit for the geometric transformation (Pollination, 2024). We manually extrude the rooms according to each level's floor plan using each level-specific height. As boundary location, we use the wall center so that no gap occurs between adjacent rooms. Finally, we store the created surface models in gbXML and HBJSON data format for further semantic enrichment steps and BEPS. In our future work, we aim to base both steps, the geometric transformation and semantic enrichment, on the IFC data formation. However, as we exclude the geometric transformation step of our evaluation, we directly derive the geometrically transformed HBJSON file from the authoring tool, while the semantic enrichment step is based on the exported IFC model.

### 5.3.3 Semantic enrichment

This section first introduces the relevant databases of space types and construction sets by NREL. Next, the matching method is described. First, the use case of matching space types to architectural rooms, and second, the use case of matching constructions with thermal properties to BIM elements is introduced. Finally, we explain the workflow of the integrated enrichment process for both use cases. Generally, for this step IFC models are used as input to semantically enrich the previously geometrically transformed BEM in HBJSON format, which will be finally used as input file for the BEPS.

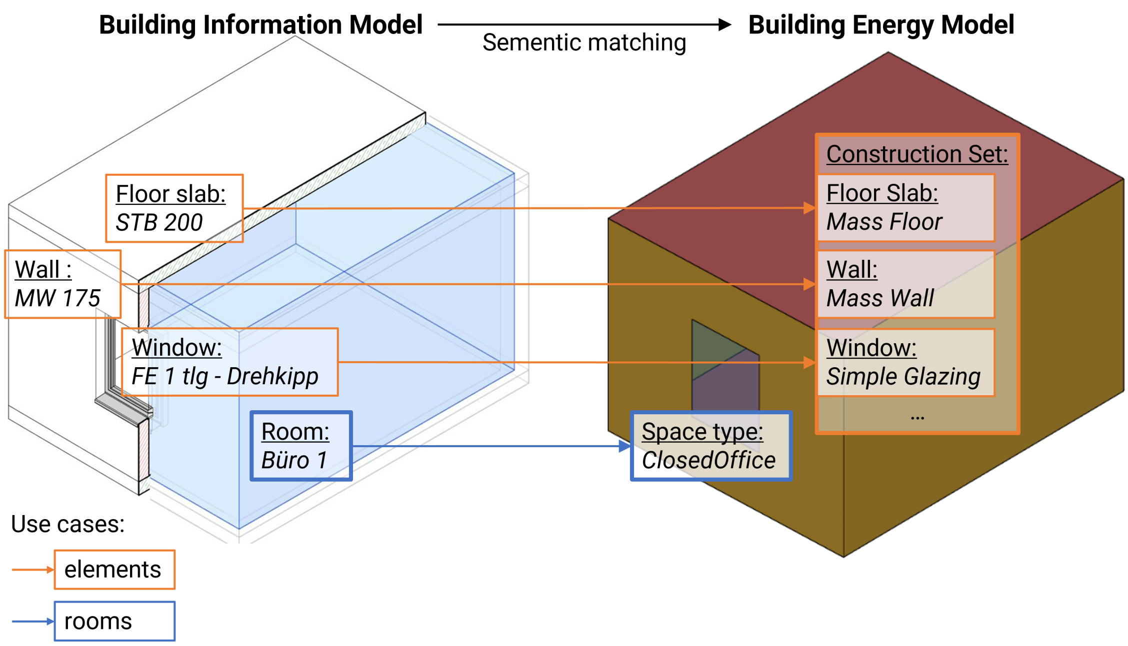



Figure 5.2: Semantic matching from BIM to BEM of space type (use case "room") and constructions sets (use case "elements") for an example room and its elements

Figure 5.2 shows exemplarily the two use cases of semantic matching from a simple Building Information Model in German to the derived Building Energy Model with English terminology. Semantic matching is the first step in semantic enrichment. It is divided into the two use cases of matching space types to architectural rooms (use case "room") and matching constructions and construction sets with thermal properties based on the BIM elements and its materials of the BIM model (use case "elements"). A construction set is a set of different construction types, such as walls, slabs, windows, etc., and will be defined in more detail in the following section.

#### Relevant databases

For enriching the BEM with room-specific space types, also known as program types or end-use load profiles, we use the standardized database by the National Renewable Energy Laboratory (NREL) of the U.S. Department of Energy (Wilson et al., 2021) providing the target terminology. Space types are necessary for building energy simulations and contain all relevant MEP settings and information related to a room or zone. It contains 23 different building types and 224 different space types, and its ontology is shown in Figure 5.3, including internal loads of lighting, infiltration, ventilation, electric equipment, and different schedules. For matching the semantically most similar space type, we use the identifier which consists of the year of the space types, e.g., 2019, the main usage or building type, such as "Large Office", and the room or zone-specific usage, e.g., "open office", "corridor", or "conference". For the enrichment step, we add all information attached to the matched identifier.

For enriching the construction sets of the building, we use the official construction set and material database by OpenStudio, which was also co-developed by NREL (NREL, 2024). The ontology of these construction sets is shown in Figure 5.4. Every construction set consists of a "WallFloorSet", a "RoofCeilingSet", a "DoorSet", and an "ApertureSet". Apertures consist of a window construction, while walls, floors, roofs, and ceilings are opaque constructions. Doors can be either opaque or window construction. Opaque construction consists of one or multiple opaque materials that have different thermal properties. Window constructions consist of a glazing system, the glazing material and the gas filling, each having different thermal, and radiant properties.

For the construction set matching, the opaque construction and opaque materials identifiers are important. For example, the construction "Typical Insulated Exterior Mass Wall-R10" consists of the materials "Stucco", "CONCRETE HW RefBldg", "Typical Insulation-R8", and "Gypsum". The thickness description in the identifier can be neglected in the matching process.

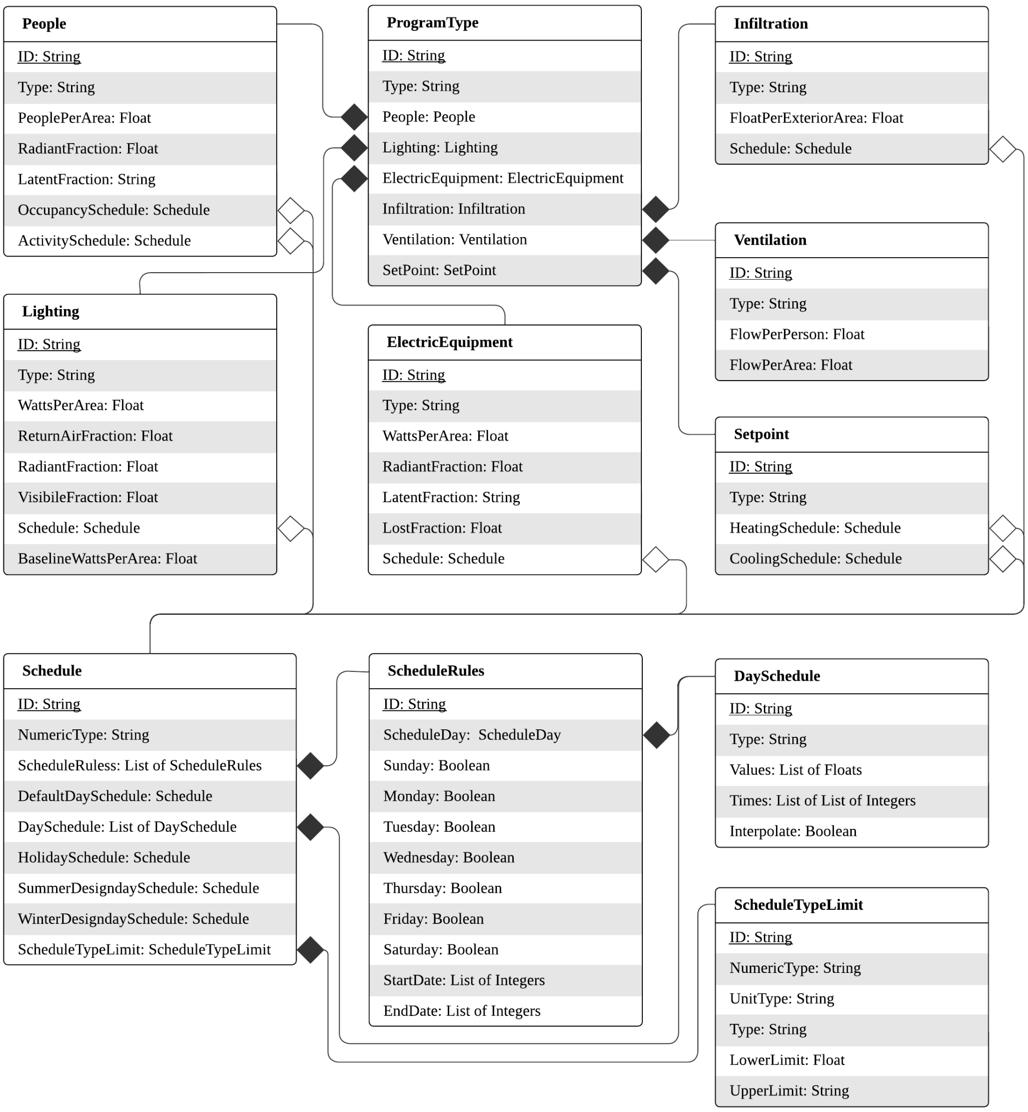



Figure 5.3: UML diagram of the space or program types by the NREL (Wilson et al., 2021)


Figure 5.4: UML diagram of construction sets, and opaque & window constructions & materials by the NREL (NREL, 2024)

### Matching method of space types to architectural rooms

As shown in Figure 5.5, the matching method consists of three main steps: Program filtering (1), similarity calculation (2), and finally, space type selection and model enrichment (3). These steps follow the same logic as in previous proposed methods for LCA enrichment (Forth, Abualdenien, & Borrmann, 2023). The filtering step is necessary to narrow down the solution space. First, the main usage is filtered consisting of several space type identifiers, as shown in Section 5.3.3. Next, for every IfcSpace, the semantic similarities of the whole expression and of the tokenized space are calculated. Different fine-tuned LLM are used for the encoding, and for the similarity calculation, the cosine similarity of the two encoded vectors is calculated. The maximum similarity of both the tokenized and the whole expression is identified and iterated through all filtered space types of the main usage. Finally, the space type with the highest similarity is selected to the energy model for each IfcSpace for further enrichment in the next step.

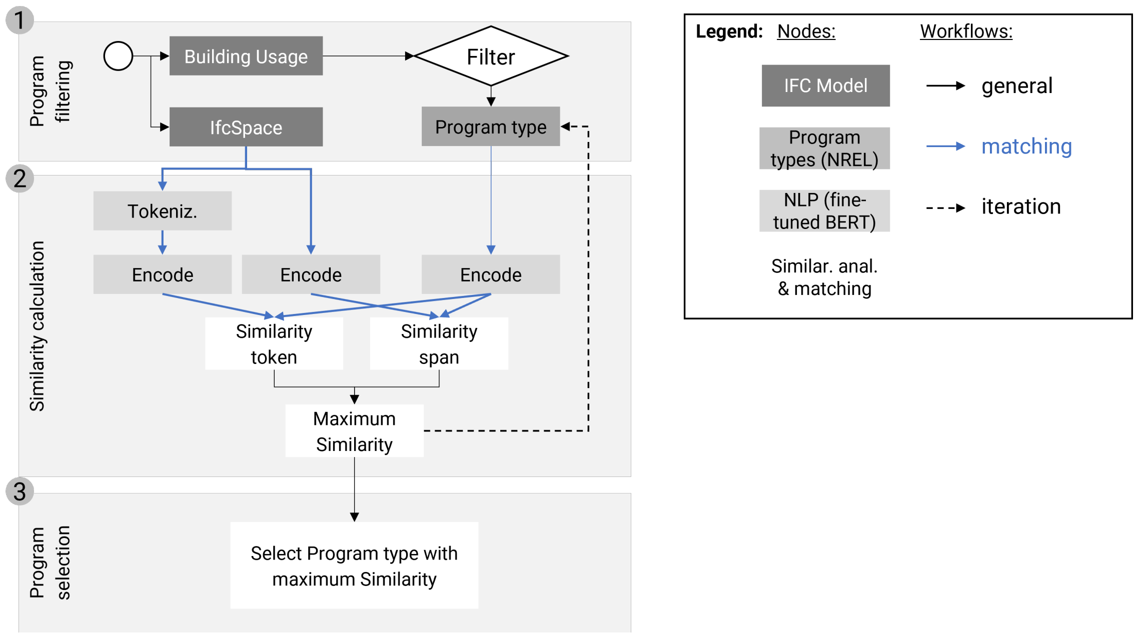



Figure 5.5: Matching space types to architectural rooms based on IfcSpaces using Semantic Textual Similarity

#### Matching method of thermal constructions to elements

Following the logic of the space type matching, see Section 5.3.3, and the LCA element matching (Forth, Abualdenien, & Borrmann, 2023), the construction matching method consists of three main steps: construction filtering (1), similarity calculation (2), and construction selection and model enrichment (3), as shown in Figure 5.6. Nevertheless, the method is more complex, as more filters need to be applied to narrow down the solution space, and the matching can happen on the construction or element level, as well as on the material level, similar to the element matching in (Forth, Abualdenien, & Borrmann, 2023).

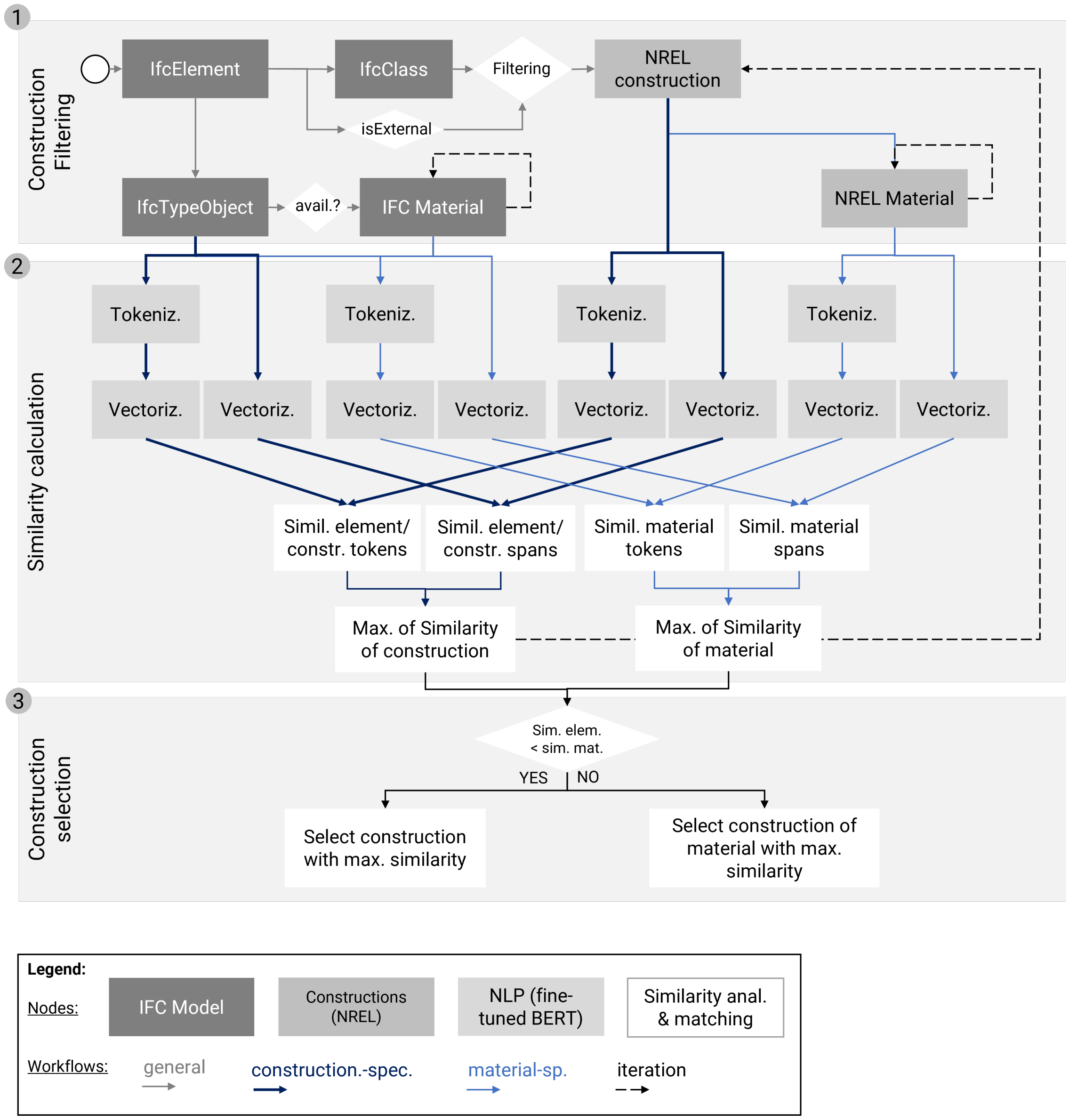



Figure 5.6: Matching thermal constructions to IfcElements using Semantic Textual Similarity based on material and element level

First, for every IfcElement, the unique IfcTypeObject is identified to decrease the calculation effort by matching only based on unique object type descriptions and materials. For example, IfcTypeObject can originate from the same family in Autodesk Revit. Next, for every IfcTypeObject, the availability of IfcMaterials is checked. If there are no IfcMaterials available in the model, the matching method can still be processed based on the IfcTypeObjects. If IfcMaterials are available, the matching method iterates through all available materials for each IfcTypeObject.

For the filtering step (1), for every IfcElement, the IfcClass is identified to narrow down the solution space. According to its internal description, we structured the NREL constructions database into the following groups, which can be directly assigned the IfcClasses:

- External Walls ("IfcWall" or "IfcWallStandardCase" or "IfcCurtainWall" and "isExternal")
- Internal Walls ("IfcWall" or "IfcWallStandardCase" and not "isExternal")
- External Floor ("IfcSlab" and "isExternal")
- Internal Floor ("IfcSlab" and not "isExternal")
- Roofs ("IfcRoof)
- Ceiling ("IfcCovering")
- Doors ("IfcDoor")
- Windows ("IfcWindow")

Window constructions already have a different schema, are identified with "IfcWindow", and don't have to be restructured. After the filtering, all relevant NREL constructions are iterated, and their related materials are iterated, too.

In the next step (2), the similarity calculation takes place, both on construction and on material level. For every expression, IfcTypeObject, IfcMaterial, NREL construction identifier, and NREL material identifier, the whole expression or spans and the tokenized one are encoded, and the similarities are calculated using the cosine similarity. The results lead to four similarities between the tokenized and the whole construction or element expressions, as well as the tokenized and whole material expressions. In the next step, the maximum similarity on the construction level and on the material level for each IfcTypeObject are identified and iterated over all filtered NREL constructions and their iterated materials. However, as in the NREL database, there are a lot of similar constructions, which only differ by the thermal resistance and its classification, for example, "Typical Insulated Exterior Mass Wall" and "Typical Insulated Exterior Mass Wall-R2". We are neglecting these classifications and excluding similar ones from the matching. Instead of 201 constructions, we take 20 different ones into account.

For the final step (3), these two similarities are compared. If the maximum similarity of the construction level is higher than the similarity on the material level, this construction is selected for enrichment in the next step. If the similarity on the material level is higher, the construction of the most similar material is selected.

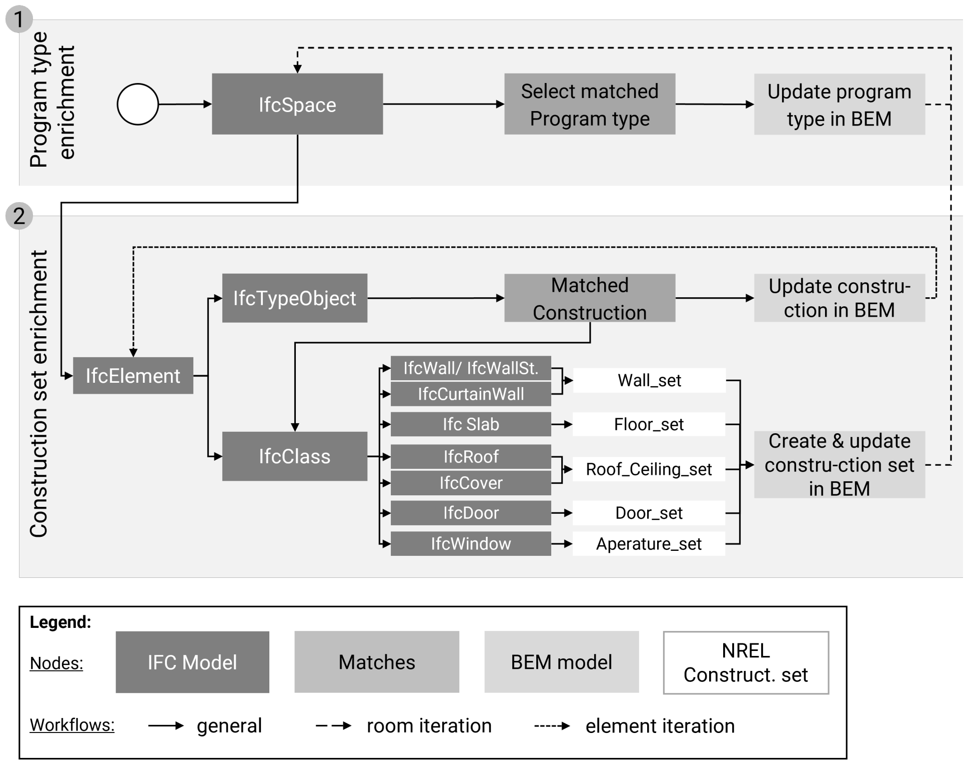



Figure 5.7: Workflow of semantically enriching BEM for space types and construction sets using previous matches

Figure 5.7 describes the workflow of enriching the Building Energy Models (BEM) using the previous matches. The BEM is iteratively enriched for each architectural room (IfcSpace) first (1.) and element afterward (2.). For every IfcSpace, the matched space type is selected, and all relevant information from the NREL database, as previously described in Figure 5.3, are enriched in the according BEM schema.

Next, each IfcElement, which is connected to each IfcSpace, is iterated, and the matched constructions of its IfcTypeObject are selected for enrichment. The construction is updated in the BEM. However, every room needs a complete construction set consisting of a wall set, floor set, roof/ ceiling set, door set, and aperture set. Therefore, we identify the IfcClass for each matched construction and assign it to the according set, as described in the previous chapter. In this step, we also distinguish between internal and external elements. However, we are excluding windows from the matching approach and only assign generic constructions due to the specific naming of its thermal constructions and materials.

By iterating through all elements of each IfcSpace, we create a step-by-step construction set by enriching the according construction. Missing constructions are replaced by generic constructions. Finally, each room gets a unique construction set assigned based on the previously matched constructions of the BIM model.

### 5.3.4 Fine-tuning LLM

In this section, we first describe the general workflow for multilingual fine-tuning pre-trained LLM for adapting domain knowledge. Next, we differentiate different strategies for adapting the fine-tuning process and improving the overall matching performance.

#### General LLM fine-tuning workflow

We propose three steps for multilingual LLM fine-tuning, consisting of (a) the dataset pre-processing, (b) the monolingual LLM fine-tuning, and (c) multilingual LLM fine-tuning. Figure 5.8 shows the general workflow for multilingual LLM fine-tuning. Generally, we used different epochs for fine-tuning: 1, 5, 10, 15, and 20. We tested even more epochs, but as the results overall decreased, we limited it to 20.

In the first step (a), we are preparing all relevant monolingual and multilingual data sets. The main input contains the German room description, element name, and material information of the BIM model, as well as the manually matched space types and construction, including their materials. This input is split into, first, the monolingual matching pairs, including the BIM information and German translation of the manual matches, and second, the multilingual translation pairs of the German and English NREL databases. Furthermore, we include further domain knowledge, such as BIM-specific abbreviations in the German language (Helmus et al., 2021), as well as multilingual material pairs by Madaster (EPEA GmbH, 2022).

In the next step (b), we are first fine-tuning the manual matches in German language only following Reimers' methodology of Sentence Embeddings using Siamese BERT-Networks (Reimers & Gurevych, 2019). We are using the German BERT model as the base model ('bert-base-german-cased') and, therefore, the BertTokenizer for the tokenizing step. As one of the fine-tuning strategies, we propose different loss functions for the mean pooling for converting token embeddings into sentence embeddings.

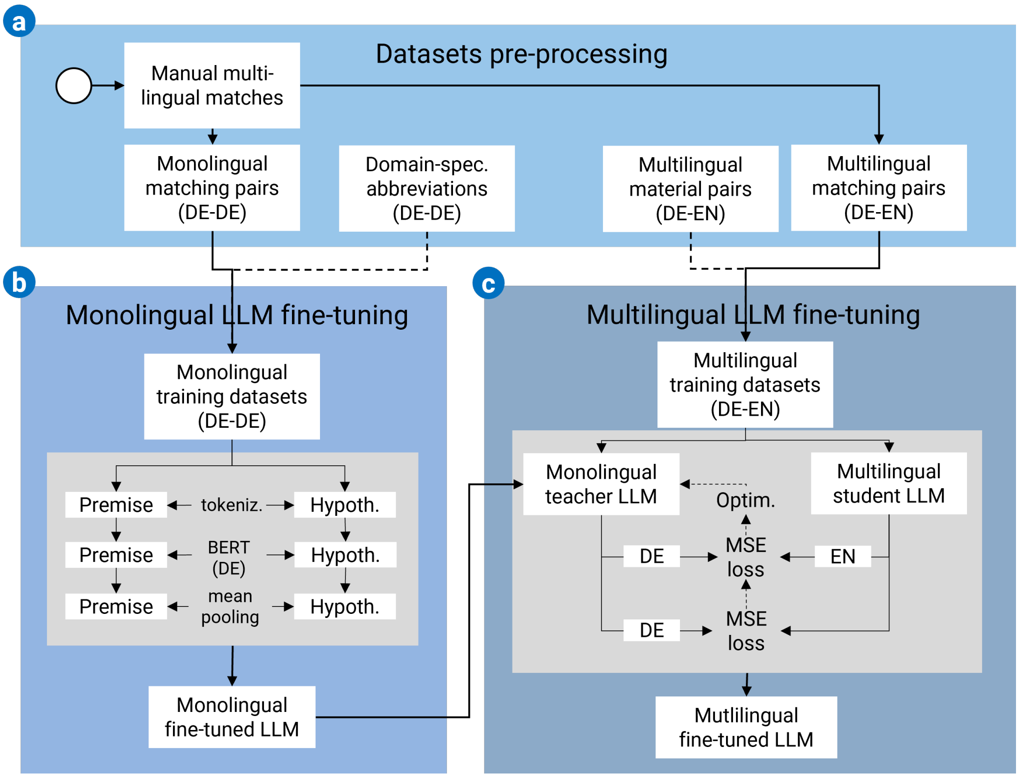



Figure 5.8: General workflow for multilingual LLM fine-tuning

In the final step (c), we proceed with the multilignual fine-tuning following Reimers' Knowledge Distillation for multilingual Sentence Embeddings (Reimers & Gurevych, 2020). The previously fine-tuned monolingual LLM serves as a monolingual student model. Furthermore, we are testing different multilingual student LLM for the fine-tuning process. This step uses the MSE loss function to optimize the matching performance.

**Strategies for fine-tuning domain-specific LLM and improving matching performance**

We propose four different strategies for the monolingual fine-tuning process and improving the overall matching performance:

1. Applying different loss functions for monolingual fine-tuning, such as
    a. Cosine Loss,
    b. Contrastive Loss or
    c. Multiple Negatives Rranking Loss (MNRL)
2. Adding negative/ contradicting word pairs (already included in MNRL)
3. Adding abbreviations, including
    a. general ones in the AEC industry and
    b. about construction and material nomenclature usually given in German BIM practice

4. Adding context, such as IfcClasses, isExternal, classifications using Softmax Loss Function and multiple labels

The first two strategies of testing different loss functions and additional contradicting word pairs are technically driven. The following two strategies, however, follow the aim of incorporating domain knowledge using abbreviations and BIM-related context information into the fine-tuning process of the LLM to increase the matching accuracy. As shown in a previous study (Forth, 2023b; Forth, Abualdenien, & Borrmann, 2023), domain-specific abbreviations were a big challenge for the STS-based matching approach. We can differentiate between abbreviations in the field of BIM, as well as general construction materials-related abbreviations. The last two strategies can't be applied to the room-specific fine-tuning, as the context is not given, and the BIM- or AEC-specific abbreviations are not applied in the room naming.

We propose two main different strategies for the multilingual fine-tuning process and improving the multilingual matching performance:

I. using different multilingual student LLM, such as
   A. LaBSE,
   B. XLM-RoBERTa and
   C. Distilluse-v2
II. adding more multilingual pairs of materials.

Similarly to the monolingual strategies, the first multilingual strategy (I) follows a pure technical motivation to select the best performing multilingual student LLM. The second strategy (II) tries to include further multilingual domain knowledge by using domain-specific material translations from other databases, but is not applied to the room-specific fine-tuning.

**Combinations of proposed strategies**

In Table 5.1 the overview of all possible combinations of different for fine-tuning LLM and improving matching performances are shown. The table is divided by mono- and multilingual fine-tuning steps, as well as by the two matching use cases of room-specific matching space types and element-specific matching of thermal constructions, and finally also the different strategies themselves.

Strategies 3 to 4 and II are not available for room-specific matching but element-specific matching only. Furthermore, all monolingual strategies can be combined with all multilingual strategies. Therefore, we select the best-performing monolingual fine-tuned LLM for further multilingual fine-tuning.

| Fine-tuning | monolingual | | | | | | | multilingual | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Matching use case | room & element | | | | element | | | room & element | | | elem. |
| Strategy | 1a | 1b | 1c | 2 | 3a | 3b | 4 | IA | IB | IC | II |
| 1a CosineL | o | | | x | (x) | (x) | | x | x | x | x |
| 1b ContrastiveL | | o | | x | (x) | (x) | | x | x | x | x |
| 1c MNRL | | | o | incl. | x | x | | x | x | x | x |
| 2 Negative pairs | x | x | incl. | o | x | x | | x | x | x | x |
| 3a AEC-Abbr. | (x) | (x) | x | x | o | x | | x | x | x | x |
| 3b BIM-Abbr. | (x) | (x) | x | x | x | o | | x | x | x | x |
| 4 Context labels | | | | | | | o | x | x | x | x |
| IA LaBSE | x | x | x | x | x | x | x | o | | | x |
| IB RoBERTa | x | x | x | x | x | x | x | | o | | x |
| IC Distil.-v1 | x | x | x | x | x | x | x | | | o | x |
| II Multil. mater. | x | x | x | x | x | x | x | x | x | x | o |

Table 5.1: Overview of combination of different mono- & multilingual strategies for fine-tuning domain-specific LLM and improving matching performance

### 5.3.5 Building Energy Performance Simulations

As the main focus is on the semantic enrichment of BEM, we define the annual energy demand simulation as the primary BEPS indicator to evaluate the impact of our matching method. The results of the annual energy demand simulation are measured by the energy use intensity (EUI) across the conditioned floor area in [kWh/sqm], which is, in our case, the same as the total model floor area. We use Ladybug Tools and Pollination (Pollination, 2024) as an interface to run EnergyPlus simulations on a cloud server (EnergyPlus Development Team, 2010). Besides the results as HTML files, providing detailed information about monthly EUI results, the overall EUI results are divided into heating, cooling, interior lighting, interior equipment, and pumps.

## 5.4 Evaluation and results

This section shows the evaluation and results of the proposed methodology. First, we introduce five real-world case studies. We use three for fine-tuning and two for testing and evaluation. Next, we are evaluating the results of the domain-specific fine-tuned LLM for the use case of matching rooms and elements using for each using different strategies. Furthermore, we are showing the results of the BEPS using the previous matching approach and the proposed enrichment workflow. Finally, we are discussing the limitations of the proposed approach.

### 5.4.1 Case studies and datasets

All case studies are new building designs, mainly for office usage, modeled by different architectural offices. As the case studies are new constructions, we can assume the latest standards for the space types and construction sets, in our case, from the year "2019". For the main usage or building type, we set "Large Office", which covers most of the room-specific usages with only a few exceptions, e.g., canteen or parking. First, we exported the model as an IFC4 design transfer view, including base quantities and second-level space boundaries, as shown in Figure 5.9. Next, we used the Revit-Plugin from Pollination (Pollination, 2024) to export a Building Energy Model from Revit as HBJSON data format only focusing on the geometric transformation steps. For the semantic enrichment step, the IFC model and its information about architectural rooms and elements, including their materials, are used.

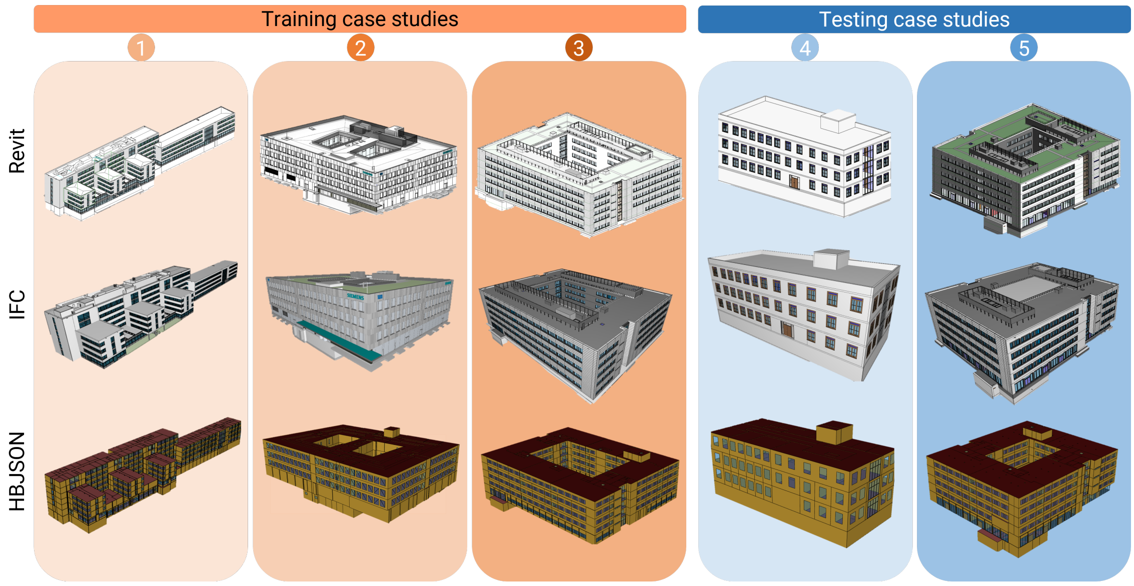



Figure 5.9: Overview of five case study buildings as Revit model (top), IFC exports (middle), and manually transformed energy model (down)

Case studies 1, 2, and 3 were used for training the LLM. Case studies 4 and 5, consisting of one small and one more complex building, were used to test the matching method, evaluate the matching accuracies and evaluate the results of the BEPS. Table 5.2 shows the overview of the extracted datasets from the training and testing IFC models for mono- and multilingual fine-tuning for the different matching use cases, such as room-specific matching of space types and element-specific matching of thermal constructions.

<table>
<tr><td colspan="3">Datasets</td><td colspan="2">training</td><td>testing</td></tr>
<tr><td>Mat. use case</td><td>Fine-tuning</td><td>Source</td><td>positive</td><td>negative</td><td></td></tr>
<tr><td rowspan="2">Room</td><td>monolingual</td><td>Case studies 2,3,4</td><td>205</td><td>3.619</td><td rowspan="2">86</td></tr>
<tr><td>multilingual</td><td>Case studies 2,3,4</td><td>129</td><td>-</td></tr>
<tr><td rowspan="6">Element</td><td rowspan="4">monolingual</td><td>Case studies 1,5</td><td>484</td><td>3.619</td><td rowspan="6">206</td></tr>
<tr><td>AEC-abbr.</td><td>571</td><td>284.114</td></tr>
<tr><td>BIM-abbr.</td><td>155</td><td>21.763</td></tr>
<tr><td>Context labels</td><td>2.065</td><td>3.619</td></tr>
<tr><td rowspan="2">multilingual</td><td>Case studies 1,5</td><td>156</td><td>-</td></tr>
<tr><td>Multil. materials</td><td>387</td><td>-</td></tr>
</table>

Table 5.2: Overview of used datasets of the case studies and domain-specific strategies for the matching use cases and divided by the mono- & multilingual strategies fine-tuning

As we used whole case studies for testing, the training datasets of the room matches are around 70%, while the testing datasets are ca. 30%. For the element matching, the ratio is slightly shifted to 66% training and 34% testing datasets. The multilingual pairs derived from the training case studies are the unique space types and the thermal constructions and materials. For the domain-specific abbreviations, we used 571 general AEC abbreviations and their descriptions (Bundesamt für Bauwesen und Raumordnung, 2021) and 155 BIM-specific abbreviations (Helmus et al., 2021), such as construction types or material abbreviations. For the multilingual material datasets, we used 387 material translations created by EPEA provided by Madaster (EPEA GmbH, 2022).

The negative pairs are derived from the positive ones and take all potential contradicting datasets into account. For the multilingual fine-tuning, no negative pairs are available for training.

The NREL databases are provided by Ladybugtools in a JSON format and, therefore, can be easily parsed without any preprocessing. Updates of these databases are also provided by Ladybugtools when installing the latest versions of Honeybee.

### 5.4.2 Evaluation of matching space types to architectural rooms

In this section, we evaluate the results of the use case of matching space types to architectural rooms. First, the monolingual fine-tuning results are evaluated, followed by the multilingual ones. Finally, we discuss the analysis results of detailed error clusters and the wrong matching using confusion matrices.

#### Evaluation of different monolingual strategies

Despite Komatsuzaki's recommendation of just using one epoch for training Large Language Models (Komatsuzaki, 2019), it usually applies to expensive training of large datasets. In our case, we are fine-tuning based on smaller datasets, so we use multiple epochs for fine-tuning. As we evaluate the matching accuracy in a second step for multiple combinations, the training with multiple epochs is still expensive. For this reason, we train in steps of five (1, 5, 10, etc.) and maximum twenty epochs, as several test with more than 20 epochs (25, 30, 50) showed lower matching accuracy results.

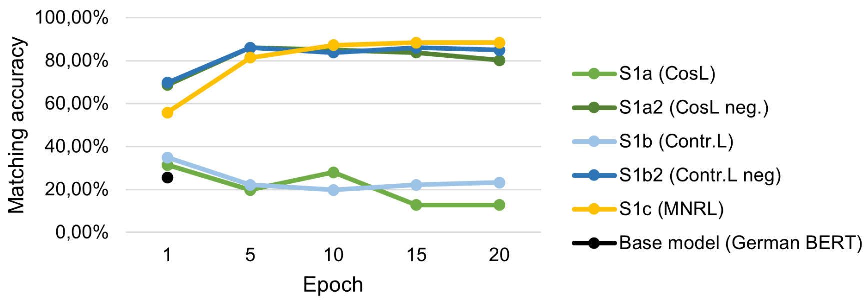



Figure 5.10: Results of the accuracy of monolingual matching architectural rooms to space types using different loss functions, negative pairs, and training epochs

Figure 5.10 shows the matching accuracy of strategies 1a, 1b, and 1c using different loss functions (CosineL, ContrastiveL, MNRL) in combination with adding negative pairs over training of several epochs. Cosine Loss and Contrastive Loss perform worse without negative pairs and lead to insufficient matching performance. Using negative pairs, the different loss functions, Cosine Loss, Contrastive Loss, and MNR Loss, perform similarly. Increasing the epochs more than one significantly improves the matching accuracy up to over 80%. The maximum is reached by MNRL (S1c) with training 15 epochs. Furthermore, MNRL is much cheaper in training compared to CosineL and ContrastiveL including all negative pairs.

**Evaluation of different multilingual strategies**

For the multilingual fine-tuning, we use the best performing monolingual, fine-tuned LLM from the previous Section 5.4.2 as the teacher mode, S1c trained on 15 epochs.

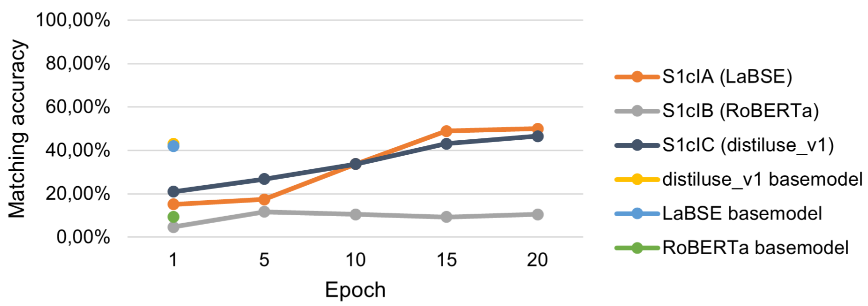



Figure 5.11: Results of the accuracy of multilingual matching architectural rooms to space types using different student models and training epochs

Figure 5.11 indicates that multilingual fine-tuning using one epoch leads to insufficient matching accuracies. Generally, XLM-RoBERTa (S1cIB) performs worse the other two student models. Furthermore, only by a higher number of epochs the multilingual fine-tuning, the matching accuracy reaches similar results compare to the other two base models, which are around 42%. The best performing multilingual, fine-tuned LLM follows the combination of S1clA with a matching accuracy of 50%, using LaBSE as the base model and being trained on 20 epochs. However, the matching accurarcy is significantly lower than the monolingual ones.

**Error analysis of fine-tuned LLM for room matching**

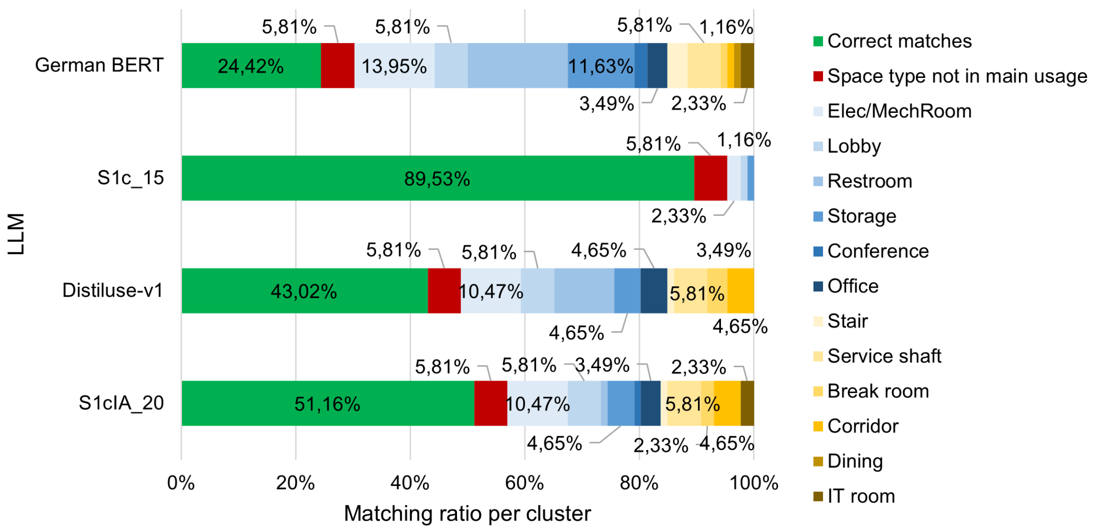



Figure 5.12: Error analysis of matching space types to rooms using different mono- and multilingual base-models and fine-tuning strategies

In Figure 5.12, we identify different error clusters for the best-performing mono- and multilingual LLM in comparison to the best-performing mono- and multilingual base models ("German BERT" and "Distiluse-v1"). In green, the correct matches for each LLM are shown and correspond with the matching accuracy of each LLM. The other colors represent wrongly matches space types matchings.

In red, for five rooms, the manual matched space types were from a different building type and not "LargeOffice", but the manual matches were "Retail", "Security Screening", or "Laundry". For this reason, no correct matching could be performed. To avoid this issue in the future, we can either consider the two main usage types or extend specific space types also for other main building usages.

For the other errors, we clustered them according to their manually matched space types. The errors in yellow represent those clusters that could be fully rectified by fine-tuning the monolingual LLM. Also, the number of blue-colored error clusters could be significantly reduced to a minimum of 5,81% in total. However, two electrical or mechanical rooms and one room each, lobby and storage, were wrongly matched.

The multilingual base model "Distiluse-v1" performs better than the monolingual base model "German BERT". However, the multilingual fine-tuning based on the best-performing monolingual LLM doesn't significantly solve any error cluster. On the contrary, the correctly matched IT rooms were wrongly matched in the fine-tuned LLM. This shows that the multilingual fine-tuning does not give reliable matching results, in contrast to the translated version of the best-performing monolingual one.

In Figure 5.13, we compare the detailed errors for each space type of the building usage "LargeOffice" using confusion matrices and comparing the monolingual base model on the left side and the best performing fine-tuned LLM on the right side. As described above, the multilingual LLM has lower accuracy than the translated monolingual. Therefore, we compare the base model (German BERT) with the LLM by following the strategies S1c trained with 15 epochs.

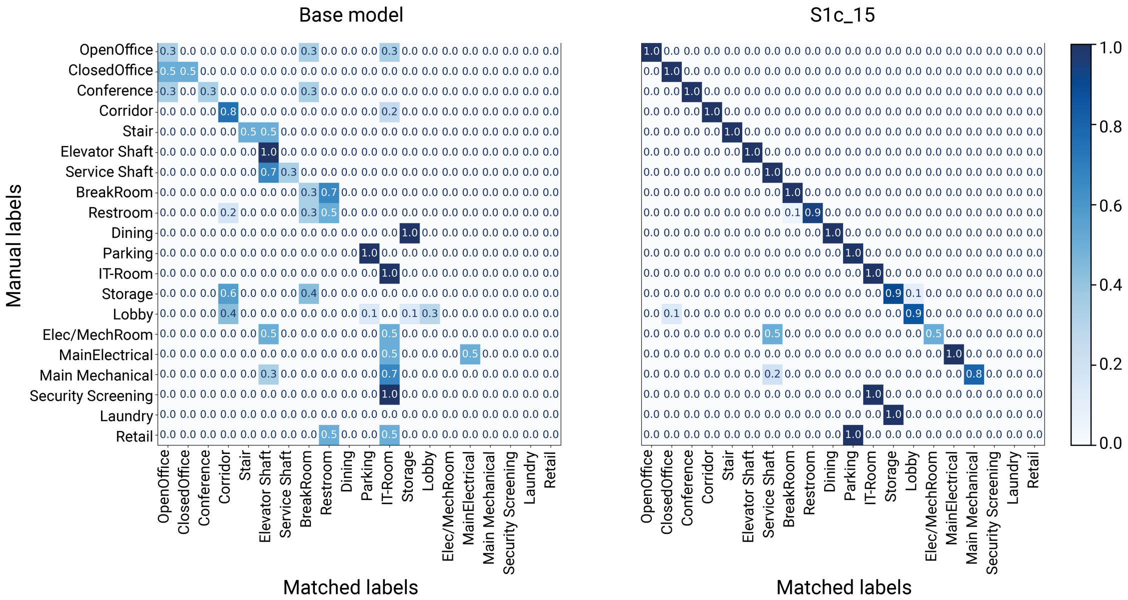



Figure 5.13: Normalized confusion matrices of base model and best performing LLM of matching space types to rooms

For the space types "Security Screening", "Laundry", and "Retail", which are not part of the building usage "LargeOffice", we can see the mismatches to "IT-Room", "Storage", and "Parking". Furthermore, significant confusion can be identified for "Elec/MechRoom" and "Main Mechanical" with "Service Shaft". Only minor confusion appears for "BreakRoom" instead of "Restroom", "Lobby" instead of "Storage", and "ClosedOffice" instead of "Lobby". Overall, the confusion matrix of S1c_15 reflects the promising results of the previous evaluations and error analysis.

### 5.4.3 Evaluation of matching thermal constructions to elements

In this section, we evaluate the results of the use case of matching thermal constructions to the elements in the BIM model based on the similarities on the element and material level. First, we discuss the monolingual fine-tuning results, followed by the multilingual ones, and finally concluding with the detailed error analysis and confusion matrices.

### Evaluation of different monolingual strategies

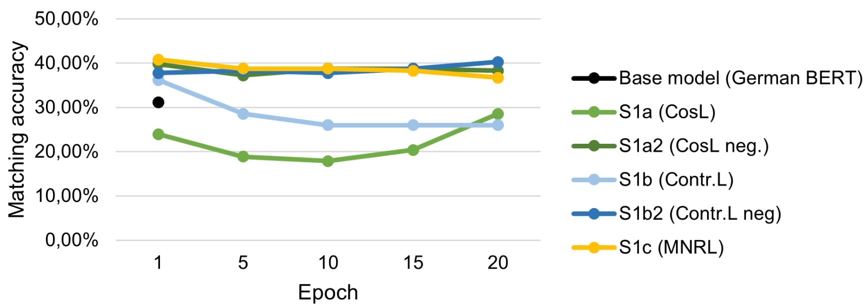



Figure 5.14: Results of the accuracy of monolingual matching BIM elements to thermal construction using different loss functions, negative pairs and training epochs

Figure 5.14 shows the matching accuracies of the monolingual fine-tuning strategies of using different loss functions for elements matching compared to the base model (German BERT) using different training epochs. Adding negative pairs significantly improves the matching accuracy for Cosine and Contrastive Loss. However, MNRLoss performs slightly better than the other loss functions and has the highest accuracy of 40,82% using one epoch for training. For this reason, we'll use MNRL for fine-tuning the other strategies in the next steps, including AEC-specific abbreviations (S3a) and typical BIM modeling abbreviations (S3b) and the combination of both (S3ab).

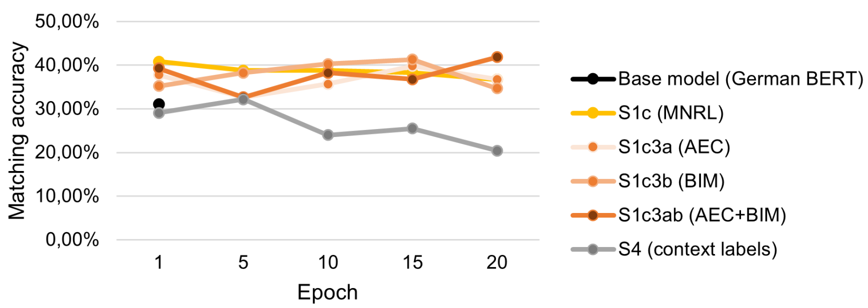



Figure 5.15: Results of the accuracy of monolingual matching BIM elements to thermal construction using MNRL,different domain-specific abbreviations and training epochs

Figure 5.15 depicts the results of matching accuracies using the above-mentioned domain-specific abbreviations (S3a-c) using MNRLoss (S1c) compared to the base model and the version of S1c without abbreviations. Adding BIM-specific abbreviations increases the matching accuracy up to 41,33% using 15 epochs for training. The overall maximum of 41,84% is using both abbreviation sources and training with 20 epochs (S1c3ab). Increasing the number of epochs higher than 20 does not improve the matching accuracy. However, compared to the use case of room matching, the accuracy are significant lower, which will be analyzed in more detail in Section 5.4.3.

#### Evaluation of different multilingual strategies

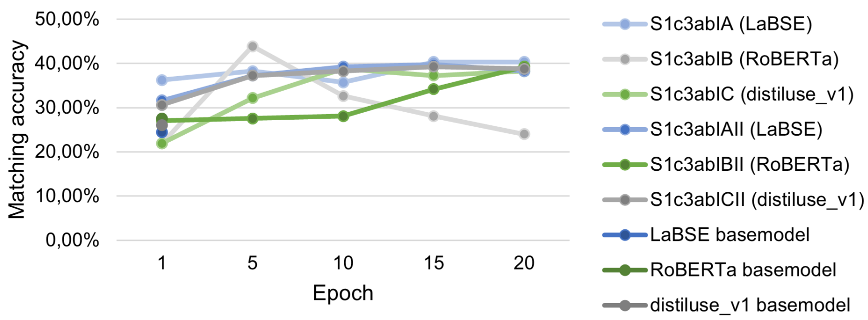



Figure 5.16: Results of the accuracy of multilingual matching elements to constructions with thermal properties using different student models and training epochs

Figure 5.16 shows the matching accuracies of multilingually matching thermal constructions to BIM elements using different fine-tuning strategies and epochs. We used S1c3ab as the best-performing monolingually fine-tuned LLM (S1c3ab) as the teacher model. In contrast to the multilingual fine-tuning of space type matchings, increasing the epoch improves the matching accuracy for all student models. The maximum accuracy of 43.88% is reached by the student model RoBERTa with five training epochs.

Strategy II, adding more multilingual pairs of materials, slightly improves the matching accuracy for the student models RoBERTa (S1c3ablB) and distiluse-v1 (S1c3ablC) but not for LaBSE (S1c3ablA). However, the fine-tuning without SII is significantly lower for five training epochs.

Overall, the multilingual LLM fine-tuning combination of S1c3ab_20_IB_5 has the highest matching accuracy with 43.88%. This one has slightly higher accuracy than the best-performing monolingual ones. However, we conduct a more detailed error analysis on the best-performing mono- and multilingual ones comparing it to the base model (German BERT).

#### Error analysis of fine-tuned LLM for element matching

In Figure 5.17, we clustered the errors of the base model and the three best-performing LLM, S1c_1, S1c3ab_20, and S1c3ab_20 combined with SIB_5 according to the element construction types (interior and exterior walls, doors, ceilings, floor slabs, roofs). We can identify the increased correct matches from 31,12% to 40.82%, 41,84%, and 43.88% for the different fine-tuned LLMs.

While the high errors of wrongly matched exterior walls were only reduced by around 6% each for S1c_1, floors, doors, and ceilings slightly decreased, and the roof was matched correctly. However, the error cluster of the interior wall significantly increased by over 15%. A similar trend can be observed for S1c3ab_20. Here, we have slightly better results for the cluster of interior walls and floors but slightly worse for exterior walls.

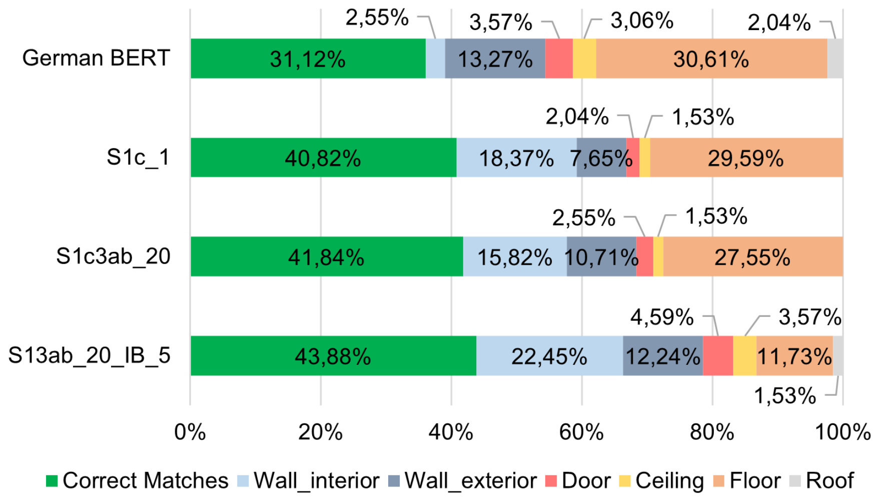



Figure 5.17: Error analysis of matching thermal constructions to elements using different mono- (and multilingual) base-models and fine-tuning strategies

For the multilingally fine-tuned LLM with the highest matching accuracy S1c3ab_20_IB_5, the worst matching error occurs for interior walls with 22.45% wrong matches and exterior walls with 12.24% of wrong matches. Nevertheless, we identify the best improvements for floors with only 11.73% wrong matches.

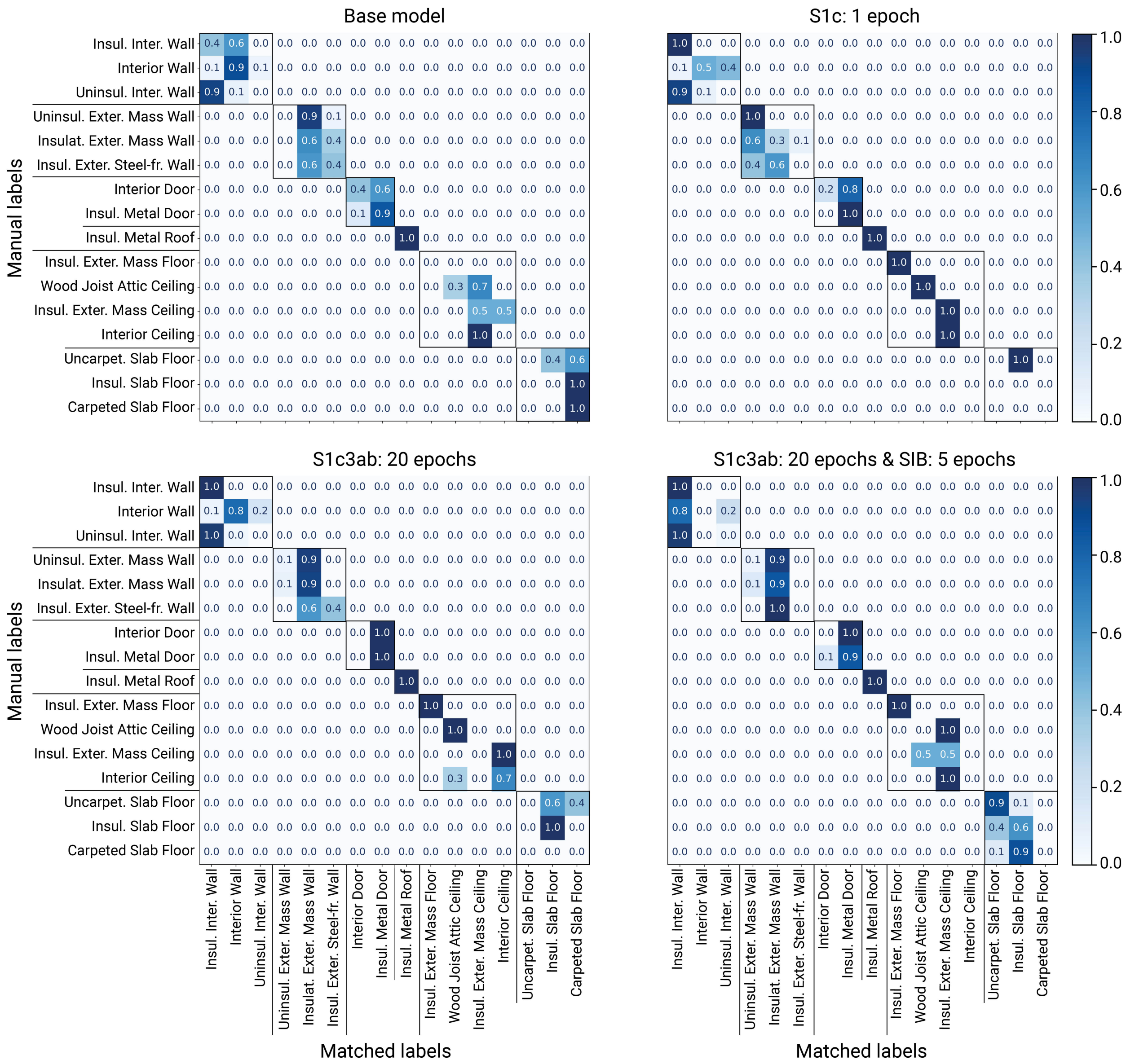



Figure 5.18: Normalized confusion matrices of the base model and three best-performing LLM of matching thermal constructions to elements

To have a more detailed analysis of each error cluster, we visualize the results in confusion matrices for the different LLMs, as shown in Figure 5.18. We grouped the confusion matrix according to the construction clusters to have a better overview of mismatches and analyze it for the three best-performing LLMs together. For interior walls in all three LLMs, we can identify an improvement for matching "Insulated Interior Wall" and only minor deterioration for "Interior Wall". However, uninsulated interior walls are mostly mismatched with insulated ones. For exterior walls, we identify different trends. While in the LLM using abbreviations, we have improvements for insulated mass walls; this is not the case for S1c_1. Nevertheless, in that LLM, we see improvements for uninsulated mass walls, which are mismatched in the other two LLMs. All three LLM show mismatches for insulated steel-framed walls; only S1c3ab_20 shows some correct matches.

For doors, most of them are matched with "Insulated Metal Door" and not with "Interior Door". As we only have one unique roof construction, all roof matches are correct. The biggest improvements in the matching accuracy are identified for ceilings in all three LLMs. While exterior mass and wood joist attic ceilings are correctly matched, insulated exterior mass ceilings and interior ceilings are often confused, depending on the LLM. For the slab floors, no LLM correctly matches carpeted ones, but mostly matches "Insulated Slab Floor". The multilingually fine-tuned LLM has the best matching performance for this cluster, as it almost always matches uncarpeted slab floors correctly.

Generally, the difference between "insulated" and "uninsulated" or "carpeted" and "uncarpeted" is often confused. This usually happens if other materials that exist in both elements are identified as semantically most similar. The insulation or carpeting material does not seem to change the matching results. Therefore, the proposed element-specific matching does not produce overall sufficient results evaluating unique constructions. The differences for the whole model and effects on the BEPS results will be evaluated in the following Section 5.4.4. One main limitation is that the NREL databases for construction with thermal properties do not differentiate between different insulation materials in their material or construction naming but rather between different thermal resistances of the constructions and materials. Therefore, generic materials are difficult to correctly identify.

| Fine-tuned LLM | Matching accuracy | F1-score (macro) |
|---|---|---|
| Base model (German BERT) | 31.12% | 27.16% |
| S1c: 1 epoch | 40.82% | 43.01% |
| S1c3ab: 20 epochs | 41.84% | 44.29% |
| S1c3ab: 20 epochs & SIB 5 epochs | 43.88% | 30.20% |

Table 5.3: Matching accuracies and F1-scores of best-performing fine-tuned LLM for thermal construction matching

Before BEM enriching, we compare the matching accuracy and F1-scores of the three best-performing LLMs in Table 5.3. As we have an imbalanced dataset, we use the macro average for the F1-score. We can identify that the multilingually fine-tuned LLM has the highest matching accuracy with 43.88%, but only a F1-Score of 30.20%. The other fine-tuned LLM S1c_1 and S1c3ab_20 have significantly better F1-scores. The most balanced fine-tuned LLM is S1c3ab trained with 20 epochs, which we will use for BEM enrichment and evaluation of the BEPS results.

### 5.4.4 Evaluation of resulting BEPS

In Figure 5.19, we show an overview of the total results of several BEPS in the shown values, based on the energy use intensity, differentiating its different end uses, such as Heating, Cooling, Electric Equipment, Interior Lighting, and Water Systems in the colors. We compare the different BEM enrichment processes, such as using generic profiles, manual enrichment, as well as automated BEM enrichment using the best-performing fine-tuned LLM for both use cases "rooms" and "elements". Generally, we differentiate between the room-specific enrichment steps, where we use the manual matches for the elements, and the element-specific enrichment of thermal construction, where we use the manual matches of the space types. Finally, we combine both use cases and perform a fully automated enrichment for both room- and element-specific enrichment. The manual enriched version is set as ground truth.

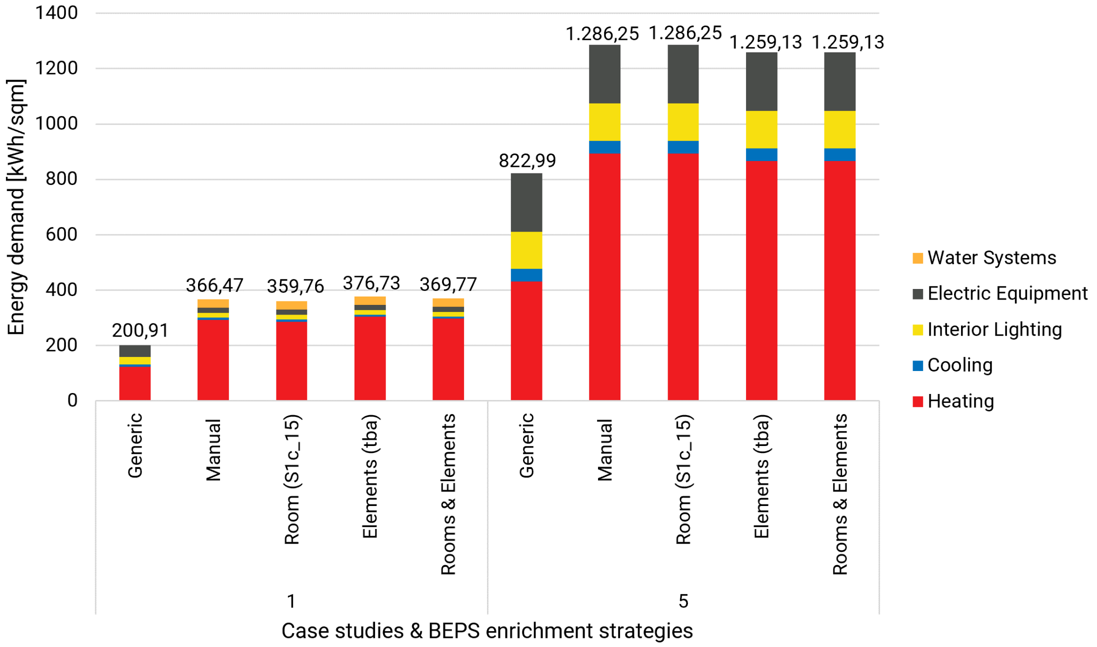



Figure 5.19: Overview of EUI results [kWh/sqm] of the two case studies using the different enrichment strategies of generic profiles, manual enrichment, and enrichment using the best performing LLM for room-specific, element-specific matching and both

Table 5.4 shows the deviations of the automatically enriched BEM models and the resulting simulated annual energy demand compared to the manually enriched ground truth. Furthermore, the table is clustered by case studies and matching use cases "rooms" and "elements" as defined in Section 5.3.3. The matching use case "generic" is when only using generic space types ("Generic Office Program") and construction sets ("Generic Construction Set"). Matching use case "both" includes the use cases "rooms" and "elements" at the same time.

| Deviation in | BEM model | | Annual energy demand (EUI) | | | |
|---|---|---|---|---|---|---|
| Matching use case | rooms | elements | generic | rooms | elements | both |
| Case study 4 | 6.25% | 28.75% | 45.18% | 1.83% | -2.80% | -0.90% |
| Case study 5 | 0.00% | 20.00% | 36.02% | 0.00% | 2.11% | 2.11% |

Table 5.4: Deviations in BEM models and simulated annual energy demand (EUI) compared to manually enriched BEM clustered by case study and use case or enrichment method

Generally, we can identify a significant difference between generic enrichment and the manual enriched BEM for both case studies, as shown in Figure 5.19 and Table 5.4. The computed annual energy demand of the generic enrichment for case study 4 differs by

45.18% and for case study 5 by 36.02% compared to the manually enriched ground truths. This underlines the necessity of more precise enrichment than only using generic profiles.

Following the high accuracy of the room-specific matching of space types, we only see a minor deviation between the LLM-based enrichment and the manual ground truth. In case study 4, we have a small deviation of 1,83% for the total EUI. Although the matching accuracy of unique room descriptions and elements is lower, we have a few rooms, like "ClosedOffice" or "Conference", which occur more often. In total, we only for 6.25% of all rooms differently matched space types for case study 4. The deviation in the EUI simulation results is even lower, with 1.83%. For case study 5, we have no deviation as the room-specific matching results are identical to the ground truth. Therefore, there is also no deviation in the simulation results of the annual energy demand.

For the element-specific matching of constructions with thermal properties, we have lower deviations to the ground truth for the whole BEM. The reason is similar to the room-specific matching, that some correct matches, such as insulated interior and exterior walls, occur in more rooms and have, therefore, a higher impact. In total, the elements differ by 28.75% for case study 4 and 20.00% for case study 5 compared to the manually enriched BEM. However, the EUI results for both case studies only slightly differ, with -2.80% for case study 4 and 2.11% for case study 5.

If we are considering both matching use cases together, the deviations of the simulated annual energy demand results are even lower. As for case study 5, only the matched elements differ from the ground truth, the total deviation is 2.11%. For case study 4, we have a difference of 0.90% of the EUI results compared to the manually matched BEM.

### 5.4.5 Limitations

The main limitation, but also the main challenge, is the availability of sufficient real-world case studies and matching datasets for training and testing. We focused on office buildings as a main usage of the buildings and need native BIM models in Revit to export BEM models.

The overall results showed that small deviations of the simulated EUI results for the room-specific matching of space types nor the element-specific matching of thermal constructions are achieved and sufficiently accurate for BEPS in early design stages. Therefore, this approach can assist with the matching process and can be integrated into a decision support tool.

Additionally, we limited the strategies for fine-tuning and improving the matching performance to testing different loss functions, including negative pairs, domain-specific abbreviations, adding context labels, different student models, and existing domain-specific material translations. We did not further develop loss functions or the basic architecture of pre-trained LLM. However, we tested different epochs for the training process.

Furthermore, we focused only on German models and abbreviations. The requirements and additional datasets might vary in different countries. Also, there might be differences in the performance of the base model of BERT in different languages due to the availability of its trained data.

In the use case of matching thermal constructions to BIM elements, we only took unique constructions into account. Those with similar names but different thermal resistances of the insulating materials were neglected and only the typical constructions were included. Furthermore, due to the specific material names of window elements and materials, we excluded these constructions from our matching approach, too, and only assigned a generic construction.

## 5.5 Conclusion and future research

In this publication, we aim to close the research gap of semi-automated enriching BEPS-related information robustly based on real-world Building Information Models in early design stages. Therefore, we proposed a methodology using Semantic Textual Similarities and different strategies for fine-tuning Large Language Models. We enrich space types based on the semantically most similar architectural rooms using the name attribute of IfcSpaces and MEP constructions including thermal properties based on IfcElements and IfcMaterials. We trained the datasets of manual matches of three case studies and tested and evaluated the matching results using two different case studies.

Our proposed methodology showed that Semantic Textual Similarity and fine-tuning of pre-trained LLM support the two multilingual matching use cases of space types to architectural rooms and thermal construction to BIM elements in early design stages to answer our first research question. However, we had to separate the fine-tuned LLM for both use cases of room-specific space type matching and element-specific matching of constructions with thermal properties. Next, we first evaluated the monolingually fine-tuned LLM and the multilingual ones afterward.

We also need to differentiate between both use cases for the second research question about which fine-tuning strategies are improving the matching accuracy, as not all strategies can be applied for room-specific matching of space types. Generally, we can see for both use cases that adding negative pairs significantly improves the accuracy, and MNRLoss leads to the highest accuracies with the least computing time. Adding domain-specific abbreviations also improved the matching of thermal constructions to BIM elements.

To answer the third research question of whether a fully automated or semi-automated matching is possible to ensure reliable BEPS results, we need to differentiate between the two use cases. For the room-specific enrichment, we reached high matching accuracy. The results of matching the constructions to elements show some confusion in case one thermal material appears in multiple constructions. Therefore, we need to develop the element-specific matching approach and the quality of the database of the construction and materials with thermal properties. However, in total, the BEPS results only slightly deviated using the fully automated enrichment compared to the manual enriched BEM based on the two testing case studies.

In our future research, we also aim to include the geometric transformation to surface BEM models based on IFC to incorporate this approach in a holistic open BIM workflow.

Additionally, we want to extend the scope to more case studies and datasets. This shall include fine-tuning on different languages, more country- and domain-specific abbreviations for BIM modeling, and different construction and material databases with thermal properties. These should have difference in the materials and not only in the thermal properties to identify more unique matches on material level.

## 5.6 Acknowledgments

The authors would like to gratefully acknowledge the support of Siemens Real Estate (SRE) and Forth Grünig Architekten for providing access to the analyzed case studies.

# Chapter 6

# BIM4EarlyLCA: An interactive visualization approach for early design support based on uncertain LCA results using open BIM

Previously published as: Forth, K.; Hollberg, A.; Borrmann, A.: *BIM4EarlyLCA: An interactive visualization approach for early design support based on uncertain LCA results using open BIM*. Developments in the Built Environment 16, 2023, DOI: 10.1016/j.dibe.2023.100263

## Abstract

To meet the European climate goals in the building sector, a holistic optimization of embodied greenhouse gas (GHG) emissions using the method of life cycle assessments (LCA) are necessary. The early design stages have high impact on the final performance of the buildings and are characterized by high uncertainty due to the lack of information and not yet taken decisions. Furthermore, most current BIM-based LCA approaches require high expertise and experience in both BIM and LCA and do not follow an intuitive visualization approach for other stakeholders and non-experts. This paper presents a novel design-decision-making approach for reducing embodied GHG emissions by interactive, model-based visualizations of uncertain LCA results. The proposed workflow is based on open BIM data formats, such as IFC and BCF, and is developed for decision support for non-LCA experts in the early design stages. With the help of a user study, the prototypical implementation is tested by 103 participants with different levels of experience in BIM and LCA based on a case study. We evaluate the proposed approach regarding the support of open BIM data formats, different LCA visualization strategies, and the intuitiveness of different approaches to visualizing uncertain LCA results. The user study results show a broad acceptance and need for open BIM data formats and model-based LCA visualization but less for visualizing uncertainties, which needs further research. In conclusion, this interactive, model-based visualization approach using color coding supports non-LCA experts in the design decision-making process in early design stages.

## 6.1 Introduction

The AEC industry, which contributes to 40% of the world's greenhouse gas emissions, needs to make significant changes to achieve the global climate goals (United Nations

Environment Programme, 2022). Recent studies have shown the increasing importance of embodied environmental impacts (Röck et al., 2020).

Life-cycle assessments (LCA) of the whole building are being used as an established method to evaluate these emissions during the design phase of buildings taking the operational and embodied emissions of buildings into account. Different environmental impact indicators, such as Global Warming Potential (GWP), are assessed, estimating the emitted Greenhouse gas (GHG) emissions. This LCA method ensures current national and international regulatory frameworks, for example, LEVEL(s), used to verify EU Taxonomy classification and report ESG conformity (European Commission & Directorate-General for Environment, 2021).

Schumacher et al. pointed out that Building Information Modeling (BIM) has significant potential for a loss-free data exchange, as well as for understandable and user-friendly communication of LCA results (Schumacher et al., 2022). Recent BIM-based approaches partially automate the LCA calculation process and reduce the assessment effort using different strategies (Wastiels & Decuypere, 2019). For automatic semantic enrichment, Sackes et al. highlighted a "combined, optimal use of topological rule inferencing and machine learning" as a foundational research challenge (Sacks et al., 2020). Fonseca et al. identified BIM-based "data retrieval and representation based on the needs of nonexperts" (Fonseca Arenas & Shafique, 2023) as a current research gap in the field of BIM-LCA integration.

However, most of the current approaches in the field of BIM-based LCA are either using closed BIM workflows or require a high level of LCA expertise to conduct and interpret the calculated LCA results. Building owners, clients, or project developers, who usually make overall decisions, often do not have the required expertise in LCA and are increasingly using open BIM models. Today's decision-making of construction and material choices in industry practice hardly considers environmental impacts. Furthermore, there is high uncertainty in early design phases, and current BIM-based LCA approaches do not communicate these to decision-makers. This gap of visualization of environmental impacts in BIM models, including uncertainties of early design stages for non-experts, has not been filled yet (Tam et al., 2022).

The main aim of this publication is to close this gap by proposing a conceptual workflow for interactively visualizing LCA results for design-decision support in early design stages using open BIM models. Different interactive visual strategies, such as model-based color-coding or box-plot diagrams, should help non-LCA experts to intuitively understand the environmental impact of different design variants and select the preferred option.

In summary, we aim to answer the following three research questions:

a) How can open BIM data formats support the design decision-making process for environmental impacts?

b) Which LCA visualization strategies support non-LCA-experts in the decision-making of elements and material variants in early design stages?

c) How can uncertainties of LCA results in early design stages be intuitively visualized?

To answer them, we first propose a workflow for visualizing LCA results for design decision support based on open BIM Standards, such as Industry Foundation Classes (IFC) and BIM Collaboration Format (BCF). Second, we test different LCA visualization strategies for decision-making and uncertainty visualization by a prototypical implementation. Finally, we test them with a user study to evaluate how they perform for different participants, differentiating, for example, by their LCA experience.

This publication is structured as follows: Section 6.2 provides an overview of the state-of-the-art of BIM-based LCA for decision-making, feedback communication using open BIM data formats, visualization of LCA results, and uncertainties. Section 6.3 presents the general research method and an approach for an interactive visualization and design decision support of LCA using open BIM. The proposed workflow is then explained in Section 6.4.1 and evaluated using a prototypical implementation described in Section 6.4.2 and a user study using a real-world project as a case study provided in Section 6.5. Finally, Section 6.6 provides the overall findings and recommendations for future research.

## 6.2 Background and Related Works

This Section describes multiple fundamental topics about BIM-based LCA calculation, model-based feedback communication, visualization strategies of LCA results, as well as visualization of uncertainties of LCA results, providing the necessary background for the following Sections.

### 6.2.1 BIM-based LCA for decision-making in early design stages

The field of LCA using BIM models has been increasing over the last few years. Thereby, it is necessary to use open BIM data formats to enable loss-free interoperability between different software tools (Borrmann et al., 2021). Industry Foundation Classes (IFC) is an open BIM data format for semantic-rich geometric building models developed and maintained by buildingSMART (buildingSMART Technical, 2023b).

Rezaei et al. proposed a BIM-based workflow for LCA calculation using closed BIM and Revit for early and detailed building design stages (Rezaei et al., 2019). They used a Monte Carlo simulation to allocate the uncertainty of materials in the early design stages. Schneider-Marin et al. focus in their approach on uncertainty analysis of LCA using BIM in early design stages (Schneider-Marin & Lang, 2020). In order to reduce the vagueness and increase the result precision, they use sensitivity analysis as guidance for design teams. However, they did not include material uncertainties in the early design stages. Kamari et al. introduce a BIM-based LCA tool for early design stages (Kamari et al., 2022). Their study showed that critical hotspots can be identified at a low level of detail at an early design stage. However, they did not implement an element-based LCA where the material with the highest contribution can be identified.

Palumbo et al. propose in their study the use of Environmental Product Declarations (EPD) in early design stages for LCA based on BIM models (Palumbo et al., 2020). In their limitation, they state a lack of harmonized and homogenous formats of EPD schemes and only focus on specific material groups, mainly of the main structure, but excluding the building envelope. Llatas et al. extends their proposed approach to life cycle sustainability analysis (LCSA) to also integrate social life-cycle assessment (sLCA) and use IFC4 schema in early design stages (Llatas et al., 2022). Nevertheless, they used Autodesk Dynamo to calculate and visualize. In the last step, the LCSA results and enrich the IFC properties and attributes using IfcPropertySet. Soust-Verdaguer et al. propose a similar approach of LCSA introducing and validating an "element method" from early to late design stages (Soust-Verdaguer et al., 2022). Although their approach uses element-specific property sets for GWP, costs, and labor effort, their process is performed manually but can be automated with an Application Programming Interface (API).

### 6.2.2 Feedback communication

As most approaches are based on closed BIM workflows, not all project stakeholders, such as clients or project developers, are involved in the decision-making process. Conversely, those methods, which are based on open BIM workflow, face the challenge of communicating the decision back to the BIM modeler and into the authoring tool.

One established communication method using open BIM workflows includes the BIM collaboration Format (BCF) (buildingSMART Technical, 2023a). Generally, BCFs help in a BIM-based collaboration project by communicating and solving issues, such as clashes, and work similarly to a ticketing service. BCF is an XML-based file format zipped with other relevant data, such as images. It consists of an issue name with a short text, a viewpoint including a screenshot of the BIM model, a GUID of the selected elements, descriptions, a history of the issue, the recipient of the message (group, person, or craft), a status of information, as well as annotations. The topic details can be directly linked to the BIM model by storing particular viewpoints and the unique identifiers of the related elements (Borrmann et al., 2021). At the time of writing, more than 70 software products implement the XML-based BCF exchange, while almost 30 software products use additionally the server-based BCF API (buildingSMART Technical, 2022).

Horn et al. propose in their method the integration of IFCXML for a bi-directional BIM-LCA integration (Horn et al., 2020). To this end, they enrich the BIM model with raw LCA results, structured by LCA phases and materials, and are linked to the reference data set. Their approach requires a complex setup, which is not applicable in broad yet.

Zahedi & Petzold introduce a minimized communication protocol specifically for the early design stages (Zahedi & Petzold, 2019). Meng et al. implemented a web-based communication platform for discussing early design stages variants(Meng et al., 2020). Different functions from a defined workflow were implemented using different data formats, such as JSON, IFC, or CSV.

### 6.2.3 Visualization of LCA results

Wiberg et al. document the progression of a visual, dynamic, and integrated approach to building LCA in their publication (Wiberg, Løvhaug, et al., 2019). They identify various methods of integration utilizing Visual Programming Languages, such as Dynamo and Revit or Rhino and Grasshopper, to address dynamic aspects. Additionally, they categorize other parametric approaches and dashboard implementations that employ Revit models or district models, typically displaying the models without utilizing them to highlight or visualize results. In their subsequent proposal, they put forward a visualization method employing Virtual Reality to enhance stakeholder engagement (Wiberg, Wiik, et al., 2019). In this approach, Revit models are employed to apply color coding based on LCA results, and VR is utilized to interact with the model. This is deemed a "good platform for communicating and visualizing complex data [...] not only for researchers but also for the general public" (Wiberg, Wiik, et al., 2019).

Utilizing BIM models to visualize LCA results has demonstrated significant potential (Mousa et al., 2016; Naneva, 2022; Röck et al., 2018a, 2018b; Tsikos & Negendahl, 2017). These approaches primarily employ color coding in authoring tools to represent the final LCA results visually. Kiss and Szalay apply a different visual technique for a detailed analysis of LCA results, utilizing model-based color coding in conjunction with a sunburst diagram to emphasize specific aspects of the results (Kiss & Szalay, 2019). For their implementation, Kiss and Szalay utilize Rhino and Grasshopper.

Miyamoto et al. (2022) present a method that suggests incorporating LCA and LCC findings to serve as a foundation for making design decisions (Miyamoto et al., 2022). Despite not utilizing BIM models, they discuss the increasing significance of integrating a spreadsheet approach with BIM workflows, albeit solely focusing on architects.

Hollberg et al. emphasize the importance of considering target users in developing their user-centric LCA tool, specifically for early planning stages (Hollberg et al., 2022). The process involved various stakeholders such as architects, sustainability engineers, consultants, and real-estate developers. However, the visualization of results is limited to fixed outcomes, and there was no provision for active interaction with the model. Nevertheless, we partially use this method for tool development using a case study and a user test, iteratively improving it with stakeholders' feedback.

In their recent review regarding the visualization of LCA results, Hollberg et al. provide an assessment of current practices and present a comprehensive overview of various strategies and potentials (Hollberg et al., 2021). The overview clusters different visualization strategies for LCA results according to its LCA goals and amount of information. We use this overview for the selection and development of different visualization strategies.

### 6.2.4 Visualization of uncertainties

The consideration of uncertainties in BIM models across varying levels of development has been overlooked for a long time. To address these aspects, Abualdenien and Borrmann (2020) propose multiple methods for visualizing geometric and semantic uncertainties of building elements during early design phases. Among the various approaches, they find that combining color value and transparency to quantify the reliability of semantics resulted in a relatively high level of intuitiveness and acceptance (Abualdenien & Borrmann, 2020).

Marsh et al. reviewed uncertainties of LCA for the built environment and the different sources of uncertainties (Marsh et al., 2023). Besides uncertainties due to the Goal & Scope, the Life Cycle Inventory, and the Life Cycle Impact Assessment, they list the data quality assessment, human error, and practitioner knowledge/ experience, as well as the comparability of carbon data sources and tools, data availability, unknown material specification at early stages, and time requirement for assessments as barriers.

In addition, Ströbele introduces a fuzzy life cycle assessment (fLCA) approach that accounts for vagueness through distribution curves instead of singular outcomes (Ströbele, 2022). Schneider-Marin et al. establish the EarlyData knowledge database for making material choices during the design stage when detailed information about specific materials is unavailable (Schneider-Marin et al., 2022). This method visualizes semantic uncertainty by assessing a wide range of potential material combinations simultaneously using box plot diagrams to represent the Global Warming Potential (GWP) ranges.

Petrova et al. propose a decision-support framework for sustainable design based on knowledge discovery from diverse building data. They employ various matching mechanisms between project data repositories and Common Data Environments (CDE), including data mining, direct semantic queries, and geometric feature matching (Petrova et al., 2019). The direct semantic queries rely on different ontologies, such as the Building Topology Ontology (BOT) or product-specific ontologies.

In summary, the discussed publications highlight the significance of using BIM models to visualize LCA results and present initial approaches. However, the investigation of integration within an open BIM workflow and the presentation of interactive result exploration are lacking. This reveals a gap in terms of an interactive design decision tool for non-LCA experts based on the open BIM method during early design stages. This publication's primary focus is to visualize uncertainties in rough model semantics and ambiguous results comprehensively.

## 6.3 Method

The approach consists of the following key features:

- Design decision support concept based on IFC models and embodied emission performance

- Feedback communication using BCF for LCA
- Visualization of uncertain LCA results using different strategies

Afterward, we briefly explain the steps for prototypical implementation before introducing the user study and its case study, set up, and the participants.

### 6.3.1 Research method and workflow

This paper aims to develop an approach for an interactive visualization approach for design decision support of embodied emissions using open BIM in early design stages. Therefore, we are proposing a workflow and evaluating it through a prototypical implementation and a user study. The scope of the embodied emissions focuses on Global Warming Potential (GWP) as the main environmental impact category.

We are following the research method of design science research (DSR) according to Pfeffers et al. (Peffers et al., 2012). Doing so, the developed approach represents the artifact supposed to answer the formulated research questions. We are prototypically implementing this workflow using a case study to evaluate it. The prototype hereby aims to demonstrate the utility and suitability of the artifact while the case study is applied to a real-world situation. Finally, we are setting up an experiment and user study for evaluating experts versus non-experts regarding the BIM and LCA experience.

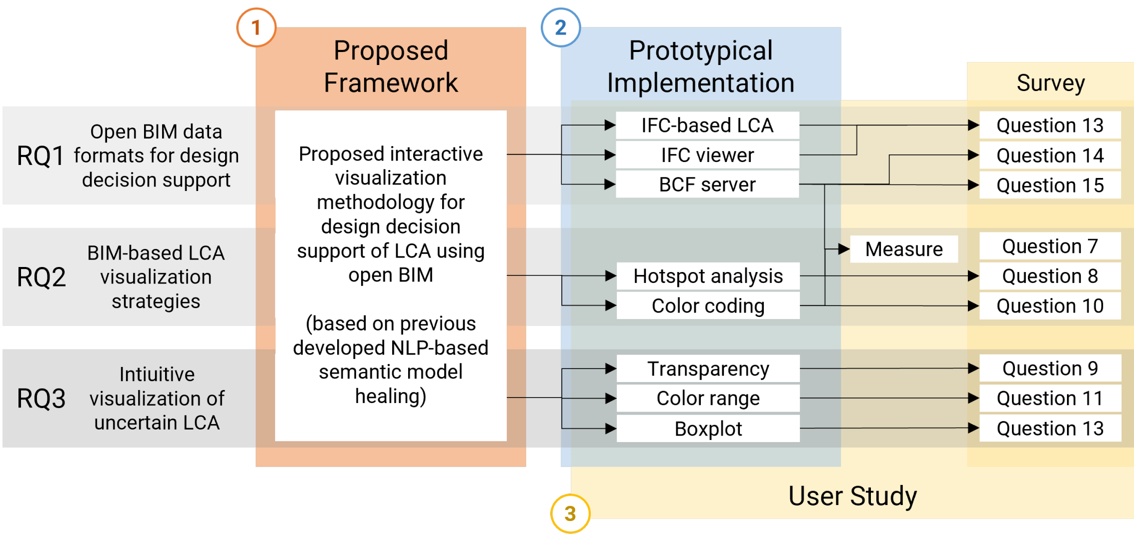



Figure 6.1: Research method and general workflow for answering the three main research questions (RQ1-RQ3) by (1) proposing a workflow, (2) prototypical implementation and (3) User study

Figure 6.1 depicts how the research method is applied to answer these questions by conceiving a workflow (see Section 6.4.1), providing a prototypical implementation, and performing a user study. The latter is performed to evaluate the prototypical implementation using a case study involving 103 participants. It will be introduced in detail in Section 6.4.2 and 6.3.4.

The semantic healing process introduced in (Forth, Abualdenien, & Borrmann, 2023) for the IFC-based LCA calculation process is implemented to answer the first research question of open BIM data formats. Furthermore, a model viewer for IFC models is implemented and tested, and the BCF server follows open BIM standards of buildingSMART International. Finally, the questionnaire of the user study proves the importance of the open BIM workflows and BCF as a standardized communication format.

For the second research question regarding different visualization strategies, we are implementing three strategies: a model-based color-coding for hotspot analysis, the color-coding of element and material-specific variants, and box plot diagrams of different design variants. Besides measuring the performance of different participants according to LCA reduction and taken time, we are additionally evaluating their feedback on these different visualization strategies using questionnaires in the user study.

The third research question is about three different visualization strategies of uncertainties, for which their intuitiveness is evaluated using a questionnaire. The three approaches include using transparency in the model viewer according to the findings of (Abualdenien & Borrmann, 2020), gradient color ranges for the different variants and box plot diagrams.

### 6.3.2 General workflow for design decision support of LCA using open BIM

The overall structure of the general workflow, illustrated in Figure 6.2, comprises four main steps and relies on the LCA knowledge database (LKdb). The LKdb includes the most typical elements based on domain knowledge and comprehensive information required for a holistic LCA, including layer-specific replacement rates, LCI datasets based on Ökobaudat, and any necessary End-of-Life scenarios. Further details regarding the LKdb can be found in (Forth, Abualdenien, & Borrmann, 2023). This paper focuses on embodied greenhouse gas (GHG) emissions, exclusively considering Global Warming Potential (GWP) as an LCA impact category. The operational part B6 is excluded from the LCA calculation. In terms of the LCA system boundaries, it encompasses the life cycle phases of production (A1-A3), replacement (B4), as well as End-of-Life (C3, C4), and benefits and loads beyond the system boundary (D). More details of the calculation process of the LCA result ranges were previously described in (Forth, Abualdenien, & Borrmann, 2023).

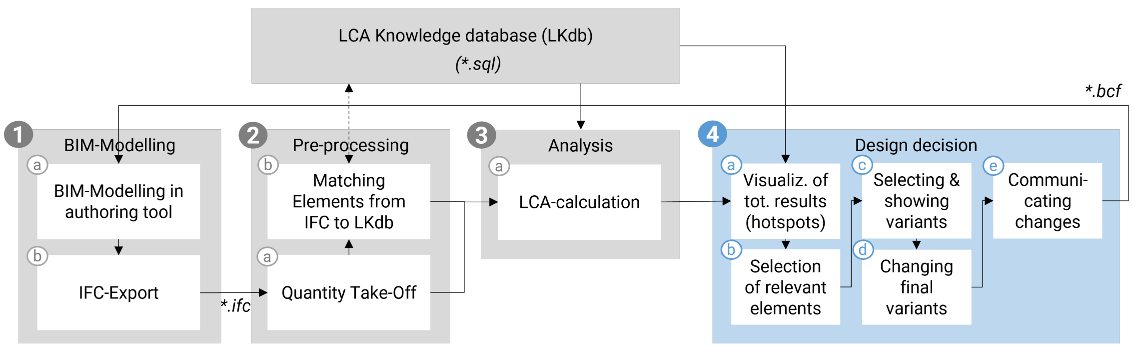



Figure 6.2: General workflow for visualizing uncertain embodied GHG emissions for design decision support in early design phases using open BIM

In the initial stage of the proposed workflow, the BIM model is created using any capable authoring software (step 1.a), followed by the export of the IFC model (step 1.b). The subsequent step involves extracting the quantity take-off and conducting element matching. The quantity take-off entails parsing all geometric and semantic information from the IFC model for LCA calculations. This includes fundamental quantities such as area, amount, layer thicknesses, or length, as well as density, materials, element names, GUIDs, and classifications (step 2.a). The expressions of materials and elements are utilized in the following step to match the IFC elements with the LCA knowledge database (LKdb) (step 2.b), which has been previously introduced and validated (Forth, Abualdenien, & Borrmann, 2022, 2023). This matching process relies on Natural Language Processing (NLP) employing a Large Language Model (LLM) to determine cosine similarities between the expressions of elements and materials in the IFC model with those in the LKdb. The most similar LKdb element is assigned to each IFC element. Previously, we identified Google's LLM BERT (Devlin et al., 2018) as the most suitable for this task (Forth, Abualdenien, & Borrmann, 2023).

Upon completing the element matching step, any missing information regarding LCA datasets, life spans, or absent layers is populated with the datasets of the matched LKdb element. Subsequently, the LCA results are computed, accounting for the uncertainty associated with the element matching (step 3.a). Depending on the level of matching (refer to Section 3.3), a range of material options for each layer of an element is considered, leading to a range of LCA results for both individual components and the entire building.

This publication focuses on the final step of the general process, specifically the design decision process (step 4.a-e). All steps are briefly described to provide an overview. The design decision-making process can be invoked after all LCA information is calculated and assembled. In the first step, 4.a, the results are visualized in the BIM model for hotspot analysis. The median values of the element-specific LCA result range are used to color-code the element in the IFC model in relation to its potential design variants. Next, those elements are selected, which still show optimization potential and can be easily detected using color coding (step 4.b). When one element variant is selected, all potential element variants based on the same classification group and IFC type from the LKdb are shown and highlighted (step 4.c). After all design choices have been made, the final variant has been changed (step 4.d), and the changes are communicated back to the authoring tool (step 4.e). To this end, an extended schema of BCF issues is automatically created and uploaded to the BCF server, as described in more detail in Section 6.4.2.

### 6.3.3 Prototypical implementation

As the first part (steps 1-3) were already previously implemented and validated (Forth, Abualdenien, & Borrmann, 2023), we are focusing on implementing the proposed decision-making approach. After defining the different steps in the design decision support concept and different visualization strategies, we implement the proposed workflow based on HTML, JavaScript, and CSS and host it on a web server. We run the previous LCA results for the case study, store all relevant information as a JSON file, and upload it with the IFC model into the web tool. The JSON file contains the following information from the previously calculated steps 1-3 from the general approach in Section 6.3.2:

- IFC model exported either as IFC2x3 or IFC4 from authoring software in step 1.b
- Quantity takeoff including element-specific information on its object type name, IFC type (e.g., IfcWall, IfcWindow), classification group, total surface area in square meters, number of all elements of the same object type and its IDs, layer-specific materials and their thicknesses from step 1.a
- Results of element matching including the level of matching (see cases in Section 6.4.1), most similar matched element, and if existing material options
- LCA results including the total GWP, the results for each layer, and the quantiles of its result distribution in [kg CO2-eq.]
- Potential element variants based on the same classification group and IFC type from LKdb including for each element variant the element name, its layer-specific, total GWP results, and the quantiles of its result distribution in [kg CO2-eq.]

To integrate the design decisions into the current BCF version, there are two options, either as BIM snippets, which are usually partial IFC files or by extending the BCF schema. We consider the second option, as for now, we only use this communication to send and store all created topics of the user study's design changes on a BCF server. After the first implementation round, we iterate and improve the tool with the first test candidates. Next, we host the prototypical design decision tool on a website, integrating it with an introductory video and the survey of the user study.

### 6.3.4 User study

The user study evaluates the prototypical implementation by setting up an experiment for testing the prototype by participants who fill out a survey. In the following section, we first briefly introduce the chosen case study, explain the overall setup of the user study, and lastly, the participants and survey.

#### Case study project

We validate the proposed workflow and prototypical implementation by applying it to a case study. To this end, an IFC model of an office building measuring 1950 m$^2$ is used. The matching results and LCA outcomes have undergone previous validation (Forth, Abualdenien, & Borrmann, 2023). Given that the project is situated in Germany, the classification adheres to the German cost groups as per the DIN 276 standard (DIN 276, 2018). The Ökobaudat database, which contains materials and elements named in

German, is utilized for this purpose (BBSR, 2021). The NLP network BERT is employed for element matching, as previously assessed in (Forth, Abualdenien, & Borrmann, 2022). The case study model encompasses 307 individual elements originating from 16 distinct object types. The cumulative surface area of all elements amounts to approximately 5824 $m^2$.

The LCA Knowledge database (LKdb) was introduced in detail in (Forth, Abualdenien, & Borrmann, 2022, 2023). When setting up the datasets for this case study, we considered 137 of the most conventional construction elements across all classification groups. These elements mainly consist of different element layers, which add up to 223 different element layers. In total, there are 127 different material categories, which add up to 343 different classification-specific material categories. The material options are directly connected to Ökobaudat (BBSR, 2021), which we manually enriched to 1000 different classification-specific material options according to its potentially related element layers.

### Set up of user study

To run the user study using the prototypical implementation, we set up a website server which hosts the user study itself and a BCF server storing the BCF issues and viewpoints. The user study itself is divided into three parts:

a) Introduction: following an explanation video (ca. 5 minutes)

b) Experiment: testing the prototype with the help of a case study by changing at least three different elements and/ or material choices (ca. 3-5 minutes)

c) Survey: filling out the final questionnaire (ca. 4 minutes)

The user study aims to investigate an interactive decision-making process for element and material variants with regard to embodied emissions using the open BIM method in early design phases. Different stakeholders from the building and planning sector with and without experience with BIM and/or LCA will be surveyed.

The questionnaire considers the current Human-Computer-Interaction (HCI) standards following the guidelines of Lazar et al. (Lazar et al., 2017). The overview of all survey questions can be found in the B.3.

### Participants

In total, 103 participants took part in the user study and the survey. Most of the participants (81%) are from Germany, while the rest of the participants work in other European countries, such as Switzerland (4), Denmark (4), Austria (3), Spain (2), and Czech Republic (2). One participant each originated from Belgium, Italy, Netherlands, Poland, Sweden, and Turkey.Most of the participants are working in the field of academia and research (27%), as planners (26%), such as architects, structural engineers, HVAC engineers, or similar, and as sustainability experts, buildings physicists, or energy consultants (25%). The rest is working as project developers and clients (15%), while only a few have a professional background in BIM Management (5%) and IT or software (2%).

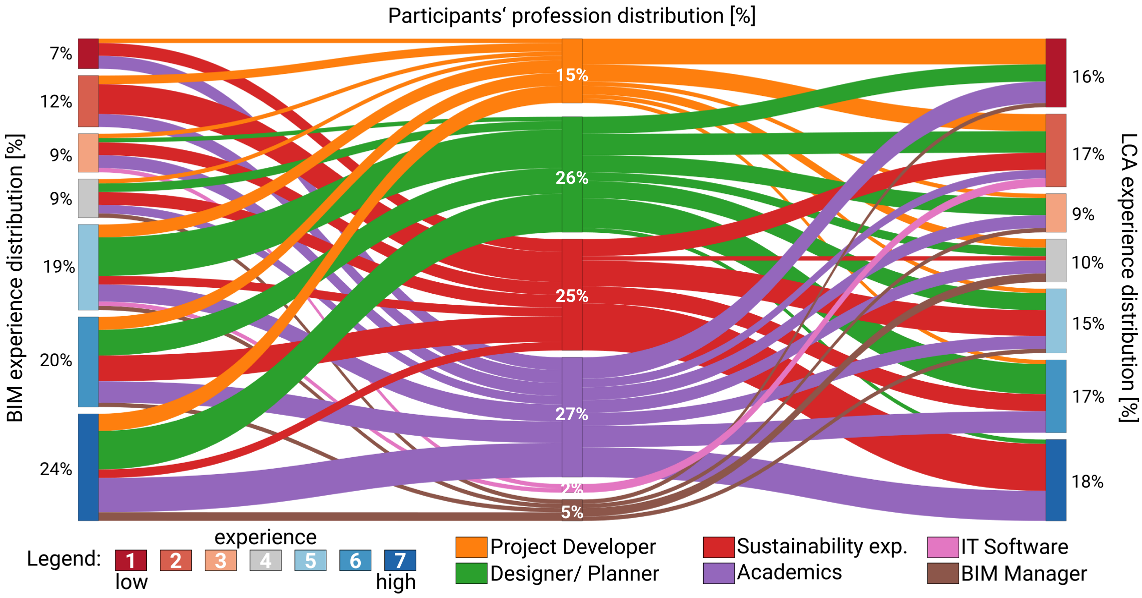



Figure 6.3: Profession of the user study participants in relation to their BIM experience (left) and LCA experience (right)

In Figure 6.3, the distribution of the user study participants and their profession is shown in correlation with their BIM experience and LCA experience. While the BIM experience is distributed almost equally amongst all professions, the LCA experience of sustainability experts is significantly higher than in other professions. Furthermore, project developers and clients seem to have the least LCA experience.

Figure B.5 shows a high correlation between those participants who have good experience with LCA to BIM-LCA experience (in total 58%) and vice-versa of those who only have little experience having no prior BIM-LCA-experience (in total 42%). This means most LCA-experienced participants already used BIM, while there is no correlation between BIM experience and BIM-based LCA experience.

## 6.4 Proposed workflow and implementation

In this Section, we first briefly introduce the proposed workflow, focusing on the design decision support concept and the selection of different visualization strategies of LCA and uncertainties. Afterward, we describe the prototypical implementation of the model viewer for hotspot analysis, the variant selection and visualization, and the feedback communication using an extended BCF schema.

### 6.4.1 Proposed workflow

The motivation of the proposed workflow is to assist stakeholders without expertise in LCA and/ or BIM in making decisions related to construction-element and material-related variants in the early design stages. First, we describe the more detailed developed design decision support concept in Section 6.4.1. Based on this concept, we discuss and introduce different visualization strategies for LCA results in Section 6.4.1. Finally, we further develop these strategies incorporating uncertainty visualization in Section 6.4.1.

#### Design decision support concept

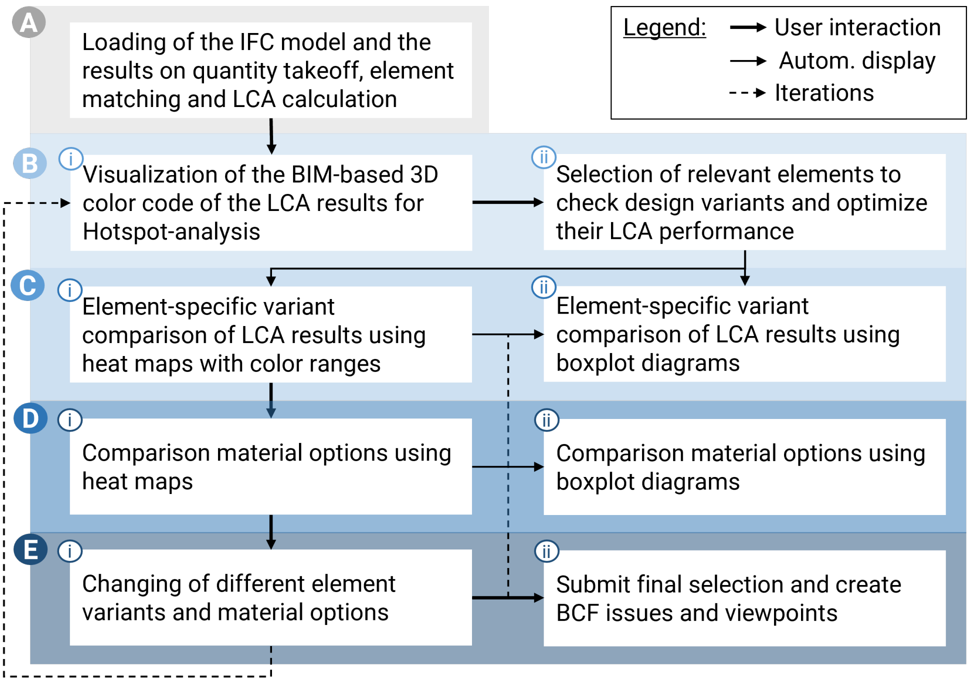



Figure 6.4: Concept of design decision making based on LCA results using BIM models and different visualization strategies

The design decision process, incorporating previous calculations and matching outcomes alongside their uncertainties, is generally divided into five steps, as depicted in Figure 6.4.

First (A), the IFC models must be loaded into the interactive decision-making platform as input. Based on this model and the previously proposed LCA Knowledge database (LKdb), the LCA results are calculated based on the IFC-based quantity takeoff and the NLP-based element matching according to (Forth, Abualdenien, & Borrmann, 2023). All this information is preprocessed and precalculated.

Next, the LCA results are presented using a model viewer and the BIM-based 3D color coding, which will be described in more detail in Section 6.4.1. According to the project LCA results, the worst-performing elements can easily be highlighted using the color-coded model as a hotspot analysis (B.i). The color of each model's element is calculated according to the LCA results of the NLP-matched element variant of the previous process and normalized according to all other element variants of the same classification group and IFC type. In the future, once benchmarks become available at the level of classification groups, e.g., according to DIN 276, the colors can be normalized based on these benchmarks. In the next step (B.ii), the user selects those elements in the model interactively to check design variants on element and material levels and thereby optimize the overall and element-specific LCA performance.

Once one element is selected, the element-specific input information and all relevant element variants of the same classification group and IFC type are shown according to the LKdb. In step, C.i, the element variants and their LCA results are depicted using heat maps with gradient color ranges. According to the compared variants in the heatmap, those selected element variants are also shown as box plot diagrams in C.ii.

If one element variant suits the user, the design decision process continues on a material level (D). All layer-specific material options of this selected element are compared using heat maps (D.i). As those elements can consist of multiple layers, all possible combinations of layer-specific material options are displayed. According to the previous design-decision level, the selected material options of the heat maps are also plotted on top of the selected element's box plot diagram, showing the specific LCA results of each material combination (D.ii).

In the next step, E.i, when the element variant and material option are decided and changed, the user can start over by selecting the next element in the model viewer, iterating the process C.i-D.ii until satisfaction. Finally, after all, design decisions are finally submitted (E.ii), these changes are communicated back to the authoring tool and BIM modeler using the BIM Collaboration Format (BCF). Therefore, BCF issues and viewpoints are created for each design change according to the workflow's step 4.e.

### Visualization strategies for LCA results

To support decision-makers during the early stages of design, a set of hierarchical visualization objectives, based on the work of (Hollberg et al., 2021), is proposed. To identify hotspot elements that require design optimization, the recommended visualization type is the 3D color code, which can be implemented using open BIM models (see Figure 6.4).

The previously described goals include identifying hotspots, comparing design options, and visualizing uncertainties. The overview of our selection of visualization strategies for LCA results is shown in the upper part of Figure 6.5. On the bottom part, the adaptation of these visualization strategies incorporating uncertainties is shown, which will be described in more detail in Section 6.4.1.

To communicate the LCA results intuitively for non-LCA-experts, we decided to use 3D BIM models, including color coding for visualization, which is also suitable for identifying hotspots. As the next visualization strategy, we propose heat maps to compare different design options and identify element variant hotspots simultaneously. As the third LCA visualization strategy, we choose the conventional approach of bar charts. However, when incorporating uncertainty information, we change them to box plot diagrams, which

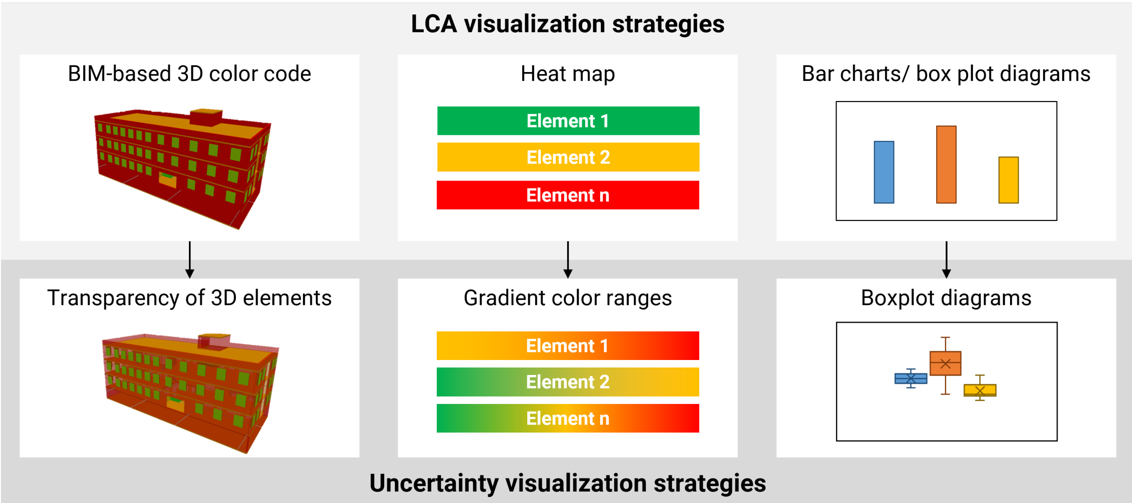



Figure 6.5: Visualization strategies of visualizing LCA hotspots as well as uncertainty

were already used in the previous study regarding this workflow (Forth, Abualdenien, & Borrmann, 2023). Therefore, we are proposing to test the following three LCA hotspot and uncertainty visualization strategies:

I Model-based 3D color code

II Heat maps

III Box plot diagrams, as they are also incorporating uncertainty information

### Visualization strategies of semantic uncertainties

Marsh et al. mentioned several sources of uncertainty in building LCAs (Marsh et al., 2023). However, our proposed workflow considers the time requirements for LCA and human error and practitioner knowledge by calculating the results automatically based on a LCA Knowledge database. Therefore, our approach mainly focuses on the unknown material specifications in the early stages as a source of uncertainty.

The previous study (Forth, Abualdenien, & Borrmann, 2023) has identified significant aspects considered in the analysis. First, the cosine similarity of the matching performance indicates the degree of resemblance upon which the matching is founded. Second, various scenarios exist regarding how the elements are matched from the IFC model to the LKdb, as outlined in the following enumeration.

IFC elements are matched to:

a) default element of the classification group in the LKdb (worst case)

b) most similar element expression, as there are no materials available

c) most similar element expression, as the material matching performs worse

d) element with the most similar material category

e) element with the most similar material option

The varying scenarios give rise to distinct levels of reliability in the obtained LCA results. Consequently, accounting for this information is imperative when visualizing semantic-related uncertainties.

As mentioned, Abualdenien et al. have previously concluded that combining color value and transparency in an element provides a highly intuitive and accepted means of visualizing semantic reliability (Abualdenien & Borrmann, 2020). Therefore, the transparency value for each element $t_e$ is determined using the following equation:

$$t_e = \frac{\sum_{l=1}^{m} c_e * cos(\theta)_{e,l}}{m} \tag{6.1}$$

where $l$ = layer number; $m$ = maximum layer number; $c_e$ = matching case of each element; and $cos(\theta)_{e,l}$ = cosine similarity of each element's layer (according to (Forth, Abualdenien, & Borrmann, 2023)).The values of the above-mentioned five matching cases for each element $c_e$ are distributed as follows: case 1 = 20%; case 2 = 40%; case 3 = 60%; case 4 = 80%; and case 5 = 100%.

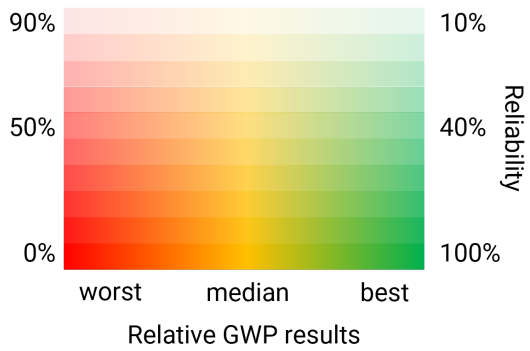



Figure 6.6: Color coding scheme for visualizing relative & uncertain GWP results

To incorporate both, the information regarding semantic-related uncertainties and the relative performance of GWP results, in the hot spot analysis, a color scheme matrix is introduced considering the gradient color range and transparency in Figure 6.6. The $x$-axis of the matrix represents the relative GWP results obtained through normalization within each classification group. The legend associated with the relative colors spans from green, representing the best-performing variant, to red, indicating the worst-performing variant. This gradient color range has been widely established and used in other research projects (Kiss & Szalay, 2019; Mousa et al., 2016; Naneva, 2022; Röck et al., 2018a; Tsikos & Negendahl, 2017).

On the $y$-axis, the transparency value corresponding to each element $t_e$ is visualized, as determined by Equation 6.1. Regarding the selection boxes of element variants and material options, the same gradient color range is employed, with 0%

As described in Section 6.2.4, the form of box plot diagrams is the most used form of visualizing uncertain LCA results. Furthermore, in a previous study, we also proposed to visualize the comparing the total and classification group-specific results to box plot benchmarks (Forth, 2023b). These benchmarks are based on a study of 50 buildings by the German Sustainable Building Council (DGNB) and are used to compare the correlating box plot GWP results (Braune et al., 2021). Nevertheless, in this paper, we left this feature out, as the calculation procedure and data of these benchmarks are not transparently available, and there is still a lack of representative benchmarks, especially in Germany, as described in more detail in (Forth, Höper, et al., 2022).

### 6.4.2 Prototypical implementation

In this Subsection, we describe the prototypical implementation, first focusing on the model viewer for hotspot analysis, followed by the element variants and material options, the variant selection and visualization part, and concluding with describing the feedback communication using BCF. Figure 6.7 shows a screenshot of the overall frontend interface of the prototypical implementation.

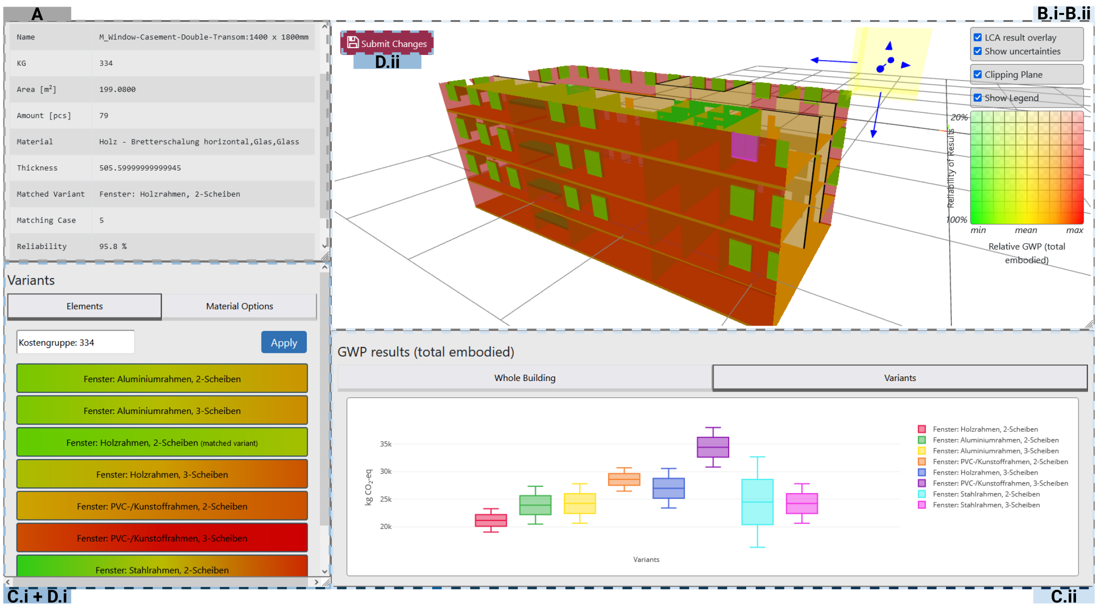



Figure 6.7: Prototypical implementation of the frontend interface according to the proposed workflow from Section 6.3.2

#### Model viewer for hotspot analysis

The prototypical implementation is based on established web development tools using HTML, JavaScript, and CSS. The web-ifc-viewer library of IFC.JS (González Viegas, 2022) is used for implementing the model viewing feature, which is a state-of-the-art open-source toolkit based on JavaScript library three.js for 3D scenes in web browsers (mrdoob, 2022). For the hotspot analysis, every element surface is colored according to its performance relative to the classification group and the mentioned color scheme of Section 6.4.1.

Depending on which variant is selected, the coloring is applied interactively and iteratively updated based on Node.js.

On the top left side of Figure 6.7, all relevant quantities and semantical information of the selected element are shown according to, such as classification group (KG), element name, amount of elements of this object type, material name, layer thicknesses, the matched element variant, and the matching case. On the right side, the 3D color-coded hot spot analysis is applied on the uploaded IFC model using the color scheme and transparency values for showing the matching-related uncertainties for step B.i.In the following step, B.ii, one highlighted element with insufficient performance, is selected to check design variants and optimize its GWP performance.

### Element variant and material option visualization

In the next step, C.i, different element variants, and material options are visualized according to the proposed visualization strategies II and III. On the bottom left side of Figure 6.7, the name of each element variant is colored according to the range of its normalized results. The normalization considers the maximum and minimum GWP results for each classification group and its LKdb-based element variants. Visualization strategy III using box plot diagrams is used on the bottom right. The results are shown on the right if the user selects multiple element variants on the left side. In case the selected element variant is sufficient, the user needs to manually apply the selection, which automatically updates the colors in the model viewer, creates a screenshot and viewpoints, and uploads all relevant BCF issues to the BCF server.

The material option tab can be selected if one selected element is detailed further. All relevant layers and material options for this element variant are color-coded according to its normalized GWP performance of the classification group, as previously described for the element-specific gradient color ranges. As every material option is connected to one pre-calculated LCA result, it is visualized as a differently colored dot mapped on top of the element-specific box plot diagram of the selected element variant as shown in Figure B.1 in B.1.1.

### Feedback communication using BCF API

As described in Section 6.2.2, BIM collaboration format (BCF) is a data format for communicating and solving issues in an open BIM workflow. In Section 6.3.3, we discussed not using BIM snippets but extending the current BCF format, consisting of a BCF topic and its related viewpoint, including a camera perspective and a snapshot.

As shown in Figure 6.8, the BCF extension "lcaSelection" consists of the selected object type as Identifier, the selected element variant, the selected material option, the time passed between selecting the object type in the model viewer and the finally applied

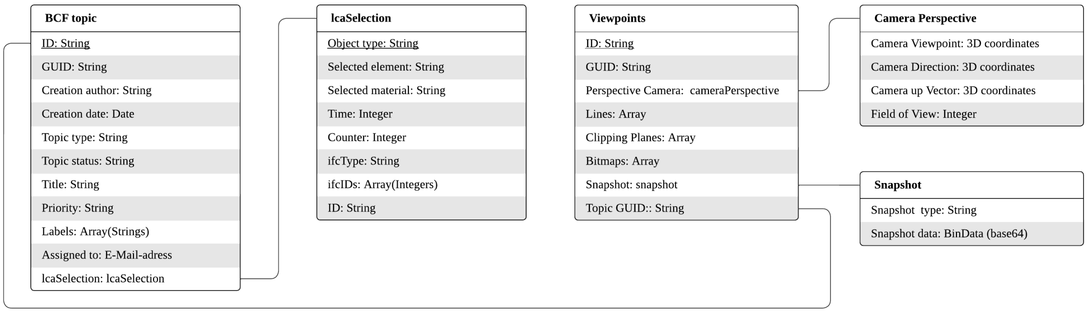



Figure 6.8: BCF topic and viewpoint schema and extension for LCA-related element and material selection

selection, a counter storing in which order the issue has been created, the IFC Type, all IFC IDs and finally the overall ID.

For implementing the feedback communication of the selected element and material variants, we use the BCF API by buildingSMART International (GitHub, 2023). For the server hosting, we use MongoDB (MongoDB, 2023).

After selecting the chosen element variant or material option, the button "Apply" temporarily holds all relevant information, according to the extended BCF schema 6.8. In the same step, the viewpoint of the model viewer is created, including the camera perspective and a snapshot. The snapshot is stored as binary code, so it can be transferred back to images, as shown in B.4. Finally, by clicking the button "Submit changes", all previously applied topics are pushed to the BCF server and are stored according to the extended schema.

## 6.5 Evaluation of the user study results

In this Subsection, we first evaluate the overall approach by quantitatively analyzing the general feedback of the participant and the measured data of their decision-making process. Next, we analyze their feedback on the topics of the three research questions in more detail about the open BIM data formats, the LCA visualization strategies, and the uncertainty visualization by analyzing the outcomes of the user study and qualitatively evaluating it according to the research questions. Finally, the limitations of the approach, implementation, and user study are discussed.

### Quantitative evaluation of the overall approach

First, we analyze the general participants' feedback on the question of how difficult (score 1) or easy (score 7) the participants rate the whole design-decision process. Afterward, we analyze the measured results on the decision-making process focusing on the GWP optimization results and the timing of their decisions.

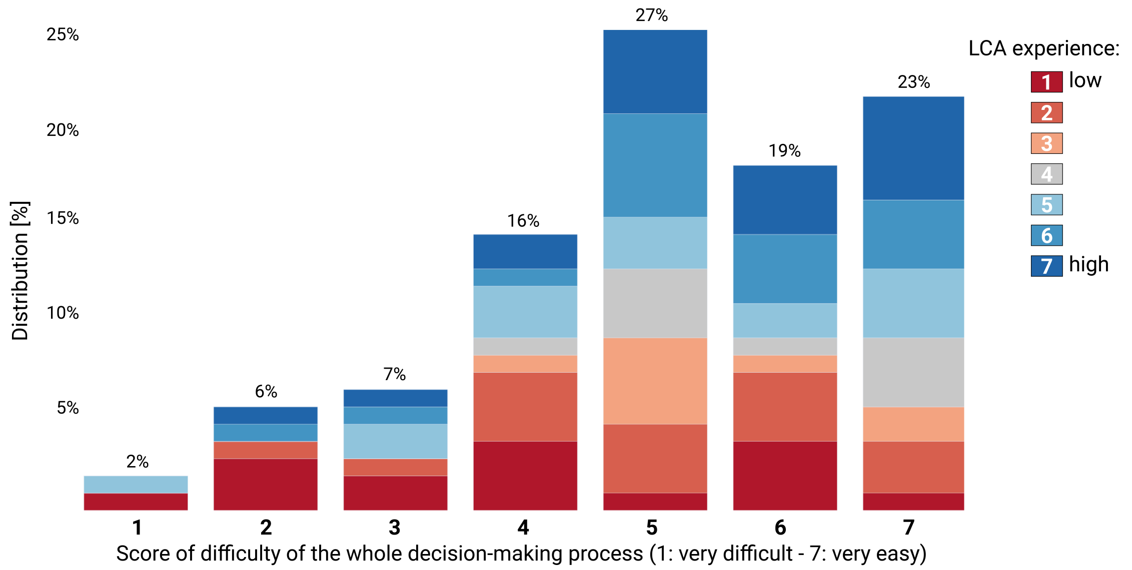



Figure 6.9: Overall feedback on the difficulty of the whole decision-making process in relation to the participants' LCA experience

The overall average score of 5.12 out of a maximum of 7 score points, which means the majority find the overall procedure relatively easy. Figure 6.9 shows the results in relation to their LCA experience, represented by the color of their LCA-experience score (1-7). Generally, a majority of 69% found the proposed workflow and the prototype of the whole design decision process rather easy, with almost a quarter scoring it as "very easy" (23%). Furthermore, no significant correlation is determinable between the participants' overall feedback and their LCA experience. Almost equally, LCA experts and non-LCA experts were rating all scores about the overall process feedback.

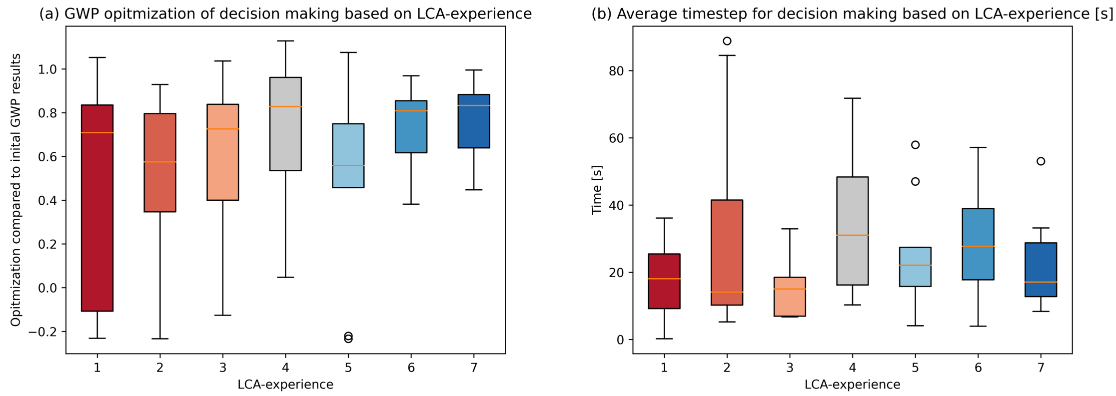



Figure 6.10: Measurement of resulting GWP optimization (a) and timestep (b) of participants' decision-making in dependency with LCA experience

Figure 6.10 shows the measurement results of the participants' decision-making process in two subfigures to analyze the difference between experts and non-experts. Generally, Subfigure 6.10 (a) only indicates a minor difference in relation to the participants' LCA experience, comparing the relative optimization of the final GWP results compared to

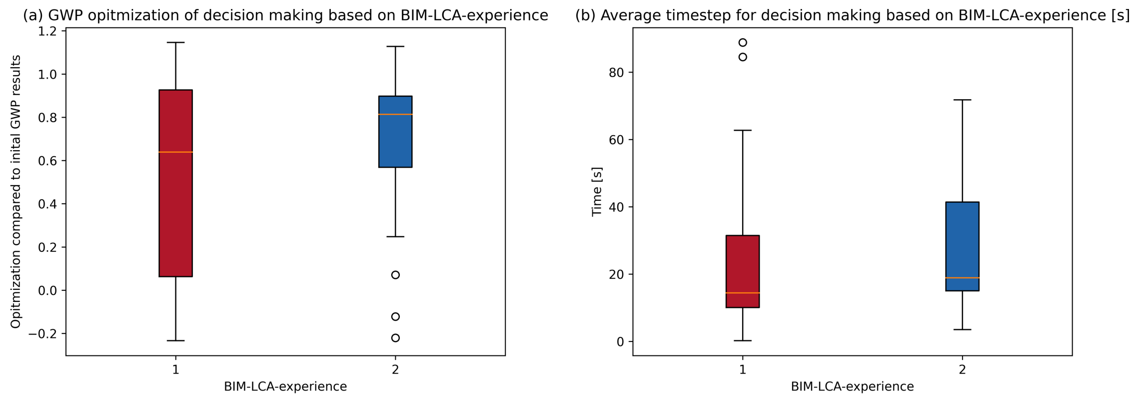



Figure 6.11: Measurement of resulting GWP optimization (a) and timestep (b) of participants' decision-making in dependency with BIM-LCA experience

the initial one. On average, those with the lowest LCA experience have slightly worse GWP optimization performances (ca. 70%) compared to those with high LCA expertise (ca. 80%). Nevertheless, the lower quartile of the box plot diagram of the lowest LCA experienced participants (around -10%), and the whiskers show less variance for the performance of the participant with high LCA experience.

Subfigure 6.10 (b) shows the average timesteps of each participant's decision-making in relation to their LCA experience, which is, on average, between 15 to 30 seconds. The box plot diagram shows no significant difference in the average time steps across all LCA experiences. This indicates the intuitiveness and overall acceptance of all participants independent of their previous expertise in LCA.

A more significant difference between experts and non-experts can be identified when considering the BIM-LCA experience, as shown in Subfigure 6.11 (a). While participants with no previous BIM-LCA knowledge have a lower median (ca. 65%) and lower quartile (ca. 10%) compared to the performance of the participants with previous BIM-LCA experience, which a median higher than 80% and the lower quartile around 60%.

When focusing on Subfigure 6.11 (b), no significant difference between those participants with prior BIM-LCA experience and those without can be detected. The average of both groups is around 18 seconds, while the upper quartile of the experienced participant is slightly higher.

Putting the described results in context with the correlation between LCA experience and BIM-LCA experience indicates that those participants with little LCA experience mostly have no prior BIM-LCA experience, showing the lowest GWP optimization performances. But considering the results from Subfigure 6.10 (a), already a little LCA experience (score 2) leads to a marginal difference in the optimization performance compared to participants with more or even high LCA experience.

### Qualitative evaluation of open BIM data formats

First, we evaluate the participants' feedback on open BIM data formats before discussing the prototypical implementation to answer the first research question and evaluate the suitability and utility of the proposed workflow. We are considering IFC and BCF as open BIM data formats.

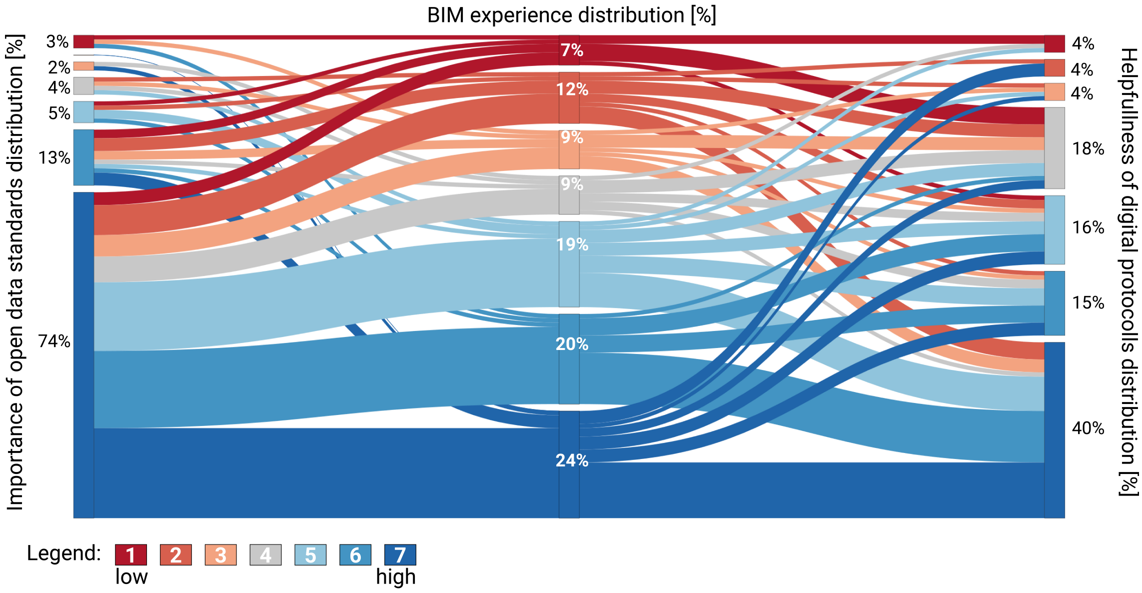



Figure 6.12: Distribution of participants' BIM experience in correlation to their opinion on the importance of open BIM standards (left), and on the helpfulness of digital collaboration protocols (right)

In Figure 6.12, the results of the participant's feedback on the importance of open BIM data formats (Question 14) and the support of the automatic creation of digital collaboration protocols are shown (question 16). 74% of the participants, independent of their BIM expertise, are considering open data standards of the BIM method (e.g., IFC models) for LCA as "very important", while only a small minority of 5% find it rather unimportant. On the right side of Figure 6.12, the correlation between BIM experience and the helpfulness of digital communication protocols is shown. A significant correlation between participants with high BIM experience and helpfulness can be determined (40% for "very helpful"). In contrast, those with little BIM experience tend to find it not helpful at all (4%) or have a neutral perspective on this topic (in total 18%).

This trend becomes even more apparent when correlating participants familiar with the BIM Collaboration Format (BCF), as shown in Figure 6.13. First, a clear correlation between BIM-experienced participants and their knowledge of BCF can be identified on the left side of the figure. More than 80% of those participants with more than average BIM experience know BCF for BIM-based issue management. Most participants with BCF knowledge also find that the implemented digital communication protocol from the prototypical implementation is helpful (ca. 80%), as shown on the right side of Figure 6.13.

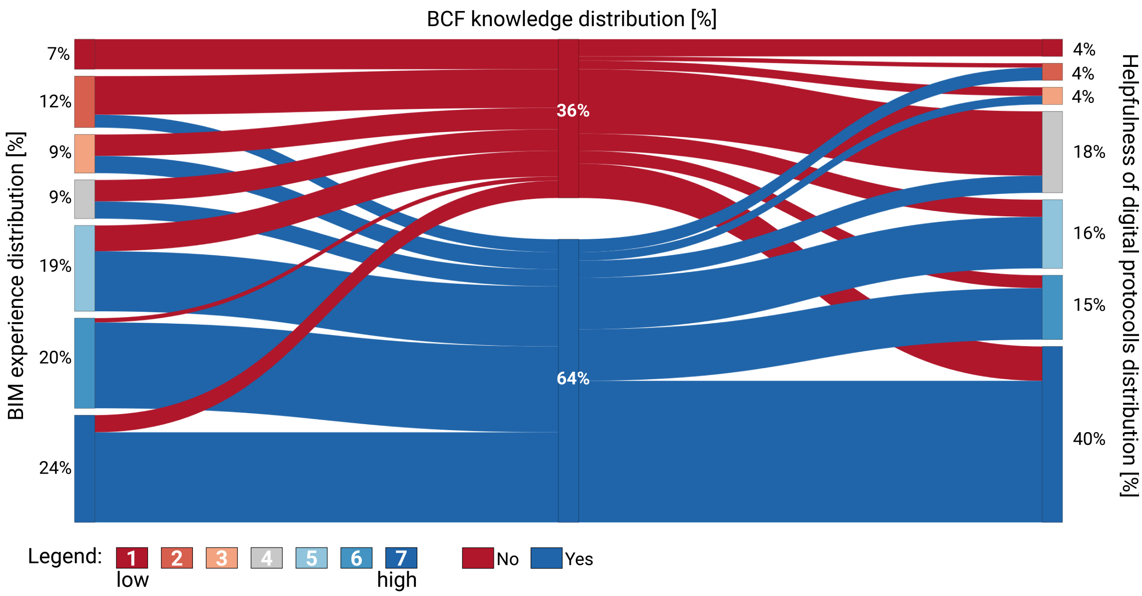



Figure 6.13: Distribution of participants' knowledge of BCF format, in correlation with their BIM experience (left), and on the helpfulness of digital collaboration protocols (right)

The suitability of IFC for semantic model healing and integration in the LCA calculation process in early design stages has already been successfully evaluated in a previous publication (Forth, Abualdenien, & Borrmann, 2023). The IFC data format was successfully used to visualize the case study in a model viewer and integrate it into the design decision-making process. In Section 6.5, we analyze in more detail the support of color-coded BIM models for decision-making.

The second open BIM data format considers BCF. As shown in Figure B.2 from B.1.2, the proposed extension according to Section 6.4.2 is successfully implemented using buildingSMART's BCF API (GitHub, 2023) and setup on mongoDB (MongoDB, 2023). After implementation, it is used in the user study for storing all issues for each design change of each participant. All necessary information is automatically stored as topics and viewpoints on the server and accessed afterward for evaluation. In total, 272 issues were created and later accessed to further assess the results. The snapshots of the viewpoints are stored as binary code using base64, which can be later transferred back to image data in PNG format. B.4 shows a representative overview of most of the snapshots.

To answer the first research question, open BIM data formats support the design decision-making process considering environmental impacts by automating the LCA calculation process using IFC for semantic model healing, visualizing LCA results interactively in an IFC model viewer, and communicating the design decision back to the designer or BIM modeler via an extended BCF schema.

### Qualitative evaluation of LCA-visualization strategies for decision making of non-experts

The decision-making of non-LCA-experts vs. LCA experts is first evaluated by the survey of the different LCA visualization strategies and afterward by tracking the improvement towards GWP that the participants achieved.

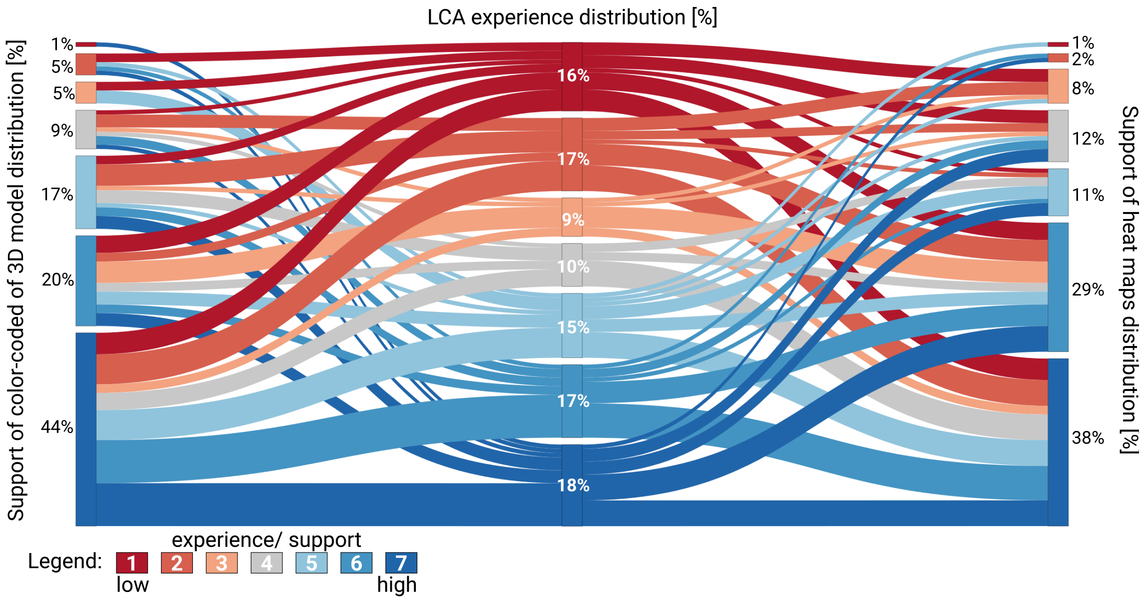



Figure 6.14: Distribution of participants' LCA experience in correlation to their support of color-coded 3D models (left), and on the support of heat maps (right)

Figure 6.14 shows the distribution of the participants' LCA experience in correlation to their support of color-coded 3D-BIM models and the support of color-coded heat maps. The left side indicates that most of the participants, independent of their LCA experience, find the colored 3D model for LCA optimization potential very helpful (44%) or more than average helpful (20% with a score of 6 and 17% with a score of 5).

A similar trend can be identified with the visualization strategy of color-coded heat maps of the element variants and material options. While only 11% of the participants tend to find this visualization strategy less helpful, the majority of around 78% find it more than average helpful. A clear correlation or dependency between the LCA experience and visualization strategy can not be identified.

Figure 6.15 presents the results on the question of how well the box plot diagrams of the LCA results helped the participants make decisions concerning their LCA experience. The average of the results is 5.11 out of a maximum of 7 points, which indicates that they found box plot diagrams rather helpful, with a tendency to a neutral middle. Only a minority of 20% rather disagreed, while generally, most of those responses were from less LCA-experienced stakeholders. Most who voted for the highest score have high or very high LCA experience.

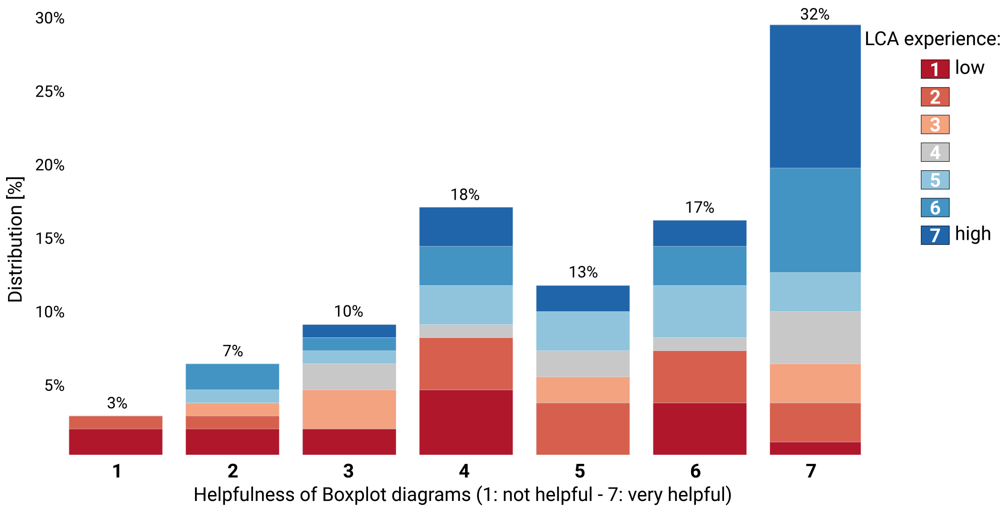



Figure 6.15: Helpfulness of the box plot diagrams in relation to the participants LCA experience

These results generally indicate that these chosen visualization strategies I and II found high acceptance across all LCA experience levels of the participants. The results on LCA visualization strategy III indicate a tendency for box plots as a visualization strategy for more advanced LCA experts and, therefore, also for visualizing uncertain LCA results.

### Qualitative evaluation of decision making based on uncertain LCA results

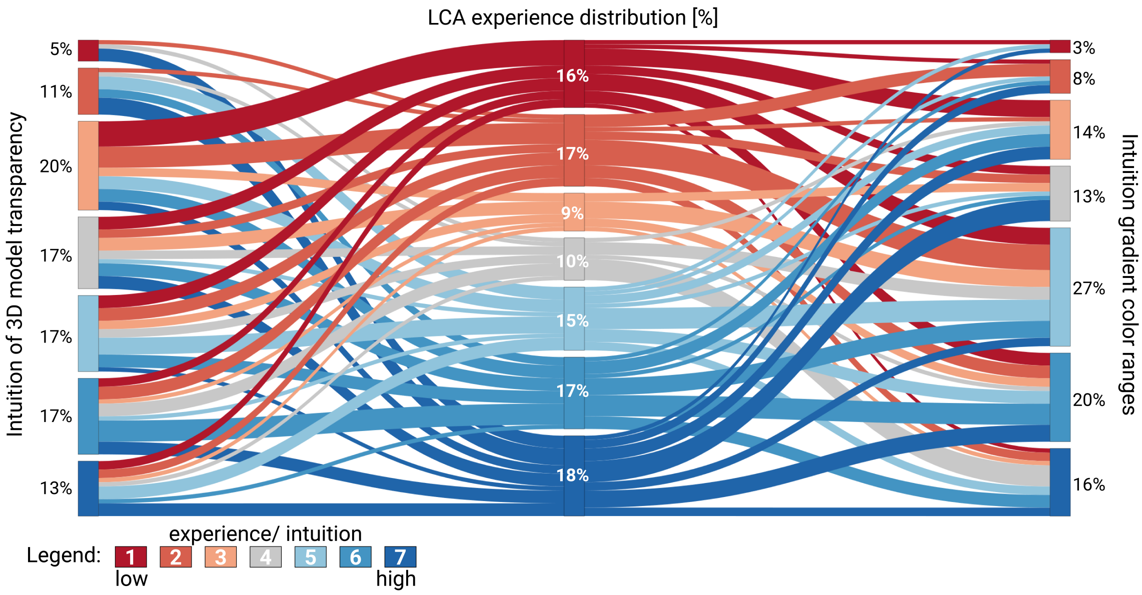



Figure 6.16: LCA experience (middle) in correlation to the intuition of uncertainty visualization considering 3D model transparency (left) and gradient color range heat maps (right)

In Figure 6.16, the correlation between the participants' LCA experience and their rating on how intuitive the uncertainty visualization of model transparency (left) and gradient coloring (right) is shown. The average score of the transparency intuition is 4.34, and of the gradient coloring, 4.77 out of a maximum of 7, so rather a neutral score. Focusing on transparency, only an overall minority of 47% found it rather intuitive, only 13% "very intuitive", and 36% rather unintuitive. These results were also independent of the participants' LCA experience. The gradient coloring for uncertainty visualization seems to have better results than the model-based transparency. A majority of 63% of the participants found it rather intuitive, with only 25% rather unintuitive. Also, no clear correlation between LCA experience and intuition can be identified here.

The third uncertainty visualization strategy is already analyzed together with the other LCA visualization strategies in Section 6.5. In general, it can be stated that these results indicate significant support for box plots for the majority of the stakeholder.

In summary, we found that the highest average scores of uncertainty visualization strategies could be found in the box plot diagrams, but rather for participants with higher LCA experience. Furthermore, a majority of more than 60% of the participants found the gradient coloring and the box plot intuitive and helpful in showing uncertain results. The model-based transparency showed relatively neutral results, independently of the participants' LCA experience. Several participants commented in the written feedback section about the intuitiveness of the transparency visualization.

### 6.5.1 Limitations

The user study and its results showed some limitations. First, the participant numbers were limited and could be extended to have even more reliable results. Next, some participants gave written feedback that combining different element variants is unrealistic and does not always make sense. We previously considered this topic with a method of "Connected Design Decision Networks" but excluded it from the scope of this project (Forth, Schneider-Marin, et al., 2022). More information on other criteria, such as costs, fire safety, etc., was suggested in the written feedback and an extension to different environmental impact categories and their related costs.

Furthermore, the case study is a small office building to simplify the elements' complexity. Testing the proposed visualization strategies with a more complex case study might produce different results. We also pre-defined the goal and scope of the LCA to have comparable system boundaries, such as period, life cycle phases, etc. More transparent information on the calculation was referred to the previous publication (Forth, Abualdenien, & Borrmann, 2023) and was not mentioned in detail, as we also wanted to include non-LCA-experts.

## 6.6 Conclusion and Future Works

This paper aimed to propose and evaluate an interactive visualization approach for design decision-making based on uncertain LCA results using open BIM in early design stages.

Three research questions were defined, including the support of open BIM data formats, which LCA visualization strategies, and which uncertainty visualization approaches are suitable and intuitive for non-LCA-experts in the decision-making process. The proposed workflow's first steps, such as the element matching approach and its LCA calculations, were described in a previous publication (Forth, Abualdenien, & Borrmann, 2023).

The objective of the suggested decision-making workflow was to support non-LCA experts in making design decisions regarding construction-element and material-related options in the early design phases. Different visualization strategies were proposed to support decision-makers during these early stages. These strategies were further developed to incorporate uncertainty visualizations, such as transparency for the 3D BIM-model color-coding, gradient color ranges for the heat maps, and box plot diagrams. Next, we prototypically implemented the proposed workflow based on open BIM data formats using IFC.js for the model viewer and buildingSMART's International BCF API for extending the BCF schema for feedback communication.

We evaluated the three research questions by evaluating the prototype through a user study and a survey. The answer to the first research question on how to open BIM data formats support decision-making includes that IFC models can automatically derive LCA results and visualize them in a color-coded model viewer. Furthermore, an extended BCF schema can be used to communicate the decision back to designers and BIM modelers. For the following research question about LCA visualization strategies, we analyzed that besides the color-coded IFC model viewer, the heat maps of design variants are found to be most supportive for non-LCA-experts. Box plot diagrams are preferred by LCA experts. The last research question is about how uncertainties of LCA results in early design stages can be visualized intuitively. Using transparency in the IFC model for visualizing uncertainties is found to b less intuitive than gradient color ranges and box plots. In contrast, gradient color ranges were rather non-LCA-expert friendly, and box plot diagrams were more intuitive for LCA experts.

The user study had limitations such as a simple office model as a case study, the focus on only embodied GHG emissions, and the number of participants, which can be extended in the future. Some element variant combinations seem unrealistic, which was neglected in this paper as it was previously discussed (Forth, Schneider-Marin, et al., 2022).

To conclude, we could show in this paper the importance and acceptance of open BIM data formats for early design decision support considering LCA. Color-coded 3D models based on IFC models and heat maps also support non-LCA-experts, such as designers, clients, or project developers, to identify LCA hotspots and make design changes to optimize the GWP performance. Therefore, we strongly recommend that LCA software developers implement an interactive model viewer, including color coding, for design optimization in early design stages. This also includes more and different stakeholders in the decision-making process, who are usually not LCA experts, such as project developers or designers.

The aspect of digital protocols, such as BCF, supports automatically communicating all decisions digitally and without information loss and can be used to close the gap of information losses in communication loops. Suppose LCA software providers consider these BCF protocols and implement these in their tools. In that case, non-BIM experts can use them intuitively without having detailed knowledge of BCFs, and participate in a fully open BIM workflow. Nevertheless, standardization of these LCA-related extensions is still needed. Furthermore, different approaches to visualizing LCA results' uncertainties were not intuitive enough for every participant and expertise. Consequently, new visualization strategies need to be researched and tested in the future.

In our ongoing research, we plan to integrate several other criteria in the proposed workflow, such as whole building simulation of the annual energy demand, thermal comfort, and daylight simulations. The limitations of the LCA knowledge database can be solved by integrating the proposed Connected Design Decision Networks and testing the prototype with more complex case studies. Furthermore, the harmonization of extending the BCF schema using BCF snippets to integrate existing LCA data standards such as the International Life Cycle Data System (ILCD) by the European Commission ("European Platform on LCA | EPLCA", 2023) is a further goal of future work.

## 6.7 Acknowledgments

The authors would like to gratefully acknowledge the support of Siemens Real Estate (SRE) for providing financial support for this research and giving access to the analyzed case study.

# Chapter 7

# Conclusion and outlook

## 7.1 Review of research questions

The previous chapters introduced the semantic enrichment methods for the different environmental analyses and the design decision support for addressing the overall thesis' objective of BIM-based semantic enrichment for environmental analyses using Large Language Models. The following chapters discuss the main findings, answer the six research questions and objectives introduced in Section 1.3.1, and highlight the limitations and outlook potential future research.

### 7.1.1 Research question 1: Degree of automation of semantic enrichment for BIM-based environmental analyses

In Chapters 3, 4 and Chapter 5, use-case-specific semantic enrichment approaches for the environmental analyses of Life Cycle Assessment (LCA), Material Passport (MP) and Building Energy Performance Simulation (BEPS) are introduced. As shown in the general framework in Figure 1.6 in Chapter 1.4.2, the overall semantic enrichment method consists of different databases, Large Language Model (LLM), and the matching approach.

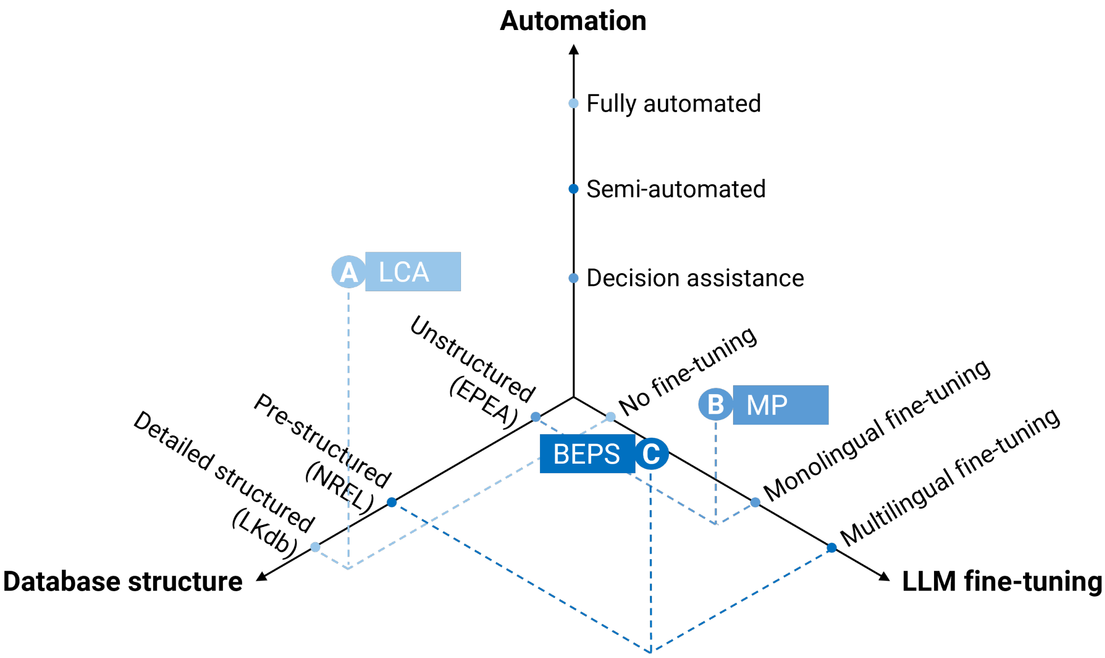



Figure 7.1: Degree of automation in comparison with the degree of hierarchically structured database and LLM fine-tuning level for the three environmental use cases LCA, MP, and BEPS

Figure 7.1 shows the correlation between the degree of automation, hierarchical database structure, and LLM fine-tuning for the three environmental analysis case studies: LCA, MP and BEPS.

A detailed hierarchically structured LCA knowledge database (LKdb) was introduced in Chapter 3 with typical elements, element layers, material categories, and material options. The LKdb is based on the German Life Cycle Inventory (LCI) database Ökobaudat, and each dataset represents one material option's life-cycle phase. Furthermore, the material category allows matching more generic material descriptions, such as *concrete*, without defining specific compressive strength classes. The enhanced database structure allows matching on either element or material level. The pre-trained German BERT model was chosen as LLM without any further fine-tuning. The enhanced LKdb enables a fully automated matching process with only approximately 1.30% wrong matches. The main limitations were wrong element classifications in the BIM model and a few missing elements in the LKdb representing specific construction types. Default elements were set for elements with too little information in the BIM models. The manually calculated LCA results are in the range of the automatically derived results. This indicates that the matching approach is sufficient for a fully automated process, ensuring reliable LCA results.

In Chapter 4, where enrichment for MP is investigated, the material database by EPEA was used with no hierarchical database structure. Therefore, different strategies for monolingually fine-tuning domain-specific LLM and improving matching accuracy were introduced, including their combinations. A combination of the best-performing strategies improved the matching accuracy up to 52.5%. However, these results of the matching accuracy are not sufficient for an automated matching approach but rather support two-step decision-making. Including the ten most similar datasets increased the matching accuracy up to 80.33%. A more detailed structure of the material database is needed to fully automate this process, such as including material classes, e.g., for concrete. Therefore, the degree of automation allows only a decision assistance for material matching.

Chapter 5 answers the research question of the level of automation for the use case of Building Energy Performance Simulation (BEPS). To ensure reliable BEPS results, two use cases of room and element matching were differentiated. As databases, the pre-structured space type and construction databases by National Renewable Energy Laboratory (NREL) were used. A fine-tuning approach was introduced, with different strategies, first for a monolingual, followed by a multilingual fine-tuning step for each use case. A high matching accuracy for the room-specific enrichment was achieved using monolingual fine-tuning but not multilingual. Therefore, a translation step was proposed. Nevertheless, a fully automated matching process of space-type enrichment could be realized.

However, the results of the element-specific matching indicate some ambiguity when a thermal material is present in multiple constructions. Consequently, there is a need to enhance the quality of the database for construction and materials with thermal properties, and the matching process only supports the semi-automatic enrichment. However, the

BEPS results show little deviations compared to the manually enriched Building Energy Modeling (BEM). Nevertheless, the results indicate a combined semi-automated matching approach for both use cases.

In general, the degree of automation of the matching method depends on the level of detail and hierarchical structure of the matching database and the degree of LLM fine-tuning. LLM fine-tuning can improve the degree of automation. However, a full automation of the matching approach is not achievable without a detailed structured database for narrowing down the solution space.

### 7.1.2 Research question 2: Semantic model healing for reliable whole building LCA in early design stages

The proposed semantic model healing approach, introduced in Chapter 3, semantically enriches all relevant and correct elements and material information to the respective model elements for a holistic and reliable LCA in the early design stages.

In the first step, each element's information, including the modeled materials, is matched to the semantically most similar one from a structured LCA knowledge database (LKdb) using Natural Language Processing (NLP), specifically pre-trained Large Language Model (LLM) and the NLP technique of Semantic Textual Similarity (STS). The cosine similarity of the encoded token and term vectors are calculated to measure the semantic similarity of every dataset from the LCA knowledge database (LKdb) with those in the IFC model. As pre-trained LLM, BERT trained on German corpus was used. The matching and similarity assessment takes place on an element or material level.

The proposed LKdb is structured according to the German cost group classification system to filter the solution space. Furthermore, the LKdb considers the two matching levels by further distinguishing between material category and product-specific material information. The LKdb contains information such as LCA datasets on a material or product level based on the German Ökobaudat, and element and material-specific service lives, including the resulting exchange rates. In the LKdb, every cost group consists of different element and material alternatives, including their composition in element layers and missing and not yet modeled material layers, for example, plastering and finishings or reinforcement steel, as well as default elements.

In the next step, all relevant information of the LKdb is enriched to the vague BIM model elements based on the matched element, including missing element layers and materials. Finally, the LCA results for every possible combination of the matched elements are calculated considering the matching level and its design alternatives. This leads to a range of LCA results representing the uncertainty of the early design BIM model.

The method was applied to one case study for LCA calculation, indicating that the calculated LCA result ranges are within the scope of the manually enriched and calculated LCA result. Therefore, the results show more reliability than only calculating the modeled materials, as missing layers were automatically added.

However, one main limitation is that this approach depends on the IFC exchange requirement of correctly assigned classification for each element. The LCA results are incorrect if the IFC element has no or incorrectly assigned classification.

Overall, NLP, especially its subtask STS, in combination with pre-trained LLM, supports automated semantic enrichment for model healing leading to holistic and reliable LCA results, including missing information.

### 7.1.3 Research question 3: Monolingual fine-tuning strategies for Material Passports

Chapter 4 introduced a semi-automated material matching approach to semantically enrich BIM models for Material Passport (MP)s. The matching method uses Semantic Textual Similarity (STS) and different strategies of domain-specific fine-tuning based on pre-trained Large Language Model (LLM).

The five strategies for fine-tuning domain-specific LLM and improving matching performance are: (1) adding domain-specific abbreviations, (2) applying different loss functions for fine-tuning (Cosine Similarity Loss (CosL), Contrastive Loss (ContrastiveL), Multiple Negatives Ranking Loss (MNRL)), (3) adding multiple labels for further context information, such as the name of the IfcTypeObject, the classification, IfcType (IfcWall, IfcWindow, etc.), and further information about load-bearing and external usage, (4) adding negative/ contradicting word pairs, and (5) filtering materials according to its previously used IfcType. The German BERT was applied as pre-trained LLM and sentence embeddings using siamese BERT networks for fine-tuning. 245 unique material-matching samples from 23 real-world case studies were used for fine-tuning, dividing 186 samples for training and 61 for testing.

Material filtering decreases the solution space and has the highest impact on improving the matching accuracy. Furthermore, using Contrastive Loss and adding negative pairs positively impacted the matching accuracy, while adding context information led only to minor improvements. Different combinations of the proposed strategies further increased the matching accuracy. The best-performing combination includes the strategies Contrastive Loss with negative pairs, adding domain-specific abbreviations, and filtering improved the matching accuracy of the test datasets from 60.66% (pre-trained LLM) to 80,33% for matching the 10 most similar material solutions. Nevertheless, including multiple similar material solutions means this workflow can only be used as a support tool instead of a fully automated approach. The main reason is the redundancy of a few material categories, such as concrete, which needs to be structured in a database more detailed in different levels.

However, the proposed fine-tuning workflow only considered one LLM network architecture (BERT). Furthermore, structuring the material database according to a classification system and including material categories, training with more real-world matching datasets, and including cross-validation needs to be realized in future research to get better and more reliable results. Generally, combining different strategies for fine-tuning LLM and structuring the material databases improves the matching accuracy for Material Passports.

### 7.1.4 Research question 4: Multilingual fine-tuning strategies for Building Energy Performance Simulations

Chapter 5 introduced a method of semantic enrichment for Building Energy Performance Simulation (BEPS) based on BIM models using Semantic Textual Similarity (STS) and different strategies of multilingually fine-tuning Large Language Model (LLM).

Two separate use cases were defined for the matching approach and semantic enrichment of Building Energy Modeling (BEM): (1) room-specific enrichment of Mechanical Electric Plumbing (MEP) space types and (2) element-specific enrichment of constructions with thermal properties. Therefore, different strategies for LLM fine-tuning over several training epochs were used for each use case separately to improve the matching accuracy. For both case studies, the pre-trained German BERT LLM was fine-tuned monolingual using strategies similar to the previously discussed use case of Material Passport (MP). Next, the best-performing monolingually fine-tuned LLM was used as a teacher model and further fine-tuned multilingually using Mean Squared Error Loss (MSEL) and Reimers and Gurevych's knowledge distillation for multilingual sentence embeddings (Reimers & Gurevych, 2020). As multilingual strategies, domain-specific material translations and three different student LLM, such as RoBERTa, distiluse-v1 and LaBSE, were evaluated.

Adding negative pairs to the fine-tuning workflow significantly improved the matching accuracy for both use cases, while the Multiple Negatives Ranking Loss (MNRL) led to the highest accuracy with the least computing time. For the fine-tuning, three manually enriched real-world case studies were used and tested with two other case studies.

For the element-specific use case, adding domain-specific abbreviations in combination with MNRL further improved the matching accuracy up to 41.84%. However, the F1-score for matching unique elements is only 44.29%, so there are no sufficient results for a fully automated matching process, although the resulting annual energy demand deviates only 2.11% compared to the manual matched.

For the room-specific use case, the best-performing combination of fine-tuning strategies, using MNRL trained with 15 epochs, improved the accuracy of matching space types to unique room names up to 89.53%. In total, the simulated results of annual energy demand deviate only a maximum of 1.83% for one real-world case study. Therefore, a fully automated matching and enrichment process can be recommended.

However, the multilingual fine-tuning step decreased the matching accuracy. As a result, the recommended workflow includes monolingual fine-tuning followed by a translation step instead of multilingually fine-tuned LLM. Other limitations are that neither different LLM architectures for monolingual fine-tuning have been tested, nor have different databases for construction with thermal properties been tested. Also, only five real-world case studies have been available, three for training and two for testing.

Overall, to improve the matching accuracy, a differentiation between the two use cases of room matching and element matching needs to be done for semantic enrichment for BEPS. A use-case-specific combination of monolingual LLM fine-tuning strategies, including a translation step, significantly improves the accuracy of room-specific matching of space types based on the NREL database for reliable BEPS results. For element-specific enrichment of thermal properties, the fine-tuning strategies have not led to sufficient matching accuracies, leading to minor deviations in BEPS results.

### 7.1.5 Research question 5: Intuitive visualization of uncertainties and LCA results for non-LCA-experts

Chapter 6 shows that combining different visualization strategies supports non-LCA-experts in the decision-making of material- and element-specific design variants.

Three visualization strategies were proposed for visualizing Life Cycle Assessment (LCA) results, such as color-coding of the mean Global Warming Potential (GWP) result in the 3D model representation, heat maps for the material and element variants, as well as bar chart diagrams. To include uncertainties, the color-coded 3D model was further developed using transparencies according to its level of uncertainty, color ranges for the heat maps, and boxplot diagrams instead of bar chart diagrams.

The proposed visualization strategies and the whole reference process were prototypically implemented and tested with a user study. 103 participants, mostly from Germany (81%) and the rest from central Europe, took part in the user study and completed the survey. The level of experience in BIM, LCA, or even both was equally distributed amongst all participants.

The results of the survey showed that the color-coded BIM model and the heat maps were most supportive in the decision-making for non-LCA experts, while LCA experts preferred boxplot diagrams. However, adding transparencies to the color-coded BIM models is less intuitive than gradient color ranges and box plots. Besides the qualitative evaluation using the survey, the overall approach was evaluated quantitatively by measuring the time of decision-making of every design change and its impact on the LCA results. The measurements showed similar results amongst all participants independently of their previous LCA expertise, indicating intuitiveness and overall acceptance.

However, the case study was a similar office building, and the method needs to be validated with a more complex case study, including other disciplines such as MEP. Also, the main focus was decreasing resulting greenhouse gas emissions, which can be extended to further impact categories.

In general, combining different visualization strategies using 3D models and color-coding supports non-LCA experts in intuitive decision-making based on uncertainties and LCA results.

### 7.1.6 Research question 6: Open BIM data formats for the design decision-making process considering environmental analyses

The answer to this research question is divided into two parts: (1) calculating environmental analysis results based on open BIM models and (2) integrating these results in a design decision-making process. First, the results of the different environmental analyses results and the three different use cases of Life Cycle Assessment (LCA), Material Passport (MP), and Building Energy Performance Simulation (BEPS) were briefly discussed.

Chapter 3 introduced a holistic methodology of calculating the environmental impacts of embodied Global Warming Potential (GWP) results based on open BIM workflow in early design stages. Uncertain BIM models, represented in the Industry Foundation Classes (IFC) data format, are sufficient to be semantically healed to derive holistic and reliable LCA results, as described in detail in Section 7.1.2. The IFC models only follow a few exchange requirements, such as base quantities, element-based classifications, and element and material names. For the second use case of MP, IFC was used as a data exchange format for the BIM models. However, the modeling and exchange requirements by Madaster were followed in this case (Frank, 2021). IFC models were partially used for the third use case in Chapter 5, as the geometric transformation from volumetric to surface model is out of scope. Nevertheless, IFC models were exported for semantic enrichment, adding second-level space boundaries as an additional exchange requirement.

To answer the second part of the research question focusing on the integration of the design decision-making part, Chapter 6 proposed the decision-making of design variants and communicating the decisions and results back to the BIM modeler based on the previously calculated LCA results and open BIM workflow in early design stages. IFC models were also used to visualize LCA results and interactively navigate the proposed decision-making platform.

Furthermore, an extension of the BIM Collaboration Format (BCF) schema was proposed to store design decisions and LCA results. Once a design variant is selected and optimized, a BCF issue is created in the extended schema and stored using the BCF server in a document database. These issues can be shown to the BIM modeler and adapted in the BIM model in the authoring tool to close the data-driven design and communication workflow based on open BIM data formats.

Overall, open BIM data formats support the design decision-making process based on environmental analyses by using IFC models with a few exchange requirements for sufficiently calculating environmental analyses and an extended BCF schema enabling cross-vendor feedback communication.

## 7.2 Outlook and future research

Generally, this dissertation used real-world case studies for training and testing, which were limited due to industry partners' limited availability. Future research will extend to more case studies for fine-tuning LLM and testing the matching approaches for several environmental analyses.

Chapter 4 discussed the material enrichment and matching method for the use case of MPs. However, to derive holistic MPs, including Building Circularity Assessment (BCA) of whole buildings, further semantic information towards connection and detachability of the building's elements is necessary. Akbarieh showed the limitations of open BIM models for circularity assessments and used Semantic Web technology to introduce a Decommissioning and Reuse Ontology (DOR) based on Building Topology Ontology (BOT), as well as Semantic Material Banks (SMB) (Akbarieh, 2023). Kaltenegger et al. highlighted the need for a more detailed approach to material information modeling (MIM) to extend BIM for further building performance assessments (Kaltenegger et al., 2024). Abu-Ghaida et al. developed a disassembly network-based approach using graphs to account for product recovery potential for LCA (Abu-Ghaida et al., 2024). Also, Kebede et al. claim ontology-based knowledge graphs as necessary for future research based on the findings of their proposed modular ontology modeling approach for DPPs (Kebede et al., 2024).

Therefore, the proposed LCA knowledge database (LKdb) from Chapter 3 shall be further developed into a knowledge graph database and be enriched with the relevant information using existing ontologies. The advantage of graph databases is that they can be extended with more data sources and information, such as detachability information needed for holistic circularity assessments (Ogunjinmi et al., 2024).

In this dissertation, the focus was mainly on BIM models of new building designs. However, a bigger challenge in the context of circularity is the existing building stock in Europe. Other digital methods, such as technical drawings and point clouds, must be geometrically and semantically processed to assess the existing building stock without manual remodeling. These semantically enriched models of existing buildings can be used for several environmental analyses, such as BCA, material passports, energy simulations, or LCA of retrofitting scenarios.

In a first research project, a methodology was introduced to use several input data for point clouds of the building envelope to automatically derive uncertain building energy models of existing buildings for calculating retrofitting scenarios (Forth, Noichl, et al., 2023). To fine-tune this approach, point clouds of the interior and identify surface materials need to be included.

Furthermore, technical drawings of floor plans of the existing buildings can be used to identify the primary structural materials, which are often hidden when using point clouds only. In the next step, conventional databases of existing structures, such as

TABULA (Loga et al., 2012) or ENOB:dataNWG (Busch & Spars, 2022), can be used to semantically enrich further elements' layers based on the primary structural material and surface materials. These existing databases can also be represented as graph databases to further link additional information, such as detachability information.

# References


Abergel, T., Dean, B., & Dulac, J. (2017). Global Status Report 2017 (United Nations Environment Programme, Ed.).

Abualdenien, J. (2023). *Consistent management and evaluation of building models across design phases* (Doctoral dissertation). Technische Universität München. https://mediatum.ub.tum.de/1687886

Abualdenien, J., Borrmann, A., & König, M. (2021). Ausarbeitungsgrade von bim-modellen. In A. Borrmann, M. König, C. Koch, & J. Beetz (Eds.), *Building information modeling: Technologische grundlagen und industrielle praxis* (pp. 165–191). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-33361-4_8

Abualdenien, J., & Borrmann, A. (2019). A meta-model approach for formal specification and consistent management of multi-LOD building models. *Advanced Engineering Informatics*, *40*, 135–153. https://doi.org/10.1016/j.aei.2019.04.003

Abualdenien, J., & Borrmann, A. (2020). Vagueness visualization in building models across different design stages. *Advanced Engineering Informatics*, *45*, 101107. https://doi.org/10.1016/j.aei.2020.101107

Abu-Ghaida, H., & Kamari, A. (2021). An Alternative Approach to Material and EPD Mapping in The Development of BIM-based LCA and LCC Tools. *Proc. of the Conference CIB W78 2021, Luxembourg*.

Abu-Ghaida, H., Ritzen, M., Hollberg, A., Theissen, S., Attia, S., & Lizin, S. (2024). Accounting for product recovery potential in building life cycle assessments: a disassembly network-based approach. *The International Journal of Life Cycle Assessment*. https://doi.org/10.1007/s11367-024-02324-8

Akbarieh, A. (2023). *Information modelling for circular transformation of built assets' end-of-life* (Doctoral dissertation). Unilu - Université du Luxembourg.

Atik, S., Aparisi, T. D., & Raslan, R. (2020). Investigating the effectiveness and robustness of performing the BIM-based cradle-to-cradle LCA at early-design stages: a case study in the UK. *Proceedings of BSO-V - 5th Building Simulation and Optimization Virtual Conference*.

Atta, I., Bakhoum, E. S., & Marzouk, M. M. (2021). Digitizing material passport for sustainable construction projects using BIM. *Journal of Building Engineering*, *43*, 103233. https://doi.org/10.1016/j.jobe.2021.103233

Bach, A., Al-Wesabi, T., & Staka, I. (2024). Datenzentrierte KI als Basis für ein zukünftiges Informationsmanagement. In S. Haghsheno, G. Satzger, S. Lauble, & M. Vössing (Eds.), *Künstliche Intelligenz im Bauwesen* (pp. 25–39). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-42796-2_2

Bahlau, S., Schumacher, R., Lambertz, M., Theißen, S., Höper, J., Borrmann, A., Forth, K., von Both, P., Ebertshäuser, S., & Horn, R. (2024). Digital Twin Footprint: Erarbeitung eines ganzheitlichen Meilensteinplans mit Handlungsempfehlungen und notwendigen Forschungsbausteinen zur zielführenden Verknüpfung der Lebenszyklusanalyse (Gebäudeökobilanzierung) und BIM-Planungsprozesse mit einem Fokus auf den frühen Planungsphasen (Bundesinstitut für Bau-, Stadt- und Raumforschung, Ed.).

BAMB. (2019). BAMB - Buildings As Material Banks (BAMB2020) - BAMB. Retrieved May 8, 2024, from https://www.bamb2020.eu/

Baumgärtel, K., & Scherer, R. (2016). Automatic ontology-based Green Building Design Parameter Variation and Evaluation in Thermal Energy Building Performance Analyses. *Proc. eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2016*.

Baustart für neues Siemens Technology Center in Garching. (2022). Retrieved October 19, 2022, from https://press.siemens.com/de/de/pressemitteilung/baustart-fuer-neues-siemens-technology-center-garching

BBSR. (2017). Nutzungsdauern von Bauteilen - Informationsportal Nachhaltiges Bauen (Bundesinstitut für Bau-, Stadt- und Raumforschung, Ed.). Retrieved December 28, 2021, from http://www.relaunch-nb.online-now.de/index.php?id=91&L=0

BBSR. (2021). ÖKOBAUDAT (Bundesinstitut für Bau-, Stadt- und Raumforschung, Ed.). Retrieved December 28, 2021, from https://www.oekobaudat.de/

BBSR. (2022). eLCA. Retrieved April 14, 2022, from https://www.bauteileditor.de/

BBSR. (2024). ÖKOBAUDAT (Bundesinstitut für Bau-, Stadt- und Raumforschung, Ed.). Retrieved June 28, 2024, from https://www.oekobaudat.de/

BIM Forum. (2020). Level of Development (LOD) specification Part I & Commentary: For Building Information Models and Data. Retrieved April 7, 2021, from https://bimforum.org/resources/Documents/BIMForum_LOD-Spec-2020.zip

BIM Forum. (2023). Level of Development (LOD) Specification. Retrieved May 14, 2024, from https://bimforum.org/resource/lod-level-of-development-lod-specification/

Bloch, T. (2022). Connecting research on semantic enrichment of BIM - review of approaches, methods and possible applications. *Journal of Information Technology in Construction*, *27*, 416–440. https://doi.org/10.36680/j.itcon.2022.020

Bloch, T., & Sacks, R. (2018). Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Automation in Construction*, *91*, 256–272. https://doi.org/10.1016/j.autcon.2018.03.018

BMI. (2015). Bewertungssystem Nachhaltiges Bauen (BNB) Neubau Büro- und Verwaltungsgebäude: Bilanzierungsregeln für die Erstellung von Ökobilanzen.

Borrmann, A., König, M., Koch, C., & Beetz, J. (Eds.). (2018). *Building Information Modeling: Technological Foundations and Industrial Practice* (1st edition 2018). Springer International Publishing; Springer.

Borrmann, A., König, M., Koch, C., & Beetz, J. (Eds.). (2021). *Building Information Modeling: Technologische Grundlagen und industrielle Praxis*. Springer Vieweg.

Bowman, S. R., Angeli, G., Potts, C., & Manning, C. D. (2015). A large annotated corpus for learning natural language inference. In L. Màrquez, C. Callison-Burch, & J. Su (Eds.), *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 632–642). Association for Computational Linguistics. https://doi.org/10.18653/v1/D15-1075

Brants, S., Dipper, S., Eisenberg, P., Hansen-Schirra, S., König, E., Lezius, W., Rohrer, C., Smith, G., & Uszkoreit, H. (2004). TIGER: Linguistic Interpretation of a German Corpus. *Research on Language and Computation*, *2*(4), 597–620. https://doi.org/10.1007/s11168-004-7431-3

Braune, A., Ekhvaia, L., & Quante, K. (2021). Benchmarks für die Treibhausgasemissionen der Gebäudekonstruktion: Ergebnisse einer Studie mit 50 Gebäuden (Deutsche Gesellschaft für Nachhaltiges Bauen –DGNB e.V., Ed.).

Braune, A., & Wellstein, I. (2024). DGNB Qualitätsstandard für Zirkularitätsindices für Bauwerke: Grundlegendes Qualitätsverständnis und DGNB Zirkularitätsindex Version 1.0. https://www.dgnb.de/de/nachhaltiges-bauen/zirkulaeres-bauen/zirkularitaetsindizes-fuer-bauwerke

Bueno, C., Pereira, L. M., & Fabricio, M. M. (2018). Life cycle assessment and environmental-based choices at the early design stages: an application using building information modelling. *Architectural Engineering and Design Management*, *14*(5), 332–346. https://doi.org/10.1080/17452007.2018.1458593

buildingSMART International (Ed.). (2024). Use Case Management. Retrieved May 10, 2024, from https://ucm.buildingsmart.org/

BuildingSMART International Limited. (2020). Industry Foundation Classes: Version 4-Addendum 2: PsetEnvironmentalImpactIndicators. Retrieved July 11, 2022, from https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2/HTML/link/pset_environmentalimpactindicators.htm

buildingSMART International Ltd. (Ed.). (2024). IfcProduct. Retrieved August 6, 2024, from https://standards.buildingsmart.org/IFC/DEV/IFC4_2/FINAL/HTML/link/ifcproduct.htm

buildingSMART Technical. (2022). Software Implementations - buildingSMART Technical. Retrieved May 15, 2023, from https://technical.buildingsmart.org/resources/software-implementations/

buildingSMART Technical. (2023a). BIM Collaboration Format (BCF) - buildingSMART Technical. Retrieved May 15, 2023, from https://technical.buildingsmart.org/standards/bcf/

buildingSMART Technical. (2023b). Industry Foundation Classes (IFC) - buildingSMART Technical. Retrieved July 10, 2023, from https://technical.buildingsmart.org/standards/ifc

Bundesamt für Bauwesen und Raumordnung (Ed.). (2021). Abkürzungsverzeichnis. Retrieved October 31, 2023, from https://bauarchivddr.bbr-server.de/bauarchivddr/archiv/dokumente/pb-abkuerzungsverzeichnis.pdf

Busch, R., & Spars, G. (2022). ENOB:dataNWG: Screening des Nichtwohngebäudebestands. Retrieved September 6, 2022, from https://www.datanwg.de/downloads/teilberichte/#c348

CAALA. (2024). CAALA Software. Retrieved April 24, 2024, from https://caala.de/features

Carvalho, J. P., Alecrim, I., Bragança, L., & Mateus, R. (2020). Integrating BIM-Based LCA and Building Sustainability Assessment. *Sustainability*, *12*(18), 7468. https://doi.org/10.3390/su12187468

Cavalliere, C., Hollberg, A., Dell'Osso, G. R., & Habert, G. (2019). Consistent BIM-led LCA during the entire building design process. *IOP Conference Series: Earth and Environmental Science*, *323*, 012099. https://doi.org/10.1088/1755-1315/323/1/012099

Cavalliere, C., Brescia, L., Maiorano, G., Dalla Mora, T., Dell'Osso, G. R., & Naboni, E. (2020). Towards An Accessible Life Cycle Assessment: A Literature Based Review Of Current BIM And Parametric Based Tools Capabilities. In V. Corrado, E. Fabrizio, A. Gasparella, & F. Patuzzi (Eds.), *Proceedings of Building Simulation 2019: 16th Conference of IBPSA* (pp. 159–166). IBPSA. https://doi.org/10.26868/25222708.2019.210634

Cavalliere, C., Habert, G., Dell'Osso, G. R., & Hollberg, A. (2019). Continuous BIM-based assessment of embodied environmental impacts throughout the design process. *Journal of Cleaner Production*, *211*, 941–952. https://doi.org/10.1016/j.jclepro.2018.11.247

Cemesova, A., Hopfe, C. J., & Mcleod, R. S. (2015). Passivbim: Enhancing interoperability between bim and low energy design software. *Automation in Construction*, *57*, 17–32.

Cer, D., Diab, M., Agirre, E., Lopez-Gazpio, I., & Specia, L. (2017). SemEval-2017 Task 1: Semantic Textual Similarity Multilingual and Crosslingual Focused Evaluation. In S. Bethard, M. Carpuat, M. Apidianaki, S. M. Mohammad, D. Cer, & D. Jurgens (Eds.), *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 1–14). Association for Computational Linguistics. https://doi.org/10.18653/v1/S17-2001

Çetin, S. (2023). Towards a circular building industry through digitalisation: 1-236 Pages / A+BE | Architecture and the Built Environment, No. 22 (2023): Towards a circular building industry through digitalisation. https://doi.org/10.7480/abe.2023.22.7284

Çetin, S., de Wolf, C., & Bocken, N. (2021). Circular Digital Built Environment: An Emerging Framework. *Sustainability*, *13*(11), 6348. https://doi.org/10.3390/su13116348

Çetin, S., Raghu, D., Honic, M., Straub, A., & Gruis, V. (2023). Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. *Sustainable Production and Consumption*, *40*, 422–437. https://doi.org/10.1016/j.spc.2023.07.011

Chan, B., Schweter, S., & Möller, T. (2020). German's Next Language Model. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 6788–6796). International Committee on Computational Linguistics. https://doi.org/10.18653/v1/2020.coling-main.598

Chandrasekaran, D., & Mago, V. (2022). Evolution of Semantic Similarity—A Survey. *ACM Computing Surveys*, *54*(2), 1–37. https://doi.org/10.1145/3440755

Chen, Z., Feng, Q., Yue, R., Chen, Z., Moselhi, O., Soliman, A., Hammad, A., & An, C. (2022). Construction, renovation, and demolition waste in landfill: a review of waste characteristics, environmental impacts, and mitigation measures. *Environmental science and pollution research international*, *29*(31), 46509–46526. https://doi.org/10.1007/s11356-022-20479-5

Chen, Z., Chen, L., Zhou, X., Huang, L., Sandanayake, M., & Yap, P.-S. (2024). Recent Technological Advancements in BIM and LCA Integration for Sustainable Construction: A Review. *Sustainability*, *16*(3), 1340. https://doi.org/10.3390/su16031340

Ciccozzi, A., de Rubeis, T., Paoletti, D., & Ambrosini, D. (2023). BIM to BEM for Building Energy Analysis: A Review of Interoperability Strategies. *Energies*, *16*(23), 7845. https://doi.org/10.3390/en16237845

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, & Peter J. Liu. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research*, *21*(140), 1–67. http://jmlr.org/papers/v21/20-074.html

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2019). Unsupervised Cross-lingual Representation Learning at Scale. Retrieved August 12, 2024, from https://arxiv.org/pdf/1911.02116

Cornago, S., Ramakrishna, S., & Low, J. S. C. (2023). How can Transformers and large language models like ChatGPT help LCA practitioners? *Resources, Conservation and Recycling*, *196*, 107062. https://doi.org/10.1016/j.resconrec.2023.107062

Costa, G., & Sicilia, A. (2020). Alternatives for facilitating automatic transformation of BIM data using semantic query languages. *Automation in Construction*, *120*, 103384. https://doi.org/10.1016/j.autcon.2020.103384

Dervishaj, A., & Gudmundsson, K. (2024). From LCA to circular design: A comparative study of digital tools for the built environment. *Resources, Conservation and Recycling*, *200*, 107291. https://doi.org/10.1016/j.resconrec.2023.107291

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. https://arxiv.org/pdf/1810.04805

de Wolf, C., Çetin, S., & Bocken, N. M. P. (2024). *A Circular Built Environment in the Digital Age*. Springer International Publishing. https://doi.org/10.1007/978-3-031-39675-5

DGNB. (2024). Der Gebäuderessourcenpass der DGNB. Retrieved May 8, 2024, from https://www.dgnb.de/de/nachhaltiges-bauen/zirkulaeres-bauen/gebaeuderessourcenpass

DGNB GmbH. (2020). DGNB System - New Construction, Buildings - Criteria Set: Version 2020 International.

Di Bari, R., Jorgji, O., Horn, R., Gantner, J., & Ebertshäuser, S. (2019). Step-by-step implementation of BIM-LCA: A case study analysis associating defined construction phases with their respective environmental impacts. *IOP Conference Series: Earth and Environmental Science*, *323*, 012105. https://doi.org/10.1088/1755-1315/323/1/012105

Di Biccari, C., Calcerano, F., D'Uffizi, F., Esposito, A., Campari, M., & Gigliarelli, E. (2022). Building information modeling and building performance simulation interoperability: State-of-the-art and trends in current literature. *Advanced Engineering Informatics*, *54*, 101753. https://doi.org/10.1016/j.aei.2022.101753

DIN 276. (2018). DIN 276:2018-12, Kosten im Bauwesen. https://doi.org/10.31030/2873248

DIN 277. (2021). DIN 277:2021-08, Grundflächen und Rauminhalte im Hochbau. https://doi.org/10.31030/3278234

DIN EN 15643-2. (2021). DIN EN 15643-2:2021-12, Nachhaltigkeit von Bauwerken - Bewertung der Nachhaltigkeit von Gebäuden – Teil 2: Rahmenbedingungen für die Bewertung der umweltbezogenen Qualität Fassung EN_15643_2:2011.

DIN EN 15804. (2022). DIN EN 15804 : 2022-03, Nachhaltigkeit von Bauwerken_- Umweltproduktdeklarationen_- Grundregeln für die Produktkategorie Bauprodukte; Deutsche Fassung EN_15804:2012+A2:2019 + AC:2021. https://doi.org/10.31030/3054172

DIN EN 15978. (2012). DIN EN 15978 : 2012-10, Nachhaltigkeit von Bauwerken_- Bewertung der umweltbezogenen Qualität von Gebäuden_- Berechnungsmethode; Deutsche Fassung EN_15978:2011. https://doi.org/10.31030/1917049

DIN EN 17412-1. (2021). Building Information Modelling – Level of Information Need: Teil 1: Konzepte und Grundsätze.

DIN EN ISO 14040. (2021). DIN EN ISO 14040 : 2021-02, Umweltmanagement_- Ökobilanz_- Grundsätze und Rahmenbedingungen (ISO_14040:2006_+ Amd_1:2020); Deutsche Fassung EN_ISO_14040:2006_+ A1:2020. https://doi.org/10.31030/3179655

DIN EN ISO 14044. (2021). DIN EN ISO 14044 : 2021-02, Umweltmanagement_- Ökobilanz_- Anforderungen und Anleitungen (ISO_14044:2006_+ Amd_1:2017_+ Amd_-2:2020); Deutsche Fassung EN_ISO_14044:2006_+ A1:2018_+ A2:2020. https://doi.org/10.31030/3179656

Ding, Y., Ma, J., & Luo, X. (2022). Applications of natural language processing in construction. *Automation in Construction*, *136*, 104169. https://doi.org/10.1016/j.autcon.2022.104169

Eberhardt, L. C. M., Birgisdóttir, H., & Birkved, M. (2019). Life cycle assessment of a Danish office building designed for disassembly. *Building Research & Information*, *47*(6), 666–680. https://doi.org/10.1080/09613218.2018.1517458

Eckstädt, E., Liersch, C., Hoch, R., & Hentschel, A. (2022). Comparing BIM2SIM Workflows for coupled building and HVAC simulation. *Proceedings of BauSIM 2022, Weimar*.

Eleftheriadis, S., Duffour, P., Greening, P., James, J., Stephenson, B., & Mumovic, D. (2018). Investigating relationships between cost and CO2 emissions in reinforced concrete structures using a BIM-based design optimisation approach. *Energy and Buildings*, *166*, 330–346. https://doi.org/10.1016/j.enbuild.2018.01.059

Ellen MacArthur Foundation (Ed.). (2015). Circularity indictators: An approache to measuring circularity.

Elma, D. (2006). *Transformable building structures. Design for dissassembly as a way to introduce sustainable engineering to building design & construction*. [s.n.]

Elnabawi, M. H. (2020). Building Information Modeling-Based Building Energy Modeling: Investigation of Interoperability and Simulation Results. *Frontiers in Built Environment*, *6*. https://doi.org/10.3389/fbuil.2020.573971

EnergyPlus Development Team. (2010). EnergyPlus engineering reference: The reference to EnergyPlus calculations.

EPEA GmbH (Ed.). (2022). EPEA Generic Dataset: For Madaster | Version 1.0. Retrieved October 30, 2023, from https://docs.madaster.com/files/en/EPEA_Generic_Dataset_Description.pdf

European Commission (Ed.). (2020). In focus: Energy efficiency in buildings. Retrieved October 30, 2023, from https://commission.europa.eu/news/focus-energy-efficiency-buildings-2020-02-17_en

European Commission (Ed.). (2021). *EU Taxonomy, Corporate Sustainability Reporting, Sustainability Preferences and Fiduciary Duties: Directing finance towards the European Green Deal*. https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=CELEX%3A52021DC0188

European Commission. (2022). Regulation of the European parliament and of the council establishing a framework for setting Ecodesign requirements for sustainable products and repealing Directive 2009/125/EC. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0142

European Commission & Directorate-General for Environment. (2021). *Level(s), A common language for building assessment*. Publications Office of the European Union. https://doi.org/10.2779/34137

European platform on lca | eplca. (2023). Retrieved May 27, 2023, from https://eplca.jrc.ec.europa.eu/ilcd.html

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. The MIT Press. https://doi.org/10.7551/mitpress/7287.001.0001

Feng, H., Liyanage, D. R., Karunathilake, H., Sadiq, R., & Hewage, K. (2020). BIM-based life cycle environmental performance assessment of single-family houses: Renovation and reconstruction strategies for aging building stock in British Columbia. *Journal of Cleaner Production*, *250*, 119543. https://doi.org/10.1016/j.jclepro.2019.119543

Figl, H., Ilg, M., & Battisti, K. (2019). 6D BIM–Terminal: Missing Link for the design of CO 2 -neutral buildings. *IOP Conference Series: Earth and Environmental Science*, *323*, 012104. https://doi.org/10.1088/1755-1315/323/1/012104

Fonseca Arenas, N., & Shafique, M. (2023). Recent progress on BIM-based sustainable buildings: State of the art review. *Developments in the Built Environment*, *15*, 100176. https://doi.org/10.1016/j.dibe.2023.100176

Forth, K., Braun, A., & Borrmann, A. (2019). BIM-integrated LCA - model analysis and implementation for practice. *IOP Conference Series: Earth and Environmental Science*, *323*, 012100. https://doi.org/10.1088/1755-1315/323/1/012100

Forth, K., Noichl, F., & Borrmann, A. (2023). LCA Calculation of Retrofitting Scenarios using Geometric Model Reconstruction and Semantic Enrichment of Point Clouds and Images. https://publications.cms.bgu.tum.de/2023_Forth_i3CE.pdf

Forth, K. (2023a). *Multilingual semantic enrichment of room-specific load profiles using BIM models for whole building energy simulation*. Ruhr-Universität Bochum. https://doi.org/10.13154/294-10093

Forth, K. (2023b). *Multilingual semantic enrichment of room-specific load profiles using BIM models for whole building energy simulation*. Ruhr-Universität Bochum. https://doi.org/10.13154/294-10093

Forth, K., Abualdenien, J., & Borrmann, A. (2022). NLP-based semantic model healing for calculating LCA in early building design stages. *Proceedings of the 14th European Conference on Product & Process Modelling (ECPPM 2022)*, 77–84.

Forth, K., Abualdenien, J., & Borrmann, A. (2023). Calculation of embodied GHG emissions in early building design stages using BIM and NLP-based semantic model healing. *Energy and Buildings*, *284*, 112837. https://doi.org/10.1016/j.enbuild.2023.112837

Forth, K., Abualdenien, J., Borrmann, A., Fellermann, S., & Schunicht, C. (2021). Design optimization approach comparing multicriterial variants using BIM in early design stages. *Proceedings of 38th International Symposium on Automation and Robotics in Construction (ISARC 2021)*, 235–242. https://doi.org/10.22260/ISARC2021/0034

Forth, K., Berggold, P., & Borrmann, A. (2024). Domain-specific fine-tuning of LLM for material matching of BIM elements and Material Passports. In B. Akinci & M. Bergés (Eds.), *Proc. of 2024 ASCE International Conference on Computing in Civil Engineering*. American Society of Civil Engineers.

Forth, K., & Borrmann, A. (2024). Semantic enrichment for BIM-based building energy performance simulations using semantic textual similarity and fine-tuning multilingual LLM. *Journal of Building Engineering*, 110312. https://doi.org/10.1016/j.jobe.2024.110312

Forth, K., Hollberg, A., & Borrmann, A. (2023). BIM4EarlyLCA: An interactive visualization approach for early design support based on uncertain LCA results using open BIM. *Developments in the Built Environment*.

Forth, K., Höper, J., Veselka, J., Theißen, S., & Borrmann, A. (2022). Towards life cycle assessment of technical building services in early design phases using building information modelling. *Proceedings of the 2022 European Conference on Computing in Construction*. https://doi.org/10.35490/EC3.2022.178

Forth, K., Schneider-Marin, P., Theißen, S., Höper, J., Svane, N. D., & Borrmann, A. (2022). Connected design decision networks: Multidisciplinary decision support for early building design LCA. *Acta Polytechnica CTU Proceedings*, *38*, 124–130. https://doi.org/10.14311/APP.2022.38.0124

Frank, C. (2021). IFC-Export für die Madaster-Plattform: Version 3.0. Retrieved October 30, 2023, from https://docs.madaster.com/files/de/Madaster%20BIM%20-%20IFC%20export.pdf

Gao, H., Koch, C., & Wu, Y. (2019). Building information modelling based building energy modelling: A review. *Applied Energy*, *238*, 320–343. https://doi.org/10.1016/j.apenergy.2019.01.032

GCR, N. (2004). Cost analysis of inadequate interoperability in the us capital facilities industry. *National Institute of Standards and Technology (NIST)*.

Gebetsroither, M., Honic, M., Kovacic, I., Löffler, C., Marx, K., Pamminger, R., Robbi, S., Sustr, C., Schützenhofer, S., & Weber, G. (2024). Erstellung eines BIM basierten

Gebäudepass. In M. Gebetsroither, M. Honic, I. Kovacic, C. Löffler, K. Marx, R. Pamminger, S. Robbi, C. Sustr, S. Schützenhofer, & G. Weber (Eds.), *Paradigmenwechsel in Bau- und Immobilienwirtschaft* (pp. 51–57). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-68276-0_8

GitHub. (2023). buildingSMART/BCF-API: Web service specification for BIM Collaboration Format. Retrieved May 15, 2023, from https://github.com/buildingSMART/BCF-API

Gomaa, W. H., & Fahmy, A. A. (2013). A Survey of Text Similarity Approaches. *International Journal of Computer Applications*, *68*(13), 13–18. https://doi.org/10.5120/11638-7118

González Viegas, A. (2022). IFCjs/web-ifc-viewer: Graphics engine and toolkit for client applications. Retrieved October 28, 2022, from https://github.com/IFCjs/web-ifc-viewer#readme

Goossens, Y., de Tavernier, J., & Geeraerd, A. (2018). The Risk of Earth Destabilization (RED) index, aggregating the impact we make and what the planet can take. *Journal of Cleaner Production*, *198*, 601–611. https://doi.org/10.1016/j.jclepro.2018.06.284

Goulouti, K., Padey, P., Galimshina, A., Habert, G., & Lasvaux, S. (2020). Uncertainty of building elements' service lives in building LCA & LCC: What matters? *Building and Environment*, *183*, 106904. https://doi.org/10.1016/j.buildenv.2020.106904

Green Building Foundation. (2021). Gbxml - an industry supported standard for storing and sharing building properties between 3d architectural and engineering analysis software. Retrieved June 14, 2022, from https://www.gbxml.org/index.html

Guo, Y., Rennard, V., Xypolopoulos, C., & Vazirgiannis, M. (2021). BERTweetFR : Domain Adaptation of Pre-Trained Language Models for French Tweets. In W. Xu, A. Ritter, T. Baldwin, & A. Rahimi (Eds.), *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)* (pp. 445–450). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.wnut-1.49

Hadsell, R., Chopra, S., & LeCun, Y. (2006). Dimensionality Reduction by Learning an Invariant Mapping. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)*, 1735–1742. https://doi.org/10.1109/CVPR.2006.100

Hamp, B., & Feldweg, H. (1997). GermaNet – a Lexical–Semantic Net for German. *In Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, 9–15.

Harter, H., Singh, M. M., Schneider-Marin, P., Lang, W., & Geyer, P. (2020). Uncertainty Analysis of Life Cycle Energy Assessment in Early Stages of Design. *Energy and Buildings*, *208*, 109635. https://doi.org/10.1016/j.enbuild.2019.109635

Heinrich, M., & Lang, W. (2019). *Materials Passports - Best Practice: Innovative Solutions for a Transition to a Circular Economy in the Built Environment*. Technische Universität München, Fakultät für Architektur.

Helmus, M., Meins-Becker, A., Feller, D. J., Hort, G., Hoffmann, H., Motzo, C., & Norrdine, A. (2021). Leitfaden für die Erstellung eines Bauwerksdatenmodells: Anlage 2: Bezeichnungsvorschläge (Bergische Universität Wuppertal, Ed.).

Henrich, V., & Hinirchs, E. (2010). GernEdiT – The GermaNet Editing Tool. *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010). Valletta, Malta*, pp. 2228–2235.

Hernández, J. L., Lerones, P. M., Bonsma, P., Delft, A. v., Deighton, R., & Braun, J.-D. (2018). An ifc interoperability framework for self-inspection process in buildings. *Buildings*, *8*(2), 32.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hollberg, A. (2016). *A parametric method for building design optimization based on Life Cycle Assessment* (PhD thesis). Bauhaus-Universität Weimar.

Hollberg, A., Genova, G., & Habert, G. (2020). Evaluation of BIM-based LCA results for building design. *Automation in Construction*, *109*, 102972. https://doi.org/10.1016/j.autcon.2019.102972

Hollberg, A., Kiss, B., Röck, M., Soust-Verdaguer, B., Wiberg, A. H., Lasvaux, S., Galimshina, A., & Habert, G. (2021). Review of visualising LCA results in the design process of buildings. *Building and Environment*, *190*, 107530. https://doi.org/10.1016/j.buildenv.2020.107530

Hollberg, A., & Ruth, J. (2016). LCA in architectural design—a parametric approach. *The International Journal of Life Cycle Assessment*, *21*(7), 943–960. https://doi.org/10.1007/s11367-016-1065-1

Hollberg, A., Tjäder, M., Ingelhag, G., & Wallbaum, H. (2022). A Framework for User Centric LCA Tool Development for Early Planning Stages of Buildings. *Frontiers in Built Environment*, *8*. https://doi.org/10.3389/fbuil.2022.744946

Hong, T., Langevin, J., & Sun, K. (2018). Building simulation: Ten challenges. *Building Simulation*, *11*(5), 871–898. https://doi.org/10.1007/s12273-018-0444-x

Honic, M., Kovacic, I., & Rechberger, H. (2019a). Concept for a BIM-based Material Passport for buildings. *IOP Conference Series: Earth and Environmental Science*, *225*, 012073. https://doi.org/10.1088/1755-1315/225/1/012073

Honic, M., Kovacic, I., & Rechberger, H. (2019b). BIM-Based Material Passport (MP) as an Optimization Tool for Increasing the Recyclability of Buildings. *Applied Mechanics and Materials*, *887*, 327–334. https://doi.org/10.4028/www.scientific.net/AMM.887.327

Honic, M., Magalhães, P. M., & van den Bosch, P. (2024). From Data Templates to Material Passports and Digital Product Passports. In C. de Wolf, S. Çetin, & N. M. P. Bocken (Eds.), *A Circular Built Environment in the Digital Age* (pp. 79–94). Springer International Publishing. https://doi.org/10.1007/978-3-031-39675-5{\textunderscore}5

Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. *To appear*, *7*(1), 411–420.

Horn, R., Ebertshäuser, S., Di Bari, R., Jorgji, O., Traunspurger, R., & von Both, P. (2020). The BIM2LCA Approach: An Industry Foundation Classes (IFC)-Based Interface to

Integrate Life Cycle Assessment in Integral Planning. *Sustainability*, *12*(16), 6558. https://doi.org/10.3390/su12166558

IEA. (2024). Buildings - Energy System - IEA. Retrieved May 13, 2024, from https://www.iea.org/energy-system/buildings

Jadwin, A., & Huang, C. (2023). Improving minBERT Performance on Multiple Tasks through In-domain Pretraining, Negatives Ranking Loss Learning, and Hyperparameter Optimization. *Stanford University*.

Jiang, J., & Conrath, D. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. *Proc. of the Int'l. Conf. on Research in Computational Linguistics*, 19–33. http://www.cse.iitb.ac.in/~cs626-449/Papers/WordSimilarity/4.pdf

Kaltenegger, J., Frandsen, K. M., & Petrova, E. (2024). A data management perspective on building material classification: A systematic review. *Journal of Building Engineering*, 109172. https://doi.org/https://doi.org/10.1016/j.jobe.2024.109172

Kamari, A., Kotula, B. M., & Schultz, C. P. L. (2022). A BIM-based LCA tool for sustainable building design during the early design stage. *Smart and Sustainable Built Environment*, *11*(2), 217–244. https://doi.org/10.1108/SASBE-09-2021-0157

Kebede, R., Moscati, A., Tan, H., & Johansson, P. (2024). A modular ontology modeling approach to developing digital product passports to promote circular economy in the built environment. *Sustainable Production and Consumption*, *48*, 248–268. https://doi.org/10.1016/j.spc.2024.05.007

Kiamili, C., Hollberg, A., & Habert, G. (2020). Detailed Assessment of Embodied Carbon of HVAC Systems for a New Office Building Based on BIM. *Sustainability*, *12*(8), 3372. https://doi.org/10.3390/su12083372

Kim, H., Shen, Z., Kim, I., Kim, K., Stumpf, A., & Yu, J. (2016). BIM IFC information mapping to building energy analysis (BEA) model with manually extended material information. *Automation in Construction*, *68*, 183–193. https://doi.org/10.1016/j.autcon.2016.04.002

Kiss, B., Kácsor, E., & Szalay, Z. (2020). Environmental assessment of future electricity mix – Linking an hourly economic model with LCA. *Journal of Cleaner Production*, *264*, 121536. https://doi.org/10.1016/j.jclepro.2020.121536

Kiss, B., & Szalay, Z. (2019). A Visual Method for Detailed Analysis of Building Life Cycle Assessment Results. *Applied Mechanics and Materials*, *887*, 319–326. https://doi.org/10.4028/www.scientific.net/AMM.887.319

Knotten, V., Svalestuen, F., Hansen, G. K., & Lædre, O. (2015). Design Management in the Building Process - A Review of Current Literature. *Procedia Economics and Finance*, *21*, 120–127. https://doi.org/10.1016/S2212-5671(15)00158-6

Kohle, L., & Jannidis, F. (2020). Domain and Task Adaptive Pretraining for Language Models. *CHR 2020: Workshop on Computational Humanities Research*.

Komatsuzaki, A. (2019). One Epoch Is All You Need. Retrieved August 14, 2024, from http://arxiv.org/pdf/1906.06669v1

Kongsgaard, C. (2024). LCAx v.2.1.2. Retrieved May 10, 2024, from https://github.com/ocni-dtu/lcax.git

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C.J. Burges, L. Bottou, & K.Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems*. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Lai, H., & Deng, X. (2018). Interoperability analysis of ifc-based data exchange between heterogeneous bim software. *Journal of Civil Engineering and Management*, *24*(7), 537–555.

Lammers, B., & Forth, K. (2022). IFC-based variant analysis considering multicriterial sustainability analysis of buildings. *Proc. of 33. Forum Bauinformatik, München*, (under review).

Lan, W., & Xu, W. (2018). Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering. https://doi.org/10.48550/arXiv.1806.04330

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction* (Second edition). Elsevier Morgan Kaufmann Publishers.

Leacock, C., & Chodorow, M. (1998). Combining local context and WordNet similarity for word sense identification. In *Fellbaum, C., ed., WordNet: An electronic lexical* (pp. 265–283).

Lee, J.-S. (2021). Life Cycle Costing for Exterior Materials on Building Façade. *Journal of Construction Engineering and Management*, *147*(7). https://doi.org/10.1061/(ASCE)CO.1943-7862.0002068

Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics. Doklady*, *10*, 707–710. https://api.semanticscholar.org/CorpusID:60827152

Li, H., Zhang, J., Chang, S., & Sparkling, A. (2023). BIM-based object mapping using invariant signatures of AEC objects. *Automation in Construction*, *145*, 104616. https://doi.org/10.1016/j.autcon.2022.104616

Li, R., Cheng, L., Wang, D., & Tan, J. (2023). Siamese BERT Architecture Model with attention mechanism for Textual Semantic Similarity. *Multimedia Tools and Applications*, *82*(30), 46673–46694. https://doi.org/10.1007/s11042-023-15509-4

Liebich, T. (2013). IFC4—the new buildingSMART standard. *IC Meeting*.

Lin, D. (1998). An information-theoretic definition of similarity. *Proceedings of the International Conference on Machine Learning*.

Llatas, C., Soust-Verdaguer, B., Hollberg, A., Palumbo, E., & Quiñones, R. (2022). BIM-based LCSA application in early design stages using IFC. *Automation in Construction*, *138*, 104259. https://doi.org/10.1016/j.autcon.2022.104259

Llatas, C., Soust-Verdaguer, B., & Passer, A. (2020). Implementing Life Cycle Sustainability Assessment during design stages in Building Information Modelling: From systematic literature review to a methodological approach. *Building and Environment*, *182*, 107164. https://doi.org/10.1016/j.buildenv.2020.107164

Lobaccaro, G., Wiberg, A. H., Ceci, G., Manni, M., Lolli, N., & Berardi, U. (2018). Parametric design to minimize the embodied GHG emissions in a ZEB. *Energy and Buildings*, *167*, 106–123. https://doi.org/10.1016/j.enbuild.2018.02.025

Locatelli, M., Seghezzi, E., Pellegrini, L., Tagliabue, L. C., & Di Giuda, G. M. (2021). Exploring Natural Language Processing in Construction and Integration with Building Information Modeling: A Scientometric Analysis. *Buildings*, *11*(12), 583. https://doi.org/10.3390/buildings11120583

Loga, T., Diefenback, N., & Stein, B. (Eds.). (2012). Typology Approach for Building Stock Energy Typology Approach for Building Stock Energy Assessment. Main Results of the TABULA project: Final Project Report. https://doi.org/10.2172/1000008

Löhr, F., Gerber, A., & Stobbe, M. (2022). Semi-automated generation of multi-zone thermal models from Building Information Modeling data. *Proceedings of BauSIM 2022, Weimar*.

Luo, X., Li, H., & Lee, S. (2023). Bridging the gap: Neuro-Symbolic Computing for advanced AI applications in construction. *Frontiers of Engineering Management*, *10*(4), 727–735. https://doi.org/10.1007/s42524-023-0266-0

MacLeamy, P. (2004). Collaboration, Integrated Information, and the Project Lifecycle in Building Design and Construction and Operation. Retrieved February 7, 2015, from http://codebim.com/wp-content/uploads/2013/06/CurtCollaboration.pdf

Madaster (Ed.). (2024). Madaster Release notes.

Marsh, E., Allen, S., & Hattam, L. (2023). Tackling uncertainty in life cycle assessments for the built environment: A review. *Building and Environment*, *231*, 109941. https://doi.org/10.1016/j.buildenv.2022.109941

Meex, E., Hollberg, A., Knapen, E., Hildebrand, L., & Verbeeck, G. (2018). Requirements for applying LCA-based environmental impact assessment tools in the early stages of building design. *Building and Environment*, *133*, 228–236. https://doi.org/10.1016/j.buildenv.2018.02.016

Meng, Z., Zahedi, A., & Petzold, F. (2020). Web-Based Communication Platform for Decision Making in Early Design Phases. In H. "Osumi (Ed.), *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*. International Association for Automation and Robotics in Construction (IAARC). https://doi.org/10.22260/ISARC2020/0138

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. https://arxiv.org/pdf/1301.3781

Miyamoto, A., Allacker, K., & de Troyer, F. (2022). Visual tool for sustainable buildings: A design approach with various data visualisation techniques. *Journal of Building Engineering*, *56*, 104741. https://doi.org/10.1016/j.jobe.2022.104741

MongoDB. (2023). MongoDB: The Developer Data Platform. Retrieved May 25, 2023, from https://www.mongodb.com/

Mousa, M., Luo, X., & McCabe, B. (2016). Utilizing BIM and Carbon Estimating Methods for Meaningful Data Representation. *Procedia Engineering*, *145*, 1242–1249. https://doi.org/10.1016/j.proeng.2016.04.160

mrdoob. (2022). three.js: JavaScript 3D Library. Retrieved October 28, 2022, from https://github.com/mrdoob/three.js/

Müller, D., Jansen, D., Treeck, C., Fichter, E., Richter, V., Lüdemann, B., Jahangiri, P., Brunkhorst, J., Dahncke, M., & Warnecke, C. (2021). BIM2SIM - Methodenentwicklung zur Erstellung von Simulationsmodellen aus Daten des Building Information Modeling : gemeinsamer Endbericht : Berichtszeitraum: 01.05.2018-31.10.2021. *BIM2SIM - Methodenentwicklung zur Erstellung von Simulationsmodellen aus Daten des Building Information Modeling*. https://doi.org/10.2314/KXP:1819319997

Naneva, A. (2022). greenBIM, a BIM-based LCA integration using a circular approach based on the example of the Swiss sustainability standard Minergie-ECO. *E3S Web of Conferences*, *349*, 10002. https://doi.org/10.1051/e3sconf/202234910002

Naneva, A., Bonanomi, M., Hollberg, A., Habert, G., & Hall, D. (2020). Integrated BIM-Based LCA for the Entire Building Process Using an Existing Structure for Cost Estimation in the Swiss Context. *Sustainability*, *12*(9), 3748. https://doi.org/10.3390/su12093748

Navigli, R., & Martelli, F. (2019). An overview of word and sense similarity. *Natural Language Engineering*, *25*(06), 693–714. https://doi.org/10.1017/S1351324919000305

Nilsen, M., & Bohne, R. A. (2019). Evaluation of BIM based LCA in early design phase (low LOD) of buildings. *IOP Conference Series: Earth and Environmental Science*, *323*, 012119. https://doi.org/10.1088/1755-1315/323/1/012119

Nizam, R. S., Zhang, C., & Tian, L. (2018). A BIM based tool for assessing embodied energy for buildings. *Energy and Buildings*, *170*, 1–14. https://doi.org/10.1016/j.enbuild.2018.03.067

NREL (Ed.). (2024). OpenStudio. Retrieved January 30, 2024, from https://openstudio.net/

OECD. (2019). *Global Material Resources Outlook to 2060*. https://doi.org/10.1787/9789264307452-en

Ogunjinmi, G. J., Forth, K., Theißen, S., & Borrmann, A. (2024). Estimating the circularity of building elements using building information modelling. In *Proc. of world sustainable built environment conference 2024*.

OpenAI. (2022). Introducing ChatGPT. Retrieved May 11, 2024, from https://openai.com/blog/chatgpt.

Palumbo, E., Soust-Verdaguer, B., Llatas, C., & Traverso, M. (2020). How to Obtain Accurate Environmental Impacts at Early Design Stages in BIM When Using Environmental Product Declaration. A Method to Support Decision-Making. *Sustainability*, *12*(17), 6927. https://doi.org/10.3390/su12176927

Parece, S., Resende, R., & Rato, V. (2024). A bim-based tool for embodied carbon assessment using a construction classification system. *Developments in the Built Environment*, *19*, 100467. https://doi.org/https://doi.org/10.1016/j.dibe.2024.100467

Park, I. (2019). CS274 Project: Semantic Textual Similarity using Transfer Learning and Embeddings. Retrieved April 24, 2024, from https://gitlab.com/ipark/cs274-wi

Peffers, K., Rothenberger, M., & Kuechler, W. (Eds.). (2012). *Design science research in information systems: Advances in theory and practice ; 7th international conference, DESRIST 2012, Las Vegas, NV, USA, May 14-15, 2012 : proceedings* (Vol. 7286). Springer. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=3070047

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, (vol. 24 no. 3), 45–77.

Petrova, E., Pauwels, P., Svidt, K., & Jensen, R. L. (2019). Towards data-driven sustainable design: decision support based on knowledge discovery in disparate building data. *Architectural Engineering and Design Management*, *15*(5), 334–356. https://doi.org/10.1080/17452007.2018.1530092

Pollination (Ed.). (2024). Export Analytical Model. Retrieved January 30, 2024, from https://docs.pollination.cloud/user-manual/revit-plugin/export-analytical-model

Potrč Obrecht, T., Röck, M., Hoxha, E., & Passer, A. (2020). BIM and LCA Integration: A Systematic Literature Review. *Sustainability*, *12*(14), 5534. https://doi.org/10.3390/su12145534

Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *Systems, Man and Cybernetics, IEEE Transactions on*, *19*, 17–30. https://doi.org/10.1109/21.24528

Raggi, E., Tavaroli, F., Zinno, A., Regidor Sànchez, M., & Gonzales Viegas, A. (2021). Analysis of BIM-BEM critical parameters and recommendations to solve the current bottlenecks.

Ramaji, I. J., Messner, J. I., & Mostavi, E. (2020). IFC-Based BIM-to-BEM Model Transformation. *Journal of Computing in Civil Engineering*, *34*(3). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000880

Reimers, N. (2023). Sentence-Transformers Losses. Retrieved November 14, 2023, from https://www.sbert.net/docs/package_reference/losses.html

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*.

Reimers, N., & Gurevych, I. (2020, April 21). Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation. https://arxiv.org/pdf/2004.09813

Reitschmidt, G. (2015). *Ökobilanzierung auf Basis von Building Information Modeling: Entwicklung eines Instruments zur automatisierten Ökobilanzierung der Herstellungsphase von Bauwerken unter Nutzung der Ökobau.dat und Building Information Modeling* (Masterthesis). Technische Hochschule Mittelhessen. Gießen.

Remmen, P., Lauster, M., Mans, M., Fuchs, M., Osterhage, T., & Müller, D. (2018). TEASER: an open tool for urban energy modelling of building stocks. *Journal of Building Performance Simulation*, *11*(1), 84–98. https://doi.org/10.1080/19401493.2017.1283539

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. *IJCAI'95: Proceedings of the 14th international joint conference on Artificial intelligence*, *Volume 1*, 448–453.

Rezaei, F., Bulle, C., & Lesage, P. (2019). Integrating building information modeling and life cycle assessment in the early and detailed building design stages. *Building and Environment*, *153*, 158–167. https://doi.org/10.1016/j.buildenv.2019.01.034

Richter, C., & Liedtke, S. (2021). *BKI IFC-Bildkommentar: Ausgewählte IFC 4 Begriffe für die BIM-Planungsarbeit gegliedert nach DIN 276*. Rudolf Müller.

Richter, V. E., Syndicus, M., Frisch, J., & van Treeck, C. (2023). Extending the IFC-Based bim2sim Framework to Improve the Accessibility of Thermal Comfort Analysis Considering Future Climate Scenarios. *Applied Sciences*, *13*(22), 12478. https://doi.org/10.3390/app132212478

Roberts, M., Allen, S., & Coley, D. (2020). Life cycle assessment in the building design process – A systematic literature review. *Building and Environment*, *185*, 107274. https://doi.org/10.1016/j.buildenv.2020.107274

Röck, M., Hollberg, A., Habert, G., & Passer, A. (2018a). LCA and BIM: Integrated Assessment and Visualization of Building Elements' Embodied Impacts for Design Guidance in Early Stages. *Procedia CIRP*, *69*, 218–223. https://doi.org/10.1016/j.procir.2017.11.087

Röck, M., Hollberg, A., Habert, G., & Passer, A. (2018b). LCA and BIM: Visualization of environmental potentials in building construction at early design stages. *Building and Environment*, *140*, 153–161. https://doi.org/10.1016/j.buildenv.2018.05.006

Röck, M., Saade, M. R. M., Balouktsi, M., Rasmussen, F. N., Birgisdottir, H., Frischknecht, R., Habert, G., Lützkendorf, T., & Passer, A. (2020). Embodied GHG emissions of buildings – The hidden challenge for effective climate change mitigation. *Applied Energy*, *258*, 114107. https://doi.org/10.1016/j.apenergy.2019.114107

Rosen, A. (2021). *Urban Mining Index: Entwicklung einer Systematik zur quantitativen Bewertung der Kreislaufkonsistenz von Baukonstruktionen in der Neubauplanung*. Fraunhofer IRB Verlag.

Sachidananda, V., Kessler, J. S., & Lai, Y.-a. (2021, September 15). Efficient Domain Adaptation of Language Models via Adaptive Tokenization. http://arxiv.org/pdf/2109.07460v1

Sacks, R., Girolami, M., & Brilakis, I. (2020). Building Information Modelling, Artificial Intelligence and Construction Tech. *Developments in the Built Environment*, *4*, 100011. https://doi.org/10.1016/j.dibe.2020.100011

Sadeghipour Roudsari, M., & Pak, M. (2013). Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. *Proceedings of the 13th International IBPSA Conference Held in Lyon, France*.

Safari, K., & AzariJafari, H. (2021). Challenges and opportunities for integrating BIM and LCA: Methodological choices and framework development. *Sustainable Cities and Society*, *67*, 102728. https://doi.org/10.1016/j.scs.2021.102728

Sanchez, B., Honic, M., Leite, F., Herthogs, P., & Stouffs, R. (2024). Augmenting materials passports to support disassembly planning based on building information modelling standards. *Journal of Building Engineering*, 109083. https://doi.org/10.1016/j.jobe.2024.109083

Santos, R., Aguiar Costa, A., Silvestre, J. D., & Pyl, L. (2020). Development of a BIM-based Environmental and Economic Life Cycle Assessment tool. *Journal of Cleaner Production*, *265*, 121705. https://doi.org/10.1016/j.jclepro.2020.121705

Santos, R., Costa, A. A., Silvestre, J. D., & Pyl, L. (2019). Integration of LCA and LCC analysis within a BIM-based environment. *Automation in Construction*, *103*, 127–149. https://doi.org/10.1016/j.autcon.2019.02.011

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. https://doi.org/10.1016/j.aei.2011.08.007

Scherz, M., Hoxha, E., Kreiner, H., Passer, A., & Vafadarnikjoo, A. (2022). A hierarchical reference-based know-why model for design support of sustainable building envelopes. *Automation in Construction*, *139*, 104276. https://doi.org/10.1016/j.autcon.2022.104276

Schneider-Marin, P., & Abualdenien, J. (2019). A framework to facilitate an interdisciplinary design process using BIM. *Proc. of the 31th Forum Bauinformatik*.

Schneider-Marin, P., Harter, H., Tkachuk, K., & Lang, W. (2020). Uncertainty Analysis of Embedded Energy and Greenhouse Gas Emissions Using BIM in Early Design Stages. *Sustainability*, *12*(7), 2633. https://doi.org/10.3390/su12072633

Schneider-Marin, P., & Lang, W. (2020). Environmental costs of buildings: monetary valuation of ecological indicators for the building industry. *The International Journal of Life Cycle Assessment*, *25*(9), 1637–1659. https://doi.org/10.1007/s11367-020-01784-y

Schneider-Marin, P., Stocker, T., Abele, O., Margesin, M., Staudt, J., Abualdenien, J., & Lang, W. (2022). EarlyData knowledge base for material 1 decisions in building design. *Advanced Engineering Informatics*.

Schumacher, R. (2021). *Potenzialanalyse zur Bewertung grauer Energie und Emissionen im digitalen Gebäudeplanungsprozess* (Bachelor Thesis). Technische Hochschule Köln. Köln.

Schumacher, R., Theißen, S., Höper, J., Drzymalla, J., Lambertz, M., Hollberg, A., Forth, K., Schneider-Marin, P., Wimmer, R., Bahlau, S., & Meins-Becker, A. (2022). Analysis of current practice and future potentials of LCA in a BIM-based design process in Germany. *E3S Web of Conferences*, *349*, 10004. https://doi.org/10.1051/e3sconf/202234910004

Siemens Deutschland. (2022). Frankfurts innovativstes Büro. Retrieved October 19, 2022, from https://new.siemens.com/de/de/produkte/services/real-estate-services/develop/the-move.html

Song, J., Kim, J., & Lee, J.-K. (2019). Spatial Information Enrichment using NLP-based Classification of Space Objects for School Bldgs in Korea. *Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC)*, 415–420.

Soust-Verdaguer, B., Bernardino Galeana, I., Llatas, C., Montes, M. V., Hoxha, E., & Passer, A. (2022). How to conduct consistent environmental, economic, and social assessment during the building design process. A BIM-based Life Cycle Sustainability Assessment method. *Journal of Building Engineering*, *45*, 103516. https://doi.org/10.1016/j.jobe.2021.103516

Soust-Verdaguer, B., Llatas, C., & García-Martínez, A. (2017). Critical review of bim-based LCA method to buildings. *Energy and Buildings*, *136*, 110–120. https://doi.org/10.1016/j.enbuild.2016.12.009

Spielhaupter, O. (2021). BIM to BEM Transformation Workflows: A Case Study Comparing Different IFC-Based Approaches.

Stenzel, V. (2020). *Wissensdatenbank für Graue Energie und Treibhauspotenzial von Baustoffen* (Masterthesis). München. Technische Universität München, Lehrstuhl für energieeffizientes und nachhaltiges Planen und Bauen.

Ströbele, B. (2022). *Ermittlung von Umweltwirkungen im Lebenszyklus von Gebäuden innerhalb der Planungsphase auf der Grundlage von unscharfen Daten* (1. Auflage). Shaker.

Suchanek, F. M., & Holzenberger, N. (2024). Knowledge-Based Language Models. Retrieved May 11, 2024, from https://suchanek.name/work/research/kb-lm/kb-lm-white-paper.pdf

Tam, V. W., Zhou, Y., Illankoon, C., & Le, K. N. (2022). A critical review on BIM and LCA integration using the ISO 14040 framework. *Building and Environment*, 108865. https://doi.org/10.1016/j.buildenv.2022.108865

Theißen, S., Drzymalla, J., Höper, J., Liermann, E., Wimmer, R., Meins-Becker, A., Henne, A., Kloster, N., & Lambertz, M. (2020). Digitalization of user-oriented demand planning through Building Information Modeling (BIM). *IOP Conference Series: Earth and Environmental Science*, *588*, 032004. https://doi.org/10.1088/1755-1315/588/3/032004

Theißen, S., Höper, J., Wimmer, R., Zibell, M., Meins-Becker, A., Rössig, S., Goitowski, S., & Lambertz, M. (2020). BIM integrated automation of whole building life cycle assessment using German LCA data base ÖKOBAUDAT and Industry Foundation Classes. *IOP Conference Series: Earth and Environmental Science*, *588*, 032025. https://doi.org/10.1088/1755-1315/588/3/032025

Tian, W., Heo, Y., de Wilde, P., Li, Z., Da Yan, Park, C. S., Feng, X., & Augenbroe, G. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, *93*, 285–301. https://doi.org/10.1016/j.rser.2018.05.029

Tobias Schnabel, Igor Labutov, David Mimno, & Thorsten Joachims. (2015). Evaluation methods for unsupervised word embeddings. *Proceedings ofthe 2015 Conference on Empirical Methods in Natural Language Processing*, (298-307).

Tomczak, A., van Berlo, L., Krijnen, T., Borrmann, A., & Bolpagni, M. (2022). A review of methods to specify information requirements in digital construction projects. *Prof. of CIB World Congress 2022*. https://publications.cms.bgu.tum.de/2022_Tomczak_CIB.pdf

Tomczak, A., Benghi, C., van Berlo, L., & Hjelseth, E. (2024). Requiring Circularity Data in Bim With Information Delivery Specification. *Circular Economy*, *1*(2). https://doi.org/10.55845/REJY5239

Tsikos, M., & Negendahl, K. (2017). Sustainable Design with Respect to LCA Using Parametric Design and BIM Tools. *Proc. of World Sustainable Built Environment Conference 2017 , Wan Chai, Hong Kong*.

United Nations Environment Programme. (2022). 2022 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector.

van Treeck, C., Wimmer, R., & Maile, T. (2018). BIM for Energy Analysis. In A. Borrmann, M. König, C. Koch, & J. Beetz (Eds.), *Building Information Modeling* (pp. 337–347). Springer International Publishing; Springer. https://doi.org/10.1007/978-3-319-92862-3_20

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. https://doi.org/10.48550/arXiv.1706.03762

VDI 2552 Blatt 9. (2022). Building information modeling - Classification systems.

Verband Beratender Ingenieure VBI (Ed.). (2016). BIM-Leitfaden für die Planerpraxis: Empfehlungen für planende und beratende Ingenieure.

Veselka, J., Nehasilová, M., Dvořáková, K., Ryklová, P., Volf, M., Růžička, J., & Lupíšek, A. (2020). Recommendations for Developing a BIM for the Purpose of LCA in Green Building Certifications. *Sustainability*, *12*(15), 6151. https://doi.org/10.3390/su12156151

Volt, J., Toth, Z., Glicker, J., de Groote, M., Borragán, G., de Regel, S., Dourlens-Quaranta, S., & Carbonari, G. (2020). *Definition of the digital building logbook – Report 1 of the study on the development of a European Union framework for buildings' digital logbook*. https://doi.org/10.2826/480977

Vossen, P. (Ed.). (1998). *Eurowordnet: A multilingual database with lexical semantic networks for european languages*. Kluwer.

Vuarnoz, D., & Jusselme, T. (2018). Temporal variations in the primary energy use and greenhouse gas emissions of electricity provided by the Swiss grid. *Energy*, *161*, 573–582. https://doi.org/10.1016/j.energy.2018.07.087

Wang, N., Issa, R. R. A., & Anumba, C. J. (2022). NLP-Based Query-Answering System for Information Extraction from Building Information Models. *Journal of Computing in Civil Engineering*, *36*(3). https://doi.org/10.1061/(ASCE)CP.1943-5487.0001019

Wang, Z., Bergés, M., & Akinci, B. (2024). Pre-trained language model based method for building information model to building energy model transformation at meta-model level. In V. Gonzalez-Moret, J. Zhang, B. García de Soto, & I. Brilakis (Eds.), *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*. International Association for Automation and Robotics in Construction (IAARC). https://doi.org/10.22260/ISARC2024/0004

Warrier, G. A., Palaniappan, S., & Habert, G. (2024). Classification of sources of uncertainty in building LCA. *Energy and Buildings*, *305*, 113892. https://doi.org/10.1016/j.enbuild.2024.113892

Wastiels, L., & Decuypere, R. (2019). Identification and comparison of LCA-BIM integration strategies. *IOP Conference Series: Earth and Environmental Science*, *323*, 012101. https://doi.org/10.1088/1755-1315/323/1/012101

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). Emergent Abilities of Large Language Models. http://arxiv.org/pdf/2206.07682v2

Wiberg, A. H., Løvhaug, S., Mathisen, M., Tschoerner, B., Resch, E., Erdt, M., & Prasolova-Førland, E. (2019). Visualisation of KPIs in zero emission neighbourhoods for improved stakeholder participation using Virtual Reality. *IOP Conference Series: Earth and Environmental Science*, *323*(1), 012074. https://doi.org/10.1088/1755-1315/323/1/012074

Wiberg, A. H., Wiik, M. K., Auklend, H., Slåke, M. L., Tuncer, Z., Manni, M., Ceci, G., & Hofmeister, T. (2019). Life cycle assessment for Zero Emission Buildings – A chronology of the development of a visual, dynamic and integrated approach. *IOP Conference Series: Earth and Environmental Science*, *352*(1), 012054. https://doi.org/10.1088/1755-1315/352/1/012054

Wilbur, W. J., & Sirotkin, K. (1992). The automatic identification of stop words. *Journal of Information Science*, *18*(1), 45–55. https://doi.org/10.1177/016555159201800106

Wilson, E., Parker, A., Fontanini, A., Present, E., Reyna, J., Adhikari, R., Bianchi, C., CaraDonna, C., Dahlhausen, M., Kim, J., LeBar, A., Liu, L., Praprost, M., Zhang, L., DeWitt, P., Merket, N., Speake, A., Hong, T., Li, H., . . . Li, Q. (2021). End-Use Load Profiles for the U.S. Building Stock. https://doi.org/10.25984/1876417

Wimmer, R., Cao, J., Remmen, P., Maile, T., O'Donnell, J., Frisch, J., Thorade, M., Streblow, R., Müller, D., & van Treeck, C. (2015). Implementation of Advanced Bim-Based Mapping Rules for Automated Conversions to Modelica. *Proceedings of Building Simulation 2015: 14th Conference of IBPSA*. https://doi.org/10.26868/25222708.2015.2424

Wolber, J., Steinbrenner, S., Sievering, C., & Haghsheno, S. (2024). Einsatz der OCR-Technologie in Kombination mit NLP-Algorithmen in der Bauindustrie. In S. Haghsheno, G. Satzger, S. Lauble, & M. Vössing (Eds.), *Künstliche Intelligenz im Bauwesen* (pp. 255–271). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-42796-2_15

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., . . . Rush, A. M. (2019). HuggingFace's Transformers: State-of-the-art Natural Language Processing. https://doi.org/10.48550/arXiv.1910.03771

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. *Automation in Construction*, *134*, 104059. https://doi.org/10.1016/j.autcon.2021.104059

Wu, Z., & Palmer, M. (1994). Verbs semantics and lexical selection. In J. Pustejovsky (Ed.), *Proceedings of the 32nd annual meeting on Association for Computational Linguistics -* (pp. 133–138). Association for Computational Linguistics. https://doi.org/10.3115/981732.981751

Xie, Q., Zhou, X., Wang, J., Gao, X., Chen, X., & Chun, L. (2019). Matching Real-World Facilities to Building Information Modeling Data Using Natural Language Processing. *IEEE Access*, *7*, 119465–119475. https://doi.org/10.1109/ACCESS.2019.2937219

Yang, X., Hu, M., Wu, J., & Zhao, B. (2018). Building-information-modeling enabled life cycle assessment, a case study on carbon footprint accounting for a residential building in China. *Journal of Cleaner Production*, *183*, 729–743. https://doi.org/10.1016/j.jclepro.2018.02.070

Yang, Y., Pan, Y., Zeng, F., Lin, Z., & Li, C. (2022). A gbXML Reconstruction Workflow and Tool Development to Improve the Geometric Interoperability between BIM and BEM. *Buildings*, *12*(2), 221. https://doi.org/10.3390/buildings12020221

Yang, Y., Cer, D., Ahmad, A., Guo, M., Law, J., Constant, N., Abrego, G. H., Yuan, S., Tar, C., Sung, Y.-H., Strope, B., & Kurzweil, R. (2019). Multilingual Universal Sentence Encoder for Semantic Retrieval. https://doi.org/10.48550/arXiv.1907.04307

Zahedi, A., Abualdenien, J., Petzold, F., & Borrmann, A. (2022). BIM-based design decisions documentation using design episodes, explanation tags, and constraints. *Journal of Information Technology in Construction*, *27*, 756–780. https://doi.org/10.36680/j.itcon.2022.037

Zahedi, A., & Petzold, F. (2019). Interaction with analysis and simulation methods via minimized computer-readable BIM-based communication protocol. *Blucher Design Proceedings*, 241–250. https://doi.org/10.5151/proceedings-ecaadesigradi2019_140

Zheng, Z., Lu, X.-Z., Chen, K.-Y., Zhou, Y.-C., & Lin, J.-R. (2022). Pretrained domain-specific language model for natural language processing tasks in the AEC domain. *Computers in Industry*, *142*, 103733. https://doi.org/10.1016/j.compind.2022.103733

Zimmermann, R. K., Bruhn, S., & Birgisdóttir, H. (2021). BIM-Based Life Cycle Assessment of Buildings—An Investigation of Industry Practice and Needs. *Sustainability*, *13*(10), 5455. https://doi.org/10.3390/su13105455

## Appendix A

# Paper I

## A.1 Classification according to DIN 276 cost groups

| Nr. | 2nd Level | Nr. | 3rd Level |
|---|---|---|---|
| 320 | Foundations | | |
| 330 | External walls | | |
| | | 331 | Load-bearing external walls |
| | | 332 | Non-load-bearing external walls |
| | | 333 | External columns |
| | | 334 | External doors and windows |
| | | 335 | External cladding units |
| | | 336 | Internal wall linings (of external walls) |
| | | 337 | Prefabricated facade units |
| 340 | External walls | | |
| | | 341 | Load-bearing interior walls |
| | | 342 | Non-load-bearing interior walls |
| | | 343 | Interior columns |
| | | 344 | Interior doors and windows |
| | | 345 | Interior cladding units |
| | | 346 | Elemental interior wall constructions |
| 350 | External walls | | |
| | | 351 | Ceiling constructions |
| | | 352 | Ceiling openings |
| | | 353 | Ceiling coatings |
| | | 354 | Ceiling claddings |
| | | 355 | Elemental ceiling structures |
| 360 | Roofs | | |

Table A.1: Classification of LCA relevant cost group 300 (Structure - construction works) according to DIN 276 cost groups

# Appendix B

# Paper IV

## B.1 Prototypical implementation

### B.1.1 Implementation of Material option selection

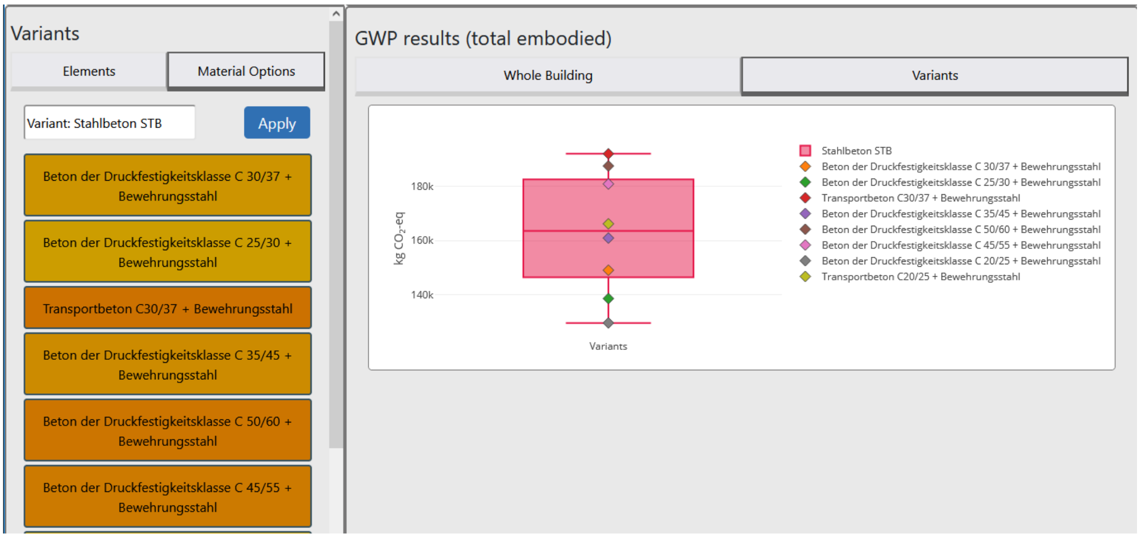



Figure B.1: Prototypical implementation of the material option comparison of reinforced concrete (step 4.c)

### B.1.2 Example BCF implementation

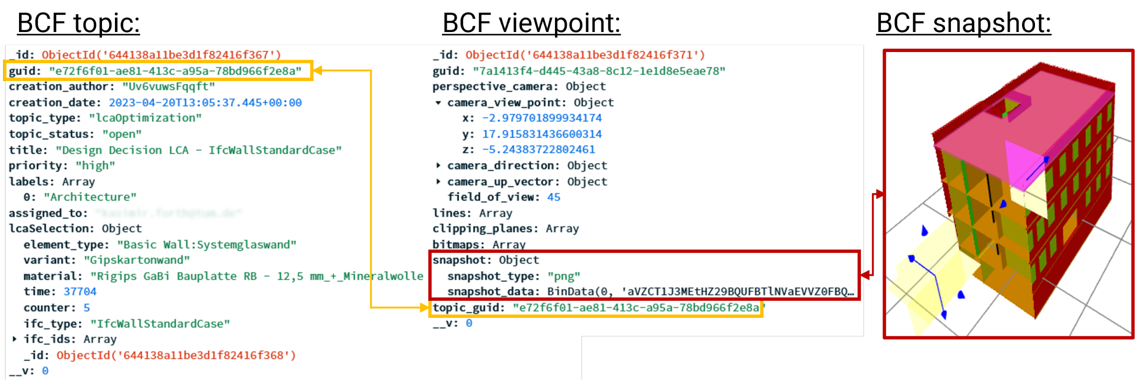



Figure B.2: Screenshot of one example BCF issue implemented using BCF API and extended

## B.2 User study

### B.2.1 Participants

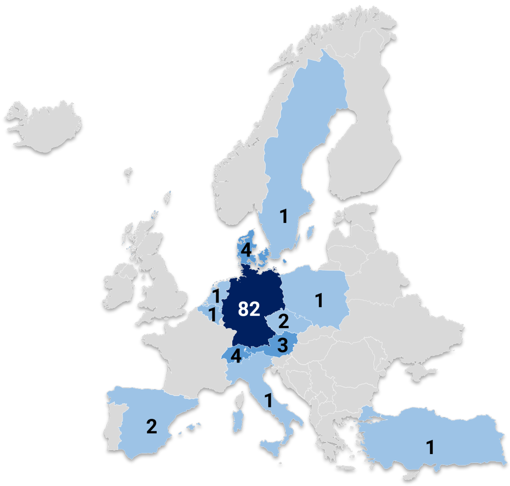



Figure B.3: Distribution of participants' work countries

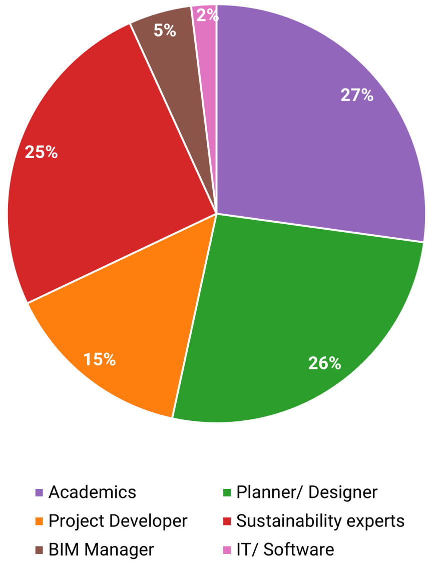



Figure B.4: Distribution of participants' stakeholder background

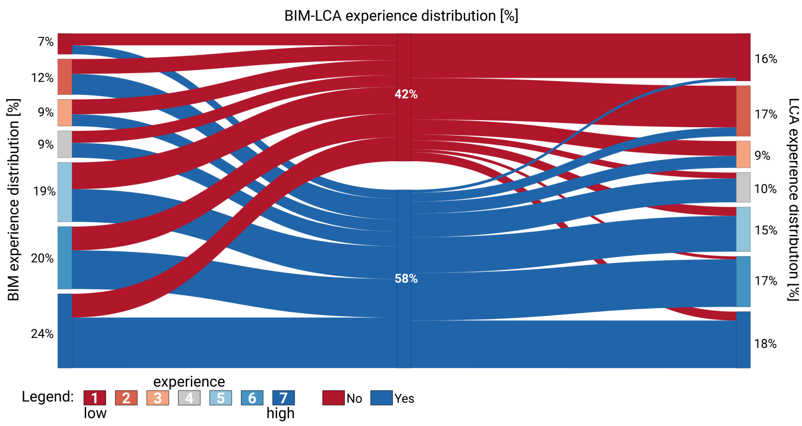



Figure B.5: Profession of the user study participants in relation to their BIM experience (left) and LCA experience (right)

## B.3 Survey questions

| Nr. | Question | Answer options |
|---|---|---|
| 1 | In which country do you work? | Germany, Austria, Switzerland, others |
| 2 | What is your professional background? | Project Developer/ Client/ Housing association, Portfolio Manager/ Investor, Planner (architect, structural engineer, HVAC/ MEP engineer, etc.), Sustainability expert/ Building physicist/ Energy consultant, Student/ researcher |
| 3 | How experienced are you with BIM workflows and models? | 1-7 |
| 4 | How experienced are you with Life Cycle Assessments (LCA) of Buildings? | 1-7 |
| 5 | Have you already gained experience with BIM-based life cycle assessments? | Yes/ No |
| 6 | Who do you think should have significant influence on component and material decision making based on environmental impact? | Project Developer/ Client/ Housing association, Portfolio Manager/ Investor, Planner (architect, structural engineer, HVAC/ MEP engineer, etc.), Sustainability expert/ Building physicist/ Energy consultant |

Table B.1: Survey questions of user study, part 1: general and background questions

| Nr. | Question | Answer options |
|---|---|---|
| 7 | In general: how easy was the task of the whole decision-making process? | 1-7 |
| 8 | How well did the colored 3D model help you identify LCA optimization potential? | 1-7 |
| 9 | How intuitive do you find the transparent display of the colored 3D elements to show the uncertainties of the results? | 1-7 |
| 10 | How well did the coloring of the element variants and material options help you in making decisions? | 1-7 |
| 11 | How intuitive do you find the gradient coloring of the element variants and material options to show the uncertainties of the results? | 1-7 |
| 12 | How well did the box plot diagrams of the LCA results help you in making decisions? | 1-7 |
| 13 | How important do you consider open data standards of the BIM method (e.g. IFC models) for LCA? | 1-7 |
| 14 | Do you know the exchange format BCF - BIM Collaboration Format for BIM-based Issue Management? | Yes/ No |
| 15 | How helpful do you find the automatic creation of digital collaboration protocols (via BCF - BIM Collaboration Format) for communicating the final decisions to the BIM modeler? | 1-7 |
| 16 | Finally: how difficult was the task of the whole decision-making process? | Yes/ No |

Table B.2: Survey questions of user study, part 2: evaluation questions

## B.4 Viewpoints

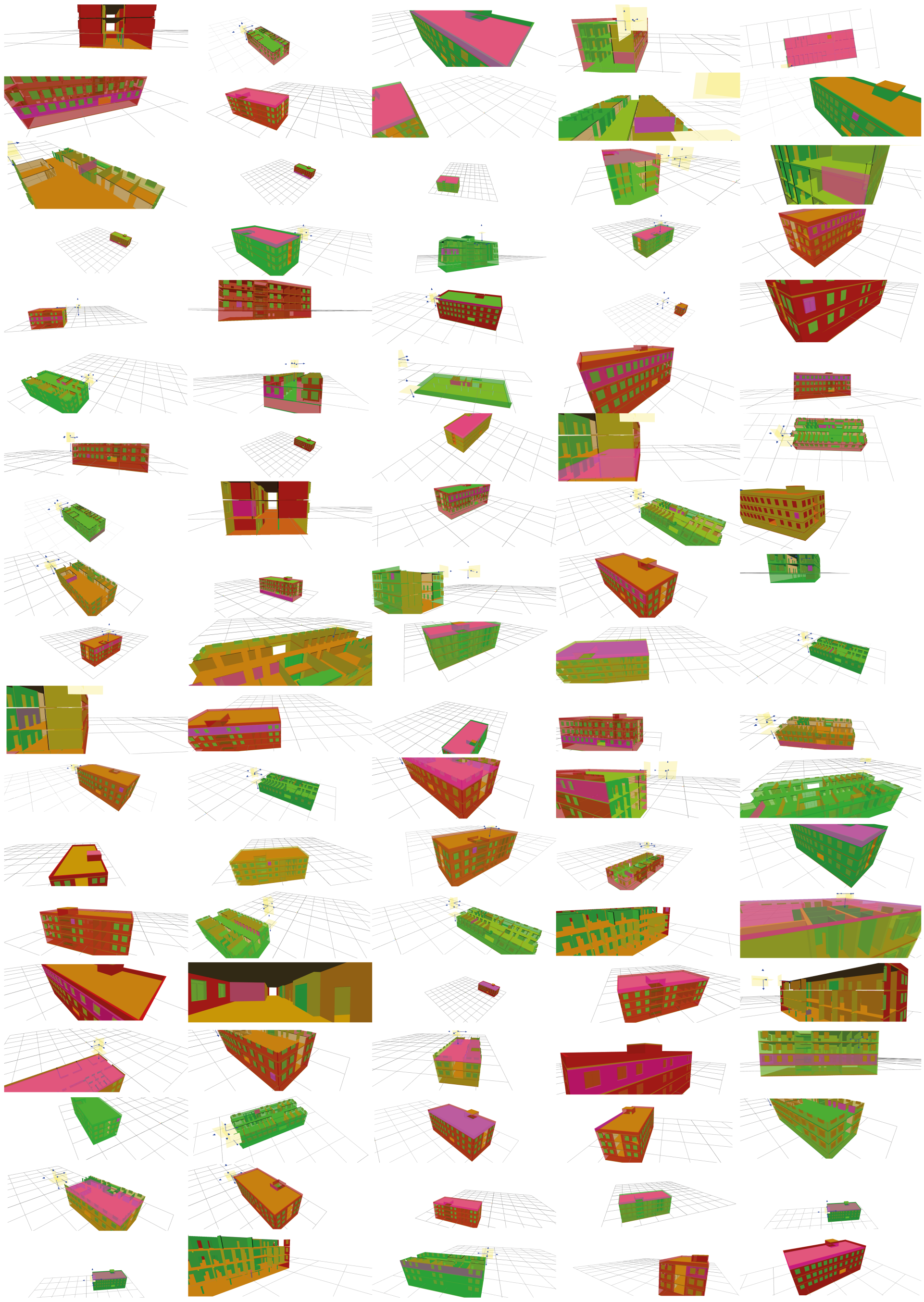


Figure B.6: Viewpoints of most BCF issues about selected variants for decision making (part 1)

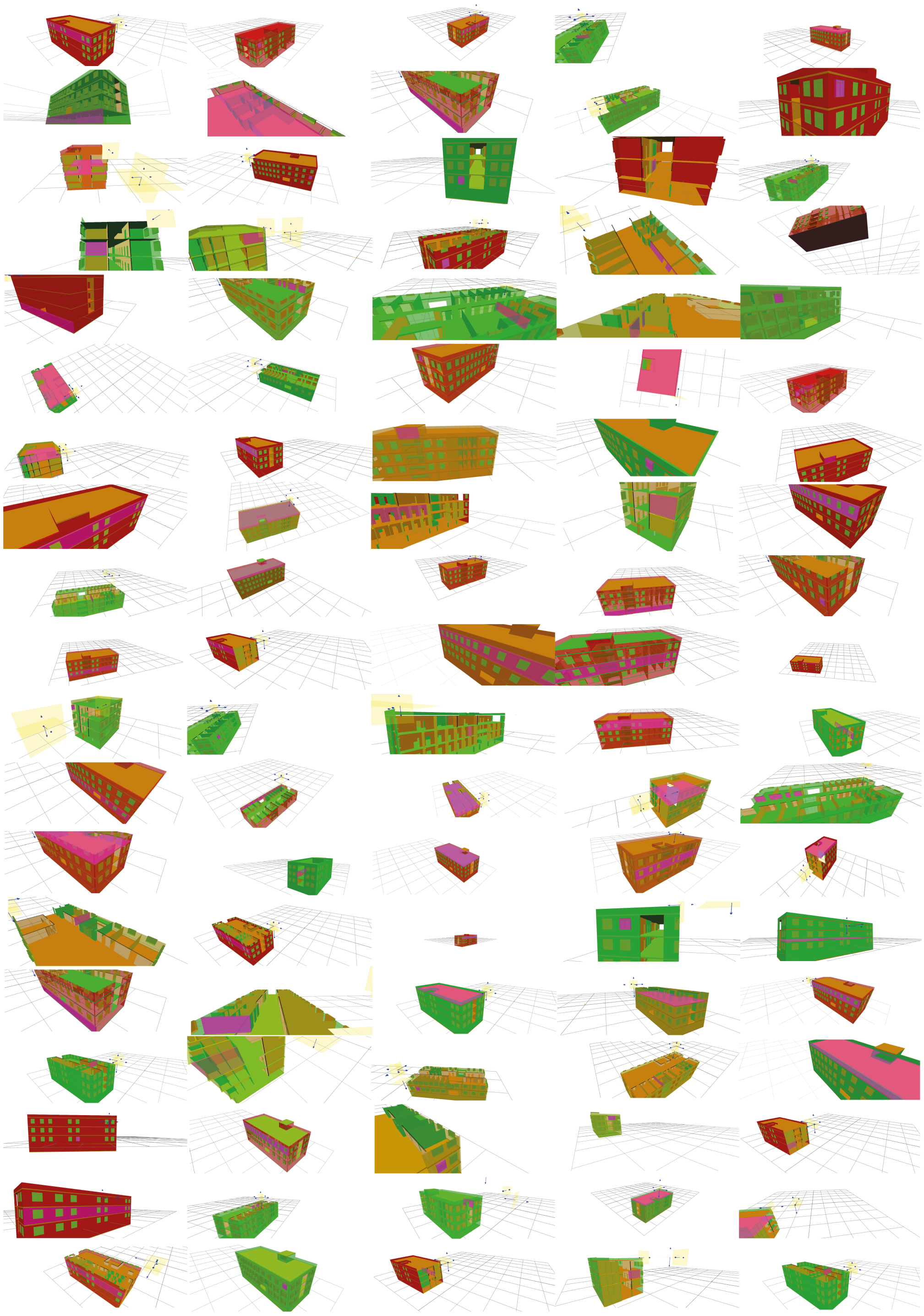


Figure B.7: Viewpoints of most BCF issues about selected variants for decision making (part 2)