The Morphological Echo of Architects Concept for a Conversational Artificial Intelligence to Support Architects during the Early Design Stages

Jessica Bielski[1], Viktor Eisenstadt[2,3], Christoph Langenhan[1] and Burak Mete[1]
[1]Technical University of Munich, Arcisstr. 25, 80333 Munich, GER
[2]German Research Center for Artificial Intelligence, Trippstadter Str. 122, 67663 Kaiserslautern, GER
[3]University of Hildesheim, Universitätsplatz 1, 31141 Hildesheim, GER
E-mail(s): jessica.bielski@tum.de, viktor.eisenstadt@dfki.de, langenhan@tum.de, burak.mete@tum.de

Abstract: *Communication* is considered an essential aspect of a successful architectural design process. Sketching plays an important role as a tool of *Communication*, while supporting the design development. Using Artificial Intelligence (AI) methods for auto-completion, accelerating and improving work, the metis projects pursue a conversational intelligent design assistant suggesting further *design steps*. Similar to human conversations, the system aims to understand the architect to provide comprehensible suggestions with coherent and appropriate explanations in the user's language at the correct time for improved Human-System-Interaction (HSI), User Experience (UX) and ultimately the effectiveness of the application. We plan to investigate explainability (XAI) and its temporal aspects to eventually generate these individually and automatically in real time.

Keywords: XAI, Explainability, Artificial Intelligence, Design Decision Support, Sequencing

1 Introduction

Communication is considered an essential part throughout all design phases of a successful architectural design process, differentiated into external communication (e.g. public participation), communication within the team, and internal communication (architect's inner dialogue) [1], [2]. Hand-drawings inhabit a special role by clearly illustrating ideas, providing tangibility. Sketching supports the designing by instantly depicting the implications of design decisions. The sketch is both tool and language of *Communication*, as well as a catalyst of the design selection process for further design development [3]. Therefore, sketching is an intuitive interaction method for Computer-Aided

Architectural Design (CAAD) naturally integrating into the design process [4]. Derived from Artificial Intelligence (AI) methods of auto-completion to solve tasks faster and more efficiently, the metis projects pursue the development of an intelligent design assistant. It suggests further *design steps* for the floor plan design during the early design stages, inspired by the 'Architecture Machine' of Negroponte [5]. The system aims to enter the architect's inner dialogue as a conversational partner by suggesting possible solutions. Similar to human conversations, the system needs to understand the user to provide comprehensible suggestions with complementary explanations in the same language at the appropriate time. These explanations are essential for enabling mutual understanding, promoting Human-System-Interaction (HSI), Human-Data-Interaction (HDI), User Experience (UX) and ultimately the effectiveness of the application through explainability as an eXplainable Artificial Intelligence (XAI).

In the following, we will present our approach for determining the appropriate timing of each design step accompanied with explanations and their respective visualisations, by deducing *when which information* is suitable to support the user's understanding of the individual suggestion. Thus, we will describe the current state of research for both current technology and literature, describe three phases of our approach, based on the previously developed and implemented methods within the metis projects accompanied by an extensive further literature review, and finally, discuss possible obstacles and other applications within the research projects.

2 Current State of Research

2.1 Artificial Intelligence and Explainability

Within the past few decades, AI methods have been applied in various capacities and varieties in different industries, research and even daily life (e.g. weather forecast). However, their use within the architecture, engineering and construction (AEC) industry is rather rare. Following the adoption of Case-Based Reasoning (CBR) as Case Based Design (CBD) in the 1990s, contemporary AI approaches focus on optimisation during later stages of the construction process from Building Information Modelling (BIM) information formalising the final building, e.g. financial, temporal and performance [6]. None of these approaches deal with the complex designs of the early design stages, which are often vague and incomplete because of a non-standardised design process, ill-defined design problems and hand-drawn sketching as a best practice for ideation. Nevertheless, the impact of the design decisions of the early stages are significant on the overall design and sustainability of the construction process and building itself. Vice versa, the later the stages fatal flaws and issues are detected, the more difficult it becomes to remedy these shortcomings and change the design [3], [7]. Thus, the early design stages are a good time for supporting the architectural design decision making to enable faster and more efficient designing while increasing architectural quality [7].

While the decisions of the first AI systems were easily understandable for the human user, much more complex contemporary models like Deep Learning (DL), including Recurrent Neural Networks (RNNs), resulted in the lost connection of the comprehension between the system's output and the internal system operations for the output – the *black box* problem. To remedy these shortcomings, *transparency* is used to externalise the internal mechanisms in a comprehensible way for the human user [8], based on principles of HCI, adopted for HSI and HDI. Current RNN-related XAI research can be divided into explainability based on the data learned by the RNN models (e.g. feature relevance methods) or based on the modifications of the RNN architecture outputting insight on the decision (e.g. local explanations) [8]. Both types use post-hoc explainability by re-formatting RNN model data into human interpretable information through different techniques (e.g. visualisations, text, explanation by example). These explanations can be generated by *hybrid transparent* and *black box* methods for creating transparent design methods, which are enhanced in different ways (e.g. constituting data, replacing output) with more interpretable models (e.g. CBR) to create explainable outputs for the user.

Wang, Yang, Abdul, *et al.* [9] shift the perspective of XAI from the system implementation to the theoretical necessities of explainability for the user, based on User-Centred Design (UCD). They propose a conceptual framework and adaptation steps for theory-driven designing of explainability for a specific target group (see Figure 1). The framework is based on the human dual process model of reasoning, where explanation goals, and inherent inquiry and reasoning, as well as possible errors and weaknesses of heuristic System 1 and analytical System 2, inform XAI strategies to support the user in the decision making process. Wang, Yang, Abdul, *et al.* [9] describe the following steps for adapting the framework, tested in a case study for medical diagnoses: clarify the user's reasoning goals and possible biases for the respective applications through a literature review, ethnography, and/or participatory design, deduce appropriate explanation ways for reaching the reasoning goals and/or mitigating cognitive biases using the pathways of the framework, and integrate the explainable intelligent system facilities to create an explainable user interface.

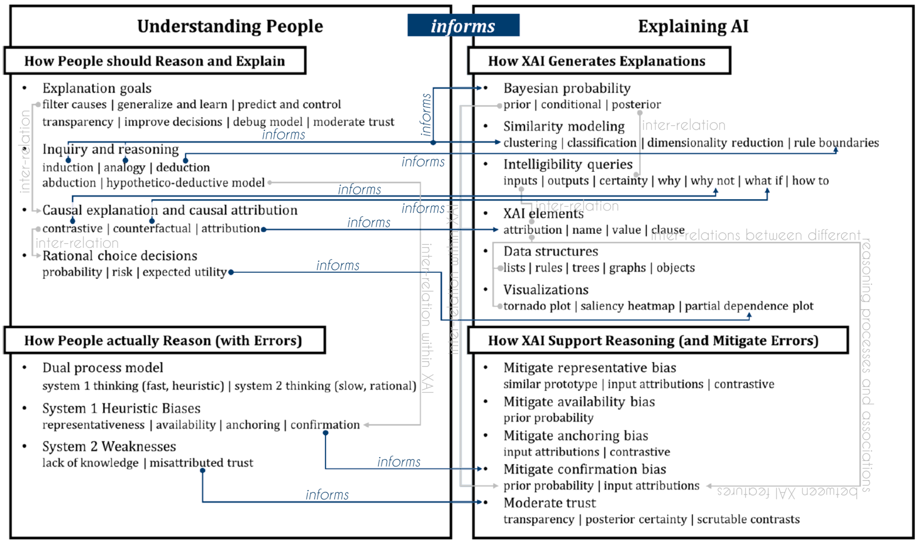

Figure 1: Framework for human reasoning informing XAI techniques [Adapted from 9, p. 4]

2.2 Design Process Protocolling and Sequencing

Design protocolling is a concept for tracking the architectural design process, which can be applied to the sketch process of architects and designers [1], [2]. It is based on the premise of mental design activities being evident through sketching and consequently, sketch protocol studies. Lawson [10] summarises two methods of segmenting the design process: *temporal* and *relational*. He recommends *relational* protocolling by clustering connected sketching activities in a sensible way. Design protocol studies are commonly divided into 'Thinking aloud' studies and retrospective (reporting) studies. Even though both protocol study types suffer from an artificial setting within a limited work environment, Lawson [10] and Suwa and Tversky [11] identify the retrospective protocol study as superior by providing a more genuine and less distorted design and sketching process. Further, the latter [11] suggest video recording the sketching to support the participant during the retrospective reporting, providing visual cues to avoid selective recall and post-hoc realisation. Sketch protocol studies results are interpreted, whereat only qualitative data is available through approaches of established literature known to the authors. Even though every design process is truly unique because of the problems, tools, stakeholders, and designers, Lawson [10] sees the potential for receiving reproducible results through sketch protocol studies in the manner of the natural science paradigm.

Different attempts at sequencing the architectural design process have been made. The nine service phases by the German 'Honorarordnung für Architekten und Ingenieure' (HOAI; eng.: Fee Structure for Architects and Engineers), and its equivalent of seven phases by the American Institute of Architects (AIA) describe the period from client brief to constructed building, which Lawson [10] calls the 'primary design problem'. However, he defines the *design problem* itself simply as a problem solved through designing. Thus, the architectural design process consists of a multitude of individual design problems to be adhered to and solved [10]. Consequently, Lawson [10] introduces the following order-less 'design activities' for solving a singular *design problem* with an overarching *Communication* for the process: *Analysis*, *Synthesis* and *Evaluation*, the *ASE* model. Laseau [2] further differentiates the *Synthesis* into *Exploration* and *Discovery*. Barelkowski [12] uses knowledge management to further separate the *Analysis* into *Knowing* and *Understanding*, resulting in the *Evaluation* as the final selection, as well as a tool for creating more information in the form of *Evaluation - (informing) Knowing*. Further, Lawson [10] segments the design process through 'design events' in sketches, whereas he deems the 'intention behind' more important to identify the mental activity. Thus, we formulate the following three *relationally* clustering categories for temporal sequences of the architectural design process: *design steps* as the finest clustering method, followed by *design intentions*, and further *design phases*.

3 Approach

In order to truly design and implement a conversational intelligent design assistant that is able to assimilate to the architect's inherent idiosyncrasies, it is necessary for the system to distinguish what *design steps* the architects takes with what *design intention*. The system aims to enter the internal dialogue of the architect and their sketch to suggest appropriate further *design steps*, as well as predict possible biases and weaknesses during the architect's design decision making process. Through an interdisciplinary approach at the interface of AI and architecture, we utilise methods of computer science, design theory and Human-Computer-Interaction (HCI) to ultimately be able to predict and thus, suggest further *design steps* through creating a closed AI pipeline, which constantly re-informs and improves itself [13]. In order to design and implement timed explainability into the existing DL pipeline of the metis project, we formulate the following research questions for architects designing through hand-drawn sketching during the early design stages:

1. When is what kind of information relevant for the user to make design decisions?
2. When are what explanation techniques needed to answer to the relevant information? When are what explanation visualisations necessary to be dominantly visible?
3. How can we implement our explainability methods, including the newly explored temporal aspect, into our existing DL pipeline?

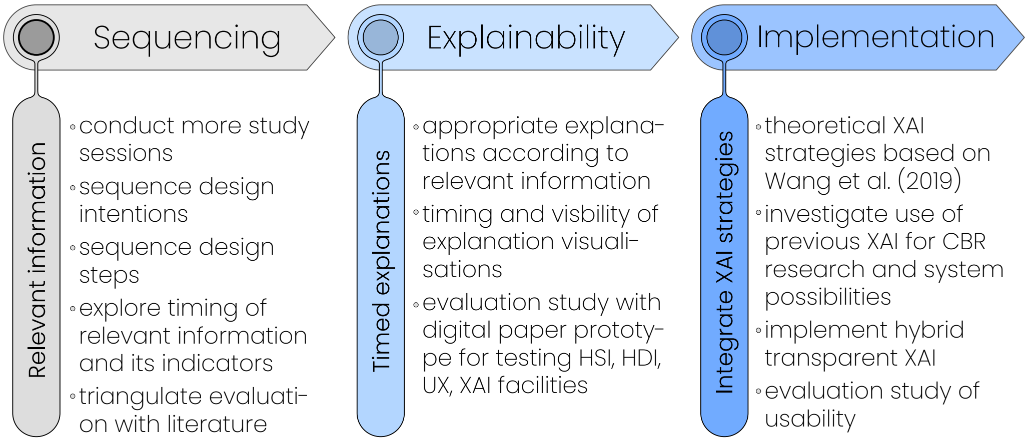

Figure 2: Approach divided by phases answering the three research questions and work steps.

We propose the following approach for answering these research questions on the backdrop of the already conducted research, incorporating their findings, results and evaluations, as illustrated in Figure 2 with three phases: Sequencing, Explainability, and Implementation. We use our successful prediction approach for sequential learning of *design phases* of a refined *ASE* model with six design phases (see Section 2.2) with an overarching *Communication*, based on sketch protocol data [14], quantified through manually assigning sequences within our open-source tool [15], within an Long Short-Term Memory (LSTM) model [16], for further investigating the segmentation of the design process (i.e. *design intention*, followed by *design steps*) by testing the appropriation and robustness of our current LSTM model. Thus, the sketch protocol study needs to be extended with new sequencing labels, while more sessions are continuously conducted to gain more training data. Using this information, we aim to explore *when what information* is relevant to the architect through *design intentions* and how these requests of different information is visible through the sketch and sketching process as *design steps*. We aim to further triangulate, derived from social sciences, these findings through a further literature research, for example Darke [17], Schön and Wiggins [18] and Rezaei [19].

Afterwards, we explore the appropriate explanations for the previously determined relevant information considering their timing, based on the framework of Wang, Yang, Abdul, *et al.* [9] (see Figure1 in Subsection 2.1) and their recommended steps for adaptation. Thus, we explore the temporal aspect of the individual explanations according to their request by the architect. Based on the necessary explanation techniques according to the architectural design process, we determine *when what explanations* need to be visible to the architect through what representation, e.g. *explanation visualisation* and/or *text*. This includes *when what kind* of interaction methods need to be presented to the user to provide *scrutinising methods* for correcting or improving feature values, and ultimately the suggestions. Through *transparency and scrutability* of the system, we plan to extend the conversational properties of an intelligent design assistant, entering the architect's internal dialogue with the sketch, by providing options for feedback to the system. In order to evaluate the findings in terms of HSI, HDI, UX and XAI, we plan to conduct a case study utilising a digital paper prototype with working architects, based on these findings [14]. Implemented as a high-fidelity clickdummy, incorporating timed explainability, we are enabled to observe the interaction and UX of the study participants.

Even though no XAI facilities are implemented in the latest AI system of the metis projects, explanation modules have been integrated in previous CBR models for retrieval of spatial configurations during the early design stages, based on the 'semantic fingerprint' [20] as a search pattern. They generate textual explanations using Natural Language Generation (NLG) [21] through case-based agents, explanation ontology, and explanation patterns as its foundational components for *transparency*, *traceability*, *relevance* and *justification* After exploring the theoretical XAI strategies proposed by Wang, Yang, Abdul, *et al.* [9] for implementing the explanation strategies to either support the design decision making process or to mitigate possible biases of architects, we investigate the possibilities and mechanisms of our system in order to determine *how we can* implement the theoretical and previous findings combined with our successful explainability for CBR. Thus, we aim to create a *hybrid transparent* system capable of producing the explanations interpretable for the user [8]. Finally, the explainability under the backdrop of its temporal aspect will be implemented into our existing DL pipeline, which will be evaluated through a case study by domain experts - architects.

4 Discussion

Creating explainability for an intelligent design assistant supporting architects through assimilating to the inherent idiosyncrasies, bears great potential to provide actual conversational properties. Based on previous experiences, we will need to take appropriate measures to be able to react to the expected limited data volume produced from the different studies of the proposed approach. Nevertheless, if the process and the implementation of our approach can be validated, it creates potential for formalising a transferable framework for explainability with temporal aspects. Further, the findings of the timing of explainability can be re-used within the metis projects for determining phases of deep concentration to minimise distraction on the user interface on a whole for the truly rewarding *Discovery* of ideas and solutions, maximising the feeling of ownership over the design and satisfaction. Meanwhile, the feedback acquired through the *scrutability* features of the system will provide us with valuable information to improve the intelligent design assistant. Ultimately, the results and findings of the proposed research are promising for creating an intelligent design assistant, which is in a constant conversation with the architect to externalise the internal dialogue for supporting the design decision making process using the silent language of architects, sketches.

Acknowledgements

We want to thank the DFG for funding the metis projects and the participants of the mentioned studies.

References

[1] B. Lawson, *How Designers Think, Fourth Edition: The Design Process Demystified*, 4th. Taylor & Francis Ltd, Jan. 2005.

[2] P. Laseau, *Graphic thinking for architects and designers*, 3rd. Wiley, 2000.

[3] B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Design*, 1st. Morgan Kaufmann, 2007.

[4] P. Leclercq and R. Juchmes, “The absent interface in design engineering”, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 16, pp. 219–227, 2002.

[5] N. Negroponte, *The Architecture Machine: Toward a More Human Environment*. The MIT Press, Jan. 1973.

[6] S. O. Abioye, L. O. Oyedele, L. Akanbi, *et al.*, “Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges”, *Journal of Building Engineering*, vol. 44, p. 103 299, 2021.

[7] T. Harputlugil, A. T. Gultekin, M. M. Prins, and Y. I. Topcu, “Architectural design quality assessment based on analytic hierarchy process: A case study”, *Metu Journal of The Faculty of Architecture*, vol. 31, pp. 139–161, 2014. DOI: 10.4305.

[8] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, *et al.*, “Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai”, *Information Fusion*, vol. 58, pp. 82–115, 2020.

[9] D. Wang, Q. Yang, A. Abdul, and B. Y. Lim, “Designing theory-driven user-centric explainable ai”, ser. CHI ’19, Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–15.

[10] B. Lawson, *What Designers Know*, 1st. Oxford: Elsevier/Architectural Press, 2004.

[11] M. Suwa and B. Tversky, “What do architects and students perceive in their design sketches? a protocol analysis”, *Design Studies*, vol. 18, no. 4, pp. 385–403, 1997.

[12] R. Barelkowski, “Learning design from knowing less. on approaches to architectural design process”, in *Knowing (by) Designing*, J. Verbeke and B. Pak, editors, 1st. LUCA Sint Lucas School of Architecture Ghent-Brussels KU Leuven Faculty of Architecture, Jan. 2013, ch. 8.1, pp. 517–526.

[13] J. Bielski, B. Mete, V. Eisenstadt, *et al.*, “Chasing the white rabbit – a case study of predicting design phases of architects by training a deep neural network with sketch recognition through a digital drawing board”, in *ICCC*, [Preprint], Pinhal de Marrocos: Association for Computational Creativity, Jun. 2022.

[14] J. Bielski, C. Langenhan, C. Ziegler, V. Eisenstadt, A. Dengel, and K.-D. Althoff, “Quantifying the intangible - a tool for retrospective protocol studies of sketching during the early conceptual de-sign of architecture”, in *CAADRIA*, vol. 1, Hong Kong: Association for Computer-Aided Architectural Design Research in Asia, Apr. 2022, pp. 403–411.

[15] C. Ziegler, *Sketch protocol analyser*, version 1.2, [Computer Software], Sep. 2021. [Online]. Available: https://github.com/KSD-research-group/sketch-protocol-analyser.

[16] B. Mete, J. Bielski, V. Eisenstadt, C. Langenhan, F. Petzold, and K.-D. Althoff, “Predicting semantic building information (bim) with recurrent neural networks”, in *ECPPM*, [Accepted], Trondheim, Norway: European Association of Product and Process Modelling, Sep. 2022.

[17] J. Darke, “The primary generator and the design process”, *Design Studies*, vol. 1, no. 1, pp. 36–44, 1979.

[18] D. A. Schön and G. E. Wiggins, “Kinds of seeing and their functions in designing”, *Design Studies*, vol. 13, pp. 135–156, 1992.

[19] M. Rezaei, *Reviewing Design Process Theories: Discourses in Architecture, Urban Design and Planning Theories* (SpringerBriefs in Architectural Design and Technology). Springer, 2020.

[20] C. Langenhan, “Datenmanagement in der architektur”, Dissertation, Technical University of Munich, Munich, 2017.

[21] V. Eisenstadt, C. Espinoza-Stapelfeld, A. Mikyas, and K.-D. Althoff, “Explainable distributed case-based support systems: Patterns for enhancement and validation of design recommendations”, in *ICCBR*, M. T. Cox, P. Funk, and S. Begum, editors, Cham: Springer International Publishing, 2018, pp. 78–94.