

# A Preliminary Survey of Explanation Facilities of AI-Based Design Support Approaches and Tools

Viktor Eisenstadt<sup>1</sup>, Klaus-Dieter Althoff<sup>1,2</sup>

<sup>1</sup>University of Hildesheim, Institute of Computer Science  
Samelsonplatz 1, 31141 Hildesheim, Germany  
viktor.eisenstadt@uni-hildesheim.de

<sup>2</sup>German Research Center for Artificial Intelligence (DFKI)  
Trippstadter Strasse 122, 67663 Kaiserslautern, Germany  
klaus-dieter.althoff@dfki.de

**Abstract** This work provides a first version of a survey of explanation facilities implemented in AI-based design support software. The goal of the survey is to provide a comprehensive analysis that examines the existing approaches, prototypes, and tools, that aim at support of conceptual or concrete design phases in domains such as architecture or CAD and possess a certain explainability feature or component. The requirement for the approaches to be included in this survey is that they should be mainly based on AI or AI-based techniques, such as artificial neural networks, case-based reasoning, multi-agent systems etc. Each of the included approaches will be shortly described, with an emphasis on its explanation facility which then will be checked for certain criteria. In this paper, only a short selection of the approaches is presented to provide a first look at it and to build a foundation for the discussion of possible suggestions, additions, or improvements during the workshop.

**Keywords:** knowledge-based design, AI and creativity, explanation, survey, computer-aided design, computational design, case-based reasoning, XAI

## 1 Introduction

Explainability of AI-based systems (XAI) is currently one of the most emerging topics in the research field of artificial intelligence, being present from the early days of this research area. Its aim is to improve the relationship between the user and the AI-based system by explaining the system's functionality, e.g., the decisions made by the implemented algorithm. Main criteria for rating of an XAI component are the understandability of produced explanations and the added value of the component itself. In the last years, many initiatives were started to examine the research topic of XAI further, the most well-known examples are the workshops, such the XAI workshop at IJCAI [1], ExACT series of workshops [12], or the Workshop on Explainable Smart Systems [8].

The idea of support of design process with AI methods is not a novelty as well, many approaches were developed that had an (explicit) explanation facility implemented. Some of them were conceptualized with an explanation facility as an integral part of the system, others were extended in the way that such a component was added at a later stage of the development or conceptualization. The applications domains, for which these approaches were developed, had a wide range and varied from architecture, product design, engineering design to the design of control systems. The approaches were also conceptualized for different phases of the design process, e.g.:

- An *early conceptual phase* where the ideas of the product or building are vague and only abstract information about the final shape is available.
- A later, more *concrete design phase*, where exact measurements and material information are already collected and can be used for further development or improvement of the design.

The main aim of this survey is to examine the explanation facilities of approaches, prototypes, and tools that contain an explanation feature or a related functionality, in order to provide an overview of the past, current and probably future directions of XAI in AI-based design. Our goal is to determine problems and emerging challenges of combination of these two research directions and to provide suggestions for their improvement. We also think that such an overview can be useful for the future explanation components of the design support tools that are knowledge-based and built upon an AI technique. In this paper however, we provide only a preliminary version of this survey, as our current primary goal is to collect opinions and suggestions for our intent, and adopt and/or adapt them in order to improve this survey.

The rest of this paper is structured as follows: in Section 2 we give a short description of the planned survey, the approaches selected for this preliminary survey will then be presented in Section 3, future work, e.g., which approaches are planned to be included and presented in the final survey, is included in Section 4.

## 2 Survey Structure

To provide a proper structure for our survey, we took inspiration from many other published surveys. The emphasis however, was on surveys that were published for the research fields of either AI or computational design, e.g., CAD (computer-aided design), CAAD (architecture), construction, control systems design, recommender systems, case-based reasoning etc. The most related works to date are the overviews of case-based approaches developed for architecture [11], the overview of cloud-based CAD services [14], or the survey of explanation facilities in recommender systems [13]. The main structure of our survey is as follows: as AI+design combination had many periods of development we decided to separate the approaches generally by the decade criteria (1980s, 1990s, 2000s, and 2010s). The approaches of each decade will be briefly described and their explanation facility will be presented more detailed. A conclusion sentence will summarize the presented system’s functionality and the explanation component.

### 3 Approaches with XAI Features

In this section, we present a selection of AI+design-related approaches for this preliminary version of the survey. As mentioned in the previous section, we separate the approaches by the decade criteria. In this section, for each decade, one approach will be summarized and its explanation facility will be evaluated with criteria of *added value* and *understandability* of explanations (verifications, transparency reports) etc. produced.

#### 3.1 Argo

Argo [6] is a system for solving of design problems by means of applying analogical reasoning and was developed at the Microelectronics and Computer Technology Corporation in Austin, TX, USA around 1988. The exemplary application domain of the system is VLSI (very-large-scale integration) circuit design. Argo consists of modules for “acquisition, storage, retrieval, evaluation, and application of previous experience” [6]. The system works with so-called *design plans* interpreted as problem-solving plans and represented by directed acyclic rule-dependency graphs, where the graphs can have different levels of abstract representations by applying graph editing to the original graph.

**Explanation Feature** Argo applies macrorules for creation of design graphs, the macrorules for a single graph are a result of *explanation-based scheme calculation*. As stated in [6]: “macrorules are built by regressing through the component rules of the plan using a variant of explanation-based generalization”. Explanation-based generalization [9] is an approach for generalization of training examples in machine learning using explanation-based learning [5], an approach for supervised learning by explaining the results of the training (e.g., what went wrong, what was correct). These rules then solve as analogy for solution of a new problem, i.e., they are used for retrieval as well as for learning from past cases.

**Conclusion** Typically for the AI approaches of this decade, Argo is based mostly on complex rule-based operations, explanation-based generation of rules is an integral system feature and adds to the overall traceability of the decisions.

#### 3.2 KOALA

KOALA [10] is a distributed, i.e., autonomous agent-based system for decision support in very early conceptual design phases in its application domain of architecture. It was developed around 1996 at Collaborative Agent Design (CAD) Research Center, CalPoly, San Luis Obispo, USA. The system consists of agents of different types that work collaboratively on solving of a design task, each of them has its own task domain: designer agent (reflects the human designer’s intentions), domain agents (represent knowledge of specific architectural domains), space agents (represent knowledge about spatial parts of the design, such as rooms), and monitor agents (detect and resolve conflicts between other agent types).

**Explanation Feature** Despite the fact that it seems like that the explanation component was eventually not implemented in KOALA, it was described in detail in the corresponding paper [10]. Based on this description, the explanation feature of KOALA was aimed at providing the agents' inference traces that lead to the given decision (how-question), answering general domain knowledge questions (what-questions), and describing goals the agent would aim in the current situation (why-questions).

**Conclusion** The KOALA system was a contemporary distributed design support system whose functionality and traceability of decisions could be probably improved even more by implementing an explanation facility described in [10].

### 3.3 WebCADET

WebCADET [4] is a web-based distributed system for design support that relies on the "AI as text" paradigm and the client-server-based software architecture. The system was developed around 2000 at the University of Cambridge, the Middlesex University, and the Napier University. It is the continuation of the previously developed CADET system. Its application domain is product design and the system is aimed at helping to improve the collaborative design process. Its mode of operation makes use of text-based descriptions that can contain conditions, history, or keywords that were inferred or recorded for the particular product design. These texts are saved in a special knowledge base and used for "design guidance, knowledge viewing, and knowledge capture" [4].

**Explanation Feature** The explanation feature of WebCADET is activated in the design guidance mode, which is comparable to the retrieval phase of other approaches, however, in WebCADET, this takes the form of a manual evaluation of the above named design text rules, where the user first should select the concrete evaluation attribute. The explanation presents justification of the design evaluation process by providing the rule-based inference traces as well as references in form of links to the relevant publications.

**Conclusion** WebCADET was one of the first web-based design support approaches that relied on human-readable linked data that was also embedded in the explanation, which in turn served as a reasonable justification for the rule-based evaluation that looked familiar for web-users, thus providing an intuitive environment for the user interface.

### 3.4 SLASSY

SLASSY [3] is an intelligent system for assistance in engineering design that makes use of its self-learning abilities, and is based on a general structure that consists of different machines that are responsible for execution of different tasks (an explanation machine inter alia is available). The system was developed at the Engineering Design Department of the University of Erlangen-Nürnberg around 2012 for the application domain of sheet-bulk metal forming. The self-learning-based structure of SLASSY allows for intelligent adaptation to the current task

by means of applying a learning algorithm for automated knowledge acquisition. SLASSY differs between primary and secondary design elements, the knowledge for assistance is then generated by means of applying a number of specific *meta-models*. These meta-models are responsible for prediction of the most suitable corresponding design part for the current task.

**Explanation Feature** The explanation component of SLASSY provides information about the predictions made by the meta model. In contrast to other explanation features presented in this paper, the explanation consists mostly of numeric values and is not constructed via, e.g., natural language generation. The information contained does not provide any justification information, however it can be considered as a means to provide more system transparency.

**Conclusion** SLASSY is a system with a modern touch of AI for analysis and synthesis of the possible design parts. Its explanation component however, does not provide understandable expressions for inexperienced users and thus has added value only for experienced ones (although this might not be a problem considering the very narrow application domain).

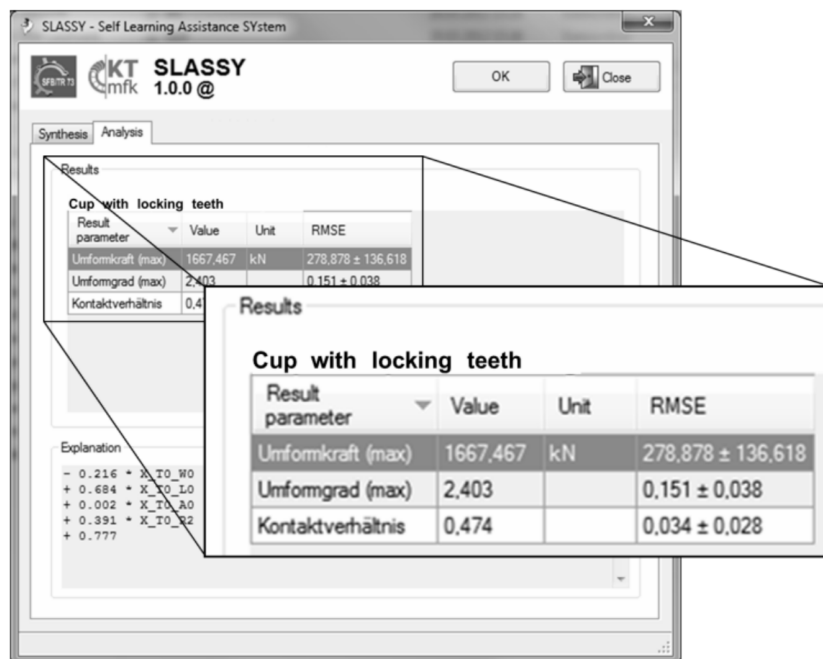


Figure 1. An exemplary explanation in SLASSY (bottom part). Figure from [3].

## 4 Discussion and Future Work

In this paper, we presented a selection of AI-based design support approaches that contain an explanation facility or a related feature that can provide more

insights into the functionality of the corresponding system, e.g., how it came to the current decision. This version of the paper is intended to be an impulse and background for a discussion about explanation facilities in the past and modern AI approaches for design support. In the planned extended (final) version of this survey more approaches will be analysed and described, some examples are ARCHIE [15], CaseBook [7], or MetisCBR [2]. An overview of possible future directions will also be presented.

## References

1. Aha, D., Darrell, T., Pazzani, M., Reid, D., Sammut, C., Stone, P.: Ijcai-17 workshop on explainable ai (xai). In: IJCAI-17 Workshop on Explainable AI (XAI) (2017)
2. Ayzenshtadt, V., Langenhan, C., Bukhari, S.S., Althoff, K.D., Petzold, F., Dengel, A.: Thinking with containers: A multi-agent retrieval approach for the case-based semantic search of architectural designs. In: Filipe, J., van den Herik, J. (eds.) 8th International Conference on Agents and Artificial Intelligence (ICAART-2016), February 24-26, Rome, Italy. SCITEPRESS (2016)
3. Breitsprecher, T., Wartzack, S.: Architecture and realization of a selflearning engineering assistance system for the use within sheetbulk metal forming. In: DS 71: Proceedings of NordDesign 2012, the 9th NordDesign conference, Aarlborg University, Denmark. 22-24.08. 2012 (2012)
4. Caldwell, N.H., Clarkson, P.J., Rodgers, P.A., Huxor, A.P.: Web-based knowledge management for distributed design. *IEEE Intelligent Systems and their Applications* 15(3), 40–47 (2000)
5. DeJong, G., Mooney, R.: Explanation-based learning: An alternative view. *Machine learning* 1(2), 145–176 (1986)
6. Huhns, M.N., Acosta, R.D.: Argo: A system for design by analogy. In: *Artificial Intelligence Applications, 1988.*, Proceedings of the Fourth Conference on. pp. 146–151. IEEE (1988)
7. Inanc, B.S.: Casebook. an information retrieval system for housing floor plans. In: *The Proceedings of 5th Conference on Computer Aided Architectural Design Research (CAADRRIA)*. pp. 389–398 (2000)
8. Lim, B., Smith, A., Stumpf, S.: Exss 2018: Workshop on explainable smart systems (2018)
9. Mitchell, T.M., Keller, R.M., Kedar-Cabelli, S.T.: Explanation-based generalization: A unifying view. *Machine learning* 1(1), 47–80 (1986)
10. Pohl, K.J., Pohl, J.G.: Koala: An object-agent design system. *Collaborative Agent Design (CAD) Research Center* p. 66 (1996)
11. Richter, K., Heylighen, A., Donath, D.: Looking back to the future – an updated case base of case-based design tools for architecture. *Knowledge Modelling-eCAADe* (2007)
12. Roth-Berghofer, T.: Explanation-aware computing exact 2012. In: *Proceedings of the Seventh International ExaCt workshop*
13. Tintarev, N., Masthoff, J.: A survey of explanations in recommender systems. In: *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*. pp. 801–810. IEEE (2007)
14. Wu, D., Terpenney, J., Schaefer, D.: Digital design and manufacturing on the cloud: A review of software and services. *AI EDAM* 31(1), 104–118 (2017)
15. Zimring, C.M., Pearce, M., Goel, A.K., Kolodner, J.L., Sentosa, L.S., Billington, R.: Case-based decision support: A case study in architectural design (1992)