

# Designing a conversational requirements elicitation system for end-users

## Research Abstract

Tim Rietz

Institute of Information Systems and Marketing (IISM)  
Karlsruhe Institute of Technology (KIT)

Fritz-Erler-Straße 23, 76133 Karlsruhe, Germany  
[https://issd.iism.kit.edu/team\\_rietz.php](https://issd.iism.kit.edu/team_rietz.php)  
[tim.rietz@kit.edu](mailto:tim.rietz@kit.edu)

**Abstract**—[Context] Digital transformation impacts an ever-increasing degree of everyone’s business and private life. It is imperative to incorporate a wide audience of user requirements in the development process to design successful information systems (IS). Hence, requirements elicitation (RE) is increasingly performed by end-users that are novices at contributing requirements to IS development projects. [Objective] We need to develop RE systems that are capable of assisting a wide audience of end-users in communicating their needs and requirements. Prominent methods, such as elicitation interviews, are challenging to apply in such a context, as time and location constraints limit potential audiences. [Research Method] The presented dissertation project utilizes design science research to develop a requirements self-elicitation system, LadderBot. A conversational agent (CA) enables end-users to articulate needs and requirements on the grounds of the laddering method. The CA mimics a human interviewer’s capability to rephrase questions and provide assistance in the process and allows users to converse in their natural language. Furthermore, the tool will assist requirements analysts with the subsequent aggregation and analysis of collected data. [Contribution] The dissertation project makes a practical contribution in the form of a ready-to-use system for wide audience end-user RE and subsequent analysis utilizing laddering as cognitive elicitation technique. A theoretical contribution is provided by developing a design theory for the application of conversational agents for RE, including the laboratory and field evaluation of design principles.

**Index Terms**—End-user, Wide Audience, Requirements Elicitation, Conversational Agent, Design Science, Laddering

### I. INTRODUCTION

Digital transformation has brought a variety of information systems into everyone’s business and private life, altering traditional work processes and society itself [1]. We observe a transformation towards a digital society, stressing the influence of the Internet on many traditional services, which advocates a power shift towards the user [2]. In the face of persistently high failure rates of IS development projects, it is imperative that an increas-

ing number of end-users is involved in RE processes, with varying degrees of technological and methodological expertise [3]. Hence, building tools that enable wide audience elicitation of user requirements is crucial for developing software that meets needs and reduces project failure rates [4]. Consequently, we observe two challenges:

**C1.** Designing and developing a system that can engage a wide audience of individual end-users while overcoming time and location barriers [5].

**C2.** Integrating end-users independently of their previous experiences with contributing requirements to development projects, hence enabling novices to interact with the (end-user friendly) elicitation system [1].

Several tools to tackle separate aspects of mentioned challenges with RE have been proposed over the years, introducing diverse approaches. AnnotatePro allows users to submit requirements by making annotations to their screens. Similarly, Moore and Shipman presented a Graphical Requirements Collector, attempting to enable end-users to provide requirements remotely [5]. Given the common problems with requirements quality, such as completeness and ambiguity, exploring natural-language (NL) based elicitation systems gained traction [6]. Pérez and Valderas combine visualization-based RE with NL to reduce ambiguity and inconsistency in end-user RE [7]. Derrick et al. evaluated an embodied conversational agent to facilitate a group workshop that used prompts to guide and assist during user story formulation [8]. However, these tools do not suffice in providing a solution to both challenges introduced: Annotation-based tools primarily enable RE for iteratively improving existing systems; NL tools commonly require a requirements engineer to facilitate the process, hence retaining a bottleneck for wide audience integration [7]; additionally, existing research rarely considers (methodological) guidance for novice end-users. Tools such as FAME [9] and ASSERT [10] cater to novices, but only on the side of a novice analyst, not novice end-users, hence neglecting

the self-elicitation of requirements. A literature gap remains in extending RE techniques to wide audiences [11].

Interviews have been used most commonly for RE and proved to be one of the most effective methods [12]. Guidance and assistance are necessary to enable the self-elicitation of information and requirements from novice end-users on top of a technological foundation that allows for scalability [13]. I utilize an NL-based conversational agent (CA) in the form of a chatbot to mimic a human interviewer’s capability to guide an elicitation interview [8]. Chatbots allow us to include a wide audience of end-users, independent of personal, time, or location restrictions.

The chatbot needs to adhere to a predefined interviewing technique to be capable of conducting an elicitation interview. Cognitive techniques for knowledge acquisition provide a natural way for end-user RE and enable to generate insights into the causes of requirements variability and ambiguity during analysis [14]. Especially the laddering interview is considered an advantageous technique for eliciting relevant information for articulating requirements [12]. Laddering produces comprehensive and structured insights due to the method’s hierarchical nature. During laddering, an interviewer would identify a seed attribute as the overarching topic and ask a series of “why...?” questions to uncover and clarify needs and related attitudes [15]. Since laddering interviews require highly trained and experienced interviewers, the availability of suitable interviewers imposes a bottleneck onto elicitation interviews [15]. Through the combination of the laddering interview as cognitive RE technique and a CA as the technological foundation, I envision a RE system capable of tackling both challenges of scalable and end-user friendly RE. Therefore, I propose the following research questions for the Ph.D. project:

**RQ1.** *How can a conversational requirements elicitation system be designed to enable wide audience RE with novice end-users?*

**RQ2.** *How can the scalable evaluation of collected data be enabled through visualization and requirements mining techniques?*

The remainder of this research abstract is as follows. In Section II, I extend the overview of foundations for this work, as well as the state-of-the-art. Section III presents the research methods that are and that will be used, alongside the evaluation strategy. Section IV introduces the current status of the system prototype – “LadderBot” and present technical challenges. In Section V, I conclude the research abstract by providing an outlook and presenting expected contributions.

## II. RELATED WORK

I classify existing tools and systems for RE along four dimensions: *Target user group*, *Scope*, *Level of development*, and *Mode of elicitation*, according to my SLR (in revision). To identify related work for C1 & C2, I classified tools that target end-users and allow for (remote) wide audience RE. Furthermore, I focused on tools that reached a proof-of-concept level of development. Therefore, I only consider systems that have been built and have seen some form of evaluation already. Table 1 provides an overview of matching tools from the RE and IS domain.

I identified five tools that operate in a problem context related to the Ph.D. research project, in two groups: Tools for annotation and feedback & Tools for forum-like communication. FAME and AnnotatePro both intend to collect requirements in the context of an already existing (host) system. Therefore, these

TABLE 1. EXISTING TOOL SUPPORT FOR LARGE-SCALE END-USER RE

<i>FAME</i>	Continuous end-user RE through form-based text or voice feedback and usage monitoring	Recordings and attachments - Usage data	Mental model improvement - Minimal user guidance on how to provide input	[9]
<i>AnnotatePro</i>	Professional tool to annotate software screens for communicating requirements	- User draws annotations on screen - Sharing of annotations via picture without additional information	- Requires an existing software tool - No formalization for analysis	[31]
<i>CRUISE</i>	Involving interested users for gathering, analyzing, validating, prioritizing and negotiating requirements	- Forum-like requirements collection - User provide free-form text input - Discussion of requirements	- No user guidance on how to provide input - Potentially resulting in trivial needs - Varying level of detail - Requires team collaboration - No formalization for analysis	[32]
<i>REfine</i>	Gamified online platform for collaborative RE	- Forum-like requirements collection - User provide free-form text input - User signal agreement or disagreement for requirements	- No user guidance on how to provide input - Potentially resulting in trivial needs - Varying level of detail - Active management required	[17]
<i>Requirements Bazaar</i>	Forum for requirements discussion	- Forum-like requirements collection - User provide free-form text input	- No user guidance on how to provide input - Potentially resulting in trivial needs - Varying level of detail - No formalization for analysis	[33]

**Filtering criteria for presented RE tools:** (*Target user group*) **End-user**; (*Scope*) **Remote**; (*Level of development*) **Proof-of-concept**

*Target user group:* For which user group is this tool developed? *Scope:* Is the tool supporting local or remote collection?

*Level of development:* Has the tool been developed and evaluated (Proof-of-concept), or does it exist on prototype level only?

tools may be used for continuous RE to maintain and incrementally improve existing systems. However, for the development of entirely new systems, disruptive changes in software or arguably also to identify delightful features, feedback or annotations may not provide a high enough level of detail [16]. Furthermore, while FAME provides minimal guidance on how to provide feedback through using a feedback form, AnnotatePro only allows free-form annotations, complicating the subsequent analysis of user input. On the other hand, CRUISE, REfine & Requirements Bazaar collect requirements by providing end-users with a forum-like interface. End-users create discussion topics, including at least a title and descriptions, and subsequently may discuss their requirements (CRUISE, Requirements Bazaar). In the case of REfine, users can signal their agreement or disagreement to proposed requirements. This forum-like method of RE bears many potential downfalls. The lack of (methodological) guidance for novice end-users increases the risk of collecting ‘trivial’ requirements [17]. Additionally, analyzing gathered requirements is complicated by the varying degree of information contained in the individual submissions, and active management of the collection process may still be required (REfine). Hence, to the best of my knowledge, there is no tool available which solves both C1 & C2 to a satisfying degree.

### III. FOUNDATIONS

#### A. The laddering interview technique for RE

Laddering is a cognitive technique with roots in personality psychology that utilizes a structured approach for data-gathering [15]. For RE, cognitive techniques, in comparison to traditional, collaborative, or contextual techniques, are commonly used to acquire knowledge. As such, requirements are not directly communicated but extracted from the structure and content of end-user knowledge based on rich enough information [11]. Herein, cognitive techniques provide the most natural interaction with end-users [11]. Furthermore, cognitive techniques enable the generation of a map of end-user knowledge, including mental representations of a domain through content and relations between attributes.

The laddering technique was introduced as a method to elicit superordinate from subordinate items and clarify the relations between constructs and attributes, with its origin in personal construct theory [14]. A series of probing questions are used to elicit the structure and content of end-user knowledge [11]. In marketing, laddering elicits knowledge in the form of ACV chains, based on the means-end theory [15]: attributes – consequences – values. *Attributes* as the least abstract level describe “concrete, physical, or observable characteristics” of products and software [18]. *Consequences* describe what a product provides a user with, either on the positive (benefits) or negative side (costs). A product can have functional or non-functional, e.g., psychosocial, consequences. *Values* represent a user’s wishes, goals, and needs and are the end state a customer is trying to achieve through, e.g., a purchase. An exemplary ACV chain in a soft-

ware development context has the following form: *Providing default values (A) – No need to fill out data repeatedly (C) – Happiness (V)* [18]. Therein, the ACV structure is related to the critical success chain (CSC) structure in information system science and used for information systems planning. CSCs follow the same structure as ACV chains and link application attributes to critical success factors and finally, to personal goals [19].

The laddering interview itself follows a straightforward structure. Participants are asked why a particular attribute is important to them, using a series of “why...?” questions while navigating through the ACV chains. E.g., an interviewer might ask, “why is starting process X from the landing page important?”. Laddering interviews may also be conducted without a human interviewer, in pencil-and-paper manner using a questionnaire [15]. Content-coding initiates the analysis process of laddering interviews [15]. Requirements analysts manually build common codes for replies of interviewees that can be related. These content codes are then used to construct a summary matrix, visualizing each chain from each participant, showing the included codes per chain. Subsequently, an aggregate implication matrix is formed, containing all direct and indirect relations between attributes, consequences, and values across interviews. Finally, the aggregate implication matrix is visualized as a hierarchy value map, a tree diagram showing either only direct or both direct and indirect relations up to a certain degree of relation (for examples, see [18], [20]).

#### B. Form and Function of Chatbots

The goal of CAs, as McTear puts it, is the “[...] effortless, spontaneous communication with a computer” [21]. Klopfenstein et al. conducted a systematic analysis of one of the instantiations of CAs, chatbots, categorizing the advantages for users and developers [22]. They find *instant availability*, a *gentle learning curve*, and *platform independence* to be among the most prominent benefits. Hence, I argue that chatbots serve as a promising form of CAs for approaching a wide audience of end-users. Instant availability and platform independence enable barrier-free interaction with the system. A gentle learning curve, resulting from an interaction modality that is familiar to novice users, chatting using natural language, creates an effortless experience [22]. Over the years, multiple variants of chatbots have seen use, which we can differentiate according to form and function [23]. The form of a chatbot describes the arrangement of aspects that do not primarily contribute to the utility of the bot (similar to non-functional requirements). For example, anthropomorphism comprises methods for making the appearance and behavior of a bot more human-like. Function describes aspects related to the general performance, such as the bot’s dialogue control strategy. A frame-based bot uses question templates to communicate with its user. These systems do not have pre-determined dialogue flows but adapt to user input, e.g., in the form of a software problem support tool [21].

Despite a renewed research interest in chatbots, due to advances in artificial intelligence [24], the integration of CAs into RE remains sparse. Derrick et al. investigated the effect of a simple scripted agent in facilitating group elicitation sessions with users [8] while other studies developed prototypes for frame-based agents in interview scenarios [25], [26]. While these studies evaluated the general applicability of CAs as facilitators of elicitation processes, to the best of my knowledge, no evaluation of chatbot-based requirements elicitation with a wide audience of end-users has been conducted, comparing the performance of a system with established processes on the basis of measures such as performance and perception [12].

### C. Research Hypotheses

I formulate the following hypotheses building on the proposed research questions and foundations of chatbots and the laddering interview technique:

**H1.** A RE system using a laddering chatbot outperforms established pencil-and-paper laddering interviews in terms of performance and perception.

**H2.** A RE system using a laddering chatbot matches the performance of traditional laddering interviews in terms of performance and perception.

**H3.** Computer-supported analysis of interview results through natural language processing increases results analysis in terms of effectiveness and efficiency.

## IV. RESEARCH METHOD

The dissertation research applies Design Science Research (DSR) to make a theoretical and practical contribution to the fields of RE and IS [27]. Herein, the research takes multiple steps: (1) exploring the state-of-the-art of data collection in end-user RE with a systematic literature review following the methodology proposed by Kitchenham and Charter & Webster and Watson [28], [29]; (2) development of a requirements self-elicitation system for end-users including its evaluation in a laboratory experiment; (3) adaptation of the system with data analysis capabilities and subsequent system evaluation in a field experiment. This research abstract outlines results for (1) in the form of challenges, research questions, and hypotheses as well as an overview of existing tool support for large-scale end-user RE. Furthermore, it introduces preliminary results for (2) by showcasing the LadderBot artifact.

For the evaluation of the system's results, I will rely on established procedures for analyzing the results of laddering interviews [15]. In a laboratory experiment, I compare LadderBot against a standard procedure for (scalable) laddering: pencil-and-paper laddering using a digital questionnaire [20]. I will conduct the study with students from a large university in Germany in an experimental lab designed for running scientific studies. As a laddering case, I recreate the laddering interview conducted by Jung (2014): exploring users' goals for smartphone use. Therein, I hope to be able to not only compare the treatments introduced in the following based on quantitative

measures, but also to evaluate the quality of my results against the findings of the original study. The experiment will apply three treatments: (1) RE through computer-based pencil-and-paper laddering, (2) the laddering questionnaire from the first treatment but enhanced with a adaptive visualization of elicited ladders, and (3) LadderBot. I will compare the treatments of the between-subject experiment regarding performance and system perception. I measure the performance by building an aggregate implication matrix and comparing values for abstractness and centrality. Furthermore, I measure the quantity of elicited consequences/values and direct/indirect links between concepts, as well as time taken. In terms of system perception, I compare treatments using a self-reporting questionnaire, measuring the following constructs: Understandability, learnability, efficiency, effectiveness, and enjoyment, which have previously been instrumentalized to evaluate RE systems [30]. Finally, I will incorporate questions from the Big Five personality test, as well as multiple control questions, to evaluate the influence of experience, age, or gender, amongst others, on the experiment results.

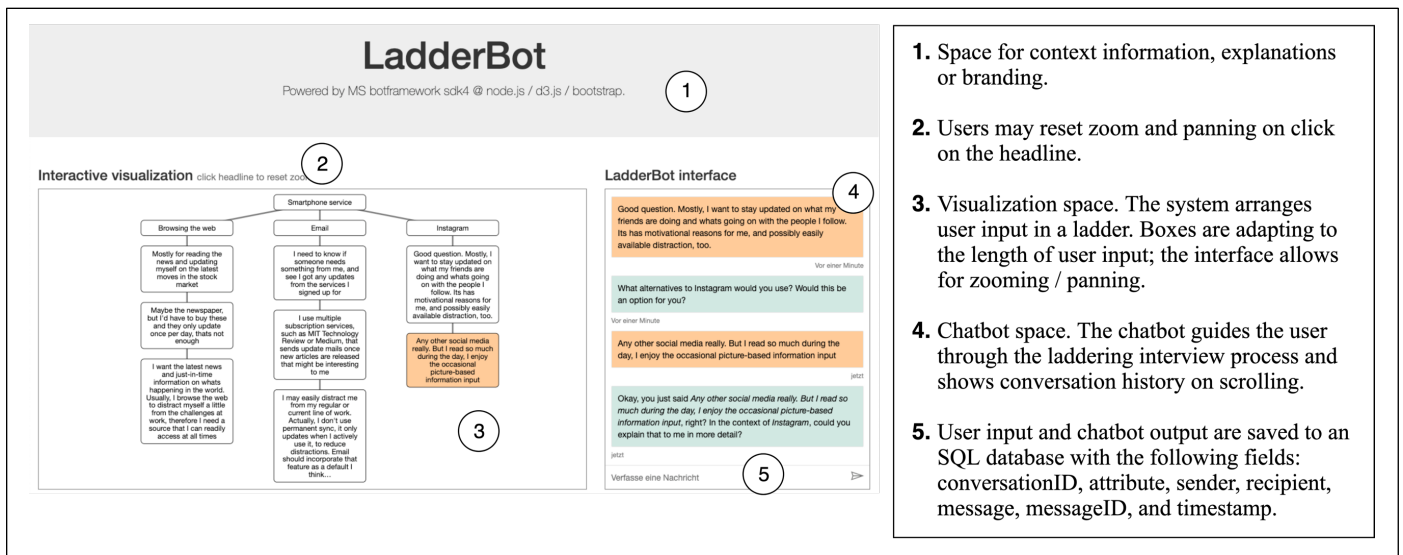
## V. STATUS QUO AND TECHNICAL CHALLENGES

I already took several steps towards answering RQ1, following the general structure of DSR: *Problem Awareness, Suggestion, Development, Evaluation, and Conclusion*, usually in an iterative fashion in multiple cycles. As part of *Problem Awareness*, an SLR has been conducted, analyzing RE techniques and tools with a focus on end-user requirements (in revision). Furthermore, to deepen the understanding of how to design the chatbot aspect of LadderBot for end-users, a study on the impact of form and functional chatbot design features on user acceptance was conducted (published at WI 2019). Finally design principles on how to combine conversational user interfaces with graphical user interfaces were developed, building a foundation for the design of LadderBot (working paper). Currently, I am finishing the development of the system artifact. A detailed description of the artifact, as well as the evaluation strategy, has been accepted to the RE@Next! track of RE 2019.

### A. The LadderBot artifact

I developed an end-user requirements self-elicitation system using the laddering interview technique, LadderBot. The system is currently in a pre-final state, with most of its functionalities working. It is capable of conducting a laddering interview with end-users and visualizing the interview structure in a graphical interface. A screenshot of the interface is shown in fig. 1.

LadderBot is configured to elicit consequences and values for three attributes. To begin the interview, LadderBot asks the user to state the three most frequently used features of a system as seed attributes for each chain. The following process is then repeated until participants constructed three chains. At the beginning of each chain, LadderBot asks an initial question to elicit the first consequence for the current attribute. This attribute is used as seed for the ACV chain until the users switch to the next attribute, for which the selection process is repeated. When ask-



1. Space for context information, explanations or branding.
2. Users may reset zoom and panning on click on the headline.
3. Visualization space. The system arranges user input in a ladder. Boxes are adapting to the length of user input; the interface allows for zooming / panning.
4. Chatbot space. The chatbot guides the user through the laddering interview process and shows conversation history on scrolling.
5. User input and chatbot output are saved to an SQL database with the following fields: conversationID, attribute, sender, recipient, message, messageID, and timestamp.

ing why-questions repeatedly, the chatbot will rely on four techniques for rephrasing questions to help and guide the user. Peer-reviewed guidelines for human interviewers on how to conduct laddering interviews inspired the utilized techniques [15]. For now, the four techniques are applied by LadderBot at random. However, no technique may be used two times in a row. The rephrasing techniques primarily incorporate the seed attribute of the current ladder into the question formulation. User replies are used for rephrasing only in the form of quotes, for ensuring that the resulting question makes sense. The visualization of the current status of the interview on the left side updates itself for each elicited consequence before LadderBot asks the next question. To end the elicitation for a specific attribute, or the interview in general, a human interviewer would need to identify when an interviewee has reached the 'end' of an ACV chain (e.g. [29]). As the current iteration of LadderBot is not capable of recognizing on its own whether a user has already described all values for a chain of consequences, the bot requires the user to indicate willingness to continue the laddering process for the current attribute, or switch to the next chain. The user can make this indication using the command 'stop'.

The current implementation of LadderBot does not impose restrictions on the length of an answer of a user, to keep the interaction with the chatbot as natural as possible. Long replies impose a challenge for LadderBot in formulating an appropriate question as a response. As such, user replies are incorporated in questions only as complete references. Furthermore, LadderBot uses the three features provided by the users at the start of the interview to formulate more direct questions, as we identified these replies to be rather short. Users are currently not capable of making changes to previous answers. However, we plan to include this functionality in future iterations.

### B. Technical challenges

The system is developed using a combination of the Microsoft Bot Framework on node.js, d3.js as visualization framework and bootstrap to structure an HTML website. The system is capable of collecting ACV chains in an interview and writing results to an SQL database. The visualization is updated dynamically, allowing for zooming and panning and will fit responsively to various screen sizes. Technical challenges for RQ1 primarily appear when comparing chatbot interviews with human interviewers. While I expect the four questioning techniques used by LadderBot to improve its performance compared to pencil-and-paper laddering, the application of specific techniques follows no strategy thus far. For the evaluation of H2, it may be necessary to develop an (algorithmic) strategy for LadderBot to select a questioning technique based on specific parameters (e.g., time taken to provide an answer). For RQ2, technical challenges arise regarding the analysis of results. Content codes have to be computed from individual answers to have the system create an aggregate implication matrix. I currently attempt this using part-of-speech (POS) tags. However, the evaluation of the applicability of POS for content code generation and subsequent comparison/aggregation in the context of laddering interviews is still at a very early stage and will require the identification of adequate solutions or workarounds for linguistic (and likely also semantic) NL processing. In the future, I will explore both methods for visual data analysis as well as requirements mining to extend LadderBot through automatically extracting requirements-relevant information from laddering knowledge.

### VI. OUTLOOK AND EXPECTED CONTRIBUTION

This Ph.D. project will make a practical contribution to the field of RE and IS by providing a ready-to-use system for wide audience RE using a cognitive technique for end-users. I developed a prototype of the system and make preparations for a

Fig. 1. Overview and explanation of the LadderBot user interface

proof-of-concept laboratory experiment providing answers to RQ1. I attempt to make a second practical contribution by showcasing the application of the laddering technique by a conversational agent, potentially paving the way for other elicitation techniques to (partially) be conducted through a conversational agent. Furthermore, I plan to make a practical contribution by enabling the end-user RE system to process data on its own to allow for the scalable analysis of elicited knowledge, concepts, and requirements. Given that the system enables wide audience requirements collection, manual methods for generating content codes introduce a bottleneck to analysis processes. Consequently, I attempt to provide feasible solutions for aggregating results from laddering self-elicitation interviews to assist requirements analysts. On a theoretical side, I will contribute to the body of knowledge by developing a design theory for building conversational agents for RE. This design theory will include guidelines on how to combine specific interface modalities, such as conversational and graphical, for particular elicitation techniques. Furthermore, I contribute to the study of the laddering method for RE by creating and analyzing requirements data in the form of ACV chains both from a laboratory experiment with students as well as a field experiment with organizational and end-users.

#### REFERENCES

- [1] K. Villela et al., "Towards ubiquitous RE: A perspective on requirements engineering in the era of digital transformation," in 2018 IEEE 26th International Requirements Engineering Conference (RE'18), 2018, pp. 205–216.
- [2] J. M. Leimeister, H. Österle, and S. Alter, "Digital services for consumers," *Electron. Mark.*, vol. 24, no. 4, pp. 255–258, 2014.
- [3] J. Jia and L. F. Capretz, "Direct and mediating influences of user-developer perception gaps in requirements understanding on user participation," *Requir. Eng.*, vol. 23, no. 2, pp. 277–290, 2018.
- [4] H. F. Hofmann and F. Lehner, "Requirements Engineering as a Success Factor in Software Projects," *IEEE Softw.*, vol. 18, no. 4, pp. 58–66, 2001.
- [5] J. M. Moore and F. M. Shipman, "Requirements elicitation using visual and textual information," in Fifth IEEE International Symposium on Requirements Engineering, 2001, pp. 308–309.
- [6] K. Li, R. G. Dewar, and R. J. Pooley, "Computer-Assisted and Customer-Oriented Requirements Elicitation," in 2005 13th IEEE International Conference on Requirements Engineering (RE'05), 2005, pp. 4–5.
- [7] F. Pérez and P. Valderas, "Allowing end-users to actively participate within the elicitation of pervasive system requirements through immediate visualization," 2009 4th Int. Work. Requir. Eng. Vis., pp. 31–40, 2009.
- [8] D. C. Derrick, A. Read, C. Nguyen, A. Callens, and G. J. De Vreede, "Automated group facilitation for gathering wide audience end-user requirements," in Annual Hawaii International Conference on System Sciences (HICSS'13), 2013, pp. 195–204.
- [9] M. Oriol et al., "FAME: Supporting continuous requirements elicitation by combining user feedback and monitoring," in 2018 IEEE 26th International Requirements Engineering Conference (RE'18), 2018, pp. 217–227.
- [10] A. Moitra et al., "Towards development of complete and conflict-free requirements," in 2018 IEEE 26th International Requirements Engineering Conference (RE'18), 2018, pp. 286–296.
- [11] T. Tuunanen, "A new perspective on requirements elicitation methods," *J. Inf. Technol. Theory Appl.*, vol. 5, no. 3, pp. 45–72, 2003.
- [12] O. Dieste and N. Juristo, "Systematic review and aggregation of empirical studies on elicitation techniques," *IEEE Trans. Softw. Eng.*, vol. 37, no. 2, pp. 283–304, 2011.
- [13] I. Mohedas, S. R. Daly, and K. H. Sienko, "Requirements Development: Approaches and Behaviors of Novice Designers," *J. Mech. Des.*, vol. 137, no. 7, pp. 1–10, Jul. 2015.
- [14] S. Dey and S. W. Lee, "From requirements elicitation to variability analysis using repertory grid: A cognitive approach," in 2015 IEEE 23rd International Requirements Engineering Conference (RE'15), 2015, pp. 46–55.
- [15] G. M. Breakwell, *Doing Social Psychology Research*, 1st ed. The British Psychological Society and Blackwell Publishing Ltd, 2004.
- [16] R. M. P. Falcão, "Improving the Elicitation of Delightful Context-Aware Features: A Data-Based Approach," 2017 IEEE 25th Int. Requir. Eng. Conf., pp. 562–567, 2017.
- [17] R. Snijders, F. Dalpiaz, S. Brinkkemper, M. Hosseini, R. Ali, and A. Özüm, "REfine: A gamified platform for participatory requirements engineering," 1st Int. Work. Crowd-Based Requir. Eng., pp. 1–6, 2015.
- [18] C. M. Chiu, "Applying means-end chain theory to eliciting system requirements and understanding users perceptual orientations," *Inf. Manag.*, vol. 42, no. 3, pp. 455–468, 2005.
- [19] T. Tuunanen and M. Rossi, "Engineering a method for wide audience requirements elicitation and integrating it to software development," in 2004 37th Annual Hawaii International Conference on System Sciences (HICSS'04), 2004, pp. 1–10.
- [20] G. Botschen, E. M. Thelen, and R. Pieters, "Using means-end structures for benefit segmentation," *Eur. J. Mark.*, vol. 33, no. 1/2, pp. 38–58, 2004.
- [21] M. F. McTear, "Spoken dialogue technology: enabling the conversational user interface," *ACM Comput. Surv.*, vol. 34, no. 1, pp. 90–169, 2002.
- [22] L. C. Klopfenstein, S. Delpriori, S. Malatini, and A. Bogliolo, "The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms," in 2017 Conference on Designing Interactive Systems (DIS'17), 2017, pp. 555–565.
- [23] T. Rietz, I. Benke, and A. Maedche, "The Impact of Anthropomorphic and Functional Chatbot Design Features in Enterprise Collaboration Systems on User Acceptance," in 2019 14th International Conference on Wirtschaftsinformatik (WI'19), 2019, pp. 1656–1670.
- [24] U. Gnewuch, S. Morana, and A. Maedche, "Towards Designing Cooperative and Social Conversational Agents for Customer Service," in 2017 International Conference on Information Systems (ICIS'17), 2017, pp. 1–13.
- [25] J. F. Nunamaker, D. C. Derrick, A. C. Elkins, J. K. Burgoon, and M. W. Patton, "Embodied Conversational Agent-Based Kiosk for Automated Interviewing," *J. Manag. Inf. Syst.*, vol. 28, no. 1, pp. 17–48, 2011.

- [26] M. Pickard, R. M. Schuetzler, J. Valacich, and D. A. Wood, "Next-Generation Accounting Interviewing: A Comparison of Human and Embodied Conversational Agents (ECAs) as Interviewers," *SSRN Electron. J.*, no. April, pp. 1–21, 2017.
- [27] A. Hevner, S. March, J. Park, and S. Ram, "Design Science Research in Information Systems," *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.
- [28] B. Kitchenham and S. Charters, *Guidelines for performing Systematic Literature reviews in Software Engineering*. 2007.
- [29] J. Webster and R. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *Manag. Inf. Syst. Q.*, vol. 26, no. 2, pp. 13–23, 2002.
- [30] C. R. Coulin, "A Situational Approach and Intelligent Tool for Collaborative Requirements Elicitation," University of Technology, Sydney, 2007.
- [31] "AnnotatePro," 2019. [Online]. Available: <http://www.annotatepro.com>. [Accessed: 10-Jun-2019].
- [32] R. Sharma and A. Sureka, "CRUISE: A platform for crowdsourcing Requirements Elicitation and evolution," 2017 10th Int. Conf. Contemp. Comput., pp. 1–7, 2018.
- [33] "Requirements Bazaar," 2019. [Online]. Available: <https://requirements-bazaar.org/>. [Accessed: 10-Jun-2019].