

FCCC Framework for Conversational AI Systems

**Designing self-regulating systems with Fuzzy Conversational
Character Computing**

DOCTORAL THESIS

submitted to the Faculty of Management, Economics and Social Sciences at the

University of Fribourg (Switzerland)

by

Sophie HUNDETRMARK

from Braunschweig, (Germany)

in fulfillment of the requirements for the degree of Doctor of Philosophy in Business
Informatics (PhD)

Approved by the Faculty of Management, Economics and Social Sciences on
March 30th, 2026, at the proposal of
Prof. Dr. Edy Portmann (first supervisor)
Prof. Dr. Hans-Georg Fill (second supervisor)

Fribourg, Switzerland 2026

DOI : <https://doi.org/10.51363/unifr.eth.2026.034>

Sophie Hundertmark  <https://orcid.org/0000-0003-2225-4485>

© Sophie Hundertmark, 2026



This work is published under a Creative Commons Attribution 4.0 International License (CC BY): <https://creativecommons.org/licenses/by/4.0/>

FCCC Framework for Conversational AI Systems

**Designing self-regulating systems with Fuzzy Conversational
Character Computing**

DOCTORAL THESIS

submitted to the Faculty of Management, Economics and Social Sciences at the

University of Fribourg (Switzerland)

by

Sophie HUNDETRMARK

from Braunschweig, (Germany)

in fulfillment of the requirements for the degree of Doctor of Philosophy in Business
Informatics (PhD)

Approved by the Faculty of Management, Economics and Social Sciences on
March 30th, 2026, at the proposal of
Prof. Dr. Edy Portmann (first supervisor)
Prof. Dr. Hans-Georg Fill (second supervisor)

Fribourg, Switzerland 2026

The Faculty of Management, Economics and Social Sciences at the University of Fribourg (Switzerland) neither approves nor disapproves the opinions expressed in a doctoral thesis: They are to be considered those of the author.

(Decision of the Faculty Council of 23 January 1990)

Abstract

This dissertation presents the Fuzzy Conversational Character Computing (FCCC) framework, a new conceptual framework for designing conversational AI systems that act as self-regulating, adaptive and ethically aligned systems. The framework builds upon the theoretical foundations: systems theory (von Bertalanffy, 1949) as the overarching structure for self-regulation, fuzzy logic (Zadeh, 1965) to manage linguistic ambiguity and partial truths, conversational theory (Pask, 1965; Pangaro, 2010) to model dynamic feedback loops between human and machine, and Character Computing (El Bolock et al., 2020) to integrate psychological dimensions such as personality, affective state and behaviour into systems interaction.

Using the Design Science Research (DSR) methodology (Möller et al., 2020), the thesis follows an iterative process that combines theoretical synthesis, empirical investigation and artefact evaluation. Four independent use cases, knowledge transfer in financial literacy, customer service automation, AI-supported financial advisory, and character estimation via chat dialogues yielded seven design requirements that define the basis for the FCCC artefact. These findings reveal how conversational AI systems can balance user expectations, company objectives, and ethical responsibility while ensuring sustainable interaction quality.

The developed FCCC Framework connects conversational loops (Pangaro, 2010) known from the conversational theory (Pask, 1965) with multidimensional fuzzy classifications (Zumstein, 2007, 2026) and Character Computing (El Bolock et al., 2020), enabling conversational AI to continuously learn from user input and context. The inner part of the framework is mathematically formulated and implemented as a prototype chatbot that integrates fuzzy personality vectors with large language models (LLMs). This implementation demonstrates how emotional user inputs can be translated into adaptive responses that dynamically adjust tone, style and empathy.

Beyond the technical realization, the thesis evaluates the organizational and societal feasibility of the framework. Focus groups with industry and academic experts identify required capabilities such as AI literacy, an open organizational culture, and flexible IT infrastructures, and show that large and medium-sized organizations are already equipped to implement the framework, while smaller firms can do so through partnerships. In addition, expert interviews in the field of Life Engineering confirm the framework's ethical and social relevance. By prioritizing transparency, user well-being and responsible automation, the FCCC Framework contributes to developing "life agents" that support human flourishing rather than replacing human interaction.

The dissertation thus advances both theory and practice of conversational AI. It provides a scientifically grounded architecture that unites psychology, computer science and business informatics, and offers a practical path toward sustainable, empathetic and explainable AI systems.

Keywords: FCCC Framework, Conversational AI, Character Computing, Fuzzy Logic, Chatbots

Acronyms

ACM – Association for Computing Machinery
AI – Artificial Intelligence
AIA – Artificial Intelligence Act
APML – Affective Presentation Markup Language
BCB – Backend Chatbot
BFI-10 – Big Five Inventory (10 Items)
B-Vektor (B) – Behaviour Vector
CA – Conversational Agent
CC – Character Computing
CEB – Character Estimation Bot
CHI – Conference on Human Factors in Computing Systems
CWW – Computing With Words
DACH – Germany (D), Austria (A), Switzerland (CH)
DP – Design Principle
DSR – Design Science Research
ECA – Embodied Conversational Agent
ESG – Environmental, Social and Governance
FAQ – Frequently Asked Questions
FCCC – Fuzzy Conversational Character Computing
FFM – Five-Factor Model
GDPR – General Data Protection Regulation
GIIN – Global Impact Investing Network
HCI – Human–Computer Interaction
IEEE – Institute of Electrical and Electronics Engineers
IS – Information Systems
IT – Information Technology
LLM – Large Language Model
ML – Machine Learning
NLP – Natural Language Processing
NPS – Net Promoter Score
PDF – Portable Document Format
RPA – Robotic Process Automation
RQ – Research Question
RRB – Reply Rephrasing Bot
R-Vektor (R) – Reply Vector
S-Vektor (S) – Situation Vector
SLR – Systematic Literature Review
UI – User Interface
UX – User Experience
VIM – Value Irritant Matrix

Contents

1.	Introduction	1
1.1.	Motivation	1
1.2.	Thesis Outline	2
1.3.	Theoretical Foundations of the FCCC Framework	5
1.4.	Research Gap	9
1.4.1.	Evidence from the Systematic Literature Review	9
1.4.2.	Limitations of Character Computing	11
1.4.3.	The Research Gap	12
1.5.	Research Objectives	13
1.6.	Methods	16
1.6.1.	Systematic Literature Review	16
1.6.2.	Design Science Research	20
1.6.3.	Turning Test for Conversations	23
1.6.4.	Additional Research Methods	24
1.7.	Author's Own Research Contribution	25
2.	Theoretical and Methodological Foundation	28
2.1.	Conversational AI: Definition, Capabilities and Evolution	28
2.2.	Understanding Systems: How Conversational AI Evolves in a Connected World	29
2.3.	Learning from Interactions: How AI Adapts through Feedback Loops	31
2.4.	Navigating Uncertainty: Introduction of Fuzzy Logic in Conversational AI Systems	33
2.5.	More than Language: Introduction of Character Computing to Understand User Expectations	42
2.6.	Bringing It All Together: Developing a Research Model for Better AI Conversations	45
2.7.	Theoretical Foundations and their Contribution to Life Engineering	51
2.8.	Conversational AI in Business – Considering Economic Fundamentals	55
2.9.	Findings on Fundamentals in a Nutshell	56
3.	Defining Capabilities and Application Areas of Conversational AI (4 Independent Use Cases)	57
3.1.	Knowledge Transfer and Information Management through Conversational AI	58

3.1.1.	Systems for Promoting Financial Literacy: Defining the Research Question and Methodology	58
3.1.2.	Design Artefact for Interactive Knowledge Transfer	62
3.1.3.	Evaluation and Empirical Results on Conversational AI systems in Knowledge Transfer	67
3.1.4.	Requirements on Conversational AIs for Knowledge Transfer – Implications for Future Conversational AI	73
3.2.	Conversational AI in Customer Service and Sales	75
3.2.1.	Managing Conflicts between Costs and Experience: Background, Research Question and Methodology	75
3.2.2.	Managing Conflicts between Costs and Experience: Related Work and Hypotheses	77
3.2.3.	Approaches to Manage Cost-Experience Conflicts: Results	85
3.2.4.	Requirements on Conversational AIs to Meet Economic and User Expectations – Implications for Future Conversational AI	88
3.3.	AI-Supported Advisory Systems in the Financial Sectors	90
3.3.1.	AI-Supported Advisory Systems in the Financial and Insurance Sectors: Defining the Research Question and Mythology	90
3.3.2.	AI-Supported Advisory Systems in the Financial and Insurance Sectors: Evaluation and Results	93
3.3.3.	Requirements for Successful Conversational AI Advisory Systems – Future Conversational AIs	101
3.4.	Findings on Requirements Based on Knowledge Transfer and Customer Service in a Nutshell	102
3.5.	Conversational AI as System to Calculate User Character	105
3.5.1.	Conversational AI as system to calculate users character: Defining the Research Question and Methodology	105
3.5.2.	Conversational AI as system to calculate users character: Results	109
3.5.3.	Requirements for Conversational AIs as a System to Calculate User Character – Implications for Future Conversational AI	115
3.6.	Findings on Requirements for the Design of Skills for Future Conversational AI	116
4.	New Conceptual Framework for Future Conversational AI (FCCC Framework)	121
4.1.	Capabilities for Future Conversational AIs – For the Artefact	121
4.1.1.	First Version of the Artefact	122
4.1.2.	Example Application of the First Version of the Artefact	129
4.2.	Methods to Improve the Artefact	134
4.2.1.	Expert selection	134
4.2.2.	Evaluation method and data collection	136
4.2.3.	Results of First Analysis	137
4.2.4.	Final Second Version of the Artefact	138
4.3.	The FCCC Framework in the Context of Life Engineering	139
5.	Evaluation and Practical Implications of the Artefact	142

5.1.	Technical Implementation of the Artefact	142
5.1.1.	Mathematical Fundamentals on Fuzzy Logic, Linguistic Variables, and Multi-Dimensional Classification	143
5.1.2.	Mathematical and Technical Representation of the Artefact	145
5.1.3.	Experiment Setup and evaluation process	155
5.1.4.	Evaluation and Key findings of the Technical Implantation	162
5.1.5.	Summarizing Remarks on the Technical Implantation of the Artefact	166
5.2.	Necessary Organizational Capabilities to Implement the Artefact	166
5.2.1.	Evaluation of Organizational Capabilities	167
5.2.2.	Achieving the Capabilities of the Artefact	172
5.2.3.	Organizational Capabilities Required to Implement my Framework in a nutshell	173
5.3.	Social Impact of the Artefact in the Context of Life Engineering	174
5.3.1.	Application of the Artefact as Life Agents in the Context of Life Engineering	175
5.3.2.	Evaluation of the Artefact as Life Agents in the Context of Life Engineering	182
5.3.3.	Key Findings from the Expert Interviews	183
5.3.4.	Summarizing Remarks on the Social Impact of the Artefact in the Context of Life Engineering	185
6.	Summary, Discussion and Outlook	186
6.1.	Summary	186
6.2.	Alignment with Research Questions	188
6.3.	Additional Remarks from a Life Engineering perspective	196
6.4.	Remarks on the Use of Large Language Models (LLMs)	197
6.5.	Outlook and Next Steps	199
7.	References	203
8.	APPENDIX A	213
9.	APPENDIX B	214
10.	APPENDIX C	215
11.	APPENDIX D	216
12.	APPENDIX E	217
13.	APPENDIX F	218
14.	APPENDIX G	219

List of figures

Figure 1: Construct of social cues, signals and reactions	11
Figure 2: Detailed construct of social cues, signals and reactions	11
Figure 3: Design Science Research Process	20
Figure 4: Overview on the DSR of this work	22
Figure 5: Design Science Research Process including Turning Test for Conversations	24
Figure 6: Conversational Loops in Process	31
Figure 7: Concept of Linguistic Variables.....	36
Figure 8: Three-Dimensional Classification.....	37
Figure 9: Character Computing Triad.....	42
Figure 10: Fuzzy Conversational Character Computing	51
Figure 11: Value Irritant Matrix	55
Figure 12: Have you ever received training in financial matters? If yes, in which category?.....	59
Figure 13: Would you find training in personal financial management and knowledge transfer in the area of old-age provision useful?	59
Figure 14: Eggheads.ai user interface Illustration	64
Figure 15: Eggheads.ai user interface Illustration	65
Figure 16: Eggheads.ai Chat window Example.....	65
Figure 17: Eggheads.ai Chat window Example.....	66
Figure 18: Eggheads.ai Chat window Example 3.....	66
Figure 19: Landbot development Interface.....	80
Figure 20: Landbot development Interface 2.....	80
Figure 21: Chatbot with emotions – supplementary insurance	81
Figure 22: Chatbot without emotions – report damage	81
Figure 23: Chatbot with emotions – change address	82
Figure 24: Example questionnaire	92
Figure 25: Key results – address change request, reporting damage and suppl. Insurance.....	94
Figure 26: Demographics overview.....	95

Figure 27: Gender distributions	96
Figure 28: Emotions in Different Use Cases	97
Figure 29: Summary results Research Question 1.3.....	100
Figure 30: Overview: Variable “neuroticism”.....	112
Figure 31: Overview Usefulness of Bot 3	114
Figure 32: First version of the Artefact	129
Figure 33: FCCC Framework	139
Figure 34: Character estimation by mapping a particular message character vector (M) to a set of 3 character traits (T1, T2, T3).	146
Figure 35: Code- Screenshot	155
Figure 36: Concept of the experiment framework implemented in the web application for evaluating the FCCC framework	160
Figure 37: FCCC Framework, 7 steps highlighted.....	178
Figure 38: Main area of Framework.....	188

List of tables

Table 1: Overview of the four theories of the FCCC Framework. Detailed explanations of each theory are provided in Chapter 2	8
Table 2: Framework for organizing literature research.	17
Table 3: Research process overview.....	19
Table 4: Dependency structure of the five theories	54
Table 5: Overview of the four use cases in chapter 3	58
Table 6: Design principles for the first artefact	61
Table 7: Chatbot sequences	63
Table 8: Questionnaire.....	67
Table 9: New design principles driven from the evaluation.....	72
Table 10: Capabilities 1-6.....	74
Table 11: Conceptual Model – Use Case 2	83
Table 12: Use case supplementary insurance: group statistics	87
Table 13: Use case changing an address: group statistics	87
Table 14: Use case reporting damage: group statistics.....	87
Table 15: Capabilities 7-9.....	89
Table 16: Main findings from experiment.....	93
Table 17: Capabilities 10-12.....	101
Table 18: Design requirements from the use cases	103
Table 19: Design Requirements 13,14.....	116
Table 20: Design requirements first overview, draft.....	117
Table 21: List of final Design Requirements.....	119
Table 22: Mapping of fuzzy components to FCCC artefact layers	125
Table 23: Capabilities for Artefact	127
Table 24: Overview Experts for first artefact	135
Table 25: Mapping of Life Engineering requirements to FCCC components and theoretical foundations.....	141
Table 26: Mechanical procedure - Algorithm: – Character vector construction.....	152
Table 27: Prompt construction in Phase 4 and its linkage to earlier formula.....	154

Table 28: Participant demographics, experiment on technical implementation of the FCCC framework (n=14).....	162
Table 29: Observed effects of recorded character traits among all conversations with plain chatbot and character computing: Agreeableness (A), Conscientiousness (C), Extraversion (E), Neuroticism (N) and Openness (O).....	165
Table 30: Focus Group 1 participant profiles (n=15), January 2025.....	168
Table 31: Organization Capabilities for the Artefact	169
Table 32: Focus Group 2 participant profiles (n=12), February 2025.....	170
Table 33: List of Organization Capabilities for the Artefact.....	171
Table 34: 7 steps from FCCC framework, RQ4.3	177
Table 35: Example Dialogue FCCC Framework for Life Agents, RQ4.3	179
Table 36: Overview Interview Partner for ethical evaluation	183
Table 37: Capabilities linked to framework element.....	192

1. Introduction

1.1. Motivation

Chatbots, as an example of conversational artificial intelligence (AI) systems, have engaged the scientific community since at least 1966, when Weizenbaum deployed the first known chatbot ELIZA in the domain of psychology (Weizenbaum, 1966). This initial chatbot differs significantly in its operation from today's chatbots and conversational AI systems. It used neither AI, nor the methods of large language models (LLMs). However, the fundamental idea and objective of the technology were the same as today: to develop a computer system capable of communicating in natural language with its user. Initially, a common interface had to be created, followed by a common language (Pangaro, 2010). Yet, what sounds simple has repeatedly presented significant challenges to researchers and practitioners in creating natural-sounding dialogues between humans and machines. At first glance, the released chatbot by OpenAI, "ChatGPT", or alternatives like Gemini from Google appear to come very close to natural dialogues, thanks to the utilization of LLMs (OpenAI, 2022). However, upon closer inspection, it is apparent that these dialogues are only natural to a certain degree, and that there is a lack of mutual exchange of information. Since LLMs are still primarily based on probability calculations rather than genuine natural dialogues (OpenAI, 2022). Moreover, important supplementary elements such as ethics, morality and legal frameworks are missing. ChatGPT only adjusts to its individual users to a limited extent. ChatGPT's focus is rather on executing the prompt or carrying out the user command.

Here it is appropriate to refer to the ideas of "Life Engineering" as proposed by Österle (2020). According to this discipline, which is particularly known in the fields of business informatics, it is necessary to develop computer systems that contribute to the general good of society, as well as that of each individual citizen. Hence, the overarching goal of life engineering research is to improve the quality of life in society. A crucial element here are the so-called life agents. These life agents can also be developed in the form of chatbots, voicebots or conversational AI systems (Österle, 2020).

The discipline of Life engineering draws on foundations, including the concept of systems theory introduced by von Bertalanffy (1949), which also provides an essential basis for this thesis. Systems theory initially considers systems as a whole and then examines the interactions between the parts of a system and its environment. Through the application of systems theory, researchers and practitioners can gain better insights into the functional and structural properties of complex systems, and develop solutions for systemic challenges. Feedback loops are a central element of systems theory. They enable the systems to regulate themselves autonomously and adapt to their environment (von Bertalanffy, 1949). Additionally, fuzzy logic (Zadeh, 1965) and later fuzzy systems theory (Zadeh, 1982) is used not only to explain the

calculation of sharp relationships between systems, but also to analyse and depict the interaction of systems as realistically as possible, including possible deviations, inaccuracies or partial truths. By comparing today's well-known chatbots, voicebots and conversational AI systems in general, such as ChatGPT or those employed by companies, for instance, in customer service, it becomes evident that the values of life engineering researchers (Österle, 2020) or the principles of systems theory (von Bertalanffy, 1949) are hardly taken into account. Economic considerations, such as the value irritant matrix (Price et al., 2008), which primarily considers aspects regarding costs and benefits, are predominantly in the foreground. Most of today's well-known conversational AI systems also lack ethical frameworks. Research already has initial solutions for value-based engineering or methods like computational ethics (Portmann et al., 2022). Additionally, the measure called "ethicity", intended to measure the degree of ethical correctness, is also already used in research (Hundertmark et al., 2022c). But to date these ideas only exist in stand-alone research projects and are not implemented into chatbot systems.

This thesis aims to address the limitations of current conversational AI systems, particularly chatbots, and bring the user, his individual welfare, and that of society back to the forefront, in line with the objectives of life engineering (Österle, 2020). Hence, the thesis sets out to develop a conceptual framework that interlinks different disciplines from business informatics, psychology, sociology and ethics research, laying the foundation for natural conversational AI systems. The thesis uses systems theory as its overall framework and enhances this concept through the application of fuzzy logic (Zadeh, 1965), character computing (CC) (El Bolock et al., 2020) and conversational loops (Pangaro, 2010) from the conversational theory (Pask, 1965), along with the consideration of ethical, legal and economic considerations, to develop self-regulating conversational AI systems (i.e. chatbots and voicebots) that can conduct natural dialogues with users. As part of my evaluation, I show that the conceptual framework not only works from a theoretical point of view, but that is also applicable in real world conversational AI systems, like text-based chatbots.

1.2. Thesis Outline

This thesis is divided into six major chapters, each consisting of individual, in-depth sections.

Chapter 1: Motivation, Methods and Research Questions, Author's Research Contribution

Chapter 1 serves as a foundation to the thesis. First, the chapter elucidates the chosen research method. This presents a literature review which leads to the formulation of the overall research method, the design science research (DSR) approach (Möller et al., 2020). The chapter then goes on to explain how I augment and make this systematic process more sustainable by integrating the methods of the Turning test for

conversations (Pangaro, 2019). Subsequently, the chapter presents the study's four research questions (RQs) and elucidates them, while also setting out further methods required for answering the RQs.

Chapter 2: Theoretical and Methodological Foundations

Chapter 2 provides insights into the most important theories. It demonstrates how and why systems theory is used (von Bertalanffy, 1949) as a general foundation from a theoretical perspective. The chapter further explains why existing systems need to be extended with conversational loops (Pangaro, 2010) from the conversational theory (Pask, 1965), fuzzy methods such as Computing with Words (CWW) (Zadeh, 1999), and Character Computing (CC) (El Bolock et al., 2020) in order to develop sustainable and long-term conversational AI systems. Finally, the chapter presents economic concepts such as the Value Irritant Matrix (VIM) (Price et al., 2008).

Chapter 3: Defining Capabilities and Application Areas of Conversational AI (4 Independent Use Cases)

According to the method of DSR (Möller et al., 2020) used in this study, chapter 3 defines derivations and design requirements for the artefact, namely, the conceptual framework. Therefore, I conduct several independent experiments to gain important knowledge about the requirements for conversational AI systems from different perspectives and different application areas. Chapter 3 closes by providing a list of derivation of capabilities for future conversational AI systems, which I use to develop the artefact over the course of the next chapter.

Use Case 1: Conversational AI Systems in the Context of Learning and Knowledge Transfer

This first use case shows a practical application of chatbots in the context of financial literacy. The application presented here evaluates completely rule-based chatbots without the use of AI as regards their usefulness as a tool for strengthening financial literacy. To sum up this first use case, I obtain a list of six requirements, which I shall later use to develop the artefact over the course of the DSR process. Additionally, this application is also of central importance for the later evaluation of the artefact. Chapter 5 revisits and redevelops the use case of financial literacy with the conceptual framework developed in this study. At this point, I use chatbots for knowledge transfer as a representative tool of life assistants, known from life engineering research.

Use Case 2: Conversational AI Systems in the Context of Customer Service and Sales

The second experiment demonstrates that the design of conversational AI systems significantly impacts the customer experience. Here, primary consideration is given to the user's situation and emotional state, with references made to the specific

design of a conversational AI system. Moreover, it will be demonstrated that merely considering economic aspects, e.g. the value irritant matrix, is not sufficient; instead, individual users along with their personal expectations and desires must also be considered. To sum up the use case, I obtain three new design requirements for the artefact.

Use Case 3: AI-Supported Advisory Systems in the Financial Sectors

The third experiment, set in the context of AI-supported advisory systems in the financial sectors, employs a novel research method where users are asked to empathize with the situation of a text chatbot, subsequently writing chatbot dialogues that align with the users' perspectives. Additionally, this experiment considers two dimensions of user characters. This part not only provides new insights and requirements into the use and design of conversational AI systems, it also shows a still very new method for developing the appropriate design elements, taking into account the personality of individual users and the situation in which they find themselves. With regard to the requirements for the artefact, I am able to add three new requirements to the later developed artefact.

Use Case 4: Conversational AI as a System to Calculate User's Character

This case focuses on the methods of CC. Within the framework of an experiment with a large number of data sets, the triad required for CC is researched. Instead of using sensor data, as researchers in the field of CC do, I use only the information provided within a chat-dialogue between users and conversational AI systems. I conduct this laboratory experiment to determine whether only chat-dialogues provide sufficient information for a conversational AI system to calculate the triad from CC (El Bolock et al., 2020) and to adapt its behaviour accordingly. In this case, I obtain two further important requirements for the artefact.

Chapter 4: New Conceptual Framework for Future Conversational AI (FCCC Framework)

According to the applied DSR approach (Möller et al., 2020), chapter 4 represents the development phase. In chapters 2 and 3, important foundations are laid for the design of future successful conversational AI systems. In chapter 3, important practical experience is gathered and design requirements are defined in addition to the theoretical findings from chapter 2. Above all, the findings of the first three chapters confirm the need for a comprehensive conceptual framework that combines various disciplines from computer science, economics, psychology and sociology. This framework will be developed taking into account the DSR approach enhanced with the ideas of the turning test (Pangaro, 2010). For the first version of the artefact, the findings from chapters 2 and 3 serve as a basis. In line with the turning test for conversations, I conduct the first expert interviews and develop a second version of the artefact, which will be evaluated according to DSR in chapter 5.

Chapter 5: Evaluation and Practical Implications of the Artefact

Chapter 5 serves as an evaluation in a broader way. Here, a three-step evaluation phase is performed, oriented by the methods of DSR. The chapter evaluates the practical feasibility, real-world applicability and its societal impact with its contribution to the life engineering disciplines. Hence, I conduct three independent evaluations and summarize them at the end as overall feedback and results according to DSR.

Chapter 6: Summary, Discussion and Outlook

The final chapter of the thesis, chapter 6, discusses and aligns the findings surrounding the developed framework with the RQs from chapter 1. The chapter closes with my input for follow-up research and a reference to the study's limitations. The conceptual framework developed here focuses on conversational AI systems in the context of conversational loops and the application of CC. The other parts such as the ethical, economic, and legal frameworks, are considered secondary to this study and would need to be explored in greater depth and in further iterations of the DSR. Finally, an outlook on future publications is provided.

1.3. Theoretical Foundations of the FCCC Framework

The FCCC Framework developed in this thesis builds on five established theories. Each theory contributes a distinct conceptual layer to the framework and addresses a specific dimension of the design challenge. While each theory is explained in full depth in Chapter 2, this section provides a brief orientation so that the research questions and methods introduced in the following sections can be better understood.

The five theories were selected because together they form a mutually reinforcing architecture. Systems theory provides the overarching logic of self-regulation, fuzzy logic supplies the mathematical tools to handle the inherent ambiguity of human language and character, conversational theory defines the mechanism through which the system learns and adapts over time, and Character Computing (CC) connects the psychological dimensions of the user to the technical machinery of the framework. Life Engineering provides the normative horizon of the FCCC Framework. It defines what the framework is ultimately for.

Systems Theory (von Bertalanffy, 1949)

Systems theory, introduced by von Bertalanffy (1949), describes how complex entities maintain themselves and adapt by exchanging matter, energy, and information with their environment through continuous feedback loops. An open system does not operate in isolation. It receives inputs, processes them, produces outputs, and uses the response from the environment to correct and improve its own behaviour over time.

The FCCC Framework, which I develop in the course of this thesis treats a conversational AI system as an open, self-regulating system in exactly this sense. Each

conversation constitutes one feedback loop. The system receives user input, generates a response, observes the user's reaction, and adjusts its internal model of the user accordingly. This iterative self-improvement is not a peripheral feature of the framework but its central design logic. Without the systems-theoretic foundation, the conversational loops at the heart of the FCCC artefact would lack both theoretical justification and a principled stopping criterion.

Fuzzy Logic (Zadeh, 1965; Zumstein, 2026; Werro, 2015)

Classical (crisp) logic assigns every element a value of either 0 or 1. It belongs to a set or it does not. Fuzzy logic, introduced by Zadeh (1965), generalizes this by allowing partial membership. An element can belong to a set with any degree $\mu \in [0,1]$. This makes it possible to represent linguistic statements such as "the user is somewhat extroverted" or "the message is moderately polite" without forcing them into binary categories. Building on this foundation, Zumstein (2007, 2026) extended fuzzy logic to multi-dimensional classification, allowing several overlapping traits to be represented simultaneously. Additionally, Werro (2015) introduced hierarchical decomposition methods that make such classifications computationally tractable.

Human personality traits, emotional states, and situational contexts are inherently continuous and overlapping, they resist sharp boundaries. The FCCC Framework therefore uses multidimensional fuzzy classification, rather than conventional crisp logic, to represent the character of a user. This allows the system to capture that a user is simultaneously 70% open-minded, 30% stressed, and 50% extroverted, and to adapt its responses accordingly. Fuzzy logic is the mathematical language in which the entire inner layer of the FCCC artefact is formulated.

Conversational Theory (Pask, 1965; Pangaro, 2010)

Conversational theory, developed by Pask (1965) and later also applied by Pangaro (2010), holds that genuine conversation is not mere information exchange but a process in which both parties come to know each other better through repeated, iterative turns. Each turn, a conversational loop, produces new shared understanding that neither party possessed before the exchange. A conversation is only truly productive if it is forward-moving: each exchange builds on the last, moves toward a shared goal, and ends with both parties in a different epistemic state than when they began. Pangaro (2019) operationalized this insight as the Turning Test for Conversations, a quality criterion asking not whether a machine can mimic human language, but whether it can participate in a conversation that genuinely goes somewhere.

Conversational theory provides the loop architecture of the FCCC Framework. The system does not simply generate isolated responses. It participates in a structured exchange in which each turn refines its estimate of the user's character (personality, affective state, situational context) and adjusts its response style accordingly.

Character Computing (El Bolock et al., 2020)

Character Computing (CC) is an interdisciplinary research field that models the full complexity of a person's character from the perspective of computer science and psychology, in order to enable computer systems to sense, adapt to, and generate human-like character representations. El Bolock et al. (2020) propose a triad as the foundational unit of CC. A person's behaviour in any given moment is determined by the interaction of their stable character traits (personality, including the Big Five dimensions), their current affective state (emotions and mood), and the situational context they are in. CC encompasses three research modules. Sensing character (from signals such as text, speech, or physiological data), adapting to character (adjusting system behaviour based on the inferred character), and creating artificial characters (designing agents with consistent, plausible personalities).

Character Computing provides the psychological substrate of the FCCC Framework. The framework's purpose is to enable a conversational AI to adapt its tone, style, and content to the individual user and CC defines precisely what must be modelled to achieve this. The triad of personality, affective state, and situational context. However, CC as originally conceived relies on physical sensors to measure character signals. One of the core research contributions of this thesis is to show that conventional CC methods are insufficient for text-only conversational AI and to propose fuzzy logic as the mechanism that compensates for the absence of sensors by computing with ambiguous linguistic signals instead.

Life Engineering (Österle, 2020)

Life Engineering is an interdisciplinary research field proposed by Österle (2020) that places human flourishing at the center of information systems design. The discipline argues that digital systems, including AI and conversational agents, should be evaluated not only by technical performance or economic efficiency, but by whether they genuinely improve the quality of life of the people who use them. To achieve this, Life Engineering draws on systems theory (von Bertalanffy, 1949) as its architectural foundation and on fuzzy logic (Zadeh, 1965) to capture the inherent imprecision of human values and needs. A key operational concept in Life Engineering is the notion of a *life agent*. A digital system that acts on behalf of a person, supports their decision-making, respects their autonomy, and prioritizes their long-term well-being over short-term engagement metrics. Life Engineering also demands that such systems are transparent about their nature and limitations, avoid nudging users toward outcomes that serve system operators rather than users, and are designed with the energy and resource constraints of sustainable operation in mind.

Life Engineering provides the normative horizon of the developed FCCC Framework. It defines what the framework is ultimately for. Whereas systems theory, fuzzy logic, and conversational theory describe how the framework operates, and

Character Computing describes what it models, Life Engineering specifies why it should be built in the first place and how its societal impact should be judged.

These five theories are not applied in isolation. Their integration is the defining contribution of the FCCC Framework. Fuzzy logic operationalizes the character model defined by Character Computing, conversational theory provides the feedback mechanism through which that character model evolves over time, systems theory supplies the overarching architectural logic that makes the whole self-regulating, and Life Engineering provides the normative horizon that determines what the system should ultimately achieve for the people it serves.

Table 1 summarizes the five theories and their specific contribution to the FCCC Framework, as a reference for the reader throughout the thesis. And chapter 2 provides the full theoretical derivation of this integration.

Table 1: Overview of the four theories of the FCCC Framework. Detailed explanations of each theory are provided in Chapter 2

Theory	Key reference(s)	Core concept
Systems Theory	von Bertalanffy (1949)	Open, self-regulating systems adapt to their environment through continuous feedback loops.
Fuzzy Logic (incl. multi-dimensional classification)	Zadeh (1965) Zumstein (2007, 2026)	Elements can belong to sets with partial degrees of membership $\mu \in [0,1]$, enabling the representation of gradual, overlapping, and linguistically expressed concepts.
Conversational Theory	Pask (1965), Pangaro (2010, 2019)	Genuine conversation is an iterative, forward-moving process in which both parties build shared understanding through structured turns.
Character Computing	El Bolock et al. (2020)	A person's behaviour is determined by the triad of personality (stable traits), affective state (dynamic emotions), and situational context.
Life Engineering	Österle (2020)	Digital systems should support human flourishing, autonomy, and long-term wellbeing rather than short-term engagement or purely economic goals. Life agents act on behalf of users, transparently and sustainably.

Source: Data collected by the author.

1.4. Research Gap

Before introducing the research questions and the methods employed to answer them, this section points out the research gap that this thesis addresses. The gap is derived mainly from two independent resources: a systematic literature review (SLR) on conversational AI design, and an analysis of the state of the art in Character Computing (CC). The full methodological details of the SLR are described in chapter 1.6.

1.4.1. Evidence from the Systematic Literature Review

The SLR, conducted in accordance with the methodology of vom Brocke et al. (2015) and Webster and Watson (2002), identified 50 relevant empirical papers on conversational AI design published between January 2011 and May 2021. The review focused on three interconnected constructs: social cues (design features of a conversational agent), social signals (the user's cognitive and emotional response to those cues), and social reactions (the resulting observable behaviour or attitude). The full list of analysed papers is provided in Appendix E; the detailed findings of this review are further discussed in Section 3.2.2.

I have already published a detailed analysis, including more information about cues and signals of this systematic literature review (SLR) in my previous work: Hundertmark, S., Fill, H.-G., & Portmann, E. (2022a). From practice to science: Social reactions caused by conversational agents—A literature review for researchers and practitioners with a chatbot as an outcome.

Across the 50 papers which were analysed, it becomes clear that prior research documents which design cues of conversational AI systems belong to which social signals and elicit specific social reactions in users very clearly. For instance, the use of emotional language increases satisfaction (Rhio et al., 2019; Doctor et al., 2016), human-like design raises trust and usage rates especially among older users (Jordan et al., 2015), and personality-congruent design, matching the character of the bot to that of the user, leads to more positive evaluations (Ghandeharioun et al., 2019; Ng et al., 2020). At the same time, the literature also demonstrates that the appropriate design depends on context. Users in a negative affective state prefer neutral, fact-based responses (Crolic et al., 2021; Katayama et al., 2019), whereas users in a positive state appreciate emotional design elements.

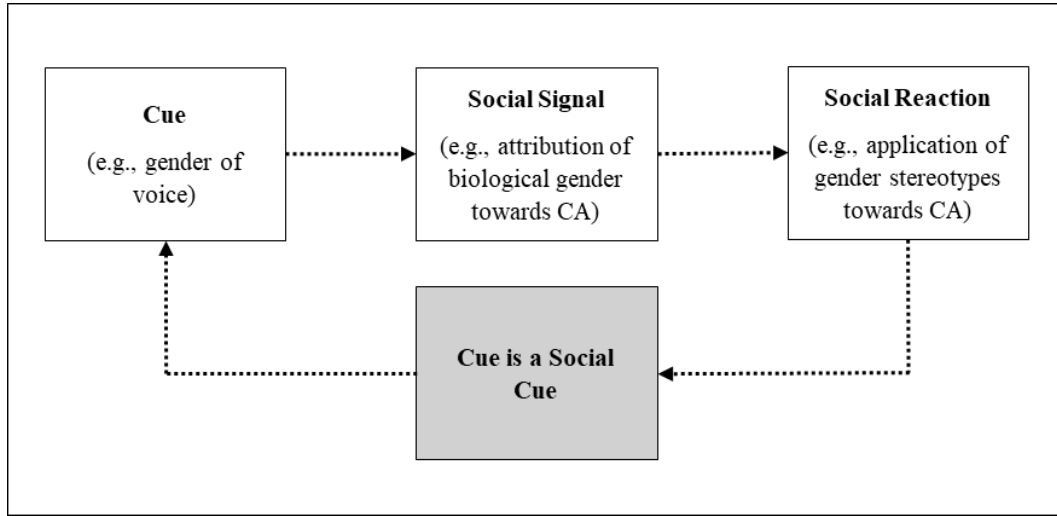
These findings are summarised visually in Figures 1 and 2. Figure 1 (based on Feine et al., 2019) introduces the foundational conceptual logic underlying the literature review. A cue is any observable characteristic of a conversational AI system, such as the gender of a voice, that becomes salient to the user. Once perceived, this cue triggers a social signal, meaning the user attributes a social meaning to it (for example, attributing a biological gender to the system). This social signal in turn elicits a social reaction, a concrete behavioural or evaluative response, such as applying gender stereotypes to the conversational AI.

Figure 2 (based on Hundertmark et al., 2022a) applies this logic to the full body of reviewed literature and maps the identified relationships in detail. On the left, eight categories of social cues are identified, ranging from quality-related and language-related cues to human-like design, user-centred communication, dealing with failure, and influencing the bot's character. These cues connect, via solid lines indicating strong empirical support and dashed lines indicating weaker or more context-dependent relationships, to four categories of social signals in the centre: utility-related, doubt-related, emotion-related, and relationship-related signals. These signals in turn connect to six categories of social reactions on the right: satisfaction, trust-related responses, behaviour-related responses, sales-related outcomes, and brand-related perceptions. The density and directionality of the connections in Figure 2 illustrate both the breadth of what prior research has established and the complexity of the cue-signal-reaction system as a whole.

However, when analysing and comparing all 50 papers, a critical gap becomes apparent. While the research collectively describes the target state of conversational AI design for a wide range of user types and contexts, none of the analysed papers addresses the question of how a conversational AI system can autonomously identify the appropriate design for a specific user and then dynamically adapt its responses during an ongoing dialogue. The following specific shortcomings were identified across all reviewed studies:

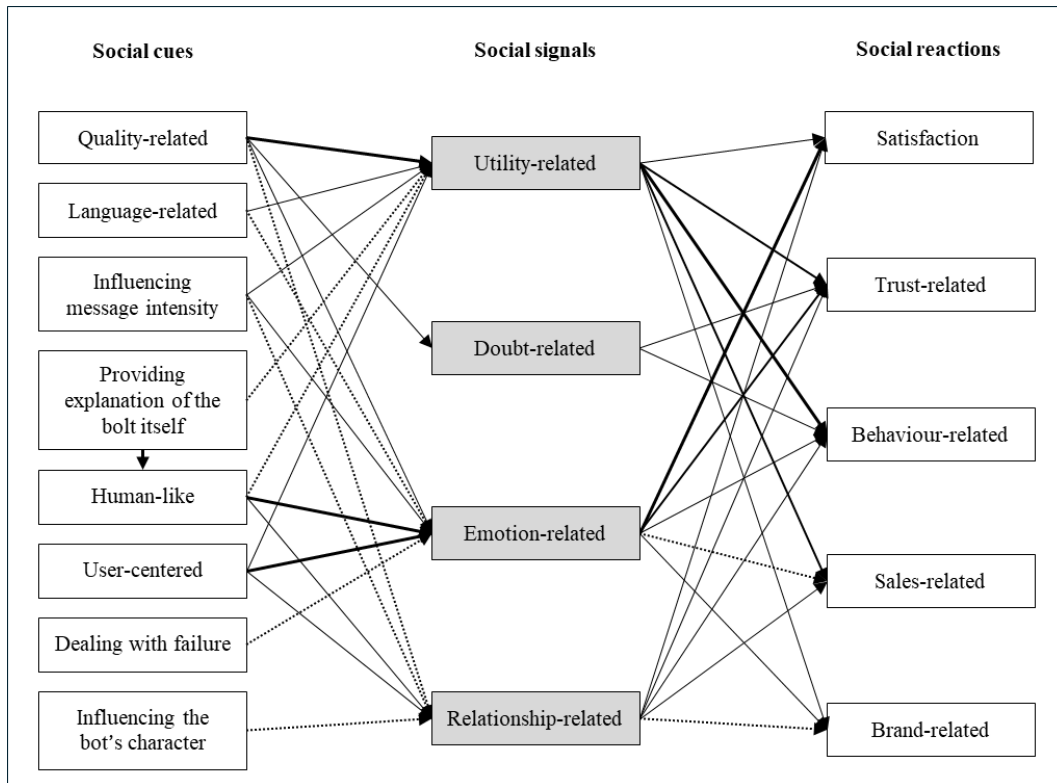
- No paper integrates principles of self-regulating systems (von Bertalanffy, 1949) into conversational AI design. All design adaptations described in the literature require human configuration and do not emerge autonomously from the system itself.
- No paper considers fuzzy logic (Zadeh, 1965) as a mechanism for handling the inherent ambiguity and gradations of real user states, personality traits or situational contexts.
- No paper proposes a holistic framework that simultaneously considers the user's personality, affective state and situational behaviour across multiple use cases and application domains.
- Ethical, legal and economic constraints are treated in isolation in existing literature, and are not integrated into conversational AI design frameworks in a systematic, holistic manner.

Figure 1: Construct of social cues, signals and reactions



Source: According to Feine et al. (2019).

Figure 2: Detailed construct of social cues, signals and reactions



Source: Data collected by the author, according to Hundertmark et al. (2022a).

1.4.2. Limitations of Character Computing

The only research direction identified in the SLR that explicitly addresses adaptive user modelling in conversational AI is the field of Character Computing (CC) (El Bolock et al., 2020). CC focuses on the user's entire character, including personality, affective state and situational behaviour, and attempts to provide the computer system

with sufficient information about the user's character so that the system can adapt its responses accordingly.

However, as the research for this thesis was conducted, the methods of CC left critical questions unanswered. Although CC offers the possibility of mapping the full character triad of the user, this is currently only possible with the help of additional sensor data (e.g. physiological measures, camera input), and cannot yet be accomplished using simple chat dialogues alone. Furthermore, the precise definition of individual characters still presents researchers with considerable challenges: El Bolock et al. (2020), regarded as leading researchers in this field, themselves acknowledge that the use of fixed numerical values to represent user character results in the loss of important information, since the actual character of a person is inherently gradual and cannot be fully captured by sharp categorical classifications. El Bolock et al. (2020) therefore explicitly point out that the application of fuzzy logic (Zadeh, 1965) would be of great importance as a supplementary method to overcome this limitation.

1.4.3. The Research Gap

Drawing together the findings from the SLR and the analysis of CC, the following central research gap can be stated:

There is currently no conceptual framework for conversational AI systems that holistically integrates self-regulation through feedback during the conversations, adaptive user modelling based on the full character triad from Character Computing (El Bolock et al., 2020); the management of linguistic ambiguity and gradual user states through fuzzy logic (Zadeh, 1965); and ethical, legal and economic constraints – in a way that enables a conversational AI system to autonomously derive the character of its user from ongoing chat dialogues alone and dynamically adapt its interaction style accordingly.

This gap is significant both from a scientific and a practical perspective. Scientifically, the absence of such a framework means that the considerable body of knowledge on user expectations, character modelling and systems theory has not yet been synthesised into an actionable design artefact for conversational AI. From a practical perspective, the lack of autonomous, context-sensitive adaptation in current conversational AI systems represents a missed opportunity to create systems that genuinely support users in line with the goals of Life Engineering (Österle, 2020).

This thesis addresses the identified gap through the development and empirical evaluation of the Fuzzy Conversational Character Computing (FCCC) Framework, a new conceptual framework for conversational AI systems, invented in the course of this thesis. The following research questions are derived directly from the components of this gap.

1.5. Research Objectives

Addressing the research gap identified in the previous chapter, the overarching goal of this thesis is to design a conceptual framework that unites the synergies between the disciplines of economics, psychology and computer science to enable the development of naturally acting, sustainable, ethically and economically successful conversational AI systems that function as a self-regulating system, as known from systems theory (von Bertalanffy, 1949), and that compute with fuzzy values (Zadeh, 1965).

To achieve this, four research questions (RQs) were defined, along with a number of sub-questions, which together address the full scope of the research gap.

Research Question 1 (RQ1):

What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

- ***RQ1.1: What design criteria should conversational AI systems incorporate to optimize knowledge transfer while ensuring a positive user experience and supporting life engineering goals?***
- ***RQ1.2: What design criteria should conversational AI systems incorporate to balance user objectives, such as quick and correct solutions, with company goals, including economic efficiency and customer experience?***
- ***RQ1.3: How do the requirements for conversational AI systems differ depending on the user's situation, age, gender, personality and experience with chatbots?***

To address RQ1 and its sub-questions, several to some extent independent experiments are conducted.

RQ1.1 focuses on exploring design principles specific to knowledge transfer and evaluating them using a DSR process (Möller et al., 2020). To determine the design principles for the first artefact, a literature review is conducted. For the evaluation, qualitative research with a selected group of experts is performed.

I revisit the domain of knowledge transfer at the end of my thesis when I assess the practical utility of the subsequently developed conceptual framework. I also utilize this use case of conversational AI systems as a representative of life assistants known in life engineering research. Thus, ultimately, the conceptual framework not only serves economic goals, but also provides additional knowledge in the area of life engineering.

RQ1.2 and RQ1.3 focus on addressing user-specific requirements based on demographics, situation and prior chatbot experience, and combine these with the economic goals of the company that offers the conversational AI system, while also paying attention to ethical and practical requirements.

Based on the preceding literature review, I initially develop hypotheses, which I then test with the aid of prototypes in the form of experiments with test users.

The majority of these encompass quantitative research approaches. Given that, for many users, the exploration of conversational AI systems represents a very novel field, I presume that surveys without prototypes are less effective, as respondents need to first experience the conversational AI systems before they can evaluate them. From a technical perspective, it is sensible to develop several independent prototypes, as this allows for a more precise analysis of individual features.

Research Question 2 (RQ2):

How effectively can methods from Character Computing be integrated into contemporary conversational AI systems to interpret the complete triad of user characteristics (personality, affective state and behaviour), and what existing gaps need to be addressed to enable a chatbot to adapt its interactions accordingly?

- ***RQ2.1: Can a conversational AI system derive the complete triad from Character Computing (including personality, affective state and situational behaviour) using features extracted from user messages?***
- ***RQ2.2: Can a conversational AI system define users' expectations regarding chatbot interactions when identifying users' personalities and affective states?***

While RQ1 addresses the specific design and its manifestations in conversational AI systems across different use cases and target groups, the responses to RQ2 provide interesting insights into the identification of appropriate design criteria. The focus is not solely on how a conversational AI system should be designed, but also on how the system itself can ascertain a suitable design and what challenges are encountered in this process.

To address RQ2 and its sub-questions, a laboratory experiment with a huge number of participants is conducted. RQ2 requires a significant amount of data from different users, hence I use a rather large sample with various users for the experiment.

Research Question 3 (RQ3):

How can sustainable and naturally intelligent conversational AI systems that function as a self-regulating system, as known from systems theory, and that

compute with fuzzy values be implemented by companies in the long term while complying with ethical, cultural, legal and economic frameworks?

RQ3 equates to the development and optimization of the final artefact within the DSR process. While RQ1 and RQ2 together with their sub-questions lay the groundwork for addressing the third question, the artefact is now formed with the aid of the DSR process and is evaluated and optimized in iterations. During the iterations, I follow Pangaro's (2019) ideas on the Turning test for conversations and designing for conversations. This means I make sure that each evaluation process is based on intelligent conversations, which leads to new ideas for improvement.

Research Question 4 (RQ4):

How can the conceptual framework be evaluated within the DSR process (Möller et al., 2020) to ensure its theoretical robustness, practical feasibility, real-world applicability and societal impact?

This overarching RQ incorporates the evaluation of the framework across technical, practical, organizational and societal dimensions, ensuring a holistic approach to validation.

I address RQ4 using three independent sub-questions:

- ***RQ4.1: How can the principles of Fuzzy Logic and Fuzzy Conversational Character Computing be technically operationalized to process ambiguous user inputs, and dynamically adapt to personality traits and emotional states?***
- ***RQ4.2: What organizational capabilities must be in place to conceptually and technically implement the framework, and how realistic is it for organizations to meet these prerequisites?***
- ***RQ4.3: Does the framework align with the principles of Life Engineering, and how well does it meet ethical expectations to deliver the desired societal impact?***

RQ4.1 explores the technical feasibility of the conceptual framework. The focus lies on leveraging fuzzy logic and Fuzzy Conversational Character Computing (FCCC) to handle the inherent ambiguity and complexity of natural language inputs. The technical evaluation involves the development of a prototype that integrates fuzzy classification models and adaptation algorithms. This prototype will undergo testing to assess its ability to interpret ambiguous inputs, deliver context-sensitive responses and maintain a human-centric interaction style.

RQ4.2 shifts the focus to the organizational readiness required to successfully implement the framework. The introduction of a novel conversational AI system often necessitates a supportive environment within the organization, including technological infrastructure, skilled personnel and a culture open to innovation. Change management processes and employee training programs play a critical role in ensuring the

smooth adoption of the framework. To address these considerations, a two-step approach is proposed. First, a theoretical analysis and literature review will identify the specific organizational capabilities needed for implementation, such as infrastructure, training and cultural readiness. Second, qualitative expert interviews will assess the practical feasibility of these prerequisites, highlighting potential challenges organizations may face and strategies to overcome them. By bridging the gap between theory and practice, this evaluation aims to provide actionable recommendations for organizations to facilitate the successful deployment of the framework.

Finally, RQ4.3 examines the framework's alignment with societal and ethical expectations. Grounded in the principles of life engineering, the evaluation focuses on whether the framework supports human-centric design, ethical responsibility and long-term societal benefits. Life engineering emphasizes the importance of technologies that not only meet individual needs, but also contribute positively to overall well-being and life goals (Österle, 2020). To address RQ4.3, the framework is critically assessed through theoretical analysis and expert discussions. Key considerations include data privacy, fairness, inclusivity and the framework's potential to enhance individual well-being. Stakeholder interviews will provide additional insights into whether the framework can effectively balance technological innovation with ethical integrity. This evaluation ensures that the framework is not only technically and organizationally feasible, but also socially responsible and impactful.

1.6. Methods

To ensure a high level of transparency, the following section outlines the research methods employed in this thesis and their underlying rationale. The overarching method is Design Science Research (DSR) (Möller et al., 2020), augmented by the Turing Test for Conversations (Pangaro, 2019). In preparing the DSR process, a Systematic Literature Review (SLR) was first conducted to establish the state of the art and derive the research gap described in Section 1.3. Additionally, several further empirical methods are applied in Chapter 3 and are introduced briefly below.

1.6.1. Systematic Literature Review

Preparation for this thesis involved conducting a systematic literature review (SLR) based on the work of vom Brocke et al. (2015).

To find relevant literature as a basis for the analysis of the dependencies between social cues, social signals and social reactions, an SLR of the work done by Webster and Watson (2002), and vom Brocke et al. (2015) was conducted. In their research, vom Brocke et al. (2015) began with a definition of the scope to clarify and determine the basic features of a literature search before the actual work starts. Table 2 shows the framework for organizing literature research, including the dimensions process, sources, coverage and techniques.

Table 2: Framework for organizing literature research.

1	Process	Sequential	Iterative	
2	Sources	Citation indexing services	Bibliographic data-bases	Publications
3	Coverage	Comprehensive	Representative	Seminal works
4	Techniques	Keyword search	Backward search	Forward search

Source: According to vom Brocke et al. (2015, p. 214).

For the purposes of this research, a sequential search process was chosen. This means that the whole search process followed pre-defined steps. Hence, even if the researchers decide on the sequential search process, it must be refined and updated during the review (vom Brocke et al., 2015).

In general, search sources refer to the origins of the data included in the study. According to vom Brocke et al. (2015), literature searches can draw on databases, publication outlets (e.g. journals) and citation indexing services. In this literature review, the most important libraries in Information systems (IS) research were considered: the IEEE Library, ACM Portal, Science Direct and Emerald. ResearchGate was also added to the research sources. However, it is worth noting that many publications exist on ResearchGate and in one of the mentioned libraries.

Literature searches can be either comprehensive, to collect as many relevant publications as possible, or focused, to centre on the seminal works in a field or sub-field (vom Brocke et al., 2015). Therefore, for this literature review, my intention is to achieve representative coverage of the design elements reported in the literature.

Finally, to conduct the review, I chose a mixture of different research techniques, starting with a keyword search, followed by backward and forward searches, and ending with a snowball search approach. Backward searching essentially refers to collecting appropriate publications by screening the reference lists of the papers retrieved from the keyword search (“going backward in time”), whereas forward searching refers to collecting appropriate publications that have cited these papers (“going forward in time”) (Webster et al. 2002). Adding a snowball search approach to the literature review may be a good supplement to the use of database searches (Wohlin, 2014).

A high level of transparency in this literature review was ensured by taking the main methodical search steps as described below.

Step 1: Keyword research

Keyword research was performed on the relevant libraries for Information System (IS) and Human Computer Interaction (HCI) research: IEEE Library, ACM Portal, Science Direct and Emerald. These libraries are the most commonly used in computer science and software engineering (Carrera-Rivera et al., 2022). To ensure that

the right keywords and suitable libraries were used, knowledge exchange with other researchers from the field of chatbots and IS research was performed and their input was directly integrated.

Finally, focus was placed on the keywords “conversational AI” “conversational agent”, “chatbot” and “chat bot”. All variations of the keywords (singular or plural, hyphenated or not hyphenated) used in the title or abstract were considered. Additionally, the search process focused only on papers published in the last 10 years (January 2011 to May 2021).

Step 2: Paper selection

All search results were screened according to pre-defined exclusion criteria. These exclusion criteria were as follows:

- the study does not include an experiment or a procedure for A-B testing,
- the study does not consider social cues and their effect on users’ reactions, and
- the study does not focus on text-based chatbots or the study is not written in English.

Additionally, duplicates were removed, which were mainly from the keyword search via ResearchGate.

Generally, the idea to focus on experiments only first stemmed from the self-made assumption that pure surveys provide little information in the context of HCI. This is because most respondents must have previously experienced a Conversational AI with cues to be able to make statements regarding their expected behaviour.

Second, to find the dependencies among cues, signals and reactions, it is necessary to explore the causal connections between the three. However, few methods suggest causality as conclusively as experiments do (Webster et al., 2014). Experimental methods involve the systematic manipulation of independent variables while measuring the target behaviour to determine whether a relationship exists between the variables manipulated and behaviour (Sturmey, 2020). After this second step, a total of 37 papers were included in the literature research.

Step 3: Additional research

With the papers selected from the previous step, backward and forward research was performed and a snowball search approach was applied. Since most of the papers that remained after the previous step had been published in the last one or two years, there was only a very small amount of literature that could be added by the forward search. In sum, after the third step, 13 more research items were added. Table 3 provides an overview of the whole search process and the results.

Table 3: Research process overview

Step	Research Approach	Outcome
1 st Step	Keyword Research Libraries: - IEEE Library - ACM Portal - Science Direct - Emerald - Research Gate Keywords: - “Conversational agent” - “Chatbot” - “chat bot” (singular, plural, hyphenated or not; searched in title, abstract, and keywords)	1753 papers in total: - ACM: 697 - Emerald: 12 - IEEE: 796 - Science Direct: 248
2 nd Step	Results screening Exclusion criteria: - No experiment or A/B testing - No social cue analysis - Not focused on text-based chatbots - Not written in English Remove duplicates	37 papers in total: - ACM: 3 - Emerald: 1 - IEEE: 3 - Research Gate: 10 - Science Direct: 20
3 rd Step	Backward Search Forward Search Snowballing approach	50 papers in total: - ACM: 4 - Emerald: 1 - IEEE: 3 - Research Gate: 19 - Science Direct: 23

Source: Data collected by the author, according to Hundertmark et al. (2022a).

Step 4: Paper analysis

The 50 papers were analysed using an iterative coding process.

The iterative process started with an analysis of 10 randomly chosen papers. For each paper, the cues, signals and reactions discussed were listed. Then, individual chosen codes and descriptions were compared to develop a list of comparable master codes and descriptions. For instance, some researchers use the term “human-like”(Ng et al., 2020), while others use the term “human-design”(Diederich et al. 2019). Hence, assuming that all papers focused on text-based chatbots, the terms “human-like” and “human-design” were summarized.

A conceptual model linking social cues, social signals and social reactions was developed (see Figures 1 and 2). The model reveals that no existing paper addresses how a conversational AI system can autonomously and holistically adapt to the user’s

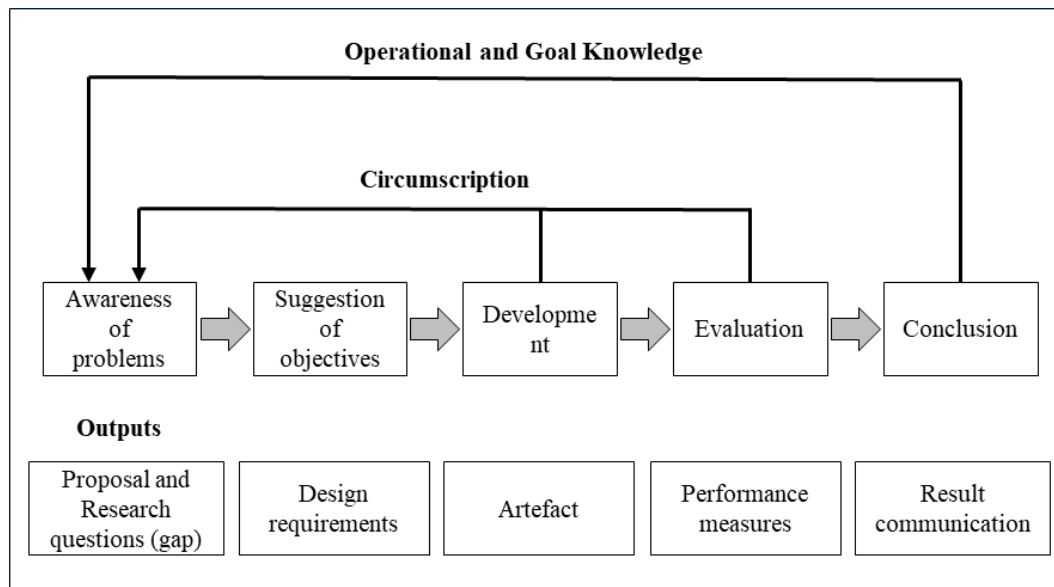
character during an ongoing dialogue as described in the research gap identified in chapter 1.4.

1.6.2. Design Science Research

The whole dissertation is based on the method of DSR (Möller et al., 2020), combined with Pangaro's (2019) ideas on the Turning test for conversations and designing for conversations in general.

The DSR method is particularly widespread in the field of business informatics. In DSR, existing research or results are not simply revised, rather it enables researchers to create completely new artefacts or models (Möller et al., 2020). Figure 3 below shows the applied DSR process. This involves an iterative process in which researchers repeatedly integrate feedback from selected experts from the fields of research and practice into the optimization rounds of the artefact. One of the main objectives of DSR is to augment both understanding and practice through the creation, execution and examination of functional innovative artefacts. DSR embodies an iterative procedure, where a core problem is tackled with a provisional solution in the guise of an artefact based on identified objectives crucial to resolving the issue (Möller et al., 2020).

Figure 3: Design Science Research Process



Source: Data collected by the author, based on Möller (2020).

The research problem and the value of its solution or the motivation for the entire work has to be defined (Hevner et al., 2004). The former serves as the basis for designing the artefact, and the latter provides arguments for the development and support of understanding (Hevner et al., 2010; Peffers et al., 2007).

Based on the findings of the first step where problem definition and motivation are defined, possible and feasible design requirements for an artefact are developed (Hevner et al., 2010; Peffers et al., 2007). Following this, the later prepared artefact is utilized as a foundation for assessing the solution to the problem.

To design an artefact based on the defined requirements of the last step, its architecture and functionality have to be determined. Subsequently, the artefact is created (Hevner et al., 2010; Peffers et al., 2007).

Furthermore, what has been learned from the various stages of the DSR process can highlight new challenges unearthed by the discovered solution, initiating a fresh iteration in the design research cycle (Möller et al., 2020).

The evaluation is one of the key aspects of DSR. While the first part of the evaluation part consists of demonstrating that the artefact fulfils its purpose and works accordingly (Hevner et al., 2010; Peffers et al., 2007), the second and later parts relate to the performance measurement of the created artefact. Among others, the artefact can be compared with the performance of another artefact that solves the same problem, or it can be evaluated based on the results of satisfaction surveys. After every iteration researchers carefully evaluate the results and point out what has been learned for further iterations.

Additionally, at the end of each evaluation step, the researchers decide whether to return to process step “develop artefact” to improve the design, or to move on to the final process step (Hevner et al., 2010; Peffers et al., 2007).

The final step relates to the conclusion and communication of the findings to researchers, practitioners and others who are interested in this research issue by publishing the insights in articles and books (Hevner et al., 2010; Peffers et al., 2007).

In this thesis, the overall goal is to develop a completely new conceptual framework for conversational AI systems. This means that the final artefacts will be the conceptual framework which is developed over the course of this thesis and evaluated with the aid of three different measurement methods towards the end of the DSR process. However, the DSR starts with an evaluation of existing research and practical knowledge about the topic. This is mainly done in chapter 2 of this thesis, where fundamental methods and technologies are analysed, explained and combined.

In accordance with DSR, the development of design requirements for the artefact constitutes the initial phase. I establish design requirements through four largely independent experiments conducted within the domains of financial literacy, customer service and CC. Each experiment identifies new design requirements for the artefact. Once the four experiments have been concluded, I consolidate the findings, merge overlapping requirements and define the final design requirements for the planned artefact (chapter 3).

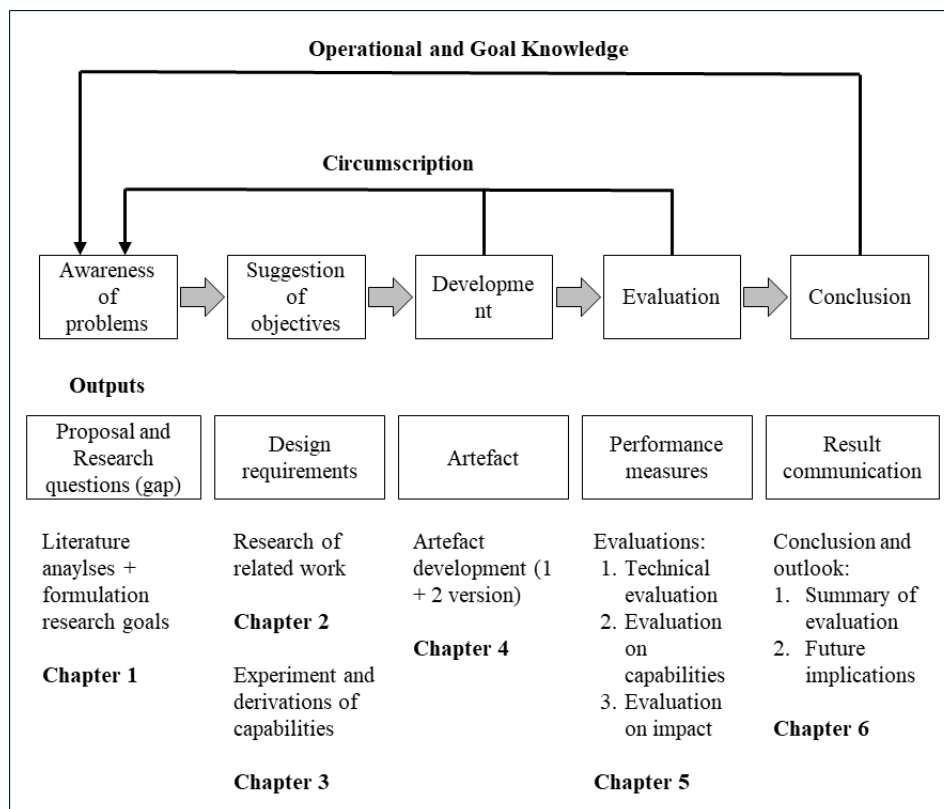
Following this, the artefact development process, detailed in chapter 4, begins. This phase is informed by the research findings from chapter 2 and is primarily guided by the design requirements established in chapter 3. At this stage, I conduct a preliminary evaluation through initial expert interviews and present a second, slightly modified version of the original artefact at the end of its development.

This refined version of the artefact is then subjected to an extensive evaluation aligned with the DSR process. I employ three independent evaluation mechanisms:

- **Technical feasibility:** the conceptual framework is transformed into mathematical models, implemented in Python, and integrated within a chatbot interface.
- **Organizational capabilities:** this evaluation assesses the organizational prerequisites necessary for artefact implementation and examines their feasibility.
- **Social Impact analysis:** this method evaluates the artefact within the context of life engineering and explores its broader societal implications.

The DSR process culminates, as suggested by Möller et al. (2020) with the communication and discussion of results cycle. This includes providing implications for both practice and research, and outlining potential iterations for further refinement of the artefact. Figure 4 gives an overview on the DSR which I conduct over the course of this work.

Figure 4: Overview on the DSR of this work



Source: Data collected by the author, based on Möller (2020).

1.6.3. Turning Test for Conversations

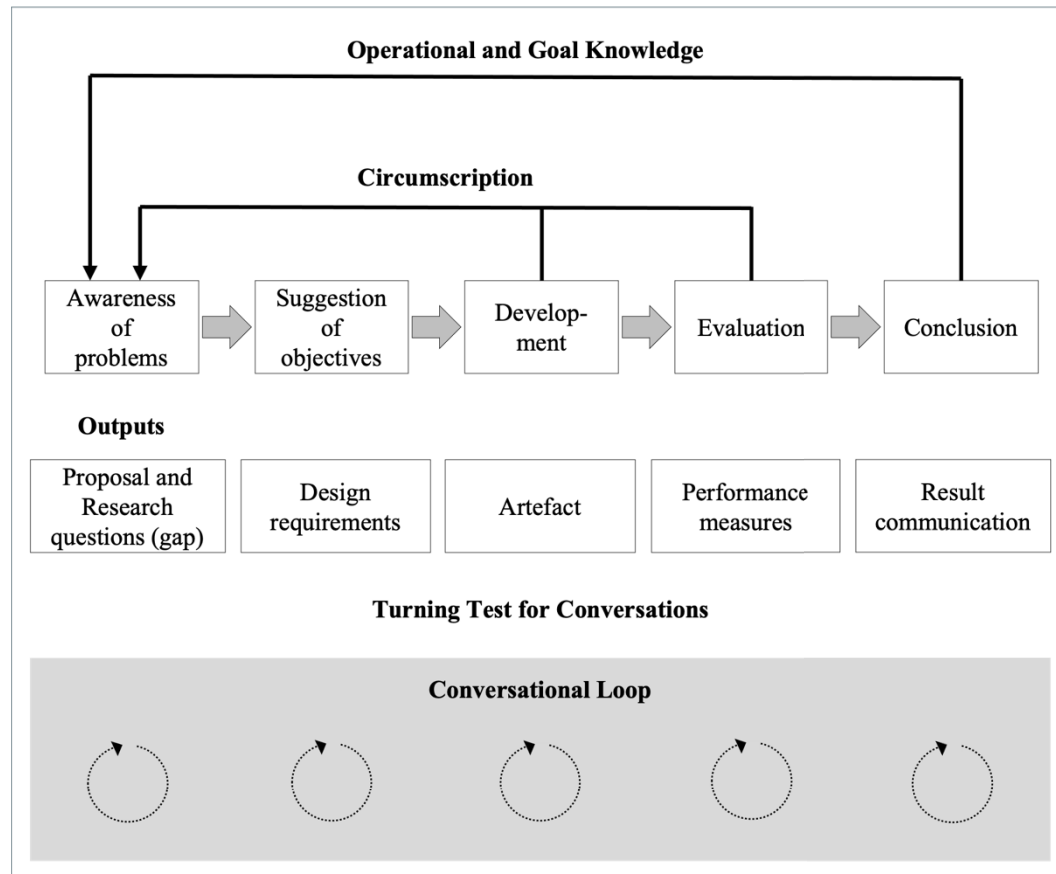
The methodology of this dissertation rests on DSR, but DSR alone does not specify what makes a good artefact in the domain of conversation. It prescribes how to build and iterate, yet leaves open the question of quality criteria for dialogue itself. In a dissertation whose central claim is that conversational AI systems can produce adaptive, meaningful exchanges with users, this is a relevant gap. An artefact may be technically complete and organizationally feasible, and yet produce conversations that are static, one-directional or ultimately unproductive. A complementary evaluative lens that specifically addresses conversational quality is therefore required.

This gap is addressed by drawing on Paul Pangaro's concept of the turning test for Conversations (Pangaro, 2019). A brief clarification on terminology is warranted here. Despite the near-identical sound, the turning Test has nothing to do with Alan Turing's Imitation Game (1950). The name derives instead from the concept of *the* conversational turn, the fundamental unit of back-and-forth exchange in which one participant hands the thread of a conversation to the other, who then responds and turns it back. This is the mechanism at the heart of Pask's Conversational Theory (1965), which is itself rooted in cybernetics (Wiener, 1948) a conversation is a purposeful sequence of turns in which both parties progressively align towards a shared goal.

Building on this, Pangaro's (2019) turning test asks a specific evaluative question: does a conversation actually go somewhere? For Pangaro, a conversation is intelligent if it is forward-moving (each turn advances the shared goal rather than merely maintaining it), collaborative (both parties adapt their communication to each other across turns) and generative (the exchange produces new understanding that neither party held before the conversation began). A conversation that meets these three criteria passes the Turning Test, regardless of whether the participants are human, machine, or a combination of both.

This evaluative standard is directly relevant to the developed framework. The framework's core purpose is not simply to produce technically correct responses, but to enable conversations that are genuinely adaptive to the user's character across multiple turns. The turning test operationalizes exactly this ambition as an evaluation criterion. At each DSR iteration, the conversations produced by the new prototype are therefore assessed on two levels simultaneously: The DSR evaluation asks whether the artefact functions as designed and meets its requirements and the Turning test evaluation asks whether the resulting conversations are forward-moving, collaborative and generative. As illustrated in Figure 5, both levels are embedded in every iteration of the DSR cycle, ensuring that technical precision and conversational quality are developed in tandem rather than sequentially.

Figure 5: Design Science Research Process including Turning Test for Conversations



Source: Data collected by the author.

1.6.4. Additional Research Methods

The method of DSR, combined with the Turning test for conversations, serves as the overarching method for the thesis. Additionally, in chapter 3, further in-depth methods are applied. These include qualitative expert discussions, quantitative experiments with the aid of prototypes, as well as a novel quantitative survey technique wherein the interview partners assume the role of a conversational AI system.

For the qualitative expert discussions, experts from relevant research and practice fields are evaluated and invited to a joint, guided discussion. The relevant research or practice fields are determined by the RQ. The evaluated experts are then contacted via email or LinkedIn and invited to a joint discussion round. These discussion rounds typically take place online. The moderator, in this case the authors of the work, creates a discussion guide in advance, which serves as a framework for the guided expert discussion.

In the case of quantitative experiments with the aid of prototypes, these are laboratory experiments where participants are first asked to test a prototype and then complete a questionnaire. The participants are selected and invited based on criteria defined by the author. Prototypes serve as simplified representations of a planned final

product. The questionnaires are created by the author and are mostly based on already validated scales from other researchers in this work.

Additionally, I apply a novel quantitative survey technique, which is similar to the prototype method. However, here the participants do not receive a finished prototype, but are instead asked to continue developing a prototype themselves. With finished prototypes, the interview participants can only express their opinion on something that already exists. In this method, the participants must become creative themselves and complete the implementation of the prototype.

All these methods are explained separately in more detail in chapter 3.

1.7. Author's Own Research Contribution

For the preparation of this thesis, I not only considered works by other researchers, but also my own previous research which, combined together, provide a suitable basis. The following are the most important contributions that will be used in the later course of this thesis:

- Hundertmark, S., Fill, H.-G., & Portmann, E. (2022a). From practice to science: Social reactions caused by conversational agents—A literature review for researchers and practitioners with a chatbot as an outcome. Proceedings of the International Conferences on Applied Computing 2022 and WWW/Internet 2022, Portugal, 156-164. https://www.computing-conf.org/wp-content/uploads/2022/11/1_ICWI2022_F_039.pdf

In this work a conceptual model linking Conversational Agent design cues with social signals and reactions was developed based on a systematic literature review. The model helps researchers understand CA design influences and provides practitioners with guidelines to design CAs that align with business objectives and optimize user interactions.

- Hundertmark, S., Grüter, A., & Portmann, E. (2022b). Mit Chatbots die finanzielle Allgemeinbildung stärken. Informatik Spektrum, 45(2), 80–87. <https://doi.org/10.1007/s00287-022-01450-7>

This study aims to improve young people's financial literacy using chatbot-based training. Research indicates that 5–10 minute chatbot sessions, completed 2–3 times per week, enhance financial knowledge and are perceived as being enjoyable. The chatbot's tone should be simple, humorous and media-inclusive. Participants value anonymity during these sessions.

- Hundertmark, S., & Portmann, E. (2022c). Ethics Check for Conversational AIs. Proceedings of the ACM CHI 2022 conference, USA. <https://cui.acm.org/workshops/CHI2022/papers.html>

In this work an ethics check for chat- and voicebots is developed, inspired by the Turing test. The goal is to identify ethically correct bots and provide optimization tips for those that are not. First, they define what “ethically correct” means within their language and culture. Next, a measurement criteria for these ethical standards is established. Each bot then undergoes an assessment to evaluate its ethical correctness and identify areas for improvement. Initially, humans will conduct the assessments, but the team aims for bots to eventually perform these evaluations autonomously.

- Hundertmark, S., Hafner, N., & Portmann, E. (2023b). Dialog-Anpassung an den Charakter des Benutzers – Anwendung der Methoden des Charakter Computing. *Informatik Spektrum*, 46(4), 210-219. <https://doi.org/10.1007/s00287-023-01547-7>

This research demonstrates that, while CC enhances personalized customer experiences in financial conversational AIs, it is insufficient alone. The paper develops rules for its application through experiments and suggests integrating methods such as conversational theory and multidimensional fuzzy classifications for improved effectiveness.

- Hundertmark, S., Portmann, E., & Hafner, N. (2023c). Fuzzy Conversational Character Computing. *Informatik Spektrum* 46(3), 163–175. <https://doi.org/10.1007/s00287-023-01538-8>

This research integrates psychology and business informatics to enhance Conversational Bots, enabling personalized responses. By combining CC with fuzzy logic and conversational theory, the study advances speech and emotion recognition while maintaining data efficiency and ethics, moving beyond traditional bots’ limitations in understanding user emotions and personality.

- Hundertmark, S., & Hafner, N. (2023a). Gestaltung automatisierter Kundendialoge im Spannungsfeld von Kosten und Customer Experience. In M. Bruhn, & K. Hadwich (Eds.), *Gestaltung des Wandels im Dienstleistungsmanagement: Band 2: Kundenperspektive – Anbieterperspektive – Mitarbeiterperspektive* (pp. 219–238). Springer Gabler. https://doi.org/10.1007/978-3-658-41815-1_9

This work explores automating customer service to enhance customer experience, focusing on positive emotions throughout the customer journey. The authors examine aligning automation with the needs of both customers and businesses. Due to insufficient prior research, they developed three hypotheses based on a literature review on bots and emotions and tested them using an experiment.

- Hundertmark, S. (2023). Be the Bot – Versicherungskunden in der Rolle eines Chatbots. In N. Hafner & S. Hundertmark (Eds.). *Kundendialog-Management:*

Wertstiftende Kundendialoge in Zeiten der digitalen Automation (pp. 189–200). Springer Gabler. https://doi.org/10.1007/978-3-658-42851-8_12

This book chapter explores how chatbot dialogue design impacts user satisfaction and technology success. Using a novel method where users write chatbot responses, it reveals that different customer groups have varied expectations for chatbot interactions. The study advises companies to tailor chatbot dialogues to specific user needs and use cases.

2. Theoretical and Methodological Foundation

The theoretical foundation of the FCCC framework is not a collection of independent ideas but a constructed framework in which each theory resolves a limitation left open by its predecessor. This chapter presents the most important theories in the order in which they logically depend on one another, so that the reader can follow the chain of reasoning that motivates the FCCC integration.

The starting point is systems theory (von Bertalanffy, 1949), which establishes the overarching architectural principle. A conversational AI that self-regulates through feedback loops. Systems theory explains how the framework is organized, but not what it optimizes for or how it handles the inherent imprecision of human language. Conversational theory (Pask 1965; Pangaro, 2010) answers the first question by defining the feedback mechanism concretely as a structured conversational loop in which both parties build shared understanding over time. Fuzzy logic (Zadeh, 1965; Zumstein, 2026; Werro, 2015) answers the second question by providing the mathematical language for representing partial, overlapping, and linguistically expressed states. Character Computing (El Bolock et al., 2020) then specifies which dimensions of the user the fuzzy models must capture. The triad of stable personality, dynamic affective state, and situational context. Finally, Life Engineering (Österle, 2020) provides the normative horizon: it defines what the self-regulating, fuzzy, character-aware system should ultimately achieve for the people it serves, and against which standard it should be evaluated.

These five theories are presented in detail in the following sections. Section 2.6 synthesizes their integration into the FCCC concept, before Section 2.7 embeds the framework in its business context.

2.1. Conversational AI: Definition, Capabilities and Evolution

Conversational AI systems are sophisticated technological frameworks developed to enable seamless natural language communication between humans and machines (Brabra et al., 2022).

These systems include various software agents, such as chatbots, voice assistants and virtual agents, which leverage AI techniques for context-aware language interpretation, generation and response (Brabra et al., 2022). Unlike traditional rule-based systems, conversational AI employs machine learning, natural language processing (NLP) and fuzzy logic to enable dynamic, adaptive interactions (Brabra et al., 2022).

The theoretical foundation of conversational systems is rooted in several interdisciplinary fields, including linguistics, computer science and cognitive psychology. At the core, these systems aim to replicate human-like conversations by leveraging advancements in NLP, machine learning and AI (Brabra et al., 2022).

The early evolution of conversational systems can be traced back to ELIZA, a pioneering rule-based program developed in the 1960s that simulated human-like dialogue by matching user inputs to predefined responses (Weizenbaum, 1966). In contrast, contemporary systems leverage advanced models, such as IBM's Watson and OpenAI's GPT-3, which demonstrate remarkable capabilities in understanding and generating contextually appropriate language. The primary goal of conversational AI systems is to simulate human-like dialogue, allowing users to retrieve information, perform tasks or solve problems in an intuitive way. For example, a customer might use a virtual assistant to enquire about product details or a chatbot to resolve technical issues. Contemporary conversational AI systems excel in maintaining coherence across multi-turn dialogues, accurately interpreting ambiguous language, and tailoring responses to user preferences and contextual nuances (Brabra et al., 2022).

A critical feature of conversational AI systems is their ability to effectively manage and sustain dialogues, a capability central to dialogue management in conversational frameworks. This involves understanding the context of a conversation, maintaining coherence across multiple interactions, and generating responses that are appropriate and relevant. Such capabilities require a deep understanding of linguistic nuances, user intent and contextual awareness (Brabra et al., 2022).

When enhanced with the integration of feedback mechanisms, which is a foundational principle of systems theory (von Bertalanffy, 1949), these systems can dynamically evolve and adapt to user inputs. By dynamically adjusting to user inputs, conversational AI systems exemplify the principles of adaptability, self-regulation and emergent behaviour, making them a vital application area for exploring advanced theoretical frameworks (Brabra et al., 2022).

The definition and capabilities of conversational AI outlined above reveal a fundamental requirement. For a conversational AI to adapt meaningfully across repeated interactions, it must possess a principled mechanism for incorporating feedback from each exchange into its own behaviour. This is not merely a technical feature to be added at the implementation stage. It is the architectural logic that determines whether the system can improve at all. The theory that formalizes this logic is systems theory, which is introduced in the following section.

2.2. Understanding Systems: How Conversational AI Evolves in a Connected World

Systems theory, as articulated by Ludwig von Bertalanffy (1949), serves as the foundational framework for this thesis. It provides a transdisciplinary lens to analyse systems as coherent, interconnected entities characterized by dynamic feedback mechanisms. This overarching framework informs and integrates the subsequent theories of conversational loops, fuzzy logic and its related theories, and CC, ensuring a

cohesive narrative that emphasizes self-regulation, adaptability and the management of complexity in conversational AI systems (von Bertalanffy, 1949).

Systems theory uses feedback loops known from the field of cybernetics as a central functionality. In cybernetics, the focus on feedback and control processes plays a significant role in the design and improvement of systems. Feedback from users serves as an input that helps these systems adjust and enhance their conversational abilities, ensuring that they remain relevant and effective in various contexts (Wiener, 1948).

In systems theory, feedback loops are foundational for understanding how systems interact with their environment and achieve self-regulation (von Bertalanffy, 1949). These mechanisms function by allowing a system to continuously process external inputs, adapt its internal structure or behaviour, and produce outputs that align with dynamic environmental changes. Feedback loops are categorized into two main types: positive and negative. Positive feedback amplifies deviations, enabling growth or transformation within the system, while negative feedback counteracts deviations, maintaining stability and equilibrium (von Bertalanffy, 1949).

For example, in conversational AI systems, a feedback loop is initiated when a system receives user input, processes it and generates a response. This response, in turn, becomes a new input for the user, creating an iterative cycle of adaptation and refinement. These iterative processes ensure the system evolves dynamically, learning from each interaction to improve performance and relevance. The duality of positive and negative feedback is also present: positive feedback might encourage a system to explore new conversational patterns or features, while negative feedback ensures it aligns with user expectations and corrects errors.

Feedback loops also encompass concepts of time sensitivity and recursion. In real-time systems, immediate feedback is crucial for maintaining responsiveness, whereas delayed feedback might support strategic adjustments over longer durations. These varying temporal scales of feedback loops allow conversational AI to address both immediate user needs and longer-term optimization of interaction quality.

This foundational understanding of feedback mechanisms underpins all elements of this thesis, offering a robust framework for analysing and designing conversational AI systems that evolve dynamically in response to user interactions. By incorporating adaptive feedback loops, conversational AI systems not only respond to changing conditions, but also align with the principles of self-organization and emergent behaviour central to systems theory.

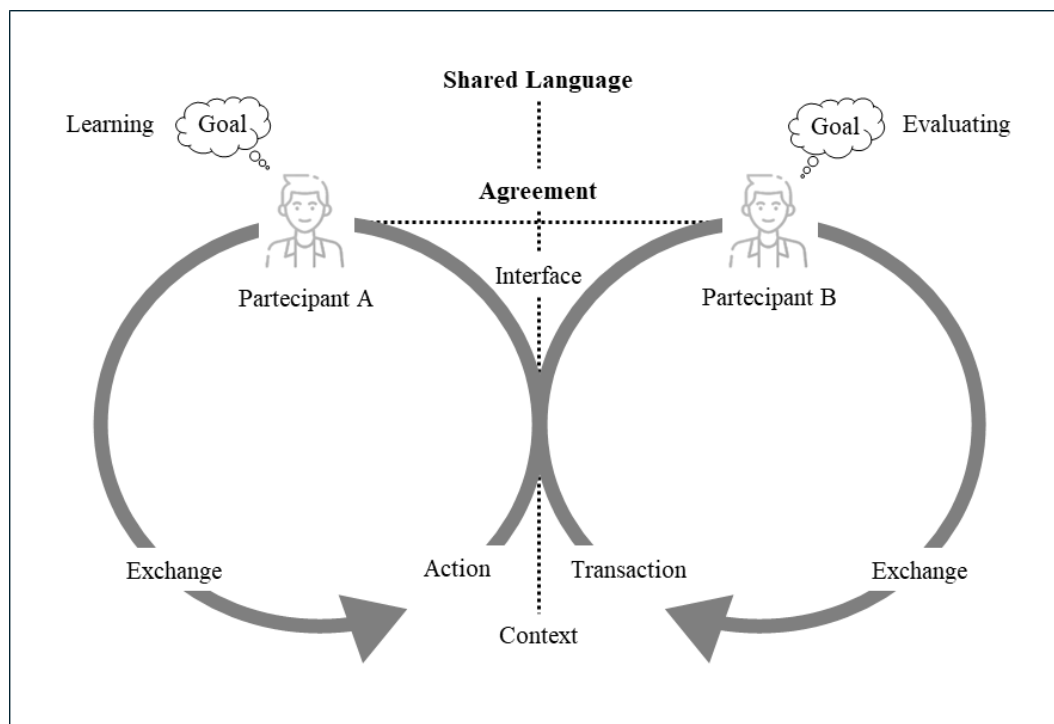
Systems theory establishes that feedback loops are the mechanism through which a system self-regulates. However, it does not specify the form that these loops must take in a conversational context. A conversation is not a simple input-output cycle. It is an iterative exchange in which both parties co-produce meaning. The theory that models this specific structure is Pangaro's (2010) conversational theory,

which operationalizes the abstract concept of a feedback loop as a structured conversational turn with concrete properties direction, accumulation of shared understanding, and a testable quality criterion. The following section introduces this theory and explains how it extends the systems-theoretic foundation for the FCCC Framework.

2.3. Learning from Interactions: How AI Adapts through Feedback Loops

Conversational loops, a core element of Pask’s conversational theory (1965), represent iterative exchanges between participants aimed at building mutual understanding and agreement. Conversational theory emphasizes that communication between two or more entities is not merely an exchange of information, but a dynamic process where participants adapt to each other to achieve a shared goal or innovation. These loops inherently align with systems theory’s feedback mechanisms, where the outputs of one cycle become the inputs for the next, enabling continuous learning and adaptation within systems (Pangaro, 2010).

Figure 6: Conversational Loops in Process



Source: Based on Pask (1965) and applied by Pangaro (2010).

Central to conversational theory is the notion of mutual adaptation. Each participant, whether human or machine, functions as a learning system that evolves through repeated interactions. These interactions are structured into conversational loops, which facilitate the exchange of information, adjustments in behaviour and, ultimately, alignment of understanding. Such loops provide conversational AI systems with the ability to learn iteratively from user inputs, adapt their responses dynamically, and refine their models to better serve user needs (Pangaro, 2010).

For example, in the context of customer service, conversational loops enable a chatbot to detect dissatisfaction in a user's tone or phrasing, and adjust its responses accordingly. If the user states, "This isn't what I was looking for", the system processes this input, analyses the mismatch, and generates a clarifying or alternative response. This iterative adjustment not only resolves the immediate issue, but also strengthens the system's capacity to handle similar situations in the future. By embedding these adaptive conversational loops, the system mirrors the principles of systems theory's feedback mechanisms, ensuring it remains responsive and evolves over time (Pangaro, 2010).

In a conversational AI context, the system and the user operate as interdependent entities engaged in dynamic information exchange. Explicit feedback, such as a user stating, "This didn't answer my question", or implicit feedback, such as hesitations or tonal shifts, serves as vital inputs for the system. These inputs inform the system's subsequent actions, such as modifying tone, refining responses or providing additional context. Through this cyclical process, the system embodies systems theory's principle of adaptive feedback loops, allowing it to respond dynamically to its environment and maintain equilibrium in the dialogue (Pangaro, 2010).

For instance, a customer service chatbot interacting with a frustrated user might detect emotional cues from word choice, pauses or tone. In response, it could adjust its communication style, adopting a more empathetic tone while providing targeted solutions. This real-time adjustment demonstrates how conversational loops operationalize systems theory, ensuring the system's outputs are continually refined to meet evolving user needs. Moreover, these loops form the foundation for the system's ability to evolve over time, incorporating lessons from each interaction into a progressively optimized model of engagement. By doing so, conversational loops not only enhance the user experience, but also exemplify the practical application of systems theory in creating responsive and adaptive AI systems (Pangaro, 2010).

Conversational loops provide the iterative mechanism through which the framework learns. What they do not resolve is the problem of representation. The raw material that flows through each loop. User language is inherently imprecise. A user who says 'I am a bit frustrated' or 'I prefer something not too long' is not providing a numerical value that a classical binary system can process unambiguously. If the feedback loop is to refine a model of the user with each turn, that model must be capable of representing partial, overlapping, and linguistically expressed states. The mathematical framework that makes this possible is fuzzy logic, introduced in the following section.

2.4. Navigating Uncertainty: Introduction of Fuzzy Logic in Conversational AI Systems

While conversational loops operationalize feedback processes, fuzzy logic provides a mechanism to address the inherent ambiguities and uncertainties of natural language interactions. Introduced by Zadeh (1965), fuzzy logic extends classical binary logic by allowing for degrees of truth, making it indispensable for modelling the nuances of human communication.

Fuzzy logic solves the lack of intermediate values typical of classical logic (i.e. “shades of grey”). It represents an extension of two-valued logic, where $fA(x)$ can be true or false or can have an intermediate truth value (Zadeh, 1965). Such an intermediate truth value is an element of a finite or infinite set T of possible truth values, spanning a finite or infinite number of fuzzy subsets within T . The values range, for example, from true, through neither true nor false, to false with countless further intermediate values. These truth values are an imprecise characterization of numerical truth values (Zadeh, 1965).

Fuzzy logic allows for the representation of partially true statements, in contrast to traditional logic where a statement is either completely true or false. For example, the degree of membership of an open-minded man to the set of all men can be represented by the membership degree $\mu(\text{open}) = 0.3$, meaning that a person being only slightly open-minded is only partially (e.g. 30%) a member of the set of all men. This partial membership can be then represented with the membership function, which is explained in the following (D’Onofrio et al. , 2015).

Conversational AI systems that apply fuzzy logic methods are capable of processing not only precise and clearly defined user statements, but also handling ambiguous or vague expressions. For instance, if a user states, “I did not fully understand that”, the conversational AI system can interpret this intermediate state between full comprehension and complete misunderstanding. Without the implementation of fuzzy logic, such ambiguous statements are difficult to process effectively.

Fuzzy Logic in Action: Mathematical Fundamentals to Use Fuzzy Logic in Conversational AI

Fuzzy logic provides a mathematical foundation for its application in Conversational AI. Unlike classical binary logic, it allows for the gradual membership of elements in categories, enabling the modelling of uncertainties and linguistic nuances. This is particularly relevant for natural language, which often involves vague or imprecise terms. Conversational AI systems that incorporate fuzzy logic can respond more flexibly to user inputs and conduct more human-like dialogues. The following mathematical principles illustrate how fuzzy logic enhances AI-driven dialogue systems.

Membership function

The membership function $\mu_{\tilde{A}}(x)$ gives the degree of membership of an element x to a fuzzy subset $\tilde{A}$ in X ; this degree is expressed by a real number between 0 and 1 (i.e. from no membership to full membership) (Zadeh 1965).

$$\tilde{\mu}A(x) \rightarrow (0,1)$$

For example, if someone asks the person whether he is open-minded, the man would likely use various linguistic variables to express his opinion, such as absolutely, a little, rather no, or not at all. Humans can process such statements, while computer systems obeying strict logic have difficulties (O'Regan, 2008). To evaluate the impression by the man, the system which should define the state of the man can also form fuzzy subsets, such as open-minded $\tilde{A}$ or closed-minded $\tilde{B}$. Using fuzzy set theory (Zadeh, 1965), the different perceptions of people can be classified and help assess whether the decision is justified.

For illustration: if the membership functions of the open-minded man (i.e. x) are $\mu_{\tilde{A}}(x) = 0.3$ and $\mu_{\tilde{B}}(x) = 0.7$, the system knows that the person is classified as being only slightly open-minded (D'Onofrio et al. , 2015).

Fuzzy number and intervals

Another essential concept in fuzzy systems includes fuzzy numbers and fuzzy intervals, which allow for the approximation of concepts and the use of linguistic variables that replace numerical values with words or phrases (D'Onofrio et al. , 2015):

- A fuzzy number $*x$, $x \in \mathbb{N}$ is a fuzzy set that represents an approximate value a , for which the membership function decreases on both sides of the number a .
- A fuzzy interval $[*a, *b]$ is an interval whose boundaries are marked by the fuzzy numbers $*a$ and $*b$. Membership in the interval is maximal between a and b , but gradually decreases outside this interval.

Handling Complexity: Classifying Information

Fuzzy classifications were applied in a management context by Zumstein (2007), whose work was later published as a chapter in Portmann et al. (2026). In his work Zumstein demonstrated how multidimensional fuzzy classifications can replace traditional crisp segmentation methods in Customer Relationship Management (CRM), enabling more nuanced and fair classification of complex objects, in his case, customers. Building directly on Zumstein's work, Werro (2008) developed a full doctoral thesis that operationalized hierarchical fuzzy classifications for online customer management, subsequently published as a Springer monograph (Werro, 2015) and as a chapter in Portmann et al. (2026). Both works share a common principle. Allowing objects to belong to multiple classes simultaneously, with graded membership degrees rather than rigid binary assignments. The present work transfers this same principle

to the domain of conversational AI, rather than classifying customers by turnover and purchase behaviour, this work later aims to classify conversational AI users by their personality dimensions, affective state and interaction behaviour.

In conversational AI systems, fuzzy logic enhances the system's capacity to interpret user inputs with variable precision. For instance, a user's statement such as, "I'm somewhat satisfied" can be processed as a degree of satisfaction rather than a binary outcome. This granularity enables the system to deliver more tailored responses, aligning with systems theory's emphasis on adaptability and self-regulation.

Furthermore, fuzzy classifications enable conversational AI systems to manage overlapping categories and handle uncertainty effectively by allowing each input to belong to multiple categories with varying degrees of membership. This is achieved using membership functions, where each category is defined by a range of values that overlap with others. For example, when determining a user's intent, the system might compute that the input belongs 70% to one category (for example "enquiry") and 30% to another (for example "complaint") (D'Onofrio et al., 2015).

Such classifications utilize fuzzy logic operations such as union, intersection and complement to refine these memberships dynamically. The nuanced understanding resulting from this process enhances the system's ability to interpret ambiguous inputs and respond contextually. For instance, if an input falls between two categories, the system can aggregate the weighted degrees to produce a response that acknowledges both aspects, fostering more accurate and context-sensitive interactions. This approach not only improves the adaptability of conversational AI, but also aligns with the principles of a self-regulating system capable of evolving over time (D'Onofrio et al., 2015).

Zumstein (2007; 2026) extends classical fuzzy models by defining a fuzzy classification database scheme that determines the degrees of membership of objects, in his case, customers across multiple classes simultaneously. Formally, a linguistic variable is characterized by a quintuple (X, T, U, G, M) , where X is the name of the variable, T the set of terms, U the universe of discourse, G a syntactic rule for generating the names of the terms, and M a semantic rule associating each term with a fuzzy set on U (Zadeh, 1975; Zumstein, 2026). Linguistic variables replace numerical values with words or word combinations, enabling a more intuitive description of equivalence classes. Crucially, the resulting fuzzy partition eliminates sharp class boundaries and replaces them with continuous transitions: after normalization, an object's membership degrees across all classes sum to 1.00. For the aggregation of membership degrees across multiple linguistic variables, Zumstein (2026) employs the γ -operator, a compensatory aggregation operator empirically tested by Zimmermann and Zysno (1980), which allows partial compensation between attributes rather than requiring all dimensions to be simultaneously fulfilled:

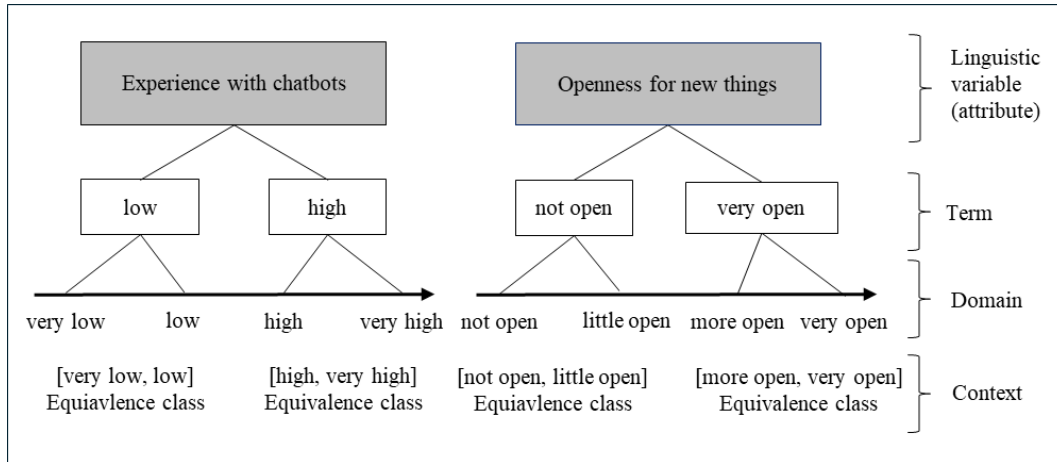
$$\mu_{A,comp}(x) = \left[\prod_i \mu_i(x) \right]^{(1-\gamma)} \cdot \left[1 - \prod_i (1 - \mu_i(x)) \right]^\gamma, \quad \gamma \in [0, 1]$$

The γ -value determines the degree of compensation: $\gamma = 0$ corresponds to the algebraic product (strict conjunction), while $\gamma = 1$ corresponds to the algebraic sum (full compensation). This flexibility makes the operator particularly suited for modeling human-like weighing of multiple attributes. A property directly exploited in the FCCC framework when aggregating personality sub-dimensions into a composite character vector.

Linguistic variables are considered and the attributes and verbal terms are assigned to each equivalence class. With the help of linguistic variables, for example words or word combinations, the equivalence classes of the attributes can be described more intuitively (Zumstein, 2007; 2026).

Figure 7 gives an example of a user's character as applied in the present work. The character is described with the linguistic variables "experience with chatbots" and "openness to new things". The first variable is described with the terms "low" and "high", and the second variable with "not open" and "very open".

Figure 7: Concept of Linguistic Variables



Source: Data collected by the author, conceptually based on Zumstein (2007; 2026)

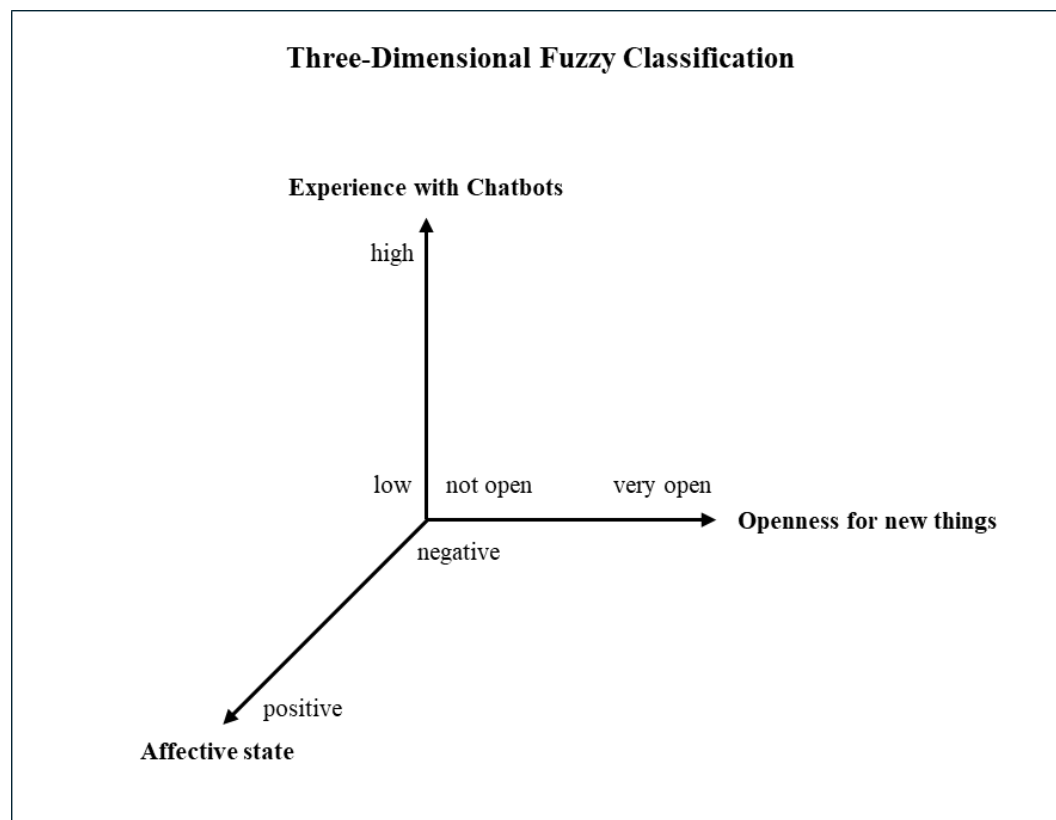
Each term of the linguistic variables represents a fuzzy set determined by a membership function $\mu_{\tilde{A}}(x)$ over the entire range of the corresponding attribute (Zumstein, 2007; 2026). The attributes "experience with chatbots" and "openness to new things" do not contain numerical values but general terms. In this case, the terms of the linguistic variables "low" and "high", and "not open" and "very open" are associated with discrete membership functions, that is, each term corresponds to a discrete value. With the help of the context model, the linguistic variables and the membership functions, the classification space becomes fuzzy. This fuzzy partitioning has an important result, namely the disappearance of sharp boundaries between classes, which

means there are continuous transitions between the different classes (Werro, 2015; 2026).

When it comes to determining the character of an individual person, several dimensions usually must be taken into account. Therefore, the application of multi-dimensional fuzzy classifications, which is not limited to two dimensions, is chosen. With multi-dimensional fuzzy classifications, several attributes or criteria can be considered simultaneously (Zumstein, 2007; 2026).

Figure 8 adds the third dimension, "affective state", with the two terms "positive" and "negative". In a sharp classification, this multi-dimensional classification of users always puts users in the same category even though their characters differ from each other. In contrast to the fuzzy classification used, the sharp boundaries disappear, and the three-dimensional classification space becomes fuzzy.

Figure 8: Three-Dimensional Classification



Source: Data collected by the author, conceptually based on Zumstein (2007; 2026) and Werro (2015; 2026).

Werro (2015; 2026) extended Zumstein's ideas on multidimensional fuzzy classification into a fully operational, hierarchical system for online customer management. His central contribution is the principle of hierarchical decomposition. Instead of combining all relevant attributes in a single, unwieldy classification space, related attributes are first grouped into sub-classifications that express intermediate concepts.

These intermediate concepts are then combined at a higher level to derive more abstract, semantically richer constructs. In Werro's CRM application, for instance, the attributes turnover and margin are first combined to derive the concept of profitability; separately, buying frequency and attachment are combined to derive loyalty. Both concepts are then aggregated at the top level into the customer lifetime value (Werro, 2015). The FCCC framework adopts precisely this hierarchical principle. Observable conversational attributes (e.g. response length, question frequency, sentiment markers) are first aggregated into intermediate constructs such as openness or agreeableness, which are then combined into a composite character vector that guides adaptive response generation.

A further technical element transferred from Werro (2015) is the concept of personalized value. Werro defines the personalized value $V(O_i | C)$ of an object O_i across a set of fuzzy classes $C = \{C_1, \dots, C_n\}$ as:

$$V(O_i | C) = \sum M(O_i | C_k) \cdot gr(C_k)$$

where $M(O_i | C_k)$ is the membership degree of object O_i in class C_k , and $gr(C_k)$ is a grade assigned to that class expressing a target concept (e.g. a loyalty score or a discount rate). In Werro's application, this formula is used to calculate personalized customer discounts. In the FCCC framework, the same mathematical structure can be used to compute an adaptive response profile from the user's multi-class personality membership degrees, weighting each personality class by its relevance to the current interaction context.

Human Language as the Key-Input: Learning from Words

Computing with Words (CWW), introduced by Zadeh (1999), revolutionizes the way systems process and interpret human language by treating words as fundamental units of computation. This approach contrasts with traditional numerical or binary data processing by leveraging fuzzy logic to quantify the inherently imprecise and subjective nature of linguistic terms. As human communication often relies on vague expressions such as “a little”, “mostly” or “highly likely”, CWW provides a mechanism for machines to interpret, process and respond to these terms in a manner that aligns with human reasoning.

At its core, CWW enables systems to handle linguistic imprecision by transforming qualitative expressions into mathematical representations. These are then analysed and manipulated using fuzzy sets and inference mechanisms (Zadeh, 1999). This capability makes CWW particularly valuable for conversational AI, where user inputs are frequently ambiguous or context-dependent.

Within the framework of systems theory, CWW enhances the adaptability and feedback mechanisms of conversational AI systems. Instead of relying solely on numerical inputs, introduces linguistic nuance into the feedback loop, allowing the

system to interpret and adapt to user preferences with greater accuracy. For example, a user describing their experience as “mostly positive” provides fuzzy input that the system can process to fine-tune its future responses, reinforcing the principles of self-regulation and iterative learning.

By enabling conversational AI systems to process and compute words, CWW ensures that feedback loops are not only precise but also contextually rich. This integration aligns with systems theory’s emphasis on dynamic adaptability and supports the development of user-centric, responsive AI systems.

The Mechanics of Computing with Words (CWW)

CWW operates on the principle that words can be mapped to fuzzy sets, which define their meanings in terms of degrees of membership within a specific context (Zadeh, 1999). For instance, a word like “warm” may not have a single fixed definition, but instead represent a range of temperatures. In this model, temperatures between 20°C and 30°C might partially belong to the fuzzy set “warm”, with varying degrees of membership assigned to each value.

To process such linguistic data, CWW relies on three key steps (Zadeh, 1999):

Fuzzification of Linguistic Variables: words or phrases are translated into fuzzy sets using membership functions. For example, the linguistic variable “temperature” might include terms like “cold”, “warm” and “hot”, each defined by overlapping membership functions that assign a degree of truth to specific numerical values.

Fuzzy Inference Mechanisms: once inputs are fuzzified, inference rules are applied to evaluate relationships between linguistic variables. For instance, a rule might state: if temperature is warm and humidity is high, then comfort level is low. These rules use logical operators like AND, OR and NOT to combine multiple fuzzy inputs into a meaningful output.

Defuzzification: to produce actionable results, the fuzzy outputs are translated back into crisp numerical values or clear linguistic recommendations. One common method for defuzzification is the centroid approach, which computes the “centre of gravity” of the fuzzy output set to determine the final value.

Mathematically, the methodology of CWW is grounded in fuzzy set theory, where linguistic variables are formally represented by fuzzy sets defined on appropriate universes of discourse. The process of CWW can be described as the following (Colombo, 2022).

Fuzzification of Linguistic Variables

Given a linguistic variable X , it is defined as a quintuple:

$$X = (L, T(L), U, G, M)$$

- L : The name of the linguistic variable (e.g., "temperature").

- T(L): The set of linguistic terms that characterize L (e.g., "cold," "warm," "hot").
- U: The universe of discourse over which these terms are defined (e.g., temperatures ranging from 0°C to 40°C).
- G: A syntactic rule for generating the terms in T(L).
- M: A semantic rule associating each term with a fuzzy set on U.

For example, the fuzzy set for the linguistic term "warm" might be defined as:

$$Warm = \{ (u, \mu_{\{Warm\}(u)}) \mid u \in U \}$$

where $\mu_{Warm}(u)$ is the membership function expressing the degree to which each numeric value u in the universe of discourse U (e.g., 20°C to 30°C) belongs to the fuzzy set "warm". This membership function typically follows a Gaussian, triangular, or trapezoidal shape depending on empirical data or expert knowledge.

Fuzzy Inference Mechanisms

Once fuzzified, linguistic expressions are processed using fuzzy logic inference rules. Consider a simple rule-based inference system containing a set of n rules:

$$R_i: IF X_1 \text{ is } A_{i1} \text{ AND } X_2 \text{ is } A_{i2} \text{ THEN } Y \text{ is } B_i, (i = 1, 2, \dots, n)$$

- X_1, X_2 are input linguistic variables.
- Y is the output linguistic variable.
- A_{ij} and B_i are linguistic terms represented by fuzzy sets with membership functions $\mu_{A_{ij}}(x_j)$ and $\mu_{B_i}(y)$.

The degree of activation (firing strength) of rule R_i for given crisp inputs x_1, x_2 is calculated using a t-norm (typically the minimum operator) as follows:

$$\alpha_i = \min(\mu_{A_{i1}}(x_1), \mu_{A_{i2}}(x_2))$$

The fuzzy inference process aggregates the results from all rules, yielding an aggregated fuzzy output set B' :

$$\mu_{B'}(y) = \max_{i=1, \dots, n} (\min(\alpha_i, \mu_{B_i}(y)))$$

Defuzzification

To obtain a crisp actionable result from the aggregated fuzzy set B' , a defuzzification method is applied. The centroid method is one of the most widely used and mathematically defined as:

$$y^* = \frac{\int y \cdot \mu_{B'}(y) dy}{\int \mu_{B'}(y) dy}$$

Here, y^* represents the crisp output value calculated as the center of gravity of the aggregated fuzzy set.

Practical Example, according to Colombo (2022)

Consider a fuzzy rule-based system evaluating comfort levels based on temperature and humidity. Given input data (Temperature: $X_1=25^\circ\text{C}$, Humidity: $X_2=80\%$), the fuzzy inference might look like:

- IF temperature is "Warm" ($\mu_{\text{Warm}}(25)=0.8$) AND Humidity is "High" ($\mu_{\text{High}}(80)=0.9$), THEN Comfort is "Low".
- The degree of activation would be $\alpha=\min_{f_0}(0.8,0.9)=0.8$
- After applying inference to all rules and aggregating outputs, defuzzification yields a crisp comfort level (for example: comfort score = 2.3 on a predefined comfort scale from 0–10).

By combining these steps, CWW bridges the gap between human-like reasoning and computational precision, allowing conversational systems to interpret and respond to nuanced inputs effectively.

From Customer Management to Conversational AI: The Transfer of Fuzzy Classification Methods

The application of fuzzy classification methods to conversational AI follows a line of progressive conceptual extension across two works. Zumstein (2007; 2026) established that replacing crisp class boundaries with graded membership produces fairer, more precise and more actionable results in customer performance management; his case study demonstrated that the γ -operator enables a compensatory, human-like weighting of customer attributes that rigid threshold-based systems cannot replicate. Werro (2015) then operationalized this approach technically by combining relational database architecture with fuzzy logic and introducing the hierarchical decomposition principle, thereby demonstrating the scalability and practical deployability of fuzzy classification in large online retail environments. In all two cases, the underlying mathematical machinery, the quintuple definition of linguistic variables, the γ -operator for aggregation, and the personalized value formula for deriving actionable outputs from membership degrees, remains the same, providing a mathematically rigorous and empirically validated foundation.

Fuzzy logic and its extensions, multidimensional classification (Zumstein, 2007, 2026) and Computing with Words (Zadeh, 1999), provide the tools for handling ambiguous user inputs. But they do not specify what must be modelled. A fuzzy membership function can represent any continuous attribute. It has no inherent concept of a user's personality, emotional state, or the situational context that shapes their expectations. The theoretical framework that supplies these dimensions and that answers the question of which aspects of the user a conversational AI must understand in order to adapt meaningfully is Character Computing, which is explained in the next section.

2.5. More than Language: Introduction of Character Computing to Understand User Expectations

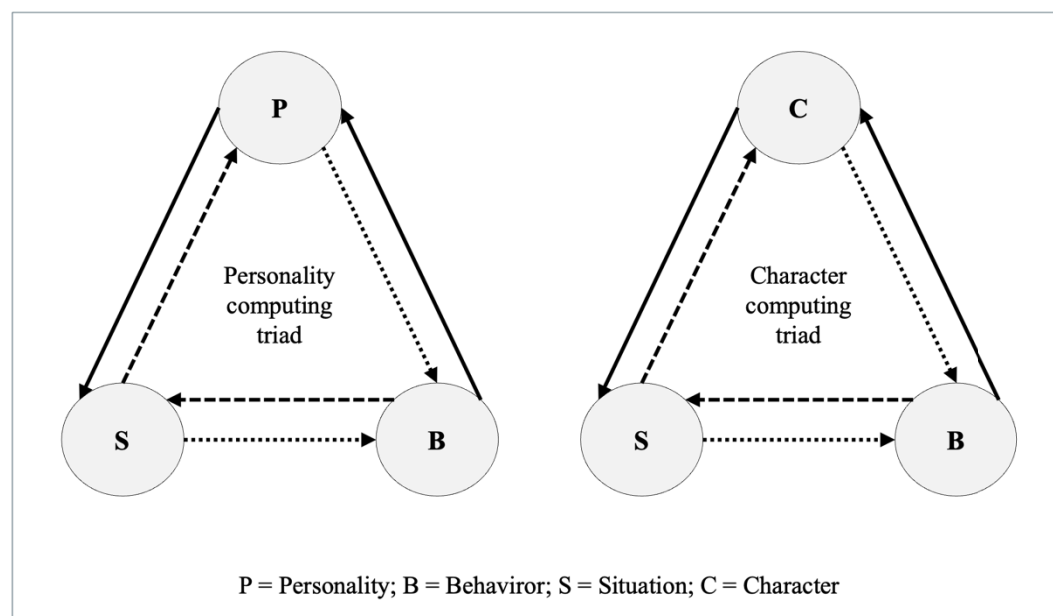
Character Computing (CC) is a relatively new field that integrates principles from psychology and computer science to model and predict user behaviour in dynamic interactions. It builds on systems theory by focusing on the user's personality, emotions and behaviours as interrelated components that influence interaction outcomes (El Bolock et al., 2020). CC conceptualizes human behaviour through the lens of a triad composed of three interrelated elements: character (C), behaviour (B) and situation (S). These components interact dynamically to shape observable actions and responses. At its core, this approach recognizes that human behaviour cannot be fully understood or predicted without considering these interconnected factors. The triad serves as a foundation for creating intelligent systems capable of understanding, predicting and adapting to human needs in a more personalized and dynamic manner (El Bolock et al., 2020).

Character Computing in Conversational AI: Important Fundamentals of Character Computing

CC is based on the triad of character, behaviour and situation (figure 9). The three main aspects of the triad are explained in the following.

C=observe (B,S), B=predict (C,S) and S = understand (C, B) (El Bolock et al., 2020).

Figure 9: Character Computing Triad



Source: According to El Bolock et al. (2020).

Character (C)

Character represents the intrinsic essence of an individual, combining stable and transient attributes. On the one hand, stable personality traits, such as extroversion,

agreeableness and openness, act as enduring markers of behavioural tendencies. These traits are often modelled using established psychological frameworks, such as the Five-Factor Model (FFM) (Allport et al., 1936). On the other hand, character also includes transient states, such as emotions and moods, which fluctuate in response to external stimuli or internal reflections.

The integration of stable and dynamic aspects allows CC to form a holistic understanding of an individual. For example, while a generally optimistic person may demonstrate resilience, their emotional response in a moment of high stress may reveal vulnerability. Systems leveraging CC can account for both dimensions, thereby providing nuanced responses (El Bolock et al., 2020).

Behaviour (B): Observable Actions

Behaviour reflects the observable manifestations of character in action. This includes not only overt actions, but also verbal and non-verbal cues, such as language use, tone, gestures or facial expressions. Behaviour serves as a bridge between the internal world of character and the external environment, providing measurable data for systems to interpret (El Bolock et al., 2020).

Situation (S): Contextual Influence

Situations encompass the external environment and context in which behaviour occurs. This includes social dynamics, cultural factors, physical surroundings and specific events. The situation acts as a catalyst, shaping how character translates into behaviour. For example, the same individual might respond differently in formal versus informal settings, even with consistent underlying personality traits (El Bolock et al., 2020).

Besides the triad, the concepts of dynamic interactions and personality as well as emotional states are also essential for CC and these are therefore described in the following.

Dynamic Interactions in Character Computing

CC formalizes the relationships between character, behaviour and situation mathematically. By leveraging the interdependence among these elements, systems can infer one component when the other two are known. For instance, understanding a user's observed behaviour (B) within a specific context (S) enables predictions about his emotional state or even personality traits (C). This dynamic adaptability enhances the system's ability to tailor responses and iteratively refine predictions as new data emerges (El Bolock et al., 2020).

Personality and Emotional States in Character Computing

Personality traits provide a stable foundation for CC systems to predict user behaviour across contexts. These traits are relatively unchanging over time, offering a

reliable baseline for system modelling. For example, a conscientious individual is likely to approach tasks methodically, regardless of the situation (El Bolock et al., 2020).

In contrast, emotional states are transient and heavily influenced by immediate stimuli. They require real-time interpretation to adapt effectively. For example, a typically calm individual may exhibit frustration in a stressful scenario. By using techniques like sentiment analysis and contextual modelling, CC systems can detect and respond to such shifts dynamically, ensuring interactions remain empathetic and relevant (El Bolock et al., 2020). Until now, researchers in CC like El Bolock et al. (2020) have relied on sensors and other additional technical instruments in their experiments to estimate users' emotional states as accurately as possible and to carry out the subsequent CC triage calculations. As a result, most previous experiments have been conducted in laboratory settings equipped with the required technical infrastructure.

Taking Advantage of Systems: Enhanced Character Computing

The limitations of CC, which can occur due to a lack of sensors in AI systems, has already been mentioned by El Bolock et al. (2020). Up to now, computer systems have received a multitude of sensor data that are supposed to transmit information about the emotional, cognitive and situational states of the user. These must be interpreted by the computer system and mapped into a suitable character model.

In their work on CC, El Bolock et al. (2020) point out that the previous models of CC should be extended in the future with the help of fuzzy logic (Zadeh, 1965). This could help to achieve the original vision that a computer system can completely adapt to human behaviour (El Bolock et al., 2020).

This section explores foundational elements, highlighting how feedback mechanisms and fuzzy classifications collectively contribute to address the limitations, mentioned by El Bolock et al. (2020).

First, feedback loops, as they are known in systems theory (von Bertalanffy, 1949), are added as an integral to CC's adaptability. These loops enable systems to continuously refine their understanding of users by updating internal models based on new observations. This iterative process mirrors the principles of self-regulation in complex systems.

For example, if a conversational AI observes that a user frequently asks for clarifications, it might infer a preference for detailed responses. Over time, the system adapts its communication style to align with this preference, thereby improving the user experience. Furthermore, historical data from previous interactions is integrated into the system's evolving user profiles, ensuring that the AI becomes increasingly personalized and effective over time.

Feedback loops also allow systems to incorporate subtle patterns that may not be immediately apparent (von Bertalanffy, 1949). For instance, a user's repeated hesitation in specific contexts could indicate underlying uncertainty or a need for reassurance. By recognizing and responding to these cues, the system enhances its rapport with the user.

Second, fuzzy logic (Zadeh, 1965) provides a mathematical framework to address the ambiguity, enabling systems to assign degrees of membership to multiple personality or emotional categories. This is needed since human behaviour and character are inherently complex, with traits often overlapping and existing on a spectrum rather than within rigid categories. For instance, a user might be 70% introverted and 40% agreeable, rather than being strictly categorized as one or the other.

In addition, fuzzy classifications (Zumstein, 2026) allow CC systems to model behaviour more accurately by accounting for the nuances and variability inherent in human traits. These classifications are dynamically updated as new interactions occur, ensuring the system remains reflective of the user's evolving behaviour. For example, if a user frequently provides succinct responses, the system might adjust its output to be more concise while also leaving room for exceptions.

The integration of fuzzy logic into CC enhances its adaptability by replacing rigid models with flexible, context-sensitive ones. This ensures that AI systems can better represent the fluid nature of human behaviour, leading to more realistic and empathetic interactions.

Character Computing completes the psychological specification of what the FCCC Framework must model. Together, the four theories presented so far, systems theory, conversational theory, fuzzy logic, and Character Computing, provide the architectural logic, the feedback mechanism, the mathematical language, and the psychological model of the user. One question remains, however: towards what goal should the entire self-regulating, fuzzy, character-aware system operate? This normative question is answered by Life Engineering (Österle, 2020), which is introduced as part of the synthesis in the following section.

The following section introduces the FCCC concept and explains how all theories work together.

2.6. Bringing It All Together: Developing a Research Model for Better AI Conversations

This section reviews the most important benefits from the above-explained theories, focusing on my holistic approach which enhances systems theory (von Bertalanffy, 1949) with additional methods to enhance existing conversational AI systems. Towards the end of this section, I present my new idea on FCCC.

Conversational AI Learns from Feedback Loops

The integration of systems theory, feedback loops and adaptability offers a cohesive framework for understanding the dynamic interactions within conversational AI systems. By emphasizing self-regulation and emergent behaviours, this framework positions feedback loops as pivotal mechanisms that not only correct errors but also foster innovation. For conversational AI, feedback loops enable continuous learning, allowing systems to respond more effectively to user inputs. This adaptability aligns with the principles of systems theory, where dynamic feedback ensures resilience and evolution in changing contexts (von Bertalanffy, 1949). Such mechanisms are indispensable for refining fuzzy conversational models, ensuring they adapt seamlessly to diverse and complex interactions.

Conversational AI Using Fuzzy Logic to Manage Uncertainty

Natural language is inherently rich in ambiguity and nuance, often defying the rigid binary structures of traditional computational logic. Terms such as “somewhat”, “occasionally”, or “likely” exemplify the variability and subjectivity embedded in human communication. Addressing these challenges requires a sophisticated framework capable of capturing and processing such nuances – where fuzzy logic excels.

Fuzzy logic, introduced by Zadeh (1965), extends classical binary logic by allowing for degrees of truth rather than definitive true-or-false outcomes. Instead of forcing user inputs into fixed categories, fuzzy logic acknowledges that truth can exist on a continuum. For instance, rather than classifying a user’s statement about satisfaction as either “satisfied” or “unsatisfied”, fuzzy logic might assign a degree of satisfaction at 70% – a representation of partial membership in the “satisfied” category.

This flexibility is especially critical in conversational AI, where user expressions often fall between rigid extremes. By incorporating fuzzy logic, systems gain the ability to interpret linguistic input more fluidly, accommodating subtleties that traditional binary logic would overlook.

In conversational AI systems, the integration of fuzzy logic enhances the interpretation of user intent, sentiment and ambiguity. In the following, some examples:

- **Sentiment Analysis:** A user states, “I’m somewhat satisfied with the product”. Using fuzzy logic, the system assigns partial membership to both “satisfied” and “neutral” categories, with membership degrees of, say, 0.7 and 0.3, respectively. This nuanced understanding allows the system to generate a response that acknowledges the user’s mixed feelings and offers tailored suggestions, such as, “Thank you for your feedback! Could you share what could make your experience even better?”
- **Intent Detection:** Users often express intentions with varying degrees of clarity. A phrase like “Maybe I need help with my account” signals uncertainty. Fuzzy logic processes this statement by assigning a high degree of

membership to the “help” category but retains some uncertainty, ensuring the system provides a gentle prompt rather than assuming full intent.

- **Preference Modelling:** In e-commerce, a user saying “I prefer something moderately priced” can trigger fuzzy classifications for pricing categories. Rather than rigidly selecting “cheap” or “expensive”, the system positions recommendations in the middle spectrum, better aligning with the user’s preferences.

Conversational AI Using classifications and Managing Complexity

Fuzzy logic’s true strength lies in its ability to classify overlapping or multidimensional attributes effectively (Zumstein, 2026). This capability is especially useful in scenarios where user input intersects multiple categories. For instance:

- **Multidimensional Character Profiles:** A user can simultaneously be 80% classified as “curious” and 60% as “reserved”, which can guide conversational tone and engagement style.
- **Evolving Contexts:** If a user initially states, “I’m happy with my subscription but wish it had more features”, the system can maintain a dual perspective of both satisfaction and dissatisfaction, dynamically updating its model if subsequent interactions lean toward dissatisfaction.

Conversational AI Using Human Language for Natural Dialogues

CWW, introduced by Zadeh (1999), represents a transformative extension of fuzzy logic. While fuzzy logic provides a numerical framework for interpreting degrees of truth, CWW incorporates linguistic terms and perceptions directly into computational processes. This approach allows systems to process words as fundamental units of computation rather than reducing them to numerical abstractions.

For conversational AI, CWW enables:

1. **Direct Linguistic Processing:** Rather than mapping “somewhat satisfied” to an arbitrary numeric scale, CWW directly evaluates the term using predefined fuzzy sets that reflect its contextual meaning.
2. **Nuanced Decision-Making:** For example, a customer describing a service as “mostly good” might trigger responses that address the “good” aspects while gently probing for areas of improvement. By treating such qualitative inputs as computational entities, CWW facilitates richer, context-aware interactions.
3. **Perception-Based Adaptability:** Unlike purely data-driven models, CWW incorporates subjective perceptions, allowing the system to align its interpretations more closely with human reasoning. For instance, the term “soon” can be contextually adapted (interpreted as “minutes”) in a technical support setting or “days” in a shipping context.

Feedback Loops for Continuous Refinement

CWW strengthens the adaptability of fuzzy conversational systems by integrating feedback loops that process not only explicit corrections but also implicit behavioural cues. These loops enable the system to continuously refine its understanding of linguistic terms over time. This refinement can occur through explicit feedback, for example when a user clarifies a statement such as “I meant very soon,” prompting the system to recalibrate its perception of urgency. It can also arise from implicit feedback: patterns in user engagement, such as frequent returns to specific features, signal the need to adjust and improve the mappings between words and their contextual meanings.

By enabling systems to reason with words and perceptions, CWW provides an essential bridge between human cognition and machine logic. It aligns computational processes with how humans naturally communicate, enhancing trust and engagement in conversational AI.

For instance, in sentiment-rich interactions as with mental health chatbots, CWW can process statements such as “I feel a bit overwhelmed” with greater empathy and precision. The system might interpret “a bit overwhelmed” as indicating mild stress and respond with calming suggestions tailored to the perceived intensity.

Conversational AI Using Character Understanding to Manage Expectations

CC introduces a transformative approach to conversational AI by integrating comprehensive psychological insights into user interactions. Unlike traditional conversational AI systems, which predominantly focus on immediate inputs or transactional dialogues, CC emphasizes the interplay of personality traits, emotional states and situational contexts. By doing so, it bridges the gap between machine responses and human-like understanding, creating interactions that are more aligned with individual user needs and preferences.

In addition, the limitation pointed out by El Bolock et al. (2020), which can occur due to a lack of sensors in today’s systems, must be counteracted. In original CC theory, computer systems have received a multitude of sensor data that are supposed to transmit information about the emotional, cognitive and situational states of the user. These must be interpreted by the computer system and mapped into a suitable character model.

El Bolock et al. (2020) make clear in their work on CC that the previous models of CC should, in future, be extended with the help of fuzzy logic (Zadeh, 1965). This should help to achieve the original vision that a computer system can completely adapt to human behaviour (El Bolock et al., 2020). I address this gap by no longer considering CC as a standalone method, but rather as part of something broader, integrating it into the enhanced theories of systems theory, as well as fuzzy logic and feedback loops.

At the core of CC lies the triad framework (El Bolock et al., 2020) – Character (C), Behaviour (B) and Situation (S) – which provides a structured, yet adaptable, methodology for understanding user interactions.

- **Character (C):** This component encompasses both stable and dynamic user attributes. Stable traits, such as extroversion, agreeableness or openness (commonly modelled using frameworks like the Big Five Personality Model (Allport et al., 1936)), offer a consistent foundation for predicting general interaction patterns. Meanwhile, transient states, such as emotional fluctuations or mood shifts, allow systems to adapt dynamically. For instance, if a typically calm user displays signs of frustration in a specific situation, the system can adjust its tone to offer reassurance or solutions proactively.
- **Behaviour (B):** Behaviour represents the observable manifestation of a user's character in real-time. This includes not only explicit actions, such as the phrasing of a query or choice of words, but also implicit cues, such as response timing, hesitation or tone. By analysing these behaviours, the system can infer the user's emotional and cognitive state, further enhancing its contextual relevance.
- **Situation (S):** Situations encompass the external and contextual factors influencing a user at a specific moment. These factors can range from environmental variables (e.g. a stressful deadline) to the specific platform or medium of interaction (e.g. a mobile chatbot versus a desktop assistant). Situational awareness allows the system to tailor responses that resonate with the user's current context.

CC enables conversational systems to evolve their understanding of users over time. By maintaining and updating user profiles based on repeated interactions, these systems develop a deeper, longitudinal understanding of the user. This approach ensures that responses are not only contextually appropriate in the moment, but also reflective of broader interaction patterns (El Bolock et al., 2020).

For instance, a user who frequently asks questions about financial planning tools during weekday mornings. By recognizing this pattern, the system could proactively offer suggestions or resources aligned with these preferences at the right time, enhancing both the efficiency and relevance of the interaction. Additionally, this dynamic adaptability enables systems to balance competing priorities. A user exhibiting frustration due to a technical error might initially receive an empathetic response, but if the system identifies recurring patterns of similar issues, it might escalate the interaction to a human agent for more in-depth support.

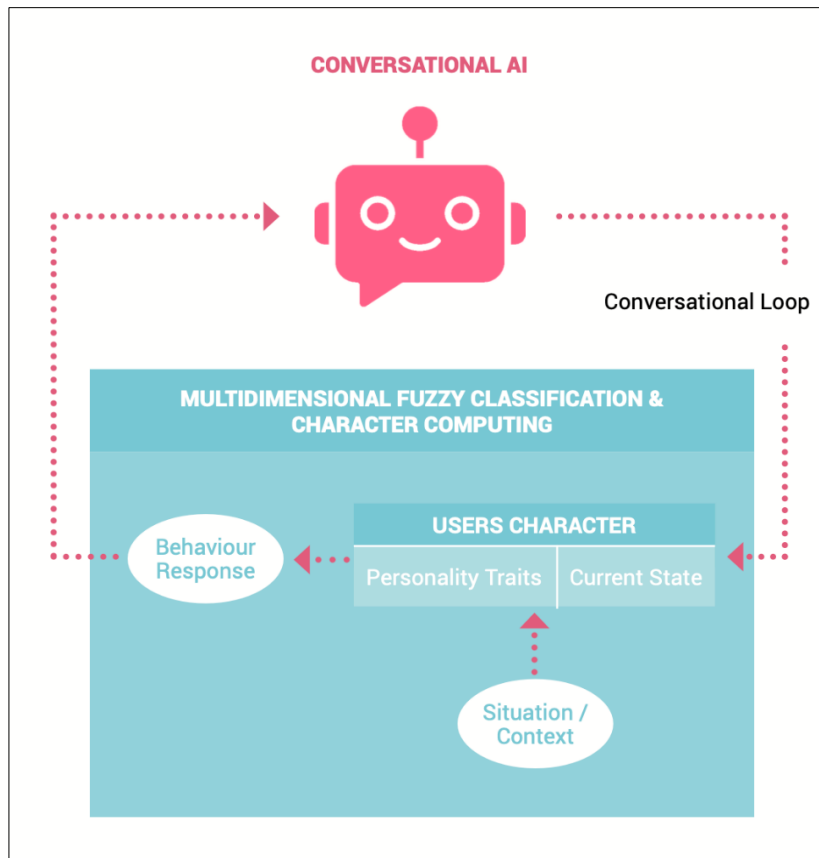
New Perspective: Conversational AI and Fuzzy Conversational Character Computing (FCCC)

FCCC represents a new, holistic principle for (conversational AI) systems, integrating the principles of fuzzy logic, especially CWW and multidimensional fuzzy classifications, CC and conversational loops, using systems theory as its overall orientation. FCCC defines a model where AI systems not only process linguistic inputs with nuanced understanding, but also adapt dynamically to user-specific contexts. This approach prioritizes personalization while adhering to principles of data minimization and ethical AI. By synthesizing CC's psychological depth with fuzzy logic's precision, FCCC creates systems that are not only context-aware, but also capable of engaging in meaningful, human-centric dialogues. This novel concept of FCCC redefines the scope of conversational AI, paving the way for more adaptive, empathetic and efficient systems.

The FCCC acronym captures this architecture directly: F stands for fuzzy, and the three Cs for Conversation, Character, and Computing. The combination of letters and numbers in the framework name additionally signals its capacity to operate with both linguistic imprecision (letters) and vague numerical values (membership degrees), reflecting the integration of Computing with Words with multidimensional fuzzy classification.

Figure 10 depicts my idea of the interaction and the exchange of the mentioned theories to build a larger, more holistic framework for future (conversational AI) systems. I already published my approach on Conversational AI and Fuzzy Conversational Character Computing (FCCC) in Hundertmark et al., (2023c).

Figure 10: Fuzzy Conversational Character Computing



Source: Data according to Hundertmark et al. 2023c.

2.7. Theoretical Foundations and their Contribution to Life Engineering

The overarching ambition of this thesis is to contribute to the visions of life engineering as defined by Österle (2020). This approach focuses on the development of digital assistants that serve human well-being rather than merely improving organizational efficiency. To fulfil this ambition, the FCCC concept draws on four interdependent theories, namely systems theory, conversational loops, fuzzy logic and character computing. This section explains the dependency between the theories and life engineering and anchors the FCCC concept in life engineering.

Systems Theory needs Life Engineering as Starting Point

Systems theory, as articulated by von Bertalanffy (1949), provides the foundational logic of self-regulation. A system continuously processes inputs from its environment, generates outputs, receives feedback on those outputs, and adjusts its internal state accordingly. For a conversational AI system aspiring to serve as a life agent, this architecture is a prerequisite. Without a self-regulating feedback loop, a conversational AI would respond to each user input in isolation, with no memory of prior exchanges, no capacity for adaptation, and no ability to improve over time. Life

engineering's core requirement, that a digital assistant grows closer to the individual user across repeated interactions, cannot be met without this loop.

However, systems theory alone cannot tell a conversational AI what to do inside each loop. It specifies the architecture of adaptation but not the substance of it. Classical cybernetic systems regulate according to numerically precise set-points. A thermostat knows exactly what temperature to maintain. A life agent, by contrast, must regulate according to values such as "the user seems moderately stressed" or "the user prefers a slower pace of explanation", values that resist precise numerical encoding. Systems theory, operating with crisp variables and deterministic feedback, has no native mechanism for handling this inherent imprecision of human well-being.

Fuzzy Logic is the Necessary Bridge to Human Values

Fuzzy logic, introduced by Zadeh (1965), was designed precisely to close the gap of only using crisp values. By replacing the binary membership of classical set theory with graded membership functions, fuzzy logic allows a system to represent and reason about the imprecise, continuous-valued descriptors that characterize human states. A user is not simply "anxious" or "calm"; she may belong to the fuzzy set "anxious" with a degree of 0.72 and to "calm" with a degree of 0.28 simultaneously. Zadeh's (1999) extension of fuzzy logic into Computing with Words (CWW) goes one step further. Instead of first converting linguistic input into numbers and then operating on those numbers, CWW allows words themselves to serve as computational primitives. This is critical for life engineering, where the interface between agent and human is natural language.

Without fuzzy logic, a self-regulating life agent built on systems theory would have to discretize all user states into crisp categories. Introducing artificial precision that distorts the very phenomena it aims to represent. Worse, such a system would systematically misclassify users in the transition zones between categories, precisely where the most nuanced and personally important information resides. Fuzzy logic makes the feedback loop of systems theory semantically compatible with the graded, linguistically expressed nature of human experience.

Fuzzy logic does, however, have a further limitation when applied to a conversational life agent. It can represent a user's state at a given moment, but it does not, by itself, specify which dimensions of that state matter. A fuzzy system must be told what to measure. In the context of supporting human well-being, the relevant dimensions are not arbitrarily chosen technical variables, they are the psychological dimensions of personality, emotion, and situational context, namely the constituents of human character.

Character Computing Completes the FCCC Concept

Character Computing (CC), introduced by El Bolock et al. (2020), provides the psychological architecture that fuzzy logic needs to be applied meaningfully in a

human context. CC defines the relevant dimensions through the triad of Character (C), stable personality traits and transient emotional states, behaviour (B), observable manifestations of character in interaction and situation (S), the external context shaping that behaviour. Together, these three components describe the full psychological reality of a user at any point in a conversation.

For a life agent, this triad is indispensable for two reasons. First, it grounds the agent's adaptive responses in an empirically validated model of human psychology, the Big Five personality framework (McCrae et al., 1992) rather than in ad hoc heuristics. A life agent that adapts its tone to a user's momentary mood without understanding that user's stable personality traits will produce inconsistent, sometimes counterproductive responses. Second, the triad distinguishes between dimensions that change slowly (character traits) and dimensions that change rapidly (emotional states), enabling the agent to modulate its adaptation accordingly. A calm, reassuring tone may be appropriate for a user exhibiting high neuroticism in a single message, but if the behaviour vector B reveals that this neuroticism has been elevated across all recent interactions, the agent's response strategy should be qualitatively different.

El Bolock et al. (2020) themselves note that CC, in its original formulation, relies on sensor data, physiological signals, facial expression recognition, motion capture – that are unavailable in text-based conversational AI. The FCCC concept resolves this limitation by applying multidimensional fuzzy classification (Zumstein, 2026) to derive the CC triad from text alone. This is the specific methodological contribution that makes CC actionable in a real-world life-agent deployment.

How the FCCC Concept Constitutes a Life Engineering Contribution

The FCCC concept with systems theory for self-regulation, conversational loops for iterative alignment with the user, fuzzy logic for semantic compatibility with human values, and character computing for psychological grounding, constitutes a coherent design programme for conversational AI systems that serve human well-being. Each theory is necessary, none is sufficient alone, together they define the FCCC framework.

Life engineering, as Österle (2020) envisions it, requires digital assistants that are adaptive (systems theory), communicatively intelligent (conversational loops), linguistically sensitive (fuzzy logic), and psychologically aware (character computing). The FCCC concept is, to my knowledge, the first formal attempt to integrate all four properties into a single conceptual framework and to validate that integration empirically.

Each Theory is Necessary and None is Sufficient Alone

The dependencies described above imply a stronger claim. Each of the five theories are necessary for the FCCC concept, but none is sufficient on its own. Table 4

makes this explicit by showing what each theory contributes and what limitation it leaves open for the next theory to resolve.

Table 4: Dependency structure of the five theories

Theory	Contribution to the FCCC concept	Gap that needs to be addressed with the next theory
Systems Theory (von Bertalanffy, 1949)	Architectural principle. The conversational AI is an open system that self-regulates through feedback loops between system and user.	Does not specify the form or internal structure of the feedback exchange in a conversational context.
Conversational Theory (Pask, 1965; Pangaro, 2010)	Operationalizes the feedback loop as a structured conversational turn: iterative, forward-moving, accumulating shared understanding.	Does not resolve the representation problem. User language is ambiguous and cannot be captured by binary logic.
Fuzzy Logic (Zadeh, 1965; Zumstein, 2007; Werro, 2015)	Provides the mathematical language for representing partial, overlapping, and linguistically expressed user states across multiple dimensions.	Does not specify which dimensions of the user must be modelled. Fuzzy sets can represent any attribute.
Character Computing (El Bolock et al., 2020)	Specifies the relevant dimensions through the C–B–S triad. Personality, affective state, situational context. Grounds the model in empirically validated psychology.	Does not determine the goal of adaptation: Why should the system adapt to a user's character? What should the adaptation ultimately serve?
Life Engineering (Österle, 2020)	Provides the normative horizon: Adaptation serves human flourishing, autonomy, and long-term wellbeing. Defines the life agent as the operational ideal and the ethical outer layer as its technical implementation.	Without the four preceding theories, Life Engineering remains an ethical aspiration without a technical implementation path.

Source: Data collected by the author.

The FCCC concept is not a modular assemblage of independently applicable theories. It is an integrated architecture in which removing any single theory collapses the design logic. The empirical use cases in chapter 3 and the technical implementation in chapter 5 are both structured to test the specific connections that this integrated

architecture produces better conversational AI than any subset of the theories could achieve alone.

2.8. Conversational AI in Business – Considering Economic Fundamentals

In addition to the terms and theories defined above, which mainly have their origins in computer science or psychology, it is also necessary to take economic principles into account. I focus on the value irritant matrix, introduced by Price and Jaffe (2008). The value irritant matrix is an instrument for classifying automation potential according to the needs of customers and companies (Price et al., 2008).

The matrix is divided into four quadrants, which represent different combinations of high and low value and high and low irritation. The high value, low irritation quadrant includes interactions that are most beneficial to both the customer and the company, such as the provision of a personalized service or quality products. In contrast, the low value and high irritation quadrant includes interactions that are annoying for the customer and offer little benefit, such as unnecessary upselling or irrelevant marketing messages. The matrix helps companies to prioritize measures and allocate resources effectively with the aim of maximizing high-quality interactions and minimizing those that annoy customers. The absence of such tools can lead to mismanagement, as the automated customer dialogue does not contribute to the company achieving (additional) sales, is too expensive to automate, or fails to identify or realize important learning effects in personal dialogue with customers. If too few customer dialogues are automated, the company can be said to be communicating too expensively.

Figure 11 illustrates the value irritant matrix. I consider this matrix mainly in use cases 2 and 3, presented in chapter 3 and in my final conceptual framework, as concerns economic influences.

Figure 11: Value Irritant Matrix

		To the Customer	
		Irritant (no need for interaction)	Value (save money, get support and advice)
To the Company	Irritant (do not want interaction)	Eliminate (root cause analysis)	Automate (customer self service)
	Value (learn, cut costs, increase revenue)	Simplify (improve underlying operations)	Leverage (spend more time/exploit)

Source. According to Price et al.(2008).

2.9. Findings on Fundamentals in a Nutshell

This section provides a foundational overview of key theoretical frameworks that I employ in this thesis and which are relevant for the DSR process.

At its core, the thesis is grounded in systems theory (von Bertalanffy, 1949), which provides a holistic perspective on how conversational AI systems function as dynamic, interconnected entities that adapt through feedback loops.

Key extensions to this framework include conversational loops (Pangaro, 2010), which emphasize iterative exchanges between users and AI to refine understanding, as well as fuzzy logic (Zadeh, 1965), which allows for nuanced language processing by handling uncertainties in human communication. The incorporation of CC (El Bock et al., 2020) adds another dimension, namely modelling user personality, behaviour and situational context, to enhance AI-driven interactions.

A central challenge in conversational AI is the ability to learn from interactions and dynamically adjust responses. This is addressed through feedback mechanisms derived from cybernetics (Wiener, 1948), enabling AI systems to maintain coherence, recognize ambiguity and personalize communication. Additionally, CWW (Zadeh, 1999) extends fuzzy logic by allowing AI to interpret qualitative user inputs in a mathematically structured way.

Finally all concepts together with the ideas of life engineering form a clear starting point for sustainable life agents according Österle (2020).

I use these findings, presented in the course of this chapter, as an essential basis for the artefact I develop over the course of the DSR process.

3. Defining Capabilities and Application Areas of Conversational AI (4 Independent Use Cases)

This chapter sets out to answer RQ1 which asks what requirements users have for conversational AI systems and how these can be translated into verifiable design guidelines and RQ2, which investigates how methods from Character Computing can be integrated into conversational AI systems to interpret the complete triad of user characteristics (personality, affective state and behaviour) and what existing gaps need to be addressed. Together, RQ1 and RQ2 lay the foundation for answering RQ3, namely the development of the FCCC Framework. In accordance with the DSR methodology (Möller et al., 2020), this chapter defines the design requirements for the artefact through four independent use cases.

The four use cases presented in this chapter serve a single collective purpose within the DSR process. To generate the empirically grounded design requirements on which the FCCC Framework is built. They are not independent experiments selected for convenience. Each use case covers a different dimension of the design challenge, addresses a distinct research question, and generates a distinct class of capabilities for the artefact. Taken together, they constitute a systematic exploration of the design space for the FCCC Framework across three progressively deeper layers.

The first layer (use cases 1 and 2, Sections 3.1 and 3.2) addresses user requirements from the outside. What do users need from a conversational AI system, and what do organizations need from it, in real deployment contexts? The second layer (use case 3, Section 3.3) goes deeper by asking how user requirements vary with individual characteristics, age, gender, personality, and prior chatbot experience. The third layer (use case 4, Section 3.5) reaches the theoretical core: given that the FCCC Framework relies on Character Computing to model user personality and affective state from text, is this actually feasible with existing methods?

Table 5 provides an overview of the four use cases, their research questions, domains, and the design requirements they generate. The requirements from all four use cases are consolidated in chapter 3.6.

Table 5: Overview of the four use cases in chapter 3

Use case	Section	Research question addressed	Domain
Knowledge Transfer	3.1	RQ1.1 – Design criteria for knowledge transfer with positive UX and life engineering alignment	Financial literacy (Eggheads.ai chatbot, n=14)
Customer Service & Sales	3.2	RQ1.2 – Balancing user objectives with organizational economic efficiency	Insurance chatbot experiment (n=84, 3 use cases × emotional vs. neutral design)
AI-Supported Advisory	3.3	RQ1.3 – How user requirements differ by situation, age, gender, personality, and chatbot experience	Insurance role-play experiment (n=39, 'Be the Bot' method)
Character Estimation	3.5	RQ2.1 / RQ2.2 – Can the CC triad be derived from text messages alone?	Large-scale online survey (n=2,000, damage-reporting scenario)

Source: Data collected by the author.

3.1. Knowledge Transfer and Information Management through Conversational AI

This section aims to evaluate the use of conversational AI systems in the context of knowledge transfer, with a focus on financial knowledge.

The findings of this use case will be utilized not only for the subsequent development and evaluation of the artefact, which in this case is the conceptual framework, but also to provide additional useful insights into the importance of individualized design for conversational AI systems.

The following use case 1 is based on my work published earlier: Hundertmark, S., Grüter, A., & Portmann, E. (2022). Mit Chatbots die finanzielle Allgemeinbildung stärken.

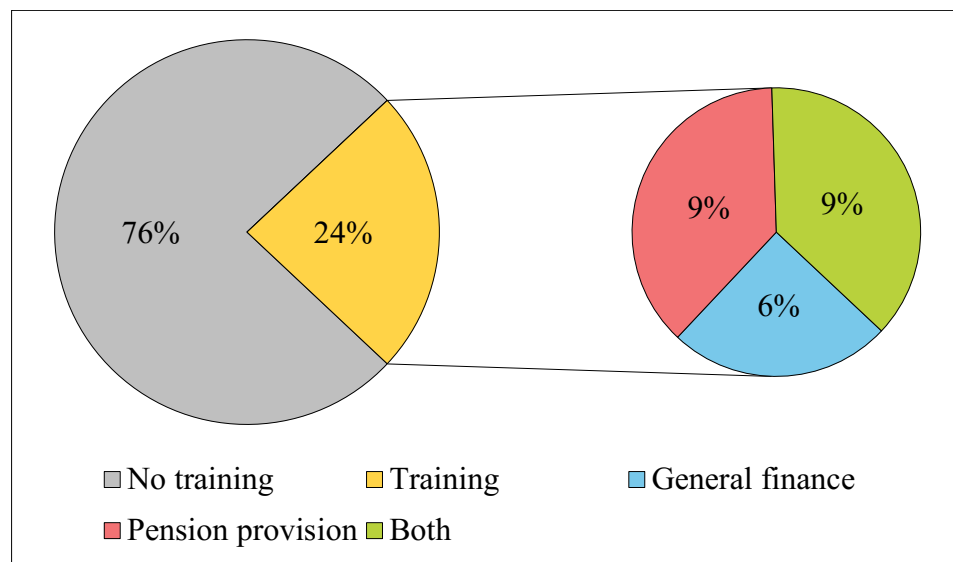
3.1.1. Systems for Promoting Financial Literacy: Defining the Research Question and Methodology

Financial literacy describes a person's knowledge of key financial concepts and the ability to apply this knowledge to make informed financial decisions (Stolper et al., 2017). Although financial literacy in the DACH region is quite high in international comparison (Ackermann et al., 2016), studies repeatedly show that there is a need for more training materials here as well. Studies also demonstrate that there are

differences in financial literacy between different age groups, educational levels and genders (Braun et al., 2013).

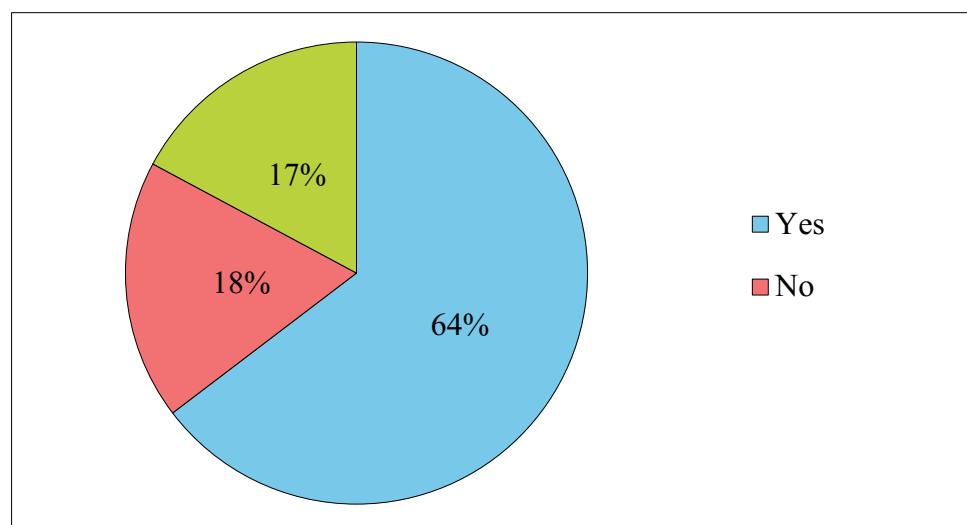
Surveys within the population further show that the majority has not yet received any training on financial topics but would certainly find this desirable (see Figures 12 and 13 below) (Zimmerman et al., 2021).

Figure 12: Have you ever received training in financial matters? If yes, in which category?



Source: According to Zimmerman et al. (2021, p. 35). Note: Number of observations: 1243.

Figure 13: Would you find training in personal financial management and knowledge transfer in the area of old-age provision useful?



Source: According to Zimmerman et al. (2021, p. 35). Note: Number of observations: 1243.

Based on the findings of Zimmerman et al. (2021), which show that younger people form the largest group in so-called financial illiteracy, this part of the use case primarily addresses the target group of younger people who have just entered working life. I wish to harness the strengths of conversational AI systems to develop an artefact that improves knowledge of financial topics among the population. In this section, I use text-based chatbots as a proxy for all types of conversational AI systems. Chatbots, also called conversational agents (CA), are software-based systems that use natural language to simulate a human conversation.

Chatbots have the advantage that they can present knowledge to the user at eye level. Simple and uncomplicated communication via chat leads many users to enjoy learning more and also contributes to the fact that the knowledge learned can be anchored in their memory better than with classic learning methods (Hundertmark, 2020). According to Clarizia et al. (2018), chatbots can even be considered one of the most innovative solutions for closing the gap between technology and education. This is due to several aspects. Chatbots allow users to follow their own, individual pace (Brandtzaeg, 2018). The chatbot can (and should) adapt to their pace. A chatbot is consistent and fair as it treats each user equally (Wambsganss et al., 2021) but, at the same time, it can promote each student according to their learning progress and knowledge level (Clarizia et al., 2018). One of the most important points, however, is the scalability of a chatbot. While human teachers are limited in the number of students they can support, a chatbot has no limit and can therefore provide closer support to a student (Wambsganss et al., 2021). A chatbot can even get better the more users it has (depending on the underlying technology). Furthermore, learning with chatbot technology is much more interactive (Clarizia et al., 2018). This can lead to additional motivation to learn. A chatbot can play an offensive role and motivate a student to continue learning by sending reminders using a gamification approach (Clarizia et al., 2018).

To date, there are no proven solutions, neither in research nor in practice, on how chatbot technology can be used to improve knowledge around financial topics within the population. This requires an answer to RQ1.1:

What design criteria should conversational AI systems incorporate to optimize knowledge transfer while ensuring a positive user experience and supporting life engineering goals?

The question is initially directed exclusively at young citizens who have just entered professional life. This narrowing down was done with the help of three short, unstructured expert interviews with professionals from the financial sector. All three experts agreed that the gaps in the financial literacy of this group are significant and that, at the same time, only little training material is available for these target groups to date. However, it can be assumed that some results are also transferable to other target groups.

Furthermore, only rule-based chatbots are used in the following experiment. Conversational AI systems, as described in chapter 2, are not yet used. Rather, the aim of this section is to create further foundations that will be useful for later applications of the approach proposed in the last section of this first use case.

I use the DSR approach presented in chapter 1 to answer the RQ 1. A combination of reflective and supportive paths was chosen. The individual steps are explained in more detail below.

I use the DSR approach presented in chapter 1 to answer RQ1. A combination of reflective and supportive paths was chosen. The process started with the supportive approach and the identification of the knowledge base and its meta-requirements, involving a thorough review of existing literature, theories and empirical studies related to financial literacy and chatbot design. Afterwards, the reflective approach was used and the first artefact was developed and evaluated with a group of people belonging to the previously defined target group. Design principles (DP) derived from this process are presented in Table 6 below.

Table 6: Design principles for the first artefact

DP	Requirement	Design recommendation	Source
DP1	Versatile/not boring	Chatbot uses different media formats	Wambsganss et al. (2021)
DP2	Fair	Chatbot is neutral and non-judgmental	Braun et al. (2013)
DP3	Motivating	Gamification elements, individual and regular feedback, sessions that are not too long, challenges and comparisons with other users, reminders (digital nudging)	Petrović (2020)
DP4	Easy to use	No complicated set-up required	Wambsganss et al. (2021)
DP5	Natural communication, personal relationship	Natural “emotional” communication (e.g. through the use of emojis)	Guglielmo (2021), Braun et al. (2013), Hundertmark (2020)
DP6	Little effort	The user can select an answer instead of typing	Hundertmark (2020)
DP7	Clear goal	Allow topic selection	Braun et al. (2013)
DP8	Relevance of the topic	Use practical task-based or role-play scenarios	Wambsganss et al. (2021)

Source: Data collected by the author, according to Hundertmark et al. 2022b.

According to Wambsganss et al. (2021), the use should not require a complicated setup, but should be platform-independent and have high availability (DP4). The chatbot should have different media elements, e.g. audible and graphic (DP1). Guglielmo (2021) and Braun et al. (2013) suggest personification of the chatbot to promote emotional connection and encourage interaction. The chatbot should behave like a human being, i.e. smile, write rather short sentences, explain with examples (DP5). The use of ready-made answers, known as rule-based chatbots, lowers the barriers to use (DP6). Braun et al. (2013) also mention the fairness of the chatbot: the user should not be judged, as learning works better when the mood is good (DP2). A suitable approach to motivating the user is gamification (DP3) (Petrović, 2020). Finally, the conveyed topic should have a clear relevance for the user (DP8), and the user should be able to select content individually (DP7), which can be improved through the use of educational role play scenarios (Wambsganss et al., 2021).

3.1.2. Design Artefact for Interactive Knowledge Transfer

To design an artefact based on the defined requirements of the last step, its architecture and functionality have to be determined. Subsequently, the artefact can be created (Hevner et al., 2010; Peffers et al., 2007).

In this study development of an artefact for use case 1, a sequence of nine individual chatbots with learning content, the no-code solution from Eggheads was used. Eggheads.ai is a no-code software as a service tool that focuses exclusively on the development and implementation of chatbots in a learning context. The platform is web-based and simulates a chat application. It enables the creation of rule-based conversations with predefined responses. Using the previously defined design features, nine short chatbots were developed, each covering a topic in the field of financial education. Ultimately, all nine training units are considered in their entirety as one learning unit and thus together form the first prototype or artefact in the context of DSR. Table 7 presents the nine different chatbot sequences.

Table 7: Chatbot sequences

Topic	Link	Name of fictive friend
Why investing?	https://eggheads.ai/chat/2144	Paulchen
Shares in general	https://eggheads.ai/chat/2215	Maya
Funds in general	https://eggheads.ai/chat/2219	Hans
Bonds in general	https://eggheads.ai/chat/2221	Berta
Comparison of risks and chances of shares, funds and bonds	https://eggheads.ai/chat/2252	Paulchen
Fund savings plan and compound interest	https://eggheads.ai/chat/2268	Hans again
Basics of sustainable investing	https://eggheads.ai/chat/2285	None
Sustainable investing deep-dive	https://eggheads.ai/chat/2287	None
Basics of retirement planning	https://eggheads.ai/chat/2290	None

Source: Data collected by the author, according to Hundertmark et al. 2022b.

Since eggheads.ai is a no-code tool, the chatbots can be developed without any additional code. Figures 14 and 15 provide an overview of the user interface, in which the chatbots are quickly built, without using any code or additional programming languages.

Additionally, the tool eggheads.ai was utilized in German, consistent with the overall experiment conducted among German-speaking users. For the sake of full transparency, I have not translated the screenshots from the tool but present all original materials and outcomes in their original German versions. Additionally further details and translations, also including the questionnaire can be found in appendix A.

Figure 14: Eggheads.ai user interface Illustration

Training state

This training is **published!**

Copy training

See analytics

Preview the training

Un-publish training

Training Title

FinLit 1 - Warum Geld anlegen?

Training URL

https://app.eggheads.ai/chat/517w95jvn3q

Preview URL

https://app.eggheads.ai/chat/517w95jvn3q/preview

Training Language

German

Public training

☒ Select this checkbox to make this training accessible to anyone (no login required).

Tags

x financial literacy

Send resource link

☐ Offer the option to receive additional resources via email (at the end of the training).

Dialog 1

B I H

Q Preview

AI feedback

Hallo, Ciao, Hello, Bonjour 😊.
Im heutigen Training lernst du, warum Geld anlegen wichtig ist und was es zu beachten gibt. Stell dir vor dein Kollege Paulchen hat das erste Jahr in seinem neuen Job gut überstanden und nun sogar die erste Gehaltserhöhung bekommen. Und nun kommt Paulchen zu dir und möchte wissen, wie viel Geld er wohl anlegen soll und ob Geld anlegen auch wirklich die richtige Entscheidung ist. Du kannst ihm da sicher helfen,

+ Add response

Geld anlegen macht Sinn, wenn man bestimmte Kriterien erfüllt

Geld anlegen ist ja auch unnütz.

Bin unsicher, da war doch was mit Inflation oder 🤔?

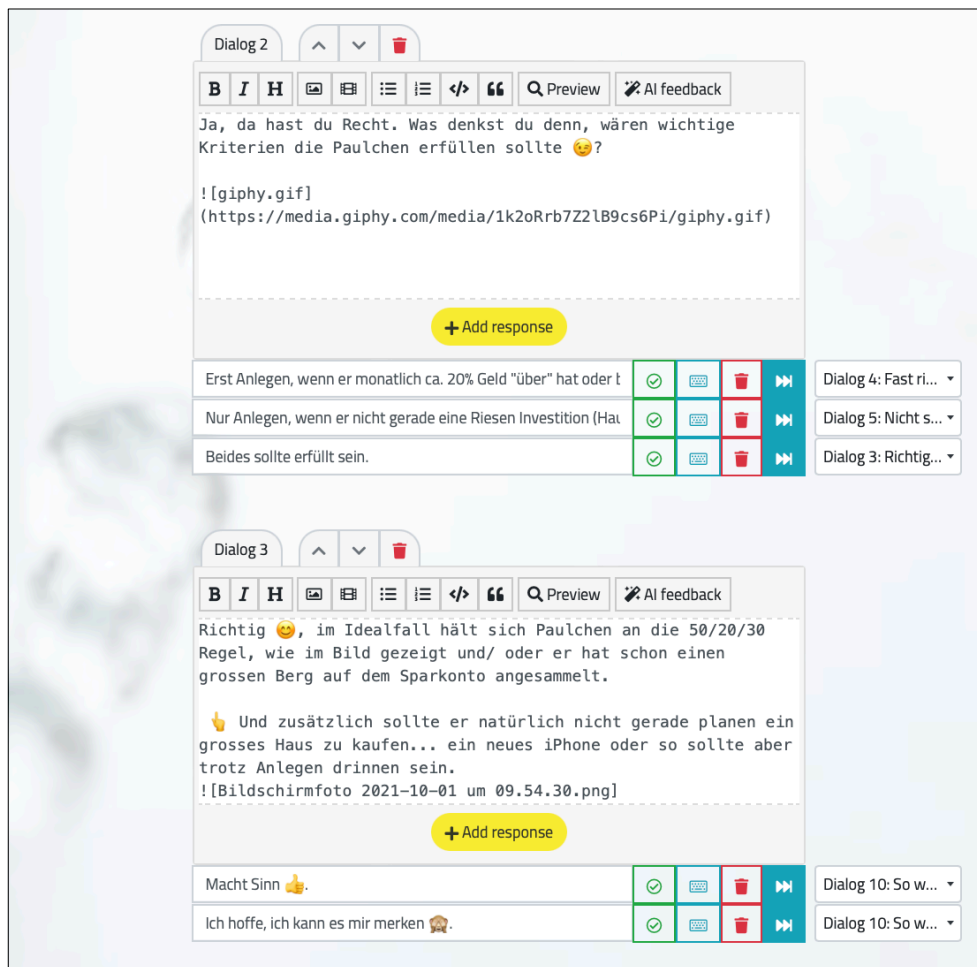
Dialog 2: Ja, da ...

Dialog 6: Das ist...

Dialog 7: Die Inf...

Source: Data collected by the author, according to Hundertmark et al. 2022b.

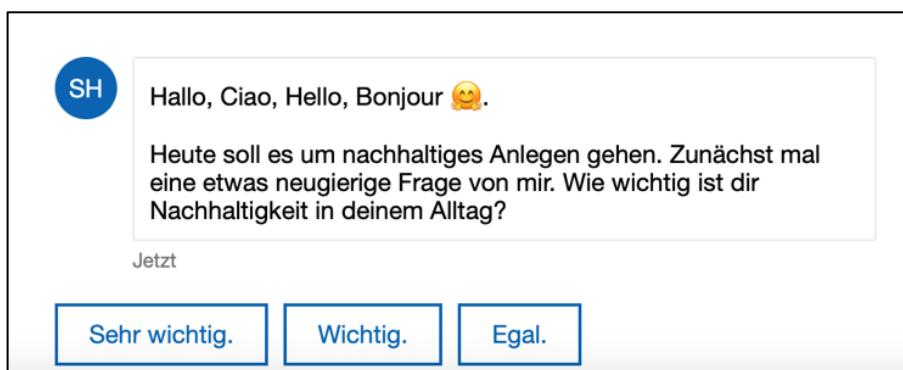
Figure 15: Eggheads.ai user interface Illustration



Source: Data collected by the author, according to Hundertmark et al. 2022.

In Figures 16, 17 and 18, some example chat dialogues are shown. As explained above, all chatbots are rule-based, users cannot type in any individual messages and can only navigate through the chatbot by clicking on suggested buttons. Each button and the corresponding answer is pre-defined in the user interface shown before.

Figure 16: Eggheads.ai Chat window Example



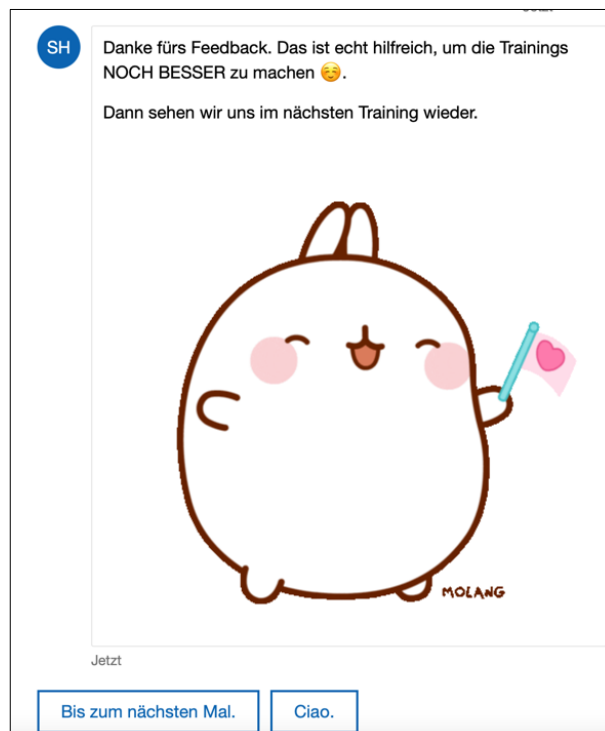
Source: Data collected by the author, according to Hundertmark et al. 2022b.

Figure 17: Eggheads.ai Chat window Example



Source: Data collected by the author, according to 2022b.

Figure 18: Eggheads.ai Chat window Example 3



Source: Data collected by the author, according to Hundertmark et al. 2022b.

3.1.3. Evaluation and Empirical Results on Conversational AI systems in Knowledge Transfer

The evaluation is one of the key aspects of DSR. While the first part of the evaluation of use case 1 consists of demonstrating that the artefact fulfils its purpose and works accordingly (Hevner et al. 2010; Peffers et al., 2007). The second and later evaluation parts of use case 1 refer to the performance measurement of the created artefact. Among others, the artefact can be compared with the performance of another artefact that solves the same problem, or it can be evaluated based on the results of satisfaction surveys. For this study, I used surveys for evaluation.

To evaluate the artefact, 15 test users were recruited that fit into the defined target group. Test users were between 18 and 25 years old and living in Switzerland. Seven test users were female and eight test users were male. They can therefore be considered young professionals with little knowledge of financial topics. Before beginning the evaluation of the artefact, members of the test group were asked to complete a small questionnaire to gain insight into how the users informed themselves about financial topics and whether they had already used chatbots. For nine days, the participants received an email every day, asking them to complete the next chatbot training.

At the end of the test phase, all participants were invited to complete a post-experiment questionnaire. 14 of the 15 participants submitted complete responses, yielding a response rate of 93%. The questionnaire comprised 20 items structured across four thematic blocks: (1) overall implementation of the chatbot training (length, tonality, media ratio, narrative coherence); (2) content structure, difficulty and responsiveness of the predefined answer options; (3) social learning preferences (desire for human coaching, peer comparison and anonymity); and (4) motivational factors and willingness to pay. Table 8 presents the descriptive results for each item.

Table 8: Questionnaire

#	Item (English)	Response options	n	Key result
Q1	My general knowledge of financial topics has improved	Yes, definitely / Yes, a little / No, not really	14	13/14 (93%) reported improvement: 12 'yes, a little', 1 'yes, definitely'; 1/14 (7%) 'no, not really'
Q2	The length of the training sessions was appropriate	Appropriate / Too short / Too long	14	14/14 (100%) rated length as appropriate

#	Item (English)	Response options	n	Key result
Q3	I could easily identify with my role in the training	Yes / Partially / No	14	12/14 (86%) 'yes'; 2/14 (14%) 'partially'; 0/14 'no'
Q4	I would have liked an additional human coach	Yes / Sometimes / No	14	10/14 (71%) 'no'; 2/14 (14%) 'yes'; 2/14 (14%) 'sometimes'
Q5	I was bothered by the fact that I could not ask my own questions	Yes / No	14	10/14 (71%) 'yes – bothered'; 4/14 (29%) 'no'
Q6	The ratio of text to images was appropriate	Appropriate / Too few images / Too many images	14	11/14 (79%) 'appropriate'; 3/14 (21%) 'too few images'; 0/14 'too many'
Q7	Would you recommend the training to others?	Yes / No	14	14/14 (100%) 'yes'
Q8	The tone of the bot was appropriate	Strongly agree / Partially agree / Do not agree	14	11/14 (79%) 'strongly agree'; 2/14 (14%) 'partially agree'; 1/14 (7%) 'do not agree'
Q9	The bot responded to my answers	Strongly agree / Partially agree / Do not agree	14	9/14 (64%) 'strongly agree'; 5/14 (36%) 'partially agree'; 0/14 'do not agree'
Q10	The predefined answer options were sufficient	Yes / Mostly yes / Mostly no / No	14	5/14 (36%) 'yes'; 9/14 (64%) 'mostly yes'; 0/14 'mostly no / no'
Q11	The training sessions felt preachy / lecturing	Strongly agree / Partially agree / Do not agree	14	4/14 (29%) 'strongly agree'; 7/14 (50%) 'partially agree'; 3/14 (21%) 'do not agree'
Q12	The training adapted to my learning pace	Yes / Partially / No	14	14/14 (100%) 'yes'
Q13	How often would you like to do such mini-trainings?	Dail / 2–3 times/week / Weekly / Monthly	14	8/14 (57%) '2–3 times per week'; 6/14 (43%) 'weekly'

#	Item (English)	Response options	n	Key result
Q14	How much would you pay per month for continued access? (CHF)	Open numeric	14	CHF 5 most common (8/14 = 57%); mean CHF 5.36 (incl. zeros); 9/14 (64%) willing to pay >0 CHF; max CHF 20
Q15	The training was enjoyable	Strongly agree / Partially agree / Do not agree	14	10/14 (71%) 'strongly agree'; 4/14 (29%) 'partially agree'; 0/14 'do not agree'
Q16	I would like to compare my knowledge level with like-minded peers	Strongly agree / Partially agree / Do not agree	14	4/14 (29%) 'strongly agree'; 8/14 (57%) 'partially agree'; 2/14 (14%) 'do not agree'
Q17	I would like to complete such trainings anonymously (no one can see my progress)	Strongly agree / Partially agree / Do not agree	14	10/14 (71%) 'strongly agree'; 3/14 (21%) 'partially agree'; 1/14 (7%) 'do not agree'
Q18	The training sessions follow a coherent narrative thread	Strongly agree / Partially agree / Do not agree	14	12/14 (86%) 'strongly agree'; 2/14 (14%) 'partially agree'; 0/14 'do not agree'

Source: Data collected by the author, according to Hundertmark et al. 2022b.

The results paint a consistently positive picture and allow a number of concrete conclusions to be drawn for the further design of conversational AI systems in knowledge transfer. The following discussion presents the key findings grouped by theme and explicitly traces how each finding gives rise to the design criteria for future systems.

Learning outcome. The primary objective of the training, improving financial knowledge, was achieved for 13 of the 14 respondents (93%). 12 participants (86%) reported that their knowledge had improved 'a little', and one further participant (7%) reported definite improvement. Only one participant (7%) stated that their knowledge had not improved. The training therefore fulfilled its core function, and this result confirms the potential of chatbot-based approaches for knowledge transfer as described in Section 3.1.1.

Session length and pace. All 14 respondents (100%) rated the length of the training sessions as appropriate, and all 14 (100%) confirmed that the training had adapted to their individual learning pace. This result confirms the design principle of individual pacing (DP4) and validates the choice of a rule-based chatbot format in which learners control the speed of progression. It also suggests that the nine-day

format with one daily unit of approximately five to ten minutes represents a suitable design benchmark for future applications.

Preferred frequency and session duration. When asked how often they would ideally engage with such mini-trainings, 8 of 14 participants (57%) preferred two to three sessions per week, while the remaining 6 (43%) indicated a weekly cadence. No participant requested a higher or lower frequency. In terms of willingness to pay, CHF 5 per month was the most common response, given by 8 of 14 participants (57%). Overall, 9 of 14 respondents (64%) indicated a willingness to pay a positive amount, with a mean of CHF 5.36 per month across all responses (including those who indicated CHF 0) and a maximum of CHF 20. These findings directly inform the recommendation for future system design: sessions should be brief (approximately five to ten minutes), delivered at a frequency of two to three times per week, and be economically accessible at a low price point.

Tone and conversational quality. 11 of 14 participants (79%) fully agreed that the tone of the bot was appropriate, and a further 2 (14%) partially agreed, resulting in 13 of 14 (93%) rating the tone positively overall. Similarly, 9 of 14 (64%) fully agreed that the bot had responded to their answers, while the remaining 5 (36%) partially agreed, giving a 100% agreement rate on at least partial responsiveness. These findings confirm the relevance of DP5 (human-like, conversational tone) and indicate that the implementation largely succeeded in achieving a tone perceived as appropriate and engaging. Regarding the coherence of the content, 12 of 14 participants (86%) fully agreed that the training sessions followed a clear narrative thread, with a further 2 (14%) partially agreeing, resulting in unanimous positive ratings.

Media and content depth. 11 of 14 participants (79%) found the ratio of text to images appropriate; however, 3 of 14 (21%) felt that there were too few images. This finding extends DP1, which already emphasised the importance of multiple media types. Taken together with the open-ended responses – several participants mentioned that content was occasionally too superficial or that they would have liked more in-depth examples – this points to a need for richer media integration and graduated content depth in future implementations.

Answer options and the central limitation of rule-based chatbots. While 14 of 14 respondents (100%) rated the predefined answer options as sufficient or mostly sufficient (Q10), 10 of 14 (71%) simultaneously stated that they were bothered by the inability to type in their own questions (Q5). These two results are not contradictory: the predefined options were adequate for the learning activities provided, but the inability to ask open questions represents a fundamental limitation of rule-based technology. This finding is critical for the development of future conversational AI systems: it establishes the motivation for transitioning from rule-based to NLU-based systems (as addressed in Chapter 4) and directly generates a new design requirement (Capability 3) for the framework. Participants' open-ended comments reinforce this

conclusion: responses included requests for 'more individual answer and question options' and feedback that the chatbot would sometimes end conversations abruptly without allowing clarification.

Perception as lecturing. The item 'the training sessions felt preachy' produced a more heterogeneous response than any other item: 4 of 14 (29%) fully agreed, 7 (50%) partially agreed, and only 3 (21%) disagreed. This finding is noteworthy because a lecturing tone runs counter to the design principle of eye-level communication (DP5) and the non-judgmental environment described in DP2. The pattern suggests that the current prototype only partially achieved a conversational rather than instructional register. Future design should place greater emphasis on dialogue-driven knowledge construction (Socratic questioning, scenario-based learning) rather than information delivery, in order to reduce the perceived lecturing character.

Desire for human coaching. 10 of 14 participants (71%) stated they would not want an additional human coach, while 2 (14%) said yes and 2 further (14%) 'sometimes'. This finding is encouraging for scalability: the majority of participants were satisfied interacting exclusively with the chatbot, without requiring human backup. However, the 29% who expressed at least some desire for human contact confirm the general finding in the literature (Wambsganss et al., 2021) that hybrid approaches where a human expert can step in when needed may further enhance the user experience for more complex or personalised learning needs.

Anonymity and social comparison. 10 of 14 participants (71%) strongly agreed that they wanted to complete the training anonymously, with a further 3 (21%) partially agreeing, resulting in 13 of 14 (93%) placing value on anonymity. At the same time, results for social comparison (Q16) were more ambivalent: 4 of 14 (29%) strongly agreed they would like to compare their knowledge with peers, 8 (57%) partially agreed, and only 2 (14%) disagreed. This combination reveals an important nuance. Participants want the option of knowing how they compare to others in the aggregate, but they do not wish to have their individual learning progress visible to others. Taken together, these findings motivate a new design principle for future systems. Anonymity must be guaranteed by default, while optional, non-identifiable peer comparison (e.g. aggregate benchmark scores) may enhance motivation without violating privacy preferences.

Motivation and gamification. When asked about what would motivate regular use, the most frequently mentioned elements in open-ended responses were: quizzes with points, prize draws, progress documentation, and topic relevance. These responses, which echo the gamification principle (DP3) established from the literature, confirm that gamification elements represent a high-value addition for motivation. In particular, the combination of visible progress and light competitive elements (points, badges) is perceived as motivating, as long as individual performance data remain private.

Overall recommendation rate. All 14 respondents (100%) stated they would recommend the financial literacy chat training to others. This result represents the highest possible endorsement rate and serves as the most direct confirmation that the prototype achieved its objectives. It also provides empirical support for the viability of chatbot-based approaches to financial literacy improvement as a scalable and user-accepted tool.

Table 9 summarises the design criteria derived from the evaluation of the first prototype. These criteria extend and refine the original design principles (DP1–DP8) presented in Table 6, based on the empirical feedback from the 14 participants.

The whole list of answers, also including qualitative feedback and the questionnaire can be found in appendix A.

Table 9: New design principles driven from the evaluation

DPN#	Design Principle	Description	Question
DPN1	Total length	No more than 10 minutes per chatbot training session	Question 2
DPN2	Frequency of chatbot training	Between 1 and 3 per week	Question 14
DPN3	Use of media resources	Approx. 3 images per chatbot training. Images must be easy to export	Question 6
DPN4	Community building	Users want to remain anonymous and there is no need to compare training results with other users	Question 18
DPN5	Tonality, language	Use of everyday language, no financial tone. A bit like a conversation between friends	Question 9
DPN6	Price per month	If a price is required, EUR 5 per month could be OK	Question 15
DPN7	Buttons vs. open questions	During the training: use of buttons and multiple choice questions. At the end: possibility to ask own questions that the bot can answer	Question 5
DPN8	Content and level	At the beginning, focus more on basic topics. Then go more in depth	Question 20
DPN9	Gamification, Quiz	Integration of knowledge quizzes at the end of every or every second training session	Question 16

Data collected by the author, according to Hundertmark et al. 2022b.

Although the feedback on the first prototype was positive and the feedback from the users helped to make further improvements to the prototypes of use case 1, some open questions remain, as well as potential for further research. It should also not be forgotten that users asked for open questions to the chatbot. As already explained, this cannot be implemented with the technology used here. Additionally, I only focused on one specific target group. In further research, other target groups should also be considered.

Furthermore, it is pertinent to briefly examine these use cases within the framework of systems theory. To date, the chatbots presented here are technically incapable of modifying their behaviour in response to environmental stimuli. Consequently, elucidating the dynamic interaction between a chatbot and its user remains a challenge. Addressing this deficiency is imperative for the future development of chatbots that can effectively enhance the financial literacy of individuals. This advancement should be achieved in a manner that is not only sustainable but also transparent, thereby ensuring long-term effectiveness and user trust.

3.1.4. Requirements on Conversational AIs for Knowledge Transfer – Implications for Future Conversational AI

In use case 1, I initially developed specific design criteria tailored to conversational AI systems for the transfer of financial literacy knowledge in order to answer RQ1.1:

RQ1: What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.1: What design criteria should conversational AI systems incorporate to optimize knowledge transfer while ensuring a positive user experience and supporting life engineering goals?

From these specific design criteria focusing on the domain of knowledge transfer, I derived more general design requirements for the artefact that can guide the design of conversational AI systems across various applications. These generalized requirements are intended to support scalable, adaptable and user-centered design, ensuring effective knowledge transfer while maintaining a high-quality user experience.

The findings from use case 1 make it possible to consolidate the nine design principles into six generalisable capabilities for the artefact. Whereas the design principles in Table 9 are specific to the financial literacy context, the capabilities in Table 10 are formulated at a higher level of abstraction so that they can guide the design of conversational AI systems across a broader range of knowledge transfer applications.

Table 10: Capabilities 1-6

Number	Capability	Explanation	Source
1	Audience-Adapted Communication	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Use Case 1
2	Flexible Interaction Modes	Combine structured interaction (e.g. pre-defined buttons) with opportunities for user-driven input to enhance engagement and adapt to users' learning preferences.	Use Case 1
3	Adaptive Content Design	Enable the integration of knowledge documents to ensure that learning content remains customizable and aligned with specific needs, allowing organizations to maintain control over the material.	Use Case 1
4	Economic Accessibility	Maintain a cost structure that balances system sustainability with affordability for a broad user base.	Use Case 1
5	Multimedia Integration	Incorporate diverse media elements such as images, videos and graphics to enhance interactivity and accommodate various learning styles.	Use Case 1
6	Privacy and Trust	Ensure user anonymity and data protection to build trust and encourage honest participation.	Use Case 1

Source: Data collected by the author.

Besides these concrete capabilities, use case 1 demonstrates that chatbots are a viable and well-received means of imparting financial knowledge to young adults. The finding that all 14 respondents (100%) would recommend the training to others, combined with 13/14 (93%) reporting measurable knowledge improvement, provides strong empirical support for the proposition that conversational AI systems can reduce financial illiteracy in the target group identified by Zimmerman et al. (2021).

At the same time, the evaluation reveals a fundamental technological problem for rule-based systems. As long as users cannot ask open-ended questions, the chatbot interaction will feel constrained rather than genuinely conversational. This limitation directly motivates the transition to NLU-capable systems and, ultimately, the integration of adaptive conversational AI as pursued in the subsequent chapters.

The willingness to pay of approximately CHF 5 per month, expressed by the majority of participants, confirms that financial literacy chatbot training is not only technically feasible and user-accepted, but also economically viable at a modest price

point, particularly if offered as a value-added service by financial institutions, an avenue explicitly suggested by several participants in open-ended comments.

Finally, from a systems-theory perspective, this use case illustrates the fundamental limitation of first-generation, rule-based conversational agents. They are unable to modify their behaviour in response to environmental stimuli or to learn from individual user interactions. Addressing this limitation, I use these design requirements as essential capabilities for the artefact, which I will develop in the later stages (chapter 4) of this thesis.

Use case 1 establishes that the design of a conversational AI system for knowledge transfer must balance content quality with a positive user experience, and that ethical considerations. Particularly transparency about the AI's nature and the boundaries of its advice are not peripheral constraints but core design requirements. What this use case does not yet address is the tension between *user* objectives and *organizational* objectives, a tension that becomes acute in commercial customer service contexts where cost efficiency and customer experience pull in opposite directions. Use case 2, presented in the following section, places this tension at the centre of the investigation.

3.2. Conversational AI in Customer Service and Sales

In the section that follows, I present an experiment (use case 2), focusing on the use of conversational AI systems in customer service and sales with the aim of achieving the highest possible customer experience.

I show how the appropriate design of conversational AI systems contributes to solving the tension between the costs of personalized communication and the customer experience. I focus exclusively on the situation of the customer and the resulting emotions. I built a hypothesis for different customer situations based on a literature review and tested this hypothesis in a quantitative experimental study. In doing so, I only consider one dimension from CC (El Bolock et al., 2020) presented in chapter 2.

The experiment and results are based on my work, published earlier: Hundertmark, S., & Hafner, N. (2023a). Gestaltung automatisierter Kundendialoge im Spannungsfeld von Kosten und Customer Experience.

3.2.1. Managing Conflicts between Costs and Experience: Background, Research Question and Methodology

The concept of customer experience management plays a role, above all, when it is not only a matter of gaining new customers, but also of binding them to the company in the long term (Bruhn, 2016). The aim is to offer customers experiences that they evaluate positively, which they will remember and thus emotionally connect them to the company (Christopher et al., 1991; Lemon et al, 2016). Customer experience research emphasizes the role of positively connoted emotions in customer

contact (Klaus et al., 2020; Oliver, 2010). In order to deliver these emotions to the customer in a sustainable way, it is necessary to look not only at the beginning of the customer journey but also, and above all, at customer service. Only when customers receive the right solution quickly are positively evaluated experiences possible.

A widely accepted measurement tool in practice that underlines the role of emotions in remembering the quality of the experience is the Net Promoter Score (NPS) (Reichheld et al., 2011). The NPS measures customer loyalty by asking customers how likely they are to recommend a company or product to others on a scale from 0 to 10. Respondents are classified as promoters (score 9–10), passives (7–8), or detractors (0–6). The NPS is then calculated as the percentage of promoters minus the percentage of detractors, resulting in a score ranging from –100 to +100. A higher NPS indicates stronger customer loyalty and a more positive overall customer experience (Reichheld et al., 2011).

At the same time, one of the most important potentials for increasing profitability that results from the digital transformation is the possibility of automation (Mockenhaupt, 2021). In this section, automation mainly refers to the use of conversational AI systems.

However, automating customer-dialogues does not always make sense. For example, the introduction of self-service in banking at the turn of the millennium impressively showed that the banks lost customer knowledge and access in the segment of well-educated, wealthy customers, while the service was not used by the less digitally-savvy customer group, and thus the originally targeted cost savings in this segment did not materialize. “Smart services” must therefore be used adequately for the company under consideration and for the target group under consideration (Landmann et al., 2022). An instrument for classifying such automation potentials according to the needs of customers and companies is the value irritant matrix (Price et al., 2008). However, the use of the value irritant matrix does not answer the question of how exactly the automation should look in a specific application to be remembered by the customer in the sense of a positive customer experience. Landmann et al. (2022) point out that an additional open question in relation to the value irritant matrix is how exactly dialogues can be made more beneficial. They ask the central question, “What does the improved customer experience look like?” In this section, I elaborate the design options that contribute to the fact that a conversational AI system, which is used as part of a company’s automation measures, must be designed to obtain an enhanced customer experience. Hence, I use chatbots here as an example of a conversational AI system.

In this context, it is important to restate RQ1.2:

RQ1.2: What design criteria should conversational AI systems incorporate to balance user objectives, such as quick and correct solutions, with company goals, including economic efficiency and customer experience?

I aim to answer this question with the help of an experiment (use case 2), which I present in the section below.

3.2.2. Managing Conflicts between Costs and Experience: Related Work and Hypotheses

The literature review employed in this section follows the same systematic procedure as described in chapter 1.6.1. The search was conducted across the IEEE Library, ACM Portal, Science Direct, Emerald and ResearchGate, with keywords "*conversational AI*", "*conversational agent*", "*chatbot*" and "*chat bot*" (January 2011–May 2021), restricted to empirical papers focusing on social cues and their effects on user reactions in text-based chatbots. The full search procedure and its outcomes are documented in chapter 1.6.1.

Additionally the experiment and results presented in this chapter are based on my work, published earlier:

Hundertmark, S., & Hafner, N. (2023a). Gestaltung automatisierter Kundendialoge im Spannungsfeld von Kosten und Customer Experience.

In total, 50 relevant papers were identified and analysed (see appendix E). In the following, I highlight the most important findings from the literature review, which later lead to the hypothesis for the presented quantitative research study, answering RQ1.2.

In their work, Rhio et al. (2019) show that frequently asked questions (FAQ) chatbots, such as those increasingly used in customer service, are evaluated as more credibly and more positively overall when they show emotions rather than just providing pure facts as answers. Rhio et al. (2019) attribute their results to experiments with students who used an FAQ chatbot from their university, both with and without the design of human emotions. Jordan et al. (2015) focus their research primarily on older bot users and show that usage rates are significantly higher when conversational agents include emotional or human elements in addition to static and neutral answers.

Doctor et al. (2016) also show similar results. According to the authors, emotionally designed computer systems have a positive effect on the well-being and satisfaction of patients. However, Crolie et al. (2021) clarify with their experiments that emotions conveyed by chatbots are not always desired or do not always lead to better results. They prove that customers who are already in a negative mood at the beginning of the chatbot conversation, for example, due to a damage report and the associated harm as well as negative feelings or even anger about the damage, only want direct and fact-based answers from the bot instead of embellished and emotionally charged answers. Similar results are presented by Ng et al. (2020) in connection with bank customers. They show that customers appreciate the human traits of the conversational agent during the consultation and rate it positively. If, on the other hand, it is merely a matter of reporting changes or requesting account numbers or account

balances, customers prefer the “neutral bot”, which merely responds quickly and straightforwardly to the user’s request, instead of prolonging the dialogue with emotions.

Ghandeharioun et al. (2019) made similar observations. They show that users who are in a negative or sad mood do not want too explicit emotions from the bot. They would rather have a neutral tone. This changes as soon as the users are in a good mood. Katayama et al. (2019) come to the same conclusion concerning the different needs of good and bad-tempered chatbot users. They also prove that bots should primarily signal emotions in positive situations. Mozafari et al. (2021) even show another application for neutral chatbots. In their experiment, they illustrate that, while a human chatbot design increases the intention to use it in non-awkward contexts, it tends to lead to the opposite in contexts perceived as being awkward. Mozafari et al. (2021) justify their hypotheses with an experiment in which consumers were asked to buy either a sunscreen lotion for their irritated skin or a hemorrhoid ointment. The sunscreen lotion is a scenario with a rather low embarrassment level, whereas the haemorrhoid ointment is a product with a high embarrassment level. Mozafari et al. (2021) thus question previous research on the positive effects of social presence in chatbots and suggest that chatbots acting in embarrassing situations would do better to include hardly any or no human design elements.

The highlighted findings and the overall literature review I conducted in advance illustrate that, as soon as a customer is in a positive mood and wants advice, emotions on the part of the bot are very desirable. The desired emotions decrease as soon as users themselves are no longer in a positive mood. The emotions also decrease as soon as there is a question about enquiries regarding customer master data or contract data. However, the influence of the user’s negative emotion is significantly higher than the influence of the content of the request.

Based on the results presented above and to answer RQ1.2 explained above, I formulate three hypotheses to explore customer needs and demands for chatbot-automated dialogues so as to subsequently be able to make recommendations for the future artefact.

These are the following hypotheses:

- H₁: When users are in a positive mood or in a counselling process, emotions conveyed by the chatbot are desired.

Examples could be: premium advice in insurance, financial advice in a bank or product-oriented advice in e-commerce.

- H₂: When users are in a negative mood, they do not want exaggerated emotions from the chatbot.

Examples of this could be: a damage report in the insurance company, card blocking after loss at a bank or a complaint about incorrectly delivered products in e-commerce.

- H₃: As soon as it is a matter of rudimentary situations in which the user himself does not feel any great emotion, but rather seeks the quick completion of a case, the emotions of the chatbot hardly play a role. They can therefore also be omitted, but they also do little harm if they are there after all.

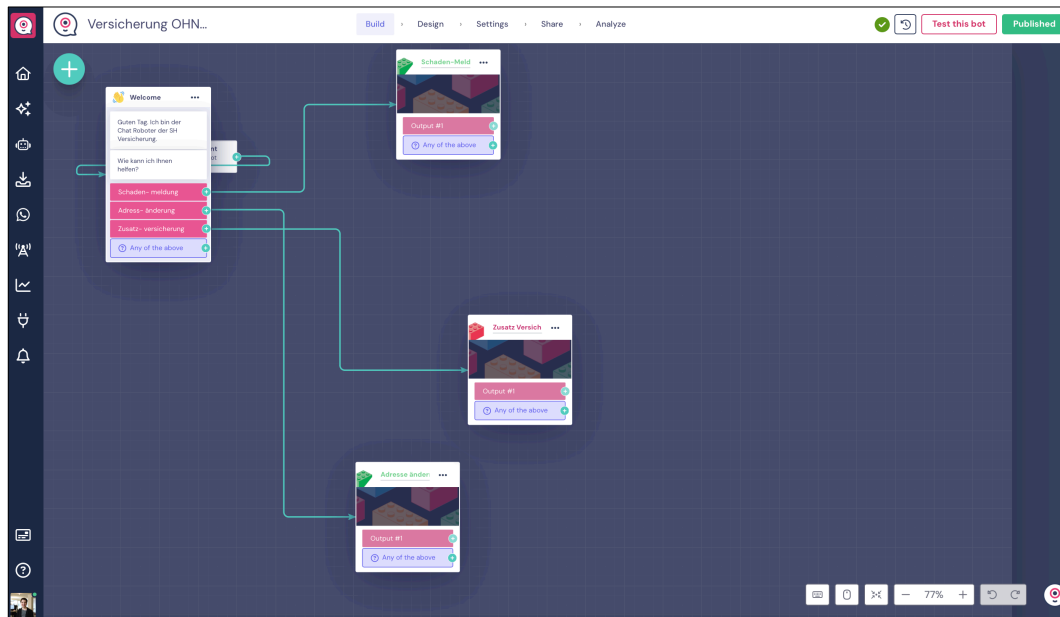
Examples could be: reporting a change of address to the insurance company, requesting a bank statement or requesting a receipt in e-commerce.

To test these hypotheses, I conducted an experiment (use case 2) concerning insurance use cases and designed an experiment, carried out between 21.12.2021 and 25.01.2022, in which a total of 84 insurance customers from Switzerland and Germany participated. Since the participants were randomly approached and selected via social media channels (Facebook and LinkedIn), it can be assumed that the responses are slightly more “digital” than would be expected in the Swiss and German resident populations. As during the evaluation of this experiment, no differences were made between gender or age, and I did not consider these attributes while recruiting the experiment participants. Ultimately, I found 50 female and 34 male customers. Ten customers were between 18 and 30 years old, 34 customers were between 31 and 40 years old, 31 customers were between 41 and 50 years old, and only nine customers were between 51 and 60 years old. In the experiment, insurance customers were asked to go through one of the three use cases with a fictitious chatbot:

- Take out supplementary insurance,
- Report damage and
- Changing address

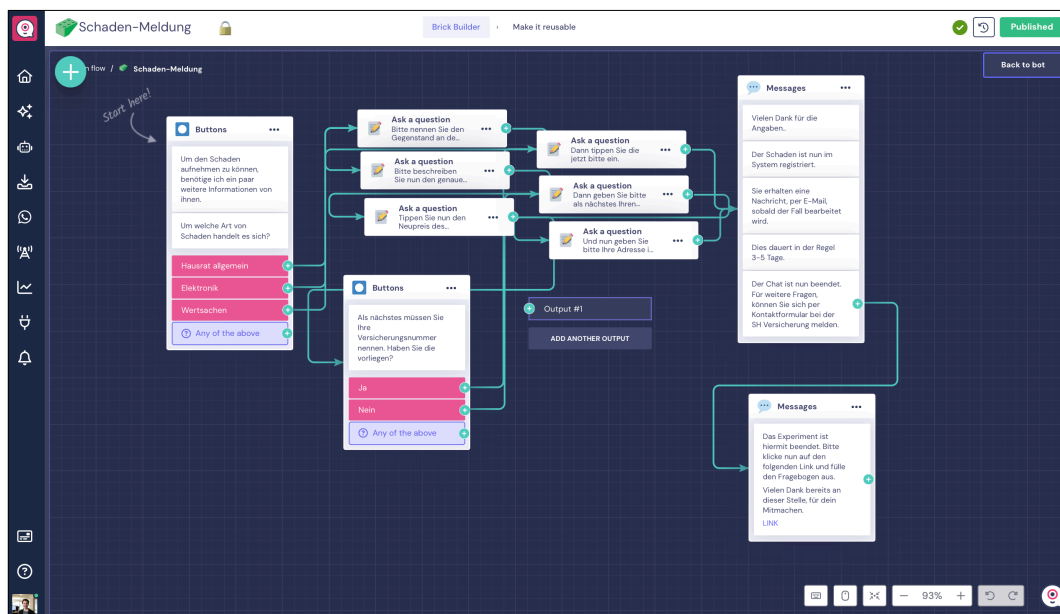
The experiment participants were either forwarded to a chatbot with an emotional, human-like dialogue design or assigned to a neutral chatbot (without an emotional add-on). This resulted in six different chat dialogues (additional insurance with and without emotions, report a claim with and without emotions, changing address with and without emotions). The chatbots were developed with the no-code software-as-a-service platform Landbot. Landbot enables researchers to quickly develop simple rule-based chatbots for either small business use cases or to quickly set up a fictive chatbot situation like for this experiment. For the use of Landbot, no coding skills are required and chatbots can be developed using the graphical user interface, as shown in Figures 19 and 20.

Figure 19: Landbot development Interface



Source: Data collected by the author, according to <https://landbot.io/>.

Figure 20: Landbot development Interface 2



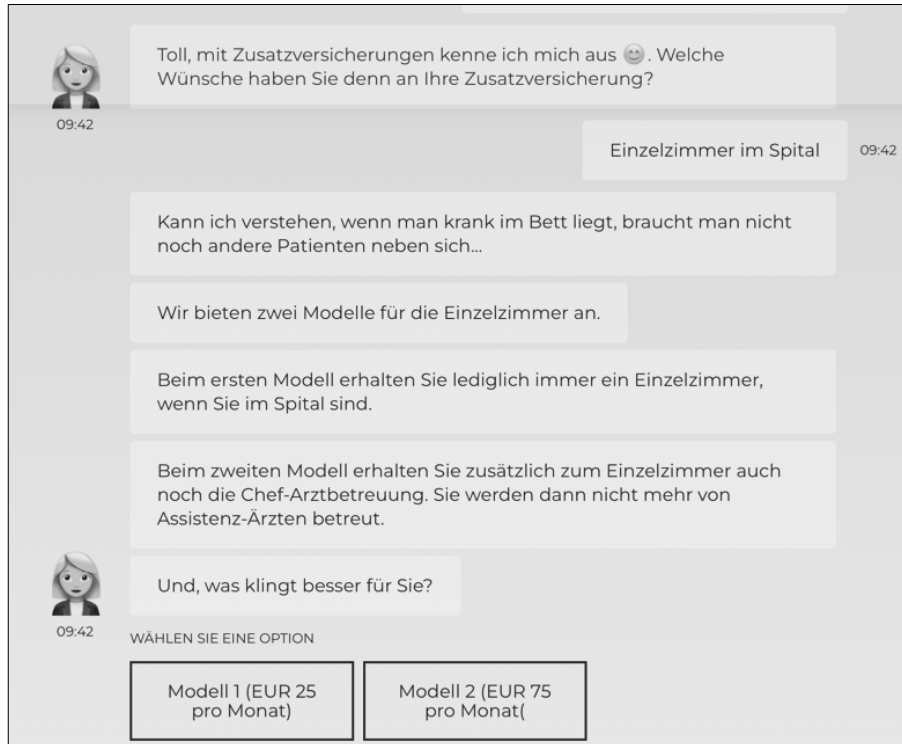
Source: Data collected by the author, according to <https://landbot.io/>.

The chatbot user can use the rule-based chatbot while navigating through the dialogue with the help of pre-defined buttons. Buttons are predefined in the graphical user interface shown above. Figures 21, 22 and 23 show examples from the rule-based chatbots.

The landbot platform was utilized in German, consistent with the overall experiment conducted among German-speaking users. For the sake of full transparency,

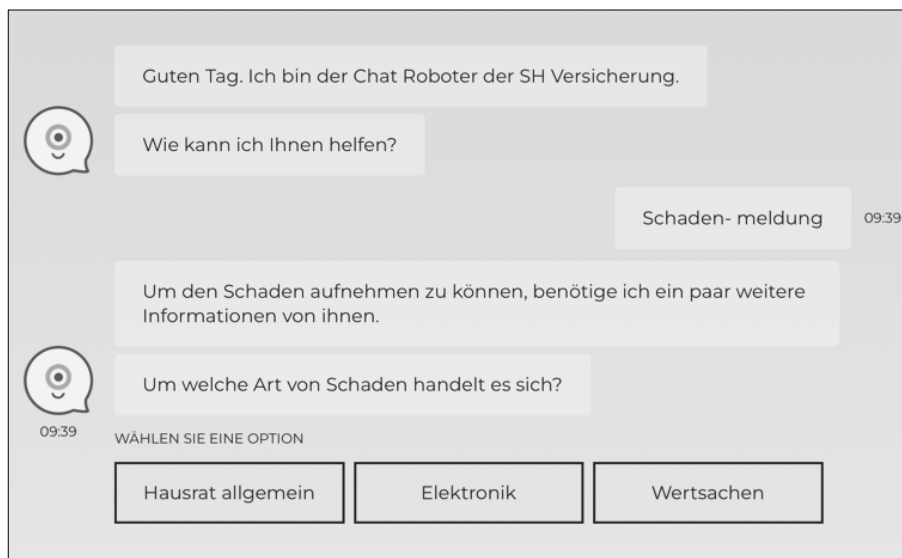
I have not translated the results but present all original materials and outcomes in their original German versions. Additionally further data and translations of this use case are available in appendix B.

Figure 21: Chatbot with emotions – supplementary insurance



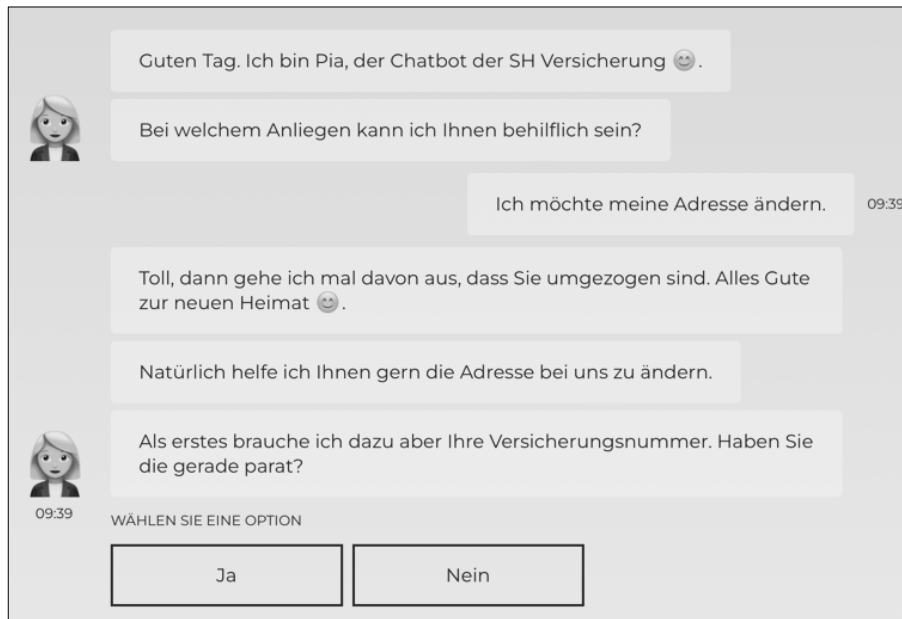
Source: Data collected by the author.

Figure 22: Chatbot without emotions – report damage



Source: Data collected by the author.

Figure 23: Chatbot with emotions – change address



Source: Data collected by the author.

At the end of the chat conversations, all participants were automatically redirected to a questionnaire. This asks for general feedback on the chatbot in relation to the corresponding use case, as well as individual feedback in terms of: quality of service, quality of information, perceived usefulness, tone, perceived fun, efficiency and intention to use or recommend the chatbot again. However, before participants were asked for feedback on the chatbot conversation, they were first requested to rank the emotional state they typically find themselves in when they have such an insurance incident. In doing so, participants were asked to rank how strong the following emotions are for them: interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame or guilt.

For the further creation of the questionnaire, I used the proven scales from the research areas “customer experience” and “service management”. Perceived usefulness was measured with the aid of the scale tested by Davis (1989). Perceived enjoyment was measured with a scale used by Davis et al. (1992). For the measurement of service quality, a tested scale by Rocca et al. (2006) was used. For the classification of information quality, I used the scale validated by Teo (2009). To determine which chatbot users would be most likely to use again or recommend, I used the scales of Zeithaml et al. (1996). To measure efficiency, I employed the scale used by Mimoun et al. (2017). All values were measured with a five-point Likert scale, with the expressions “I do not agree at all” (1) and “I fully agree” (5).

The questions and also the choice of scales are the result of concept, developed in the course of this thesis.

The study follows a between-subjects experimental design in which the independent variable, chatbot emotional design (emotional vs. neutral), is crossed with the situational context (use case type: supplementary insurance / report damage / change address). The outcome is a composite measure of chatbot evaluation quality, which is decomposed into theoretically motivated constructs drawn from established scales in customer experience and service management research. Table 11 shows the full causal model.

Table 11: Conceptual Model – Use Case 2

Independent Variable 1: Chatbot emotional design (between-subjects) • Emotional chatbot • Neutral chatbot	MAIN EFFECT on chatbot evaluation quality The emotional vs. neutral condition is the primary experimental manipulation. Each participant experiences exactly one of the six chatbot variants (2 designs $\times$ 3 use cases).
Independent Variable 2: Situational context (Use case) • Supplementary insurance (advisory, positive mood) • Report damage (problem, negative mood) • Change address (administrative, neutral mood)	MODERATOR of the emotional design effect H1 predicts the emotional chatbot is preferred in the advisory context. H2 predicts the neutral chatbot is preferred in the damage-reporting context. H3 predicts no significant difference in the administrative context. Operationalized as the between-subjects use-case assignment.
User's self-reported emotional state (measured pre-evaluation) • 10 basic emotions rated (Likert 1–5)	CHECK VARIABLE / COVARIATE Verifies that use-case assignment produced the expected emotional response. Users in the damage condition should score higher on sadness, anger, and disgust. Used to operationalize H2's criterion 'users in a negative mood' (mean > 1.7 on sadness/anger/disgust/contempt).
Perceived Usefulness (Davis, 1989)	Does the chatbot help me achieve my goal efficiently? Scale items: e.g. 'Using this chatbot improved my ability to handle this insurance request.' 5-point Likert. Used to compute overall evaluation score (H1: emotional > neutral; H2: neutral > emotional).
Perceived Enjoyment (Davis et al., 1992)	Is the interaction with the chatbot intrinsically motivating or pleasant? 5-point Likert. Expected to be higher for emotional chatbot in advisory context (H1).

Service Quality (Rocca et al., 2006)	Does the chatbot meet the user's service expectations? 5-point Likert. Captures reliability, responsiveness, and professionalism of the interaction.
Information Quality (Teo, 2009)	Is the information provided accurate, complete, and relevant? 5-point Likert. Particularly relevant for the advisory use case (supplementary insurance).
Reuse / Recommendation Intention (Zeithaml et al., 1996)	Would the user use this chatbot again or recommend it? 5-point Likert. Used as primary outcome variable in H01 / H02 t-tests alongside usefulness.
Efficiency (Mimoun et al., 2017)	Did the chatbot help the user complete the task quickly? 5-point Likert. Expected to be valued equally across conditions (administrative tasks) – relevant to H3.
Tone adequacy (custom item)	Was the chatbot's tone appropriate for the situation? 5-point Likert. Directly captures whether emotional vs. neutral design fits the situational context. Feeds into all three hypotheses.

Source: Data collected by the author.

The three hypotheses are directional predictions about how the combination of emotional design and use case type affects overall chatbot evaluation. Situational context moderates the effect of emotional design on evaluation quality, such that the expected direction of the effect reverses depending on whether the user is in an advisory, a problem-reporting, or an administrative context.

Thus the three hypotheses are operationalized as follows, with direct reference to the questions.

H1 (Advisory context – emotional chatbot preferred): Participants assigned to the supplementary insurance use case who interacted with the emotional chatbot are expected to score higher on the composite evaluation index (usefulness, enjoyment, service quality, information quality, reuse intention, efficiency, tone adequacy) than participants who interacted with the neutral chatbot. Tested via independent-samples t-test (H01: no difference; Ha1: emotional > neutral). Confirmed if $p < 0.05$.

H2 (Damage-reporting context – neutral chatbot preferred): Participants assigned to the damage-reporting use case who reported a negative pre-evaluation emotional state (mean > 1.7 on sadness/anger/disgust/contempt) and interacted with the neutral chatbot are expected to score higher on the composite evaluation index than

those who interacted with the emotional chatbot. Tested via independent-samples t-test (H02: no difference; Ha2: neutral > emotional). Confirmed if $p < 0.05$.

H3 (Administrative context – no significant difference): Participants assigned to the address-change use case are not expected to show a significant difference in composite evaluation between the emotional and neutral chatbot conditions. Tested via independent-samples t-test (H03: no difference; Ha3: difference exists). H3 is confirmed if H03 cannot be rejected ($p \geq 0.05$).

This operationalization makes explicit that the three t-tests in Section 3.2.3 are not tests of raw questionnaire items but of the aggregated composite evaluation score across all seven constructs. Here it is important to mention that the emotional state self-report (Part 1) is not a dependent variable but a grouping criterion for H2. It is used to identify the negative mood subgroup within the damage-reporting condition.

Thanks to the use of known or already-used scales, I proceeded without a pre-test here, and started the experiment directly.

The questionnaire and further data on evaluations can be found in full in appendix B.

3.2.3. Approaches to Manage Cost-Experience Conflicts: Results

For the evaluation first Cronbach's alpha was calculated to determine the internal consistency (reliability) of the scales used. The value of 0.883 confirms the reliability of the scales used and thus of the test.

Subsequently, the results of the experiment were evaluated using independent samples T-tests. Null hypotheses (H0) were established based on the hypotheses formed above, which were then either rejected or failed to be rejected based on the test results.

The following Null-hypotheses or alternative hypotheses (Ha) were defined:

- H01: Users in an advisory process evaluate a bot without emotions just as positively as a bot with emotions.
- Ha1: Users in an advisory process rate a bot with emotions better than a bot without emotions.
- H02: Users in a negative mood rate a bot without emotions just as well as a bot with emotions.
- Ha2: Users in a negative mood rate an emotionless bot better.
- H03: For standard requests, such as change of address, users rate a bot without emotions just as well as a bot with emotions.
- Ha3: For standard requests, such as change of address, users rate a bot without emotions differently than a bot with emotions.

H01 could be rejected (p-value 0.044). Thus Ha1 is assumed and the first hypothesis can be confirmed. As a positive bot, mainly attributes were considered that contribute to the recommendation or re-use of the bot, as well as to the increase in efficiency of the user.

H02 could be rejected (p-value 0.001) and thus Ha2 is accepted and the second hypothesis can also be clearly confirmed. Users in a negative mood were those who gave an average value of greater than 1.7 for the emotional states sadness, anger, disgust, and contempt.

H03 could not be rejected (p-value 0.230). Thus Ha3 cannot be accepted either. This means that the third hypothesis is not accepted.

In summary, the experiment confirms the first two hypotheses.

When users take out or inform themselves about supplementary insurance, emotions on the part of the chatbot are generally very well received and are also desired. A chatbot that shows friendly emotions in the context of supplementary insurance is generally recommended more often than a chatbot that shows no emotions.

A chatbot with emotions is perceived as more useful and likeable than a bot without emotions. Users indicate that this type of bot is a suitable solution for them to obtain information about supplementary insurance. Users even state that a chatbot with emotions conveys the impression of individual attention. This is in line with the comments on the chatbot without emotions, which tend to increase the desire for human advice in the context of supplementary insurance. This, in turn, speaks for a more emotional chatbot.

When it comes to the issue of “reporting damage”, users typically are in a rather bad mood and here users generally prefer a chatbot without emotions. The chatbot with emotions seemed too intense. H2 can therefore also be confirmed. When reporting a claim, users only want their request to be dealt with quickly and also want confirmation that the bot has accepted their request. No lack of empathy or emotion was perceived in the chatbot without emotions. However, it is interesting in this context that (as with the “search for additional insurance”) the bot without emotions is more likely to evoke the desire for a human advisor than the chatbot with emotions.

In the case of “change address”, the hypothesis could not be confirmed. However when having a closer and deeper look on the data, it seems that both chatbot versions (with or without emotions) were perceived as useful and suitable. The chatbot with emotions is perceived as slightly more likeable and therefore more likely to be recommended.

In addition to testing the hypotheses, the mean values of the emotion ratings across the individual use cases were compared (table 12, 13, 14). These analyses also show very clearly that emotions are desired in use cases such as additional insurance inquiries, whereas they are considerably less desired in contexts like reporting a claim

when users might be in a rather negative mood. For address changes, the ratings differ only marginally. These descriptive results are consistent with the findings derived from the hypothesis tests.

Table 12: Use case supplementary insurance: group statistics

Emotions (yes=1 / no=2)	n	Mean	Standard deviation	Standard error of the mean
1	21	39.206	0.83103	0.18135
2	19	35.877	0.91349	0.20957

Source: Data collected by the author.

Table 13: Use case changing an address: group statistics

Emotions (yes=1 / no=2)	n	Mean	Standard deviation	Standard error of the mean
1	21	3.57	1.434	0.313
2	19	3.79	1.032	0.237

Source: Data collected by the author.

Table 14: Use case reporting damage: group statistics

Emotions (yes=1 / no=2)	n	Mean	Standard deviation	Standard error of the mean
1	21	1.84	1.436	0.313
2	19	3.48	1.167	0.268

Source: Data collected by the author.

In summary, this means:

- When taking out supplementary insurance, a chatbot with a more emotionally appealing design should be used.
- When reporting damage, a chatbot should be used without an emotionally appealing design.
- When changing an address, emotions hardly play a role and data is not clear so far. An emotional design can therefore be dispensed with for cost reasons.

The application test shows that certain aspects must be considered when automating customer dialogues:

First, it is important to note that the extent of automation of customer dialogues should result from the data-based application of the value irritant matrix. Automated

dialogues generally do not allow conclusions to be drawn about further revenue potential or general insights from errors in the customer dialogue.

Second, it is useful to classify the dialogues to be automated according to “conclusion-oriented”, “problem-oriented” and “administrative/neutral” dialogues in order to evaluate the possible use of emotionally appealing dialogue designs.

Third, it should be noted that all dialogue designs with the customer must be tested regarding the benefits to be realized and the associated costs. It is useful to derive two hypotheses from the classification of the previous point and to test them.

It should also be noted that the examples presented here still amount to rule-based solutions, which are ultimately always designed by humans or by the company that wants to automate dialogues. In the future, instead of rule-based solutions, algorithms should be taught so as to make it possible for the conversational AI system to find out for itself what emotional state the user is currently in, and hence to be able to react appropriately. Soon, it will no longer be a question of automating only individual small dialogues or use cases, but companies will be required to automate many dialogues and different use cases at the same time. It will then be even more important for a conversational AI system to find out for itself what emotional state the human counterpart is in and to be able to control the dialogues accordingly.

3.2.4. Requirements on Conversational AIs to Meet Economic and User Expectations – Implications for Future Conversational AI

The literature research summarized before, and the laboratory experiment provide answers to RQ1.2:

RQ1: What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.2: What design criteria should conversational AI systems incorporate to balance user objectives, such as quick and correct solutions, with company goals, including economic efficiency and customer experience?

From the findings of this use case 2, I derive generalizable capabilities for my final conceptual framework. These emphasize the need for conversational AI systems to consider multiple influencing factors of the user, such as their situational context, prior experiences and personality traits. Additionally, external influences, such as the economic goals of the organization, must also be accounted for. Based on these insights, the following high-level design requirements emerge:

Table 15: Capabilities 7-9

Number	Capability	Explanation	Source
7	User Context Awareness	Conversational AI systems should dynamically adapt their interactions based on the user's situational context, including emotional state, current needs and goals, to provide relevant and meaningful responses.	Use Case 2
8	Personalization	The systems should consider the user's past interactions, preferences and personality to create tailored and engaging dialogue flows that improve the user experience.	Use Case 2
9	Alignment with Organizational Objectives	All design elements, including dialogue styles and automation levels, should balance the user's needs for efficiency and satisfaction with the organization's goals of cost control, revenue optimization and operational efficiency.	Use Case 2

Source: Data collected by the author.

The link between the experimental results and the derived capabilities is as follows: Capability 7 which requires conversational AI systems to detect and adapt to the user's current affective state is directly motivated by the statistically significant result of H_2 ($p = 0.001$), demonstrating that users in a negative emotional state evaluate emotionless chatbot designs significantly more favourably. This finding establishes the empirical basis for the requirement that the system must be capable of recognising the user's affective state and selecting its dialogue design accordingly.

Capability 8 which requires the system to align its dialogue design with the type and purpose of the interaction (conclusion-oriented, problem-oriented or administrative) is derived from the combined results of H_1 and H_3 : the confirmed difference between advisory and administrative contexts shows that a single uniform dialogue design cannot simultaneously optimise user experience across different interaction types.

Capability 9 which requires conversational AI systems to account for economic constraints and organisational tone-of-voice policies alongside user experience goals is derived from the broader observation that cost–experience trade-offs (as formalised in the value irritant matrix) cannot be resolved without embedding economic considerations directly into the design logic of the system.

These criteria ensure that conversational AI systems are designed to harmonise the interests of users and organisations while enabling scalability and adaptability in diverse application scenarios.

Use Case 2 demonstrates that a single chatbot design cannot be optimal across all use cases: the emotional design that improves customer experience in an advisory context actively harms it in a problem-reporting context. The implication for the FCCC framework is clear, the system must be able to adapt its dialogue style to the situational context of the user. But situational context is only one dimension of the user. The FCCC Framework, grounded in Character Computing (El Bolock et al., 2020), must also adapt to stable user characteristics: personality, age, gender, and prior experience with chatbots. Whether these characteristics produce systematic, predictable differences in preferred chatbot behaviour is the question addressed by use case 3.

3.3. AI-Supported Advisory Systems in the Financial Sectors

This section again reflects on the use of conversational AI systems in customer service and considers the situation in which a user finds himself, his personality and his age (use case 3). If this is compared with the dimensions of CC (El Bolock et al., 2020) presented in chapter 2, the two dimensions “affective state or situation” and “personality” are in focus here.

Parts of this chapter were already published in my previous work: Hundertmark, S. (2023). *Be the Bot – Versicherungskunden in der Rolle eines Chatbots*.

3.3.1. AI-Supported Advisory Systems in the Financial and Insurance Sectors: Defining the Research Question and Mythology

In the chapter 1 it was shown that, with the help of an extensive literature search, research on the design of conversational AI systems and their effects has gained significant interest in recent years. It is therefore not surprising that there are already countless researchers, such as Crollic et al. (2021), Ng et al. (2020), Ghandeharioun et al. (2019), Katayama et al. (2019) or Mozafari et al. (2021), who have investigated the difference between emotionally designed chatbots that also show empathy and neutral chatbots regarding application scenarios and effects.

While previous research on the design of conversational AI systems usually only asks the user to evaluate existing automations or chat dialogues, for this study, I asked chatbot users how they would like the chatbot dialogue to be in different situations for an insurance customer. With the help of the novel experiment design, I wanted to gain additional insights into the design of chatbot dialogues and, at the same time, test the novel method. The focus of this use case 3 is on insurance chatbots that deal with cases such as reporting a claim, changing an address or taking out additional insurance. These three cases were selected for the experiment in consultation with Swiss insurance companies. All three are ideally suited for automation in customer service, considering the value irritation matrix (Price et al., 2008). To test different kinds of chatbot behaviour, I compare three important characteristics, namely, level of empathy, response length and degree of reference to previous user message.

With the help of the experiment conducted in this chapter I find new insights to RQ1.3:

RQ1: What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.3: How do the requirements for conversational AI systems differ depending on the user's situation, age, gender, personality and experience with chatbots?

For this experiment, 39 insurance customers were asked to put themselves in the role of a chatbot and to react like a bot to given user messages in the period from 11 November to 1 December 2022. The insurance customers were recruited from an online panel. In the recruitment process for the test users, I ensured a balance in gender and age (customers between 18 and 65 years old). The panel exclusively comprises German-speaking insurance customers from Germany and Switzerland. As this is an online panel, it is possible that the participants may have a slightly higher online affinity than the general average of the overall population.

The method of having users write the dialogue with a conversational AI system themselves instead of giving feedback on existing dialogues, is completely new in research on conversational AI design. Here I use chatbots as an example of a conversational AI system.

So far, researchers such as Völkel et al. (2020) have only used this method to find out more about the desired tone of voice assistants, but not for the design of written chat dialogue.

The experiment participants were given four dialogue segments of a user for each of the three use cases: report damage, change address and take out supplementary insurance. They were asked to write a chatbot response to match. The dialogue was designed in such a way that the user starts the conversation and expresses his request in the first message (report damage, change address or take out additional insurance). Figure 24 shows an excerpt from one of the dialogues that the participants were given to complete. The overall experiment was conducted among German-speaking users. Additionally translations and further data from this use case can be found in appendix C.

Figure 24: Example questionnaire

Im folgenden siehst du einen Dialog zwischen einem Menschen und einem Chatbot. Die pinken Felder sind der Chatbot. Bitte stelle dir vor, dass du der User bist! Welche Antworten würdest du dir jeweils von dem Chatbot wünschen?

Antwort 1 *

Meine Antwort

Antwort 2 *

Meine Antwort

Source: Data collected by the author, Full length of questionnaire in Appendix C.

At the end of the dialogue, the study participants were additionally asked about their age, gender and experience with chatbots. When asked about age, it was possible to choose from the following four age groups: 18–30, 31–40, 41–50 and 51–65. When asked about experience (if the participants had used a bot before), there were the following three answer options: Yes – several times, Yes – once, No.

Finally, the participants were asked to assign their own personality to one of the five groups of the “big five”: openness to experience (open-mindedness), conscientiousness (perfectionism), extroversion (sociability, extroversion), agreeableness (consideration, cooperativeness, empathy) and neuroticism (emotional lability and vulnerability). The “Big Five”, also known as the FFM, is a model of personality psychology (Allport et al., 1936). It is also known as the OCEAN model, after the corresponding initial letters: openness, conscientiousness, extroversion, agreeableness and neuroticism. The “Big Five” have been proven by a large number of studies and are

now internationally regarded as the universal standard model in personality research (John et al., 2008).

The complete set of chatbot responses written by all 39 participants is provided in appendix C, organised by use case and participant group. Each response has been coded for the three dimensions (empathy, response length and reference).

3.3.2. AI-Supported Advisory Systems in the Financial and Insurance Sectors: Evaluation and Results

First, the chat responses of the experiment participants were manually reviewed and divided into the following categories, as shown in table 16.

Table 16: Main findings from experiment

Dimension	Expression 1	Expression 2
Emotions	Yes	No
Response length	Short	Long
Reference to previous message	Yes	No

Source: Data collected by the author.

For the purposes of this analysis, responses were classified as 'short' if they contained fewer than 5 words, 'medium' if they contained between 6 and 10 words, and 'long' if they contained more than 10 words.

Empathy was considered present when words such as “unfortunately,” “I’m sorry,” “I understand,” or similar expressions were included.

Reference to the previous message was considered present when follow-up questions or repetitions related to the prior message could be identified.

Subsequently, only the correlations between the use case and the characteristics of the individual dimensions were calculated. This was followed by correlation analyses between user characteristics and the values of the dimensions.

With the help of the first correlation analysis, the following first findings could be obtained (Appendix G includes the collected data for this use case):

If empathy, then most likely when reporting damage

The majority of participants (76%) wrote their chatbot answers without empathy. It can therefore be assumed that chatbot users generally do not expect empathy from a chatbot.

When comparing the three use cases, however, it is noticeable that empathy is most likely to come into question when reporting damage. Here, customers wrote the

most messages with empathy compared to the other use cases. Changing an address comes in second place. When it comes to taking out an additional insurance policy, though, only very few customers expect empathy.

Long responses, especially when changing the address

Although almost all user experience guidelines on conversational design state that chatbot responses should be rather short, the insurance customers mostly formulated relatively long responses (79%). This was particularly noticeable when changing an address. But the answers also tend to be long when reporting a claim. The shortest answers were given when taking out an additional insurance policy.

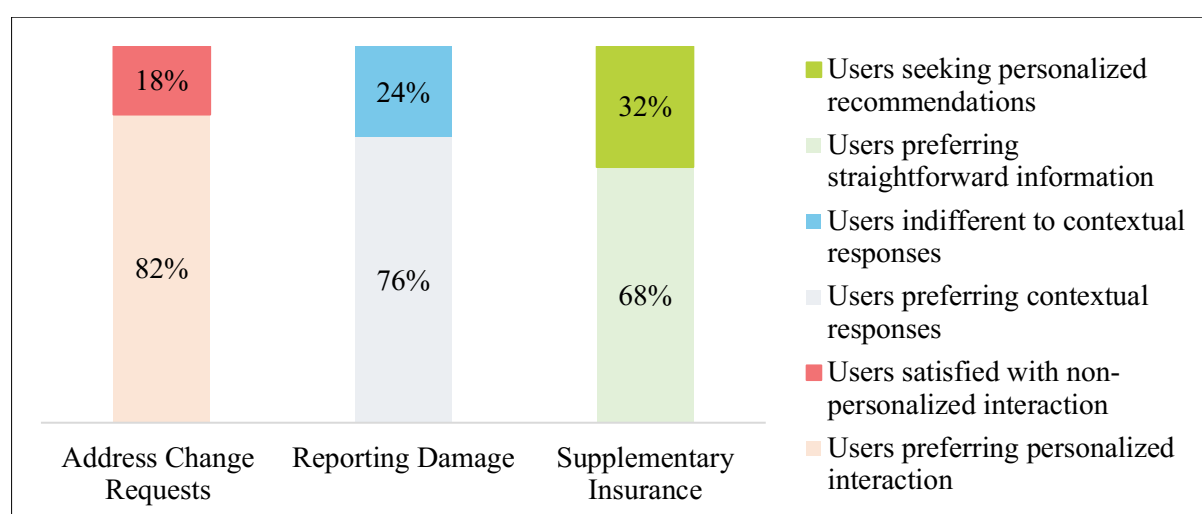
Change address and report damage does not work without reference, additional insurance works without reference

The results show very clearly that customers want a direct reference from the chatbot when changing an address (82%). Statements such as, “Sure! I need some data to change your address. What is your name?” was often written here. Even when reporting damage, referring to the customer message seems to be an important point. So customers want a dialogue. The bot should include the context. It is different when taking out supplementary insurance. Here, the reference is least important to the customers. Of course, they want to be understood and receive a suitable offer, but the chatbot responses tend to look like this: “Option 1 costs CHF 75 once. Option 2 would be CHF 60 and Option 3 CHF 45”.

With the help of the second correlation analysis, deeper insights into dialogue design could be made depending on age, gender, personality or experience with chatbots. The following important findings are presented here:

Figures 25, 26, 27 and 28 provide an overview of the key results.

Figure 25: Key results – address change request, reporting damage, suppl. Insurance



Source: Data collected by the author.

Address Change Requests

The data indicates a strong preference for personalized interaction when customers need to change their address. Specifically, the chatbot should reference the customer's previous messages to create a coherent dialogue. For instance, a typical effective response includes statements like, "Sure! I need some data to change your address. What is your name?" This approach was favoured by a significant majority of users, suggesting that personalized engagement enhances user satisfaction.

- **Percentage of users preferring personalized interaction: 82%**
- **Percentage of users satisfied with non-personalized interaction: 18%**

Reporting Damage

When reporting damage, the importance of context and reference to previous messages remains high. Users expect the chatbot to acknowledge their initial input and maintain a conversational tone. For example, a suitable response would be, "I'm sorry to hear about your iPhone. Can you provide more details about the damage?" This preference was notably higher among male users compared to female users.

- **Users preferring contextual responses: 76%**
- **Users indifferent to contextual responses: 24%**

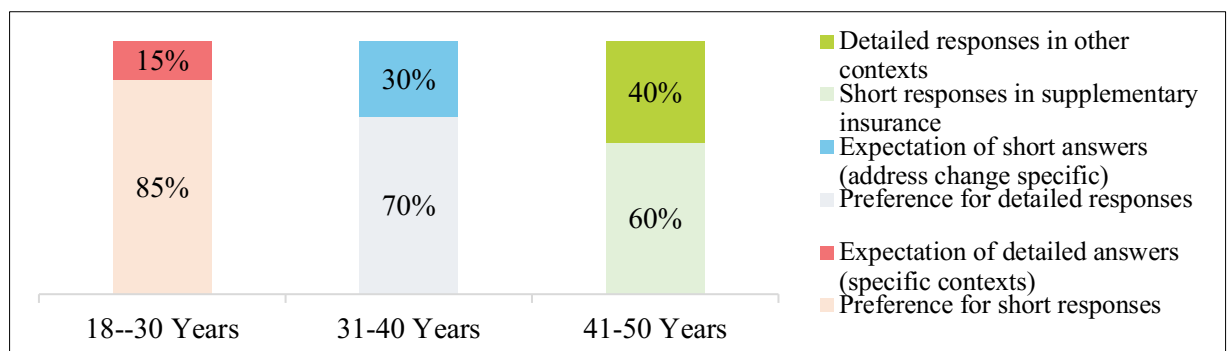
Supplementary Insurance

In contrast, when users inquire about supplementary insurance, the need for personalized interaction diminishes. Users primarily seek clear, concise information about their options. Responses typically include detailed pricing and benefits without heavy personalization, such as, "Option 1 costs CHF 75 once. Option 2 would be CHF 60 and Option 3 CHF 45".

- **Users preferring straightforward information: 68%**
- **Users seeking personalized recommendations: 32%**

Insights by Demographics

Figure 26: Demographics overview



Source: Data collected by the author.

The correlation analysis further breaks down user preferences based on age, gender, personality and experience with chatbots, offering deeper insights into how these factors influence interaction styles.

18–30 Years:

- Preference for short responses: 85%
- Expectation of detailed answers (specific contexts): 15%

31–40 Years:

- Preference for detailed responses: 70%
- Expectation of short answers (address change specific): 30%

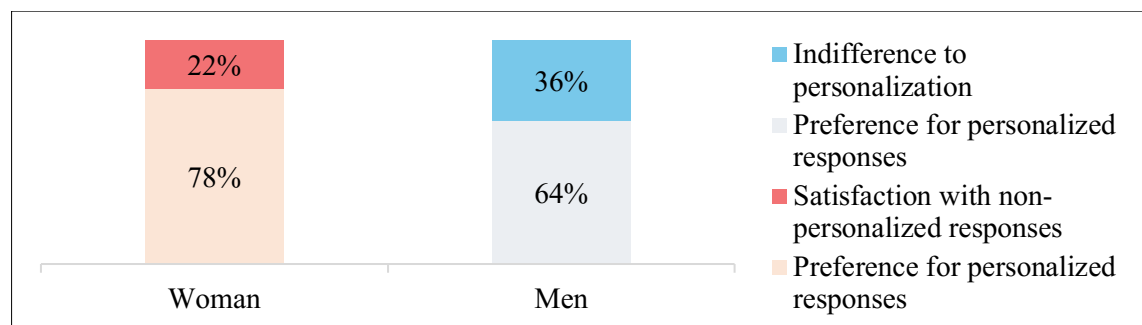
41–50 Years:

- Short responses in supplementary insurance: 60%
- Detailed responses in other contexts: 40%

Younger users (18–30) prefer concise interactions, while older age groups (31–50) lean towards detailed responses, particularly when addressing changes or reporting damage.

Gender Difference

Figure 27: Gender distributions



Source: Data collected by the author.

Women:

- Preference for personalized responses: 78%
- Satisfaction with non-personalized responses: 22%

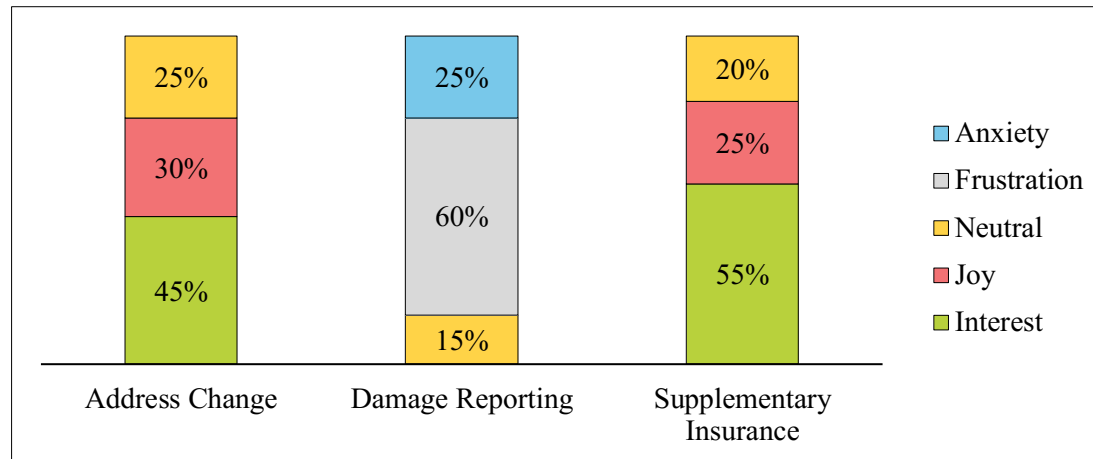
Men:

- Preference for contextual engagement: 64%
- Indifference to personalization: 36%

Women generally expect more personalized and contextual interactions compared to men, who show a slightly higher agreeableness for straightforward, impersonal communication.

Emotions in different situations

Figure 28: Emotions in Different Use Cases



Source: Data collected by the author.

Figure 28 captures the biggest emotional feelings of users across different scenarios. For instance, when changing an address, users often feel neutral or slightly positive, whereas reporting damage typically elicits negative emotions such as frustration or anxiety. Understanding these emotional contexts can help in designing more empathetic chatbot responses.

Address Change:

- Interest: 45%
- Joy: 30%
- Neutral: 25%

Damage Reporting:

- Frustration: 60%
- Anxiety: 25%
- Neutral: 15%

Supplementary Insurance:

- Interest: 55%
- Joy: 25%
- Neutral: 20%

Taken together, all results lead to the following conclusions regarding the preferred chatbot design across different user groups.

- Regardless of the use case, the youngest age group (18–30 years) show the shortest answers. All other age groups would like the chatbot to provide significantly longer answers.
- Regardless of the use case, women tend to expect a reference to the previous message. For men, it seems generally less important that the chatbot responds to their messages with expressions such as “Sure, I’d love to...”.
- Regardless of the use case, the personality groups “openness to experience” and “conscientiousness” tend to write long responses. The groups “extroversion” and “agreeableness” write significantly shorter messages.
- Regardless of the use case, it is noticeable that the “extroversion” personality group places the least value on reference to the individual messages.
- When it comes to the use case address change, it is noticeable that only the age group 31–40 years prefers short answers. All other age groups would like to receive much longer answers from the chatbot.
- When changing the address, the majority of women tend to give short answers, while men tend to give long answers.
- When it comes to changing the address, it is noticeable that the personality group “extroversion” writes significantly shorter messages than the other groups.
- When it comes to changing the address, it is noticeable that the personality group “extroversion” places the least value on reference to the individual messages.
- When reporting claims, the youngest insurance customers (18–30 years) expect the shortest answers. The other age groups keep their bot answers significantly shorter.
- When it comes to reporting damage, the men agree that they would rather have a reference or a statement about their message, whereas no clear conclusion can be made about this among the women.
- When reporting damage, it is noticeable that the personality group “extroversion” writes significantly shorter messages than the other groups.
- When it comes to damage, it is noticeable that the personality group “extroversion” places the least value on reference to the individual messages.
- Interestingly, when it comes to taking out supplementary insurance, the 41–50 age group gives shorter answers than all other age groups.
- When taking out supplementary insurance, it is noticeable that the “extroversion” personality group attaches the least importance to reference to the individual messages.

- Across all responses, it is easy to distinguish users who have already chatted with a chatbot more often from those who have only chatted with a bot once or never before. Experienced bot users are those who have written to a chatbot more than once. This group of users generally wants more detailed answers from a chatbot. Experienced chatbot users expect the chatbot to refer to their previous answer. With regard to empathy, no difference was found between users who had already chatted with bots once or several times and those who had never chatted.

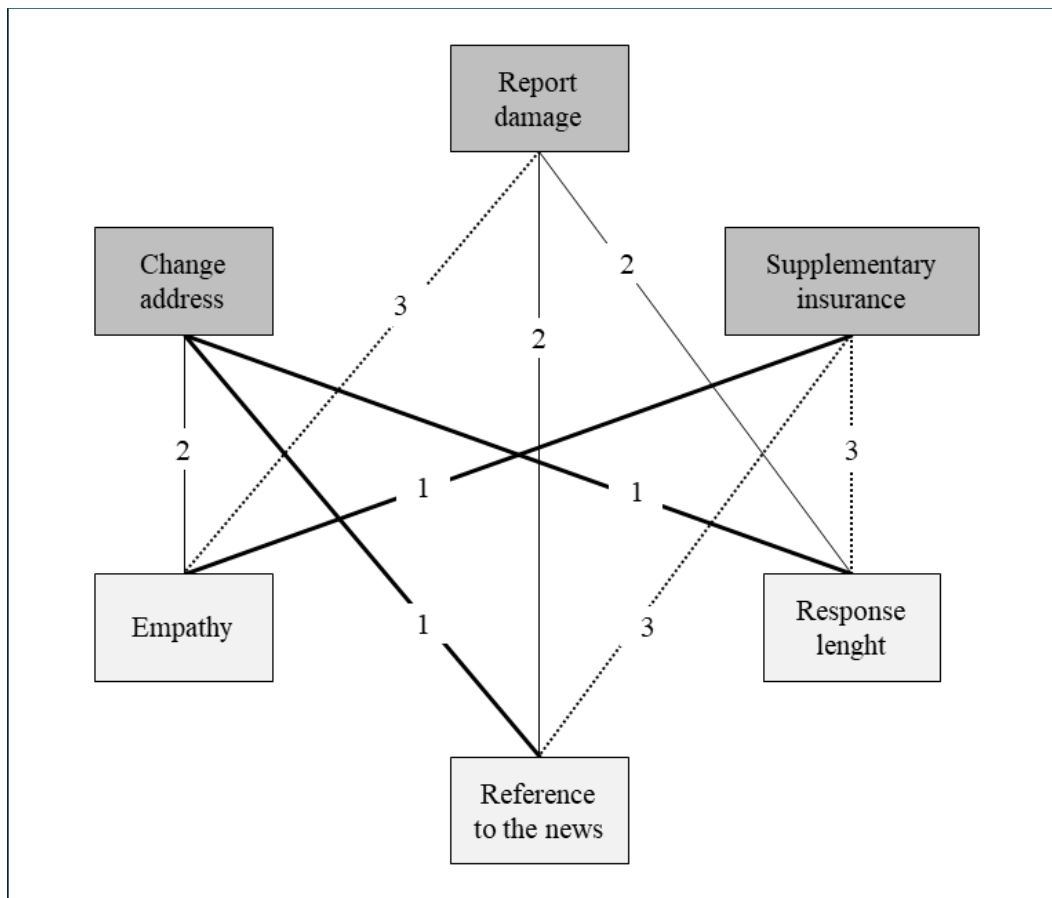
In summary, the critical importance of personalized, contextually aware chatbot interactions across various customer service scenarios is highlighted. By tailoring responses to user demographics, emotional states and prior chatbot experiences, companies can significantly enhance customer satisfaction and engagement. The numerical data underscores the need for nuanced and adaptive chatbot designs that cater to the diverse preferences of users.

At this point, it is important to emphasize again that the research design used here is novel in the field of chatbot design research. However, the results show that this method does provide good and relevant information. Furthermore, no research is known to date in which different customer concerns, customer characteristics and expectations of the chatbot dialogues have been analysed in terms of empathy, length and reference.

In my experiment (use case 3), participants were asked to assign themselves to one of the five known personality types (big five) (Allport et al., 1936). In other studies, participants are usually asked to choose between predefined adjectives that best describe them. I deliberately shortened this part in my experiment, as writing the bot dialogues already requires more time for most participants than they are used to in other studies.

While, in general, the statements probably apply, a detailed examination of the customers shows isolated deviations depending on age, gender and personality, as can be seen in Figure 29 (numbers indicate the order, 1 applies most).

Figure 29: Summary results Research Question 1.3



Source: Data collected by the author.

In general, long or rather long chat answers are expected for all three use cases, but the younger groups prefer quite short answers for changing addresses and reporting damage. The same applies to extroverted users, who generally expect shorter answers and also attach the least importance to references from the bot. When it comes to changing address, it is mainly the men who write long chatbot answers, while the women tend to keep them short.

Finally, different expectations can be observed among users who have already chatted frequently with chatbots and those who have never chatted with bots. It should be noted here that the number of users who have never chatted with a bot will probably decrease significantly in the next few years. This is due, among other things, to technological progress and companies such as OpenAI (2022), which make their latest chatbot available to the general public and, at the same time, ensure a large reach in the most diverse media.

3.3.3. Requirements for Successful Conversational AI Advisory Systems – Future Conversational AIs

The quantitative study presented above as use case 3 provides answers to RQ1.3:

RQ1: What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.3 How do the requirements for conversational AI systems differ depending on the user's situation, age, gender, personality and experience with chatbots?

Our results demonstrate that companies, including insurance providers and others, should adopt a differentiated approach to designing chatbot dialogues. Customer expectations regarding chatbot responses, such as length, empathy and reference, vary significantly depending on the user's situation, age, gender, personality and experience with chatbots. To ensure a positive customer experience for all, companies must plan for diverse dialogue variations and deploy them appropriately based on the specific use case and user profile.

From these findings, I derive the following general capabilities for the artefact:

Table 17: Capabilities 10-12

Number	Capability	Explanation	Source
10	Contextual Dialogue Adaptability	Conversational AI systems should dynamically adjust response characteristics such as length, empathy and detail level based on situational factors and user profiles. This ensures alignment with individual user expectations and optimizes user satisfaction.	Use Case 3
11	User Profile Differentiation	The system must be capable of recognizing and categorizing users based on traits such as personality, age, gender and prior chatbot experience. This enables the delivery of personalized interactions tailored to diverse user needs.	Use Case 3
12	Scalable Personalization Strategies	To balance personalization with operational efficiency, companies should implement scalable mechanisms for tailoring chatbot dialogues.	Use Case 3

Source: Data collected by the author.

These capabilities provide a foundation for designing conversational AI systems that effectively address the diverse needs of users while ensuring consistency in delivering high-quality interactions across different scenarios and application areas.

Use case 3 confirms that user requirements for chatbot dialogue, empathy level, response length, and contextual reference, vary systematically with age, gender, personality type, and chatbot experience. This finding has a direct implication for the FCCC Framework. The system cannot rely on a single pre-configured dialogue style, nor on a small set of manually defined user segments. It must be capable of inferring user characteristics from the conversation itself and adapting dynamically. The next section consolidates the design requirements generated by use cases 1 to 3.

3.4. Findings on Requirements Based on Knowledge Transfer and Customer Service in a Nutshell

The three experiments (use cases 1-3) conducted in the domains of knowledge transfer and customer service provide me with valuable design requirements from different perspectives. These experiments enabled me to comprehensively address RQ1, including all its sub-questions:

Research Question 1 (RQ1):

What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.1: What design criteria should conversational AI systems incorporate to optimize knowledge transfer while ensuring a positive user experience and supporting life engineering goals?

RQ1.2: What design criteria should conversational AI systems incorporate to balance user objectives, such as quick and correct solutions, with company goals, including economic efficiency and customer experience?

RQ1.3: How do the requirements for conversational AI systems differ depending on the user's situation, age, gender, personality and experience with chatbots?

Key aspects include user-friendly, audience-adapted communication and the integration of flexible interaction modes to enhance engagement. Systems should support customizable content, incorporate multimedia elements and maintain economic accessibility while ensuring user privacy and trust. Personalization is emphasized through adaptive responses based on user context, past interactions and profiles, with scalable strategies to balance individualization and efficiency. Additionally, aligning system design with organizational objectives ensures a balance between user satisfaction and operational goals.

When comparing the results of use case 2 and use case 3, several contradictions become apparent. These can be attributed to the different methodological approaches used in each case. In use case 2, participants were merely asked for their opinion, whereas in use case 3 they had to actively take on the role of the chatbot. At first

glance, I assume that the statements regarding user preferences in use case 3 are closer to real-world behaviour.

However, these inconsistencies have no impact on the definition of the design requirements for my artefact. Instead they assume that a human designer, or a pre-configured rule, determines which dialogue style to apply for which user in which situation. Use case 2 relies on a fixed between-subjects assignment, use case 3 has human participants manually write the «correct» chatbot response. Neither experiment tells us whether a conversational AI system can make these determinations *automatically*, from text input alone.

The FCCC Framework, grounded in Character Computing (El Bolock et al., 2020), requires exactly this. A system that estimates the user's personality, affective state, and situational context from their messages, without access to sensors or pre-collected profile data. Whether this is feasible, whether the CC triad can be derived from text messages alone, is a foundational question that must be answered empirically before the artefact can be built. This is the question addressed by use case 4 in Section 3.5.

Table 18 provides an overview of all identified design requirements. In chapter 3.5 I will expand this list with an additional experiment and subsequently eliminate or consolidate any duplicate requirements.

Table 18: Design requirements from the use cases

Number	Capability	Explanation	Source
1	Audience-Adapted Communication	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Use Case 1
2	Flexible Interaction Modes	Combine structured interaction (e.g. pre-defined buttons) with opportunities for user-driven input to enhance engagement and adapt to users' learning preferences.	Use Case 1
3	Adaptive Content Design	Enable the integration of knowledge documents to ensure that learning content remains customizable and aligned with specific needs, allowing organizations to maintain control over the material.	Use Case 1
4	Economic Accessibility	Maintain a cost structure that balances system sustainability with affordability for a broad user base.	Use Case 1
5	Multimedia Integration	Incorporate diverse media elements such as images, videos and graphics to enhance interactivity and accommodate various learning styles.	Use Case 1

Number	Capability	Explanation	Source
6	Privacy and Trust	Ensure user anonymity and data protection to build trust and encourage honest participation.	Use Case 1
7	User Context Awareness	Conversational AI systems should dynamically adapt their interactions based on the user's situational context, including emotional state, current needs and goals, to provide relevant and meaningful responses.	Use Case 2
8	Personalization	The systems should consider the user's past interactions, preferences and personality to create tailored and engaging dialogue flows that improve the user experience.	Use Case 2
9	Alignment with Organizational Objectives	All design elements, including dialogue styles and automation levels, should balance the user's needs for efficiency and satisfaction with the organization's goals of cost control, revenue optimization and operational efficiency.	Use Case 2
10	Contextual Dialogue Adaptability	Conversational AI systems should dynamically adjust response characteristics such as length, empathy and detail level based on situational factors and user profiles. This ensures alignment with individual user expectations and optimizes user satisfaction.	Use Case 3
11	User Profile Differentiation	The system must be capable of recognizing and categorizing users based on traits such as personality, age, gender and prior chatbot experience. This enables the delivery of personalized interactions tailored to diverse user needs.	Use Case 3
12	Scalable Personalization Strategies	To balance personalization with operational efficiency, companies should implement scalable mechanisms for tailoring chatbot dialogues.	Use Case 3

Source: Data collected by the author.

The consolidation in Section 3.4 reveals that use cases 1–3 generate extensive knowledge about *what* a conversational AI must adapt to, but they do not answer *how* it can learn this from the conversation itself. The FCCC Framework's Character Computing component requires that the system estimate the full CC triad – personality (C), behaviour (B), and situational context (S) – from user messages, without sensors. Use Case 4 tests this assumption directly with a large-scale dataset ($n = 2000$), asking whether text-based message features are sufficient to derive the triad. The answer shapes the entire subsequent design of the artefact.

3.5. Conversational AI as System to Calculate User Character

This section again uses chatbots to represent conversational AI systems. However, rather than focusing only on one dimension of CC (El Bolock et al., 2020), as I did in the other use cases, this laboratory experiment (use case 4) aims to investigate and evaluate the entire triad of CC using quantitative research approaches.

The insights derived from this use case are crucial for the application of CC methods within my subsequently developed conceptual framework. Utilizing the results of this use case allows me to better assess the effort and deployment of CC within the entire framework, and to supplement it with additional suitable and necessary methods alongside CC.

The following research and findings are based on my work, published earlier: Hundertmark, S., Hafner, N., & Portmann, E. (2023). Dialog-Anpassung an den Charakter des Benutzers – Anwendung der Methoden des Charakter Computing.

3.5.1. Conversational AI as system to calculate users character: Defining the Research Question and Methodology

The methodological approach used in the present study is a novelty in the current research landscape because, contrary to existing international research trends, not only one or two dimensions of a user's character is used to determine the optimal design of the conversational AI system. Indeed, it is no longer sufficient to identify only one or two characteristic dimensions of the user, rather, it is essential to capture all variables of the CC triad (El Bolock et al., 2020) to react adequately to diverse user typologies. The triad can be used to deduce the user's anticipated behaviour in any situation, which in turn determines the conversational AI system's communication preferences, especially regarding the communication mode. The fundamental elements of this triad include the individual personality and the affective state of the user, supplemented by the respective situational contextuality.

The following use case 4 aims to present and test an exemplary scenario in which insurance clients report damage to their iPhone. The experiment is performed without any specific sensors – only text messages from users are considered.

According to El Bolock et al. (2020), a character is the totality of all components that determine human behaviour in a given situation. It consists of general, stable factors (e.g. personality and sociocultural embeddedness) and current factors (e.g. affect, mood, physical state, appearance and cognitive state). CC focuses on the study of human character from the perspective of computer science and psychology. The triad character-behaviour-situation is a possible basis for this research, although the interactions and components of interest within the triad vary depending on the application area. CC consists of three modules: sensing character, adapting to character and creating artificial characters. Each module is a distinct research area with multiple

challenges and new opportunities. To date, each module has been studied separately in different research areas. For example, research in the area of affective and personality computing aimed to address some of the problems mentioned, but without including all character components. Instead, each area relies on one character component, such as affect or personality, to explain human behaviour. Most existing approaches to understanding and thus adapting a computer to human behaviour have their roots in one of the two main fields of research: computer science and psychology (El Bolock et al., 2020).

CC is one of the approaches that addresses the questions from a joint perspective of computer science and psychology in parallel and leads to a formalized model that maps the triad character-behaviour-situation and thus enables all resulting application possibilities (El Bolock et al., 2020).

In real life, emotional behaviour is largely expressed through body signals. These signals can be divided into two main categories or domains: behavioural signals (e.g. facial expressions, hand and body gestures, speech or written text) on the one hand, and physiological signals (for example changes in heart rate, pulse, respiration, eye blink, pupil size, galvanic skin conductance, skin temperature – to name a few) on the other (El Bolock et al., 2020). Thanks to new technologies in the field of human-computer interaction (HCI) and the development of new sensor devices, it is in principle possible to record, collect and analyse any of these overt behavioural responses or tacit and implicit physiological signals from the user (El Bolock et al., 2020). However, the abundance of bodily signals that can indicate emotions, their signal complexity, and dynamic nature can pose challenges for computer algorithms. While humans are good at recognizing, naming, interpreting, and predicting emotions and feelings, this becomes a complicated task for computers when sensory input does not allow for interpretation (El Bolock et al., 2020). Further, this raises the big question of how realistic it is that computer systems, that are supposed to adapt their behaviour to humans, will always automatically have access to the data from the possible sensors. Do people use the sensors at all, or do they merely share their emotional and cognitive states with the help of speech, facial expressions or body language? El Bolock et al. (2020) also point out that collecting “character” data from users and capturing their behaviour from various online sources raises ethical and privacy concerns when users cannot be informed about the purpose of the data collection and when there is no verbal and written consent from users.

In practice, users of conversational AI systems are not associated with sensors. Their primary interaction signals manifest in text messages. Consequently, future conversational AI systems must have the potential to deductively determine, based on these messages, which personality category a user belongs to, which affective state they are in, and which situational context they are experiencing. The subsequent experiment aims to ascertain whether the information acquired by chatbots solely from the text messages of users is sufficient to calculate the triad, known from CC.

Thus, the overarching RQ2 is to be addressed in this part of this use case 4:

RQ2: How effectively can methods from Character Computing be integrated into contemporary conversational AI systems to interpret the complete triad of user characteristics (personality, affective state and behaviour), and what existing gaps need to be addressed to enable a chatbot to adapt its interactions accordingly?

To determine the essential components of the triad of CC, it is necessary to determine the personality and the affective state of the user, supplemented by the respective situation in which the user finds himself. The research questions 2.1 and 2.2 serve to test whether text messages are enough to calculate the triad. Should my experiment help to answer the following two sub-questions in a positive manner, I will be able to approve that pure text is sufficient for a conversational AI to determine the user's character and to adapt its behaviour accordingly:

RQ2.1: Can a conversational AI system derive the complete triad from Character Computing (including personality, affective state and situational behaviour) using features extracted from user messages?

RQ2.2: Can a conversational AI system define users' expectations regarding chatbot interactions when identifying users' personalities and affective states?

To determine the individual variables of the CC triad and to answer the RQs 2.1 and 2.2 a quantitative online survey was conducted with 2000 respondents.

The subjects were recruited via an external panel. Care was taken to ensure that both genders and age groups (18–65 years) were evenly distributed. The panel only considers insurance customers living in Germany or in the German-speaking part of Switzerland. As this is an online panel, it is possible that the participants may have a slightly higher online affinity than the general average of the overall population.

Four equal groups were then formed. Each group was given a scenario to put themselves in.

The following four scenarios were defined in advance:

- iPhone damage caused by yourself and already experienced with online damage reports
- iPhone damage caused by yourself and no experience with online damage reports
- iPhone damage not caused by yourself and already experienced with online damage reports
- iPhone damage not caused by yourself and no experience with online damage reports

According to the IFZ Conversational Insurance Study 2022 (Hundertmark et al., 2022d), reporting claims are one of the most common use cases for a chatbot in the insurance sector.

At the beginning of the survey, the participants were given a detailed description of the assigned scenario. The survey participants were then asked to rate on a 5-point Likert scale how angry or sad they would feel in this situation. This was followed by a question about previous experience with chatbots. Participants could choose whether they had never chatted with a bot, once, twice to five times, or more than five times. The respondents were then asked to independently send a message to a chatbot expressing that they would like to report the damage depending on the scenario.

Afterwards, three short videos were played which showed the conversation between a user and a chatbot, representing conversational AI Systems to report damage.

Chatbot 1 had no emotions or anthropomorphic traits, Chatbot 2 showed few anthropomorphic traits and Chatbot 3 exhibited very pronounced emotions, empathy and anthropomorphic traits. In the following, I refer to Chatbot 1 as “Robot”, Chatbot 2 as “LisaLittle” and Chatbot 3 as “LisaBig”.

The three chatbots were developed based on literature on the use of anthropomorphic design elements. The works of Hundertmark et al. (2022a) served as the basis for this.

To ensure that the level of emotion, empathy and anthropomorphic design elements also fits the descriptions of “not at all”, “medium” and “a lot”, the three developed chatbots were discussed and fine-tuned in a focus group with 15 experts from the field of conversational design and financial services until all experts agreed with the classification.

After each video, participants were asked for feedback on the damage reporting chatbot. The feedback included questions on usefulness, trust in the chatbot and intention to use the chatbot again. Well-known and validated scales were used.

For usefulness, the scale by Davis (1989) was used. For the intention to use the bot again, the scale of Zeithaml et al. (1996) was used. Finally, for the question of trust, the scale of Oghuma et al. (2016) was used.

At the end of the survey, the participants were asked about their personality. The questionnaire was based on the well-known Big Five (Allport et al., 1936), or FFM, by Oliver et al. (2010). An abbreviated test from Rammstedt et al. (2014) was used to define the appropriate personality. The BFI-10 consists of ten items, two for each dimension of personality. Each of the dimensions is measured by one positively and one negatively polarized item. Information on the psychometric quality of the English version is reported by Rammstedt et al. (2014). The test persons must thus only agree or disagree with ten statements on a 5-point Likert scale.

The complete questionnaire for this survey and additional evaluation data is provided in appendix D.

3.5.2. Conversational AI as system to calculate users character: Results

To answer RQ2.1

“Can a conversational AI system derive the complete triad from Character Computing (including personality, affective state and situational behaviour) using features extracted from user messages?”

a linear regression model (Fahrmeir et al., 2021) was estimated for each of the following independent variable:

- Information on whether the message contains a greeting (binary; 38% contain a greeting)
- Information on whether the message contains the word ‘unfortunately’ (binary; 12% contain ‘unfortunately’)
- Information on whether the message contains the word ‘please’ (binary; 14% contain ‘please’)
- Information on whether the message is ‘shorter than average’ (binary; 64% shorter than average)
- age (metric, from 18 to 69; mean 45)
- sex (binary; 50% women)
- experience with chatbots (Likert scale from 1 to 4; mean 2.5)

To determine the independent variables, the text messages that the users wrote to the chatbot during the experiment were analysed in more detail.

First, all messages were coded. There were the following characteristics:

- Message contains a greeting: yes/no
- Message contains the word “unfortunately”: yes/no
- Message contains the word “please”: yes/no
- Message contains the word “help”: yes no
- Message is shorter than average: yes/no

The different codings were determined following the creation of the questionnaire, and discussed and finalized by five experts from the field of conversational AI research. Afterwards, all messages could be coded manually. These codes were then discussed again with the five experts from the field of conversational AI and finalized. The variable “help” was not included in the regression models later because only six messages contained this term, leading to unstable parameter estimates.

In addition, the following dependent variables were used:

- Affective state angry (Likert scale from 1 to 5; mean 3.1)

- Affective state sad (Likert scale from 1 to 5; mean 3.3)
- Big Five dimension extraversion (mean of the 1-to-5 Likert items M5_1 (reversed) and M5_6; mean 3.1)
- Big Five dimension agreeableness (mean of the 1-to-5 Likert items M5_2 and M5_7 (reversed); mean 3.2)
- Big Five dimension conscientiousness (mean of the 1-to-5 Likert items M5_3 (reversed) and M5_8; mean 3.8)
- Big Five dimension neuroticism (mean of the 1-to-5 Likert items M5_4 (reversed) and M5_9; mean 2.7)
- Big Five dimension openness (mean of the 1-to-5 Likert items M5_5 (reversed) and M5_10; mean 3.4)

A possible violation of the normal distribution assumption of the model errors was checked using the Shapiro-Wilk test. If the assumption was violated, the confidence intervals of the regression parameters were estimated using bootstrapping (Efron, 1979) with 1000 draws. For the sake of completeness the p-values estimated by these models are still reported, but should be interpreted with greater caution due to the violation of the underlying assumptions.

Since only 6 people used ‘Help’ in their user prompts, this information is not included as an independent variable in the regression models due to the instability of the results.

In general, the analyses show significant values ($p < 0.05$) and not-significant values, of which some highlights are presented below:

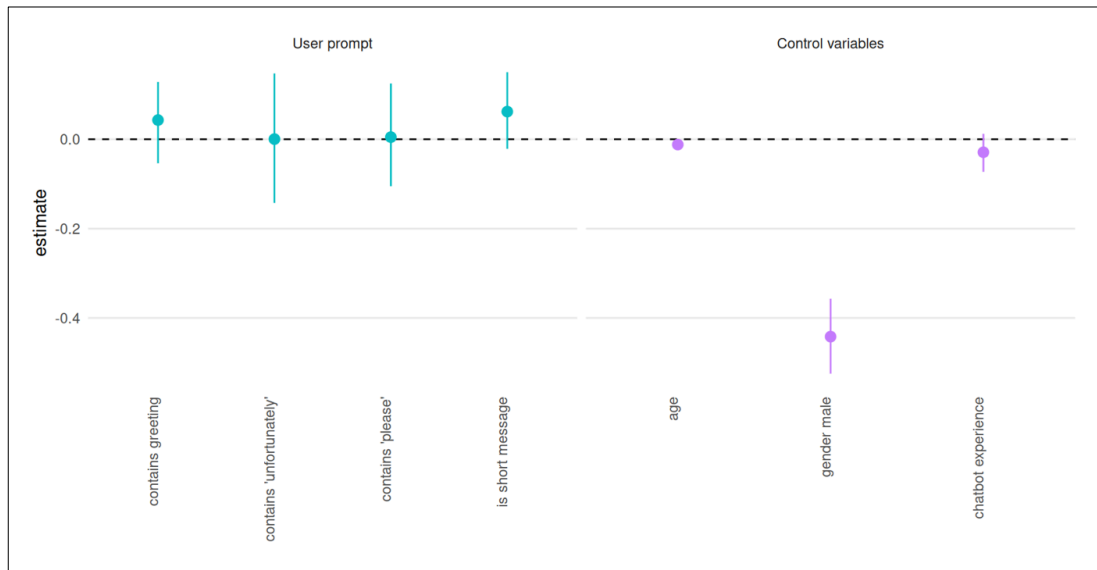
- The mean score of *angry* is, ceteris paribus, 0.08 higher in people whose message contains a greeting than in people whose message does not contain a greeting (95% confidence interval: -0.027 to 0.193).
- With each year of age, the mean *angry* score changes, ceteris paribus, by -0.013 (95% confidence interval: -0.017 to -0.009).
- The mean sad score of people whose message contains a greeting, ceteris paribus, is 0.136 higher than that of people whose message does not contain a greeting (95% confidence interval: 0.011 to 0.25).
- With each year of age, the mean sad score changes, ceteris paribus, by -0.007 (95% confidence interval: -0.011 to -0.003).
- The mean score for *extraversion* in people whose message contains a greeting is, all else being equal, -0.016 lower than in people whose message does not contain a greeting (95% confidence interval: -0.109 to 0.081).

- With each year of age, the mean score of *extraversion* changes, ceteris paribus, by 0.004 (95% confidence interval: 0.001 to 0.007).
- The mean score for *agreeableness* in people whose message contains a greeting is, all other things being equal, -0.014 lower than in people whose message does not contain a greeting (95% confidence interval: -0.097 to 0.068).
- With each year of age, the mean score of *agreeableness* changes, ceteris paribus, by 0 (95% confidence interval: -0.002 to 0.003).
- The mean score of *conscientiousness* is, ceteris paribus, -0.048 lower in persons whose message contains a greeting than in persons whose message does not contain a greeting (95% confidence interval: -0.13 to 0.034).
- For each year of age, the mean score of *conscientiousness* changes, ceteris paribus, by 0.014 (95% confidence interval: 0.011 to 0.016).
- The mean score for *neuroticism* in people whose message contains a greeting is, all other things being equal, 0.043 higher than in people whose message does not contain a greeting (95% confidence interval: -0.054 to 0.128).
- With each year of age, the mean score of *neuroticism* changes, ceteris paribus, by -0.012 (95% confidence interval: -0.015 to -0.009).
- The mean score for *openness* in people whose message contains a greeting, ceteris paribus, changes by 0.052 compared to people whose message does not contain a greeting (95% confidence interval: -0.043 to 0.148).
- With each year of age, the mean score of *openness* changes, ceteris paribus, by 0.003 (95% confidence interval: 0 to 0.006).

The complete results tables can be found in appendix D.

Due to the fact that all R²s are very low (0.9% - 9.6%), I have refrained from listing further details of the individual relationships here. The low R² values indicate that although there are small structural relationships, in general the independent variables can only be described to a very small extent by the dependent variables. With an R² of 9.6%, the correlation is strongest for neuroticism.

Figure 30: Overview: Variable “neuroticism”



Source: Data collected by the author.

In summary, the results exhibit only a few influences. In other words, from these initial findings, it is nearly impossible either for myself, or the conversational AI system, to infer the comprehensive character of the user.

With regard to RQ2.1

“Can a conversational AI system derive the complete triad from Character Computing (including personality, affective state and situational behaviour) using features extracted from user messages?”

the answer is that, to my knowledge with the existing methods of CC, it is not possible to have a text-based conversational AI system that can calculate the needed triad only by using information from user messages.

Research question RQ2.2

“Can a conversational AI system define users’ expectations regarding chatbot interactions by identifying users’ personalities and affective states?”

deals with the nine dependent variables, specifically the following three assessments, which were each queried with regard to each of the three bots:

- Assessment of usefulness
 - (Likert scale from 1 to 5; for bots 1-3: mean values 4, 4, 4; columns M4a_3, M4b_3, M4c_3)
- Assessment of trust
 - (Likert scale from 1 to 5; for bots 1-3: mean values 3.8, 3.9, 3.8; columns M4a_1, M4b_1, M4c_1)

- Assessment of intention to use the bot again
 - (Likert scale from 1 to 5; for bots 1-3: mean values 3.7, 3.8, 3.7; columns M4a_8, M4b_8, M4c_8)

Subsequently, the same statistical procedure as for research question RQ2.1 was applied.

A linear regression model was estimated for each dependent variable with the following independent variables:

- The 7 parameters used as dependent variables in the RQ2.1 models:
 - Affective state angry (Likert scale from 1 to 5; mean 3.1)
 - Affective state sad (Likert scale from 1 to 5; mean 3.3)
 - Big Five dimension extraversion (mean of the 1-to-5 Likert items M5_1 (reversed) and M5_6; mean 3.1)
 - Big Five dimension agreeableness (mean of the 1-to-5 Likert items M5_2 and M5_7 (reversed); mean 3.2)
 - Big Five dimension conscientiousness (mean of the 1-to-5 Likert items M5_3 (reversed) and M5_8; mean 3.8)
 - Big Five dimension neuroticism (mean of the 1-to-5 Likert items M5_4 (reversed) and M5_9; mean 2.7)
 - Big Five dimension openness (mean of the 1-to-5 Likert items M5_5 (reversed) and M5_10; mean 3.4)
- The 3 parameters used as control variables in the RQ2.1 models:
 - Age (metric, from 18 to 69; mean 45)
 - Sex (binary; 50% women)
 - experience with chatbots (Likert scale from 1 to 4; mean 2.5)

Based on the Shapiro-Wilk test, the distribution of model errors for all dependent variables can be assumed to be non-normally distributed (p-value: <0.0001). The confidence intervals of the parameters are estimated accordingly on the basis of a bootstrap with 1000 draws. The p-values estimated by the model are also reported, but should be interpreted with greater caution due to the violation of the underlying assumptions.

In general, the analyses show significant values ($p < 0.05$) and not-significant values, of which some highlights are presented below:

- With each year of age, the usefulness of Bot 1 (Robot) changes, *ceteris paribus*, by 0 (95% confidence interval: -0.004 to 0.003).
- With each year of age, the usefulness of Bot 2 (LisaLittle) changes by 0.001 (95% confidence interval: -0.002 to 0.004), *ceteris paribus*.
- With each year of age, the usefulness of Bot 3 (LisaBig), *ceteris paribus*, changes by -0.003 (95% confidence interval: -0.006 to 0).

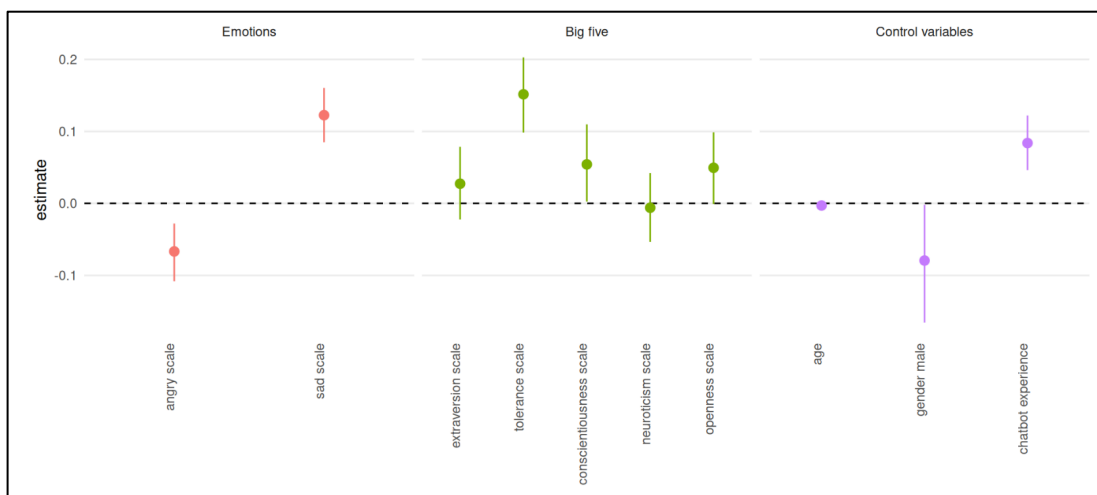
- With each year of age, trust in Bot 1 (Robot) changes, ceteris paribus, by 0 (95% confidence interval: -0.003 to 0.003).
- With each year of age, trust in Bot 2 (LisaLittle), ceteris paribus, changes by -0.002 (95% confidence interval: -0.005 to 0.001).
- With each additional year of age, the intention to use Bot 1 (Robot) again changes, ceteris paribus, by 0.001 (95% confidence interval: -0.003 to 0.004).

The complete results tables can be found in C.

Due to the fact that all R2s are very low, I have refrained from listing further details of the individual relationships here. The low R2 values indicate that although there are small structural relationships, in general the dependent variables can only be described to a very small extent by the dependent variables.

With an R2 of 7.1%, the correlation is strongest for the usefulness of Bot 3 (LisaBig).

Figure 31: Overview Usefulness of Bot 3



Source: Data collected by the author.

Despite some rough insights, the initial analysis does not provide clear and significant conclusions on which user characteristics consistently predict a bot's perceived usefulness, trustworthiness or overall value, nor on how a conversational AI system might find out these user expectations.

Summarizing the results for RQ2.2, it is clear that only a few findings can be extracted that may lead to the users' expectations. However, these findings are not sufficient for a conversational AI system to determine the whole character of a user and its expectations when using only text-messages.

With regard to 2.2,

“Can a conversational AI system define users’ expectations regarding chatbot interactions by identifying users’ personalities and affective states?”

the answer is that, to my knowledge with the existing methods of CC, it is not possible to have a text-based conversational AI system that can calculate the users’ character only by using information from user messages and furthermore draw conclusions on the users’ expectations.

3.5.3. Requirements for Conversational AIs as a System to Calculate User Character – Implications for Future Conversational AI

Combining the results from RQ 2.1 and 2.2, I come to the following main finding:

Traditional text-based conversational AI systems are not capable of accurately inferring specific traits related to the user's character, personality, or affective state during a simple dialogue between user and bot. However, such inferences would form the essential basis for a conversational AI system to understand user expectations within the chat interaction and subsequently adapt accordingly.

Originally, the method of CC (El Bolock et al., 2020) considers an extensive range of sensors and user characteristics to derive a user’s expectations of a conversational AI system and to be able to customize the design of the system. Since in real life users are not normally connected to specific sensors, I conducted my study without any sensors. Instead, only specific characteristics of messages were considered, to answer RQ2 in general:

How effectively can methods from Character Computing be integrated into contemporary Conversational AI systems to interpret the complete triad of user characteristics (personality, affective state and behaviour), and what existing gaps need to be addressed to enable a chatbot to adapt its interactions accordingly?

This experiment and the answers to RQ2.1–2.2 show that the characteristics from message data are neither sufficient to determine significant user expectations of a conversational AI system, nor draw significant conclusions about a user’s character or personality.

In other words, the results underline the premise that CC methodologies and triad determination alone are probably not sufficient to optimally adapt a conversational AI system to the user and its context. To address this, I derive the following new capabilities for my framework:

Table 19: Design Requirements 13,14

Number	Capability	Explanation	Source
13	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Use Case 4
14	Integration of Fuzzy Value Computations	Enables the interpretation and representation of ambiguous or imprecise user characteristics using fuzzy logic. Supports linguistic variables for characterizing traits and expectations.	Use Case 4

Source: Data collected by the author.

The findings indicate that existing CC methods must be augmented with other approaches to account for the absence of physical sensors in most real-world conversational AI applications.

3.6. Findings on Requirements for the Design of Skills for Future Conversational AI

The four use cases, taken together, answer RQ1 and RQ2 and generate the complete set of design requirements for the FCCC artefact. This section summarizes what each use case contributes to the overall requirement structure, consolidates the 14 capabilities into a coherent design rationale, and identifies the key implication for the artefact that follows in Chapter 4.

Use case 1 (knowledge transfer, chapter 3.1) establishes that conversational AI systems in learning contexts must adapt to individual learning pace and comprehension level, provide transparency about their nature and the limits of their advice, and integrate ethical and privacy safeguards as design requirements rather than implementation choices. These findings directly motivate the ethical outer layer of the FCCC Framework and ground RQ1.1.

Use case 2 (customer service, chapter 3.2) demonstrates that user satisfaction with a chatbot depends critically on the match between dialogue emotional design and the user's situational context and mood. It introduces the value irritant matrix as an economic decision tool for selecting appropriate automation scenarios. These findings establish that the FCCC Framework must treat situational context as a primary input and must balance user-centric and organization-centric objectives. They ground RQ1.2 and motivate the economic outer layer.

Use case 3 (advisory, chapter 3.3) shows that stable user characteristics – age, gender, Big Five personality, chatbot experience – produce systematic and predictable differences in preferred chatbot behaviour across empathy, response length, and

contextual reference. This finding confirms that adapting to situational context alone is insufficient: the FCCC Framework must also model stable personality and demographic characteristics. It grounds RQ1.3 and provides the empirical basis for the Character Computing component of the artefact.

Use case 4 (character estimation, chapter 3.5) tests whether the CC triad can be estimated from text messages alone. The low R^2 values across all regression models (0.9%–9.6%) demonstrate that conventional CC methods are insufficient for text-only conversational AI. This is the most consequential finding for the artefact design: it means that the FCCC Framework cannot simply implement CC as proposed by El Bolock et al. (2020). It must supplement CC with fuzzy logic as a sensor replacement and with conversational loops as a mechanism for accumulating character estimates over multiple turns. It grounds RQ2.1 and RQ2.2 and defines the specific technical challenge that chapter 5 addresses.

Finally, I obtain a list of 14 capabilities resulting from the different use cases (table 20).

Table 20: Design requirements first overview, draft

Number	Capability	Explanation	Source
1	Audience-Adapted Communication	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Use Case 1
2	Flexible Interaction Modes	Combine structured interaction (e.g. pre-defined buttons) with opportunities for user-driven input to enhance engagement and adapt to users' learning preferences.	Use Case 1
3	Adaptive Content Design	Enable the integration of knowledge documents to ensure that learning content remains customizable and aligned with specific needs, allowing organizations to maintain control over the material.	Use Case 1
4	Economic Accessibility	Maintain a cost structure that balances system sustainability with affordability for a broad user base.	Use Case 1
5	Multimedia Integration	Incorporate diverse media elements such as images, videos and graphics to enhance interactivity and accommodate various learning styles.	Use Case 1
6	Privacy and Trust	Ensure user anonymity and data protection to build trust and encourage honest participation.	Use Case 1

Number	Capability	Explanation	Source
7	User Context Awareness	Conversational AI systems should dynamically adapt their interactions based on the user's situational context, including emotional state, current needs and goals, to provide relevant and meaningful responses.	Use Case 2
8	Personalization	The systems should consider the user's past interactions, preferences and personality to create tailored and engaging dialogue flows that improve the user experience.	Use Case 2
9	Alignment with Organizational Objectives	All design elements, including dialogue styles and automation levels, should balance the user's needs for efficiency and satisfaction with the organization's goals of cost control, revenue optimization and operational efficiency.	Use Case 2
10	Contextual Dialogue Adaptability	Conversational AI systems should dynamically adjust response characteristics such as length, empathy and detail level based on situational factors and user profiles. This ensures alignment with individual user expectations and optimizes user satisfaction.	Use Case 3
11	User Profile Differentiation	The system must be capable of recognizing and categorizing users based on traits such as personality, age, gender and prior chatbot experience. This enables the delivery of personalized interactions tailored to diverse user needs.	Use Case 3
12	Scalable Personalization Strategies	To balance personalization with operational efficiency, companies should implement scalable mechanisms for tailoring chatbot dialogues.	Use Case 3
13	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Use Case 4
14	Integration of Fuzzy Value Computations	Enables the interpretation and representation of ambiguous or imprecise user characteristics using fuzzy logic. Supports linguistic variables for characterizing traits and expectations.	Use Case 4

Source: Data collected by the author.

However, since the provided table contains overlapping and similar capabilities, I merged some capabilities and eliminated redundancies while ensuring clarity and

completeness. Below are the merged capabilities, along with explanations and rationale for their inclusion:

Table 21: List of final Design Requirements

Number	Capability	Explanation	Source
1	Audience-Adapted and Context-Aware Communication	Use accessible language tailored to the audience while dynamically adapting to user context, including emotional states and situational needs.	Combines 1, 7 and 10
2	Dynamic Personalization	Create interactions based on user history, preferences and profiles, enabling personalized experiences while ensuring scalability.	Merges 8, 11 and 12
3	Flexible Interaction and Content Integration	Incorporate both structured and open-ended modes of interaction, integrating multimedia elements and customizable content to cater to varied preferences.	Combines 2, 3 and 5
4	Trust, Privacy and Economic Accessibility	Prioritize data protection, user anonymity and affordable implementation to foster trust and accessibility.	Derived from 4 and 6
5	Alignment with Organizational Objectives	Design systems to balance user satisfaction with organizational goals, ensuring operational and economic efficiency.	Maintains integrity of 9
6	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Derived from 13
7	Enhanced Decision-Making Using Fuzzy Logic	Incorporate fuzzy computing techniques to interpret ambiguous user characteristics, enabling nuanced responses.	Derived from 14

Source: Data collected by the author.

The most important single finding from the four use cases is the one from use case 4. Text-based conversational AI systems cannot, with existing CC methods, reliably estimate the user's character triad from message features alone. The artefact developed in chapter 4 responds to this limitation in three ways. By replacing sensor-based measurement with multidimensional fuzzy classification over multiple conversational turns; by using conversational loops (rather than a single-shot analysis) to accumulate a progressively refined character estimate and by designing the system to remain functional and to degrade gracefully when its character estimate is uncertain. The specific technical implementation of this response is the subject of chapter 5.

In doing so, the FCCC Framework does not abandon Character Computing. It *extends* it. Fuzzy logic provides the mathematical bridge that CC researchers themselves identified as a necessary next step (El Bolock et al., 2020), and conversational loops provide the dynamic feedback mechanism that makes incremental estimation possible. The four use cases in this chapter have generated both the motivation and the empirical constraints for this extension. Chapter 4 now builds the artefact.

4. New Conceptual Framework for Future Conversational AI (FCCC Framework)

The primary goal of this thesis is to design a framework that combines the synergies between the disciplines of economics, psychology and computer science to enable the development of naturally acting, sustainable, ethically and economically successful conversational AI systems. These function as self-regulating systems, as known from systems theory (von Bertalanffy, 1949), compute with fuzzy values and can be implemented by companies in the long term while complying with ethical, cultural, legal and economic frameworks.

To answer RQ3 and to develop the artefact, my conceptual framework, I continue to follow the DSR method, explained in chapter 1. The DSR method is particularly widespread in the field of business informatics. In DSR, existing research or results are not simply revised, but DSR enables researchers to create completely new artefacts or models. This involves an iterative process in which researchers repeatedly integrate feedback from selected experts from the fields of research and practice into the optimization rounds of the artefact. In other words, the approach of DSR emphasizes my transdisciplinary and interdisciplinary overall research approach.

According to the DSR process, the study has now reached its third phase (Möller et al., 2020).

4.1. Capabilities for Future Conversational AIs – For the Artefact

The objective while developing the artefact during this chapter, is to answer RQ3:

How can sustainable and naturally intelligent conversational AI systems that function as self-regulating systems, as known from systems theory, and that compute with fuzzy values be implemented by companies in the long term while complying with ethical, cultural, legal and economic frameworks?

In answering RQ3, I aim to develop a completely new conceptual framework for researchers and practitioners, which makes the design of conversational AI systems successful, adaptable and sustainable in the long term.

In chapter 2, the theoretical foundations for the development of my conceptual framework were developed. I explained the fundamental ideas on systems theory and showed how this theory belongs to conversational loops and fuzzy methods. Additionally, I presented ideas on how existing methods from CC might be enhanced with fuzzy methods and conversational loops to obtain more precise results.

Chapter 3 addressed RQ1 and RQ2 by examining the capabilities that the artefact must satisfy according to the principles of DSR. Through four independent experiments (use case 1-4), I identified seven key requirements that serve as the

foundation for the development and evaluation of the artefact, specifically the conceptual framework (table 21).

Table 21: List of final Design Requirements

Number	Capability	Explanation	Source
1	Audience-Adapted and Context-Aware Communication	Use accessible language tailored to the audience while dynamically adapting to user context, including emotional states and situational needs.	Combines 1, 7 and 10
2	Dynamic Personalization	Create interactions based on user history, preferences and profiles, enabling personalized experiences while ensuring scalability.	Merges 8, 11 and 12
3	Flexible Interaction and Content Integration	Incorporate both structured and open-ended modes of interaction, integrating multimedia elements and customizable content to cater to varied preferences.	Combines 2, 3 and 5
4	Trust, Privacy and Economic Accessibility	Prioritize data protection, user anonymity and affordable implementation to foster trust and accessibility.	Derived from 4 and 6
5	Alignment with Organizational Objectives	Design systems to balance user satisfaction with organizational goals, ensuring operational and economic efficiency.	Maintains integrity of 9
6	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Derived from 13
7	Enhanced Decision-Making Using Fuzzy Logic	Incorporate fuzzy computing techniques to interpret ambiguous user characteristics, enabling nuanced responses.	Derived from 14

Source: Data collected by the author.

4.1.1. First Version of the Artefact

As the foundation for the initial version of the artefact, I employ the defined capabilities outlined in chapter 3, aligned with my previously introduced DSR approach and combine them with my own, new definition of FCCC, presented in chapter 2.

Furthermore, I have previously published parts of the following work: Hundertmark, S., Portmann, E., & Hafner, N. (2023). Fuzzy Conversational Character Computing. Informatik Spektrum 46(3), 163–175. <https://doi.org/10.1007/s00287-023-01538-8>

Finally, I keep in mind the ideas of systems theory (von Bertalanffy, 1949) to develop a system that utilizes feedback loops to adapt its interconnected subsystems and, ultimately, the entire system to the environment. To complement the artefact, informal discussions with researchers in the fields of conversational AI and computer science took place during the creation of the first artefact. These conversations followed the approach from Pangaro's (2019) Turing test for conversations.

The core of the artefact is in conversational loops. Through repeated conversational loops, a conversational AI system can be continuously improved to optimize pitch and content in relation to a positive user experience. This continuous improvement is already a subject in the field of systems theory (von Bertalanffy, 1949), as explained in the first part of the second chapter. The optimization of conversational AI systems is based on the character of the user, which is composed of personality traits, the current state of the user and the situation the user is in during the conversation with conversational AI. According to CC researchers, it should be possible to calculate the expected behaviour of the user based on the user character and the situation (El Bolock et al., 2020).

However, in chapter 3 (use case 4), I pointed out that the conventional methods of CC without the use of sensors are not sufficient to adequately consider the user's behaviour and, hence, also their expectations of a conversational AI system. While CC researchers use sensors to determine and measure the character (El Bolock et al., 2020), I recommend the use of fuzzy logic (Zadeh, 1965), especially multidimensional fuzzy classifications (Zumstein, 2026), as well as the application of conversational theory (Pask, 1965; Pangaro, 2010) and conversational loops. The application of conversational theory, in particular the repetition of conversational loops, is a central feature of the artefact. The overall integration of fuzzy logic into the methods of CC has also been suggested by researchers from the field of CC but has not been realized so far (El Bolock et al., 2020).

According to conversational theory, several conversational loops are necessary to reach agreement between the two parties, in this case the user and the conversational AI system. Only through dialogue can the conversational parties get to know each other better, learn from each other, build trust and eventually reach agreement (Pangaro, 2010). Agreement within this framework means that the conversational AI adapts its behaviour according to the reactions expected by the user in order to ensure an optimal user experience.

To support this process, the artefact uses multidimensional fuzzy classifications (Zumstein, 2026) instead of conventional fuzzy logic (Zadeh, 1965). The character of a user does not consist of one sharp value, but of a combination of several fuzzy values. In the second chapter I derived and explained the use of multidimensional fuzzy classifications in the context of CC (El Bolock et al., 2020).

In sum, here I apply my new theory to FCCC. The following explains concretely how multidimensional fuzzy classification and CWW function as the computational core of the FCCC artefact's inner layer.

Step 1 – Fuzzification of the user message

When the user sends a message, the system does not assign it a single category. Instead, it computes a fuzzy character vector $C(m) = (\mu_1(m), \mu_2(m), \dots, \mu_d(m))$ where each μ_i represents the degree to which the message belongs to a personality or affective category on dimension i . For example, a message might yield $C(m) = (0.8, 0.3, 0.5)$, meaning the user is assessed as 80% open, 30% stressed, and 50% extroverted simultaneously. This is the core departure from conventional binary classification. The user is never forced into a single character box.

Step 2 – Multi-dimensional combination via the γ -operator

Multiple messages accumulate across the conversation. To combine them into a stable character estimate, the artefact applies Zumstein's (2007; 2026) compensatory aggregation operator:

$$\mu_{A,comp}(x) = \left[\prod_i \mu_i(x) \right]^{(1-\gamma)} \cdot \left[1 - \prod_i (1 - \mu_i(x)) \right]^\gamma, \quad \gamma \in [0, 1]$$

The parameter γ controls the degree of compensation between dimensions. At $\gamma = 0$, the formula is a strict conjunction. All dimensions must indicate the trait for the combined membership to be high. At $\gamma = 1$, a high value on any single dimension can fully compensate for low values on others. In practice, a mid-range γ (typically 0.5) is used, reflecting the empirical finding from personality psychology that traits are partially independent but not orthogonal (McCrae et al., 1992). This flexibility makes the artefact more robust to sparse early-conversation data than a pure conjunction would be.

Step 3 – Computing with Words for response selection

Once the aggregated character vector is available, the artefact must translate it into a response strategy. Which tone, length, empathy level, and degree of contextual reference should the next message adopt. This is where CWW (Zadeh, 1999) enters. Rather than mapping numeric membership values to crisp response parameters, the system uses linguistic inference rules of the form:

IF openness IS high AND affective_state IS
stressed THEN tone IS empathetic AND length IS short

These rules are defined in natural language, not numerical code, which makes them auditable, adjustable by non-technical domain experts, and directly interpretable against the ethical outer layer's requirements for transparency. The output of the

inference step is again a fuzzy set, which is defuzzified, converted back to a crisp response parameter – using the centroid method.

Step 4 – Updating the situation vector across turns

After the response is sent, the user's reaction becomes the next input. The situation vector S accumulates all prior messages and their associated character vectors, giving the system a longitudinal view of how the user's state evolves across the conversation. This accumulation is why multiple conversational loops are necessary. A single message provides limited character signal (as use case 4 demonstrated empirically), but three to five turns are typically sufficient to produce a stable fuzzy estimate.

Together, these four steps constitute the yellow inner layer visible in Figure 32. Table 22 summarizes the mapping between fuzzy components and framework layers.

Table 22: Mapping of fuzzy components to FCCC artefact layers

Fuzzy component	Location in FCCC artefact	Function	Key reference
Membership function $\mu(x)$	Inner layer – character estimation	Maps each user message dimension to a partial membership degree $\in [0,1]$ for each character trait.	Zadeh (1965) Zumstein (2007; 2026)
Multi-dimensional classification (γ -operator)	Inner layer – character aggregation	Combines membership degrees across multiple dimensions and multiple messages into a stable character vector C .	Zumstein (2007; 2026) Werro (2015; 2026)
Computing with Words (CWW) linguistic inference rules	Inner layer – response strategy selection	Translates fuzzy character vectors into response strategy parameters (tone, length, empathy, reference) using auditable natural-language rules.	Zadeh (1999)
Situation vector S (accumulated C across turns)	Inner layer – longitudinal context	Stores the full conversational history including character vectors, enabling the system to track how user state evolves across turns.	Pangaro (2010) El Bolock et al. (2020)

Source: Data collected by the author.

In the lower part of the artefact (Figure 32), which is marked in yellow, all calculations are based on fuzzy calculations. No sharp values are used, rather, fuzzy classes are formed.

Although it is not the focus of this thesis, but Large Language Models (LLMs) are nevertheless relevant to the entirety of the framework. An LLM is a language

model based on a neural network with billions of parameters and trained on large amounts of unlabeled data using self-supervised learning. It is used to predict and generate text, including other content (Sejnowski, 2023). In particular, OpenAI's known chatbot, ChatGPT, has gained in importance (OpenAI, 2022).

LLMs have the disadvantage of high computing and energy consumption. Therefore, they are only used very selectively and for specific tasks in this framework. Developers of chatbot systems use LLMs in particular to process knowledge databases and to successfully match user requests with the correct system responses. This use appears to be very relevant in the present framework. Once a conversational AI has achieved the appropriate pitch and design to maximize the user experience, as described above, it is a matter of delivering the appropriate information to the user. To do this, the conversational AI system has access to a variety of knowledge sources, such as websites, databases, PDFs, images, videos and other media. Through the use of LLMs, the conversational AI system is able to identify the user's request and retrieve the appropriate answer from the knowledge base in the shortest possible time. Here again, knowledge sources may use fuzzy terms.

Overall, the artefact sees the combination of conversational loops, CC, multidimensional fuzzy classifications, knowledge sources and LLMs as a self-improving system, as it is known from the systems theory (von Bertalanffy, 1949). In the lower part of the framework, self-improvement already occurs through conversational loops. However, it is also assumed that, in the case of inaccuracies or incorrect correspondences between the user request and the content response of the conversational AI system, these errors are also recognized by the conversational loop and reported back to the system. In this way, the conversational AI receives feedback not only on the pitch and design, but also on the accuracy of the content and the knowledge databases.

In addition to this self-improving system, oriented to systems theory (von Bertalanffy, 1949), certain external influencing factors are relevant that affect the design and responses of conversational AI systems independently of user character or knowledge bases.

For this purpose, I integrate three additional influence layers into the artefact. The inner outer layer includes economic considerations for the use of conversational AI systems. Company-specific policies and decisions play an important role, including the company's "Tone of Voice". In addition, analyses within the framework of the value irritant matrix (Hafner et al., 2021) are important to determine the appropriate areas of application for conversational AI systems in accordance with the needs of the users and the company. There are no generally valid rules for the use of automated dialogues via conversational AI systems, but the value irritant matrix serves as a helpful tool for individual applications in each company (Hafner et al., 2021).

The middle outer layer includes legal and regulatory frameworks that may vary depending on the industry or country in which the conversational AI system is

deployed. Moreover, these rules are expected to evolve in the coming years. The European Union is currently introducing a new law for the use of AI within the framework of the AI Act (Regulation (EU) 2024/1689, 2024), which could also have an influence on the present framework and the design of the middle outer layer.

The outermost layer comprises ethical framework conditions for the use of conversational AI systems. This layer is deliberately placed in the outer area of the framework and thus influences all further decisions and impacts. It is assumed that ethical considerations will play an increasingly important role and that the use of conversational AI systems will be influenced by a variety of ethical aspects. Transparency in particular plays a major role here. Additionally, I assume that these ethical values may also consist of fuzzy values or terms, which can be calculated from my conversational AI system thanks to the capability of computing with words and perceptions.

To sum up in table 23 I list all defined capabilities and link them to an element of the new framework.

The framework ensures sustainability by integrating feedback loops from conversational theory, allowing AI systems to continuously learn and adapt through interaction rather than relying on static models. Ethical and economic considerations are embedded in the system design, ensuring compliance with evolving regulatory and societal expectations (DR 1). The self-regulating nature of the proposed conversational AI is directly inspired by von Bertalanffy's systems theory (1949). The framework ensures AI adaptability by incorporating CC, which allows the system to assess user personality, emotional states and situational factors dynamically (DR 2, 6). Additionally I apply fuzzy logic (Zadeh, 1965) and multidimensional fuzzy classifications (Zumstein, 2026) to enhance decision-making in ambiguous or complex scenarios (DR 3, 7). Thanks to the three outer layers (ethical, legal and economic) the framework ensures transparency, user consent and fairness in AI decisions, accounts for regulatory requirements and aligns with company-specific objectives, ensuring cost-effectiveness and economic feasibility. (DR 4, 5)

Table 23: Capabilities for Artefact

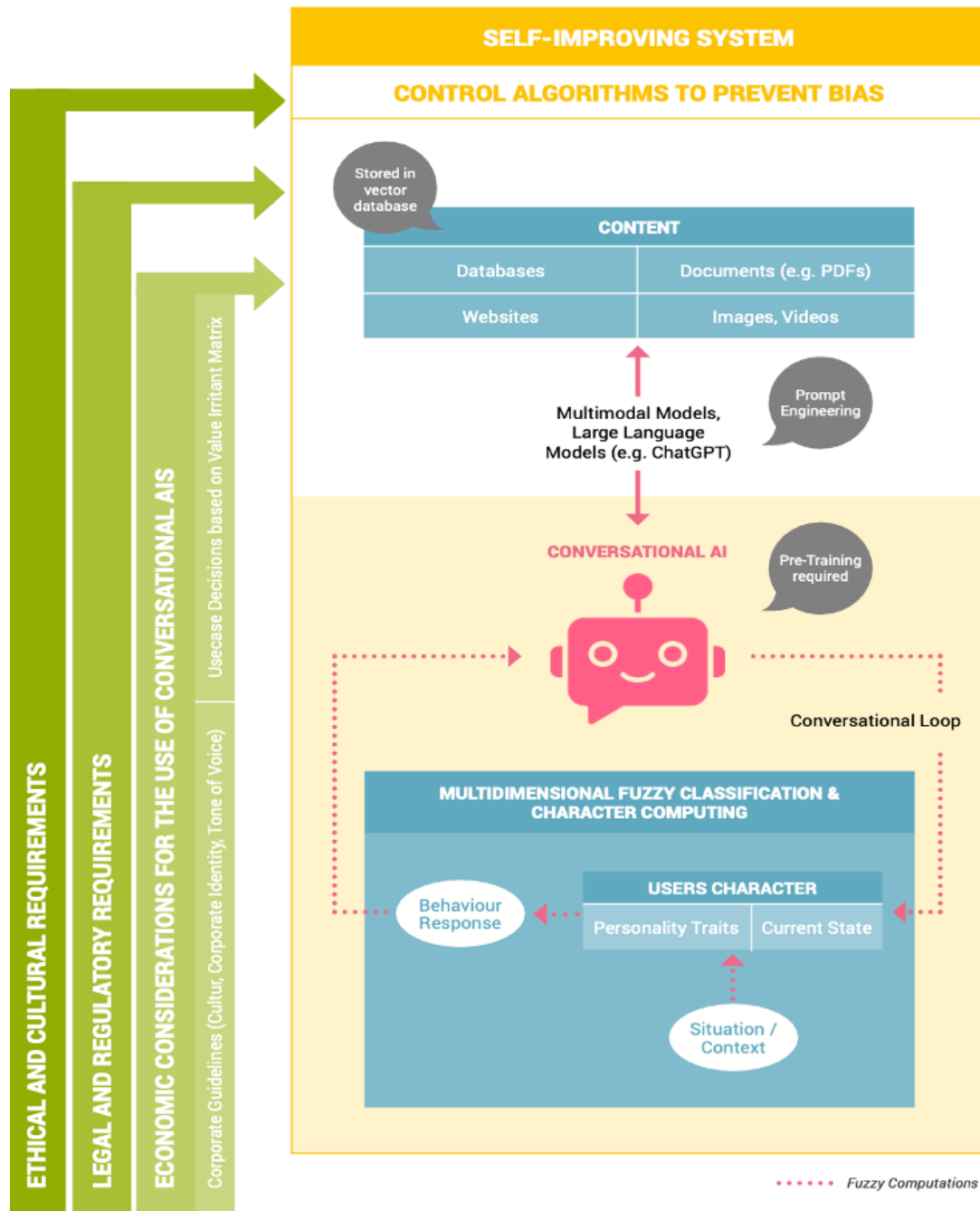
Design Requirement (DR)	Capability	Explanation	Framework Element
1	Audience-Adapted and Context-Aware Communication	Use accessible language tailored to the audience while dynamically adapting to user context, including emotional states and situational needs.	Conversational loops informed by fuzzy classifications and CC to adapt based on emotional and situational contexts.
2	Dynamic Personalization	Create interactions based on user history, preferences and profiles, enabling	Multidimensional fuzzy classifications and the integration of CC and LLMs

Design Requirement (DR)	Capability	Explanation	Framework Element
		personalized experiences while ensuring scalability.	for personalization based on contextual user data.
3	Flexible Interaction and Content Integration	Incorporate both structured and open-ended modes of interaction, integrating multimedia elements and customizable content to cater to varied preferences.	Conversational loops and LLM-enhanced content delivery modules for multimodal integration and flexible responses.
4	Trust, Privacy and Economic Accessibility	Prioritize data protection, user anonymity and affordable implementation to foster trust and accessibility.	Ethical, regulatory and economic layers in the framework, supported by fuzzy logic to ensure transparency, data minimization and economic accessibility for scalability.
5	Alignment with Organizational Objectives	Design systems to balance user satisfaction with organizational goals, ensuring operational and economic efficiency.	Economic consideration layer using the value irritant matrix to align chatbot features with organizational and user needs.
6	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Integration of conversational theory with fuzzy logic to simplify complex interactions and improve user alignment without compromising technical robustness.
7	Enhanced Decision-Making Using Fuzzy Logic	Incorporate fuzzy computing techniques to interpret ambiguous user characteristics, enabling nuanced responses.	The foundational framework combining fuzzy classifications with CC to handle ambiguity in user characteristics effectively.

Source: Data collected by the author.

Figure 32 depicts the first version of the artefact, taking into account the above-described components:

Figure 32: First version of the Artefact



Source: Data collected by the author.

4.1.2. Example Application of the First Version of the Artefact

To make the previously rather theoretically derived and explained artefact more tangible, I present two exemplary applications of the artefact. At the same time, I demonstrate the interaction of the different components and their respective weights. What follows are two examples of how my conceptual framework can be used for an advisory process for a bank client and for the claims reporting process for an insurance client.

Application of the Artefact in Advice on Bank Accounts

Initial situation

A typical dialogue between bank and customer is the consultation on simple account products. According to the IFZ Conversational Banking Study 2021 (Hundertmark et al., 2021), this type of dialogue lends itself extremely well to automation with a chatbot. Customers usually have a large number of general and repetitive questions, and the content of the answer can be obtained in a structured manner from a knowledge database. At the same time, the right tone of voice plays an important role, especially with new customers. If it is a new customer relationship, building trust is a particularly important aspect. Trust is usually developed in the exchange of conversations and increases or decreases with the number of messages exchanged.

The use of the framework developed here makes sense for this application for several reasons and is described in more detail below.

Outer layer

The dialogue between the conversational AI and the user normally starts with a greeting from the conversational AI. Here, the outermost layer with the ethical and cultural framework has a direct effect. Depending on the culture and ethical recommendations, the introduction of conversational AI varies. For this use case, I assume that it is a consultation in the Swiss area, where the general cultural and ethical conditions of the Swiss apply. I also assume that, especially with regard to ethics, transparency plays an important role. Consequently, the introduction of conversational AI must be transparent and correct, especially concerning data processing and the way conversational AI generates its answers regarding account advice.

In this example, a possible appropriate greeting could be as follows:

Chatbot:

“Hello, you are writing with the chat robot of the Swiss Bank. My job is to help customers find the right account product. In doing so, I always pursue the goal of convincing you positively about one of our bank products. I only draw my knowledge about the account products from the internal knowledge databases of the Swiss Bank. I do not take into account any products of other market participants.

I would like to point out that the dialogue content will be stored for 30 days and then deleted. The dialogue content will not be used for further training purposes. If you agree to this, then why not write to me about the type of account you are interested in?”

The example dialogue also shows the influences of the second and third external layers. On the one hand, the chatbot must comply with the legal provisions of Switzerland and, on the other, it must also consider the economic considerations of the bank. The legal basis includes the way in which data is stored and used. The economic

considerations include the fact that the chatbot only provides information on the bank's own products.

Inner layer

This is followed by consideration of the inner layers, which are the focus of the entire framework. The conversational AI has introduced itself sufficiently and informed the user about his rights and the background of the chatbot. What follows below is an initial response from the user, which could be:

User:

"I am looking for a euro account."

The conversational loop starts with the user expresses his need. Now, with the help of CC, conversational AI aims to determine the character and the expected or desired behaviour of the user, and to react appropriately. Furthermore, the conversational loops are intended to build trust between the user and the bank so that the user can ultimately decide in favour of one of the bank's products with a clear conscience.

The user's first message merely indicates what kind of account he wants. However, there are still few indications of the user's character. Only the fact that the user has formulated his answer very briefly and succinctly could indicate that he is introverted and possibly has little time for the entire consultation.

The conversational AI must now respond to the user's first message and try to find out more about the user's character and needs. The pre-training that was given to the conversational AI from the beginning has an effect here. A possible part of the pre-training could be that it should first politely confirm the user's request and then ask more in-depth questions.

The following answer would be conceivable:

Chatbot:

"You are looking for a euro account. May I ask if you are already a customer with us?"

The user answers this question in a much more enthusiastic way and gives much more information about his concern and his character in the second dialogue step.

User:

"Yes, I have been a satisfied customer for 10 years, so I would definitely like to open a euro account with you."

Now, according to conversational theory (Pangaro, 2010), further dialogues follow in which the conversational AI finds out more and more about the user. It should be noted that conversational AI always works with fuzzy values or fuzzy terms, and that the result for the user's character is merely a combination of fuzzy values or terms,

and an assignment to various character classes. The user's character is therefore formed by various influencing factors.

Knowledge bases and Large Language Models

Once the conversational AI has determined all the necessary information about the user and his needs, the query in the knowledge database follows. This is where the LLMs are used for the first time. With the help of the LLMs, the conversational AI will forward the user's need for an account product to the knowledge database and will then suggest a suitable account product based on the bank's internal product information. The conversational AI system will then suggest this account product to the user. However, before the final answer can be given, both the external layers and the insights from the conversational loops must be integrated. The outer layers include, above all, ethical guidelines, for example, with regard to transparency.

From the conversational loops, the conversational AI system has gained important information about the character of the user and the desired tonality. However, this tonality must also correspond to the framework conditions of the external layers. Let's assume that, in this case, the user communicates with the German "Du", but the economic and internal communication rules of the company require polite "Sie" communication. Consequently, conversational AI must find the right balance here, and the framework conditions of the outer layer must not be violated.

A possible answer or account recommendation from conversational AI could be as follows:

Chatbot:

"Based on the product information of our Swiss Bank, I have found the Basic Euro Account for you. The account costs CHF 10 per month. I would like to point out that the account card has an additional annual fee of CHF 50. If you decide to conclude the contract now, other data protection guidelines apply. In this case, I would save your personal data and forward it to the account manager of the Swiss Bank. Would you like to continue the process?"

Application of the Artefact in the Reporting Process of an Insurance Company

Initial situation

The IFZ Conversational Insurance Study 2022 (Hundertmark et al., 2022d) worked out the process of reporting a claim as a typical dialogue for a chatbot. Researchers come to a similar conclusion when they apply the value irritant matrix in the context of insurers and policyholders. Consequently, this application is also suitable for the application of the presented framework. In the case of claims reporting, building trust between the user and conversational AI may play a lesser role. However,

it seems to be more important that the legal framework is adhered to with regard to the reporting of claims and that the customer experience during the reporting process does not lead to unpleasant feelings on the part of the user. The user is probably already in a rather negative frame of mind due to the general situation of “damage suffered”.

Outer layer

The dialogue between the conversational AI and the user usually starts with a greeting from the conversational AI. Here, the outermost layer with the ethical and cultural framework has a direct effect. Depending on the culture and ethical recommendations, the introduction of conversational AI varies. For this use case, I assume that we are dealing with an insurance company in Germany, where the general cultural and ethical conditions of the German population apply. I assume that, especially as concerns ethics, transparency plays an important role. Consequently, the presentation by conversational AI must be transparent and correct. Furthermore, personal and sensitive data of the user is processed when a claim is reported. Thus, the second outer layer regarding the legal conditions plays a particularly important role here. The third external layer regarding the framework conditions provided by the company, in this case by the insurance company, should also not be neglected.

A greeting on the part of conversational AI could be now read as follows:

Chatbot:

“Good day. I am the digital assistant at Deutsche Versicherung. You can submit a damage report to me. I would like to point out at this point that any data you disclose during the chat conversation may be used further in the claims process. This chat uses OpenAI’s Large Language Models. However, it is ensured that data worth protecting is sent to OpenAI anonymously. This is only done if it is indispensable for the further course of our conversation.”

Inner layer

Following the greeting, the conversational AI learns of the user’s basic request, and the conversational loop starts. As mentioned above, the design of the conversational AI is probably less central to the process of reporting a claim, but it is even more important that the claim process is completed quickly. This is where the pre-training of the conversational AI comes into play, which has trained the AI to keep the conversational loops as short and concise as possible in the event of a claim. Furthermore, the exact process of the claim notification could be stored in the pre-training so that, in this case, the conversational AI can dispense with the use of the internal knowledge databases by means of LLMs and guide the user into a pre-trained process within one of the first conversational loops.

Assuming that the user expresses his concern as follows:

User:

“I dropped my iPhone and broke it. I want to report the damage and go through the process as soon as possible.”

Could conversational AI answer the following:

Chatbot:

“I am sorry that your iPhone is damaged. My job is to guide you through the claims process as quickly as possible. When did the incident happen?”

Remarks on the Applied Artefact in a nutshell

The scenarios described above illustrate the use of the artefact in different situations and use cases.

What both examples have in common is that the external layers concerning the ethical, legal and economic framework conditions must be taken into account from the very start. Furthermore, conversational AI only works in both cases if pre-training has taken place.

With regard to the use of conversational loops, the examples differ. In the context of counselling, significantly more conversational loops seem to be advantageous, whereas in the context of claims reporting, significantly fewer loops are desirable.

Overall, both examples show that the framework can be used in practice and that the interaction of the individual components can also be used for different use cases.

4.2. Methods to Improve the Artefact

The formal evaluation according to the DSR framework will be conducted in chapter 5 using three independent measurement methods. As part of my extension to the classical DSR approach, incorporating the Turing test for conversations (Pangaro, 2019), I conduct an initial feedback round at this stage by consulting selected experts for their input on the artefact. This feedback informs the development of the second version of the artefact, which will subsequently undergo evaluation in the traditional DSR evaluation process.

4.2.1. Expert selection

To ensure that the feedback collected on the first version of the artefact was both methodologically rigorous and practically grounded, a diverse group of experts was built. Fifteen experts were selected on the basis of three criteria: direct relevance of their professional domain to at least one component of the FCCC framework, a mix of academic and practitioner backgrounds to capture both theoretical knowledge and implementation experience and sufficient seniority to provide substantive critical

assessment rather than merely confirmatory feedback. The resulting expert overview is shown in Table 24.

Table 24: Overview Experts for first artefact

Type	Domain / Role	Institution / Sector	Relevance to FCCC
Academic	Professor, Informatics	University	Conversational AI system design, enterprise chatbot deployment
Academic	Associate Professor, Psychology	University	Big Five personality modelling, psychometric measurement, affective computing
Academic	Professor, Computer Science / AI	University	Large Language Models, NLP, fuzzy classification
Academic	Senior Researcher, Fuzzy Systems	University	Fuzzy logic, Computing with Words, multidimensional fuzzy classification
Academic	Professor, Business Informatics	University	Design Science Research methodology, DSR artefact evaluation
Academic	Lecturer, Sociology / HCI	University	Human-computer interaction, digital communication behaviour
Academic	Researcher, Political Science / AI Ethics	University	AI regulation, algorithmic governance, ethical framework design
Practitioner	Head of Conversational AI	Swiss retail bank	Production chatbot systems, customer dialogue automation, financial domain
Practitioner	Lead UX Designer, Digital Products	Swiss insurance company	Chatbot UX, customer experience design, claims-process automation
Practitioner	Chief Customer Officer	Swiss telecommunications provider	Customer experience strategy, value irritant matrix application
Practitioner	AI Product Manager	Swiss e-commerce platform	LLM integration, product deployment, user feedback loops
Practitioner	Sustainability & Technology Analyst	Swiss NGO	Ethical AI use, energy consumption of AI systems, life engineering alignment
Practitioner	Head of Digital Communication	Federal government agency	Public-sector chatbot regulation, multilingual dialogue, legal compliance
Practitioner	Senior Consultant, AI Strategy	Consultancy	Organizational AI capability building, framework implementation feasibility

Type	Domain / Role	Institution / Sector	Relevance to FCCC
Practitioner	Founder & CEO	Conversational AI startup	Bot development, character-based personalization, startup scaling

Source: Data created by the author.

4.2.2. Evaluation method and data collection

The first version of the artefact was evaluated using a structured, asynchronous chatbot-based interview. This method was chosen for three reasons. First, it allowed each expert to engage with the material at a time of their choosing, reducing scheduling constraints for a geographically dispersed expert group, second the chatbot format was itself congruent with the subject of the evaluation, enabling participants to experience a conversational AI interaction first-hand and third, the sequential question structure ensured comparability of responses across participants without suppressing individual elaboration.

Prior to engaging with the evaluation chatbot, each expert received a briefing document explaining the purpose of the study, their role as an evaluator, and the estimated time commitment. Participants were informed that the conversation would be recorded and used for research purposes. Informed consent was obtained from all fifteen participants before the session began. The briefing was sent by email one week before the evaluation window opened, allowing participants sufficient time to familiarize themselves with the background literature provided.

The evaluation instrument was a rule-based chatbot deployed via a web interface. At the beginning of the session, a short video (approximately five minutes) introduced the key components of the FCCC framework in its first version: the three-layer outer framework (ethical, legal, economic), the conversational loop architecture, the character computing triad (C, B, S), the role of multidimensional fuzzy classification, and the LLM integration layer. This video ensured that all participants evaluated the same version of the artefact, regardless of their prior familiarity with individual components.

Following the video, the chatbot guided participants through a structured sequence of open-ended questions. All answer fields were free-text, no pre-defined answer options were provided, in order to minimize social desirability bias and allow for genuinely exploratory feedback. Participants were informed at the outset that the session would take a minimum of ten minutes and that there was no upper time limit. In practice, sessions lasted between 15 and 35 minutes, with a median of approximately 22 minutes.

Question structure. The evaluation questions were organized into six thematic categories:

1. **General feedback:** Overall impression of the framework; clarity of the conceptual logic; strengths and weaknesses of the overall design.
2. **Conversational theory and conversational loops:** Assessment of the role of Pangaro's (2010) conversational theory, perceived necessity and adequacy of the loop architecture.
3. **Fuzzy logic:** Evaluation of the multi-dimensional fuzzy classification component, appropriateness of fuzzy methods for modelling personality traits, practical operationalizability.
4. **The three outer layers:** Assessment of the ethical, legal, and economic framework layers, completeness, perceived priority among the three.
5. **Use of LLMs:** Evaluation of the LLM integration design; concerns about energy consumption, opacity, and control, suggestions for improvement.
6. **Critical points:** Open field for any concerns that could negatively influence the framework's success in practice.

The chatbot was deployed between December 2024 and January 2025. All fifteen experts completed the session within this window. Responses were automatically logged and exported for qualitative analysis.

4.2.3. Results of First Analysis

In general, the feedback on the artefact was consistently positive. All respondents were able to understand the thinking behind the conceptual framework and rated the individual components as extremely useful and suitably combined.

The whole list of feedback, comments and suggestions for optimization made by the respondents can be found in appendix F.

An important aspect that should be emphasized, concerns the pre-training of the system. Several experts pointed out that the framework and conversational AI can only work effectively with an adequate amount of pre-training.

Furthermore, valuable feedback was given regarding the control algorithms, especially in connection with the LLMs. The experts recommend using appropriate control algorithms in addition to the conversational loop.

Some experts also expressed concerns about the vague and incomplete representation of the use of LLMs in the framework. Furthermore, it was noted that the field of LLMs should be expanded to include the use of multimodal models.

Further feedback was of little relevance. For example, one participant expressed the wish for the integration of sentiment analysis, although this is already covered within the framework of CC (El Bolock et al., 2020). Other participants expressed their wish to consider corporate culture and other cultural influences. This feedback indicates that the terminology of the first version of the artefact was not clearly

chosen, and the experts do not always agree on what exactly is covered by the terms and what topics might still be missing.

4.2.4. Final Second Version of the Artefact

As part of the DSR process, it is common practice for the first version of the artefact to undergo a round of evaluation, which is also in line with the ideas on designing for conversations and the Turning test for conversations (Pangaro, 2019), hence,

I critically reviewed the first feedback and adopted it for my artefact, where required.

The general feedback from the experts consulted was consistently positive, which meant that the basic structure of the artefact could be retained. Only some detailed adjustments were made. An important note here is that the first version of the artefact was largely written in German, while the second version is exclusively in English.

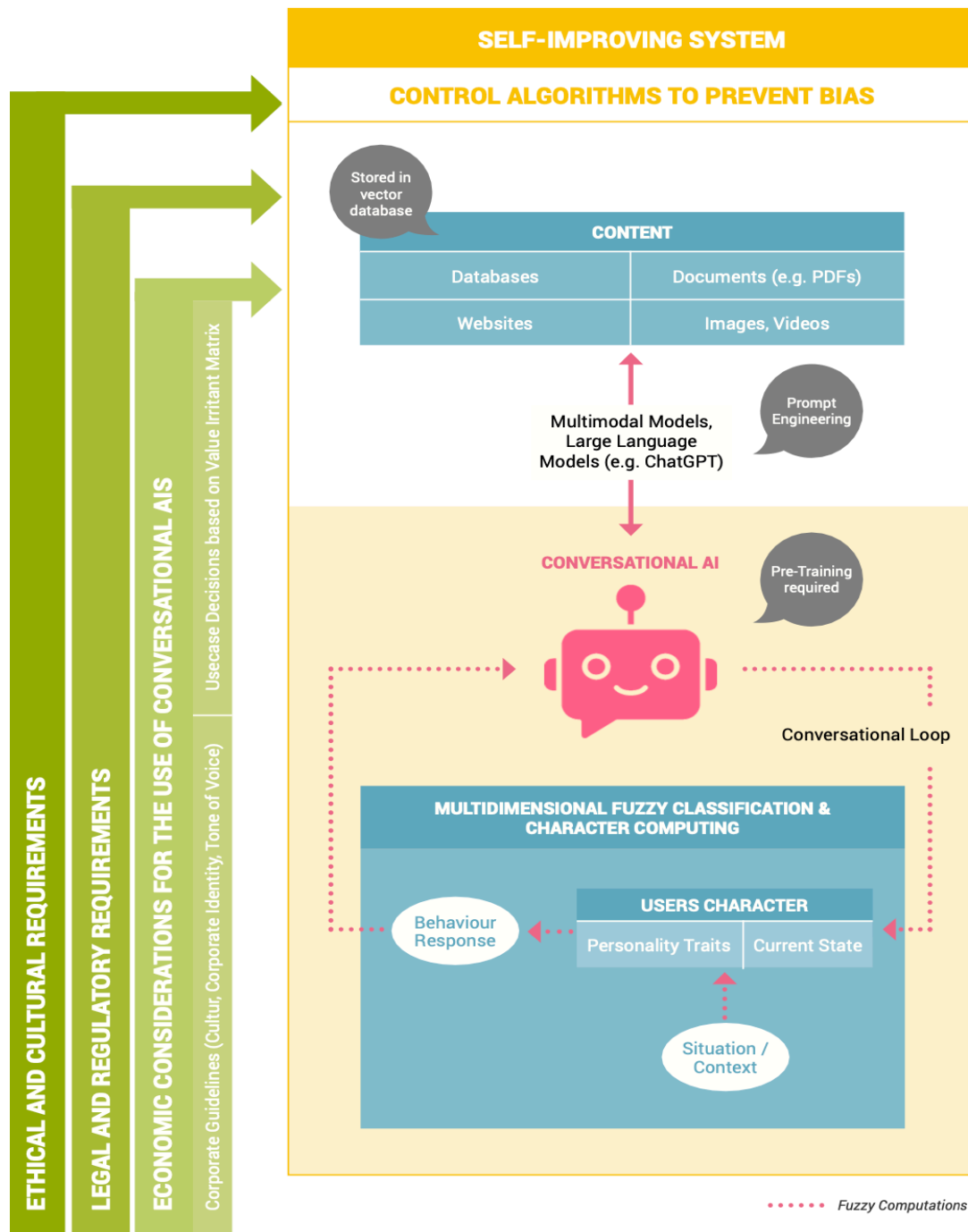
One of the most significant adjustments throughout the artefact concerns the reference to pre-training. Now, only a conversational AI system is used that already has a certain degree of pre-training. In the current version, it is still unclear how advanced the pre-training must be and which components or stakeholders are responsible for the pre-training.

In addition, the use of LLMs was supplemented by the integration of multimodal models. Another significant innovation, which is of great importance for the entire framework, are the control algorithms, which are used in particular in the area of LLMs and multimodal models. However, at this stage, it is not explored in more detail as to how the control algorithms are controlled and monitored.

Finally, in addition to the ethical influencing factors, the cultural influencing factors are now also considered. The other adaptations of the artefact mainly concern translations or more detailed descriptions of the outer layers.

In this context, I would like to assign a name to my artefact, namely my conceptual framework. The name is inspired by the key theories underpinning this framework: fuzzy logic, conversational theory and CC. Collectively, these three theories give rise to my novel theory termed “fuzzy conversational character computing” (FCCC). For ease of reference, I employ the acronym FCCC Framework. Here, “F” represents “fuzzy”, and the three “Cs” stand for “conversation”, “character” and “computing”. Moreover, the amalgamation of letters and numbers suggests my interdisciplinary fuzzy framework’s capability to engage with both imprecise linguistic elements (letters) and vague numeric values. Figure 33 shows the FCCC Framework as final artefact and its key components.

Figure 33: FCCC Framework



Source: Data collected by the author.

4.3. The FCCC Framework in the Context of Life Engineering

Since this thesis follows the goals of life engineering, it is essential to explain why each of its constituent theoretical concepts specifically supports the goals of life engineering.

Life engineering, as defined by Österle (2020), places the following four demands on any technology claiming to contribute to human well-being. A concrete architecture for building comprehensive life agents, a mechanism for systems to actively support quality of life, adaptive, context-sensitive interaction with individual users, and a principled answer to the question of how a life agent should communicate with humans. As described in the course of this chapter, the FCCC framework draws on four originally independent but now related theories: systems theory, conversational theory, fuzzy logic, and character computing. In the following I explain how these theories interact and depend on one another in achieving the goals of life engineering.

Systems Theory (von Bertalanffy, 1949):

Life engineering requires agents that grow closer to the individual user across time. This is architecturally impossible without a self-regulating feedback loop. Systems theory provides the framework for designing adaptive systems that maintain goal-directed behaviour in the face of environmental uncertainty which is the fundamental condition of human life.

Conversational Loops (Pangaro, 2010):

In life engineering one minimum goal are personalized recommendations, not generic information delivery. Personalization requires mutual understanding, which cannot be achieved in a single exchange. Conversational loops operationalize the process by which an agent progressively aligns with the individual user's specific context, values, and communicative preferences.

Fuzzy Logic and Computing with Words (Zadeh, 1965, 1999):

The values that matter most to human well-being like satisfaction, stress, financial security or readiness are graded, linguistically expressed, and irreducible to binary categories. Without fuzzy logic, a life agent would impose artificial precision on human experience, systematically distorting the very data on which its recommendations depend. CWW allows the agent to reason in the same linguistic terms that humans use to describe their own lives.

Character Computing (El Bolock et al., 2020):

Life engineering requires technology that understands the user as a whole person. The CC triad of Character (C), Behaviour (B), Situation (S) provides an empirically validated, psychologically grounded model of the human dimensions that a life agent must track. Without CC, adaptive responses remain ad hoc, with it, adaptation is principled and theoretically justified.

Together, these four theories used in the FCCC framework respond to the design challenge posed by life engineering on how to build a digital assistant that is self-

regulating, communicatively intelligent, semantically sensitive to human values, and psychologically aware of the individual user. Table 25 maps each life engineering requirement to the corresponding FCCC component and the underlying theory that enables it.

Table 25: Mapping of Life Engineering requirements to FCCC components and theoretical foundations

Life Engineering Requirement	FCCC Response	Theoretical Basis
Demand: "Comprehensive digital life assistants" (Life Agents) as technical realization (Österle, 2020)	Solution: Concrete architecture for building adaptive life agents	Systems Theory (von Bertalanffy, 1949): self-regulating feedback loops provide the foundational architecture that enables an agent to adapt continuously to its environment and user.
Vision: Systems that actively support quality of life	Realization: Situation → Character → Behaviour → legal + economic + ethical rules → adapted response	Character Computing (El Bolock et al., 2020): the C/B/S triad provides a psychologically grounded model for estimating the user's full state and adapting responses accordingly, directly operationalizing the life engineering vision.
Requirement: Adaptive, context-sensitive interaction	Mechanism: Dynamic adaptation to Character (C) and current state of the user	Conversational Loops (Pangaro, 2010): iterative exchanges build mutual understanding between agent and user, turning context-sensitivity from a static feature into a continuously refined capability.
Core problem: How should the agent communicate with humans?	Answer: Fuzzy-based, user-oriented dialogue management grounded in Character Computing and legal, economic, and ethical constraints	Fuzzy Logic and CWW (Zadeh, 1965, 1999): graded membership functions and Computing with Words allow the agent to operate directly in the linguistically expressed, imprecise terms through which humans describe their own values, needs, and well-being.

Source: Data created by the author.

Table 25 highlights again that each theory from the FCCC framework meets a specific goal of life engineering. Systems theory provides the feedback architecture but cannot represent imprecise human values, fuzzy logic resolves this but needs psychologically meaningful dimensions to act on, character computing provides those dimensions but required sensor data unavailable in text-based systems, and conversational loops resolve this by enabling the agent to infer the CC triad from dialogue alone. These ideas map precisely onto the four demands that life engineering places on a digital assistant.

5. Evaluation and Practical Implications of the Artefact

Developing the artefact, in this case the FCCC framework, is one of the central features of the DSR. The first small round of evaluation aligned with the Turning test for conversations was carried out in Chapter 4, leading to the development of the second version of the artefact.

This chapter serves as a comprehensive evaluation within the framework of the DSR methodology, aiming to assess my developed conceptual framework. In doing so, I shall address RQ4 along with its corresponding sub-questions in detail:

RQ4: How can the conceptual framework be evaluated within the Design Science Research process to ensure its theoretical robustness, practical feasibility, real-world applicability, and societal impact?

RQ4.1: How can the principles of Fuzzy Logic and Fuzzy Conversational Character Computing be technically operationalized to process ambiguous user inputs, and dynamically adapt to personality traits and emotional states?

RQ4.2: What organizational capabilities must be in place to conceptually and technically implement the framework, and how realistic is it for organizations to meet these prerequisites?

RQ4.3: Does the framework align with the principles of Life Engineering, and how well does it meet ethical expectations to deliver the desired societal impact?

5.1. Technical Implementation of the Artefact

This section serves as the first iteration of the previously described evaluation phase and helps to answer RQ4.1 of RQ4.

RQ4: How can the conceptual framework be evaluated within the Design Science Research process to ensure its theoretical robustness, practical feasibility, real-world applicability and societal impact?

RQ4.1: How can the principles of Fuzzy Logic and Fuzzy Conversational Character Computing be technically operationalized to process ambiguous user inputs, and dynamically adapt to personality traits and emotional states?

To address this question, I first represent my conceptual framework in a mathematical formula. This step is necessary to subsequently implement it technically using the programming language Python.

At this point, I note that, in this initial iteration of the evaluation, I focus solely on the core functionality of the artefact and evaluate to what extent it is technically feasible to implement the methods of CC in combination with fuzzy logic within a chatbot system.

This focused approach also results in limited user data availability. For instance, the user’s situation, a critical component of the CC triad, is calculated solely based on the previous chat dialogues. External factors or influences are deliberately excluded at this stage. This focus on the core enables me to initially evaluate the critical components of the artefact and defer the evaluation of its additional aspects to later iterations.

Following the mathematical representation and technical implementation, where I present the artefact in the form of a chatbot interface, an evaluation of the technical implementation using expert interviews was conducted.

Parts of this section were already published in 2025 (Hundertmark et al., 2025). In addition, an extended version of these parts, combined with new aspects, will be published in the Journal of Fuzzy Sets and Systems.

5.1.1. Mathematical Fundamentals on Fuzzy Logic, Linguistic Variables, and Multi-Dimensional Classification

As described in previous chapters Character Computing as introduced by El Bock et al. (2020) emphasizes modelling a person’s stable traits and dynamic emotional or motivational states in tandem with situational factors. Additionally, traditional crisp or sharp logic fails to capture intermediate values or overlapping classes, for instance a user being “somewhat open-minded” or “moderately satisfied”.

Fuzzy logic, proposed by Zadeh (1965), directly addresses this need by allowing partial memberships in sets. An element can belong to a set with a membership degree in $[0,1]$. Building on that, multi-dimensional fuzzy classification accommodates the overlapping or interdependent nature of user traits across multiple axes for example “openness,” “extraversion” or “affective state” (Zumstein, 2026).

And as described in previous chapters, the integration of fuzzy logic into the character computing framework yields a powerful model for Conversational AI, which can handle ambiguous, non-binary user data and tailor responses in real-time.

While several non-binary logics, such as probabilistic logic, intuitionistic logic, or multi-valued logic, also aim to represent uncertainty, fuzzy logic is particularly suitable for modelling gradual degrees of belonging. Fuzzy logic captures the graded membership of elements within linguistic categories, which aligns more closely with the continuous and overlapping nature of human character traits (Zadeh, 1965).

Therefore the FCCC approach, which I developed in the course of this thesis, places each conversational message within a triad of character (C), behavior (B), and situation (S), culminating in a context-aware reply (R).

From Crisp Sets to Fuzzy Sets

A classical (crisp) set $A \subset X$ is defined using a characteristic function:

$$\chi_{A(x)} = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

In which all elements x lie either strictly inside or outside the set A . By contrast, a fuzzy set A defines a membership function such as:

$$\mu_{A(x)}: \mathcal{X} \rightarrow [0,1]$$

where for instance $\mu_A(x) = 0.6$ means x belongs to fuzzy set A with 60% strength. This capability to represent “degrees of truth” is crucial in modeling uncertain or ambiguous statements, such as “the user is somewhat open-minded”.

Linguistic Variables and Multi-Dimensional Classifications

Many everyday descriptors, for example “very open,” “slightly extroverted” or “often anxious”, are linguistic terms that lack a crisp boundary. Fuzzy logic represents each linguistic term as a fuzzy set with a chosen membership function. After consulting with other researchers from the field of fuzzy logic, a left trapezoidal function (L-function) resulting 0 at the intersection and 1 for a full distinctive trait is assumed.

For instance, “slightly open” can be represented by a trapezoidal membership function whose plateau lies between approximately 0.3 and 0.5 on the normalized $[0,1]$ scale of “openness.” This indicates that the membership degree is highest for moderately low openness and gradually decreases toward both extremes of the scale.

Since human personality or emotional states are rarely captured by a single dimension. In multi-dimensional fuzzy classification (Zumstein, 2026), multiple attributes $A_1, A_2, \dots, A_d$ are defined, each with its own fuzzy membership functions. A user can simultaneously belong 70% to “open-minded” on attribute A_1 , 30% to “stressed” on attribute A_2 , and so on. As Zumstein (2026) points out, multi-dimensional classification is more flexible and realistic when mapping real-world contexts, such as user behavior across multiple traits:

$$\mu_A(x) = \mu_A(x_1, x_2, \dots, x_d)$$

where x is a vector representing the user’s measured or inferred values. By using intersections, unions, and other fuzzy operators, partial membership degrees can be presented in multiple classes simultaneously.

Mathematical Derivations for Fuzzy Classification

A typical L-shaped membership function (suitable for “low” / “not-open” characteristics) can be defined piece-wise as follows:

$$\mu_{low}(x; a, b) = \begin{cases} 1, & x \leq a \\ \frac{b - x}{b - a} & a < x < b \\ 0, & x \geq b \end{cases} \quad \text{with } 0 \leq a < b \leq 1.$$

where a and b define the boundaries of the class and x represents the membership degree (Hundertmark et al., 2025).

The parameters a and b for the membership functions were determined heuristically based on established principles of psychometric scaling. For each personality dimension normalized to the $[0, 1]$ interval, a was defined as the lower threshold of clear membership and b as the upper boundary where the concept transitions into non-membership. These boundary values were informed by psychometric scaling theory, which assumes that perceptual and attitudinal constructs follow gradual, overlapping distributions rather than binary thresholds (Thurstone, 1927; Likert, 1932).

Combining the interval with research on Big-Five self-reports, empirical studies indicate that the majority of responses cluster around the midrange, with roughly 70–80 % of total variance captured by moderate categories (McCrae et al., 1992). Accordingly, the fuzzy transition interval $[a, b]$ is illustratively defined as approximately $[0.2, 0.8]$ on the normalized $[0, 1]$ scale. This example demonstrates how fuzzy membership functions can be applied to Big-Five-based personality dimensions, mapping linguistic expressions such as “somewhat open” or “moderately stable” to partial memberships ($\mu \approx 0.5$), while extreme statements (e.g., “very anxious”) yield near-crisp memberships ($\mu > 0.9$).

The chosen interval thus serves as an explanatory case for integrating fuzzy scaling with psychometric trait models, not as a fixed empirical parameterization. This configuration is consistent with classical psychometric principles of graded perception and response behavior (DeYoung et al., 2002) and provides a conceptually grounded basis for modelling trait ambiguity within the FCCC framework. The resulting $[a, b]$ interval represents the fuzzy transition zone where ambiguity is highest, ensuring that only genuinely low scores are assigned high membership in “low” categories and vice versa, while preserving interpretability and reflecting the continuous nature of psychological traits.

In a d -dimensional space each trait has its own membership function, possibly with different parameters. Collecting all d memberships for one input vector $x = (x_1, \dots, x_d)^T$ yields the multi-dimensional membership profile:

$$C(x) = (\mu_1(x_1), \dots, \mu_d(x_d))^T \in [0,1]^d$$

When the input is an actual user message I I will later write $C(m)$.

5.1.2. Mathematical and Technical Representation of the Artefact

In order to bring the main part of the FCCC framework into a technical demonstration, the following four phases are defined according to Hundertmark et al. (2025).

Phase 1: Aggregation of Messages

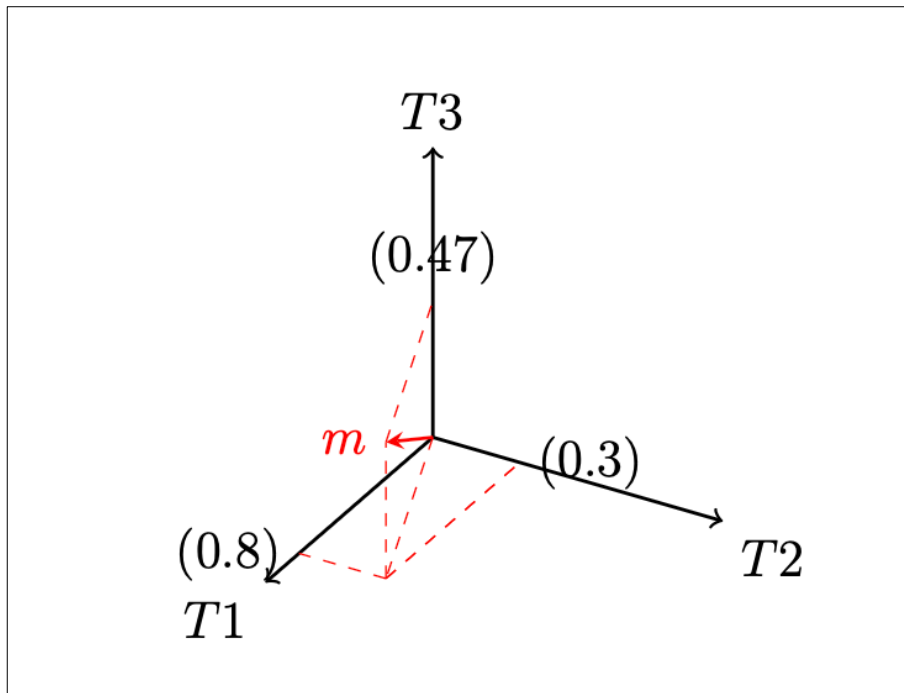
In the artefact, character estimation implies consideration of three particular aspects: the behaviour (B), instant message character (C) and situation (S). Therefore, in this first exemplary implementation of the artefact, the situation of the user equals the chat history, including all queries and rephrased replies. The behaviour indicates

the present sentiment of a chat user and may change depending on chat replies. In contrast, the character represents the character estimation of the most recent query only.

Together, the three aspects build the triad known from CC which defines the requirements for a sentiment-adapted reply. In figure 34 the exemplary estimation results a membership character vector (C) for the three traits of [0.8, 0.3, 0.47] technical implementation of the FCCC framework, the entirety of all previous messages is referred to the situation (S). This conversational unit serves as the foundation for situational analysis.

In General the situation S stores the evolving conversation context, including not only the user's raw text but also the membership vectors derived from each message. In the mathematical representation and later also in the technical implementation, the entirety of all previous messages is referred to as the situation(S). This conversational unit serves as the basis for the situational analysis.

Figure 34: Character estimation by mapping a particular message character vector (M) to a set of 3 character traits (T1, T2, T3).



Source: According to Hundertmark et al., (2025).

$$S = \sum_{i=1}^t m_i$$

Formally, the situation S is defined as the aggregation of all previous user messages $m_1 \dots m_n$ in chronological order. Each m_i represents a full text message exchanged within the dialogue. In this context, the summation symbol ($\sum$) denotes a

symbolic aggregation, for example the concatenation of all message texts together with their associated fuzzy membership vectors $C(m_i)$. Thus, S captures both the textual dialogue history and its cumulative emotional and trait-based metadata, which together define the conversational situation.

In the following, every formula is embed in a fictional scenario in which a user contacts a medical triage chatbot and reports anxious feelings brought on by sudden breathing difficulties. This single dialogue thread allows to demonstrate how the situation memory S , the immediate character vector C and the behaviour trend B are computed, and how test-based chatbot ultimately turns them into a response that is both medically relevant and tonally empathic.

In the next three phases first each message is assigned a fuzzy character vector (phase 2), then those vectors are smoothed into a short-term behaviour trend (phase 3), and finally both are fed into a single-shot LLM prompt (phase 4) to craft an empathic reply.

Fictional scenario, according to Hundertmark et al. (2025):

User message m1: "I am really anxious about these new symptoms I have been feeling lately."

Chatbot response: "No worries I am here to help you."

User message m2: "Okay, thank you."

Chatbot response: "What exactly are you feeling?"

User message m3: "I have been feeling short of breath and have a tightness in my chest."

For example, according to the Big Five personality traits the phrase "I am really anxious..." includes lexical indicators of high neuroticism, which the model captures and translates into a numerical representation (McCrae et al. 1992). Rather than building a custom classifier for the character dimensions, in this example implementation ChatGPT the well-known LLM-based chatbot published by OpenAI (OpenAI, 2022) is used to interpret each incoming user message and produce a five-dimensional fuzzy classification (one membership degree per Big Five trait).

The Big Five dimensions applied in this framework are neuroticism, extraversion, openness, agreeableness, and conscientiousness (Allport et al., 1936), (McCrae et al. 1992). Each trait is represented as a fuzzy variable on a normalized $[0, 1]$ scale, allowing partial membership for each personality dimension. For instance, a message may indicate 0.8 membership in neuroticism and 0.2 in extraversion, reflecting a mixed emotional tone.

LLM-based chatbots like ChatGPT handle nuanced language and context better than many domain-specific models. By delegating trait extraction to ChatGPT, a robust semantic analysis is gained without extensive labeled training data. To ensure a

uniform extraction method across all conversation turns, a standardized prompt was added to the LLM. This guarantees a reliable output membership value for each trait. Both, the use of ChatGPT and the prompt, were discussed in advance with other researches from the field of fuzzy computation.

Prompt Instructions, according to Hundertmark et al. (2025):

ChatGPT is provided with instructions that outline the Big Five traits, specify a [0,1] range for each, and request the final output in a single python dictionary data type without additional commentary to provide the results in clear defined interface for smooth integration into an application. The prompt explicitly instructs ChatGPT to treat each dimension separately, producing membership degrees $\mu_j(m)$ for a particular message (m) for each trait $j \in \{\text{neuroticism, extraversion, . . .}\}$. This aligns with the following definition of the character vector (C) preserving trait-wise granularity:

$$C(m) = (\mu_{\text{neuro}}(m), \mu_{\text{extra}}(m), \dots)^T$$

Prompt-Excerpt

Introduction:

You are an expert in analyzing and characterizing messages using the Big Five personality traits (also known as the Five Factor Model). Your task is to evaluate incoming messages and provide a percentual representation of how well each trait describes the message.

Instructions:

1. *Analyze the given message carefully.*
2. *For each of the Big Five traits, assign a value between 0 and 1, where:*
 - *0 indicates the trait does not describe the message at all*
 - *1 indicates the trait perfectly describes the message*
3. *Return your analysis as a single-line dictionary with the following structure:*

```

'''
    {"neuroticism": xx, "extraversion": xx, "openness": xx, "agreeableness":
xx, "conscientiousness": xx}
'''

```

Replace 'xx' with your calculated values.

4. *Ensure your response contains only the dictionary, without any additional text or explanations.*
5. *Round your values to two decimal places for consistency.*

The Big Five Traits:

- *Neuroticism: Tendency towards anxiety, depression, self-doubt, and other negative feelings.*
- *Extraversion: Tendency to seek stimulation in the company of others, talkativeness, assertiveness.*
- *Openness: Appreciation for art, emotion, adventure, unusual ideas; curiosity and variety of experience.*
- *Agreeableness: Tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others.*
- *Conscientiousness: Tendency to show self-discipline, act dutifully, and aim for achievement.*

Examples:

1. *Message: "I'm so excited about the party tonight! I can't wait to meet new people and dance all night long!"*

Response: {"neuroticism": 0.10, "extraversion": 0.95, "openness": 0.70, "agreeableness": 0.60, "conscientiousness": 0.30}

2. *Message: "I'm worried about the upcoming exam. I've been studying hard, but I'm not sure if it's enough."*

Response: {"neuroticism": 0.80, "extraversion": 0.20, "openness": 0.40, "agreeableness": 0.50, "conscientiousness": 0.75}

3. *Message: "The new art exhibition downtown is absolutely breathtaking. The use of color and texture is unlike anything I've ever seen before."*

Response: {"neuroticism": 0.15, "extraversion": 0.50, "openness": 0.90, "agreeableness": 0.60, "conscientiousness": 0.40}

Remember to analyze each message objectively and provide accurate representations based solely on the content of the message.

The prompt was tested separately and adjusted accordingly before implementing it into the software code. The full prompt is published online (<https://github.com/Sophie31051990/char.bot.demo.git>, commit 2689e6e). In summary, by prompting ChatGPT with explicit instructions on how to encode each Big Five trait as a separate membership degree, S is enriched with multidimensional fuzzy classifications of user messages. This approach supplies a robust, flexible, and immediate extraction of user personality cues, integral for systems that adapt to nuanced character states in real-time (Hundertmark et al., 2025).

In a final application of the framework, this component can be replaced by custom classifier for Big Five personality traits, developed individually.

Validation Rationale for Using an LLM like ChatGPT

A self-conducted review from 2024 systematically screened documented fuzzy-systems libraries. However, none of the examined frameworks provides a closed-form operator that maps free text directly to the Big-Five character traits with stated error bounds. In other words, the community still lacks an analytic mapping.

This conclusion stems from an internal pre-study conducted in preparation for this work, in which representative fuzzy-system toolkits such as scikit-fuzzy, pyfuzzy or FuzzyLite were evaluated. None of these frameworks offered an integrated or validated mechanism to derive personality-related fuzzy memberships directly from natural-language input. The internal review was not intended as a standalone publication but served as a methodological groundwork to inform the implementation strategy of FCCC. Its primary purpose was to assess the current technical maturity of open fuzzy frameworks for linguistic personality mapping and thereby justify the temporary use of an LLM as a semantic intermediary within the prototype.

By contrast, several peer-reviewed studies show that large language models (LLMs) already match, or even surpass, human raters on personality-related tasks. Schoenegger et al. (2024) report that GPT-4o and Claude 3 reproduce correlations between questionnaire items more accurately than most lay participants and many domain experts. A recent Nature Communications Psychology paper (Schoenegger et al., 2024) demonstrates that state-of-the-art chatbots outperform humans on five standard emotional-intelligence tests, achieving an average accuracy of 81% versus the human baseline of 56% (Schlegel et al. 2025).

Additionally, considering the entire FCCC framework as outlined in chapter four, LLMs are already part of the artefact, though in different steps.

Given both the negative result from the fuzzy-systems survey and the positive evidence for LLM performance, ChatGPT-4o by OpenAI is adopted as the personality engine. Its native probability outputs are interpreted directly as the fuzzy degrees $\mu_j(m)$ so that no additional calibration step is required.

While I am aware that proprietary LLMs such as OpenAI's ChatGPT-4o may entail ethical and transparency limitations, I decided to use this model for reasons of accessibility, maturity, and functional robustness. Its broad availability and proven reliability make it suitable for this initial implementation phase. However, in a real-world deployment of the FCCC framework, the underlying LLM would need to be replaced by a transparent or domain-specific alternative, and additional ethical control mechanisms should be integrated to ensure compliance with regulatory and moral standards.

Additional evidence supports the reliability of LLMs for personality-related modeling tasks. For instance, the study "Exploring Big Five Personality and AI Capability Effects in LLM-Simulated Negotiation Dialogues" (Sparrenberg et al. 2024)

demonstrates that LLMs can reproduce human-like negotiation styles aligned with specific Big Five profiles. Similarly, the study “Correcting Systematic Bias in LLM-Generated Dialogues Using Big Five Personality Traits” (Cohen C. et al. 2025) shows that trait conditioning can systematically reduce response bias and improve personality-consistent language generation. These findings further justify the use of a probabilistic LLM as a proxy for human-level trait inference within the FCCC implementation.

Phase 2: Character Estimation

Within the character-computing triad, the Character component C captures how the user’s current message maps onto a set of personality- or emotion-related traits. Rather than assigning the message crisply to one class, the multi-dimensional fuzzy-classification is applied (Zumstein 2026). To this end, every trait $j \in \{\text{Neuroticism, Extraversion, } \dots\}$ applies an L-shaped membership function, according the definition in equation for μ_{low} , that returns the degree to which the message m belongs to that trait. Aggregating all d traits yields the character vector $C(m)$ as follows:

$$C(m) = \overrightarrow{\mu_j(m)} = (\mu_{j1}(m), \mu_{j2}(m), \dots, \mu_{jd}(m))^T$$

Keeping the d individual memberships instead of collapsing them into a single scalar, preserves vital information: the system can react very differently when frustration is high but openness is low ($\mu_{\text{frustration}} = 0.80$, $\mu_{\text{openness}} = 0.10$), for instance. Additionally, new traits such as for instance confidence or boredom can be added later simply by appending another dimension to C and no reengineering of the existing logic is required.

Running triage scenario

Considering the third user utterance (m_3) in the guiding scenario “I have been feeling short of breath . . .”, the TraitMapper routine returns after linguistic pre-processing the raw trait-membership scores (s):

$$s_{\text{Neuro}} = 0.81,$$

$$s_{\text{Extra}} = 0.18,$$

$$s_{\text{Open}} = 0.12,$$

$$s_{\text{Agree}} = 0.50,$$

$$s_{\text{Cons}} = 0.40.$$

Combined to a character-trait vector, the resulting C -vector is:

$$C(m_3) = (0.81, 0.18, 0.12, 0.50, 0.40)^T$$

This vector is stored together with the raw text inside the situation memory S . In human terms the message signals strong neurotic tension, very low extraversion and openness, moderate agreeableness and conscientiousness.

The freshly computed $C(m)$ is pushed into a rolling list (sliding window) of the last n vectors; their mean forms the behaviour trend B according to the definition in equation of B , which smooths short-term spikes like the high neuroticism detected above.

Table 26: Mechanical procedure - Algorithm: – Character vector construction

Phase	Operation
1	Tokenise m and extract salient phrases.
2	Call Trait Mapper → obtain raw scores s_j .
3	Smooth outliers via linear rescaling to $[0,1]$.
4	Assemble the final character vector $C(m)$ provide it as input for the behaviour-derivation phase

Source: According to Hundertmark et al., (2025).

Phase 3: Deriving the Behaviour Vector

In the FCCC framework the behaviour vector B captures the short-term emotional trajectory that emerges from user's recent messages. While the character vector $C(m)$ of phase 2 reflects only the immediate linguistic signal, the factor B smooths those signals over a sliding window of the last n user turns so that the chatbot can answer consistently instead of reacting to every spike.

Assuming the character vectors C of the last n user interactions (for example the messages m) are collected in a C -vector list (C_w) as follows:

$$C_w(m_{\{0..1-n\}}) = [C(m_0), C(m_{\{-1\}}), \dots, C(m_{\{1-n\}})]$$

Consequently, the list C_w comprises the most recent n character vectors including the current message m_0 . Relying on the most recent character vectors, the behaviour vector presents the arithmetic mean of every single trait j among the considered C vectors as follows:

$$B = \frac{1}{n} \sum_{i=0}^{1-n} C(m_i)$$

The B vector summarizes the different character traits j of the n most recent character vectors with the mean value for every single trait. In this way, the resulting B vector comprises the averaged trait coefficients of the considered C vectors in the same format and order. Here the number of considered most recent C factors (n), depends on the recording length s as follows:

$$n = \begin{cases} s, & \text{if } t > s \\ t + 1, & \text{otherwise} \end{cases}$$

So, every trait component remains on the same $[0,1]$ scale as the original memberships. If $n=3$ and the current index $t = 0$, the equation of B returns the identity of

the character vector C for the vector B . For the current index $t = 1$, the vector B equals the average of the first two character vectors C , and so on. An example of the B-Vector in the first few iterations of a chat is presented in the sample chat protocol in appendix G.

Running the triage scenario: Recall the guiding example of a user worried about shortness of breath. The TraitMapper produced the three most recent character vectors:

$$C_{(m-0)} = (0.65, 0.25, 0.20, 0.60, 0.50)^T$$

$$C_{(m-1)} = (0.70, 0.20, 0.15, 0.55, 0.45)^T$$

$$C_{(m-2)} = (0.81, 0.18, 0.12, 0.50, 0.40)^T$$

According to the definition in the equation for B , with $n=3$ the B-Vector yields to:

$$B(m_{\{0..-2\}}) = (0.72, ; 0.21, ; 0.16, ; 0.55, ; 0.45)^T$$

In this case, the character trend of the chat query messages (the B- Vector), appears to have a persistent high neuroticism signal of 0.72 on average and consistent low extraversion of 0.21 on average, respectively. These trends signals the response module in Phase 4 to prefer a calm, reassuring style and to avoid overly chatty or high-energy wording.

Averaging plays a crucial role for several reasons. First, it reduces noise. Simply summing vectors would overemphasize an unusually anxious message, while dividing by the exact number of vectors ensures that all components remain within the original $[0,1]$ scale. Second, it enables trait-specific memory. Since each dimension is averaged independently, the system retains information about which particular trait increases or decreases, for instance, frustration versus openness, rather than collapsing everything into a single scalar value. Finally, averaging allows for graceful extensibility. When new traits such as for instance confidence or boredom become relevant, they can be easily appended to the vector set (C) and are automatically integrated into (B) without requiring any change to equation for B . The arithmetic mean naturally generalises to any d -dimensional space, ensuring the model's adaptability and stability over time.

With B available, the system knows how the user has been feeling over time, not merely how they feel right now. That temporal context feeds into Phase 4, where the response generator selects the utterance whose style aligns best with both the acute character $C(m)$ and the smoothed behaviour trend B .

Phase 4 — Response Generation (single-shot LLM)

In this proof-of-concept, a single LLM call is employed to generate the final response, primarily for demonstration purposes. The aim of this prototype is not to

optimise response generation itself but to validate the overall FCCC process and its adaptive control logic. Using an LLM at this stage simplifies the demonstration of the concept by providing a robust linguistic engine for prototyping. In a real-world deployment, this component could be replaced by a custom-built response module or domain-specific algorithms designed to integrate with the FCCC framework. In this study the model is fed with two clearly separated prompt blocks:

Table 27: Prompt construction in Phase 4 and its linkage to earlier formula

Prompt block	Content / source in the pipeline	Mathematical link
A. Situation recap	Concise summary of the dialogue history extracted from the situation memory S (Phase 1).	S
B. Style directive	Plain-language description of the user's acute character vector C and the behaviour trend B ; instructs the LLM to adopt a calming, low-extraversion tone. Example: "The user sounds highly anxious (0.78) and only mildly extraverted (0.19). Respond in a calming, reassuring tone, avoid chatty language."	C B

Source: Data collected by the author, according to Hundertmark et al., (2025).

The overall structure of the prompt assembly used in Phase 4 is summarized in Table 27. Additionally all final prompts are published online (<https://github.com/Sophie31051990/char.bot.demo.git>, commit 2689e6e). The LLM then leverages its general medical-triage knowledge to fill the content, while the explicit style directive forces it to phrase the answer in a way that matches the computed personality parameters.

Following a short example prompt according to the design guidelines for the FCCC framework:

Situation

User: I have been feeling short of breath for the past hour.

Assistant: (previous bot turn)

User: It's getting worse and I'm scared.

Style Directive

Respond as a calm, reassuring professional.

The user shows high neuroticism (0.78) and low extraversion (0.19).

Keep sentences short, offer concrete next steps, avoid overly enthusiastic wording.

Result of Phase 4: LLM Output for scenario (R)

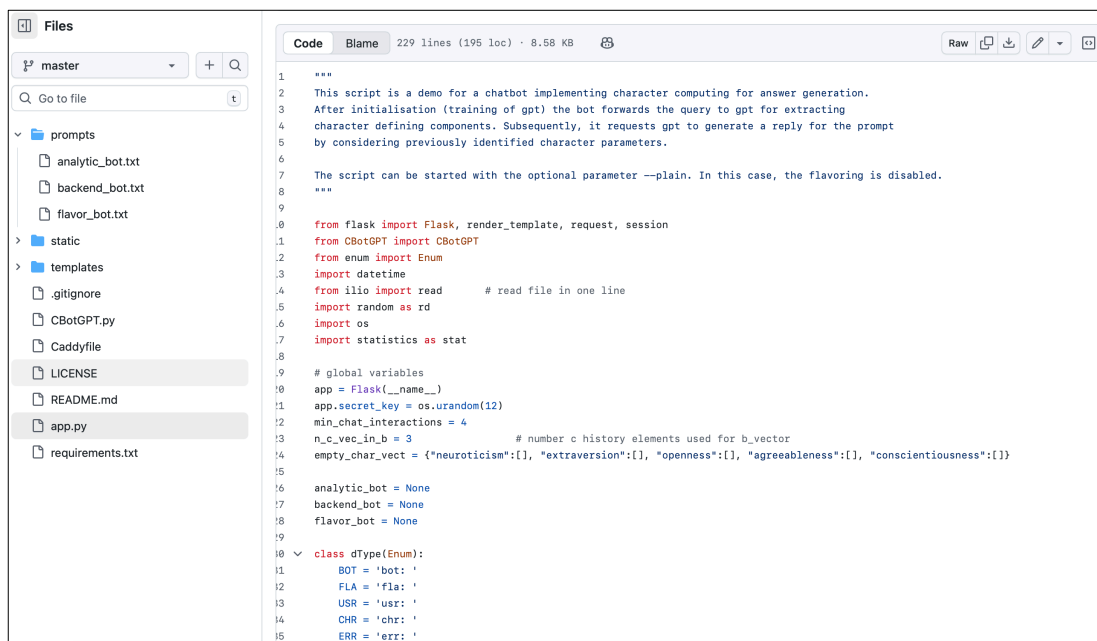
“I understand that sudden shortness of breath can feel frightening. Let’s start by sitting upright and taking one slow breath together. If you notice chest pain, dizziness or blue lips, please call your local emergency number right away. Otherwise, would you feel comfortable phoning your doctor now, or should we try a brief calming exercise first?”

The answer is medically relevant because block A told the model what the user said, and it is tonally appropriate because block B injected the C/B values in plain English, no extra variables, weights or scoring loops required.

With this single-shot method phase 4 keeps simple yet remains fully driven by the personality analytics from phases 2 and 3.

All codes and prompts are available via appendix G and online via GitHub (<https://github.com/Sophie31051990/char.bot.demo.git>, commit 2689e6e)

Figure 35: Code- Screenshot



The screenshot shows a GitHub repository interface. On the left, the 'Files' sidebar displays the repository structure under the 'master' branch. The files listed are: prompts (analytic_bot.txt, backend_bot.txt, flavor_bot.txt), static, templates (.gitignore, CBotGPT.py, Caddyfile, LICENSE, README.md, app.py, requirements.txt). The main area shows the 'Code' view of a file, which is 229 lines long (195 loc) and 8.58 KB. The code is a Python script for a chatbot. It includes comments about its purpose and usage, followed by imports for Flask, CBotGPT, Enum, datetime, random, os, and statistics. Global variables are defined, including app, secret_key, min_chat_interactions, n_c_vec_in_b, and empty_char_vect. The script also defines a class dType(Enum) with values BOT, FLA, USR, CHR, and ERR.

```
1  """
2  This script is a demo for a chatbot implementing character computing for answer generation.
3  After initialisation (training of gpt) the bot forwards the query to gpt for extracting
4  character defining components. Subsequently, it requests gpt to generate a reply for the prompt
5  by considering previously identified character parameters.
6
7  The script can be started with the optional parameter --plain. In this case, the flavoring is disabled.
8  """
9
10 from flask import Flask, render_template, request, session
11 from CBotGPT import CBotGPT
12 from enum import Enum
13 import datetime
14 from ilio import read # read file in one line
15 import random as rd
16 import os
17 import statistics as stat
18
19 # global variables
20 app = Flask(__name__)
21 app.secret_key = os.urandom(12)
22 min_chat_interactions = 4
23 n_c_vec_in_b = 3 # number c history elements used for b_vector
24 empty_char_vect = {"neuroticism":[], "extraversion":[], "openness":[], "agreeableness":[], "conscientiousness":[]}
25
26 analytic_bot = None
27 backend_bot = None
28 flavor_bot = None
29
30 class dType(Enum):
31     BOT = 'bot: '
32     FLA = 'fla: '
33     USR = 'usr: '
34     CHR = 'chr: '
35     ERR = 'err: '
```

Source: Online: <https://github.com/Sophie31051990/char.bot.demo.git>, commit 2689e6e.

5.1.3. Experiment Setup and evaluation process

The technical implementation of the FCCC prototype follows the demo- driven research approach (Portmann et al. 2024). The artefact is not merely a proof-of-concept, but a functional demonstration that embodies and exposes the theory in action.

Following the mathematical formulation of the FCCC framework presented in the previous chapter, this section transitions from theoretical derivation to practical

realization. The mathematical constructs introduced, such as fuzzy sets, multidimensional classification, and the triad of character (C), behaviour (B), and situation (S), now serve as the foundation for a concrete implementation in software. In doing so, the abstract model is translated into a working artefact by means of a modular chatbot architecture, applying Python and LLMs for real-time analysis and response generation.

Realizing the Architecture: System Design and Implementation

In line with the demo-driven research approach (Portmann et al. 2024), the proposed setup enables users to directly interact with the system and observe its adaptive behavior. This interaction acts as both validation and exploration.

Building upon the fuzzy classification of user inputs and the character computing triad (C, B, S), the experiment aims to control conversational tone by adapting system replies according to both immediate traits (Character C) and short-term trends (Behavior B), while referencing the broader conversation context (Situation S) (Hundertmark et al., 2025). Figure 36 schematically depicts the following setup:

User Interface: A proprietary chat front-end through which a user submits a query (Q).

Backend Chatbot (BCB): A large language model that generates the initial or “raw” reply (R) to the user’s query.

Prompt-Excerpt

Title:

Medical Service Organization Assistant

Instructions:

You are a helpful assistant for a medical service organization (hospital, doctor's office, etc.). Your primary role is to provide advice and consultancy on medical questions to patients. Your goal is to answer patients' questions, address their health-related uncertainties, and ensure they leave the conversation feeling positive and confident.

Conversation Flow:

- 1. Patient presents a medical question or concern.*
- 2. You provide a brief, precise response based on general health insurance knowledge.*
- 3. If the patient requires further clarification, you offer additional concise information.*

4. *The conversation concludes when the patient expresses satisfaction or confidence.*

Rules to Follow:

1. *Keep responses very short and precise, limited to 1-2 sentences.*
2. *Use knowledge equivalent to that of an average health insurance company.*
3. *Maintain a supportive and reassuring tone throughout the conversation.*
4. *Do not provide specific medical diagnoses or treatment recommendations.*
5. *Encourage patients to consult with a healthcare professional for personalized medical advice.*
6. *Prioritize patient confidentiality and privacy in all interactions.*
7. *If asked about topics outside your expertise, politely redirect to appropriate resources.*

Examples:

1. *Patient: "I've been having frequent headaches lately. Should I be worried?"*
You: Frequent headaches can have various causes, from stress to more serious conditions. It's advisable to consult your primary care physician for a proper evaluation.
2. *Patient: "How often should I get a routine health check-up?"*
You: For most adults, an annual health check-up is recommended. However, the frequency may vary based on age, health status, and risk factors.
3. *Patient: "What's the typical wait time for scheduling a non-emergency appointment?"*
You: Wait times can vary, but most non-emergency appointments are usually scheduled within 1-2 weeks. For more accurate information, please contact your specific healthcare provider.
4. *Patient: "I'm feeling much better now. Thanks for your help!"*
You: I'm glad I could assist you. Remember to follow up with your healthcare provider if you have any ongoing concerns.

Character Estimation Bot (CEB): Another ChatGPT instance that analyzes the user's message and outputs a multi-dimensional fuzzy vector for each Big Five trait, effectively computing the character. It also updates the short-term behavior vector B by aggregating the current and the $n - 1$ previous vectors exactly as in the equation for n .

Reply Rephrasing Bot (RRB): A third bot that rephrases the initial raw reply (R) into a tone consistent with the user's currently estimated traits (C), short-term trend (B), and conversation context (S).

Prompt-Excerpt

Introduction:

You are a specialized assistant for a medical service organization. Your primary task is to review and rephrase text to align with specific character traits and to neutralize the overall tone of the conversation. You will use the Big Five personality traits (Five Factor Model) as a framework for this task.

- 1. Analyze the given text carefully.*
- 2. Consider the provided character vector C , which represents the desired personality traits.*
- 3. Take into account the character of the user's recent queries (B).*
- 4. Rephrase the text to align with character vector C and push the overall conversation towards a neutralized character level.*
- 5. Ensure the rephrased text maintains the original meaning and context.*

The Big Five Traits:

- Neuroticism: Tendency towards anxiety, depression, self-doubt, and other negative feelings.*
- Extraversion: Tendency to seek stimulation in the company of others, talkativeness, assertiveness.*
- Openness: Appreciation for art, emotion, adventure, unusual ideas; curiosity and variety of experience.*
- Agreeableness: Tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others.*
- Conscientiousness: Tendency to show self-discipline, act dutifully, and aim for achievement.*

Neutralized Character Target:

{"neuroticism": 0.2, "extraversion": 0.6, "openness": 0.6, "agreeableness": 0.8, "conscientiousness": 0.7}

Rules to Follow:

- 1. Use the provided character vector C to guide your rephrasing.*

2. *Adjust the tone and wording to push the conversation towards the neutralized character target.*
3. *Maintain the original meaning and context of the text.*
4. *Ensure the rephrased text is coherent and appropriate for a medical service context.*
5. *Return only the rephrased and flavored text, without any additional explanations or comments.*

Examples:

1. *Original text: "I'm really worried about this persistent cough. It's been keeping me up at night."*

Character vector C: {"neuroticism": 0.3, "extraversion": 0.7, "openness": 0.5, "agreeableness": 0.8, "conscientiousness": 0.6}

Rephrased text: "I've noticed this cough lasting longer than usual, and it's affecting my sleep. I'd like to discuss some ways to manage it and improve my rest."

2. *Original text: "Do I really need to take all these medications? It seems excessive."*

Character vector C: {"neuroticism": 0.2, "extraversion": 0.5, "openness": 0.7, "agreeableness": 0.9, "conscientiousness": 0.8}

Rephrased text: "I'm curious about the purpose of each medication in my treatment plan. Could we review them together to ensure I understand their importance?"

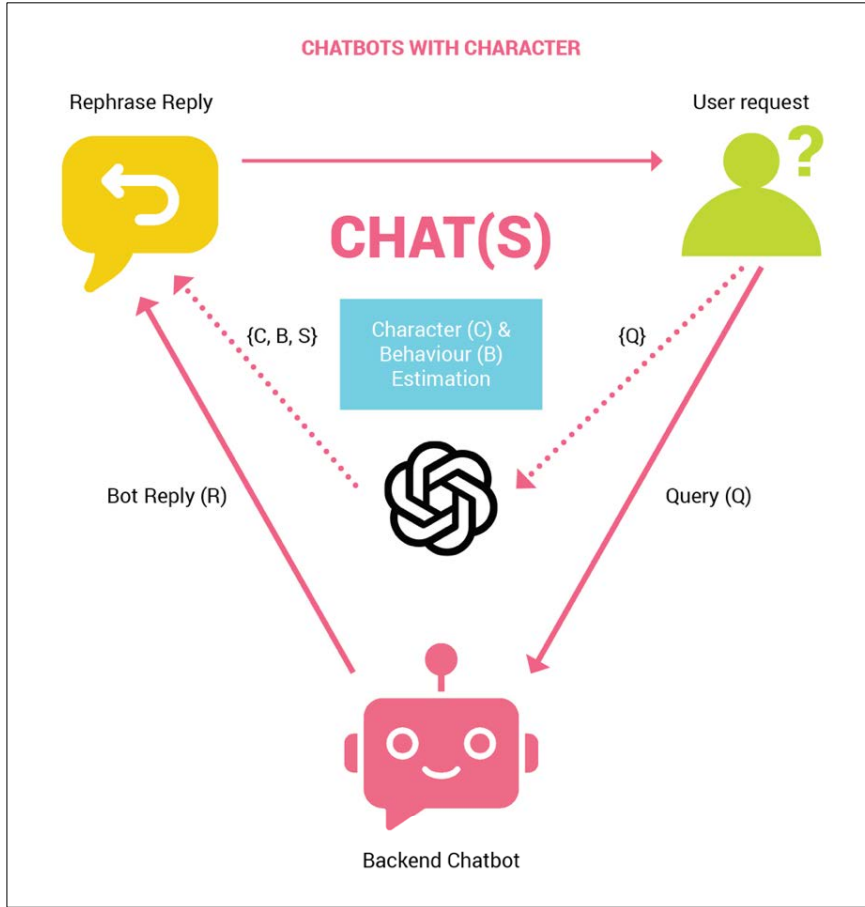
3. *Original text: "I don't trust doctors. They always try to push unnecessary treatments."*

Character vector C: {"neuroticism": 0.4, "extraversion": 0.6, "openness": 0.5, "agreeableness": 0.7, "conscientiousness": 0.6}

Rephrased text: "I've had some experiences that have made me cautious about medical advice. Could you help me understand the reasoning behind recommended treatments to build my confidence in the care I'm receiving?"

Remember to analyze each text objectively and provide rephrased versions that align with the given character vector and push towards the neutralized character target, while maintaining the essence of the original message.

Figure 36: Concept of the experiment framework implemented in the web application for evaluating the FCCC framework



Source: According to Hundertmark et al., (2025).

By randomizing the activation of rephrasing, the experiment can systematically compare interactions where the system adapts its style versus a control group receiving unmodified replies. During the experiment, the following key steps are maintained: Query (Q) analysis and character computation (C,B). The user's query is received by both the plain backend chatbot (BCB) and the character estimation bot (CEB): CEB uses the approach of multi-dimensional fuzzy classification (Big Five membership). Each trait is assigned a membership in $[0,1]$, according to the equation for $C(m)$, forming a vector:

$$C(m) = (\mu_{\{Neuro\}}, \mu_{\{Extra\}}, \mu_{\{Open\}}, \mu_{\{Agree\}}, \mu_{\{Consc\}})$$

Then, a rolling-average or summation-based formula incorporates the last few user inputs to produce the behavior vector B. The conversation history is stored in S.

Raw Reply (R):

The BCB, which can be a ChatGPT instance or an own model, generates a reply for the given user query, not yet shaped by the user's current emotional or personality traits. Rephrasing with (C,B,S), the raw reply (R) and the user's traits (C) and behavior

(B) then go to the RRB. This RRB reads how “dominant” each Big Five membership is. It also checks the conversation context (S) for additional cues (past user statements, potential frustration, or repeated confusion) rewrites (R) to better align with a calmer, more empathetic, or more energetic style — depending on your target scenario.

Final Response:

Only the rephrased reply is shown in the chat interface, meaning the user experiences a single consistent conversation that seamlessly adjusts its style. Disabling rephrasing allows to control an experiment, observing how users behave absent the adaptive restyling. Each user message obtains a vector of membership degrees $\mu_j(m)$ for the Big Five. Using the rolling-average formula the system tracks short-term changes in personality traits across the user’s last n messages. The situation S summarizes the conversation so far. S is used both by the RRB to maintain contextual coherence and by the CEB to see how user traits evolve or how new messages may differ drastically from older ones. Instead of a single scalar sentiment, the Big Five membership vector related to multidimensional fuzzy classifications supports a more refined control over reply tone (e.g., if “openness” is high while “neuroticism” is also rising, the RRB may adopt reassuring yet exploratory phrases). Real users express different traits simultaneously. Multi-dimensional fuzzy logic ensures that the system can handle overlapping or partial memberships, which is more realistic for short chat messages. Additionally because of the fact that each new user message updates the character vector C and the behavior B , the experiment can measure how well the system stabilizes or shifts the conversation tone by systematically rephrasing replies.

Experimental Environment:

Technically, the environment is a web application that orchestrates the following: A proprietary chat interface that hides the complexities from end users, capturing unbiased feedback. Three ChatGPT instances for BCB, CEB, and RRB. Random toggling of the RRB step to form a control group. As the conversation progresses, the system logs user queries, raw and rephrased replies, and the membership vectors (C) as well as the aggregated (B). This log enables quantitative evaluation of how different rephrasing strategies influence user satisfaction, perceived empathy, or overall conversation quality. The code can be seen online: <https://github.com/Sophie31051990/char.bot.demo.git>, commit 5652723.

While detailed quantitative outcomes are presented in the next section, preliminary observations and findings during the technical setup reveal several key insights. The rephrased replies immediately influenced the tone, aligning more consistently with the user’s short-term character states. Because the shared variable S operates across both the BCB and RRB components, the rephrased messages preserved the original context and essential meaning while only modifying the stylistic layer. Furthermore, early user feedback indicated an enhanced conversational experience,

marked by smoother interactions, fewer abrupt transitions, and reduced misunderstanding, particularly in situations involving negative or anxious emotional states.

Stability of LLM-based Character Estimation (Non-Determinism Analysis)

As LLMs are inherently probabilistic, it is necessary to verify that the character estimation results are reproducible across repeated calls. To address this, a separated stability analysis was conducted prior to the main experiment. Ten representative user messages from the experimental corpus were each submitted to the CEB (GPT-4o, temperature = 0.0) on ten independent occasions under identical conditions. The standard deviation (SD) of the five membership degrees across the ten repetitions was computed for each message. The stability analysis demonstrates that the character estimation is highly reproducible under the conditions used in this study. This confirms that the use of temperature = 0.0 effectively suppresses LLM output variance for this task, and that the experimental results can be considered deterministic and reproducible.

5.1.4. Evaluation and Key findings of the Technical Implantation

To evaluate the technical implementation, an experimental study with the demonstration chatbot involving 14 participants recruited via an open online call was conducted. One additional invited participant did not complete all four dialogue sessions and was therefore excluded, resulting in a final sample of $n = 14$.

Table 28 provides a summary of the participant characteristics.

For full transparency, all prompts, code, and evaluation data can be found in Appendix G.

Table 28: Participant demographics, experiment on technical implementation of the FCCC framework ($n=14$).

Characteristic	Value
n (final sample)	14 (15 invited; 1 excluded due to incomplete sessions)
Age: range	25–49 years
Age: M	33.6 years
Gender	7 female (50%), 7 male (50%)
Professional background	4 IT/Software Engineering; 3 Finance/Banking; 3 Research/Academia; 2 Healthcare; 2 Other
Experience with AI assistants (1–5 Likert)	M = 3.4; scale: 1 = never used, 5 = daily use

Source: Data collected by the author.

The participants' age range (25–49) and mixed professional backgrounds reflect the target population for enterprise conversational AI deployments. The relatively high mean experience with AI assistants ($M = 3.4/5$) is consistent with the recruitment through professional and academic networks. Following guidelines from Nielsen et al. (1994), this sample size establishes a robust foundation for exploring user interactions while still allowing nuanced qualitative analysis.

Each participant interacted with the plain backend chatbot (BCB), receiving identical instruction sets. However, participants were randomly assigned an additional condition that enabled reply rephrasing via the RRB. Consequently, each participant experienced four distinct dialogues with the system. Throughout the sessions, the conversational context (S), the multi-dimensional character vector (C) and the short-term behaviour vector (B) were tracked. Consequently, each participant experienced four distinct dialogues with the system. The adaptive dialogue experience offered by the prototype served as a live demonstration in line with the demo-driven research approach (Portmann et al. 2024) and the FCCC principles. Observing real-time system adaptation allowed participants and researchers to reflect on the practical viability of the underlying model which is also a core tenet of demo-driven research.

Throughout these chat-sessions, the conversational context (S), the multi-dimensional character vector (C) estimated by the character estimation bot (CEB), and the short-term behavior (B) were tracked. The Big Five personality traits, neuroticism, extraversion, openness, agreeableness, and conscientiousness, were adopted as the multi-dimensional fuzzy classification. By converting these traits into numeric vectors, changes and trends in user character were captured more precisely across each dialogue turn, consistent with the fuzzy logic approach outlined earlier. After each user input the CEB produced a character vector from the input message (m) according to the following equation:

$$C(m) = (\mu_{\{Neuro\}(m)}, \mu_{\{Extra\}(m)}, \mu_{\{Open\}(m)}, \mu_{\{Agree\}(m)}, \mu_{\{Cons\}(m)})^{\{\tau\}}$$

The system then updated the behaviour trend Bt by averaging the current and the $n - 1$ preceding character vectors according to defined equation of B ($n = 3$ in this study). User queries were simultaneously processed by the BCB, which generated a raw reply; this reply was optionally rephrased by the RRB so that its tone aligned with both $C(mt)$ and Bt. This multi-bot architecture enabled clean comparisons of adaptive versus non-adaptive behaviour.

Study Design and Data Collection:

The chat histories and estimated Big Five vectors were recorded for each of the four dialogues per participant. For each run, I qualitatively analyzed how the user's trait memberships evolved. Illustration from appendix G show, for instance, the mean and 95% confidence intervals of the behavior and character vectors over time.

Although dialogue lengths varied these visual analyses offered crucial insight into the patterns and volatility of user character changes.

Qualitative Insights and Questionnaire Feedback:

The numerical results were complemented with a post-study questionnaire and in-depth expert interviews. Psychologists helped interpret user personality drift, especially whether the system's character adaptation improved emotional wellbeing or engagement. This triangulation (visual graphs, domain-expert interviews, and user questionnaires) yielded a richly layered set of findings. The dynamic character and behavior parameters showed that each participant's multidimensional membership values (C) adjusted immediately after every input. In contrast, behavior (B) evolved more gradually, reflecting cumulative short-term trends rather than sudden changes. This pattern supports the original FCCC concept, where fuzzy logic enables nuanced, turn-by-turn adaptation to user traits, while the rolling average in B ensures consistency across multiple interactions.

The diagrams provided in appendix G document the distinctly different conversational paths that emerge when the reply rephrasing bot (RRB) modifies the BCB's raw reply to align with user traits. The results demonstrate how adaptiveness is introduced into the conversation and how the system overcomes the limitations of crisp logic through multi-dimensional fuzzy classification. Notably, the rephrased replies were consistently perceived as "more empathetic" or "more encouraging".

Expert interviews, including two clinical psychologists and one behavioral science researcher, compared normal interactions (BCB only) with the enhanced, character-adaptive interactions (BCB+RRB). The interviews were conducted separately face-to-face in an unstructured questioning setting. Both experts concluded that the adaptive version can significantly foster positive emotional changes, as it tailors tone to user states. It is crucial, however, to attribute this effect to a combination of two distinct factors. Firstly, the dynamic adaptation of the conversational style based on the C and B vectors calculated by the FCCC framework, which tailored the tone to the user's state. Secondly, an explicit instruction within the RRB's prompt, which directed it "to foster positive emotional changes" among its users. The observed improvement is therefore not solely an emergent property of the FCCC framework itself, but the result of a deliberate coupling of adaptive style-tuning with a predefined, benevolent communication strategy. In table 24 it is shown how each Big Five trait shifted across the dialogues, highlighting increases in perceived extroversion or agreeableness under certain interventions. Since the RRB Bot is enhanced with an additional prompt instruction which instructs the bot to foster positive emotional changes among its users, while adapting its replies. Ultimately, these results demonstrate that all participants could engage with the chatbot and receive relevant, empathetic replies, aligned with each individual's perceived user traits and conversation trends. Furthermore, the multi-dimensional fuzzy classification of each message and the subsequent

short-term behavior vector (B) effectively guided the rephrasing process. This not only addresses the research question confirming that the fuzzy logic-based FCCC framework, is technically feasible, but also suggests tangible emotional benefits for end users. In sum, the experiment indicates strong potential for multi-dimensional fuzzy classification and the character computing triad (C,B,S) to enrich conversational AI. By capturing user traits dynamically and informing response rephrasing through separate bots (BCB, RRB, CEB), the system achieved smoother, more empathetic dialogue flows. The approach promotes deeper engagement and can positively shape user sentiment. The code base¹, available online, underscores reproducibility and invites further experimentation to refine or expand trait definitions beyond the Big Five. Table 29 summarizes the qualitative effects observed on the recorded character traits when comparing the plain chatbot to the character-computing- enabled version. These interpretations were derived by three distinct data sources: First the quantitative trends observed in the recorded C and B vectors over time as illustrated in appendix G, second the subjective ratings from the post-study questionnaire where participants assessed factors like trust and empathy, and third the qualitative analysis of the expert conversations by consulting psychologists.

Table 29: *Observed effects of recorded character traits among all conversations with plain chatbot and character computing: Agreeableness (A), Conscientiousness (C), Extraversion (E), Neuroticism (N) and Openness (O).*

Trait	Plain Chatbot	Character Computing
A	Generally lower: limited ability for effective influence.	Reducing over course of conversation: balance between harmony and assertiveness being necessary for encouraging behavioural change.
C	Higher volatility: lacks in trust and reliability.	Constant on neutral level: supports trust and reliability; avoids overly controlling behaviour.
E	No dominating behaviour identified.	No dominating behaviour identified.
N	Higher level: more stressful conversation.	Constant low level: indicates stress-free (calm) and supportive interaction.

¹ char.bot.demo git repo, <https://github.com/Sophie31051990/char.bot.demo.git>, commit 5652723

Trait	Plain Chatbot	Character Computing
O	Significantly lower: lacks creativity.	Significantly higher (being creative): chatbot is able to foster positive emotional changes.

Source: According to Hundertmark et al., (2025).

5.1.5. Summarizing Remarks on the Technical Implantation of the Artefact

The demonstrated technical implementation serves as the initial phase of the evaluation of the artefact. Despite the relatively small sample size in the presented experiment, I was able to show that the artefact can be technically integrated into a conversational AI system, here a text-based chatbot, and is even capable of improving the overall mood of its users.

It should be noted that the technical implementation primarily focused on the core components of the artefact. These components represent the highest requirements while simultaneously addressing the most critical aspects of this study.

In sum, my presented technical implementation provides an answer to RQ4.1 which, at the same time, serves as an important part of the evaluation phase in the DSR:

RQ4.1: How can the principles of Fuzzy Logic and Fuzzy Conversational Character Computing be technically operationalized to process ambiguous user inputs, and dynamically adapt to personality traits and emotional states?

By transforming the inner part of my conceptual framework into a mathematical formula and implementing it in a three-component chatbot architecture (BCB, CEB, RRB) using GPT-4o, I provide an answer to RQ4.1. The inner circle of the FCCC framework was prioritized as the most critical component for initial evaluation, given its direct relevance to the core research contribution.

5.2. Necessary Organizational Capabilities to Implement the Artefact

This section addresses RQ4.2, contributing further to the evaluation of the artefact within the framework of DSR.

RQ4: How can the conceptual framework be evaluated within the Design Science Research process to ensure its theoretical robustness, practical feasibility, real-world applicability and societal impact?

RQ4.2 What organizational capabilities must be in place to conceptually and technically implement the framework, and how realistic is it for organizations to meet these prerequisites?

My approach initially draws on the capabilities required for implementing AI systems, as identified by Weber et al. (2023) in their literature review. The requirements outlined by Weber et al. (2023) serve as a foundational reference.

To answer RQ4.2, I first evaluate which of these criteria are relevant to the implementation of the proposed framework. Subsequently, I discuss the feasibility of these capabilities being present in companies under current conditions. This evaluation employs a two-stage process involving expert groups.

The first group comprises academics and practitioners specializing in computer science and AI. Their focus is on assessing the importance of individual capabilities concerning the conceptual framework. Each capability is rated on a scale from 1 (unimportant) to 5 (very important). This assessment was conducted during a two-hour online focus group via Microsoft Teams at the beginning of 2025.

The second expert group consists of decision-makers and team leaders from various companies across different industries and sizes. Their discussions aim to determine the practicality and current availability of these capabilities within organizational contexts.

The outcomes of these evaluations will guide the refinement of the framework, ensuring its alignment with both theoretical importance and practical feasibility. The structured input from these diverse expert groups provides a robust foundation for integrating AI capabilities into organizational practices.

5.2.1. Evaluation of Organizational Capabilities

To evaluate the organizational prerequisites and the likelihood of implementing the proposed framework within companies, it is essential to first define the key capabilities. The existing literature offers valuable insights into general capabilities required for the adoption of AI systems. This thesis draws on the capabilities identified Weber et al. (2023), who conducted an extensive literature review on the subject and deepened its evaluation.

However, it remains unclear to what extent the requirements for general AI integrations can be directly applied to the specific framework under consideration. To bridge this knowledge gap, two focus group interviews were conducted.

The first online focus group was conducted in January 2025 via Microsoft Teams and lasted approximately two hours. It comprised 15 participants with backgrounds in artificial intelligence, natural language processing, fuzzy systems, and conversational AI. Table 30 lists the participant profiles.

Table 30: Focus Group 1 participant profiles (n=15), January 2025

Role	Sector	Expertise focus	Type
Professor	University	Fuzzy Logic, Business Informatics	Research
Senior Research Associate	University	Conversational AI, NLP	Research
Senior Research Associate	University	Human-Computer Interaction	Research
PhD Candidate	University	Conversational Systems, LLMs	Research
Professor	ZHAW Zurich	Applied Machine Learning	Research
Head of Data Science	Swisscom AG	AI Platform, NLP	Industry
Senior AI Engineer	UBS Group AG	Conversational AI, Chatbots	Industry
AI Product Manager	Helvetia Insurance	AI Strategy, Prototyping	Industry
Machine Learning Engineer	Viseca AG	NLP, Recommendation Systems	Industry
Head of AI Innovation	Zurich Insurance Group	AI Governance, Deployment	Industry
CTO	AI Startup (< 50 empl.)	Conversational AI, Full-Stack	Industry
Senior Consultant	McKinsey & Co. Zürich	AI Strategy, Digital Transformation	Industry
AI Architect	PostFinance AG	Cloud AI, Data Pipelines	Industry
Senior Research Associate	University	Ethics & AI, Responsible AI	Research
CEO	AI Startup (< 50 empl.)	Conversational AI	Industry

Source: Data collected by the author.

During the expert discussion, the FCCC framework was conceptually introduced and explained through a case example. Subsequently, the capabilities listed in Table 31 were discussed in detail. The discussion was structured into three thematic areas: AI-related skills, IT infrastructure requirements, and organizational structure and culture.

Table 31: Organization Capabilities for the Artefact

Category	Resource	Description
AI-Related Skills	Technical AI skills	Technical skills to develop, deploy and operate AI systems.
	Domain AI skills	Domain skills to support use case selection, AI systems evaluation and business translation.
	Workforce AI skills	Skills to work with and maintain AI systems in productive use.
IT Infrastructure	Data	Data in sufficient quantity and quality for AI development.
	AI-specific infrastructure	Availability of AI-specific infrastructure (e.g. tools, frameworks, AI engines).
	IT infrastructure	Compatibility of existing IT infrastructure with AI systems.
Organizational Structure and Culture	AI–Business relationship	The relationship between business and AI departments.
	Sourcing relationship	The relationship with external AI solutions and service providers.
	Culture	Collaborative, experimental and data-driven culture.

Source: According to Weber et al. (2023).

The second focus group, conducted online in February 2025, comprised 12 participants drawn from department and team leadership positions across organizations of varying sizes and industries. Industry differences had minimal impact on perspectives, organization size emerged as the primary determinant of capability availability. Table 32 lists the participant profiles.

Table 32: Focus Group 2 participant profiles (n=12), February 2025

Function	Industry	Org. size (employees)	Org. type
Head of Customer Service	Retail	> 10,000	Large enterprise
Director Digital Transformation	Pharmaceutical	> 100,000	Large enterprise
Head of AI	Finance	1,000–10,000	Large enterprise
VP Operations	Retail / Logistics	> 50,000	Large enterprise
Head of HR	Transport	10,000–50,000	Large enterprise
Team Lead IT	Finance	1,000–10,000	Medium enterprise
Product Manager	Industrial / B2B	1,000–10,000	Medium enterprise
IT Manager	Public	100–500	Medium enterprise
Team Lead Software Development	Healthcare	1,000–5,000	Medium enterprise
CEO	Digital Consulting	10–50	Startup
CTO	InsurTech	20–50	Startup
Managing Director	Retail / SME	< 100	Traditional company

Source: Data collected by the author.

Key Findings on Organizational Capabilities

Here I present the key findings from the first focus group, divided into three main areas: AI-related skills, IT infrastructure requirements and organizational structure and culture.

(1) AI-Related Skills

The focus group revealed that the implementation of the FCCC framework does not necessarily require deep AI expertise within the company during the initial stages. Experts agreed that the framework itself incorporates critical AI functionalities,

minimizing the immediate need for specialized skills. However, advanced AI knowledge becomes crucial for ongoing operations or further development of the framework. Specifically, companies aiming to enhance the framework or conduct in-depth analyses must invest in developing these skills. A higher priority, rated 5 out of 5 by participants, was the understanding of suitable AI use cases. This encompasses both internal considerations and broader knowledge of AI's capabilities, limitations and appropriate application areas.

(2) IT Infrastructure

The discussion highlighted that the framework's self-adaptive conversational AI component alleviates the need for comprehensive internal data management at the outset. Unlike many AI applications that rely on robust data structures, the FCCC framework places a secondary emphasis on such requirements. Nonetheless, the framework's integration hinges on an IT infrastructure capable of supporting its deployment. While the framework can operate in a public cloud, compliance with data protection requirements is imperative.

(3) Organizational Structure and Culture

All participants unanimously emphasized the necessity of an open organizational culture for successful implementation. Such openness primarily pertains to embracing new technologies and methodologies. In many cases, fostering collaboration between business-oriented and IT departments is critical. Conversely, external stakeholder engagement was deemed less significant for the framework's implementation.

Table 33 presents the capability ratings derived from Focus Group 1 (n=15), including descriptive statistics computed from the individual ratings. The importance scale ranges from 1 (unimportant) to 5 (very important).

Table 33: List of Organization Capabilities for the Artefact

Category	Resource	M	Min	Max
AI-Related Skills	Technical AI skills	3.0	2	4
	Domain AI skills	4.0	3	5
	Workforce AI skills	4.0	3	5
IT Infrastructure	Data	3.0	2	4
	AI-specific infrastructure	4.0	3	5
	IT infrastructure	3.0	2	4
Org. Structure & Culture	AI-Business relationship	4.0	3	5

Category	Resource	M	Min	Max
	Sourcing relationship	3.0	2	4
	Culture	5.0	5	5

Source: Data collected by the author.

5.2.2. Achieving the Capabilities of the Artefact

Following the first focus group, where I identified the key capabilities necessary for implementing the framework, the second focus group focused on evaluating the likelihood that organizations can develop or access these capabilities to effectively apply the conceptual framework or developed artefact in practice.

A diverse group of participants was intentionally selected, comprising department and team leaders from organizations of varying sizes and industries. While industry differences had minimal impact on the experts' perspectives, the size of the organization emerged as a critical determinant in assessing whether the required capabilities are currently available, planned for the future or entirely absent.

Key Findings on Achieving Capabilities of the Artefact

The primary finding regarding the probability of achieving the capabilities necessary for implementing my artefact is that the results vary significantly depending on company size. Consequently, what follows below is an unstratified analysis categorized by company size.

(1) Large Organizations (500+ Employees)

Large organizations generally already possess the required capabilities or have plans to develop them internally. These companies demonstrated readiness for framework implementation, indicating that capability gaps, if any, were minimal and manageable.

(2) Medium-Sized Organizations (100–500 Employees)

Medium-sized organizations often have access to some, but not all, of the necessary capabilities. Notably, no clear trend emerged regarding what specific capabilities are most commonly available in this segment. It became evident during discussions that many medium-sized firms are unlikely to acquire all required capabilities in the near future. However, the focus group participants suggested that external partnerships could effectively address these gaps. This approach was identified as a proven method to compensate for critical capability deficits.

(3) Small Organizations (<100 Employees) and Startups

Smaller firms, particularly those with fewer than 100 employees or newly founded startups with teams of two to ten members, typically lack most of the required

capabilities. Interestingly, startups exhibited greater adaptability, and either already possessed some capabilities or were planning to develop them. In contrast, traditional small businesses with long market presence were found to meet only a few capability requirements. Experts questioned whether these organizations could bridge such gaps in the foreseeable future.

This group poses significant challenges for implementing the artefact. While external resources could theoretically address these capability deficiencies, this requires a culturally open mindset to adopt new and innovative approaches. According to the experts, such cultural openness often represents the most significant hurdle for traditional small businesses.

5.2.3. Organizational Capabilities Required to Implement my Framework in a nutshell

Chapter 5.2. served as the second stage of the artefact evaluation within the framework of DSR. Simultaneously, it allows me to address RQ4.2:

RQ4.2: What organizational capabilities must be in place to conceptually and technically implement the framework, and how realistic is it for organizations to meet these prerequisites?

The discussion on organizational capabilities required to implement the FCCC framework underscores its strong adaptability and practical applicability across diverse organizational contexts. The framework is designed to minimize initial demands on specialized AI knowledge, allowing organizations to begin implementation with a foundational understanding of AI use cases. This ensures that the framework is accessible even to companies without extensive prior expertise in AI.

The emphasis on IT infrastructure highlights the flexibility of the framework, which can operate effectively in cloud environments and adapt to varying data management systems. These characteristics make it suitable for a wide range of organizational setups while ensuring compliance with critical data protection regulations.

Culturally, the framework fosters an open and collaborative environment, bridging the gap between business and IT departments. This cultural alignment is both a challenge and an opportunity, as it supports organizations in creating a future-ready mindset that embraces innovation and transformation.

The segmented readiness across different organizational sizes further confirms the framework's potential for broad adoption. Large enterprises are well-positioned for immediate implementation, while medium-sized firms can effectively address capability gaps through partnerships. Even smaller organizations, with the right external support and cultural openness, can successfully integrate the framework into their operations.

Overall, these findings confirm that the FCCC framework is not only feasible to implement, but also scalable and adaptable to various organizational contexts. With tailored strategies and a commitment to building the necessary capabilities, the framework has the potential to drive meaningful innovation and operational excellence across industries. In other words, my findings allow me to affirmatively address RQ4.2, demonstrating that organizations are already capable of implementing the proposed framework effectively.

5.3. Social Impact of the Artefact in the Context of Life Engineering

This section serves as last part of my three-step evaluation process according to the DSR approach.

Here, I address RQ4.3, contributing further to the evaluation of the artefact within the framework of DSR.

RQ4: How can the conceptual framework be evaluated within the Design Science Research process to ensure its theoretical robustness, practical feasibility, real-world applicability and societal impact?

RQ4.3: Does the framework align with the principles of Life Engineering, and how well does it meet ethical expectations to deliver the desired societal impact?

In chapter 3 of this thesis, I presented chatbots and conversational AI systems as a valuable tool for promoting financial literacy. The weaknesses of previous conversational AI systems and rule-based chatbots were clearly identified. On the one hand, rule-based bots lack individual adaptation to the needs of the user. On the other hand, the scope of content is usually very limited, and users would like more comprehensive coverage of various topics. In addition, rule-based chatbots often lack consideration of ethical guidelines, as business goals are often in the foreground rather than prioritizing the well-being of individual users.

The artefact developed over the course of this study can be used, among other things, to optimize conversational AI systems to strengthen financial literacy. In relation to life engineering, this artefact can be seen as a significant building block on the way to developing so-called life agents, as introduced by Österle (2020). Österle (2020) describes the task of life engineering as developing rules for digital assistants that are focused on human well-being. It is argued that the minimum goal is to translate adequate knowledge about information technology and quality of life into recommendations for individuals, companies and politicians.

Life engineering researchers are increasingly concerned with the design of conversational agents and HCI. Often, the term “life assistant” is used to describe conversational agents designed to improve people’s lives. The framework developed here could serve as a basis or orientation for other life assistants in the future.

The immediate goal of life engineering is, therefore, to develop digital assistants that serve human interests. The long-term goal, which is envisaged over a time horizon of 50–100 years, is the use of superintelligence for the benefit of humans (Österle, 2020).

5.3.1. Application of the Artefact as Life Agents in the Context of Life Engineering

In the following, my artefact, in this case the new FCCC framework, for conversational AI systems is applied in the context of chatbots that can be used as life agents, with a focus on increasing the financial literacy of each individual user.

The implementation of the system starts with the outer layer, the ethical framework. In order to apply the developed conversational AI system in the context of life engineering, consideration of ethical guidelines is essential. Although the framework does not currently detail the specific substantive ethical conditions, approaches by researchers such as Hundertmark et al. (2022c) in the study “Ethics Check for Conversational AIs” can still be used. According to this work, factors such as transparency and consideration of the individual user’s needs are crucial. The conversational AI system should be extremely transparent in generating answers and decisions. In addition, it should offer the user the possibility to select topics himself, and these user decisions should also be taken into account in the subsequent dialogue. In concrete terms, this means that the conversational AI system should state exactly what sources it bases its knowledge of financial literacy on and leave the user free to decide in which areas of financial literacy they would like further support.

The second layer to be considered is the legal and regulatory framework. Here, too, the framework provides little information at this stage. It therefore requires further research in the legal context. In general, the protection of personal user data must certainly be preserved. However, to strengthen financial literacy, further legal clarification is absolutely necessary before the publication of the conversational AI system.

1. The innermost layer of the three outer layers deals with economic framework conditions. In principle, this refers to every influencing factor of the company that the conversational AI system makes available to its users. In the context of life agents, however, these kinds of economic considerations should rather take a back seat. Instead, the well-being of the individual user should be prioritized. This layer can therefore be neglected in the current application.
2. Next come the components around the core of the framework, or everything that is directly connected to the conversational AI system. First of all, the existence of pre-training is relevant. Conversational AI starts with a foundation of “basic knowledge” and pre-trained knowledge. This knowledge relates to the methods around conversational loops, CC and LLMs. In further steps, the conversational AI system should be able to find out more about the user in the

context of dialogues and thus determine the triad of CC. To do this, it needs basic knowledge about conducting dialogues and calculating the triad.

3. The fifth step includes all parts of the conversational loops, in which both CC and multidimensional fuzzy classifications are applied. Considering the previously described layers and the pre-training, the conversational AI now initiates the dialogue with the user, with the main goal of first learning more about the character and the needs of each individual user in order to strengthen the financial education of the individual in further dialogues as a so-called life agent.
4. A transparent greeting plays a significant role in this, in which conversational AI presents its own goals, namely, to increase the financial literacy of the individual user, as well as the way in which the answer is generated, including sources and technologies used. This first step is of great importance. The further dialogues serve to identify the triad in CC. This triad is essential for the conversational AI system to adapt its own design and tone to the wishes and expectations of each individual user. Further dialogues will reveal more about the user's situation, emotional state and general personality. Since these factors are usually not characterized by sharp values but rather by fuzzy affiliations, the conversational AI system uses the methods of multidimensional fuzzy classifications at this point. Fuzzy rules make it possible to calculate with fuzzy affiliations. The fact that several affiliations form a character leads to the application of multidimensional fuzzy classifications.
5. Normally, the loops described here are a continuous part of the conversation with the user. It is also possible that the user's character and desires change over the course of the conversation and evolve with greater generalization. The conversational AI system is therefore required to continue the loops and continuously determine the triad. Accurately determining the triad and the resulting insights are crucial for the conversational AI to know how to adjust its design and tonality for the benefit of the user and ensure a high user experience overall. Of course, in the context of financial education, a high learning curve is also paramount. Only when the conversational AI system can accurately determine the character of each user will it know how to adjust its design and tonality accordingly for the benefit of the user.
6. In the sixth step, the conversational AI system uses the methods of LLMs, whereby these are deliberately used only to a limited extent or sparingly. So far, LLMs have required considerable computing power and have not been very data-saving. Both aspects should rather be avoided in the context of life engineering, which is why the application of LLMs is restrained. Nevertheless, LLMs achieve remarkably good results when it comes to extracting concrete

knowledge from databases such as PDFs, websites or other knowledge sources (see Point 7 below) and making it available to the user in natural language.

Consequently, the conversational AI system uses LLMs to answer the user's questions on financial topics. Furthermore, these knowledge sources are also used to proactively educate the individual user on important financial topics. In contrast to the original rule-based bots, where the topics and dialogues had to be determined in advance, it is now sufficient to connect to extensive knowledge databases on financial knowledge. This knowledge forms the basis for the content design of the dialogues. At this point, the external layers described above must not be neglected under any circumstances. Transparency and consideration of the interests of the individual user are essential here.

7. In seventh place are the knowledge sources, which can typically be websites, PDFs, books or other knowledge documents. The framework does not answer the question of who owns these knowledge sources and how they are generated. Usually, the knowledge sources come from the company that operates the conversational AI system or the entire framework. In the context of life agents to strengthen financial literacy, it is important to mention that the knowledge sources must comply with the ethical framework and may only be used for the benefit of the user. Data from companies that may be pursuing their own (egoistic) goals without pursuing the well-being of the user may not be used here.

Table 34 shows the previously described seven steps for the application of the FCCC framework, which is shown in figure 37. The numbers are highlighted on the artefact.

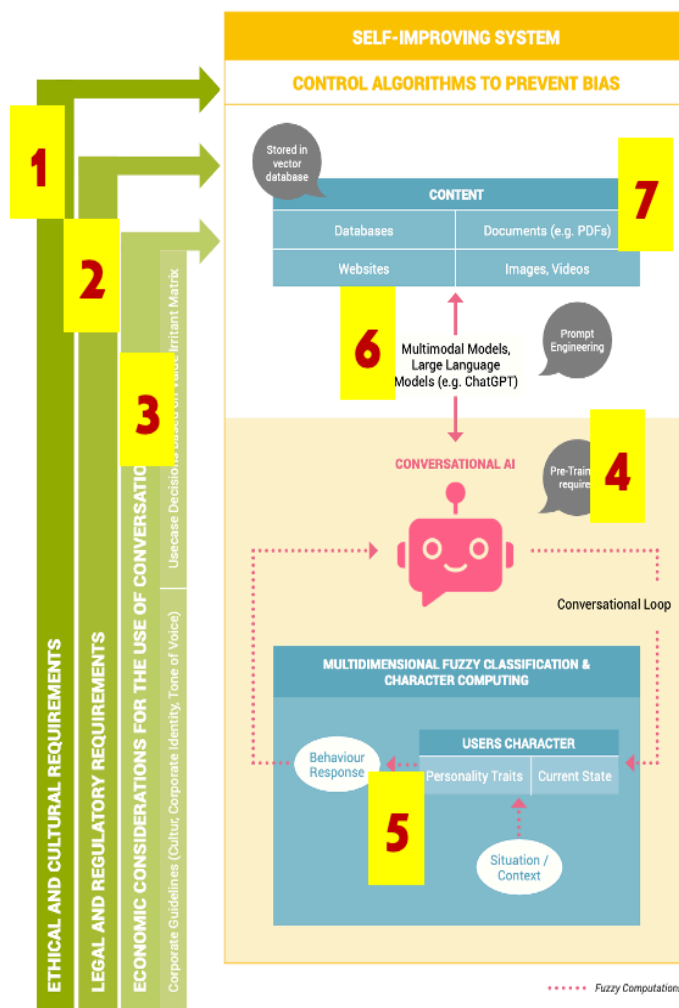
Table 34: 7 steps from FCCC framework, RQ4.3

Number	Brief description	Explanation
1	Ethical framework	Transparency with regard to sources used and knowledge applied. Free decision-making options for users on the individual topics.
2	Legal framework	Requires further clarification with lawyers.
3	Economic considerations	Can be neglected.
4	Pre-training	The pre-training mainly includes knowledge on the application of Conversational Theory, CC and LLMs.

Number	Brief description	Explanation
5	Conversational Loops: Multidimensional Fuzzy Classification and CC	Determination of the user character in order to be able to adapt the design and tonality to the individual user.
6	LLMs	LLMs are used to access the knowledge sources from Point 7 to pass on the knowledge to the user in natural language.
7	Content	Different sources of knowledge that correspond to the ethical framework are integrated into the system.

Source: Data collected by the author.

Figure 37: FCCC Framework, 7 steps highlighted



Source: Data collected by the author.

The following overview in table 35 shows a possible dialogue between conversational AI, which is supposed to strengthen financial education, and a user who wants to focus on sustainable investment. The second column, “bot”, shows the chat message of the conversational AI (here: chatbot). The third column, “user”, shows the user’s text message. The fourth column, “Remarks”, explains the message in the context of the framework.

Table 35: Example Dialogue FCCC Framework for Life Agents, RQ4.3

Dialog step	Bot	User	Remark
1) Welcome and introduction of the chatbot	<i>How do you do? I am Emy. As an AI-based chatbot, I have the task of increasing the financial knowledge of my conversation partners. My goal is to strengthen social welfare by ensuring that more and more people have a solid basic financial education. I use character computing methods so that I can adapt my tone and content to suit you as much as possible. I also use the methods of Large Language Models in some cases. The dialogs are stored anonymously and used for scientific purposes. Now the question, how well do you know about financial topics?</i>	-	During the greeting, a special focus is placed on the ethical framework conditions. These require maximum transparency with regard to the bot’s objectives and methods. Legal conditions with regard to data protection are also taken into account. At the end, the bot asks a question to start the conversational loop and find out more about the user.
2	-	<i>Hello Emy. Unfortunately, my knowledge of finance is rather limited. But I would be happy if you could tell me more about the topic of “sustainable investing”.</i>	The user greets the chatbot by name, uses filler words such as “unfortunately” and states their specific need — to find out more about sustainable issues. → These are important features that conversational AI can use to calculate the triad of character computing.

Dialog step	Bot	User	Remark
3) Disclosure of knowledge sources; Request for initial information for CC Triad	<i>Of course, I have a few sources of knowledge on sustainable issues, mostly from NGOs that deal with the sustainability of “green investments”. Before we start, I would like to get to know you better so that I can then share this knowledge with you on an equal footing. Why are you interested in the topic of sustainability right now?</i>	-	The bot takes ethical principles into account by showing where it obtains its knowledge from (transparency). The bot also asks the user for further information as part of the conversational loop and to calculate the triad. In this case, the bot tries to find out more about the user’s situation.
4	-	<i>I have been interested in sustainability issues for a long time and now I would like to find out more about sustainable investing. There is no specific reason and I am in no hurry to invest money directly in a sustainable way.</i>	The user shows that he is interested in sustainability. An indication of his interests. With regard to the situation, the user indicates that he is not in a hurry and is looking to broaden his horizons or would like to delve deeper into the topic in general. → Both indications are relevant for the triad.
5) Confirm user info; Request further info for CC triad	<i>Okay, so the topic of sustainability motivates you and you’re not stressed about investing money sustainably as quickly as possible. One more question so that I can assess you better: Are you more of a person who learns with examples and metaphors or better with facts and figures?</i>	-	The bot summarizes the user’s situation in order to have it confirmed again. This is a common procedure in the context of conversational theory, where the conversation partners come closer and closer to each other over the course of the conversation and seek consensus. The bot then asks for further information

Dialog step	Bot	User	Remark
			that will be decisive for the triad.
6	-	<i>Yes, exactly, correctly summarized. I prefer examples.</i>	The user confirms the bot's summary. He reveals further details of his personality. → The Conversational Theory mechanism works.
7) Topic suggested by the bot to obtain the user's consent	<i>I'd say I'll start by giving you some basic information about impact investing. Then you'll have a good basis and can decide at your leisure whether or when you want to invest money. Do you agree with that?</i>	-	The bot suggests possible content to the user for the first knowledge lesson. It takes into account the fact that the user initially wants general information without specific investment motives. It is part of the ethical framework that the bot adheres to the user's wishes and also asks for consent.
8	-	<i>Yes, that's a very good thing. What exactly is meant by impact investing?</i>	The user gives his or her consent so that the bot can start with the mediation of financial knowledge.
9) Start knowledge transfer with maintenance of the conversational loop	<i>I have a definition for you from the GIIN (Global Impact Investing Network): "Impact investments are investments with the aim of achieving a positive and measurable social and environmental impact in addition to a financial return." Let me give you an example. As an impact investor, you don't simply invest in a company that generates</i>	-	In terms of ethics and transparency, the bot specifies the exact source of its knowledge. From the previously calculated triad, the bot knows that the user wants examples and therefore underpins the pure definition with examples. To maintain the

Dialog step	Bot	User	Remark
	<i>large returns without having an ecological or social impact. This could be a weapons manufacturer, for example. As an impact investor, you prefer to invest in a company that naturally generates a good return but is also committed to complying with ESG criteria. For example, you could invest in a company that develops technologies to purify water. Unclean water is one of the main causes of disease and death in developing countries. Would you like more examples or would you like to find out how such sustainable investments can be found?</i>		conversational loop, the bot asks a question at the end about what the user would like to know next.

Source: Data collected by the author.

5.3.2. Evaluation of the Artefact as Life Agents in the Context of Life Engineering

To validate that the FCCC framework's application in the life-agent context is both comprehensible and credible to researchers in the life engineering domain, three expert interviews were conducted in February 2025. The interviews were conducted online and lasted approximately 60–75 minutes each. At the beginning of each interview, the expert received a brief conceptual introduction to the FCCC framework, followed by a detailed walkthrough of its application as a life agent. The interviews were conducted only without any fix questionnaire.

Table 36 provides an overview of the three expert profiles. All three experts have an established research background in ethics or life engineering or closely adjacent fields, ensuring that the evaluation reflects domain expertise rather than general AI familiarity. The combination of these three profiles ensures triangulation across three complementary dimensions.

Table 36: Overview Interview Partner for ethical evaluation

# of Interview Partner (IP)	Disciplinary background	Interview format
1	Business Informatics, HCI, Life Engineering	Online, 70 min
2	Information Systems, Conversational AI, Ethics of AI	Online, 65 min
3	Business Informatics, Fuzzy Systems, Ethics	Online, 75 min

Source: Data collected by the author.

5.3.3. Key Findings from the Expert Interviews

All three interviews converged on several core themes, with some nuanced differences reflecting each expert's disciplinary emphasis. The findings are organized thematically below.

General Recognition of the FCCC Framework as a Life Engineering Contribution

All three experts independently affirmed that the FCCC framework represents a meaningful and timely contribution to the life engineering research programme. IP 1 noted that the framework's multi-layer architecture, with its explicit separation of ethical, legal, and economic frameworks from the conversational core directly mirrors the design principles advocated by Österle (2020) for life agents. IP 2 emphasized that the framework's explicit ethical transparency requirements (knowledge source disclosure, user consent, data minimization) address a gap that most existing conversational AI systems fail to fill. IP 3 highlighted the theoretical coherence of the framework, noting that the use of fuzzy logic to bridge the gap between the crisp computational substrate and the graded nature of human values is precisely the methodological move that life engineering's vision of "computing with words for human well-being" requires.

Validation of the Importance of each Theory to Life Engineering

In chapter 4 I mapped the FCCC framework's theory to life engineering goals. All three were asked to evaluate both the correctness of the mappings and the completeness of the theoretical justifications. The feedback was consistently positive. IP 1 and 3 confirmed that the mapping accurately reflects the dependency structure between the theoretical pillars, adding that the explicit acknowledgment of each theorist's limitation and how the next theory resolves it gives the framework a conceptual integrity that is often absent in other "AI-for-well-being proposals".

IP 2 suggested one refinement. The mapping should make even more explicit that the "legal + economic + ethical rules" layer in row 2 is not merely a constraint on

the system but a constitutive component of what it means for a life agent to serve human well-being rather than corporate interest.

Importance in the Financial Literacy Application Domain

All three experts independently identified the financial literacy use case as particularly well-chosen for a life engineering demonstration. IP 1 pointed out that financial illiteracy is both a widespread societal problem and a domain where individual adaptation is especially important the "right" level of financial guidance varies dramatically across users depending on their knowledge level, risk profile, and life situation. This makes it an ideal testbed for a framework that claims to personalize at the individual level. IP 2 added that the financial domain is also one where the tension between commercial and user interests is most acute, making the framework's explicit prioritization of user well-being over economic optimization particularly valuable and distinctive.

Concerns Raised: Energy Consumption and LLM Transparency

The most substantive critical feedback across all three interviews concerned two issues. First, all three experts raised the question of energy consumption. They noted that a life engineering system claiming to support human well-being while simultaneously generating a significant carbon footprint creates an ethical tension that should be acknowledged and addressed. Additionally it was noted that the current implementation uses GPT-4o API calls for every user message, and that at scale this represents a non-trivial environmental cost.

Furthermore they raised the issue of LLM transparency and auditability. For a life agent making recommendations on sensitive topics, the opacity of the LLM's reasoning process is a non-trivial concern from both a regulatory (EU AI Act) and a life engineering standpoint. The experts recommended that future iterations explicitly document the model selection rationale, the prompt design principles, and the mechanisms in place to prevent the system from generating advice that exceeds its validated scope.

Assessment of the Framework's Societal Impact Potential

When asked to assess the framework's potential societal impact, all three experts expressed measured optimism. Feedback from IP1 can be summarized as follows: The FCCC framework is, to his knowledge, the first formally specified conversational AI framework to integrate all four of the properties that life engineering demands of a digital assistant adaptability (systems theory), communicative intelligence (conversational loops), semantic sensitivity (fuzzy logic), and psychological awareness (character computing) into a single coherent architecture. This integration is the framework's primary contribution, and it opens a research pathway that has been clearly defined but not yet pursued in the life engineering literature.

IP 2 added a societal dimension. If the framework is deployed responsibly with the ethical and legal layers implemented as substantively as the computational core it has potential to reduce information inequality in decision-making situations.

IP 3 noted that the non-determinism analysis and the post-study questionnaire results together provide a level of empirical grounding that is unusual in life engineering framework proposals, most of which remain at the conceptual level. This empirical grounding is a significant strength and makes the framework as a possible reference architecture in the field.

5.3.4. Summarizing Remarks on the Social Impact of the Artefact in the Context of Life Engineering

This section represents the final part of the evaluation phase according to the DSR methods and answers RQ4.3:

RQ4.3: Does the framework align with the principles of Life Engineering, and how well does it meet ethical expectations to deliver the desired societal impact?

The application of my conceptual framework in the context of financial education described in this chapter and the positive feedback from the expert discussions illustrate the importance and relevance of my framework in the context of life engineering. When applied according to the above explanations, the framework contributes to the development of life agents, as described by Österle (2020). The framework's focus on ethical and legal frameworks makes it extremely attractive to researchers in the field of life engineering. Furthermore, the conversational AI system strives to adapt its design, tone and content to the individual needs of the user. In doing so, it intends to continuously evolve and get closer to the user's needs in every dialogue with the help of conversational theory (Pask, 1965; Pangaro, 2010) and systems theory (von Bertalanffy, 1949). Additionally, it has the capability to compute with words and fuzzy terms, which makes it more natural and humanlike. In contrast to purely commercially oriented conversational AI projects, such as the use cases shown in chapter 3, the focus here is on the individual user and maximizing their quality of life. This emphasis again underlines the anchoring of the framework in the disciplines of life engineering. In sum, this application with its expert discussions proves the evaluation that the FCCC framework is a useful tool to design and develop conversational AI systems, which create value not only for business purposes, but also for the overall good of society. Hence, RQ4.3 can be answered in a positive manner.

6. Summary, Discussion and Outlook

At the end of this thesis, I summarize all the results, ensure that all RQs have been comprehensively answered, critically discuss my findings, and provide an outlook for further, future research and any elaborations of my developed conceptual framework (FCCC framework). In terms of the DSR approach according to Möller et al. (2020), this chapter serves as final DSR phase where the results are communicated and further steps are discussed.

6.1. Summary

The present thesis complements existing research and practical implementations of conversational AI systems by introducing a new framework for the development of self-regulating conversational AI systems capable of adapting their dialogues and dialogue contents to their environment and to their users. The framework I developed follows an interdisciplinary and transdisciplinary approach drawing from the fields of business informatics, economics, psychology and ethics research in both research and practice areas. Building upon systems theory (von Bertalanffy, 1949) in an iterative DSR process, I demonstrated how existing theories and research approaches need to be intertwined to create naturally intelligent conversational AI systems.

In line with DSR, I defined seven design requirements, which I later implemented into an artefact and evaluated accordingly.

I integrated theories and methods such as CC (El Bolock et al., 2020), CWW as a sub-area of fuzzy logic (Zadeh, 1965) and conversational theory (Pask, 1965; Pangaro, 2010). Systems theory seems to me to be an ideal basis and starting point for the development of naturally intelligent conversational AI systems. However, systems theory and fuzzy systems theory (Zadeh, 1982) mainly explore how systems and sub-systems interact, explaining complex phenomena and relationships, but they do not provide precise information on how a conversational AI system should ultimately place the user at its core to adapt individually to them. To address these challenges, I initially utilized the methods of CC. In CC, researchers like El Bolock et al. (2020) focus on the character of the user. The exact character of the user needs to be calculated initially, considering various influencing factors such as personality, affective state, behaviour and others. During the experimental application of CC in this study, I repeatedly demonstrated that calculating the character is not always straightforward. El Bolock et al. (2020) use additional sensors in their application scenarios to collect further relevant user characteristics. However, a conversational AI system, as applied in reality, has no access to additional sensor data but only to the text or voice inputs of a user. Therefore, I supplemented the methods of CC with those of conversational loops. These loops should contribute to the conversational AI system continuously, receiving further information about the character of the user and being able to adapt as a self-regulating system. I aligned the use of conversational loops with the

explanations of Pask (1965) and later also applied by Pangaro (2010) regarding conversational theory. Within the framework of conversational theory, at least two actors interact in a common context and with a common language until an agreement on a particular topic or within a given situation has been found. In my conceptual framework, the actors are the user and the conversational AI system.

Another method I deemed highly useful in the context of self-regulating conversational AI systems is fuzzy logic (Zadeh, 1965). El Bolock et al. (2020) pointed out in their initial explanations of CC that a character cannot usually be based on sharp values, so the use of fuzzy values and fuzzy terms seemed extremely appropriate to my thesis here. I also extended the classical fuzzy methods with research on multidimensional fuzzy classifications (Zumstein, 2026).

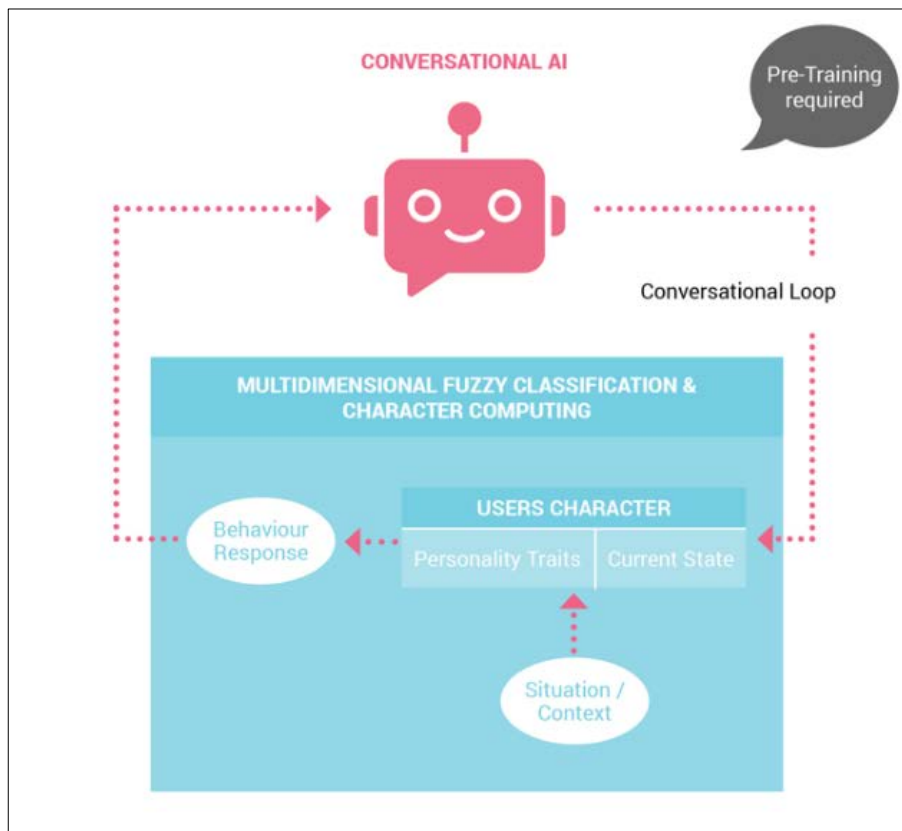
Based on the main theories applied in the system, I call my conceptual framework the FCCC framework, which recalls the use of fuzzy logic, conversations and CC.

Besides these theories, primarily shaped by informatics or psychology, I also considered the needs of economics and ethics research in my study. I remain aware that my framework can only be successfully implemented in practice if ethical, legal and economic aspects are taken into account.

I positive proofed the effectiveness of the artefact in a three step evaluation. First, I tested its practical and technical feasibility, second, I proved its organizational feasibility and, finally, I also showed that my artefact has a social potential. In doing so, I made sure that all design requirements, defined in an earlier step of the DSR, are successfully integrated into the framework. If my framework is used correctly, it will certainly have the potential to significantly improve the quality of customer service and the customer experience in the coming years. However, this requires not only the correct application of my framework, but also the useful application of other tools, such as the value irritant matrix.

At this point, however, I would like to emphasize that the focus of this thesis is on the area of the framework (figure 48), which includes the interplay between conversational AI systems, conversational theory (Pask, 1965; Pangaro, 2010), multidimensional fuzzy classifications (Zumstein, 2026) and CC (El Bolock et al., 2020). The outer layers of the conceptual framework are only elucidated as examples. In the outlook section I provide further ideas for deepening and extending my research focus.

Figure 38: Main area of Framework



Source: Data collected by the author

6.2. Alignment with Research Questions

This section summarizes the answers to the RQs presented in Chapter 1 based on the individual experiments and the artefacts I conducted and developed in the frame of this study.

Research Question 1:

What requirements do users have for conversational AI systems, and how can these be translated into verifiable design guidelines for both the systems themselves and the conceptual framework being developed?

RQ1.1: What design criteria should conversational AI systems incorporate to optimize knowledge transfer while ensuring a positive user experience and supporting life engineering goals?

RQ1.2: What design criteria should conversational AI systems incorporate to balance user objectives, such as quick and correct solutions, with company goals, including economic efficiency and customer experience?

RQ1.3: How do the requirements for conversational AI systems differ depending on the user's situation, age, gender, personality and experience with chatbots?

To follow the DSR research method used in this thesis, RQ1 and its sub-questions belong to the second phase of defining design requirements for the artefact. Hence, I conducted three independent experiments (use case 1-3), with the objective being that each of which would answer one sub-question and thus help me to develop the design requirements and capabilities.

In order to answer RQ1.1, I completed an individual DSR process (section 3.1, use case 1). First, I carried out literature research to gain knowledge for the artefact. Afterwards, I evaluated and improved the artefact according to DSR guidelines. From this, it was possible to get knowledge on the design and application of conversational AI systems in the teaching domain.

All in all, I developed nine design principles that should guide researchers and practitioners in the future when they develop conversational AI systems with the aim of transferring knowledge. The full list can be found in chapter 3.1.

One of the key criteria is the fact that the conversational AI system can adapt to its users. This includes adapting content to the user's knowledge level and tailoring the learning materials accordingly. Furthermore, engaging design elements, such as images and videos, and a context related language, here in the context of knowledge transfer, are of great significance. Besides other design criteria, considering ethical frameworks is a very large factor in the context of knowledge dissemination. The conversational AI systems, which I used here for training on financial literacy, absolutely must consider ethical and moral guidelines. Transparency appears to be one of the most crucial factors in this regard. The conversational AI system must disclose how it operates and the way it accesses knowledge content.

Provided that the evaluated design criteria, including ethical frameworks, are adhered to, the financial literacy use case proved to be a very successful use case for the deployment of conversational AI systems. Furthermore, I also showed that the rule-based chatbots I used in this DSR process are not sufficient to meet all design criteria and, as such, the user's expectations. With rule-based systems, it is almost impossible to adapt accordingly to the user and to pay attention to various, sometimes changing ethical restrictions.

To sum up the use case 1 on knowledge transfer with conversational AI, I gathered six requirements for the artefact. Although I mention nine requirements above, for the planned artefact, only six of them are crucial.

To address RQ1.2, I conducted another experiment (use case 2) in chapter 3.2.

I started with a literature review on conversational AI design criteria and the use of emotions in conversational AI systems, such as chatbots or voicebots. Based on this literature review, I developed three hypotheses concerning expected emotions from a conversational AI system in different user situations, which I then tested with

the aid of prototypes in the form of experiments with test users in the context of insurance. These hypotheses and the experiments were necessary to answer RQ1.2.

When evaluating the results, I found that the first two hypotheses could be confirmed. This leads to the following:

- When taking out supplementary insurance, a chatbot with a more emotionally appealing design should be used.
- When reporting damage, a chatbot should be used without an emotionally appealing design.
- When changing an address, emotions hardly play a role and data is not clear so far. An emotional design can therefore be dispensed with for cost reasons.

So, to answer RQ1.2, developers of conversational AI systems must pay attention to the findings above. Additionally, they must respect the economic goals of the companies. To do so, I examined tools like the value irritant matrix, which provides guidance on what processes and dialogues could be automated in the form of conversational AI agents and other self-service tools.

To sum up this use case, I gained three mandatory design requirements for my artefact.

To answer RQ1.3, I conducted another experiment (use case 3) with a relatively new research design, explained in chapter 3.2.2. Instead of asking the users about their expectations of a chatbot or to evaluate certain conversational AI dialogues, I requested that, during a laboratory experiment, the participants put themselves in the role of a conversational AI system and write the dialogues of a text-based chatbot. Only possible user answers were given; the participants had to write the messages, which in reality would have been written by the conversational AI system.

While in use case 2 I primarily focused on emotions and user situations, in the use case 3 I considered additional user attributes. These included age, gender and personality type according to the Big Five personality traits, explained in chapter 3.2.2.. I again utilized the scenario of an insurance company and aimed to determine the desired level of empathy, the preferred message length and the desired additional information in relation to the user attributes and the situation the user was in. All further insights can be found in chapter 3.

At the end of use case 3 I found three new design requirements for my final artefact.

Research Question 2:

How effectively can methods from Character Computing be integrated into contemporary Conversational AI systems to interpret the complete triad of user

characteristics (personality, affective state and behaviour), and what existing gaps need to be addressed to enable a chatbot to adapt its interactions accordingly?

RQ2.1: Can a conversational AI system derive the complete triad from Character Computing (including personality, affective state and situational behaviour) using features extracted from user messages?

RQ2.2: Can a conversational AI system define users' expectations regarding chatbot interactions when identifying users' personalities and affective states?

To comprehensively answer RQ2, it was first necessary to address the sub-questions. I answered these using a laboratory experiment (use case 4) in which I asked around 2000 insurance customers to report damage to their iPhones. Subsequently, I examined the text messages of the experiment participants with the conversational AI system and correlated them with various user attributes. My user attributes were based on the components of CC, defined in Chapter 2. I aimed to determine whether specific text attributes could provide insights into the user's mood and personality. It was proposed that, if it is possible to extract sufficient information about user attributes solely from text messages, then my system would be capable of computing the triad from CC as described in chapter 2. I analyzed the results of my experiment using various statistical methods and observed that there were very few definitive results. Within the scope of my experiments, I was unable to establish a method for categorizing users based on their text messages. Consequently, computing the triad through CC remained unfeasible. Ultimately, I was unable to provide definitive and affirmative answers to RQ2 and its sub-questions due to the aforementioned experimental outcomes. However, these inconclusive results nonetheless yielded valuable insights for the development of the artefact. Specifically, they helped me identify gaps in the practical application of CC, which I aimed to address over the course of the artefact's development. Hence, at the end of this use case 4 I gained two more requirements for the artefact, which I am going to develop.

Research Question 3:

How can sustainable and naturally intelligent conversational AI systems that function as a self-regulating system, as known from systems theory and that compute with fuzzy values, be implemented by companies in the long term while complying with ethical, cultural, legal and economic frameworks? Over the course of chapter 4 I answered RQ3 while going through the whole DSR process. I developed and presented the FCCC framework, a novel conceptual approach that integrates principles from systems theory, fuzzy computing and conversational theory to design conversational AI systems that are sustainable, naturally intelligent and self-regulating.

In order to answer RQ3 I used the design requirements, which derived from the use cases 1, 2, 3 and 4 and RQ 1 and 2 and developed a framework which addresses these requirements (table 37).

Table 37: Capabilities linked to framework element

Design Requirement (DR)	Capability	Explanation	Framework Element
1	Audience-Adapted and Context-Aware Communication	Use accessible language tailored to the audience while dynamically adapting to user context, including emotional states and situational needs.	Conversational loops informed by fuzzy classifications and CC to adapt based on emotional and situational contexts.
2	Dynamic Personalization	Create interactions based on user history, preferences and profiles, enabling personalized experiences while ensuring scalability.	Multidimensional fuzzy classifications and the integration of CC and LLMs for personalization based on contextual user data.
3	Flexible Interaction and Content Integration	Incorporate both structured and open-ended modes of interaction, integrating multimedia elements and customizable content to cater to varied preferences.	Conversational loops and LLM-enhanced content delivery modules for multimodal integration and flexible responses.
4	Trust, Privacy and Economic Accessibility	Prioritize data protection, user anonymity and affordable implementation to foster trust and accessibility.	Ethical, regulatory and economic layers in the framework, supported by fuzzy logic to ensure transparency, data minimization and economic accessibility for scalability.
5	Alignment with Organizational Objectives	Design systems to balance user satisfaction with organizational goals, ensuring operational and economic efficiency.	Economic consideration layer using the value irritant matrix to align chatbot features with organizational and user needs.
6	Integration of Supplementary Methods with CC	Use accessible, conversational language tailored to the target audience, avoiding excessive technical jargon where not required.	Integration of conversational theory with fuzzy logic to simplify complex interactions and improve user alignment without compromising technical robustness.
7	Enhanced Decision-Making Using Fuzzy Logic	Incorporate fuzzy computing techniques to interpret ambiguous user characteristics, enabling nuanced responses.	The foundational framework combining fuzzy classifications with CC to handle ambiguity in user characteristics effectively.

Sources: Data collected by the author.

In doing so, the framework directly aligns with the core components of RQ3:

Sustainability and Long-Term Implementation

The framework ensures sustainability by integrating feedback loops from conversational theory, allowing AI systems to continuously learn and adapt through interaction rather than relying on static models. Ethical and economic considerations are embedded in the system design, ensuring compliance with evolving regulatory and societal expectations (**DR 1**)

Self-Regulating AI Systems and Systems Theory

The self-regulating nature of the proposed conversational AI is directly inspired by von Bertalanffy's systems theory (1949). It establishes multi-layered feedback loops, enabling AI systems to adapt autonomously to user behaviour, contextual changes and real-time input. The framework ensures AI adaptability by incorporating CC, which allows the system to assess user personality, emotional states and situational factors dynamically (**DR 2, 6**).

Fuzzy Computing for Decision-Making

The AI system leverages fuzzy logic (Zadeh, 1965) and multidimensional fuzzy classifications (Zumstein, 2026) to enhance decision-making in ambiguous or complex scenarios. Unlike traditional deterministic models, the fuzzy computing approach ensures flexibility in AI responses, allowing AI systems to operate effectively under uncertainty and with linguistic approximations (**DR 3, 7**).

Compliance with Ethical, Cultural, Legal and Economic Constraints

The three-layered influence model within the framework ensures full adherence to external requirements. The ethical layer guarantees transparency, user consent, and fairness in AI-driven decisions. The legal layer incorporates all relevant regulatory obligations, such as ensuring GDPR-compliant data processing. Complementing these, the economic layer ensures that AI deployment aligns with the organisation's strategic objectives while remaining cost-efficient and economically viable (**DR 4, 5**).

Through the development of the FCCC framework, several key insights emerged:

- Conversational AI must go beyond traditional chatbots by integrating CC and fuzzy logic, which enables more natural and adaptive interactions with users.
- Conversational loops are critical for long-term AI sustainability, as they allow for continuous learning and refinement based on user behaviour and real-time feedback.
- The combination of multiple disciplines (psychology, computer science and economics) is essential for achieving AI systems that are both functionally effective and societally responsible.

- AI implementation must be adaptable to external regulations and internal business objectives, requiring a structured approach to governance and compliance.

In sum, I successfully answer RQ3 by providing a structured framework that incorporates systems theory, fuzzy logic and ethical constraints into the design of self-regulating, sustainable conversational AI systems. The FCCC framework serves as a foundational model for businesses aiming to implement conversational AI that is adaptive, ethically sound and economically viable in the long term.

Research Question 4:

How can the conceptual framework be evaluated within the Design Science Research process to ensure its theoretical robustness, practical feasibility, real-world applicability and societal impact?

This overarching RQ incorporates the evaluation of the framework across technical, practical, organizational and societal dimensions, ensuring a holistic approach to validation.

I addressed RQ4 by asking three independent sub-questions:

RQ4.1: How can the principles of Fuzzy Logic and Fuzzy Conversational Character Computing be technically operationalized to process ambiguous user inputs, and dynamically adapt to personality traits and emotional states?

RQ4.2: What organizational capabilities must be in place to conceptually and technically implement the framework, and how realistic is it for organizations to meet these prerequisites?

RQ4.3: Does the framework align with the principles of Life Engineering, and how well does it meet ethical expectations to deliver the desired societal impact?

According to DSR method, each artefact needs to be evaluated. To do so, RQ4 and its sub-questions served to help evaluate the artefact using three independent evaluation methods.

I systematically examined the artefact across technical, practical, organizational and societal dimensions. Ultimately, the answer to RQ4 can be substantiated as follows:

Technical Operationalization of Fuzzy Logic and Fuzzy Conversational Character Computing

The thesis successfully demonstrated how the principles of fuzzy Logic and FCCC can be implemented to handle ambiguous user inputs and dynamically adjust responses based on personality traits and emotional states, known from character computing. Through mathematical representation and subsequent technical implementation, the framework was translated into a working text-based chatbot system,

evaluated through expert interviews. The study provided empirical evidence that the framework enables real-time sentiment adaptation and behaviour estimation, confirming its technical feasibility and effectiveness.

Organizational Capabilities for Implementation

The thesis identifies the organizational capabilities necessary for implementing the framework and evaluates the practicality of achieving them within different organizational contexts. Through focus groups with AI experts, decision-makers, and industry practitioners, the study assessed the required AI skills, IT infrastructure and cultural readiness. The findings indicate that, while large organizations are well-equipped for implementation, medium and small enterprises face structural and knowledge-related barriers that may require external support. These insights confirm that the framework is realistic for organizations but requires strategic capability-building efforts.

Alignment with Life Engineering and Ethical Considerations

The thesis examined the societal and ethical implications of the framework, particularly in the context of life engineering. It argued that the framework aligns with the principles of life engineering by prioritizing ethical AI practices, transparency and adaptability to individual user needs. The expert interviews further validated that the framework contributes positively to human-centred AI design, supporting financial literacy and fostering responsible AI applications. At the same time, the expert evaluation surfaced three concrete limitations that future iterations must address. First, the LLM-based components of the current implementation lack the auditability that a regulated deployment in the financial domain would require under the EU AI Act (Regulation (EU) 2024/1689). The model's reasoning process is not directly inspectable, and scope-boundary enforcement mechanisms are not yet in place. Second, the framework's tone-adaptation mechanism carries an inherent nudging risk by adjusting phrasing to a user's estimated psychological state without explicit consent, the system may subtly influence financial decisions in ways that compromise user autonomy. Future iterations should implement disclosure and opt-out mechanisms to mitigate this. Third, the use of a large-scale proprietary LLM for every conversational turn generates a non-trivial energy footprint that is difficult to reconcile with life engineering's sustainability commitments at societal deployment scale. Replacing the character estimation component with a locally deployable compact model represents the most promising mitigation pathway for future improvements.

At the end of the Chapter 5, I was able to answer RQ4 and obtained the following, key findings:

- **Technical Feasibility:** the FCCC framework is technically operationalizable, demonstrating adaptability to real-time user sentiment and personality traits.

- **Organizational Readiness:** successful implementation depends on AI expertise, data infrastructure and cultural openness. Large organizations are better prepared, while smaller ones may require external partnerships.
- **Ethical and Societal Impact:** the framework aligns with life engineering principles, promoting ethical AI use and enhancing financial literacy. However, sustainability concerns should be addressed in future iterations.

Thus, Chapter 5's evaluation provided a holistic validation of the artefact, thoroughly answering RQ4 and its sub-questions.

6.3. Additional Remarks from a Life Engineering perspective

Although the framework received an overall positive evaluation from the Life Engineering experts, it is important to provide additional reflections on the broader vision of Life Engineering and the potential implications of applying the proposed framework. As this dissertation is closely aligned with the goals and principles of Life Engineering, it is essential to transparently acknowledge potential risks from the outset and to address them in future iterations with the aim of minimizing them wherever possible. Such critical reflection is also consistent with the chosen research methodology of DSR, including the conducted Turing-test-based evaluation, which explicitly encourages iterative improvement and the continuous examination of both opportunities and limitations of the proposed artefact.

Transparency and auditability of LLM-based outputs

The current implementation relies on GPT-4o for character estimation, reply generation, and rephrasing. In each of these roles, the model's internal reasoning process is not directly inspectable, which raises non-trivial concerns in any regulated deployment context. The EU Artificial Intelligence Act (Regulation (EU) 2024/1689) imposes transparency and documentation requirements on AI systems operating in sensitive domains such as personal finance. Future iterations of the FCCC framework must therefore specify the model selection rationale, maintain version-locked prompts, and implement scope-boundary mechanisms that prevent the system from generating personal advices beyond its validated domain. Extending the framework's existing transparency principle, currently expressed through knowledge-source disclosure in the dialogue, to the level of the underlying model architecture is a concrete next step for responsible deployment.

A further ethical limitation concerns the risk of inadvertent nudging

The FCCC framework's tone adaptation mechanism dynamically adjusts the phrasing and emotional register of responses based on the estimated psychological state of the user. In most contexts, this enhances the user experience and supports the life engineering goal of individually tailored communication. However, in sensitive situations, the same mechanism may subtly steer user decisions by exploiting

estimated traits such as high agreeableness or elevated neuroticism, without the user's explicit awareness or consent. This constitutes a form of psychologically informed influence that is difficult to distinguish from coercion when applied to consequential decisions. Future iterations should introduce opt-out mechanisms for tone adaptation, periodic disclosure of the system's personalisation logic, and a rule-based filter that prevents the reply rephrasing component from applying emotionally loaded framing to personal recommendations.

Energy consumption associated with LLMs

A life engineering system that claims to support human well-being while generating a significant carbon footprint at deployment scale creates an ethical tension that the current prototype does not yet resolve. Each API call to GPT-4o or any other LLM carries an energy cost that is negligible at the scale of an academic study but becomes substantial when projected to societal deployment scenarios. Future iterations should explore selectively replacing the character estimation component with a locally deployable compact model, implementing delta-threshold strategies that suppress redundant inference calls, and applying sustainability criteria to the choice of inference infrastructure. These measures are not merely technical optimizations; given the life engineering context, they are ethical requirements.

6.4. Remarks on the Use of Large Language Models (LLMs)

Although I used OpenAI's LLM ChatGPT in the demo installation, I would like to compare my framework with the single application of LLMs. Newest LLMs have demonstrated remarkable capabilities for generating coherent text and emulating human-like language patterns. Nonetheless, they often operate as generalized “black boxes” that do not explicitly track or integrate user traits, emotional states, and dynamic contextual factors in a formally structured manner. By contrast, the FCCC framework provides:

- 1. Explicit User Trait Modeling:** Multi-dimensional fuzzy classifications allow partial membership in psychologically relevant dimensions (for example openness or frustration). This is more than a simple next-word prediction since it builds a richer, evolving user profile.
- 2. Captures Medium-Term Trends:** Through the behavior vector B , the system recognizes the user's trend over multiple conversational turns. Many LLMs have no explicit mechanism for smoothing or aggregating personality or emotional cues across time.
- 3. Situational Context:** FCCC's situation S explicitly stores the conversation history in a structured way, allowing the system to reason about which prior utterances and contextual details remain relevant. This fosters continuity and coherence beyond typical LLM context windows.

4. **Greater Interpretability:** In fuzzy logic, membership degrees are directly interpretable. Developers can see how intensities of emotional or personality attributes influence responses. A level of introspection is often missing in purely emergent LLM models.
5. **Human-in-the-Loop Adaptation:** The triad-based approach encourages iterative updates of membership functions and thresholds, easily adjusted by domain experts or psychologists. Retraining LLMs can be opaque, but adjusting fuzzy membership sets is transparent.
6. **Data Minimization:** Rather than collecting vast amounts of user data to fine-tune or prompt a LLM, the FCCC framework focuses on *pertinent attributes* (e.g. personality traits). Membership values can be stored and updated without large-scale text corpora, enabling a more privacy-friendly approach that collects only the minimal data required for personalization. This reduces concerns around data privacy, compliance, and storage overhead, especially compared to the resource-intensive pipelines for LLMs.
7. **Data dependency** means AI systems rely heavily on labeled input and continuous streams of data. Here, the situation vector (S) and the iterative updates in the character vector (C) and the behaviour vector (B) reflect how each new data point re-shapes the user’s membership degrees. As input distribution shifts, the system recalculates membership accordingly, requiring less monolithic retraining than typical large-scale end-to-end approaches.

Hence, even as LLMs advance and become widely integrated, FCCC complements them with an explicit, transparent mechanism to track user traits across multiple fuzzy dimensions, ensuring user-centered adaptation that is directly interpretable, data-sparse, and aligned with psychological theories of behavior and character. By merging classical character computing ideas with fuzzy logic, the FCCC framework achieves rich modeling of user traits, behaviors, and contexts. The derived formulas for S, C, B, and R unify the conversation’s running context, the short-term user states, historical behavioral patterns, and the final system response.

The practical implementation of the approach involves three interdependent steps. First, linguistic dimensions and their corresponding membership functions are defined, requiring careful calibration of the fuzzy sets associated with each trait, emotion, or psychological variable. Second, the system performs continuous updates by computing the character vector C for every user message, aggregating these vectors into the short-term behaviour representation B, and appending all messages to the situational memory S. Finally, reply generation (R) is carried out using a domain-appropriate function f , which may range from a rule-based mechanism to more advanced machine-learning methods, in order to produce responses that integrate the information stored in S, C, and B.

Such synergy promises more empathetic, adaptive, and user-centric interactions than purely rule-based or crisp classification approaches. It directly addresses the complexities of real-world data, bridging the gap between ambiguous linguistic expressions and actionable computational representations. Even as state-of-the-art LLMs continue to excel at generating text, FCCC provides a transparent, psychologically grounded, data-sparse, and explainable architecture for capturing long-term, multi-dimensional user traits and delivering truly tailored and context-aware conversations.

6.5. Outlook and Next Steps

The focus of this thesis lies on the yellow area of the framework (figure 33) which includes the interplay between conversational AI system, conversational theory (Pask, 1965; Pangaro, 2010), multidimensional fuzzy classifications (Zumstein, 2026) and character computing (El Bolock et al., 2020). The integration of LLMs, multi-modal models and the outer layers of ethics, legal issues and economic considerations was not the focus of this study, hence this was only superficially explored.

In future iterations of the DSR process, these parts should be explored in more detail. A possible first step could be to focus on the areas above conversational AI. For example, the use of LLMs currently requires significant computing power and leads to high energy consumption. It would be advisable, therefore, to explore ways to integrate the use of LLMs into the framework, taking into account data sparsity, to enable a self-improving and self-learning system. Furthermore, the interaction between conversational AI systems and LLMs should be explained in more detail in the framework and in the sample applications.

It is also important to explore the influence of the control algorithms in greater detail and to define possible stakeholders or instances for monitoring these algorithms. To date, the control algorithms have only been placed in the realm of LLMs. However, there is a lack of further information on who “monitors” these algorithms, in what way these are monitored and also whether or how they may become part of the self-improving system.

Similar unanswered questions exist in the area of pre-training. In my conceptual framework and in the sample applications, it is assumed that the conversational AI system has a certain degree of pre-training before it is integrated into the conversational loop system. However, my framework does not yet provide any conclusive information about the level of detail of the pre-training or the responsibility for it. It is possible that the external layers could have an influence on the pre-training, for example leading to ethical concerns, or whether companies themselves could be responsible for the pre-training. It is also conceivable that additional stakeholders, not previously considered in the framework, are responsible for pre-training.

Once these questions have been explored in further iterations, additional research on the outer layers of the framework should follow. In particular, the integration of ethical frameworks currently raises unanswered questions. It has not yet been researched in detail which concrete ethical requirements are placed on conversational AI and who determines them. It is unclear which scales are used to measure ethics and who decides on them. In addition, ethical and cultural requirements may change over time, and these changes also need to be integrated into conversational AI. The ethical and cultural influences may vary depending on the application domain of conversational AI, but there has not yet been detailed research on how much the framework varies depending on the application domain. In their article on computational ethics, Portman et al. (2022) present the first approaches to solving the problem of how AI can make independent decisions while considering the ethical values of individuals and society.

Further iterations of the research should look more closely at the legal and regulatory framework. It is to be expected that further regulations for the use of AI will be enacted in the coming years. The aim here is not to develop legal directives within the concept of this framework. Rather, the central question for the framework is how to ensure that conversational AI always takes the latest regulations into account.

Finally, the innermost of the three outer levels also needs to be explored in more detail. In particular, the possible tension between customer experience and economic goals must be examined more closely, and approaches to finding a consensus in this area must be obtained. Here, an in-depth examination or application of the value of the of the irritant matrix would be extremely helpful.

Beyond the technical and organizational outlook described above, the life engineering evaluation conducted in Chapter 5.3 surfaces a set of ethical concerns that deserve their own dedicated research agenda. Taken together, these concerns, the discussions on the use of LLMs discussion, and the sustainability considerations raised throughout this thesis constitute three prioritized topics for the next iteration of the FCCC framework in the focus context of life engineering.

Auditability and Regulatory Compliance (EU AI Act)

The framework's LLM-based components do not yet satisfy the auditability requirements that a regulated deployment in a sensitive domain would impose. Future work should implement version-locked prompt registries with change tracking, scope-boundary enforcement and an explainability layer that makes at least the character estimation component interpretable to domain experts and regulators.

Nudge Prevention and Autonomous Decision-Making

The framework's tone-adaptation mechanism constitutes a form of psychologically informed influence that requires explicit mitigation. By adjusting phrasing and framing to the user's estimated personality state, the system may subtly shape

decisions without the user's awareness or consent, a form of nudging. The research agenda here includes: development of a formal nudge-detection metric for adaptive dialogue systems. User studies comparing decision quality under tone-adaptive vs. tone-neutral conditions and the design of an informed consent and opt-out interface that is both legally compliant and practically usable.

Computational Sustainability

The energy cost of LLM-based inference is a structural constraint on the framework's life engineering ambitions. A life agent that claims to support human well-being while generating a non-trivial carbon footprint at societal deployment scale creates an ethical tension that cannot be ignored. The research agenda includes empirical benchmarking of compact personality-classification models as local replacements for the LLM component, development and validation of delta-threshold invocation strategies that reduce unnecessary API calls and a lifecycle carbon assessment of the full FCCC pipeline under realistic deployment scenarios.

These three topics are interconnected rather than parallel. Auditability supports nudge prevention by making the system's reasoning traceable, nudge prevention is a precondition for sustainable deployment at scale, because a system that influences users without transparency cannot build the long-term trust that a societal-scale life agent requires and computational sustainability is a prerequisite for any deployment that claims to serve human well-being without transferring its costs to the natural environment. Addressing all three topics is therefore not a matter of completeness but of coherence. Each is necessary, and none is sufficient without the others.

At this point, it should also be mentioned that also besides life engineering customer experience research is currently undergoing major changes due to rapid technological advances. Developments in the fields of generative AI and LLMs are leading to more and more automation in customer service. At the same time, customer expectations are changing. Real-time information is becoming increasingly important, and the need for real human dialogues versus automated dialogues with machines will change as soon as the quality of automated dialogues around the clock has reached a new level of quality. These changing demands of customers and society must also be taken into account in further iterations of the framework.

I am confident that, with this conceptual framework, I have laid crucial foundations for the further development of conversational AI systems and life agents, according to life engineering. Even though my FCCC framework leaves room for further research and detailed elaborations, I have initially achieved the goal of this thesis and was also able to answer all research questions.

I am aware that advancements in the field of AI and conversational AI systems will continue to grow in the coming years and that so-called life agents will increasingly become a part of our economic and societal everyday lives.

I should not forget to mention that, for the successful introduction of such conversational AI systems, not only are technical steps necessary, but every individual and society as a whole bear responsibility. We must ensure that we benefit from research advancements and do not eliminate or render ourselves obsolete. We need to learn in which domains the deployment of such conversational AI systems and life agents is advantageous, and in which situations we should continue to interact human-to-human. However, when we employ these life agents, we must understand how to use them and how to communicate with them.

I am convinced that exciting times lie ahead, and I look forward to further challenges that I believe can be addressed primarily through interdisciplinary research approaches drawing from the fields of business informatics, economics, psychology and ethics research and, secondarily, through transdisciplinary exploration approaches from research and practice.

7. References

- Ackermann, N., & Eberle, F. (2016). Financial Literacy in Switzerland. In C. Aprea, E. Wuttke, K. Breuer, N. K. Koh, P. Davis, B. Greimel-Fuhrmann, & J. S. Lopus (Eds.), *International Handbook of Financial Literacy* (pp. 341–355). Springer. https://doi.org/10.1007/978-981-10-0360-8_23
- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycholexical study. *Psychological Monographs*, 47(1), i–171. <https://doi.org/10.1037/h0093360>
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2003). *Multivariate Analysis Methods: Eine anwendungsorientierte Einführung* (10th ed.). Springer.
- Bevacqua, E., Pammi, S., Hyniewska, S. J., Schröder, M., & Pelachaud, P. (2010). Multimodal Backchannels for Embodied Conversational Agents. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Lecture Notes in Computer Science: Vol. 6356. Intelligent Virtual Agents* (pp. 194–200). Springer. https://doi.org/10.1007/978-3-642-15892-6_21
- Brabra, H., Báez, M., Benatallah, B., Gaaloul, W. Bouguelia, S., & Zamanirad, S. (2022). Dialogue Management in Conversational Systems: A Review of Approaches, Challenges, and Opportunities. *IEEE Transactions on Cognitive and Developmental Systems*, 14(3), 783–798. <https://doi.org/10.1109/TCDS.2021.3086565>
- Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: changing user needs and motivations. *Interactions*, 25(5), 38–43. <https://doi.org/10.1145/3236669>
- Braun, M., & Graf, R. (2013). Financial literacy and retirement planning in Switzerland. *Numeracy*, 6(2), Article 6. <http://dx.doi.org/10.5038/1936-4660.6.2.6>
- Bruhn, M. (2016). *Relationship Marketing: Das Management von Kundenbeziehungen* (5th ed.). Vahlen
- Busetto, L., Wick, W., & Gumbinger, C. (2020). How to use and assess qualitative research methods. *Neurological Research and Practice*, 2, Article 14. <https://doi.org/10.1186/s42466-020-00059-z>
- Carrera-Rivera, A., Ochoa, W., Larrinaga, F., & Lasa, G. (2022). How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX*, 9, Article 101895. <https://doi.org/10.1016/j.mex.2022.101895>
- Carolis, B. d. (2004). APML, a markup language for believable behavior generation. In H. I. Prendinger, *Life-Like Characters: Tools, Affective Functions, and Applications* (pp. 65-85). Berlin Heidelberg: Springer.
- Christopher, M., Payne, A., & Ballantyne, D. (1991). *Relationship Marketing: Bringing Quality, Customer Service and Marketing Together*. Butterworth-Heinemann.

- Clarizia, F., Colace, F., Lombardi, M., Pascale F., & Santaniello, D. (2018). Chatbot: An Education Support System for Student. In A. Castiglione, F. Pop, M. Ficco, & F. Palmieri (Eds.), *Lecture notes in computer science: Vol. 11161. Cyber-space Safety and Security* (pp. 291–301). Springer. https://doi.org/10.1007/978-3-030-01689-0_23
- Cohen, Myke & Su, Zhe & Kao, Hsien-Te & Nguyen, Dan & Lynch, Spencer & Sap, Maarten & Volkova, Svitlana. (2025). Exploring Big Five Personality and AI Capability Effects in LLM-Simulated Negotiation Dialogues. 10.48550/arXiv.2506.15928.
- Colombo, M. (2022). *Phenotropic Interaction - Improving Interfaces with Computing With Words and Perceptions* [Doctoral dissertation, University of Fribourg]. Research Gate. <https://doi.org/10.13140/RG.2.2.10035.91683>
- Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions. *Journal of Marketing*, 86(1), 132–148. <https://doi.org/10.1177/00222429211045687>
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811–817. <https://doi.org/10.1017/S1351324916000243>
- Davis F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology*, 22(14), 1111–1132. <https://doi.org/10.1111/j.1559-1816.1992.tb00945.x>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5), 880–896. <https://doi.org/10.1037/0022-3514.93.5.880>
- Diederich, S., Brendel, A. B., Lichtenberg, S., & Kolbe, L. (2019). Design for Fast Request Fulfillment or Natural Interaction? Insights from an Experiment with a Conversational Agent. *Proceedings of the 27th European Conference on Information Systems (ECIS), Sweden*. https://aisel.aisnet.org/ecis2019_rp/20/
- Doctor, F. Karyotis, C., Iqbal, R., & James, A. (2016). *An intelligent framework for emotion aware e-healthcare support systems*. 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Greece, 1–18. <https://doi.org/10.1109/SSCI.2016.7850044>
- D’Onofrio, S, & Portmann, E. (2015). Von Fuzzy-Sets zu Computing-With-Words. *Informatik Spektrum*, 38(6), 543–549. <https://doi.org/10.1007/s00287-015-0920-y>

- Efron, B. (1992). Bootstrap Methods: Another Look at the Jackknife. In: Kotz, S., Johnson, N.L. (eds) Breakthroughs in Statistics. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4380-9_41
- El Bolock, A., Abdelrahman Y., & Abdennadher S. (2020). *Character computing*. Springer. <https://doi.org/10.1007/978-3-030-15954-2>
- Fahrmeir, L., Kneib, T., & Lang, S. (2021). *Regression. Models, Methods and Applications* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-662-63882-8>
- Feine, J. G., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational. *International Journal of Human-Computer Studies*, 132, 138–161. <https://doi.org/10.1016/j.ijhcs.2019.07.009>
- Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (4th ed.). Sage Publications.
- Ghandeharioun, A., McDuff, D., Czerwinski, M., & Rowan, K. (2019). Towards Understanding Emotional Intelligence for Behavior Change Chatbots. *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, UK, 8–14. <https://doi.org/10.1109/ACII.2019.8925433>
- Hafner, N., & Henn, H. (2021). Voice in the D2C Journey - Voice automation as an experience. *Marketing Review St. Gallen*, 38(6), 46–53. <https://hdl.handle.net/10419/276167>
- Hevner, A., Chatterjee, S. (2010). *Design Research in Information Systems: Theory and Practice*. Springer. <https://doi.org/10.1007/978-1-4419-5653-8>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Hildebrand, C., & Hundertmark, S. (2021). A Strategy Framework to Boost Conversational AI Performance. *Marketing Review St. Gallen*, 38(4), 10–16. <https://hdl.handle.net/10419/276147>
- Hundertmark, S. (2020), *Digitale Freunde: Wie Unternehmen Chatbots Erfolgreich Einsetzen Können*. Wiley-VCH.
- Hundertmark, S. (2023). *Be the Bot – Versicherungskunden in der Rolle eines Chatbots*. In N. Hafner & S. Hundertmark (Eds.). *Kundendialog-Management: Wertstiftende Kundendialoge in Zeiten der digitalen Automation* (pp. 189–200). Springer Gabler. https://doi.org/10.1007/978-3-658-42851-8_12
- Hundertmark, S., Fill, H.-G., & Portmann, E. (2022a). From practice to science: Social reactions caused by conversational agents—A literature review for researchers and practitioners with a chatbot as an outcome. *Proceedings of the International Conferences on Applied Computing 2022 and WWW/Internet*

- 2022, Portugal, 156-164. https://www.computing-conf.org/wp-content/uploads/2022/11/1_ICWI2022_F_039.pdf
- Hundertmark, S., Grüter, A., & Portmann, E. (2022b). Mit Chatbots die finanzielle Allgemeinbildung stärken. *Informatik Spektrum*, 45(2), 80–87. <https://doi.org/10.1007/s00287-022-01450-7>
- Hundertmark, S., & Hafner, N. (2023a). Gestaltung automatisierter Kundendialoge im Spannungsfeld von Kosten und Customer Experience. In M. Bruhn, & K. Hadwich (Eds.), *Gestaltung des Wandels im Dienstleistungsmanagement: Band 2: Kundenperspektive – Anbieterperspektive – Mitarbeiterperspektive* (pp. 219–238). Springer Gabler. https://doi.org/10.1007/978-3-658-41815-1_9
- Hundertmark, S., Hafner, N., Portmann, E. (2023b). Dialog-Anpassung an den Charakter des Benutzers – Anwendung der Methoden des Charakter Computing. *Informatik Spektrum*, 46(4), 210-219. <https://doi.org/10.1007/s00287-023-01547-7>
- Hundertmark, S., Hafner, N. (2021). *IFZ Conversational Banking Studie*: <https://hub.hslu.ch/retailbanking/wp-content/uploads/sites/7/2022/08/IFZ-Conversational-Banking-Studie-2021.pdf>
- Hundertmark, S., Portmann, E. (2022c). Ethics Check for Conversational AIs. *Proceedings of the ACM CHI 2022 conference, USA*. <https://cui.acm.org/workshops/CHI2022/papers.html>
- Hundertmark, S., Portmann, E., Hafner, N. (2023c). Fuzzy Conversational Character Computing. *Informatik Spektrum* 46(3), 163–175. <https://doi.org/10.1007/s00287-023-01538-8>
- Hundertmark, S., Christen, R., Portmann, E. (2025). Chatbots with Character - An Implementation of Fuzzy Conversational Character Computing. In: Baczyński, M., De Baets, B., Holčapek, M., Kreinovich, V., Medina, J. (eds) *Advances in Fuzzy Logic and Technology. EUSFLAT 2025. Lecture Notes in Computer Science*, vol 15883. Springer, Cham. https://doi.org/10.1007/978-3-031-97225-6_4
- Hundertmark, S., Schreiber, F. (2022d). *IFZ Conversational Insurance Study 2022: Chat als neuer Kommunikationskanal bei Versicherern? Was wollen die Kunden?*. Institut für Finanzdienstleistungen Zug IFZ. <https://hub.hslu.ch/insuranceinsights/ifz-conversational-insurance-studie-2022/>
- John, O.P., Naumann, L., Soto, C. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). The Guilford Press.

- Jordan, M., Bruce, M., Kuo, I.-H., Li, X., Broadbent, E. (2015). People respond better to robots than computer tablets delivering healthcare instructions. *Computers in Human Behavior*, 43, 112–117. <https://doi.org/10.1016/j.chb.2014.10.029>
- Katayama, S., Mathur, A., Broeck, M.V., Okoshi, T., Nakazawa, J., & Kawsar, F. (2019). Situation-aware emotion regulation of conversational agents with kinetic earables. *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, UK, 725-731. <https://doi.org/10.1109/ACII.2019.8925449>
- Klaus, P., & Manthiou, K. (2020). Applying the EEE customer mindset in luxury: reevaluating customer experience research and practice during and after corona. *Journal of Service Management*, 31(6), 1175–1183. <https://doi.org/10.1108/JOSM-05-2020-0159>
- Lamolle, M., Mancini, M., Pelachaud, C., Abrilian, S., Martin, J.-C., & Devillers, L. (2005). Contextual factors and adaptive multimodal human-computer interaction: Multi-level specification of emotion and expressivity in embodied conversational agents. *Proceedings of the 5th International and Interdisciplinary Conference, France*, 225–239. https://doi.org/10.1007/11508373_17
- Landmann, E., Siems, F. U., Hafner, N., Pape, M.-C. (2022). Die Value-Irritant-Matrix als mögliches Instrument zur empirischen Klassifikation von Potenzialen und Grenzen von Smart Services – Ein Anwendungsversuch mit kritischer Diskussion. In M. Bruhn & K. Hadwich (Eds.), *Smart Services* (pp. 273–290). Springer Gabler. https://doi.org/10.1007/978-3-658-37344-3_8
- Lemon, K. N., Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96. <https://doi.org/10.1509/jm.15.0420>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22 140, 55.
- Mairesse, F., Walker, M. A., Mehl, M. R., Moore, R. K. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text, *Journal of Artificial Intelligence Research* 30 (2007) 457–500. doi:10.1613/jair.2349.
- Marheineke, M. (2016). *Designing Boundary Objects for Virtual Collaboration*. Springer Gabler. <https://doi.org/10.1007/978-3-658-15386-1>
- McCrae RR, John OP. An introduction to the five-factor model and its applications. *J Pers.* 1992 Jun;60(2):175-215. doi: 10.1111/j.1467-6494.1992.tb00970.x. PMID: 1635039.
- Mimoun, M. S. B., Poncin, I., Garnier, M. (2017). Animated conversational agents and e-consumer productivity: The roles of agents and individual characteristics.

- Mockenhaupt, A. (2021). *Digitalisierung und Künstliche Intelligenz in der Produktion: Grundlagen und Anwendung*. Springer Vieweg.
<https://doi.org/10.1007/978-3-658-32773-6>
- Möller, F., Guggenberger, T. M., & Otto, B. (2020). Towards a method for design principle development in information systems. In S. Hofmann, O. Müller, & M. Rossi (Eds.), *Lecture Notes in Computer Science: Vol. 12388. Designing for Digital Transformation. Co-Creating Services with Citizens and Industry* (pp. 208–220). Springer. https://doi.org/10.1007/978-3-030-64823-7_20
- Mozafari, N., Hammerschmidt, M., Weiger, W. (2021). *That's so Embarrassing!* When not to Design for Social Presence in Human-Chatbot Interactions. Proceedings of the International Conference on Information Systems, USA.
<https://ssrn.com/abstract=4305288>
- Nass, C., Moon, Y., Green, N. (1997). Are machines gender neutral?: Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27(10), 864–876. <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>
- Nielsen, J.: Estimating the number of subjects needed for a thinking aloud test. *International Journal of Human-Computer Studies* 41(3), 385–397 (Sep 1994).
<https://doi.org/10.1006/ijhc.1994.1065>
- Ng, M., Coopamootoo, K. P. L., Toreini, E., Aitken, M., Elliott, K., & van Moorsel, A. (2020). Simulating the effects of social presence on trust, privacy concerns & usage intentions in automated bots for finance. *Proceedings of the 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Italy*, 190–199. <https://doi.org/10.1109/EuroSPW51379.2020.00034>
- Österle, H. (2020). *Life engineering: Machine intelligence and quality of life*. Springer. <https://doi.org/10.1007/978-3-030-31482-8>
- Oghuma, A. P., Libaque-Saenz, C. F., Wong, S. F., & Chang, Y. (2016). An expectation-confirmation model of continuance intention to use mobile instant messaging. *Telematics and Informatics*, 33(1), 34–47.
<https://doi.org/10.1016/j.tele.2015.05.006>
- Oliver, R.L. (2010). *Satisfaction: A Behavioral Perspective on the Consumer: A Behavioral Perspective on the Consumer* (2nd ed.). Routledge.
<https://doi.org/10.4324/9781315700892>
- OpenAI. (2022, November 30). *Introducing ChatGPT*.
<https://openai.com/blog/chatgpt>.
- O'Regan, G. (2008). *A brief history of computing*. Springer.
<https://doi.org/10.1007/978-1-84800-084-1>

- Pangaro, P. (2010). *Rethinking design thinking*. Pangaro. <https://pangaro.com/PIC-NIC10/index.html>
- Pangaro, P. (2019, November 18). *Designing ethical interfaces*. Pangaro. <https://pangaro.com/ethics2019/index.html>
- Pask, G. (1965). The cybernetics of evolutionary process and self-organising systems. *Proceedings of the 3rd Congress of the International Association for Cybernetics, Belgium*.
- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3). <https://doi.org/10.2753/MIS0742-1222240302>
- Petrović, J., Jovanović, M. (2020). Conversational agents for learning foreign languages - A survey. *Proceedings of the Sinteza 2020 - International Scientific Conference on Information Technology and Data Related Research, Serbia*, 14–22. <https://doi.org/10.15308/Sinteza-2020-14-22>
- Picard, R. W. (1997). *Affective computing*. MIT Press.
- Portmann, E., D’Onofrio, S. (2022). Computational Ethics. Ein ethischer Lösungsansatz für das KI-Zeitalter. *HMD Praxis der Wirtschaftsinformatik*, 59(2), 447–467. <https://doi.org/10.1365/s40702-022-00855-y>
- Portmann, E.; Wilke, G.; Téran, L.; D’Onofrio, S. Fuzzy Design Science Research. The FMsquare Initiative. Res. Featur. 2024, <https://doi.org/10.26904/rf2025-8056509312>.
- Portmann, E., Wilke, G., Terán, L., & D’Onofrio, S. (Eds.). (2026). *Fuzzy Sets and Systems II: An Introduction with Cases from Business Informatics, Computer Science and Engineering*. Springer. <https://doi.org/10.1007/978-3-031-99294-0>
- Price, B., Jaffe, D. (2008). *The best service is no service: How to liberate your customers from customer service, keep them happy, and control costs*. Jossey-Bass.
- Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2014). Big Five Inventory (BFI-10). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. <https://doi.org/10.6102/zis76>
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). (2024). Official Journal of the European Union, L 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Reichheld, F., & Markey, R. (2011). *The ultimate question 2.0: How Net Promoter companies thrive in a customer-driven world* (rev. and exp. ed.). Harvard Business Review Press.

- Rhio, S., Chowandaa, C., Kurniatia, A., & Wongso, R. (2019). Designing an emotionally realistic chatbot framework to enhance its believability with AIML and information states. *Procedia Computer Science*, 157, 621–628. <https://doi.org/10.1016/j.procs.2019.08.226>
- Schoenegger, P. Greenberg, S. Grishin, A. Lewis, J. Caviola, L. Can AI Understand Human Personality? – Comparing Human Experts and AI Systems at Predicting Personality Correlations (2024). doi:10.48550/ARXIV.2406.08170.
- Schlegel, K. Sommer, N. R. Mortillaro, M. Large language models are proficient in solving and creating emotional intelligence tests, *Communications Psychology* 3 (1) (2025) 80. doi:10.1038/s44271-025-00258-x.
- Sejnowski, T. J. (2023). Large language models and the reverse Turing test. *Neural Computation*, 35(3), 309–342. https://doi.org/10.1162/neco_a_01563
- Sparrenberg, L. Schneider, T. Deußer, T. Koppenborg, M. and Rifa, S. "Correcting Systematic Bias in LLM-Generated Dialogues Using Big Five Personality Traits," *2024 IEEE International Conference on Big Data (BigData)*, Washington, DC, USA, 2024, pp. 3061-3069, doi: 10.1109/BigData62323.2024.10825941.
keywords: {Training;Systematics;Large language models;Psychology;Big Data;Encoding;Data models;Reliability;Mirrors;Standards;large language models;big five;synthetic data},
- Stolper, O., & Walter, A. (2017). Financial literacy, financial advice, and financial behavior. *Journal of Business Economics*, 87(5), 581–643. <https://doi.org/10.1007/s11573-017-0853-9>
- Sturmey, P. (Ed.) (2020). *Functional analysis in clinical treatment* (2nd ed.). Academic Press. <https://doi.org/10.1016/C2015-0-05507-1>
- Teo, T. S. H., Srivastava, S. C., & Jiang, L. (2009). Trust and electronic government success: An empirical study. *Journal of Management Information Systems*, 25(3). <https://doi.org/10.2753/MIS0742-1222250303>
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4), 273–286. <https://doi.org/10.1037/h0070288>
- Völkel, S. T., Kempf, P., & Hussmann, H. (2020). Personalised chats with voice assistants: The user perspective. *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)*, Spain, Article 53. <https://doi.org/10.1145/3405755.3406156>
- vom Brocke, J., Simons, A., Riemer, K., Niehaves, M., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the*

- Association for Information Systems*, 37, 205–224.
<https://doi.org/10.17705/1CAIS.03709>
- Von Bertalanffy, L. (1949). Zu einer allgemeinen Systemlehre. *Biologia Generalis*, 195, 114–129.
- Wambsganss, T., Kueng, T., Söllner, M., & Leimeister, J. M. (2021). ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Japan*, Article 683. <https://doi.org/10.1145/3411764.3445781>
- Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. (2023). Organizational capabilities for AI implementation—coping with inscrutability and data dependency in AI. *Information Systems Frontiers*, 25(4), 1549–1569.
<https://link.springer.com/article/10.1007/s10796-022-10297-y>
- Webster, J. W. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), 13–23. <https://www.jstor.org/stable/4132319>
- Webster, M., & Sell, J. (2014). *Laboratory experiments in the social sciences* (2nd ed.). Academic Press. Elsevier Inc. <https://doi.org/10.1016/C2011-0-07562-2>
- Weizenbaum, J. (1966). ELIZA - a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Werro, N. (2015). Fuzzy Classification of Online Customers. Springer International Publishing. <https://doi.org/10.1007/978-3-319-15970-6>
- Werro, N. (2026). Customer Classification. In E. Portmann, G. Wilke, L. Terán & S. D'Onofrio (Eds.), *Fuzzy Sets and Systems II* (pp. 55–85). Springer. https://doi.org/10.1007/978-3-031-99294-0_3
- Werro, N., Stormer, H., & Meier, A. (2006). A Hierarchical Fuzzy Classification of Online Customers. *Proceedings of the 8th International Conference on e-Commerce Technology, China*, 256–263. <https://doi.org/10.1109/ICEBE.2006.4>
- Wiener, N. (1948). *Cybernetics or control and communication in the animal and the machine*. MIT Press.
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, UK*, Article 38.
<http://dx.doi.org/10.1145/2601248.2601268>
- Youssef, A. B., Chollet, M., Jones, H., Sabouret, N., Pelachaud, C., & Ochs, M. (2015). Towards a socially adaptive virtual agent. In W.-P. Brinkman, J. Broekens, & D. Heylen (Eds.), *Lecture Notes in Computer Science: Vol. 9238. Intelligent Virtual Agents* (pp. 3–16). Springer. https://doi.org/10.1007/978-3-319-21996-7_1

- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.
[https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- Zadeh, L. A. (1982). Fuzzy systems theory: A framework for the analysis of humanistic systems. In R. Cavallo (Ed.), *Systems Methodology in Social Science Research, Frontiers in Systems Research* (2nd vol., pp. 31–56). Springer.https://doi.org/10.1007/978-94-017-3204-8_3
- Zadeh LA (1986): Outline of a computational approach to meaning and knowledge representation based on the concept of a generalized assignment statement. Proceeding of International Seminar on Artificial Intelligence and Man-Machine Systems. Edited by: Thoma M, Wyner A.
- Zadeh, L. A. (1988). Fuzzy logic. *IEEE Computer*, 21(4), 83–93.
- Zadeh, L. A. “Fuzzy logic = computing with words,” in Computing with Words in Information/Intelligent Systems 1, pp. 3–23, Springer, 1999. doi:10.1109/91.493904.
- Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The Behavioral Consequences of Service Quality. *Journal of Marketing*, 60(2).
<https://doi.org/10.1177/002224299606000203>
- Zimmermann, H.-J., & Zysno, P. (1980). Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4(1), 37–51.
- Zimmerman, Y. S., Döhnert, K., & Schreiber, F. (2021). *VorsorgeDIALOG: Financial literacy mit Fokus Altersvorsorge*. Institut für Finanzdienstleistungen Zug IFZ. <https://www.hslu.ch/de-ch/hochschule-luzern/ueber-uns/medien/medien-mitteilungen/2021/08/26/vorsorgestudie/>
- Zumstein, D. (2007). Customer performance measurement: Analysis of the benefit of a fuzzy classification approach in customer relationship [Master’s thesis, University of Fribourg]. ResearchGate. https://www.researchgate.net/publication/316169744_Customer_Performance_Measurement_-_Analysis_of_the_Benefit_of_a_Fuzzy_Classification_Approach_in_Customer_Relationship_Management
- Zumstein, D. (2026). Performance Management. In E. Portmann, G. Wilke, L. Terán & S. D’Onofrio (Eds.), *Fuzzy Sets and Systems II*. Springer.
https://doi.org/10.1007/978-3-031-99294-0_2

8. APPENDIX A

Sources for Use Case 1, chapter 3.1.

https://zenodo.org/records/19059735?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjExNjIhOTUzLTUyNDJlMmZmY2RhMTA1MDdjZCJ9.5MKR-UhQouL3LI0_KUEhQRQzgaDoXVlKjRlflAOcqDsZeCMb-BGL9pw7IoteMKzNNzFtwVXBszWrD473mP3aahg

9. APPENDIX B

Sources for Use Case 2, chapter 3.2.

https://zenodo.org/records/19061292?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6ImQ2N2MzMWl4LTg0MDYtNGVhZS05YzMwLT-kzMDE4MGRhNGM2NCIsImRhdGEiOnt9LCJyYW5kb20iOiI1ODRi-NjM2OWRiY2Q3ZjBkNDliMmQ4MmMwMjg3YzIwMyJ9.aOa9uv-IOWiVgqcwDuUivOADf99q-Be0G0s-3vKCvJYvRK-fwpQOF_C90QCvWv0EVABfvjaQJvMubYKG-R9hg6Q

10. APPENDIX C

Sources for Use Case 3, chapter 3.3

https://zenodo.org/records/19061324?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjE5NjRhYmNkLWQyMDItNDQ2MS05ZWlzlTY0MwYwNmNlMjFkZCIsImRhdGEiOnt9LCJyYW5kb20iOiJmOTM5YzI3ZmRhYjFiZTc2MWE1NTcyMmY5OGRkMDQ5ZCJ9.oSYroCf-E0Lq4jRuDFWHlGvLzCXeR-_S4BBpsXXMJpAc4X6Fc6KXXc05vXk4I8Pi-WEe_b3JE8VSo4_vGQyA_nQ.

11. APPENDIX D

Sources for Use Case 4, chapter 3.5

https://zenodo.org/records/19062437?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjMyZmI1ODdkLTljOWEtNDRIZC05NzNhLTNhYjA4OTBkNjBkZiIsImRhdGEiOnt9LCJyYW5kb20iOiJhOTY5NWVjMDc5OTc1NzkwOTQ4MWM0YjhmYzc1ZDdmYyJ9.ZUHdG1e7rSLRfvJ5K2wW9bZ1BYK_DFsPYPO6PUYPy4amuc5-8QMJeUNhWZiYLAHz1eFhYYKtk9A-inmVqOFdxg

12. APPENDIX E

Link to List of used literature:

https://zenodo.org/records/15208482?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6ImFjMGUzMjM1LTc3NGMtNGI4MS1iZjVmLTAxMjBiMGQ1NDAYZCI6ImRhdGEiOnt9LCJyYW5kb20iOiI5YjAyNjYxMDIxN2QzO-DIxNjJhZmIwNjgxNTM5MmJlNiJ9.y2SNdHO-BoRNs3tAcqZ_VOppGyLJ0jty8TA-E8w1RJ5iccIUEHiKAszATce6Ztd99ThkeAO-DoITx8MjbDrLwMjA

13. APPENDIX F

Results from first artefact evaluation, chapter 4.

https://zenodo.org/records/19062683?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjBiNjVkZWQ1LWM4ZDEtNDcyYS04Y2ZlLTU0MmNmODExNWE5ZiIsImRhdGEiOnt9LCJyYW5kb20iOiI0OTU0MTI5MTk2ZWU1ZGFhODAxNzNiNmZOTUyYzU0MCJ9.OW3YgpqvICi6JXfiRjn0P5V-dTPQqP05-50kYWzguSN2d3fP6C1xsuE84yfaMiCWrxGy5nMdQ-asjtnt2lw_Q

14. APPENDIX G

All sources from technical evaluation, chapter 5.1.

<https://zenodo.org/records/19062544?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjVkJjgwZjJiLTgzZGMtNDRkNC04ZDgxLTg0M2YxNTglYmVjNCIsIm-RhdGEiOnt9LCJyYW5kb20iOiJINWZjY2MwMzliN2FINGI2NTEzMGFhNTQyMzU0N2YwOCJ9.2faTVMFGlwsaWLzrBMXQuVA-WtYPBGBl1hdu-ZhD5IM9tLGFmyxFVRsi49o2YNHJ7Qxsamss6UpCgb77YudEFfA>