University of Brasília

Exact Sciences Institute
Computer Science Department

Gamified Chatbot Management Process: A way to build gamified chatbots

Arthur Rocha Temporim de Lacerda

Dissertation submitted in fulfillment of
the requirements to Professional Master's Degree in Applied Computing

Advisor
Prof. Dr. Sergio Antônio Andrade de Freitas

Brasília
2024

University of Brasília

Exact Sciences Institute
Computer Science Department

Gamified Chatbot Management Process: A way to build gamified chatbots

Arthur Rocha Temporim de Lacerda

Dissertation submitted in fulfillment of the requirements to Professional Master's Degree in Applied Computing

Prof. Dr. Sergio Antônio Andrade de Freitas (Advisor)
Universidade de Brasília

Prof. Dr. Giovanni Almeida Santos
Universidade de Brasília

Prof. Dr. Charles Andrye Gavão Madeira
Universidade Federal do Rio Grande do Norte

Prof. Dr. Gladston Luiz da Silva
Coordinator of Graduate Program in Applied Computing

May 28th, 2024

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Rocha Temporim de Lacerda, Arthur
131g Gamified Chatbot Management Process: A way to build gamified chatbots / Arthur Rocha Temporim de Lacerda; orientador Sergio Antônio Andrade de Freitas. -- Brasília, 2024.
166 p.

Dissertação(Mestrado Profissional em Computação Aplicada) -- Universidade de Brasília, 2024.

1. Gamification. 2. Chatbot. 3. GCMP. 4. GQM. 5. Development Process. I. Antônio Andrade de Freitas, Sergio, orient. II. Título.

Abstract

Chatbot development frameworks offer diverse construction methods, but established processes, like the Chatbot Management Process (CMP), lack activities specifically designed to boost user engagement. This thesis proposes the Gamified Chatbot Management Process (GCMP), an extension of the CMP that incorporates and adapts activities to enhance user engagement with the chatbot. Three versions of the GCMP were developed, each incorporating improvements guided by the Goal-Question-Metric (GQM) approach. This iterative approach facilitated the evaluation and evolution of the process. Real-user experiments demonstrated positive engagement, with 100% of participants achieving the proposed objectives. Additionally, the average deployment time decreased by 66% between the first and final versions. User evaluations were also awarded top marks for the quality of chatbot-generated responses. These findings highlight the effectiveness of the proposed GCMP. The results of the experiment suggest a positive correlation between the use of the GCMP and the improvement in the development of gamified chatbots. The observed improvements in both chatbot functionality and gamification techniques offer promising indicators for the widespread adoption of the GCMP as a robust and effective process for the development of gamified chatbots.

Keywords: Gamification, chatbot, GCMP, GQM, development process, user engagement, Rasa

Resumo

Processo de Gerenciamento de Chatbots Gamificados: Uma forma de construir chatbots gamificados.

Frameworks de desenvolvimento de chatbots oferecem diversos métodos de construção, mas processos estabelecidos, como o Chatbot Management Process (CMP), carecem de atividades especificamente projetadas para aumentar o engajamento do usuário. Esta dissertação propõe o Gamified Chatbot Management Process (GCMP), uma extensão do CMP que incorpora e adapta atividades para aprimorar o engajamento do usuário com o chatbot.

Três versões do GCMP foram desenvolvidas, cada uma incorporando melhorias orientadas pela abordagem Goal-Question-Metric (GQM). Essa abordagem iterativa facilitou a avaliação e evolução do processo.

Experimentos com usuários reais demonstraram um engajamento positivo, com 100% dos participantes alcançando os objetivos propostos. Além disso, o tempo médio de implantação diminuiu em 66% entre a primeira e a versão final. As avaliações dos usuários também receberam as melhores notas pela qualidade das respostas geradas pelo chatbot. Esses resultados destacam a eficácia do GCMP proposto.

Os resultados do experimento sugerem uma correlação positiva entre o uso do GCMP e a melhoria no desenvolvimento de chatbots gamificados. As melhorias observadas tanto na funcionalidade dos chatbots quanto nas técnicas de gamificação oferecem indicadores promissores para a adoção generalizada do GCMP como um processo robusto e eficaz para o desenvolvimento de chatbots gamificados.

Palavras-chave: Gamification, chatbot, GCMP, GQM, processo de desenvolvimento, engajamento do usuário, Rasa

Contents

0.1 Acknowledgments 1

1 Introduction **2**

1.1 Professional Contribution and Motivation 3

1.2 Research Question 4

1.3 Objectives 4

1.4 Methodology 4

1.4.1 Literature Review and Framework Development 4

1.4.2 Experimental Application and Evaluation 5

1.5 Dissertation Structure 6

2 Theoretical foundation **7**

2.1 Chatbot 7

2.1.1 Chatbot development 9

2.1.2 Chatbot application 12

2.2 Gamification 14

2.2.1 Gamification Development 15

2.2.2 Gamification Applications 25

2.3 GQM approach for validating software processes 28

2.4 Literature review 29

2.4.1 Literature review application 29

2.5 Papers discussion 30

2.5.1 CiboPoliBot 31

2.5.2 Voice-based Apps 31

2.5.3 ScratchThAI 32

2.5.4 Chatbot-Based assessment 34

2.5.5 Charlie 34

2.5.6 Chatbot-Based learning media 36

2.5.7 Intelligent recruitment with chatbot 36

2.5.8 Artificial intelligence in education 38

2.5.9 Prevention of obesity . . . 38
2.5.10 Chatbot for weight loss . . . 38
2.5.11 Playable Cities . . . 39
2.5.12 Xploro Digital . . . 40
2.5.13 Chatbot on financial . . . 40
2.5.14 Vote Goat . . . 41
2.5.15 Elena+ . . . 41
2.6 Literature review discussion . . . 42

3 Methodology 44
3.1 Gamified chatbot development process proposal . . . 44
3.1.1 Phase 1: Literature Review and existing solutions . . . 44
3.1.2 Phase 2: Evaluation and adaptation . . . 45
3.1.3 Phase 3: Integration and process development . . . 45
3.2 Experiment . . . 45
3.2.1 Experiment Design: Evaluating the process with developers . . . 45
3.2.2 Participant selection and onboarding . . . 46
3.2.3 Project management and data collection . . . 46
3.2.4 Data Analysis and recommendations . . . 47
3.2.5 GQM application . . . 47
3.3 Data collection and analysis . . . 48

4 The Gamified Chatbot Management Process (GCMP) 49
4.1 GCMP evolution . . . 50
4.1.1 First version: GCMP as CMP extension . . . 51
4.1.2 Second version: GCMP improved . . . 52
4.1.3 Third version: GCMP generalized . . . 54
4.2 GCMP . . . 55
4.2.1 GCMP generated by applying the experiment . . . 55
4.2.2 (A) Gamification . . . 56
4.2.3 (B) Build . . . 58
4.2.4 (C) Analyze and Delivery . . . 60
4.2.5 GCMP discussion . . . 62

5 Experiment 63
5.1 Experiment context . . . 63
5.1.1 Experiment team . . . 64
5.2 Experiment files and repositories . . . 64

5.3 Development Process Assessment 69
5.3.1 Planning 70
5.3.2 Definition 71
5.3.3 Data collection 76
5.3.4 Interpretation 76
5.3.5 Remarks 76

6 Data collect and analysis **78**
6.1 Introduction 78
6.2 Goal 1 - Analyze the chatbot generated by GCMP for the purpose of understanding with respect to assertiveness from the researcher's point of view in the context of the research laboratory 78
6.2.1 Q1. How assertive is the chatbot? 79
6.2.2 Q2. What is the evolution of chatbot content? 79
6.3 Goal 2 - Analyze the chatbot generated by GCMP for the purpose of understanding with respect to chatbot gamification from the researcher's point of view in the context of the research laboratory 80
6.3.1 Q3. What is the impact of gamification on user engagement? . . . 80
6.3.2 Q4. What is the evolution of gamification in the chatbot? 81
6.4 Goal 3 - Analyze the GCMP for the purpose of understanding with respect to quality from the researcher's point of view in the context of the research laboratory 82
6.4.1 Q5. What is the adherence to the process? 82
6.4.2 Q6. How is the chatbot evolving using GCMP? 83
6.4.3 Q7. What is the developer's level of focus regarding the scope, chatbot, and gamification each week? 84
6.4.4 Discussion 86

7 Conclusion **88**
7.1 Future Works 89

References **91**

Appendix **101**

I Literature review **102**

II GCMP manual **109**
II.1 Phases and steps description . . . 110
II.1.1 (A) Gamification . . . 110
II.1.2 (B) Build . . . 111
II.1.3 (C) Analyze and Delivery . . . 112

III GQM data **114**
III.1 Planning . . . 114
III.1.1 GQM Plan . . . 114
III.1.2 Measurement Plan . . . 116
III.2 Analysis Plan . . . 116

IV Questionaries **119**
IV.1 Team forms . . . 120
IV.1.1 Questionary . . . 120
IV.1.2 Questionaries responses . . . 126
IV.2 Users gamification profile . . . 132
IV.2.1 Questionary . . . 132
IV.2.2 Questionaries responses . . . 145
IV.3 Chatbot users usage . . . 152
IV.3.1 First phase . . . 152
IV.3.2 Second phase . . . 154

List of Figures

2.1 Chatbot project layers overview. 8
2.2 Chatbot Management Process. 10
2.3 EvaTalk's average confidence by month. 10
2.4 EvaTalk's training examples at the end of each month. 11
2.5 Proposed Chatbot Design. 13
2.6 FLOSS chatbot project characteristics. 13
2.7 Project overview - technologies, artifacts, and stakeholders. 14
2.8 Chatbot team with experts (a), (b), (c) and non-experts with project reuse (d), (e), (f). 15
2.9 The Octalysis Graphical Representation. 16
2.10 Gamification User Types Hexad. 20
2.11 5W2H Framework. 21
2.12 The Gamification Design Framework (GDF). 23
2.13 MARC gamification framework. 23
2.14 Block Diagram for the Methodology. 25
2.15 Core Drivers Averages per School Grades. 26
2.16 Octalysis Distribution for all the School Grades. 26
2.17 The Graph of Gap Analysis. 27
2.18 CiboPoli High-level Architecture. 31
2.19 An Architectural Design of ScratchThAI. 33
2.20 Screenshot of the mental health checkup. 35
2.21 Affinity diagram to sort cues that shape Charlie's personality. 36
2.22 Chatbot with Gamification Architecture. 37
2.23 Overview of the STOP platform. 39
2.24 System Architecture. 41
2.25 Architectural overview. 42
2.26 Elena+ conceptual model. 43

4.1 First GCMP version. 51
4.2 Second GCMP version. 53

4.3 Third GCMP version. 54
4.4 Gamified Chatbot Management Process (GCMP). 56

5.1 GitHub repositories. 65
5.2 GitHub latte-chatbot repository. 66
5.3 Latte chatbot conversation on Telegram. 67
5.4 Gamification board on GitHub project. 69
5.5 First goal of GQM. 72
5.6 Second goal of GQM. 73
5.7 Third goal of GQM. 74

6.1 M12. Comparison of the execution of activities expected versus performed. 83
6.2 M15, M16 and M17 - Focus on phase 1 of the experiment. 84
6.3 M15, M16 and M17 - Focus on phase 2 of the experiment. 85
6.4 M15, M16 and M17 - Focus on phase 3 of the experiment. 85

II.1 GCMP 109

III.1 GQM 118

List of Tables

2.1 Chatbot and Gamification Intersection literature review results 30

5.1 Metrics table. 77

6.1 Chatbot metrics by version. 79
6.2 M7 and M8. User interview and messages/user. 81
6.3 M9. Octalysis Core Drivers values by version. 81
6.4 M10. Gamification Techniques by version. 82
6.5 M13. Development time by each version 83

I.1 Searched articles 107
I.2 Selected articles 108

III.1 Measurement table 117

0.1 Acknowledgments

I dedicate this master's thesis, with immense gratitude, to God, for giving me strength and wisdom to complete this important cycle of my life.

I would also like to thank the love of my life, Clarice, for her unconditional support and kind words were essential for me to maintain perseverance in the most challenging moments.

To my advisor, Professor Doctor Sérgio, I express my most sincere thanks. From the beginning of my master's degree, you welcomed me and guided me along this path of learning and research. Finally,

I would like to thank the University of Brasília for providing me with the opportunity to pursue my master's degree.

This work is the result of effort and dedication. To everyone, my deepest and most sincere thanks.

Chapter 1

Introduction

The demarcation between traditional gaming frameworks and their applications in unconventional settings is increasingly indistinct, a shift attributed to the broad integration of gamification techniques. This evolution, as evidenced in scholarly work, involves the deliberate embedding of gaming elements and foundational design concepts into domains beyond mere entertainment [1], [2]. This blend not only underscores the significance of motivational design theories but also showcases the capacity of gamification to substantially boost engagement across diverse sectors. Such a transformation reflects a changing perspective on the role of games, extending their value to educational, healthcare, and commercial fields, thereby emphasizing their importance in fostering engagement and motivation.

Within the rapidly evolving digital ecosystem, chatbots, powered by artificial intelligence to create sophisticated conversational interfaces, have emerged as pivotal tools for crafting highly personalized user experiences. The true potential of these AI-driven interfaces, however, is argued to be fully realized when integrated with gamification principles, creating a synergy that amplifies user engagement and interaction. This combination is explored in the design of social chatbots with gamification for user profiling and smoking trigger detection, emphasizing the effectiveness of gamification in enhancing chatbot interactions [3].

This scholarly inquiry embarks upon an explorative analysis of gamified chatbots, with a focus on identifying and evaluating methodologies to effectively leverage gaming elements to enhance user interaction quality and chatbot operational efficiency. The discourse navigates through the established Chatbot Management Process (CMP) [4], which provides a delineated framework for the systematic creation of chatbots. This examination further extends to the introduction of the Gamified Chatbot Management Process (GCMP), a novel innovation meticulously designed to facilitate the seamless integration of gamification elements into the developmental lifecycle of chatbot technologies.

The academic endeavor herein not only seeks to elucidate the theoretical underpinnings and practical applications of incorporating game design principles into chatbot development but also aims to contribute to the extant body of knowledge by offering a comprehensive framework that amalgamates the strategic objectives of gamification with the technological advancements inherent in AI-driven conversational interfaces. Through this inquiry, the GCMP is posited as a pioneering methodology, poised to redefine the conventional approaches to chatbot development and, by extension, enhance the interactive digital experiences afforded to end-users.

1.1 Professional Contribution and Motivation

With experience in a software consultancy company, catering to a diverse global clientele, has honed my expertise in crafting conversational assistants for customer service using a spectrum of chatbot frameworks. Despite the criticality of engaging chatbot content, a notable void in standardization across the development process has been observed, leading to diminished user engagement with even the most meticulously designed chatbots.

During the development of chatbots, it became evident that a chatbot's engagement level does not always correlate with positive user feedback. This discrepancy highlights the need for tools specifically designed to enhance user engagement. Moreover, the quality of the user experience significantly influences engagement, underscoring the necessity of implementing a comprehensive and effective chatbot solution. The objective is to enhance overall engagement through strategic improvements in both the chatbot's design and user experience.

This realization ignited a fervor to bridge this gap, culminating in my master's thesis on the Gamified Chatbot Management Process (GCMP)—a methodology poised to augment chatbot development user engagement. Through the application of GCMP's standardized procedures, developers are empowered to achieve desired outcomes, enhancing their capability to engage users effectively.

While the fusion of gamification with chatbots is not uncharted, illustrated by instances like CiboPoliBot, Charlie, and Elena+ [5], [5], [6] which underscore their potential, a conspicuous absence of standardized processes and methodologies for gamified chatbot development persists. Among these exemplars, solely CiboPoliBot integrates a gamification framework during its development phase, whereas other endeavors tend to superficially embed gamification elements, often culminating in suboptimal solutions. Thus, the GCMP endeavors to fill this void in gamified chatbot development methodologies.

1.2 Research Question

This study examines the outcomes of integrating gamification strategies into the chatbot development workflow. It seeks to understand the impact of incorporating gamification elements on user engagement, motivation, and the general effectiveness of chatbots. Through an analysis of the effects of a gamified methodology, this research aims to offer significant findings to the field of chatbot creation, marking the advent of more engaging and user-focused chatbot interactions. The core question driving this investigation is: What outcomes are derived from applying a gamified chatbot development process?

1.3 Objectives

The main objective of this research is to formulate and assess a gamified chatbot development process, with the following specific aims:

1. Propose a process for gamified chatbot development.
2. Validate the proposed process's effectiveness through controlled experimentation.
3. Assess the implementation of the proposed development process.
4. Evaluate the process's impact on application.

1.4 Methodology

The methodology employed in this study is delineated to systematically address the research question and achieve the outlined objectives. This methodology is structured to ensure a rigorous and replicable approach to the development and evaluation of the Gamified Chatbot Management Process (GCMP). The following steps comprise the core of our methodological framework:

1.4.1 Literature Review and Framework Development

1. **Comprehensive Literature Review**: A review of existing literature is conducted to map the current state of chatbot development processes and the integration of gamification. This review spans across multidisciplinary fields to gather on best practices, challenges, and gaps in the current methodologies.
2. **GCMP Development**: Drawing on from the literature review, the GCMP is developed as a structured process for integrating gamification into chatbot development.

This process is designed to be iterative and flexible, accommodating various types of chatbot projects and gamification strategies.

1.4.2 Experimental Application and Evaluation

3. **Prototype Development**: Utilizing the GCMP, a prototype gamified chatbot is developed. This phase not only involves applying the prescribed activities and guidelines within the GCMP but also designing the experiment in accordance with the Goal Question Metric (GQM) approach [7].

4. **Controlled Experimentation**: To validate GCMP, the prototype undergoes controlled experimentation, structured around the GQM approach. This phase is planned to observe, measure, and analyze user interaction with the gamified chatbot, with each experimental variable and outcome mapped against the goals and questions defined by the GQM framework. This structured experimentation aids in capturing relevant data for assessing the impact of gamification principles applied during chatbot development.

5. **Data Collection and Analysis**: Data concerning user engagement, interaction quality, and chatbot performance are collected throughout the experimentation phase, employing the Goal Question Metric (GQM) [7] approach to systematically evaluate the efficacy of gamification in chatbot development. This analytical framework allows for the assessment of the impact of gamification on chatbot effectiveness and user experience by defining specific goals, formulating relevant questions, and identifying appropriate metrics for measurement. The data garnered through this analysis inform the iterative refinement of the GCMP, ensuring its alignment with empirical evidence and user feedback.

6. **Feedback Loop**: Observations and feedback from the experimental application are utilized to refine and adjust the GCMP. This feedback loop ensures that the process is continuously improved based on empirical evidence and user responses.

7. **Theoretical Contribution**: The findings from the application and evaluation of the GCMP are contextualized within the broader theoretical frameworks of gamification and chatbot development. This step aims to contribute to the academic discourse by providing validated insights into the integration of gamification in chatbots.

8. **Practical Guidelines**: Based on the research findings, practical guidelines for the application of the GCMP in various contexts are developed. These guidelines aim to

assist practitioners in leveraging gamification to enhance chatbot projects, ensuring that the process is accessible and actionable.

1.5 Dissertation Structure

This dissertation is organized into seven chapters:

1. **Introduction**: Establishes the research motivation, objectives, and outlines the dissertation's scope, setting the stage for an investigation into gamified chatbot development.

2. **Theoretical Foundation**: Explores the conceptual underpinnings of gamification and chatbots, providing a critical review of relevant literature and defining key terms. 3.

3. **Methodology**: Describes the research design and methodological approach, detailing the procedures for developing and evaluating the Gamified Chatbot Management Process (GCMP).

4. **The Gamified Chatbot Management Process (GCMP)**: Introduces the GCMP, explaining its structure, components, and the rationale behind its design for integrating gamification into chatbot development.

5. **Experiment**: Details the application of the GCMP in a real-world setting, including the development process, deployment, and user interaction with the gamified chatbot prototype.

6. **Data Collection and Analysis**: Discusses the methods used for collecting and analyzing data from the experimental phase, focusing on user engagement, feedback, and the effectiveness of the GCMP.

7. **Conclusion**: Summarizes the research findings, evaluates the success of the GCMP against the stated objectives, and offers recommendations for future research and practical applications of gamified chatbots.

Chapter 2

Theoretical foundation

This chapter lays the groundwork about the research by exploring the theoretical foundations upon which it is built. We delve into two key areas: gamification and chatbots. The first section delves into the world of chatbots, analyzing their capabilities and potential impact on user experiences. The second section examines the core concepts and frameworks surrounding gamification, exploring how these elements can be harnessed to motivate and engage users. The third section bridges these two disciplines, synthesizing existing research on the intersection of gamification and chatbots. This comprehensive review will inform the development of our gamified chatbot development process, ensuring it leverages the power of gamification to enhance chatbot effectiveness.

2.1 Chatbot

In the last century, the question "Can machine think?" by Turing in his so-called imitation game' which challenges someone in a dialogue between 3 actors if an actual human or a computer is providing the answer. At that time, talking with a machine was a very tricky thing, but nowadays, it is getting hard to keep away from chatting with robots.

Until recent years, the term chatbot was not widely known, but this technology that simulates conversations has been growing [9]. Even though it appears to be a current concept, its history is ancient. The origin comes from the creation of Eliza and its use in the Doctor program [10]. Eliza is a "family of programs" created in 1966 by Weizenbaum, which searches for keywords and, when it finds them, responds to the sentence according to rules associated with a keyword script. The Doctor was the first program to use Eliza to imitate a psychiatrist.

The definition of a chatbot is a computer program designed to simulate conversation with human users, especially over the internet [11].

There are examples of chatbots like Siri, Apple's virtual assistant capable of interacting with voice commands, Microsoft's Cortana, and Google Assistant [12]. These are great examples, and part of their success is due to using speech as a way of interaction [13].

Other chatbots are also applied to different contexts, which shows that their versatility can be explored. An example is the AgronomoBot, an intelligent chatbot applied to agriculture, which was developed in a partnership between USP and IFBA to seek and present data collected from wireless sensors implanted in a vineyard [14]. Another example is the study by Dutta [15], which shows the viability of using chatbots in information security. The study points out that it is possible to use conversational agents to inform more about the subject.

It is possible to generalize chatbot solutions in two parts: Natural Language Understand (NLU) and Natural Language Generator (NLG), as shown in Figure 2.1.

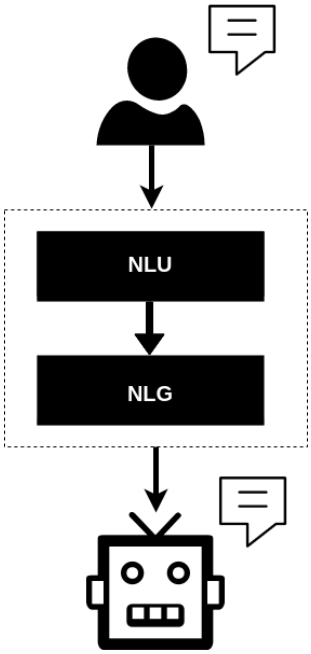


Figure 2.1: Chatbot project layers overview (Source: [16]).

1. **Natural Language Understand**

   A chatbot's natural language understanding (NLU) layer is responsible for understanding the meaning of user input. The NLU layer typically uses a combination of machine learning and rule-based techniques to understand user input. Machine learning techniques allow the NLU layer to learn from data, such as example conversations, to improve its ability to understand the meaning of user input. Rule-based techniques allow the NLU layer to understand user input based on pre-defined rules, such as grammar and semantics.

The NLU layer is typically implemented as a neural network, a type of machine learning model well-suited for tasks such as text classification and intent detection. The neural network is trained on a large dataset of text and code, allowing it to learn human language patterns.

2. **Natural Language Generator**

   A chatbot's natural language generator (NLG) layer generates human-like text responses to user input. It is the inverse of the natural language understanding (NLU) layer, which is responsible for understanding the meaning of user input.

   The NLG layer typically uses a combination of machine learning and rule-based techniques to generate text. Machine learning techniques allow the NLG layer to learn from data, such as example conversations, to improve its ability to create natural and engaging text. Rule-based techniques allow the NLG layer to generate text based on pre-defined rules, such as grammar and style rules.

   The NLG layer is typically implemented as a neural network, a type of machine learning model well-suited for tasks such as text generation. The neural network is trained on a large dataset of text and code, allowing it to learn human language patterns.

   Once the NLG layer is trained, it can generate text responses to user input. The NLG layer takes the output of the NLU layer, which is the intent and entities of the user input and generates a text response that is consistent with the user's intent and satisfies the entities.

   The NLG layer is a critical component of chatbots, as it generates the text responses that users see and interact with. A well-designed NLG layer can generate natural, engaging, and informative text. This can help to improve the user experience and make chatbots more effective in achieving their goals.

Is important to say that are more solutions to both layers, and each chatbot can implement both layers with many frameworks, it is even possible to combine different software in each layer to implement a conversational agent.

2.1.1 Chatbot development

The chatbot development context is an area that is still in progress, with frameworks, processes, and tools in the creation and testing phases. A chatbot development process broadly used by the industry is not a reality now.

Chatbot Management Process (CMP)

A Conversation-Driven Approach for Chatbot Management [4] Presents a chatbot development process for some machine learning solutions. The study says much information can impact the framework choice, like project scope, complexity, team knowledge and expertise, and time.

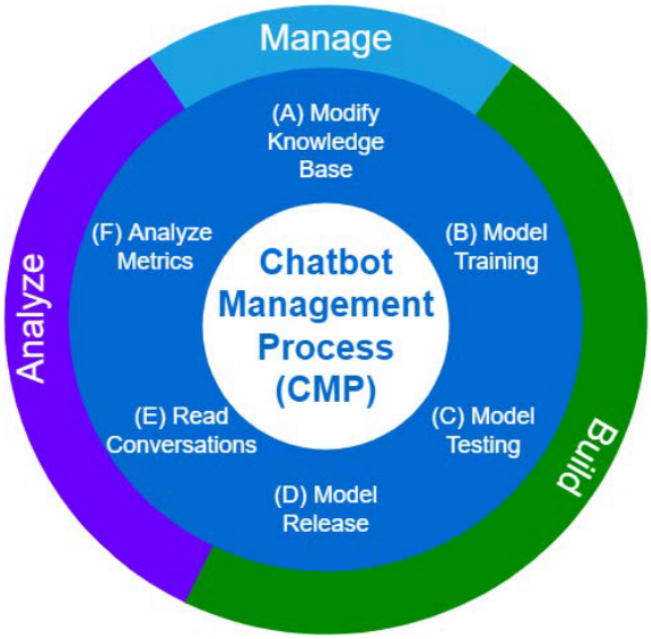


Figure 2.2: Chatbot Management Process (Source: [4]).

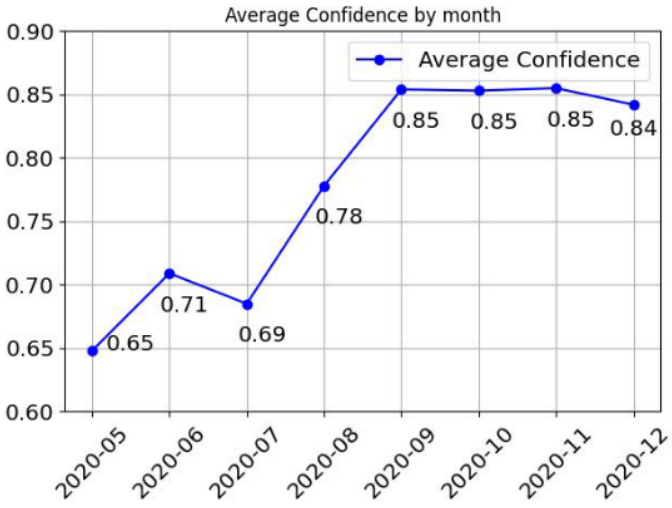


Figure 2.3: EvaTalk's average confidence by month (Source: [4]).

Figure 2.4: EvaTalk's training examples at the end of each month (Source: [4]).

The Chatbot Management Process (CMP) has 3 phases and 6 activities that aim to solve the lack of patterns in the chatbot development process, and this information is presented in Figure 2.2. This process was validated with robust data submitted with the application of the process in an educational chatbot project called EvaTalk. The study shows that with the application of CMP, the chatbot accuracy and knowledge base increase shown in Figure 2.3 and Figure 2.4.

CMP is a good proposal with valuable validation. However, like the authors shown, it still has some flaws, like the process planned for a specific kind of chatbot solution, the chatbots based on machine learning, and the activities that cannot be applied in all chatbot contexts, like the read conversation activity that is not possible to do in a chatbot that has a significant message volume.

Conversation Driven Development (CDD)

CDD is the chatbot development process to analyze the user's use and create insights with this information to improve the AI assistant. Rasa, one AI chatbot framework, created this approach [17].

CDD is divided into six actions:

1. **Share** the chatbot solution to the user through the defined message application every time that new content or behavior is needed to be added to the chatbot;

2. **Review** analyzes the user conversation data aiming to understand the user behavior and creating insights to fit the user needs better;

3. **Annotate** messages and select user interactions that can be used as training data aiming to improve the chatbot model;

4. **Test** the chatbot answers and behavior to ensure that the changes made had the desired result;

5. **Track** the assistant behavior and fails, measuring its performance;

6. **Fix** how the assistant handles wrong situations;

CDD is an AI chatbot development process with actions like Annotate that do not fit every chatbot framework.

2.1.2 Chatbot application

Chatbot solutions, in general, are easy to access and use due to the conversion of an objective to a chat structure. This solution is applied to contexts like financial, education, health, and others. There are some processes and guidelines aiming to achieve reasonable solutions. One option that stands out is the Chatbot Management Process, CMP, which has defined activities aiming to have a chatbot solution as an outcome.

Chatbot with emotional recognition

The article presented aims to develop an intelligent chatbot using natural language processing and Telegram API. One emotion recognition layer was performed on the recorded chats, too. In Figure 2.5 is shown the proposed chatbot design[18].

This design is a vague chatbot development process; the phases have many sub-activities, and the authors do not describe each activity.

This work contributes to the application of emotion recognition, and the chatbot has two parts: emotion recognition and emotional responses. This approach could be applied in gamified chatbots, switching the emotion layer to a gamification layer or a third gamification layer.

FLOSS FAQ Chatbot

According to the study presented, FAQ chatbot requirements are similar and domain-specific, and projects can benefit from the reuse of Open-Source Software (OSS). This article presents how OSS FAQ chatbot projects can benefit from reuse at the project level

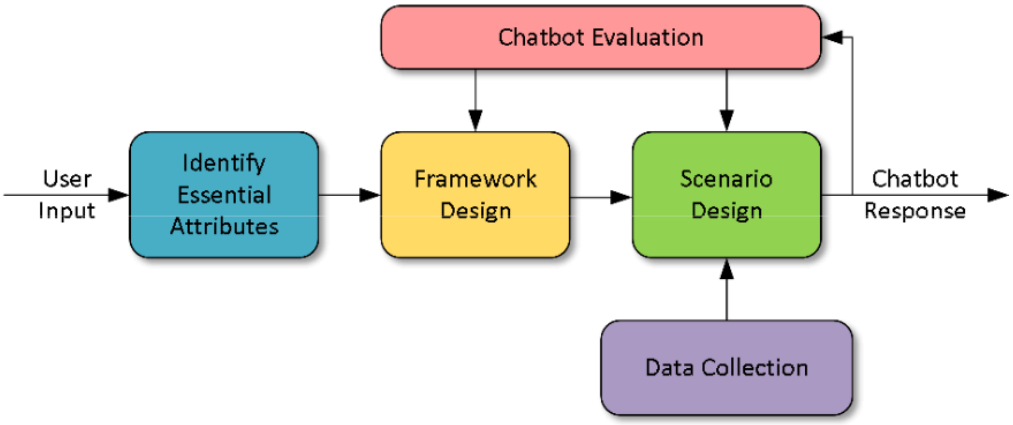


Figure 2.5: Proposed Chatbot Design (Source: [18]).

(black box reuse). A case study of a FLOSS FAQ chatbot project developed in Portuguese for an e-government service in Brazil is presented[16].

Team Members from all needed expertise	14 members
Releases in the last 12 months	14 releases
Number of Intentions	72 annotated intents
Number of Utters	147 utters
Original Size of the FAQ	35 questions

Figure 2.6: FLOSS chatbot project characteristics (Source: [16]).

Figure 2.6 presents the characteristics of the FLOSS chatbot project, showing a 1-year project with 14 releases made in total. It also shows that 72 intents and 147 responses, called "utters", were created for the chatbot from 35 FAQ questions.

This article discusses how automation, pre-configuration, and templates can help beginners develop chatbots in Portuguese without the need for specialized skills required by chatbot architecture tools. Figure 2.7 presents an overview of the project with the relationship between technologies, artifacts, and stakeholders. This architecture is provided through a code repository, where a developer can start a new chatbot project with this ready-made project base.

While the boilerplate project simplifies development, it's important to remember the importance of having different profiles for building a chatbot. Figure 2.8 showcases the various team roles needed. Traditionally, a "Chatbot Team" requires developers, DevOps, UX specialists, and data scientists. However, the use of the boilerplate base project can

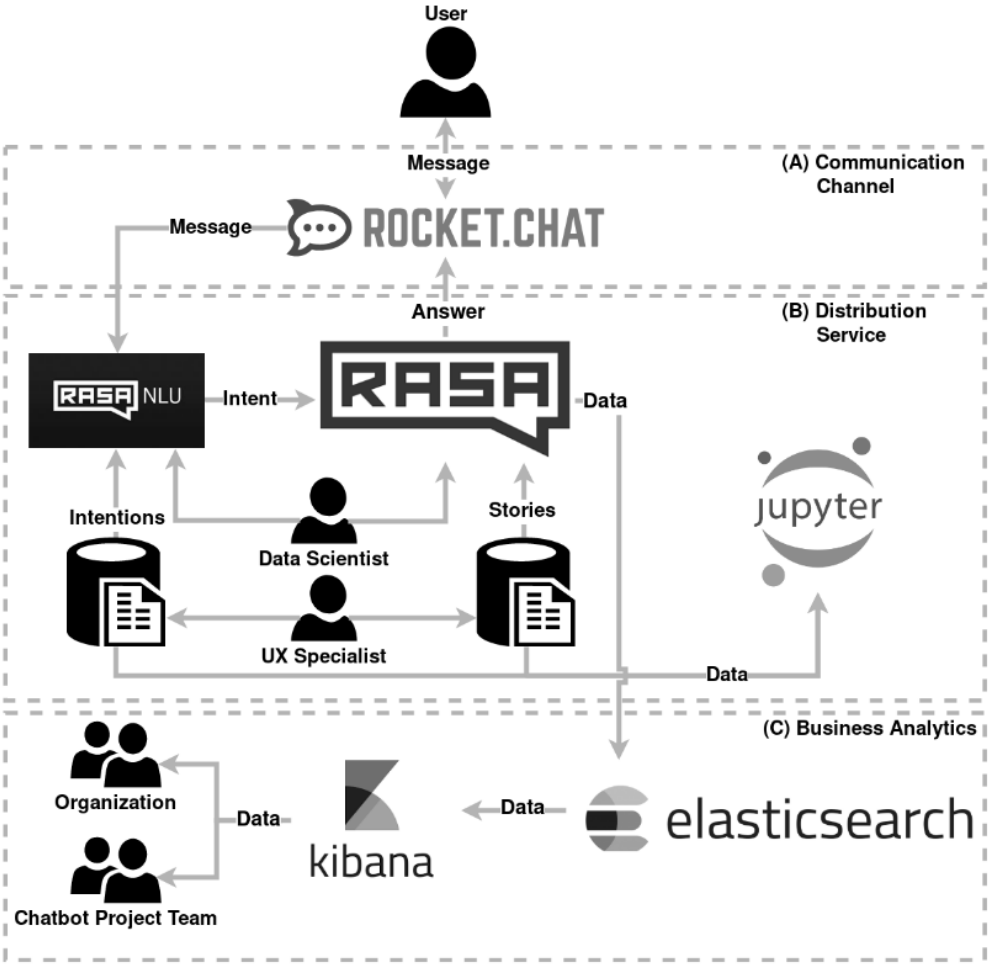


Figure 2.7: Project overview - technologies, artifacts, and stakeholders (Source: [16]).

streamline the process. These pre-built components can replace the need for dedicated DevOps and Data Scientist roles, allowing beginners to focus on core development tasks with just developer expertise.

2.2 Gamification

Systems are increasingly looking for a motivational design to engage users toward the task they are attempting to accomplish through the use of an engaged system [19]. It was predicted that most organizations would eventually implement some form of motivational design into their systems [20].

Gamification transforms activities, systems, services, products, or organizational structures to afford gameful experiences [21]. It can also be described as an informal umbrella term for using video game elements in non-gaming systems to improve user experience

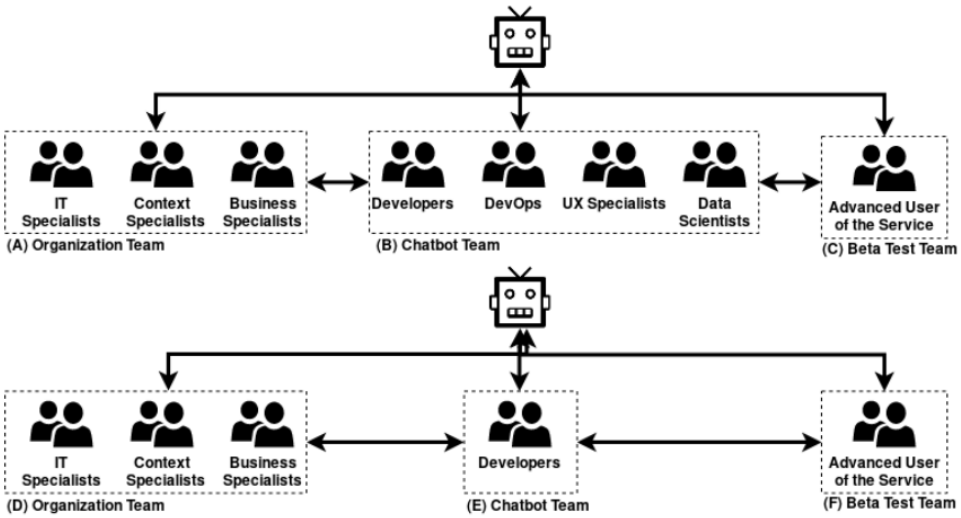


Figure 2.8: Chatbot team with experts (a), (b), (c) and non-experts with project reuse (d), (e), (f) (Source: [16]).

[22]. Even for gamification to enhance engagement, motivational design is challenging to implement as it requires the command of several disciplines, such as psychology and game design, beyond software development [23].

Gamification is one of the most challenging areas of software engineering; designing it requires the command of disciplines such as psychology, game design, and narratology, making its development difficult for traditional software developers. A study was made to understand what is needed to develop a gamified software solution [24]. The method has eight main activities: Project preparation, context analysis, user analysis, ideation, design, implementation, evaluation, and monitoring. When applying this method, the goal is to create a gamification project.

2.2.1 Gamification Development

In gamification, frameworks, processes, and guidelines exist to help its application. Some of them are focused on specific contexts. In contrast, others are more generalist, but all aim to achieve a gamification project. The proper application of gamification can raise user engagement, while a bad application can have the opposite effect. Choosing the proper framework to start a gamification project is important and can impact the final solution. The following sections present gamification frameworks:

Octalysis

The Octalysis Framework, developed by Yu-kai Chou, is a popular gamification framework that identifies eight core drives that motivate human behavior [25]. These core drives are categorized into four distinct sections, as illustrated in Figure 2.9. The core drives divided into four quadrants: the left side is associated with more logical decisions, the right side tends to intrinsic motivations, and the white hat is motivation elements that make us feel powerful. The black hat makes us feel obsessed, anxious, and addicted [26].

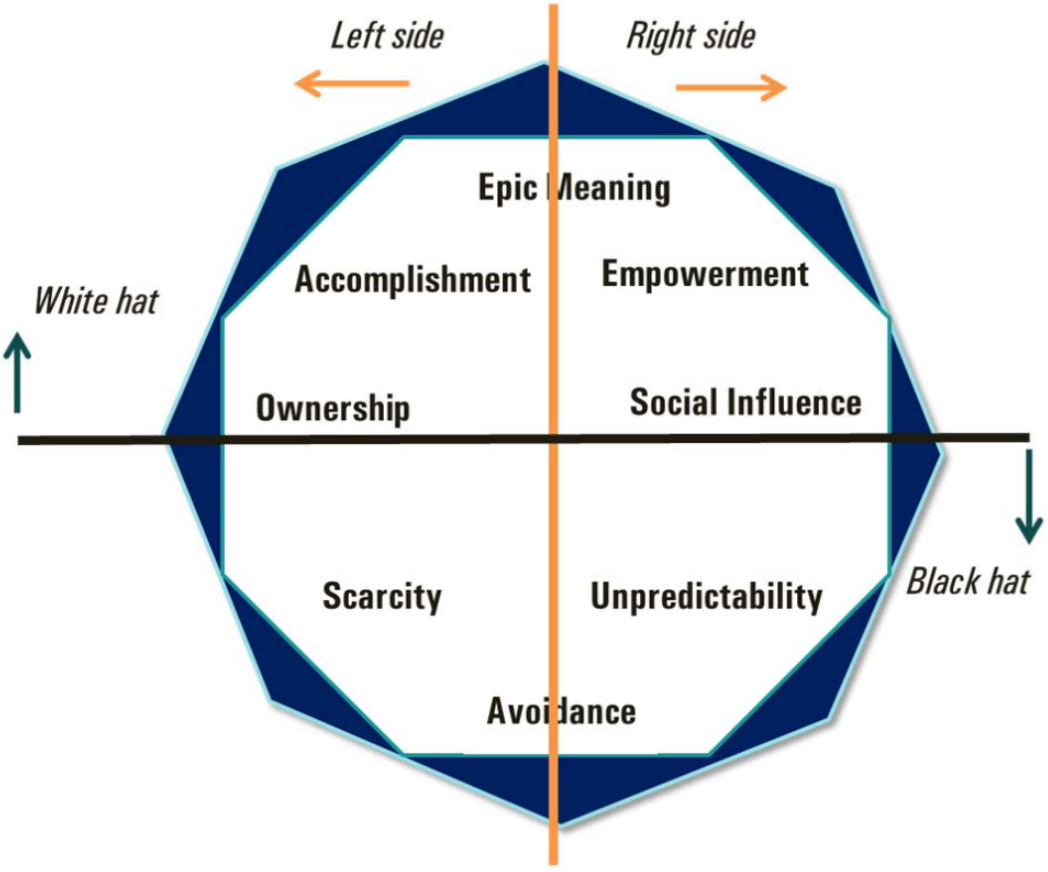


Figure 2.9: The Octalysis Graphical Representation (Source: [26]).

The eight core drives of the Octalysis Framework can be further explained as follows:

1. **Epic Meaning & Calling**: when the player believes he is doing something more remarkable for a greater good or has been chosen to do something transcendental.

2. **Accomplishment & Development**: when the player observes their progress, skill development, and overcoming challenges.

3. **Empowerment of Creativity & Feedback**: when the player is involved in a creative process where she repeatedly has to discover things and try different combinations.

4. **Ownership & Possession**: when the player is motivated because she has a sense of ownership or ownership of something.

5. **Social Influence & Relatedness**: when the player is motivated by social elements influencing people, including orientation, acceptance, social responses, companionship, competition, and envy.

6. **Scarcity & Impatience**: when the player is motivated by the desire for something they cannot have.

7. **Unpredictability & Curiosity**: when the players are motivated by wanting to figure out what will happen next. If they do not know what will happen, their brain is involved and often thinks about it.

8. **Avoidance & Loss**: when the player is motivated by the prevention of something negative that may occur.

While the eight core drives of the Octalysis Framework provide a foundation for understanding user motivation, the framework goes a step further by introducing the concept of 4 Experience Phases. These phases map out the user's journey within a gamified experience, highlighting how different core drives can be applied at each stage.

1. **Discovery**: this is where users initially encounter the gamified experience. The goal here is to spark interest and attract them to participate. Appealing to core drives like Epic Meaning & Calling (highlighting a greater purpose) or Unpredictability & Curiosity (teasing hidden features) can be effective strategies during this phase.

2. **Onboarding**: here, the focus shifts towards getting them started. Clear explanations of the rules, core mechanics, and how to participate. Leveraging drives like Accomplishment & Development (showing a clear path for progress) and Ownership & Possession (introducing initial rewards or virtual assets) can help incentivize users to take the first steps.

3. **Scaffolding**: this is where users actively engage with the gamified experience. The key here is to keep them motivated and provide a sense of continuous progress. Different core drives can be employed depending on the specific design. For instance, Empowerment of Creativity & Feedback can encourage experimentation, while Social Influence & Relatedness can foster competition or collaboration. Introducing

limited-time events can tap into the Scarcity & Impatience drive, creating a sense of urgency.

4. **Endgame**: this deals with what happens when users reach the end of the initial experience or achieve a significant milestone. The objective here is to encourage continued engagement and potentially guide users towards a new challenge. Appealing to drives like Epic Meaning & Calling (introducing a new overarching goal) or Avoidance & Loss (emphasizing the benefits of maintaining progress) can be helpful strategies during this phase.

By understanding these phases and how they connect with the core drives, designers can create a more holistic gamified experience. This ensures users stay engaged over time, experiencing a well-rounded journey within the gamified system. It's important to remember that the four Experience Phases are not always linear. Some experiences might have overlapping phases or revisit them with new twists as users progress. The key takeaway is to understand the user's journey and leverage the Octalysis Framework strategically at each stage to maximize its effectiveness.

GAMIFY-SN

The GAMIFY-SN meta-model presents a comprehensive approach to gamification within social networks, addressing the growing need for effective user engagement and behavior motivation in various domains such as education, healthcare, and marketing. Unlike rigid tools or techniques, GAMIFY-SN functions as a flexible framework, offering a conceptual structure to guide the thoughtful design and implementation of gamification initiatives. Specifically tailored for social networks, it harnesses the intrinsic power of social interaction and collaboration to amplify the impact of gamification. The process unfolds in four phases: Planning, Deployment, Monitoring, and Evaluation [27].

In the Planning phase, organizations define their target audience, articulate clear goals, and strategically select gamification elements aligning with objectives and user demographics. The Deployment phase involves seamlessly integrating chosen game mechanics into the social network platform, establishing rules and rewards, and launching engaging campaigns to stimulate user participation. During Monitoring, organizations track user activity, analyze performance data, and make real-time adjustments to optimize the gamification strategy. Finally, the Evaluation phase involves a comprehensive assessment of the initiative's impact, measuring outcomes through quantitative and qualitative data to draw informed conclusions and identify areas for improvement.

Key benefits of adopting GAMIFY-SN include its structured approach, providing organizations with a clear roadmap throughout the gamification lifecycle. By reducing

complexity, the framework facilitates easier project management and implementation, ultimately leading to improved effectiveness in achieving desired outcomes. GAMIFY-SN integrates seamlessly with existing platforms, leveraging social network dynamics, enhancing the overall gamification experience. In conclusion, the GAMIFY-SN meta-model emerges as a valuable resource for organizations seeking to implement gamification successfully within social networks, promising heightened user engagement, goal attainment, and a more gratifying online experience through its structured and socially integrated approach.

Hexad

The Hexad Gamification Framework is a model used to understand user motivation in gamified systems. It categorizes users based on six different personality types and the motivations that drive them to engage with gamified elements [28].

It identifies six core user types:

1. **Socializers**: Motivated by relatedness and connection with others. They enjoy social interaction and collaboration within gamified systems.

2. **Achievers**: Motivated by mastery and a sense of accomplishment. They enjoy challenges, goals, and progressing through levels.

3. **Philanthropists**: Motivated by purpose and a desire to make a positive impact. They enjoy contributing to a cause or helping others through gamification.

4. **Free Spirits**: Motivated by autonomy and self-expression. They enjoy creative freedom and the ability to personalize their experience within a gamified system.

5. **Players**: Motivated by rewards and recognition. They enjoy points, badges, leaderboards, and other extrinsic motivators.

6. **Disruptors**: Motivated by change and a desire to break the rules. They enjoy pushing boundaries and exploring unconventional approaches within gamified systems.

Each user type is driven by specific motivations. The framework helps developers understand what motivates different users and tailor the gamification elements accordingly. The Figure 2.10 shows how HEXAD is organized.

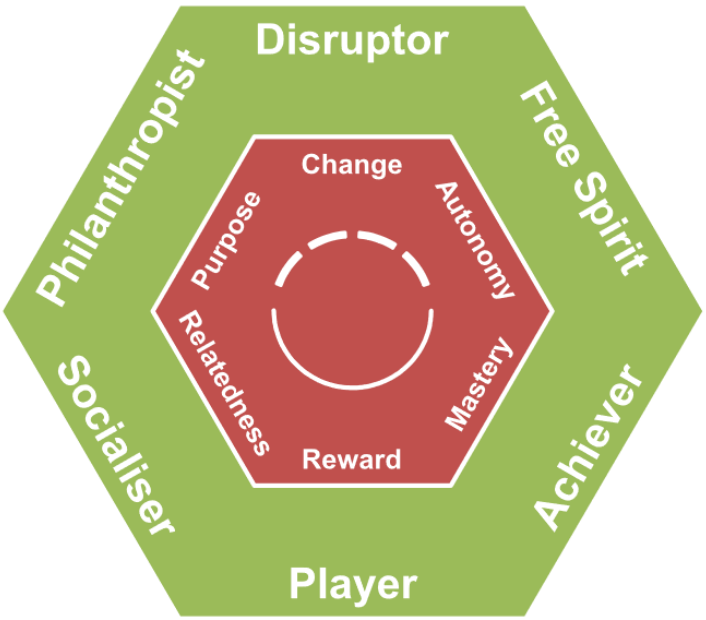


Figure 2.10: Gamification User Types Hexad (Source: [28]).

5W2H

The 5W2H framework is a helpful tool for planning and implementing gamification initiatives. It helps you clearly define who, what, where, when, why, and how your gamification efforts will occur [29]. The Figure 2.11 shows how each dimension is organized.

1. **Who?** Who is the target audience? This could be employees, customers, students, or any other group you want to engage. Who will be responsible for developing and implementing the gamification initiative? This may involve a team of designers, developers, and marketing professionals. Who will be involved in monitoring and evaluating the initiative? This might include stakeholders, users, and analysts.

2. **What?** What are the specific goals of the gamification initiative? This could be anything from increasing engagement to improving performance or changing behavior. What game mechanics will be used? Examples include points, badges, leaderboards, challenges, and rewards. What platform or technology will be used to implement the gamification initiative? This could be a dedicated app, a website, or an existing platform like a learning management system.

3. **Why?** Why are you implementing a gamification initiative? Be clear about the desired outcomes and how gamification can help achieve them. Why did you choose specific game mechanics and rewards? Ensure they are aligned with your target

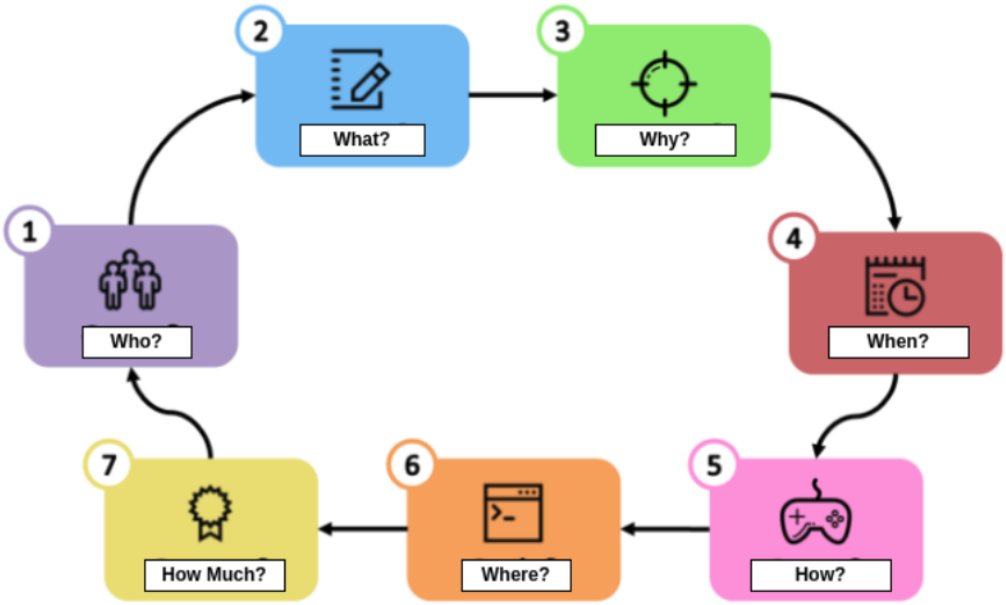


Figure 2.11: 5W2H Framework (Source: [29]).

audience and goals. Why is it important to monitor and evaluate the initiative? This helps you understand its impact and make necessary adjustments.

4. **When?** When will the gamification initiative launch? Setting a clear timeline for development, testing, and launch is important. When will different phases of the initiative occur? This could involve specific milestones for onboarding users, launching challenges, and distributing rewards. When will the initiative be evaluated? Regular monitoring and adjustments for success.

5. **How?** How will you develop and implement the gamification initiative? This includes defining processes, assigning roles, and setting budgets. How will users access and participate in the gamification experience? Provide clear instructions and make it easy for users to engage. How will data be collected and used to improve the initiative? Establish a system for data collection, analysis, and reporting.

6. **Where?** Where will the gamification initiative take place? This could be online, offline, or a combination of both. Where will users access the game mechanics and rewards? This could be through a specific portal, embedded within existing systems, or via mobile devices. Where will data be collected and analyzed? This may involve a dedicated data management platform or integration with existing systems.

7. **How much?** How will you develop and implement the gamification initiative? This includes defining processes, assigning roles, and setting budgets. How will users

access and participate in the gamification experience? Provide clear instructions and make it easy for users to engage. How will data be collected and used to improve the initiative? Establish a system for data collection, analysis, and reporting.

The benefit of using the 5W2H framework for gamification is improved planning and execution, which helps ensure a clear roadmap for your gamification initiative. Increased clarity and communication that facilitates communication and collaboration among stakeholders. Enhanced effectiveness: Helps you design and implement a gamification initiative more likely to achieve your desired outcomes. By applying the 5W2H framework, you can develop and launch a successful gamification initiative that engages your target audience and helps you achieve your goals.

GDF

Gamification Design Framework (GDF). It is a model-driven engineering framework for designing and implementing gamified applications. Its key features are Model-driven: GDF uses models to represent the various components of a gamified application, such as game mechanics, rules, and rewards. This makes it easier to design and implement complex gamification systems. Multi-level modeling: GDF provides a multi-level modeling approach, allowing different levels of abstraction to be used at various stages of the design process. Automatic code generation: GDF can generate code automatically from the models, saving time and effort during development. Flexibility and customization: GDF is flexible and can be customized to meet the specific needs of different applications [30] Figure 2.12 shows each GDF layer.

An application of GFD is education: GDF can be used to create gamified learning experiences that are more engaging and effective. Healthcare: GDF can be used to develop gamified applications for behavior change and health promotion. Marketing: GDF can be used to create gamified marketing campaigns that are more engaging and effective. Employee engagement: GDF can be used to develop gamified applications that can improve employee engagement and motivation.

The benefits of using GDF are reduced development time and cost: GDF can help to reduce the time and cost of developing gamified applications. Improved quality: GDF can help to improve the quality of gamified applications by making them more consistent and reliable. Increased effectiveness: GDF can help to increase the efficacy of gamified applications by making them more engaging and motivating for users.

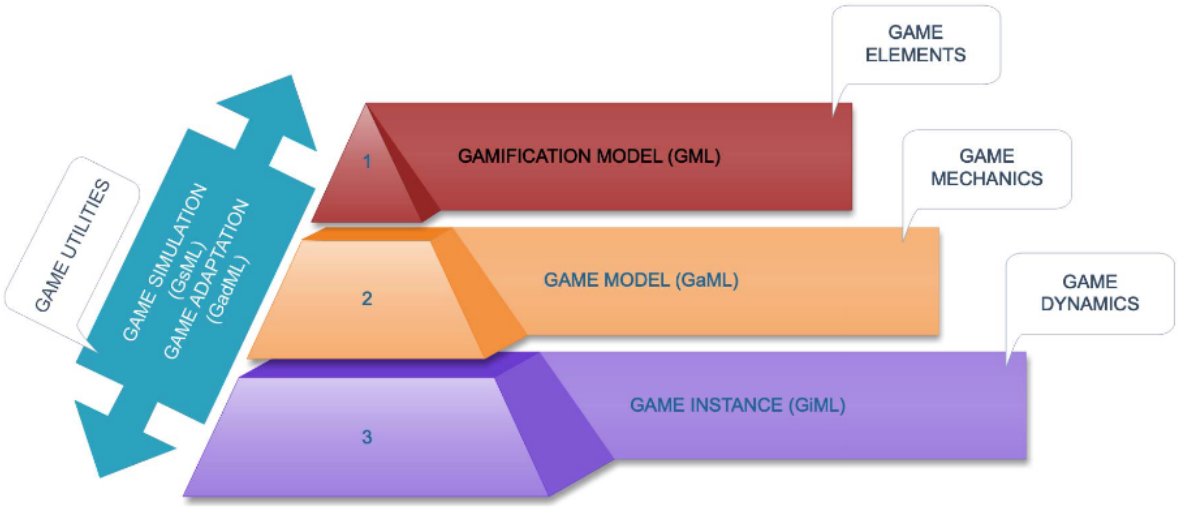


Figure 2.12: The Gamification Design Framework (GDF) (Source: [30]).

MARC

The MARC Gamification Framework. It is a framework designed explicitly for gamifying Massive Open Online Courses (MOOCs) to enhance student engagement, motivation, and learning outcomes [31] Figure 2.13 shows MARC.

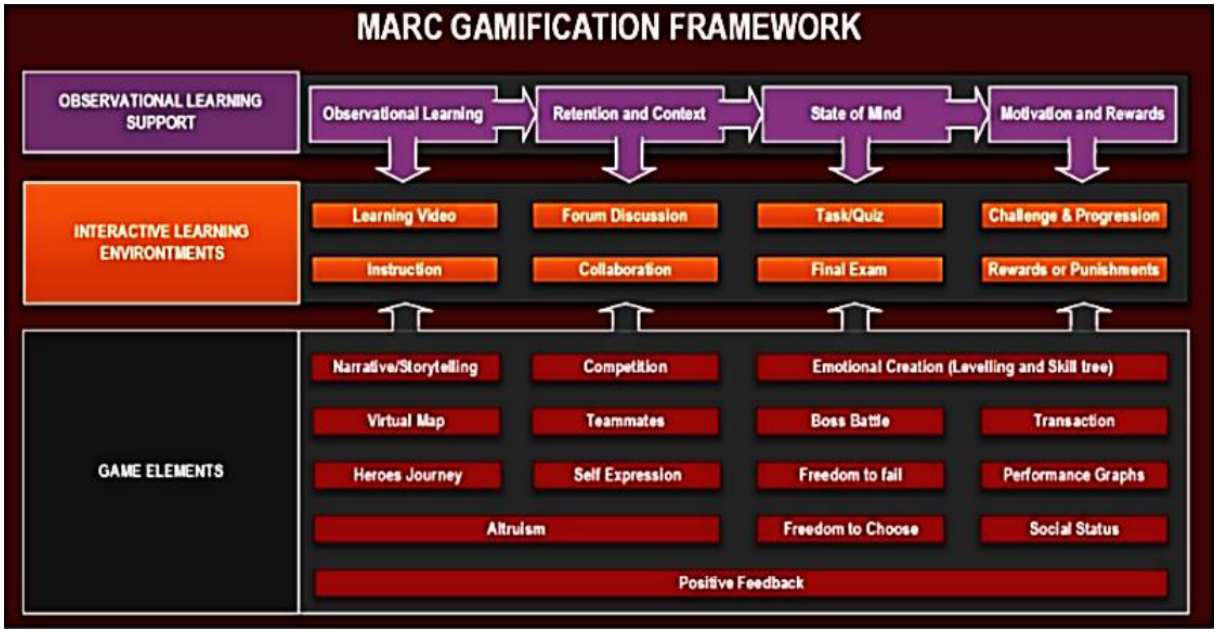


Figure 2.13: MARC gamification framework (Source: [31]).

The features of MARC are motivation: the framework addresses students' intrinsic and

extrinsic motivations for learning. It incorporates various game mechanics and elements to cater to different motivational needs. Action: MARC encourages active participation and interaction through tasks, challenges, and collaborative activities. It aims to move beyond passive learning towards a more engaged and action-oriented learning experience. Recognition: The framework acknowledges and rewards student achievements through points, badges, leaderboards, and other forms of recognition. This helps to motivate students and promote a sense of accomplishment. Community: MARC encourages creating a strong online community within the MOOC. It fosters collaboration, peer support, and student interaction, enhancing the learning experience. Components of MARC are:

1. **Motivational Triggers:** Curiosity: Using interesting content, storytelling, and problem-solving activities. Challenge: Offering tasks with appropriate difficulty levels and increasing complexity over time. Autonomy: Providing students with choices and control over their learning process. Relatedness: Fostering community and belonging through social interaction and collaboration. Progress: Providing clear feedback and tracking progress towards goals to motivate students.

2. **Action Mechanics:** Points: Awarding points for completing tasks, participating in discussions, and achieving goals. Badges: Offering badges for specific achievements or milestones to signify progress and recognition. Leaderboards: Ranking students based on points, badges, or other metrics to promote competition and healthy rivalry. Challenges: Setting time-bound challenges with specific goals to encourage active participation and problem-solving. Quests: Creating longer-term challenges with a narrative or storyline to motivate sustained engagement.

3. **Recognition Systems:** Public recognition through leaderboards and badges. Private feedback and personalized rewards. Opportunities for self-reflection and goal setting.

4. **Community Building:** Online forums and discussion boards for peer-to-peer interaction. Collaborative activities and team-based challenges. Social media integration to promote sharing and discussion.

The benefits of using MARC are increased student engagement and motivation. Improved learning outcomes and knowledge retention. Enhanced student satisfaction and course completion rates. Creation of a supportive and collaborative learning community.

2.2.2 Gamification Applications

Gamification in Education

The article presented presents a methodology to identify the students' gamification profiles for educational projects. The methodology consists of 6 activities shown in Figure 2.14 [26].

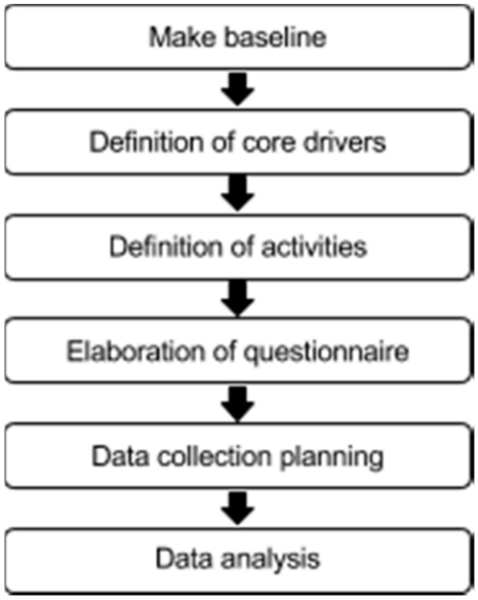


Figure 2.14: Block Diagram for the Methodology (Source: [26]).

The proposed methodology was applied, and the results show that the students have a similar impact from different core drives in the same class. Still, it varies with other students from other classes and ages, Figure 2.15 presents this information. An Octalysis was created Figure 2.16 reflecting the experiment's results applied to gather the target user gamification profile.

The article presented show a way to find the Octalysis core drivers most relevant to a specific target user. With this information, a gamification plan can be implemented with better validity. Not necessarily using all game techniques from all 8 core drives is suitable for gamification. When applying each game technique, you should focus on the core drives most relevant to the target user and less on the core drivers with less impact on the player.

Gamification in Health

Evaluation of the student core drives on e-learning during the covid-19 with octalysis gamification framework [32], aims to use gamification octalysis framework to analyze the

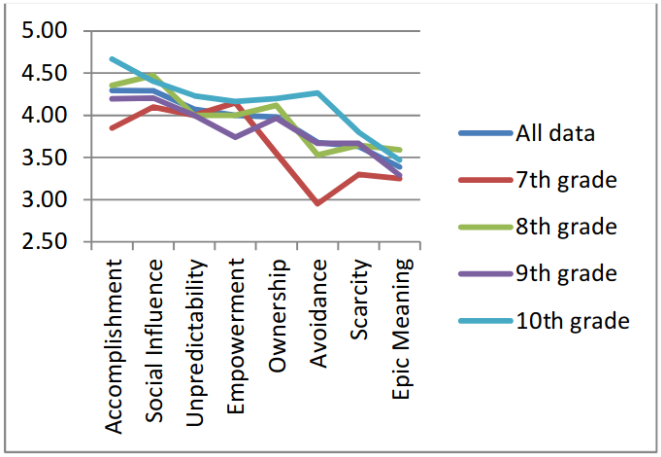


Figure 2.15: Core Drivers Averages per School Grades (Source: [26]).

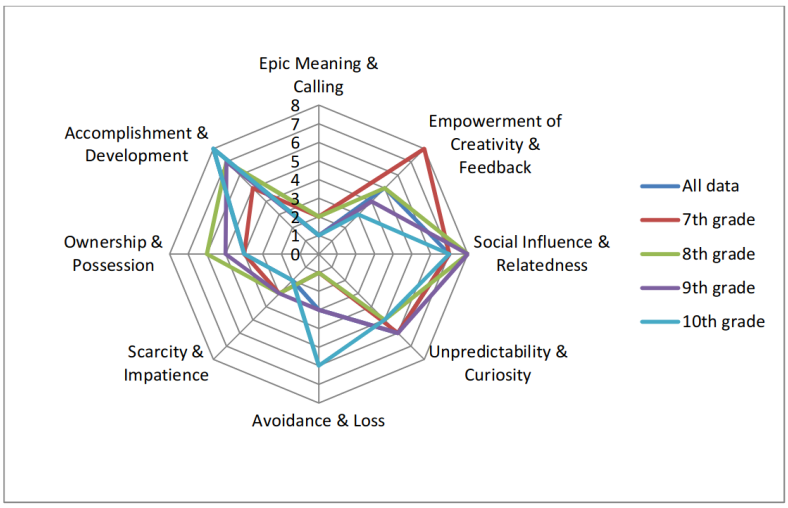


Figure 2.16: Octalysis Distribution for all the School Grades (Source: [26]).

extent of the role of gamification in the learning process and measure the amount of student motivation in online learning activities.

The paper is shown the application of the methodology in the context of e-learning pandemic students in Java Island, a questionnaire with 8 questions, one for each octalysis core drive, was applied, and after the results were collected, the Figure 2.17 was created to show the actual gamification profile. The gap to the author's desired results.

The application of the questionnaire, the study of each core drive, and the explanation of each core drive result and classification. It was unclear why the octalysis target is the maximum value, 10, for all 8 core drivers. All 8 core drivers explain why it is essential to improve the value, but there is no explanation on why it is a good target to have all 8 core drives in maximum values.

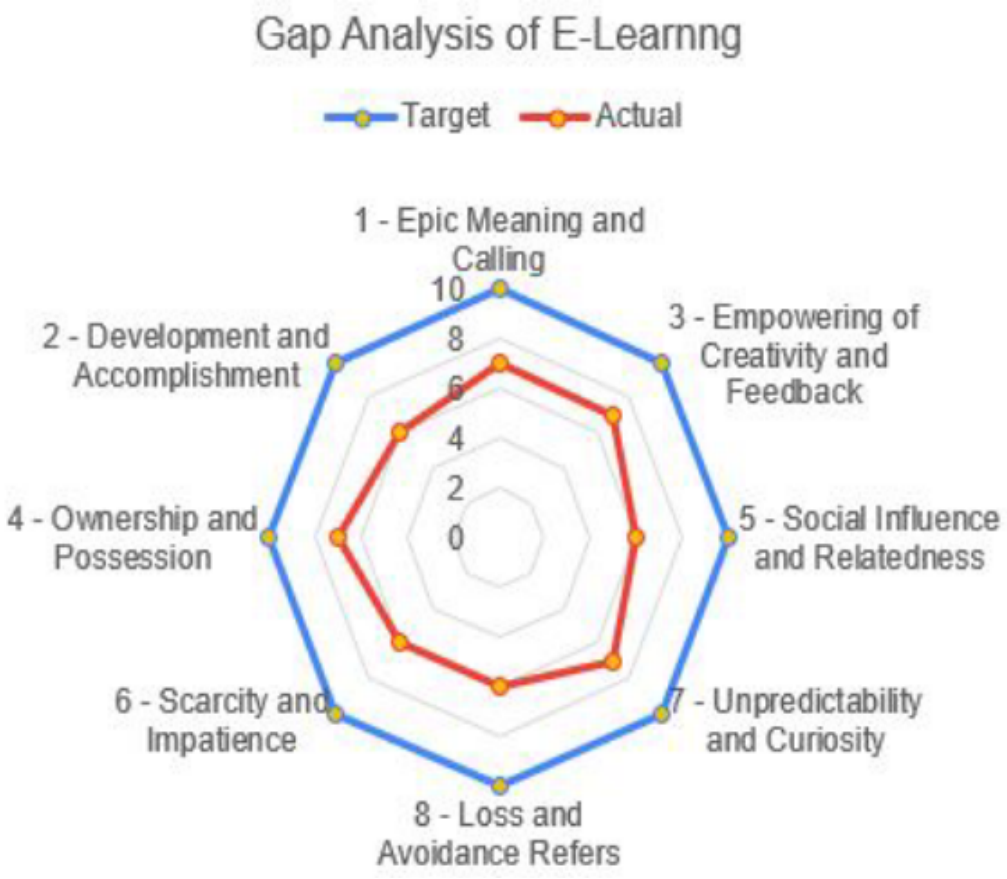


Figure 2.17: The Graph of Gap Analysis (Source: [32]).

Other gamification frameworks

There are others gamification frameworks available. Researchers have proposed various frameworks to guide the incorporation of game elements into non-game contexts. For instance, the Human-centered Design for Games (HCDG) framework[33] emphasizes understanding user needs and motivations throughout the design process. In contrast, the User-Experience-Focused Gamification Design (UXGD) framework [34] focuses on creating a positive user experience through gamification elements. Additionally, the Engagement Mechanics, Dynamics, and Aesthetics (MDA) framework [35] provides a more technical lens for analyzing game mechanics and their impact on player engagement. These are just a few examples, and selecting the most appropriate framework depends on the specific goals and target audience of the gamified experience.

2.3 GQM approach for validating software processes

The Goal-Question-Metric (GQM) approach, provides a structured framework for defining measurable objectives and evaluating software engineering processes. This literature review section will focus on the theoretical underpinnings of it and how it can be applied to validate software development methodologies[36].

This approach consists of three key components that work together to define and measure the success of a software engineering process:

1. **Goals**: These represent the high-level objectives for the process. Goals should be clear, concise, and aligned with the overall project objectives. Basili et al. recommend defining goals from multiple perspectives, such as the product, process, and resources.

2. **Questions**: These operationalize the goals by translating them into specific, measurable questions. Questions should be directly linked to the goals and provide a clear understanding of what aspects of the process need to be evaluated. According to Kitchenham et al., well-defined questions should be answerable through data collection and analysis.

3. **Metrics**: These are the quantitative measurements used to answer the defined questions. Metrics should be relevant, reliable, and sensitive to changes in the process. Following this approach principles and selecting metrics that can be objectively measured and provide a clear indication of achieving the desired goals.

The Goal-Question-Metric (GQM) approach offers a structured framework for researchers and practitioners in software engineering to effectively evaluate software development processes. This approach fosters targeted evaluations, objective assessments, and rigorous analysis.

This approach guides researchers to define clear and concise goals for the process evaluation. These goals encompass various perspectives, such as the product's quality, the efficiency of the development process, and the optimal use of resources. By focusing on specific goals, researchers can tailor their evaluation to assess the most impactful aspects of the process.

Objective assessments translate these goals into measurable questions. These questions directly address the goals and provide a clear understanding of what needs to be evaluated. For instance, a goal might be to "improve developer productivity". A corresponding question could be, "What is the average time developers take to complete a specific coding task using the new process?" By collecting data that answers these questions, researchers can objectively assess whether the process achieves the desired goals.

Finally, it emphasizes the selection of relevant, reliable, and sensitive metrics. Metrics are the quantitative measurements used to answer the defined questions. Reliable metrics ensure consistent results, while sensitive metrics readily reflect changes in the process. For example, a metric for the previous question might be "average task completion time in hours". By analyzing this metric, researchers can determine if the new process leads to a measurable reduction in development time.

While GQM offers significant benefits, it's important to acknowledge its limitations. Defining clear goals, questions, and metrics can be a complex task requiring careful planning and consideration. Additionally, even though GQM promotes objectivity, some aspects, like goal definition, may involve a degree of subjectivity from the researcher. Furthermore, GQM primarily focuses on quantitative data, potentially overlooking qualitative aspects like user experience or team dynamics.

In conclusion, the GQM approach provides a valuable framework for validating software engineering processes. By fostering targeted evaluations, objective assessments, and rigorous analysis, GQM allows researchers to objectively assess the effectiveness of a process in achieving its intended outcomes. Although limitations exist, the GQM approach remains a powerful tool for improving software development methodologies.

2.4 Literature review

This literature review examines the current landscape of research on chatbots and gamification. To gain a comprehensive understanding of this field, the review is divided into three sections. The first section focuses on the literature review application employed by various studies. The second section delves into the key findings of the reviewed papers. Finally, the literature review discussion will synthesize the insights gleaned from the reviewed research. This section will identify overarching themes, emerging trends, and potential gaps in the current knowledge base.

2.4.1 Literature review application

A comprehensive literature review was conducted to identify relevant research on gamifying chatbots and conversational agents. The details of this review can be found in Annex I. Table 2.1 summarizes the search results from three major academic databases: IEEE, SCOPUS, and Web of Science. The table details the search string used: "(CHATBOT OR CONVERSATIONAL AGENT) AND (GAMIFICATION)". It also shows the number of articles identified in each database, along with the number included and excluded for further analysis.

Below is described the selection criteria for the researched articles:

Base	Total	Removed	Selected and read
IEEE	6	0	6
SCOPUS	25	12	13
Web of Science	29	13	16
Total	60	25	35

Table 2.1: Chatbot and Gamification Intersection literature review results

1. **Gamification not presented**: This criterion indicates that the article should not be selected if it does not mention or discuss the gamification concept. In other words, the article should have some content related to gamification to be considered.

2. **Chatbot not presented**: This criterion specifies that the article should be excluded from the selection if it lacks any reference to chatbots. The article must feature some discussion or content related to chatbot technology.

3. **Gamification and Chatbot not presented**: This criterion highlights that an article should be disregarded if it lacks gamification and chatbot-related content. In other words, the article should contain at least one of these elements to be considered relevant for the study.

4. **Not an article**: This criterion filters out materials not qualifying as traditional articles, such as blog posts, forum discussions, or other non-academic sources.

5. **Not accessible**: This criterion is employed when the article is inaccessible or unavailable for review, potentially due to access permissions, publication status, or unavailability of the full text.

6. **Duplicated**: When multiple instances of the same article are encountered, this criterion excludes duplicate entries and streamlines the selection process, ensuring each unique source is considered only once.

The literature review references two tables to provide a comprehensive overview of the research literature. Table 1 (see Table I.1) offers a complete list of articles considered for this study, along with brief descriptions of their content. Table 2 (see Table I.2) focuses specifically on the articles that directly informed this research. This table indicates which articles were ultimately included and, for those excluded, provides a rationale for their omission.

2.5 Papers discussion

Below, the most relevant papers are described.

2.5.1 CiboPoliBot

According to Fadhil and Villafiorita, the first gamified chatbot is CiboPoliBot, a specialized chatbot that teaches children about a healthy lifestyle through an interactive social game environment.

It proposes an adaptive gamification approach to learning about a healthy diet and food waste management for kids between 8 and 14 years old. To accomplish this goal, a chatbot solution is chosen because of the ease of use and gamification applied to improve engagement from kids.

The authors present preliminary results about the gamification profile of the target users according to the Hexad Model. The result of a questionnaire shows that 97% of the kids use social networks like Telegram and WhatsApp, enabling the chatbot application through these message applications.

Figure 2.18 shows the project architecture. A gamification and conversation layer were created.

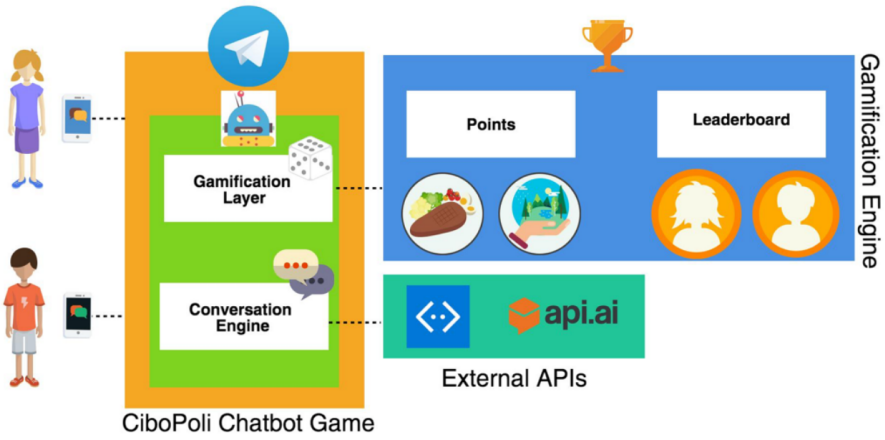


Figure 2.18: CiboPoli High-level Architecture (Source: [5]).

This paper is the first one that relates gamification and chatbots and uses a gamification Framework, too. But is an initial study not giving results, the paper includes steps in future work. The gamification framework's application shows this concept's more robust application.

2.5.2 Voice-based Apps

Voice-based Apps from Xu and Warschauer, this paper presents a framework to evaluate the educational design features of voice-based apps through a gamified.

The paper's research question is "What are the common and missing educational design features of literacy-focused voice-based apps on the market targeting young children aged 3 to 6 years?".

During the article, the authors also want to discover the common design features in learning chatbots for children. Parents and educators can use the proposed framework, and developers to choose applications or develop solutions. Another contribution is the analysis of applications in this context, providing a study on them.

According to the authors, frameworks exist for applications, computers, etc. However, they do not exist for voice applications. Showing a gap that can be explored in future studies. The proposed framework has 7 dimensions:

1. Learning Activity
2. Goal clarity
3. Interactivity
4. Gamification
5. Conversational prompts
6. Feedback
7. Scaffolding

After applying the framework with an analysis of 535 applications, it was possible to reach some conclusions, and it is important to be clear in the age range of potential users, leverage conversation technologies to promote interactive learning, encourage less restricted verbal expressions, provide feedback, provide menus to support conversation, introduce closure mechanisms.

During the study, 271 applications were found that provided a high level of interactivity; almost half implemented one or multiple gamification techniques.

The authors identified in this article that it is important to take advantage of the existing knowledge in the areas of education, human interaction, and computer development in constructing gamified conversational interfaces.

2.5.3 ScratchThAI

Katchapakirin and Anutariya created ScratchThAI with the goal of this article to overcome the lack of teachers developing a tutorial chatbot for ScratchThAI and use gamification to increase engagement. Scratch Is a colorful and interactive block-based programming language.

The Chatbot has 5 main iterations:

1. Designation custom dimensions
2. research practices
3. questionnaire resolution
4. Mission assessment
5. and report generator

Figure 2.19: An Architectural Design of ScratchThAI (Source: [39]).

The chatbot can also deliver missions according to the strengths and weaknesses of each student, in addition to offering opportunities for students to clarify and familiarize themselves with the concepts according to their pace. The conversational agent has 3 syntax processor components, a semantic processor and a dialog manager. Dialog.flow

was used as a technology for the project, but after collecting enough data, the following framework should be RASA, an OSS dialog engine. Figure 2.19 show the project architecture and layers.

2.5.4 Chatbot-Based assessment

Hungerbuehler et al. aim is to explore whether a text-based chatbot is a feasible approach to employee engagement and motivation to complete a workplace mental health questionnaire.

Applying a questionnaire made in a chatbot to mental health research appears to be highly engaging and effective in collecting anonymous mental health data. The number of responses to the chatbot can be compared to Face-to-Face interviews. In the authors' knowledge, no system is currently being used in work environments in mental health in Brazil.

The sample of an industrial plant in São Paulo, Brazil, totaling 120 employees who participated in the interview between October and November 2019. The technologies used to implement the chatbot were Ruby and JavaScript. They also added gamification features to increase engagement as they added levels of challenge points, progress feedback, and rewards.

Chatbot Viki applies the survey in conversation format and offers messages of encouragement. The entire survey takes about 15 minutes. Immediately after completing the questionnaire, participants receive personalized feedback and recommendations.

The following steps are a validation of the study to assess the effect of using chatbots for the application of psychometric questionnaires.

2.5.5 Charlie

Charlie: A chatbot to improve the elderly's quality of life and make them more active in fighting their loneliness. Charlie's chatbot can recall appointments and medications, remotely connect with doctors and family, and entertain elderly people. The authors comment on gamification strategies in the implementation of the solution but do not present what, leaving shallow the solution presented about gamification [41].

The research question investigates the communication characteristics of chatbots and interaction strategies that support their use and acceptance, especially for the elderly.

A personality was also implemented for the Chatbot Charlie, designed to promote easy connections with older users. The bot-coin concept has been created where Charlie monitors whether his suggestions have been followed and delivers coins to the user as a reward.

Figure 2.20: Screenshot of the mental health checkup (Source: [40]).

A structured interview was adopted based on the Unified Theory of Acceptance and Use of Technology *UTAUT) to assess whether the proposed solution had the desired impact.

The chatbot was efficient and helpful, and other users described the chatbot as imaginative and creative and reported that the chatbot had a warm and energetic heart. But some participants mentioned that the number of messages the chatbot sent caused a perception that the chatbot was compulsive.

Future work is to plan the gamification and implementation of the chatbot focused on the profile of users and to implement adaptations according to each user profile. Eventually, with the user's emotions, the behavior of the solution is different.

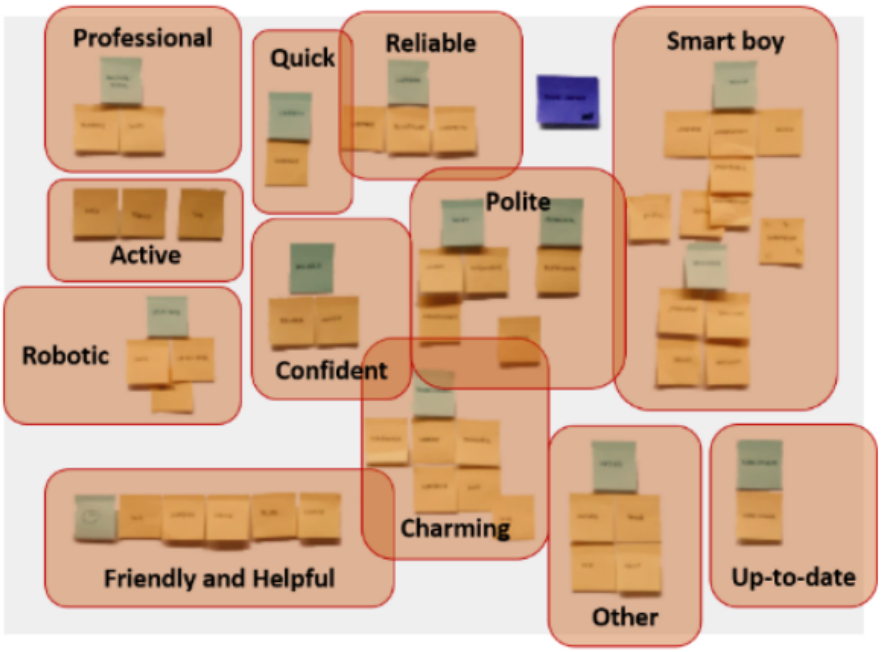


Figure 2.21: Affinity diagram to sort cues that shape Charlie's personality (Source: [41]).

2.5.6 Chatbot-Based learning media

Hidayatulloh et al. presents a review in the context of chatbots and gamification. Messengers and chatbots have become more popular in communicating with users as the mobile device industry have grown in recent decades (I haven't seen enough data to agree with this correlation). Chat systems have been extensively used in various legal, military, business, and educational areas. Additionally, there are more specialized chatbots with expertise in subjects that is best for comprehensive conversations. Chatbots have also improved student engagement and the learning process.

They also identified that studies on creating a general chatbot architecture for the overall learning system are still limited. The choice of architecture will depend on the type of system to be developed.

Chatbots with natural language processing and deep learning want a high computational process.

2.5.7 Intelligent recruitment with chatbot

This study show an analyzes how digital technologies can contribute to improving the stages of recruitment processes [43]. Recruitment chatbots can:

1. may contact potential candidates by sending messages;

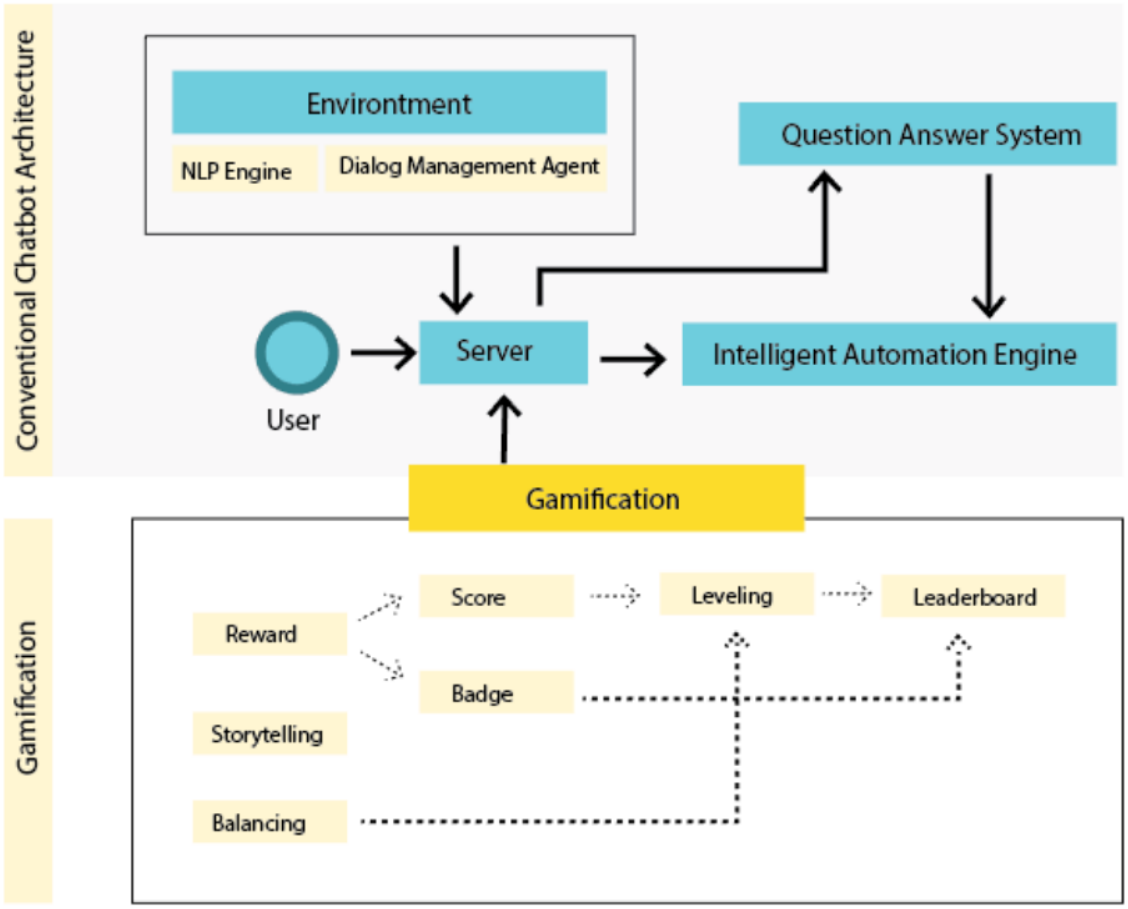


Figure 2.22: Chatbot with Gamification Architecture (Source: [42]).

2. talk to candidates and make an initial analysis using the simple criterion;

3. answer candidates' questions related to work and the company;

4. plan to prepare the interview;

5. candidates what are the next steps in the recruitment process.

Chatbots allow contact with 100% candidates and reduce the recruitment process cost. There are dozens of recruitments chatbots in the market which the ability to automate parts of the recruitment process by improving quality and reducing costs, this makes them very attractive to multinational groups or companies with large recruitment volumes.

The authors present Ari, a conversational robot dedicated to recruitment developed by a company called TextRecruit. The main purpose of this chatbot is to automate all the simple and repetitive tasks that the recruiter has to do that generate little value related to the time and energy dedicated to completing them.

2.5.8 Artificial intelligence in education

This paper Intends to address the use of artificial intelligence as a basis for modern online education, an interesting point of this study is in the context of the COVID-19 pandemic that was marked by lockdown from March 2020 [44].

The use of artificial intelligence, it is possible to make adaptive and personalized learning where the learning process should adapt to the student according to the amount of knowledge of each individual. Another important factor for the adaptation of learning is the pace of study of each individual.

The authors also cite good examples of applied gamification in the context of language learning, among them the final Duolingo. Smart Learning Analytics allows teachers to continuously improve educational content according to student progress and monitor student performance in this way teachers can adapt classes according to the analysis of the final data.

2.5.9 Prevention of obesity

This paper presents a discussion of the application of recent technologies to create a city platform [45].

The (STOP) project aims to establish the data and knowledge ecosystem as the basis for the portal to enable professional treatment and decision support in the feedback analysis of health information to optimize nutrition. The authors cite features such as Progress bar badges, rewards, and ranking as a form of gamification to increase engagement. Still, it is not in-depth which gamification techniques framework or how to use these gamification characteristics in the final solution. Dialog.Flow was chosen as a solution and mentioned Rasa as an option.

2.5.10 Chatbot for weight loss

Chew created a study on using artificial intelligence-based chatbots for weight loss. The focus is reviewing AI chatbot use cases for weight loss and identifying essential components to prolong user engagement.

The author concludes that AI chatbots should be designed to be human-like, personalized, contextualized, immersive, and enjoyable and to enhance user experience, engagement, behavior change, and weight loss. These require the integration of health metrics, personality and preferences, circumstantial behaviors, and emotional states to deliver personalized and practical recommendations for weight loss. Gamification is presented as another feature capable of improving user engagement.

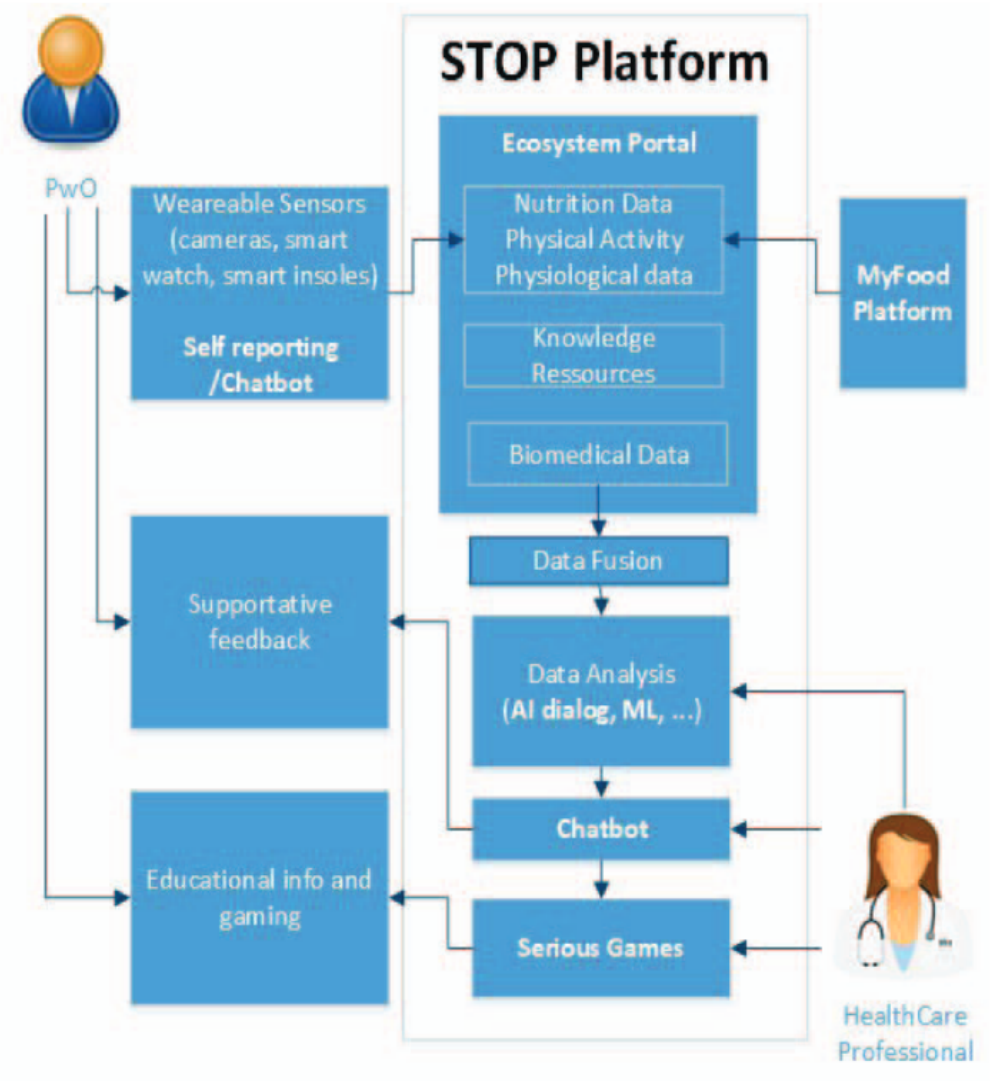


Figure 2.23: Overview of the STOP platform (Source: [45]).

The results suggest that although gamification may improve weight loss knowledge, user engagement, and intention toward health behavior change, it was insufficient to impact any current weight loss.

However, the type of gamification and the quality of gamification implementation in the articles studied are not presented, leaving doubt about the application of a structured and based gamification that could generate better results also in weight loss.

2.5.11 Playable Cities

Nijholt explores the potential of digital technology to make cities not just "smart" but also "playable." While cities are increasingly collecting data on various aspects of urban life,

the focus has primarily been on efficiency and optimization. This paper asks how this data and the underlying technology can be used to create more enjoyable living experiences for citizens.

The author proposes that sensors and actuators embedded within the urban environment can be leveraged to design playful applications. These applications could take the form of interactive digital installations or urban games, transforming mundane activities like commuting or housework into entertaining experiences.

Beyond just entertainment, these playful applications offer a valuable benefit for urban planners and designers. By observing how citizens interact with and respond to these playful elements, they can gain valuable insights into public behavior and preferences. This information can then be used to inform future design decisions and create even more engaging and user-friendly urban environments.

2.5.12 Xploro Digital

Bray et al. presents an evaluation study before and after aims to assess the acceptability of Xploro DTx and examine its impact on children and the procedural knowledge of their parents, procedural anxiety, and experiences reported when attending a hospital for a planned procedure.

A digital therapeutic platform was developed with children to provide health information through gamification, serious games, a chatbot, and an augmented reality avatar.

A total of 80 children and their parents participated in the study. The children were between 8 and 14 years and were attending the hospital for several procedures. The children in the intervention group reported a significantly lower level of anxiety before the procedures than in the standard group. The parents also reported lower levels of anxiety when using Xploro. The children told them they liked to use Xploro and that it was fun and easy.

2.5.13 Chatbot on financial

With the objective to provide financial assistance by supporting budget management tasks and then use that awareness of the users' goals and actions to provide analytical advice and suggestions to increase their financial literacy over time with a chatbot solution [49]. It uses gamification to improve engagement in keeping finances right, and the chatbot brings convenience for use:

Integrating these chatbots with social media applications like Facebook, Twitter, and Instagram increases their accessibility and convenience, which are desirable attributes for just-in-time education and nudging.

The users' feedback was positive to use the solution again, when asked if they would use such a chatbot, 75% said yes, 21% maybe, and 6% no. When asked if they think such a chatbot would improve their financial education and habits, 83% said yes, 13% maybe, and 4% no.

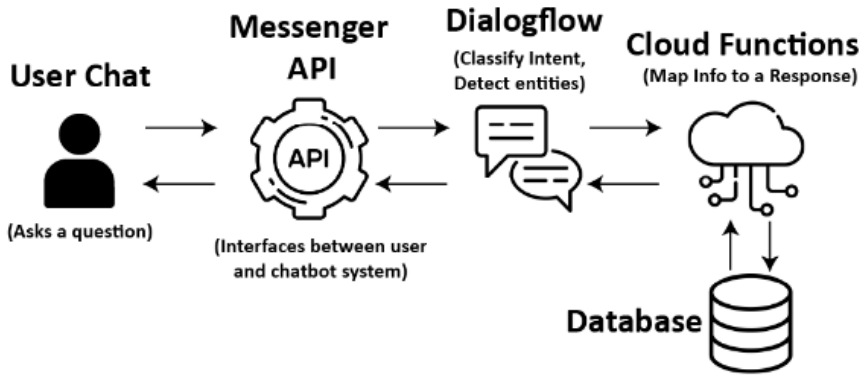


Figure 2.24: System Architecture (Source: [49]).

The article only presents the gamification concept but does not present how it was applied to the solution.

2.5.14 Vote Goat

Dalton et al. presents a movie recommendation chatbot called Vote Goat. The objective is to accelerate the research and Development of new conversational agents of recommendation for reusing components. Dialog.Flow is used as a solution development framework. The gamification aspect of using leaderboards should be extended to prevent abuse by limiting the number of ratings and detecting spam users—limitations with Dialog.Flow related to lack of support in using knowledge bases is also presented.

The authors cite gamification techniques as leaderboard ranking statistics but do not present how each of the gamification techniques was made.

2.5.15 Elena+

Ollier et al. presents a paper about a cell app that uses a Conversational Agent (CA) with various intervention components such as psychological education focused on topics: Information about COVID-19, mental health, sleep and diet, and nutrition. It can be downloaded as an IOS app and Android.

The Name Elena is named after an Italian nurse photographed exhausted in the treatment of COVID-19. About gamification, concepts by rewarding individuals with badges to

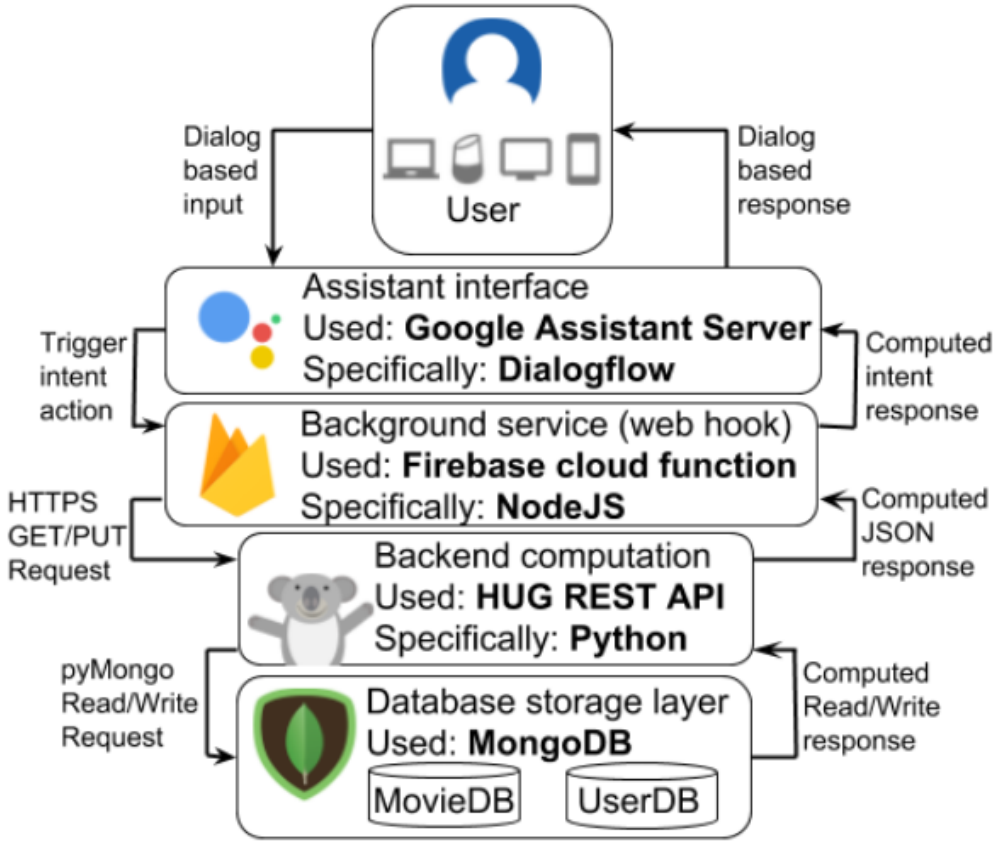


Figure 2.25: Architectural overview (Source: [50]).

symbolize progress through the App and awarding hearts for coaching session/assessment question completion, helping to evoke behavioral economic aspects related to avoiding losses and maximizing gains. The project also creates the background for future projects because it is Open Source.

It created social media accounts on Facebook, Twitter, Instagram, and LinkedIn, and a separate Source under Apache 2 license software platform. During the study's application, data on users' progress and evolution were collected and presented after each session.

The project focuses on monitoring and evolution of users in knowledge and COVID-19, and the change of indicators researchers found relevant. The application of gamification was simple, using specific techniques. The chatbot is the core product within the App and the way the solution was presented.

2.6 Literature review discussion

The literature review identified the potential of personalizing chatbot behavior based on user gamification profiles and adapting dialogue flow based on user interaction [41], [49].

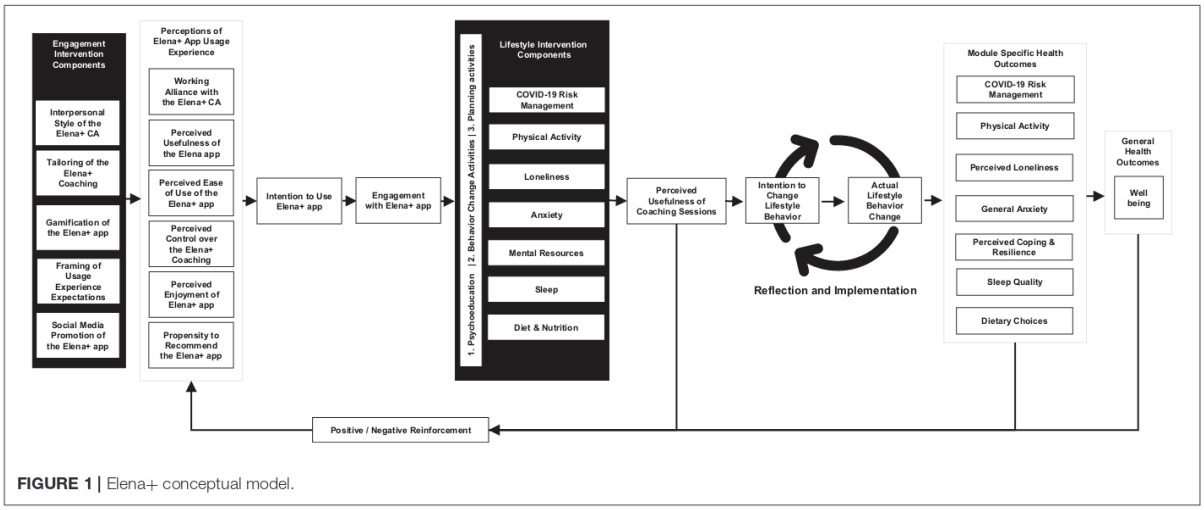


FIGURE 1 | Elena+ conceptual model.

Figure 2.26: Elena+ conceptual model (Source: [6]).

These findings suggest promising avenues for enhancing user engagement. However, a critical gap remains in the lack of consensus on best practices, activities, and desired outcomes for developing gamified chatbots.

This limited research on the intersection of chatbots and gamification highlights the need for a more systematic approach. Notably, existing studies primarily focus on the combined application of these concepts to offer quick solutions (chatbot) and increased engagement (gamification). However, none provide a structured process for implementing gamified chatbots with specific activities and targeted outcomes.

This project aims to bridge this gap by introducing the Gamified Chatbot Management Process (GCMP). This novel framework will outline a structured approach to developing gamified chatbots, fostering user engagement, and achieving desired results. Additionally, the project's open-source nature paves the way for future research and refinement within this domain.

Chapter 3

Methodology

The methodological procedure in this study delineates the approach undertaken to accomplish the specified objectives.

3.1 Gamified chatbot development process proposal

This section proposes a novel development process for creating gamified chatbots. The process is structured into three distinct phases. The first phase, Literature Review and Existing Solutions, lays the groundwork by examining relevant research and exploring existing gamified chatbots. This phase helps to identify effective gamification techniques and potential challenges that might arise during development.

Following this, phase two, Evaluation and Adaptation, focuses on tailoring the identified gamification frameworks to the specific context of the chatbot being developed. This phase involves evaluating the target audience, desired user experience, and the chatbot's functionalities to ensure the integration of gamification elements. Finally, phase three, Integration and Process Development, tackles the technical aspects of incorporating gamification into the chatbot. This phase involves the creation of a Gamified Chatbot Management Process (GCMP), which outlines the specific mechanics, dynamics, and aesthetics that will drive user engagement within the gamified chatbot.

3.1.1 Phase 1: Literature Review and existing solutions

The first phase of this project involves conducting a comprehensive literature review. This review will focus on identifying existing chatbot development processes, particularly those with a focus on user engagement. We will also explore current research on gamification techniques and their application within chatbot development.

The goal of this phase is to identify any existing "ready-made solutions" or established approaches that align with our desired outcome – creating a gamified chatbot development process. This includes exploring existing frameworks or methodologies used for chatbot development, with a focus on those that prioritize user engagement.

3.1.2 Phase 2: Evaluation and adaptation

Following the literature review, we will evaluate the identified solutions based on their effectiveness, adaptability, and alignment with our project goals. The aim is to determine if any existing process can be readily adopted or serve as a strong foundation for our proposed Gamified Chatbot Process.

If a suitable existing process is identified, we will explore the possibility of adapting it to incorporate gamification elements. This adaptation will involve integrating best practices and techniques gleaned from the literature review on gamification in chatbots.

3.1.3 Phase 3: Integration and process development

Should no suitable existing process be found. This phase will involve leveraging the knowledge gained from the literature review to define a structured approach for building gamified chatbots.

The process should outline a series of activities that guide developers through the process, ensuring the integration of gamification elements to enhance user engagement. This will likely involve defining user profiles, incorporating gamification mechanics, designing engaging interactions, and establishing clear goals and desired outcomes.

3.2 Experiment

This section details the experiment conducted to evaluate the effectiveness of the Gamified Chatbot Process with developers.

3.2.1 Experiment Design: Evaluating the process with developers

The experiment took place at the University of Brasília and involved a development team of students from the software engineering program. To equip the development team for success, it should be provided with the following initial materials:

1. **Process Manual**: This document served as a comprehensive guide, outlining the process framework in detail. It explained the various mechanics, dynamics, and aesthetics that could be incorporated into the gamified chatbot.

2. **Problem Statement & Objective**: A clear statement defining the specific problem the chatbot aimed to address and the desired outcome of building a gamified solution was provided.

3. **Gamification Plan**: A detailed plan outlining the specific gamification elements chosen for the chatbot, along with their intended impact on user engagement, was included.

The experiment uses the Goal-Question-Metric (GQM) approach to assess the effectiveness of the GCMP development process. This approach involves defining specific goals, then formulating questions to address those goals, and finally identifying measurable metrics to answer the questions. Periodic meetings were also scheduled with the development team to monitor progress, address any challenges, and provide ongoing support.

Using the GQM approach, the evaluation focused on observing whether the process facilitated the successful development of a functional gamified chatbot and whether the desired results outlined in the gamification plan were achieved. The evaluation metrics would be based on the specific questions formulated within the GQM framework.

This design allows for a structured assessment of the GCMP process. By observing the GQM results, we can gain valuable insights into the effectiveness of this approach for creating engaging and functional gamified chatbots.

3.2.2 Participant selection and onboarding

The process began with administering a questionnaire to participating developers. This questionnaire aimed to establish their baseline knowledge and experience, allowing for the creation of a well-matched development team.

Following the initial assessment, participants received a comprehensive onboarding package. This included reference documents, a project kickoff meeting, and access to the process manual. The kickoff meeting and reference materials provided developers with the necessary information and context to begin building the chatbot in accordance with the process guidelines.

3.2.3 Project management and data collection

To guide and track progress throughout the project, a schedule should be provided to the participants. Weekly interviews were then conducted with each developer. These

interviews employ a structured questionnaire to gather relevant data on their progress and experiences using the process. Additionally, researchers documented their own observations and impressions throughout the development process.

3.2.4 Data Analysis and recommendations

Upon completion of the final chatbot launch, the researchers focused on analyzing the collected data. This analysis aimed to assess the effectiveness and impact of the process on the development process, including its role in facilitating the creation of the gamified chatbot. Additionally, the analysis aimed to identify valuable lessons learned and potential recommendations for improving the GCMP framework for future projects.

This revised section provides a clearer structure and highlights the key aspects of the experiment design, data collection approach, and the dual role played by the researchers.

3.2.5 GQM application

The Goal Question Metric (GQM) approach [36] is a systematic method to ensure that measurements are goal-oriented, providing tangible value to the software engineering process. This approach operates on three levels, each contributing to formulating and realizing clear measurement objectives.

- **Goal Level**: At this primary level, the main objectives of the measurement are defined. These goals are often broad and relate to the more significant outcomes desired from the software process or product. For instance, an objective might be to "Improve the efficiency of the gamified chatbot development process".
- **Question Level**: Derived from the defined goals, specific questions are formulated to evaluate the extent to which the goals are being achieved. Using our previous goal as an example, a corresponding question might be, "How has the introduction of gamification impacted the development time of chatbots?"
- **Metric Level**: For each question posed at the operational level, concrete metrics are defined to provide quantitative answers. Concerning the abovementioned question, a potential metric could be "Average development time (in hours) for chatbots before and after gamification implementation".

By diligently applying the GQM approach, we ensure that our metrics are not just numbers but valuable insights driving purposeful improvements. Throughout our research, we have sought to align our measurement activities with this structured methodology, ensuring that our findings are relevant and actionable.

In the context of this study, GQM is applied to define measurement objectives related to the use of the process in the development of gamified chatbots.

3.3 Data collection and analysis

This section describes the approach used to analyze the data collected during the experiment and evaluate the effectiveness of the gamified chatbot process.

The analysis will leverage the pre-defined Goal-Question-Metric (GQM) plan established for the project. This GQM plan outlines specific metrics that will be used to answer research questions directly tied to the project's objectives. By analyzing these metrics, we can assess the impact of the process on various aspects of the development.

In addition to the quantitative data collected through the GQM plan, the analysis will also consider the qualitative data captured throughout the experiment. This includes researcher notes, observations documented during interviews, and any feedback provided by developers. Analyzing this qualitative data allows for the generation of hypotheses for further improvements and potential future evolutions of the process.

By combining the insights gleaned from both quantitative and qualitative data, this analysis will provide a evaluation of the process effectiveness. This will allow for the identification of strengths and weaknesses within the framework, ultimately informing future refinements and iterations of the gamified hatbot process.

Chapter 4

The Gamified Chatbot Management Process (GCMP)

After applying the literature review Annex I, one article stood out, the CMP [4]. It presents a chatbot development process with well-organized activities and robust validation. After analysis, the CMP was chosen as the base process for the creation of the Gamified Chatbot Management Process (GCMP). This way, the process can start with well-defined and validated activities.

The CMP is an iterative incremental process with well-defined steps. I was drawn to the addressing of relevant chatbot development activities and their organization into stages and phases. Below is a description of the CMP, as illustrated in Figure 2.2. It covers three phases and six stages, each detailed in the following sections.

A **Modify Knowledge Base:** The knowledge base is the data used to train the chatbot model; it is the raw material used to create the chatbot behavior and knowledge. Adjustments to this data are needed to ensure that the chatbot will have the desired dialogue actions.

B **Model Training:** This phase is responsible for generating a conversational model that can be used for a chatbot to be able to have conversations. Normally, it is a shell command executed with desired parameters.

C **Model Testing:** After the model is generated, it is needed to ensure that the resultant model has the desired behavior, so testing it in a chat and applying model evaluation metrics should be done in this step of CMP.

D **Model Release:** The release of the chatbot model to the end user is the main objective of this phase. The end user or defined stakeholders can use the developed solution after release.

E **Read Conversations:** In chatbot solutions, the user can have an unexpected behavior other than planned. Reading and extracting metrics of chatbot conversations is an important step in gathering the information needed in future chatbot adjustments and CMP interactions.

F **Analyze Metrics:** The metrics collected in the previous activity, jointly with other gamification or system metrics, should be analyzed in this step, raising a hypothesis that aims to improve the chatbot behavior and the gamification application.

The Chatbot Management Process (CMP) is designed with a focus on Rasa framework chatbots. This leads to limitations when applied to other frameworks. Here's a breakdown of some key differences:

Knowledge Base Management (Step A): The CMP emphasizes modifying a knowledge base, a step irrelevant for frameworks where content construction and model training are integrated. In these frameworks, datasets are managed directly for training purposes.

Model Training (Step B): Decision tree chatbot frameworks typically don't have a separate training step for the Natural Language Generation (NLG) layer. They focus solely on training the NLU layer to identify the most suitable intent. This specific step of the CMP might not be applicable universally.

Model Testing (Step C): The "Model Testing" step in the CMP might not be directly relevant for decision tree chatbots because they lack a separate NLG model. The chatbot's behavior is entirely determined by the implemented decision tree. Renaming this step to "Chatbot Behavior Testing" would better reflect its purpose across different frameworks.

Model Release (Step D): The CMP simplifies model release for Rasa, often requiring a CI/CD pipeline for other frameworks. CI/CD (Continuous Integration/Continuous Delivery) automates the release process, which may be more complex in these cases.

Read Conversations (Step E): The "Read Conversations" step becomes impractical in large-scale chatbot implementations. Reading vast volumes of conversations is time-consuming and inefficient. Identifying relevant conversations for improvement becomes challenging. Users might end up reading many irrelevant conversations, hindering the analysis process.

4.1 GCMP evolution

Three iterations of the GCMP were developed throughout the experimental applications, each incorporating modifications to rectify identified deficiencies and enhance the GQM process implementation.

4.1.1 First version: GCMP as CMP extension

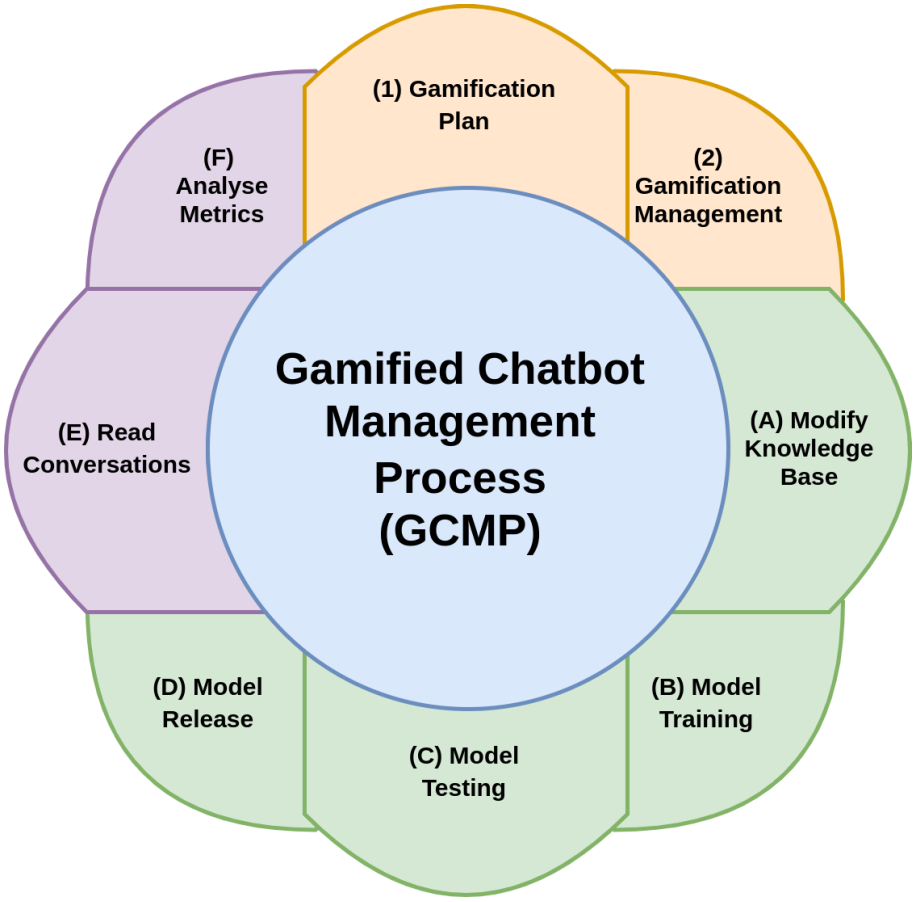


Figure 4.1: First GCMP version.

The initial version, as depicted in Figure 4.1, was characterized by a more constrained approach, emphasizing the sequential execution of each activity. Additionally, the steps were not well-defined, necessitating the manual's description improvements. The configuration of the entire project architecture, organization, and infrastructure posed a challenge during the initial phase, thereby impacting the execution of activities.

Due to the developers' lack of experience in the context, activities experienced delays in execution. Moreover, clusters of activities were identified, prompting the inclusion of phases encompassing the procedural steps.

In this first version of the GCMP, the expectation was that adding the steps (1) Gamification Plan and (2) Gamification Management, along with adjustments to the other CMP activities, would be sufficient for the implementation of a gamified chatbot. However, during the experiment, it was noticed that steps (A) Modify Knowledge Base, (B) Model Training, and (C) Model Testing were being executed more times per version than the other activities. Since it is an iterative incremental process, the expectation is

that the activities will be executed in a sequence and going through all of them, but the development activities were executed much more than the others.

Another point identified was the difficulty of the team in managing the requirements and implementing the gamified chatbot in the initial phases of the project, also indicating improvements and better directions in the GCMP manual. In the 2 versions delivered in this first version of the GCMP, the activities (E) Read Conversations and (F) Analyze Metrics were not executed at all, which indicated the need for improvements in these activities, along with indications of having tools and ready-made configurations for analyzing user conversation content available before the start of the project.

Finally, this first version of the GCMP did not prove to be positive for the construction of gamified chatbots, even using the CMP as a basis, the proposed adjustments were not enough to indicate a good application of the process. To meet the needs raised, a second version was created.

It is important to note that during the 2 releases of the first version of the GCMP, many tasks related to the initial configurations and project infrastructure configurations were necessary, inflating the weeks needed for releases.

4.1.2 Second version: GCMP improved

Due to the need to execute similar steps repeatedly identified in the first version of the GCMP, 4 phases were created and the 8 steps were divided into these phases, as in the CMP where 3 phases and 6 activities are suggested. With the application of the first version of the GCMP, more evidence was created of the need to divide the activities close to the development of gamified chatbots into phases. A change to highlight is the standardization of all steps with numbers, different from the first version, where the gamification activities were referenced by the letters A and B while the other activities by numbers.

The second version, illustrated in Figure 4.2, introduces a significant enhancement by incorporating phases encompassing distinct steps. Notably, the Gamification phase comprises steps 1, Gamification Plan and 2, Gamification Management. The Build phase encompasses steps 3 Modify Knowledge Base, 4 Model Training and 5 Model Testing. Following this, the Delivery phase introduces step 6 Model Release and finally, the Analysis phase integrates steps 7 Read Conversations and 8 Analyze Metrics.

Including phases enhances the clarity of executing steps within the same phase. Concurrently, the GCMP manual was updated to articulate this refinement and to provide improved descriptions for each step. However, challenges were observed during the analysis phase, where neither of the two analysis activities was executed.

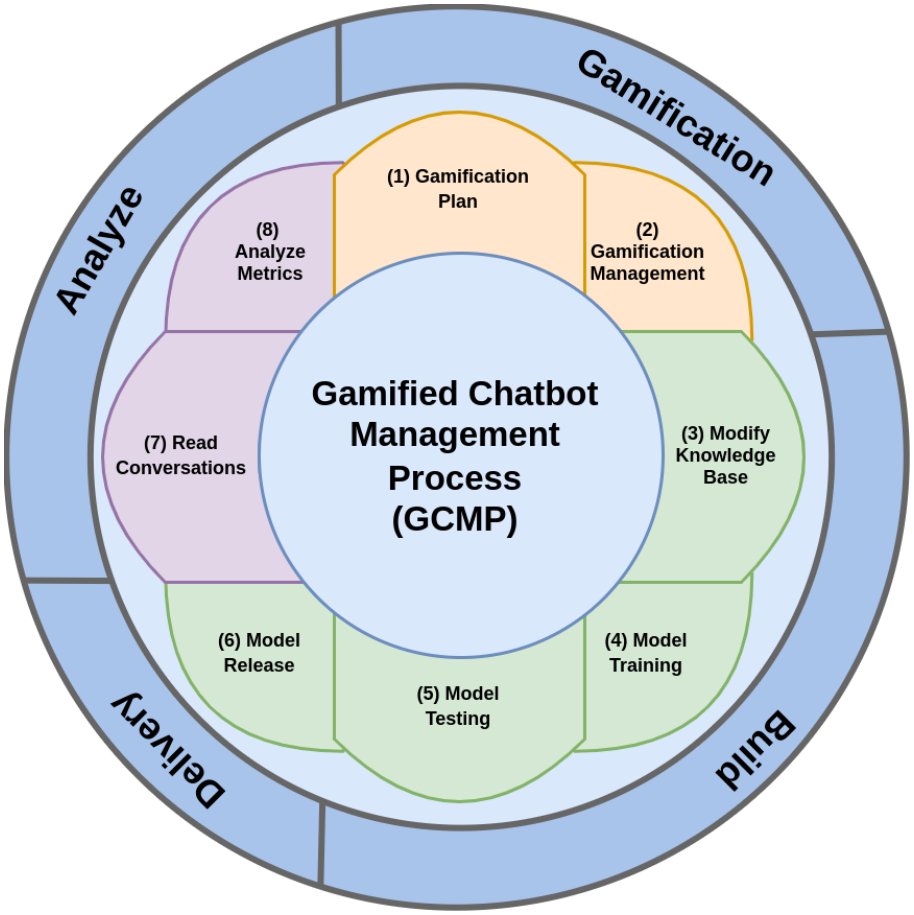


Figure 4.2: Second GCMP version.

During the four chatbot releases executed in this version, there was a noticeable improvement in the team's work rhythm, but this improvement was also due to the fact that the basic project implementations and initial configurations were already ready. The average number of weeks per release dropped from 6.5 weeks to 3.5 weeks, indicating an improvement in the time needed to implement gamified chatbot content and deliver it to production.

However, activities (7) Read Conversations and (8) Analyze Metrics have not yet been executed, partly because the number of real user conversations is close to 0 due to the project context, but also because there are gaps in the process in terms of not reinforcing the importance of these activities. With that in mind, a third version of the GCMP was created to address the new gaps identified.

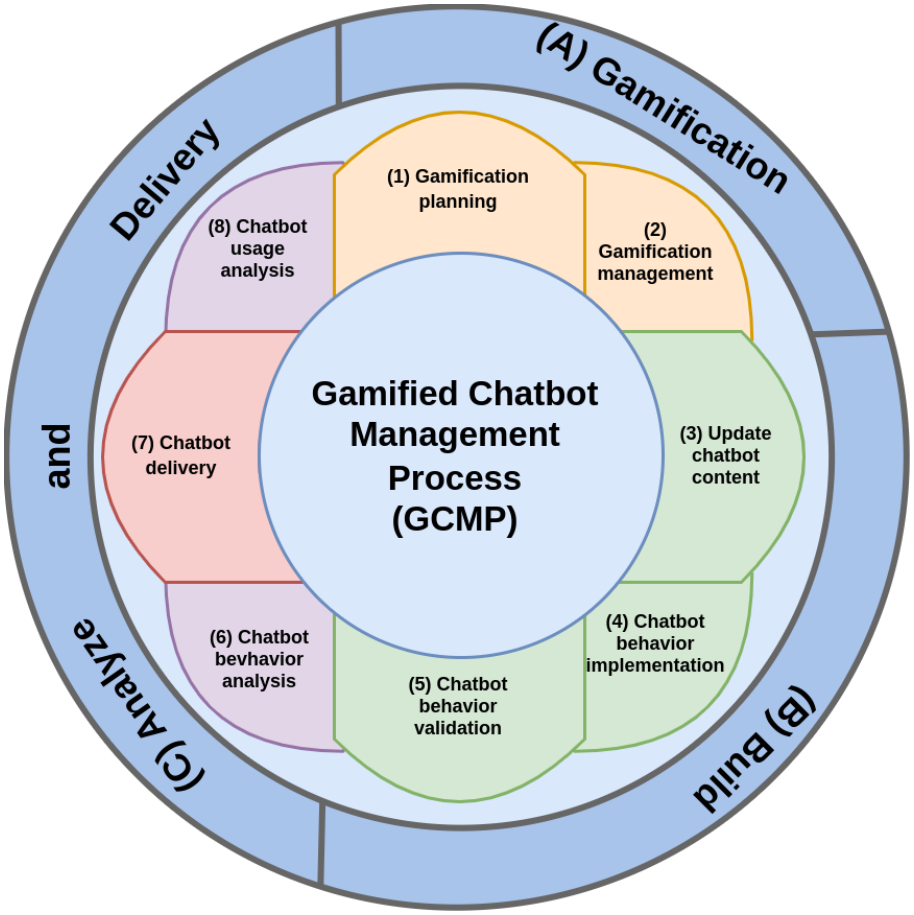


Figure 4.3: Third GCMP version.

4.1.3 Third version: GCMP generalized

The final version of the GCMP, as depicted in Figure 4.3, was developed. Commencing with Phase A Gamification it comprises steps 1 Gamification Planning and Gamification Management. In Phase B Build a comprehensive update was applied to all three steps within this phase, aimed at generalizing them to be less intricately linked to chatbot activities crafted explicitly with the Rasa framework. This adjustment was reflected in the GCMP manual, ensuring adaptability for experiments conducted with chatbots beyond the Rasa framework. Steps 3, 4, and 5 were accordingly updated to 3 Update Chatbot Content 4 Chatbot Behavior Implementation and 5 Chatbot Behavior Validation.

Within the novel Phase C Analysis and Delivery the sequence of steps underwent revision. The initial step is now 6 Chatbot Behavior Analysis. Upon a positive analysis, Step 7 Chatbot Delivery is initiated. Notably, this step was also modified to accommodate frameworks wherein the delivery activity does not solely involve deploying a model, allowing for alternative approaches to dialogue management. Finally, Step 8 Chatbot

Usage Analysis concludes the cycle, wherein the overall usage of the chatbot is analyzed, encompassing not only conversation reading but also recommending the utilization of conversational indicators, conversation reading, and an analysis of the gamification aspect of the solution.

4.2 GCMP

The Gamified Chatbot Management Process (GCMP) has a comprehensive manual available in Annex II. This manual provides detailed descriptions of each phase and step of the process, along with practical suggestions for each phase.

4.2.1 GCMP generated by applying the experiment

This section introduces the GCMP, tailored specifically for the conducted experiment. While the core principles remain consistent, the activities, files, and guidelines have been thoughtfully adapted to align with the experiment's scope as defined by the GCMP's provisions. Additionally, each activity description incorporates insights and experiences gained from applying the process during the experiment (detailed in the following chapter).

A comprehensive GCMP manual, outlining the refined steps, activities, and guidelines, is provided in the Annex II for future reference.

The ensuing premises, identified within the scope of the experiment, bear direct relevance to the adaptation of the Gamified Chatbot Management Process (GCMP), as expounded upon in the subsequent section.

1. **P01 - Octalysis**: Octalysis is the chosen gamification framework;
2. **P02 - Rasa**: Rasa is the chosen chatbot framework;
3. **P03 - Team**: The chatbot scope should fit the time, the team experience and external factors impacting the final solution.

The Gamified Chatbot Management Process (GCMP) extends the Chatbot Management Process (CMP) by incorporating gamification steps, with each step aligned to fulfill its objectives throughout the iterative process. The ensuing description delineates the requisite actions and anticipated outcomes within the eight steps. Rooted in the CMP process, the GCMP comprises three phases and eight steps, as illustrated in Figure 4.4, providing a visual representation of the organizational structure of the GCMP.

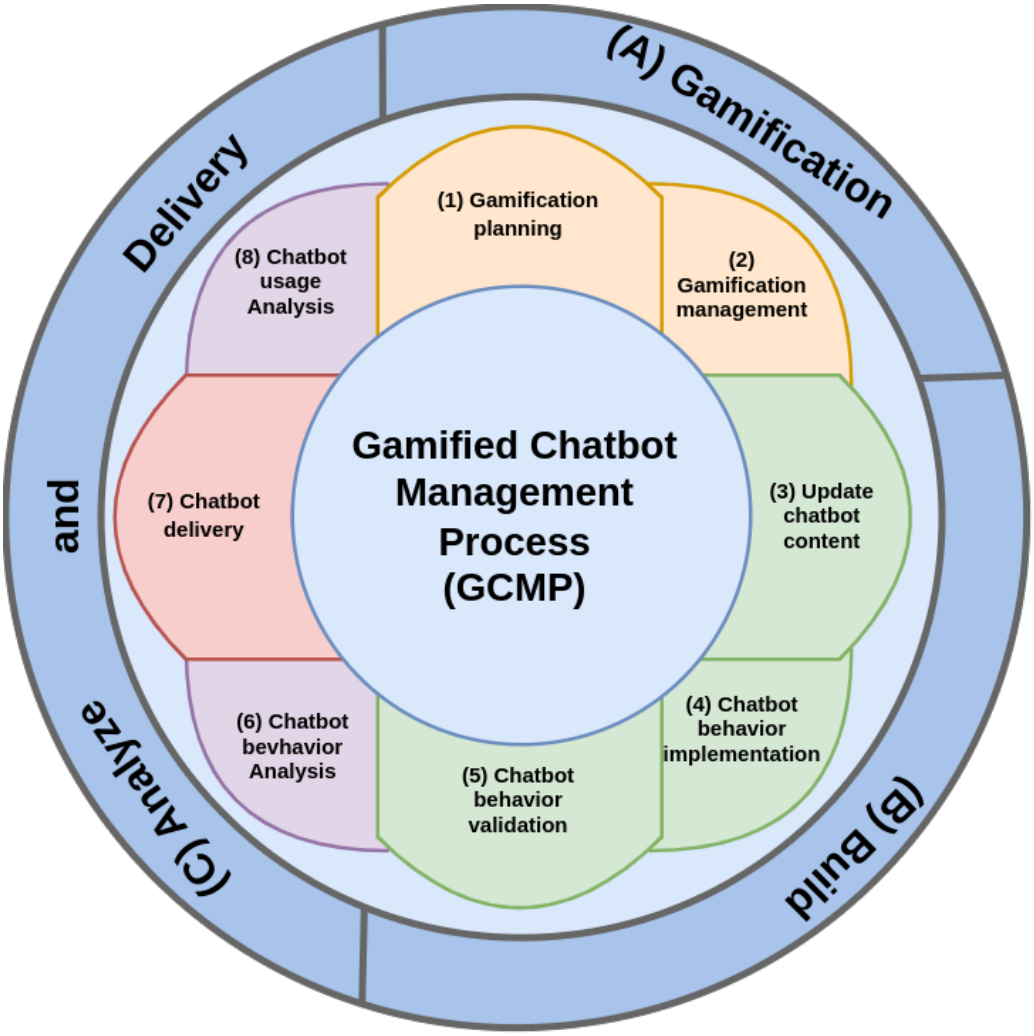


Figure 4.4: Gamified Chatbot Management Process (GCMP).

4.2.2 (A) Gamification

The gamification phase aims to address the development and evolution of the application of gamification in the chatbot, ensuring that gamification techniques are implemented in a manner that aligns with the intended format and purpose.

(1) Gamification planning

The Gamification planning step is applied in each GCMP iteration. Still, the first iteration will bring new information about the context, which should be updated in future iterations as needed. It is important to manage the following information:

1. **Business Metrics which lead to Game Objective**: many chatbots have a FAQ behavior. Stakeholders may want the user to find the solution and avoid the

actuation of a human, so increasing chatbot retention could be a game objective.

2. Users, which lead to Players: Defining the target users and raising the user gamification profile is significant to ensure that the designed gamification will impact each user.

3. **Desired Actions, which lead to win-states**: to define a win-state, the desired actions that the player should achieve are needed. All chatbot messages are essential to guide the user to the desired point.

4. **Feedback Mechanics, which lead to triggers**: feedback mechanics must be designed to tell the user that their conversation is meaningful. All feedback mechanics should become triggers to promote the desired actions.

5. **Incentives, which lead to Rewards**: Finding what the chatbot can give users when they commit the desired actions and arrive at the Win-state is the objective of this item.

Each of the gamification techniques that significantly influence the identified core drives must be meticulously cataloged, with detailed plans for their integration into the chatbot solution. These plans should undergo continual review and adaptation across subsequent iterations, informed by the gamification data and insights gleaned during solution development.

Throughout the experimental phase, this process was iterated ten times, with the primary tool being the gamification plan. Within this plan, tasks included managing gamification elements, refining techniques, and strategizing the implementation of each technique within the chatbot framework.

With every iteration of the GCMP, careful evaluation of gamification application was conducted, accompanied by necessary adjustments, as outlined in the experiment section. It is crucial to emphasize that the strategic planning of gamification technique implementation within the chatbot interface constitutes a blend of gamification and chatbot activities. This synthesis is essential, particularly regarding the integration of gamification into dialogue behavior within the context of gamified chatbots.

(2) Gamification management

With the initial or rolling gamification plan as income, managing the user's gamification profile to ensure the gamification will have the desired result. Updates on it should impact the usage and add or remove game techniques.

The management of all four game phases is important here:

1. **Discovery**: the first interactions that your chatbot sends to the user presenting itself and trying to summarize what it can do.

2. **Onboarding**: Presenting the tools and rules of the conversation is essential in this phase. This is at the beginning of the chat, usually presenting a menu with known subjects of the chatbot and using creativity to give this information with buttons, carrousels, and other chat interactions that make the onboarding easier.

3. **Scaffolding**: this is the dialogue part where the user seeks the information or achieves the objective, usually going through dialogue fluxes and asking questions to the chatbot. The chatbot should give the answers in a relevant way to the user, with personal information if possible (sending the user's name or sharing with the user that this conversation is here to help him to achieve the final goal).

4. **Endgame**: when the user gets to the end of the conversation, this achievement should be evident and celebrated, ensuring the user receives positive feedback and giving the player the want to come back and start over more conversations with the chatbot.

All the phases, target player profile, and game techniques should be clear to drive the chatbot conversation behavior on the following GCMP activities.

The GCMP's gamification management activity proved crucial for ensuring successful implementation. This step involved adjustments to the initial planning phase whenever discrepancies arose between the intended gamification experience and the chatbot's actual behavior. In essence, the planned gamification techniques (hypotheses) needed to be refined to align with the chatbot's capabilities and achieve the desired user experience.

An initial confusion existed between implementing the four game phases and developing the chatbot's gamified behaviors. This was rectified by realizing that the entire chatbot flow needed to be implemented, allowing dynamic access to each game phase based on the user's profile and conversation history.

This activity highlights the interconnected nature of gamification management. It encompasses adjustments to both the gamification plan (layer) and the chatbot's gamified behaviors themselves. In some cases, it even involves developing new behaviors or removing those that aren't functioning as intended. This iterative process ultimately leads to a more cohesive and effective gamified chatbot experience.

4.2.3 (B) Build

The build phase primarily focuses on constructing and implementing content within the chatbot, whether it is gamified or not. Additionally, adjustments are made in line with

the hypotheses generated during the analysis activity, along with the inclusion of relevant content tailored to the chatbot's context.

(3) Update chatbot content

The chatbot's content update involves adding, removing, or editing content in the Rasa chatbot files, both in the natural language understanding (NLU) and natural language generation (NLG) components. It also requires updating custom actions used for integrating with other services and implementing dialogue logic-based behaviors.

The "Update Chatbot Content" activity within GCMP prioritizes the chatbot's implemented behaviors. However, since gamification planning and management occur beforehand, they significantly influence how these behaviors are realized in the chatbot. During the experiment, frequent adjustments were made to the Natural Language Understanding (NLU) and Natural Language Generation (NLG) layers. These adjustments ensured the chatbot's behaviors aligned with the planned gamification and ultimately achieved the desired in-chat experience.

The iterative nature of GCMP allows for course correction. If a behavior or gamification technique doesn't function as expected, the gamification plan can be re-evaluated. This might involve adjusting the plan or even removing the technique entirely if it proves infeasible.

The experiment successfully demonstrated the value of this activity. The observed evolution in the chatbot's conversational capabilities highlights the effectiveness of this activity in achieving the intended outcome.

(4) Chatbot behavior implementation

Following the content update in the files, it is necessary to implement these contents into the chatbot solution. In the case of Rasa chatbots, this is accomplished through training a dialogue model. After this training, a model file is generated and can be loaded and provided to users as the solution.

The Chatbot Behavior Implementation activity in GCMP acts as a vital complement to content updates. In the experiment with a gamified Rasa chatbot, this activity proved essential for two reasons. First, after updating content, a new content model needed to be trained to ensure the chatbot recognized and processed the revised text. Second, custom actions were necessary to manage the dialogue flow and achieve the gamified experiences planned earlier.

The importance of custom actions became clear during the experiment. For instance, one custom action implemented a dynamic opening message based on the user's profile, further integrating the planned game phases into the chatbot interaction. This exemplifies

how the specific implementation details of this activity can vary depending on the chosen chatbot and gamification frameworks.

In the Rasa framework context, this activity can be streamlined. Running a single "rasa train" command initiates the training process, generating logs that detail the model's learning progress. These logs were then analyzed to identify any issues requiring adjustments.

It's important to note that this activity can vary significantly depending on the chosen chatbot framework. Some frameworks offer built-in training abstractions, further simplifying the validation process.

(5) Chatbot behavior validation

After model training, validation is necessary to ensure that the new content, behaviors, and messages are correctly presented to the users. It is also essential to verify that the gamification techniques are behaving as expected to maintain a positive user experience.

The Chatbot Behavior Validation activity within GCMP serves a critical purpose: ensuring the implemented behaviors function as intended. This activity involves interacting with the chatbot to trigger the newly developed functionalities. Researchers then evaluate the chatbot's responses to confirm they align with expectations.

Within the experiment's context, utilizing the Rasa framework, validation was achieved through terminal interaction with the chatbot. Researchers assessed the displayed logs, verifying the accuracy of the chatbot's messages and classifications.

It's important to acknowledge that different chatbot frameworks might require adjustments to this activity. In most cases, validation will likely involve direct conversation with the chatbot, potentially within a development environment. Additionally, the specific methods for evaluating and checking behavior logs may vary depending on the chosen framework.

4.2.4 (C) Analyze and Delivery

This phase is responsible for delivering the gamified chatbot solution to the user through the delivery and analysis of the chatbot. This analysis includes examining the conversations to understand the chatbot's behavior and evaluating the chatbot's model metrics.

(6) Chatbot behavior analysis

Before proceeding with the delivery, verifying the correctness of the chatbot model metrics is essential. If any issues are identified, it is imperative to assess the severity of the problem and, if necessary, make the necessary corrections before delivery.

Within the Rasa framework, the Chatbot behavior analysis activity leverages built-in tools for efficient evaluation. The "rasa test" command generates a folder containing various metrics that assess the model's health and quality. These metrics provide valuable insights for researchers, allowing them to gauge the model's performance and identify potential issues.

Further analysis can be conducted using the "rasa-model-reports" project. This tool generates simplified reports that offer a clear picture of the chatbot's health. This facilitates informed decision-making: if problems arise, researchers can determine whether adjustments are necessary or if the issue can be addressed in future iterations.

As demonstrated in the experiment chapter, this activity is relevant in monitoring the chatbot's development. The evaluation process highlighted a continuous evolution in content and model health, emphasizing the importance of ongoing validation. Skipping this analysis can lead to a "snowball" effect, accumulating issues that become difficult to resolve later in the gamified chatbot project.

(7) Chatbot delivery

This activity involves delivering the chatbot solution to the users, encompassing both the deployment of the Rasa service and the Rasa SDK to ensure that all behaviors and custom actions are provided to the users.

The Chatbot Delivery activity within GCMP serves as the final step, where the developed chatbot is deployed for user interaction. In the experiment, this activity involved a virtual machine (VM). The VM's internal code and the chatbot service itself were updated using Docker containers.

This approach contrasts with some chatbot solutions like Dialogflow or IBM Watson, which typically abstract the delivery process for users. These frameworks often provide the entire chatbot service as a managed offering.

Regardless of the chosen framework, thorough validation remains crucial after deployment. Researchers should interact with the delivered chatbot and analyze logs to ensure everything functions as expected. This step verifies successful deployment and identifies any potential issues.

(8) Chatbot usage analysis

The analysis activity is focused on examining user behavior when using the chatbot. Therefore, it is necessary to read randomly selected conversations or seek specific behaviors, in addition to monitoring metrics, data, and charts such as evaluating user interactions with the chatbot, assessing specific user behaviors and patterns, and tracking relevant metrics, data, and graphs. The goal is to gain insights into how users interact

with the chatbot, identify any areas for improvement, and assess the solution's overall effectiveness.

1. Triggered intent ranking;
2. Triggered action ranking;
3. Activated dialog flows;
4. Triggered gamification techniques ranking;

4.2.5 GCMP discussion

The application of GCMP in building the Latte chatbot proved instrumental in refining the framework itself. The initial version lacked the depth and clarity needed to achieve the process' intended goal: the construction of effective gamified chatbots. Through periodic meetings, analysis of metrics, and the formulation of corrective hypotheses, the GCMP underwent a significant evolution.

Initially, the assumption was that simply adding gamification activities to an existing Chatbot Mechanics Process (CMP) would suffice. However, this proved to be an inadequate approach. A more holistic revision of the entire process, encompassing each stage and activity, became necessary. This involved not only integrating gamification more seamlessly but also enhancing the overall chatbot construction experience.

The final version of GCMP reflects this iterative development. Adjustments were made to generalize the process, allowing for its application with other chatbot development frameworks beyond Rasa. Additionally, gaps were identified in the original CMP activities, which were heavily focused on Rasa usage. While these adjustments hold promise for broader applicability, the current experiment's limited data cannot definitively confirm their efficacy across different frameworks.

This experience highlights the importance of an iterative approach in developing frameworks like GCMP. By actively testing, analyzing results, and adapting based on real-world application, we can create tools for building exceptional gamified chatbots. Further research with a larger data set is recommended to fully validate the GCMP's generalizability to various chatbot development frameworks.

Chapter 5

Experiment

This chapter details the experiment conducted about the Gamified Chatbot Management Process (GCMP). The chapter is structured into four sections. The first section, Experiment Context, outlines the setting and overall context of the experiment. This includes details about the specific problem the chatbot was designed to address and the desired user experience.

The first section, Experiment Files and Repositories, describes the materials provided to the development team. This section will specify the documentation, resources, and tools used throughout the development process.

The second section, the Development Process Assessment section delves into the evaluation methodology. We will outline the GQM approach used to assess the effectiveness of the GCMP framework, including the metrics employed and any specific challenges encountered during the development phase.

Finally, a conclusion section about the experiment is presented. By examining these sections, this chapter will provide a comprehensive understanding of the experiment design, its execution, and the methods used to evaluate the GCMP framework.

5.1 Experiment context

This experiment was conducted within the Master's program at the University of Brasília. A researcher and their supervisor designed the experiment, involving two student participants from the university's Software Engineering program.

The experiment focused on developing a gamified chatbot to assist researchers and students in their academic pursuits. The chatbot was designed for the Brazilian audience and built to be accessible on the Telegram messaging platform, with all messages in Portuguese.

The choice of an academic context for the chatbot was deliberate. Firstly, it provided a shared area of knowledge and interest for all project stakeholders. Secondly, the developed chatbot could serve as a valuable tool for researchers and students within the university.

Beyond the academic context, the experiment held significant relevance for the researcher involved. The knowledge gained through the exploration of gamification within chatbots could be directly applied to improve existing chatbots developed for the researcher's consultancy company.

5.1.1 Experiment team

The experiment consisted of using a team of 2 developers, both students of the Software Engineering course at the University of Brasilia (UnB). Initially, a form was sent to them so that they could answer and that their initial information could be measured. The questionnaire consists of demographic questions, questions related to knowledge in the development of solutions in general, gamification questions and knowledge of chatbot development. Both developers are between 21 and 30 years old, attending a higher education course, with 2 to 5 years of experience as a developer, have knowledge of the Python language, which is the most relevant for the experiment and have 2 to 5 years of experience of gamification, both stated that the knowledge in relation to chatbots is low and both also had knowledge of the Octalysis gamification framework. The divergence was identified related to gamification knowledge where one of the students indicated that the knowledge is High while another indicated that the knowledge was neutral.

At the end of the 10-month experiment, the same questionnaire was sent again, where one of the students indicated that his knowledge in relation to chatbot development went from Low to High, however his gamification knowledge remained neutral in the 2 responses. The other student indicated that his knowledge of chatbot development went from Low to Neutral, and his gamification knowledge remained high.

During the execution of the experiment, it was identified that the students divided the activities according to affinity and knowledge, which justifies each one having taken on activities more related to the parts they knew best, justifying why both evolved but one more focused on the chatbot and another on the gamification. It is important to note that it is not within the scope of the project for all team members to evolve in both areas of knowledge mutually.

5.2 Experiment files and repositories

The experiment was organized and managed through repositories on GitHub, with the aim of documenting and making each part of the experiment accessible. Figure 5.1 shows

the main repositories used in the project.

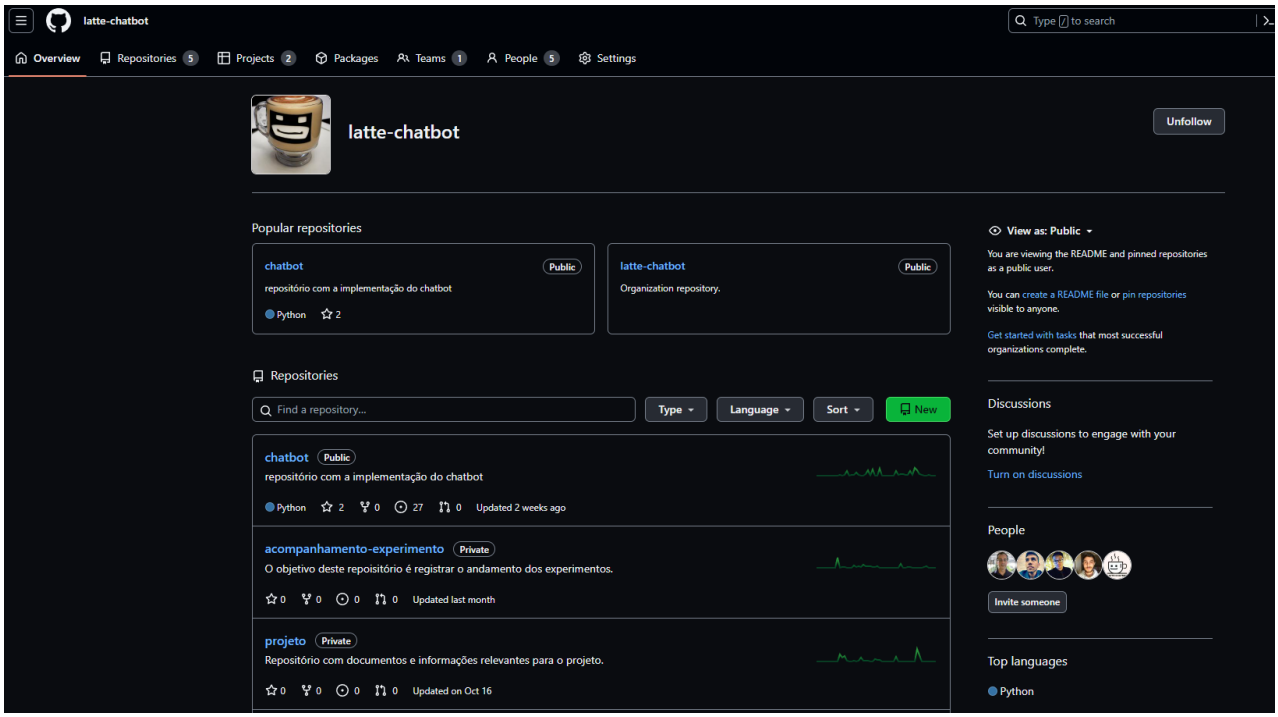


Figure 5.1: GitHub repositories.

1. **"chatbot":** This repository contains all the code and behaviors of the Latte chatbot. Pull requests are made here, the datasets used by the training model are maintained and updated, so that the evolutions are reviewed in code reviews and the history is maintained in Tags, so that it is possible to access previous versions of the solution;

2. **"acompanhamento-experimento":** This repository contains the documents related to the experiment and the research on the application of the GCMP. This repository is only accessible to the project researcher and the master's advisor, the students who are developing the Latte chatbot do not have access to this repository so as not to be contaminated with the annotations and documents with the history of the experiment;

3. **"projeto":** This repository maintains the entire project organization and scope management. Important documents such as the Gamification Plan and the GCMP manual are maintained and updated in this repository.

In the first week of the experiment, access to the chatbot and project repositories was provided, and a file called "statement" was also provided, where a brief description of the project objective is made. Additionally, the gamification planning document is

created with the initial project information and context and sent to the students. Finally, the chatbot repository was created but without any chatbot content, this repository is public and accessible for consultation at the following link: https://github.com/latte-chatbot/chatbot. Figure 5.2 shows the project repository.

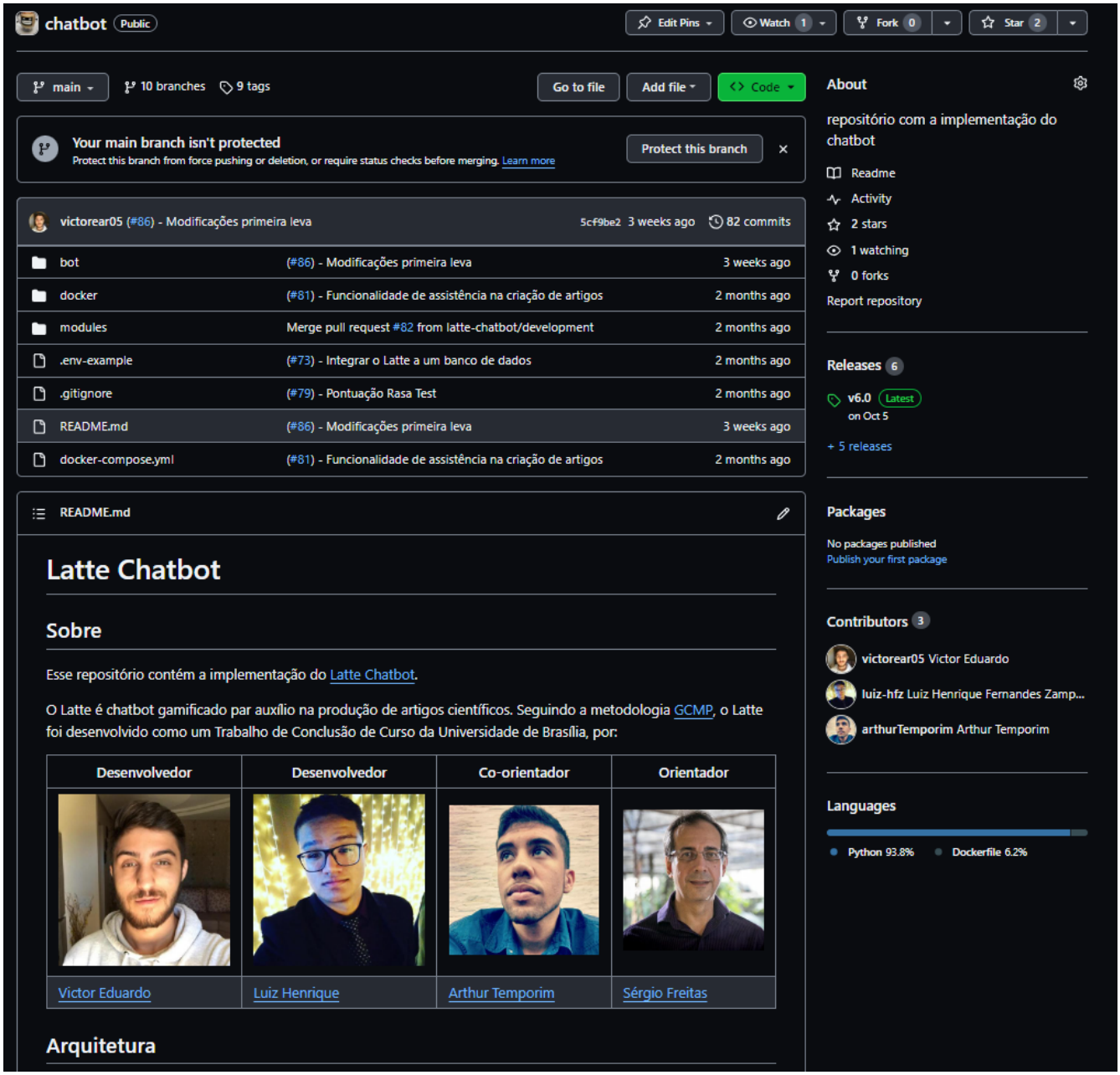


Figure 5.2: GitHub latte-chatbot repository.

The chatbot has been implemented and can be accessed by sending a message to the username @latte_chatbot on the Telegram messenger. An example of a conversation can be seen in Figure 5.3.

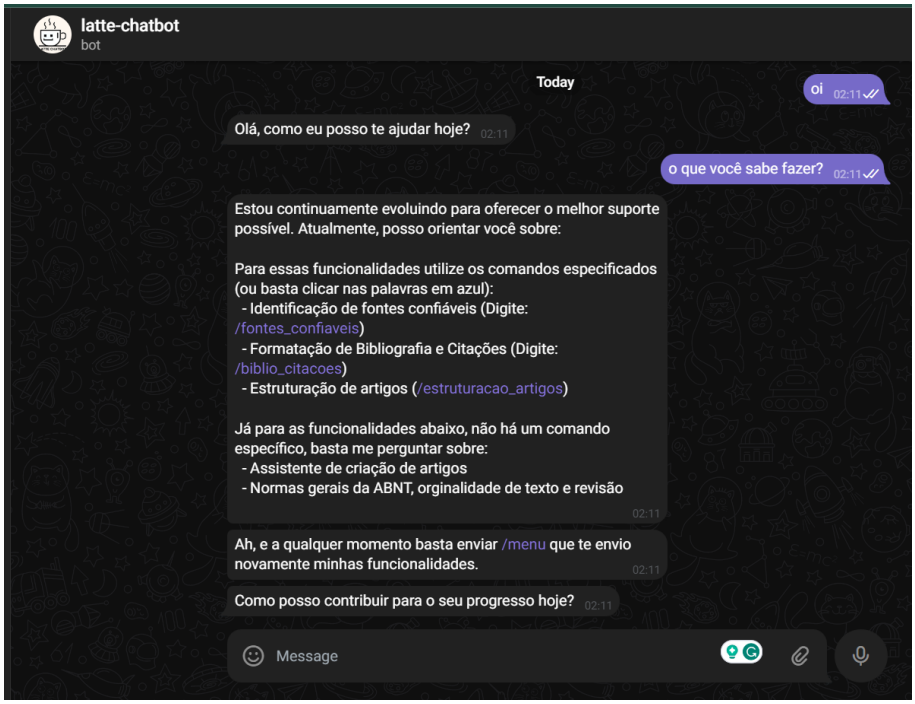


Figure 5.3: Latte chatbot conversation on Telegram.

The use of Octalysis [25] as a gamification framework and Rasa [51] as a chatbot development framework was defined. Octalysis, offered an approach to understanding the drivers behind human engagement. This framework consists of eight core drivers that express the player's motivations.

The Rasa chatbot framework was an open-source software development tool designed for building and implementing conversational AI agents. This framework offered a comprehensive set of tools and libraries to facilitate chatbot application creation, training, and operationalization. Rasa consists of two main components: Rasa NLU and Rasa Core, also known as Rasa Open Source [51].

Rasa NLU performed the critical function of natural language understanding, where it interpreted the meaning present in user inputs and extracted relevant intentions and entities from these text messages. Using machine learning algorithms, Rasa NLU-trained models can adequately categorize user intentions and extract pertinent information from the provided text.

On the other hand, Rasa Core has taken on the responsibility of orchestrating the flow of the conversation and making informed decisions about the appropriate responses based on user input and the current context of the dialogue. Using reinforcement learning, a machine learning paradigm, Rasa Core trained models with predictive capabilities, allowing the selection of the most appropriate action for the chatbot agent in any situation.

The experiment conducted in this study involved the participation of a development team made up of two members, both Software Engineering students with experience in gamification projects. The experiment aimed to implement a gamified chatbot based on the GCMP and the established gamification plan.

The experiment was conducted by sending three documents to the development team. The first document, the Statement outlined the objective to be achieved by implementing the gamified chatbot. This document provided guidelines on the desired outcomes.

The second document sent was the GCMP, a manual for the development process of the gamified chatbot. The development team used this document as their primary reference, following the steps and guidelines presented to ensure proper implementation aligned with best practices.

The third document, Gamification Plan contained initial project information and details about the adopted gamification strategy. This plan served as a basis for the development team to understand the project's overall vision and update it according to the specific context's needs.

To assist in project management and planning, a 32-week schedule was established, with expectations for the delivery of various activities throughout this period. This timetable provides a temporal perspective, allowing the development team to organize their tasks according to the established goals.

During the experiment, it was also identified that the developers had difficulty managing the project requirements, in order to identify the necessary activities, document them and monitor the progress of each implemented feature. During the experiment, in addition to the role of researcher, I also acted as Product Owner, helping to direct the project demands.

Using the GitHub platform, two boards were created to manage the project activities: one board focused on the development of the chatbot, including content, features, custom actions, and relevant documentation. The second board was created to manage the gamification of the project, mapping the gamification techniques and the progress of their implementation. Figure 5.4 shows the board. It has the following columns:

- Backlog: The complete list of gamification techniques and related activities are listed here.
- To-do: Activities that are prioritized for execution in the week are listed here.
- In Progress: Activities that are being executed during the week are listed here.
- Done: All completed cards are listed here.

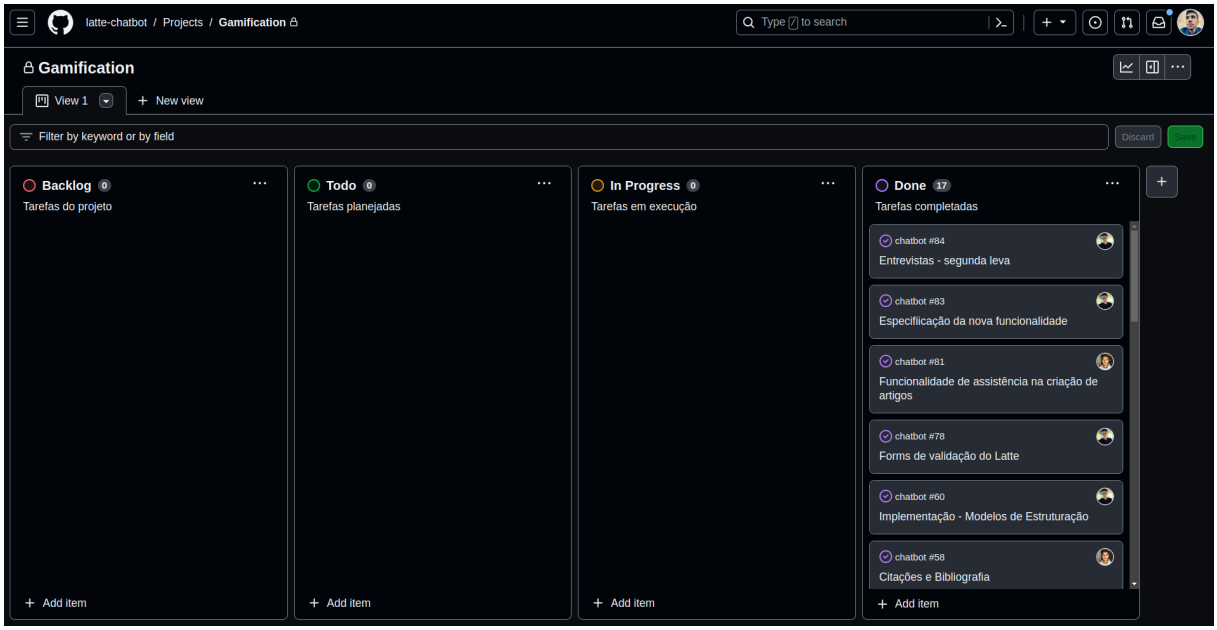


Figure 5.4: Gamification board on GitHub project.

The use of these boards facilitated the management and progress of the project, but the management of requirements is not within the scope of the GCMP.

Through this experiment, it was possible to evaluate the practical application of the GCMP, the gamification plan, and the development team's ability to follow the process and meet the project's expectations. The results obtained throughout the experiment provided valuable insights into the GCMP as a process for developing gamified chatbots and the feasibility and applicability of the established gamification plan in the project's specific context.

5.3 Development Process Assessment

This section delves into the evaluation methodology employed to assess the effectiveness of the GCMP framework. This evaluation utilizes the Goal Question Metric (GQM) approach, which provides a structured framework for analyzing software development processes. GQM focuses on defining four key aspects:

1. **Planning**: the planning stage involved defining the overall goal of the experiment: to assess the effectiveness of GCMP in facilitating the development of a functional and engaging gamified chatbot. This goal was further broken down into specific objectives, such as evaluating the clarity of the GCMP documentation, the ease of integrating gamification elements, and the overall time and resource efficiency of the process.

2. **Definition**: following the application, the GQM approach requires defining specific questions that need to be answered to achieve the overall goal. These questions should directly address the effectiveness of GCMP in aspects of the development process.

3. **Data Collection**: once the key questions are defined, the GQM approach focuses on identifying the data collection methods that will be used to answer them.

4. **Interpretation**: the final stage of the GQM approach involves interpreting the collected data to answer the defined questions. This stage requires analyzing the data, drawing conclusions, and ultimately evaluating the effectiveness of GCMP in facilitating the development process.

Each of these activities is described in subsequent sections.

5.3.1 Planning

The initial stage of the GQM application involved planning. Following the GQM framework, we defined three overarching goals for the experiment. These goals focused on evaluating the effectiveness of the GCMP framework in achieving specific objectives. For each objective, we formulated a series of targeted questions. These questions were designed to gather metrics that would allow us to assess the GQM goals.

The next step in the planning phase involved defining metrics for each of the formulated questions. A metrics table was meticulously constructed, outlining the following details for each metric:

- **Formal Description**: A clear and concise definition of the metric and what it measures.

- **Data Collector**: The person responsible for gathering the data associated with the metric.

- **Value Interval**: The range of possible values the metric could take (e.g., numerical values, yes/no responses).

- **Collection Interval**: The frequency at which the data would be collected (e.g., daily, after each development milestone).

- **Collection Method**: The specific method used to gather the data (e.g., surveys, process logs, user testing).

Finally, the planning stage also involved establishing a clear framework for data interpretation. This included defining how we would analyze the collected data to answer the formulated questions and ultimately assess whether the experiment achieved its three overarching goals.

The GQM approach is iterative by nature. We anticipated the need to potentially adjust the questions and metrics during the experiment based on the initial data gathered. This flexibility allows for refinement and ensures the GQM approach remains focused on effectively evaluating the GQM goals.

5.3.2 Definition

The first overarching goal shown in Figure 5.5 is focused on analyzing the gamified chatbot developed using the GCMP framework. This analysis aimed to assess the chatbot's capabilities from the researcher's perspective within the context of a research laboratory. Specifically, we were interested in understanding:

Chatbot Assertiveness

Question 1: How assertive is the chatbot in its communication? Metric 1 (M1): NLU score by each version of the chatbot. This metric measures the chatbot's ability to accurately understand user intent. A higher NLU score indicates better comprehension and potentially more assertive responses. Metric 2 (M2): Core score by each version of the chatbot. This metric (details about the specific core scoring system used would be helpful here) can provide insights into the overall quality and effectiveness of the chatbot's responses, potentially reflecting its level of assertiveness.

Content Evolution

Question 2: How did the chatbot's content evolve throughout the development process? Metric 3 (M3): Number of utterances by each version of the chatbot. This metric tracks the total number of unique phrases or sentences the chatbot can generate. Metric 4 (M4): Number of intents by each version of the chatbot. Intents represent user goals or requests the chatbot can understand. An increase in intents suggests a more comprehensive range of topics the chatbot can address. Metric 5 (M5): Utterance usages by each version of the chatbot. This metric tracks the frequency with which specific utterances are used by the chatbot. It can reveal patterns in the chatbot's communication style. Metric 6 (M6): Number of intent examples by each version of the chatbot. This metric indicates the variety of examples provided for each user intent, potentially influencing the chatbot's ability to understand and respond assertively. By analyzing these metrics in conjunction

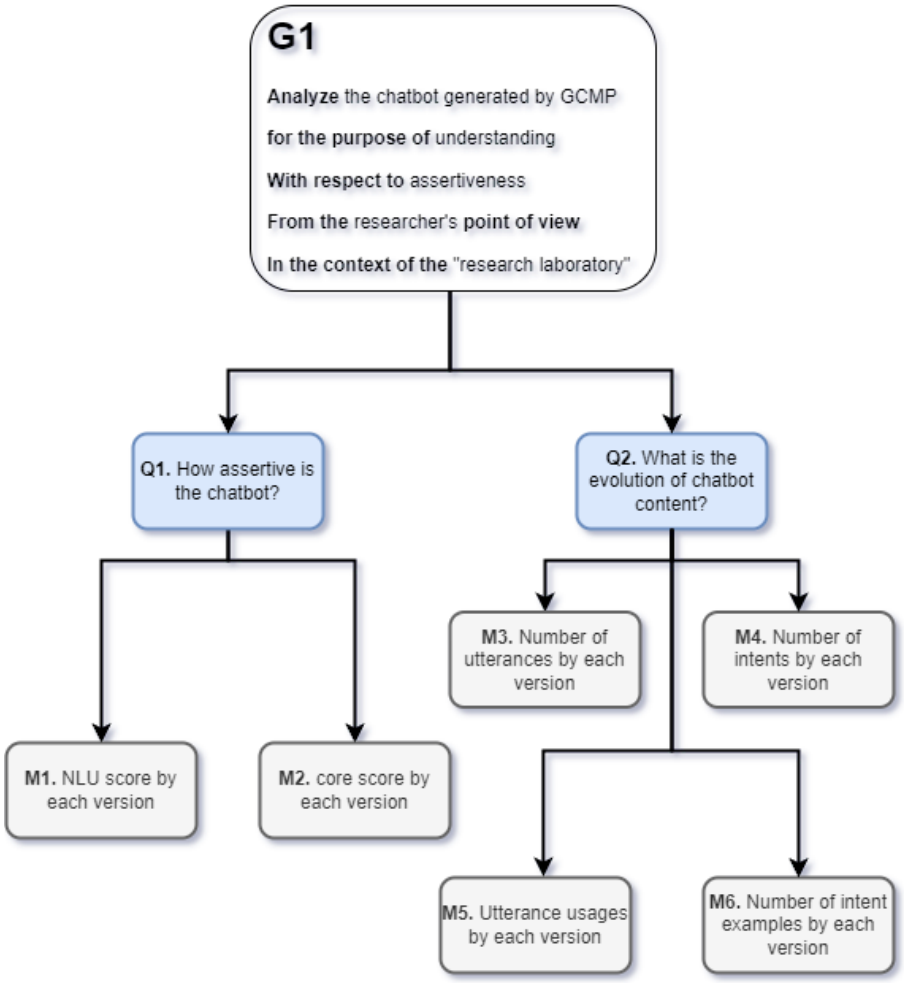


Figure 5.5: First goal of GQM.

with the research laboratory context, we aimed to gain insights into the effectiveness of GCMP in generating a chatbot with the desired level of assertiveness and the evolution of its content throughout the development process.

The second overarching goal shown in Figure 5.6 of the GQM evaluation shifted its focus towards understanding the effectiveness of gamification elements implemented within the chatbot. Here, we aimed to assess the chatbot from the researcher's perspective within the research laboratory context. Specifically, we were interested in evaluating:

User Engagement with Gamification

Question 3: What impact did gamification have on user engagement with the chatbot? Metric 7 (M7): User interviews. Conducting in-depth interviews with researchers who interacted with the chatbot allowed for gathering qualitative data on their experience with the gamification elements and their overall level of engagement. Metric 8 (M8): Number

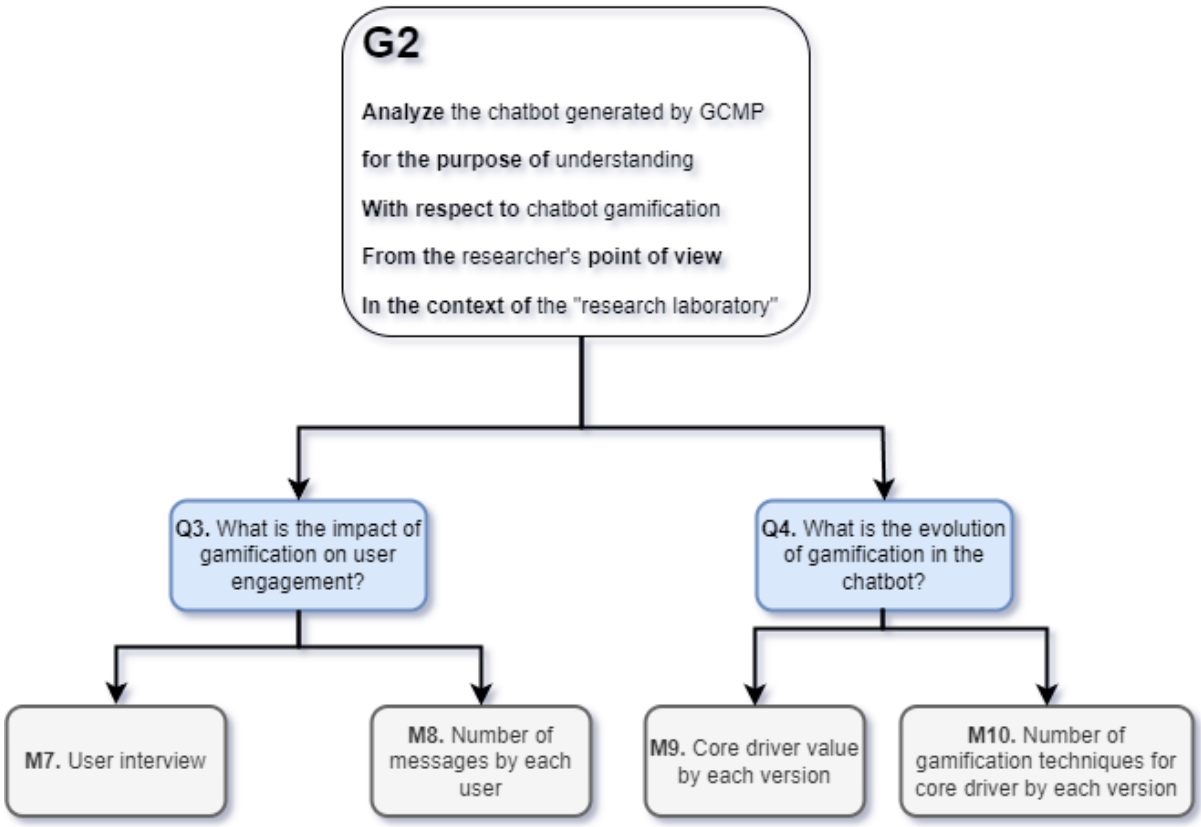


Figure 5.6: Second goal of GQM.

of messages by each user. This metric provides a quantitative measure of user interaction with the chatbot. An increase in messages could potentially indicate higher user engagement, though further analysis would be needed to distinguish between gamification-driven interaction and other factors.

Evolution of Gamification in the Chatbot

Question 4: How did the gamification elements evolve throughout the development process using GCMP? Metric 9 (M9): Core driver value by each version. It refers to a scoring system used within GCMP to assess the prominence of different core drivers within each version of the chatbot. Tracking these values can reveal how the focus on specific core drivers shifted during development. Metric 10 (M10): Number of gamification techniques for the core driver by each version. This metric tracks the quantity of specific gamification techniques employed to address each core driver within each version of the chatbot. An increase in techniques could indicate a more elaborate and potentially more engaging gamification experience.

By analyzing these metrics, we aimed to gain insights into how gamification elements, implemented through the GCMP framework, impacted user engagement and how the gamification approach itself evolved during the development process.

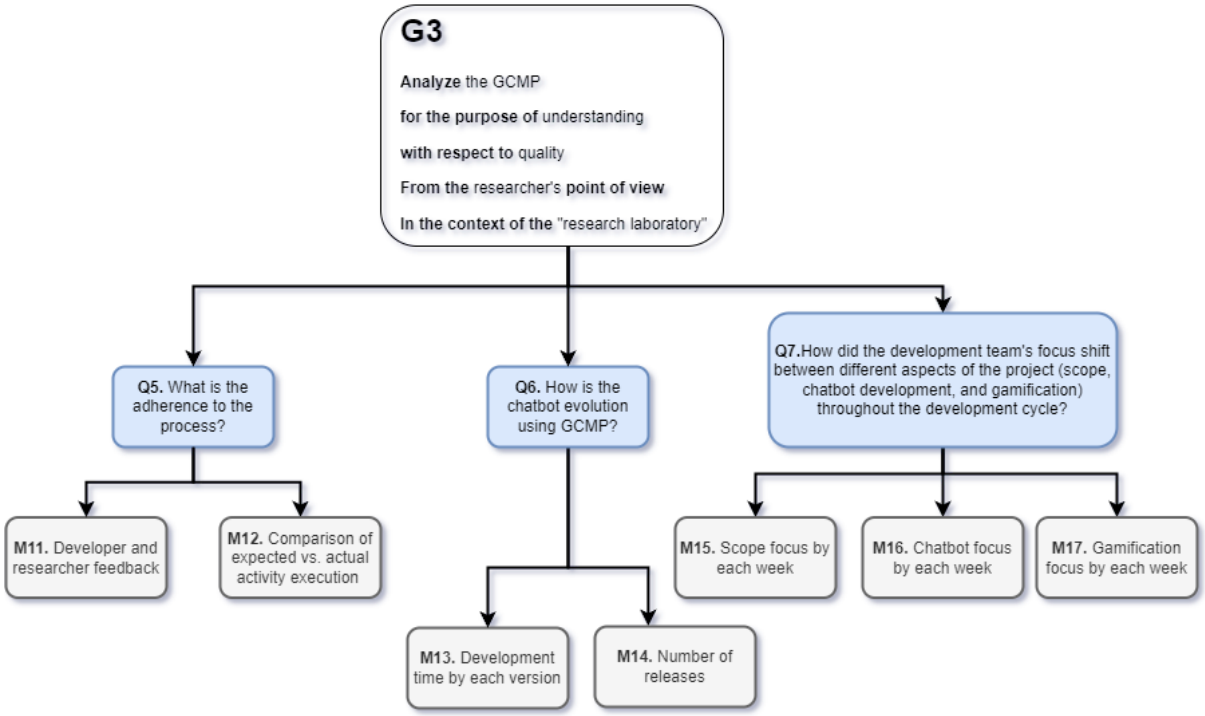


Figure 5.7: Third goal of GQM.

The third overarching goal shown in Figure 5.7 of the GQM evaluation focused on assessing the GCMP framework itself from the researcher's perspective within the context of the research laboratory. Here, we aimed to understand the quality and effectiveness of GCMP as a development process for gamified chatbots.

Adherence to the GCMP Process

Question 5: Did the development team effectively follow the GCMP process throughout chatbot creation?

- **Metric 11 (M11)**: Feedback from the developers and researcher. Gathering feedback through interviews or surveys with both the development team and the researcher allows for qualitative insights into their experience using GCMP. This feedback can reveal areas where the process was clear and helpful, as well as aspects that may have been confusing or difficult to implement.
- **Metric 12 (M12)**: Comparison of the execution of activities expected vs. performed. This metric involves a detailed comparison between the GCMP activities

outlined in the framework and the actual activities undertaken by the development team. Analyzing these discrepancies can highlight areas where the process deviated from the intended path and identify potential improvements for future iterations of GCMP.

Development Process Efficiency

Question 6: How did GCMP influence the evolution of the chatbot throughout the development process? Metric 13 (M13): Development time by each version. Tracking the time taken to develop each version of the chatbot allows for an assessment of the overall development efficiency using GCMP. A significant reduction in development time across versions could suggest a streamlining effect of the framework. Metric 14 (M14): Number of releases. The number of releases produced during development can indicate the iterative nature of the process. Frequent releases potentially demonstrate the GCMP framework's ability to facilitate a continuous improvement cycle for the chatbot.

Developer Focus and Workflow

Question 7: How did the development team's focus shift between different aspects of the project (scope, chatbot development, and gamification) throughout the development cycle? Metric 15 (M15): Scope focus by each week. This metric likely involves tracking the amount of time or effort dedicated specifically to managing the project scope during each development week. It can reveal potential scope creep or areas where the GCMP process could provide better guidance. Metric 16 (M16): Chatbot focus by each week. Similar to M15, this metric tracks the focus dedicated to core chatbot development activities throughout each week. Analyzing this data alongside M15 can highlight the balance between scope management and core development. Metric 17 (M17): Gamification focus by each week. Tracking the focus on gamification elements during each week provides insights into the time and effort required for successful gamified chatbot development using GCMP. It can also reveal potential bottlenecks or areas where the GCMP framework could offer more efficient guidance on implementing gamification.

By examining these metrics, we aimed to evaluate the overall quality of GCMP as a development process for creating gamified chatbots. Analyzing adherence to the process, development efficiency, and developer focus enables us to identify strengths and weaknesses of the framework, paving the way for potential improvements in future iterations.

These metrics were selected to provide an objective analysis of the process of developing gamified chatbots using the GCMP, allowing the evaluation of different aspects, from the efficiency of using the process to the engagement and satisfaction of users.

5.3.3 Data collection

This phase involves experimenting with the participation of two software engineering students. Initially, a questionnaire was developed and applied to the participants to identify their profiles and knowledge; this questionnaire is in annex IV. Then, a frequency of follow-up meetings was established, including a meeting to discuss the progress and evolution of the project and an interview to evaluate specific aspects related to scope, gamification, and the chatbot.

Table 5.1 presents each of the metrics, its formal description, the person responsible for data collection, the metric value range, the collection interval, and the collection method. Each column was created according to the guidelines of the GQM.

5.3.4 Interpretation

The interpretation phase analyzes the data about the defined metrics to generate answers to the questions and evaluate if the measurement objectives were achieved. Based on the analysis of the results obtained, actions are taken to address problems or identify opportunities for improvement. These actions may include changes in the development process or team training.

Finally, iterative refinement is regularly applied, involving reviewing and improving established objectives, questions, and metrics. The flexibility of the measurement process allows adaptations to be made according to the specific needs and priorities of the project, ensuring a more effective and suitable approach for obtaining the necessary data.

5.3.5 Remarks

The experiment proved instrumental in refining the Gamified Chatbot Management Process (GCMP). By actively applying GCMP throughout the development process, we were able to identify unforeseen gaps and challenges that were not apparent during the initial design phase. These insights informed subsequent iterations of GCMP, with each version specifically tailored to address the issues encountered. This iterative approach, facilitated by the experiment, has demonstrably strengthened the overall effectiveness of GCMP as a development process.

The application of the Goal Question Metric (GQM) approach proved essential to the structured and organized evaluation of GCMP's effectiveness. GQM facilitated the development of clear goals, targeted questions, and measurable metrics. This structured framework ensured that the experiment gathered relevant data to address the overarching objectives. By providing a systematic approach to data collection and analysis, GQM

Metric	Formal Description	Data Collection Responsible	Value Range	Collection Interval	Collection Method
M1	NLU Score per version	Researcher	0 to 10	Per version	Manual Extraction
M2	Overall score per version	Researcher	0 to 10	Per version	Manual Extraction
M3	Number of utterances per version	Researcher	0 to N	Per version	Manual Extraction
M4	Number of intentions per version	Researcher	0 to N	Per version	Manual Extraction
M5	Utterance usage per version	Researcher	0 to N	Per version	Manual Extraction
M6	Number of intent examples per version	Researcher	0 to N	Per version	Manual Extraction
M7	User interview	Developers	1 to 5	Per phase	Forms
M8	Number of messages per user	Developers	0 to N	Per phase	Forms
M9	Core driver value by each version	Researcher	0 to 10	Per version	Manual Extraction
M10	Number of gamification techniques for core driver by each version	Researcher	0 to N	Per version	Manual Extraction
M11	Developer and researcher feedback	Researcher	Qualitative	Weekly	Forms
M12	Comparison of expected vs. actual activity execution	Researcher	0 to N	Per version	Manual Extraction
M13	Development time per version	Researcher	N days	Per version	Manual Extraction
M14	Number of releases	Researcher	0 to 100	Final Count	Manual Extraction
M15	Scope focus each week	Researcher	0 to 10	Weekly	Forms
M16	Chatbot focus each week	Researcher	0 to 10	Weekly	Forms
M17	Gamification focus each week	Researcher	0 to 10	Weekly	Forms

Table 5.1: Metrics table.

allowed us to gain valuable insights into the strengths and weaknesses of GCMP as a development process.

Chapter 6

Data collect and analysis

6.1 Introduction

This chapter delves into the analysis of data collected during the experiment evaluating the Gamified Chatbot Management Process (GCMP). The analysis utilizes the Goal-Question-Metric (GQM) approach, which provides a structured approach for assessing software development processes.

Following the GQM approach, the experiment was designed with three overarching goals. Each goal focused on a specific aspect of GCMP's effectiveness. This chapter will present the findings for each goal in sequence. For each goal, we will revisit the associated questions formulated during the GQM planning phase. We will then analyze the relevant metrics collected throughout the experiment to answer these questions.

The final section of the chapter will provide a comprehensive analysis of GCMP itself, drawing insights from the findings across all three goals. This analysis will identify the strengths and weaknesses of GCMP as a development process for creating gamified chatbots.

6.2 Goal 1 - Analyze the chatbot generated by GCMP for the purpose of understanding with respect to assertiveness from the researcher's point of view in the context of the research laboratory

Table 6.1 presents the data collected from metrics M1 to M6. These metrics can be used to answer questions Q1 and Q2 below.

Table 6.1: Chatbot metrics by version.

	v1.0	**v2.0**	**v3.0**	**v4.0**	**v5.0**	**v6.0**	**v7.0**	**v8.0**	**v9.0**
M1. NLU score	10	10	10	10	10	8.37	X	X	X
M2. core score	2.1	1.33	4.35	10	9.47	9.17	10	10	10
M3. utterances	12	3	7	25	38	70	68	71	71
M4. intents	7	9	9	10	23	68	56	59	59
M5. intent examples	60	85	85	133	161	233	227	143	137
M6. utterance usages	32	4	20	46	48	48	84	80	80

6.2.1 Q1. How assertive is the chatbot?

The assertiveness of the chatbot is indicated by metrics M1 and M2, which range from 0 to 10 and are extracted by the rasa-mode-reports tool. In the case of M1, which is the assertiveness score of the NLU layer of the chatbot, the score remained at its maximum value in the first 5 versions, but dropped from 10 to 8.37 in the sixth version. In this version, there was a three-fold increase in the amount of chatbot content, and a 30% increase in the number of intent examples. When new content is added or existing content is updated, assertiveness can be impacted. To avoid losses, the activity "6 Chatbot behavior analysis" was created to evaluate this metric and make corrections if necessary.

It is possible to identify a drop in the number of examples of M5 intentions in the last 3 versions of the chatbot, this is due to adjustments made due to the application of activity 6 Chatbot behavior analysis which allowed the identification of intentions that were not being used correctly. Unfortunately, the metrics extraction software stopped working in the last 3 versions of the process application. To indicate this, an "X" was added to the table.

6.2.2 Q2. What is the evolution of chatbot content?

The evolution of the chatbot content can be seen in metrics M3 to M6.

The chatbot's Natural Language Understanding (NLU) score (M1) consistently maintained the maximum score of 10 across the initial five versions, even as the NLU data substantially expanded from 12 to 38 intents. However, in the latest release, the NLU score experienced a decline of 1.63 points from 10 to 8.37, coinciding with a nearly three-fold increase in intents from 23 to 68 and a rise in intent examples from 161 to 233. These outcomes underscore a notable discrepancy in the validation of the newly introduced content, raising the concern of ensuring continuous expansion of the NLU data without compromising its accuracy.

The Core Score (M2) exhibited an upward trajectory, rising from 2.1 to 10 across the initial four iterations. However, this trend reversed in the subsequent two versions,

wherein the Core Score declined from 10 to 9.17. Correspondingly, the chatbot's core layer has expanded, with the number of utterances increasing from 12 to 70 and utterance usages growing from 32 to 48.

The decline observed between the first and second versions of the Core Score can be attributed to the project's utilization of a boilerplate Rasa chatbot project as a foundational basis. Superfluous content was identified and removed in the second version, accounting for decreased Core Score, NLU, and core content.

Remarkably, the chatbot's Natural Language Understanding (NLU) and core data experienced a significant expansion in size. Throughout this growth, the NLU score consistently surpassed the 8-point threshold. Additionally, the core score exhibited a continual upward trajectory, remaining consistently above 9 points across the last three versions. These observations underscore the affirmative influence of the process on the system's performance and sustainability.

6.3 Goal 2 - Analyze the chatbot generated by GCMP for the purpose of understanding with respect to chatbot gamification from the researcher's point of view in the context of the research laboratory

6.3.1 Q3. What is the impact of gamification on user engagement?

Table 6.2 summarizes the results of user interviews (metric M7) to evaluate the chatbot's effectiveness in achieving the defined goal. It presents data for 10 users, including whether they reached the goal, the number of messages exchanged (metric M8), and their feedback on the user experience. All 10 users successfully achieved the goal using the chatbot. The number of messages required ranged from 2 best case to 4 worst case, demonstrating that the chatbot facilitated goal achievement with minimal interaction. This suggests a high degree of efficiency in the chatbot's design.

The time taken by users to reach the goal varied slightly, with a minimum of 54 seconds and a maximum of 77 seconds. The difference in time 23 seconds is relatively small, with an average completion time of 64 seconds. This indicates a consistent performance by the chatbot across different users. While the table doesn't directly display user feedback, the text accompanying it should mention if the feedback was generally positive or highlight any recurring themes from the feedback section.

Table 6.2: M7 and M8. User interview and messages/user.

User	Goal	Time	Messages	Feedback
1	Yes	54 sec	3	4
2	Yes	71 sec	3	4
3	Yes	62 sec	3	5
4	Yes	58 sec	3	5
5	Yes	65 sec	3	4
6	Yes	77 sec	3	4
7	Yes	63 sec	3	4
8	Yes	64 sec	4	3
9	Yes	51 sec	2	5
10	Yes	74 sec	4	5

6.3.2 Q4. What is the evolution of gamification in the chatbot?

Metric M9, as shown in Table 6.3, presents the core driver values used in implementing Octalysis as a gamification framework. This configuration remained consistent across the initial two versions. Subsequently, an adjustment was made in response to administering a questionnaire to the target demographic. This questionnaire aimed to scrutinize the proposed gamification's concordance with the user profiles. Following analysis of the post-questionnaire data, the gamification core drivers are updated to fit the user profile.

It was ascertained that the initial four core drivers exerted a more pronounced influence on the target audience than those within the 5 to 8 range. It was determined to devise gamification techniques tailored to the four core drivers that exhibited greater relevance to the target users.

Table 6.3: M9. Octalysis Core Drivers values by version.

	v1.0	v2.0	v3.0	v4.0	v5.0	v6.0	v7.0	v8.0	v9.0
CD1	3	3	6	6	6	6	6	6	6
CD2	6	6	8	8	8	8	8	8	8
CD3	6	6	8	8	8	8	8	8	8
CD4	3	3	6	6	6	6	6	6	6
CD5	0	0	5	5	5	5	5	5	5
CD6	3	3	4	4	4	4	4	4	4
CD7	6	6	5	5	5	5	5	5	5
CD8	4	4	5	5	5	5	5	5	5

Table 6.4 presents the gamification techniques applied to each core driver across various chatbot versions. In the initial two versions, only one gamification technique was incorporated. However, the number of gamification techniques steadily increased, reaching six techniques implemented in version six of the chatbot, with two techniques allocated to the first core driver, two to the second, and two to the third.

Table 6.4: M10. Gamification Techniques by version.

	v1.0	**v2.0**	**v3.0**	**v4.0**	**v5.0**	**v6.0**	**v7.0**	**v8.0**	**v9.0**
GT1	1	1	1	2	2	2	2	2	2
GT2	0	0	1	2	2	2	2	2	2
GT3	0	0	0	0	1	2	2	3	3
GT4	0	0	0	0	0	0	0	0	0
GT5	0	0	0	0	0	0	0	0	0
GT6	0	0	0	0	0	0	0	0	0
GT7	0	0	0	0	0	0	0	0	0
GT8	0	0	0	0	0	0	0	0	0

The application of gamification techniques also contributed to the chatbot's size expansion, as elucidated in Goal 1. The integration of each technique necessitated the implementation of additional content and behaviors. Furthermore, it proved a non-trivial undertaking, given that adopting gamification techniques for chatbots represents a novel endeavor. This novelty impacted both the quantity of techniques implemented and the pace of their incorporation. Consequently, it was not feasible to implement gamification techniques correlated with the fourth core driver despite its status as one of the four core drivers exerting the most significant influence on the target audience.

6.4 Goal 3 - Analyze the GCMP for the purpose of understanding with respect to quality from the researcher's point of view in the context of the research laboratory

6.4.1 Q5. What is the adherence to the process?

A weekly questionnaire was distributed to the developers to assess adherence to the process, and their responses were gathered and examined. Furthermore, Metric M10 was collected and is illustrated in Figure 6.1. This figure delineates the execution frequency for each activity concerning the anticipated frequency, which was one interaction with the GCMP per version. Given that six versions were developed, the expectation was that each activity would be executed six times.

It is discernible that Activities 2, "Gamification Management", 3, "Modify Knowledge Base", and 4, "Model Training", were the most frequently executed, with Activity 3 being the most commonly performed, totaling 14 executions, which is more than twice the expected frequency. Additionally, Activity 5, "Model Testing", was executed five times,

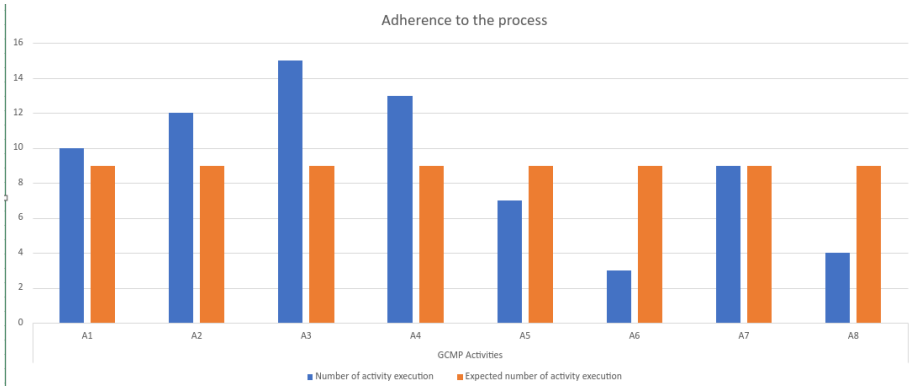


Figure 6.1: M12. Comparison of the execution of activities expected versus performed.

one less than anticipated. On the other hand, Activities 7, "Read Conversations", and 8, "Analyze Metrics", were not executed at all. This aligns with the original CMP process, where these two activities rely on real user conversation data, which falls outside the scope of the current experiment.

However, it became evident that analyzing the chatbot's behavior was necessary, as highlighted in Goal 1. The NLU and core scores declined with the chatbot's content expansion in the last two versions. The current process lacks a dedicated activity for effecting these corrections, thereby underscoring an area for further refinement and development.

6.4.2 Q6. How is the chatbot evolving using GCMP?

Table 6.5 presents the duration in weeks per version. The initial two versions were the lengthiest, spanning seven and six weeks. This extended timeline can be attributed to the necessity of infrastructure configurations, integrations that the team had to undertake during the project's initial phase, and the need to implement rasa custom actions related to the chatbot behaviors. Versions three through six also exhibited variation in the number of weeks per iteration, with the shortest lasting two weeks and the longest extending to five weeks.

Table 6.5: M13. Development time by each version

	v1.0	**v2.0**	**v3.0**	**v4.0**	**v5.0**	**v6.0**	**v7.0**	**v8.0**	**v9.0**
Weeks	7	6	2	3	5	3	1	2	2

Consequently, the chatbot's evolution is continuous but not uniform, with fluctuations in the duration required for each deployment contingent upon the complexity of the behaviors being implemented, both within the chatbot itself and regarding gamification aspects.

In total 9 releases were made during the experiment, this is the value of M14.

6.4.3 Q7. What is the developer's level of focus regarding the scope, chatbot, and gamification each week?

Weekly, one of the questionnaire items addresses the developers' focus concerning the project's scope, gamification activities, and chatbot-related tasks. These questions are subjective and intended to gauge the relevance of each axis within the application of the GCMP. The project was divided into two phases, the first spanning from February to July 2023 and the second from August to October 2023.

Figure 6.2 illustrates the consolidated focus of the developers following the researcher's consolidation of the first phase. In contrast, Figure 6.3 pertains to the second phase, and Figure 6.4 to the third phase.

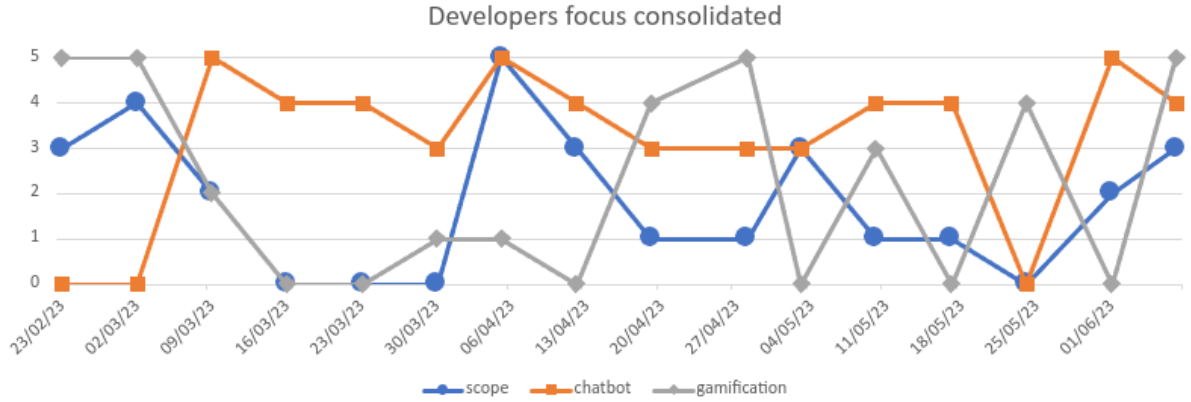


Figure 6.2: M15, M16 and M17 - Focus on phase 1 of the experiment.

In the first phase, a more significant variation in focus is discernible compared to the second phase. Upon scrutinizing the responses submitted by the developers, it becomes apparent that challenges encountered at the project's outset, including infrastructure configuration, familiarization with the Rasa chatbot development framework, adaptation of gamification within the context of chatbots, and adherence to the GCMP, collectively influenced the project's focus and progress, resulting in weekly fluctuations across the three axes.

Conversely, in the second phase, there was less variability compared to the first phase. Notably, the focus on gamification registered as 0 during the deployment weeks while exceeding 4 in the weeks leading up to deployment. This observation suggests an effort in the planning and managing gamification preceding deployment, with less emphasis on gamification during the actual chatbot deployment phase. In the second phase, gamification emerged as the primary focus among the three axes. This observation aligns with

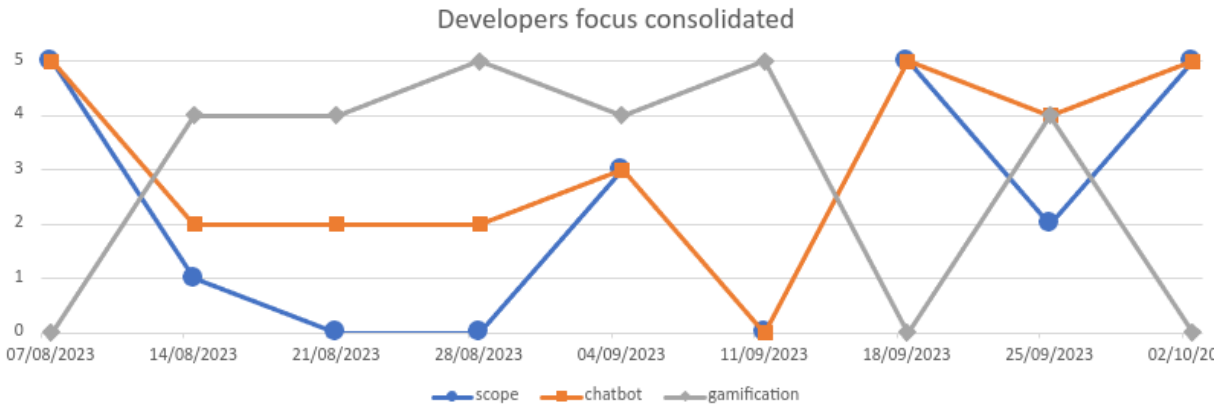


Figure 6.3: M15, M16 and M17 - Focus on phase 2 of the experiment.

Goal 2, wherein a progression from 2 to 6 gamification techniques implemented was noted during the project's second phase.

The analysis of developer focus in the third phase shown in Figure 6.4 reveals a distinct shift in priorities. During the first two weeks, gamification focus remained relatively high, with values of 3 and 5. However, this focus dropped to 0 in the final four weeks. This can be attributed to two key factors. Firstly, a greater emphasis was placed on chatbot adjustment activities, likely including bug fixes, performance optimization, and final deployment preparation. Secondly, much of the core gamification planning had already been completed in the initial weeks, reducing the need for dedicated focus in the later stages.

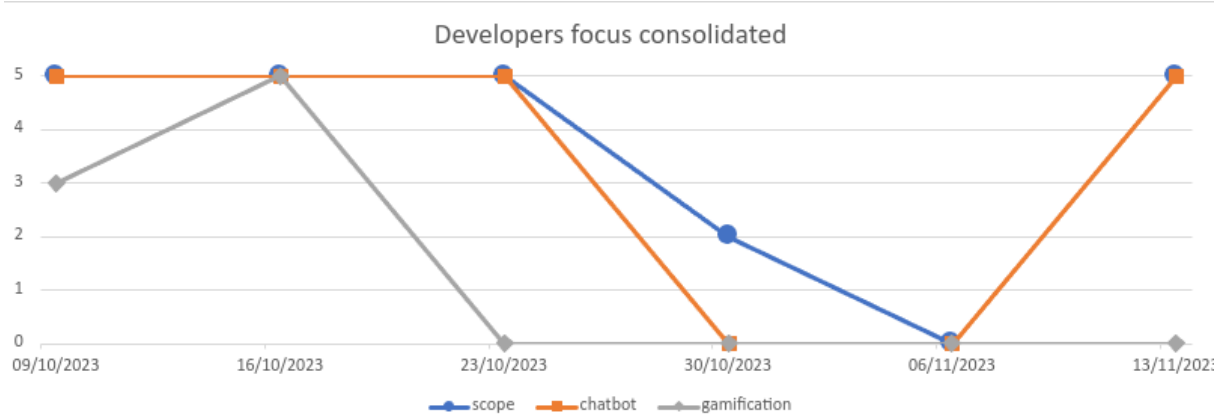


Figure 6.4: M15, M16 and M17 - Focus on phase 3 of the experiment.

Interestingly, the focus on the chatbot and project scope exhibited a similar pattern in phase 3. This suggests a close connection between these two aspects. As new content was implemented in the chatbot, its impact on the overall project scope became more

readily apparent. This potentially reflects the need for ongoing adjustments to the scope throughout the development process to accommodate the evolving chatbot functionalities.

Finally, the analysis also highlights a smaller variation in focus across all three areas (gamification, chatbot, scope) during phase 3 compared to the previous phases. This potentially indicates a more stable development stage, where the development team had established a more consistent workflow for managing the various aspects of the gamified chatbot creation process.

6.4.4 Discussion

This experiment investigated the outcomes of applying the Gamified Chatbot Management Process (GCMP) as a development process. The initial research question centered on the potential benefits of using such a process.

The results were promising. By employing GCMP, we were able to develop a chatbot (Latte) that demonstrated positive user engagement in its initial testing phase. Furthermore, the process provided valuable structure for the development team, effectively guiding them through both core chatbot development activities and gamification implementation.

As detailed in Annex 4, user testing revealed that participants were able to achieve the objectives defined for the gamified chatbot. This achievement serves as positive evidence for the effectiveness of GCMP in creating engaging and functional chatbots.

The experiment also successfully achieved all of its pre-defined objectives:

1. **Propose a Gamified Chatbot Development Process**: The GCMP process was successfully established, building upon the existing Chatbot Mechanics Process (CMP) by incorporating gamification elements.

2. **Validate Process Effectiveness**: The controlled experiment provided valuable data to assess the effectiveness of GCMP in facilitating chatbot development.

3. **Assess Process Implementation**: The experiment monitored the implementation of GCMP by the development team, allowing for ongoing evaluation and refinement.

4. **Evaluate Process Impact on Application**: The impact of GCMP on the final chatbot application (Latte) was analyzed through data collected from the chatbot itself, developer feedback, and user experience evaluations. This multi-faceted analysis confirms the validity of GCMP as a chatbot development process.

The GCMP process itself underwent an iterative development process throughout the experiment. By leveraging the Goal-Question-Metric (GQM) approach, we were able

to systematically evaluate the application of GCMP. This structured analysis provided valuable insights into the effectiveness of specific decisions and adjustments made during the development process. The GQM data, combined with user testing and developer feedback, informed the creation of three distinct versions of GCMP, each addressing gaps and improvements identified in the previous iteration.

Chapter 7

Conclusion

This study has provided compelling evidence for the efficacy of the Gamified Chatbot Management Process (GCMP) as a process for developing engaging chatbots. Building upon the established foundation of the Chatbot Mechanics Process (CMP), GCMP successfully integrates gamification elements, fostering user engagement within chatbot projects.

Furthermore, the experiment yielded valuable insights that informed improvements to GCMP activities, not just for gamified chatbots but for chatbot development processes in general. This iterative refinement, facilitated by the GQM approach, resulted in three distinct versions of GCMP, each addressing gaps identified in the previous iteration.

The GQM analysis revealed significant improvements across various metrics. The third version of GCMP boasted a remarkable 66% reduction in average deployment time compared to the first version. Additionally, it incorporated seven gamification techniques, implemented nearly six times more responses, and achieved the maximum core score of 10. User interviews further corroborated these findings, with 100% of participants successfully achieving the intended results using the chatbot.

These findings, viewed through the lens of various contextual factors relevant to gamified chatbot development, provide strong evidence for the validity of GCMP as a development process. However, it is essential to acknowledge limitations within the experiment's scope. The development team comprised only two individuals with prior gamification experience, operating within an academic context. Future research should explore the applicability of GCMP by varying team profiles, sizes, contexts, and development tools to gain a more comprehensive understanding of its effectiveness.

The scope of the experiment is limited to a team of two developers with experience in gamification and little experience in developing chatbots. This context limits the results, showing the need for more variety in the team's profile, size and experience.

During the design and evolution of GCMP, it became apparent that the underlying CMP process was heavily focused on chatbots developed using the Rasa framework. While the final version of GCMP incorporated adjustments and improvements to facilitate broader applicability, further research is needed to validate its effectiveness with other chatbot development frameworks. This would involve conducting experiments with distinct frameworks, teams, and contexts to assess GCMP's performance in diverse scenarios.

In conclusion, the integration of gamification within the chatbot development process represents a significant stride forward in our quest to enhance user experience and efficacy in digital interactions. As we continue to explore this promising avenue, we hope that the principles and findings of this study will serve as a valuable resource for both researchers and practitioners in the ever-evolving landscape of chatbot technology.

7.1 Future Works

The research conducted in this study has laid a solid foundation for the GCMP framework and its potential applications. However, avenues for future research remain unexplored.

The experiment conducted in this study involved a relatively small development team with specific backgrounds and objectives. To gain a more comprehensive understanding of GCMP's generalizability, future research should explore its application with teams of varying sizes, backgrounds, and objectives. This would involve studying how the process activities and guidelines function in different team dynamics, providing valuable insights into its adaptability and potential for wider adoption.

The user experience evaluation conducted in this study provided valuable initial insights into the engagement and perception generated by the gamified chatbot. However, to gain a deeper understanding of the long-term impact of GCMP on user experience, future research should involve a more extensive, periodic, and comprehensive evaluation. This would involve collecting data over a longer period, tracking changes in user engagement and perception over time, and incorporating a wider range of evaluation methods to capture a more holistic view of the user experience.

Throughout this study, it was not possible to determine whether the proposed process is superior or inferior to the CMP, which served as the foundation for the GCMP. Additionally, there was insufficient data to evaluate the quality of the developed chatbot comprehensively. Future research aimed at comparing the GCMP with the CMP, as well as other chatbot development processes, approaches, or frameworks, is warranted. Such comparative studies is important to assess the contributions and effectiveness of the GCMP relative to established solutions.

By exploring these future research directions, researchers can further refine and validate the GCMP framework, expanding its potential impact on the development of engaging and effective gamified chatbots.

References

[1] Yi-Jen Ho, Siyuan Liu, and Lei Wang. Fun shopping: A randomized field experiment on gamification. *Information Systems Research*, 34(2):766–785, 2023. 2

[2] Angel Torres-Toukoumidis and Mario Mäeots. Implementation of gamification strategies for the enhancement of digital competences. 03 2019. doi: 10.21125/inted.2019.2356. 2

[3] Bosschaerts, Kyana and Stragier, Jeroen and Steenwinckel, Bram and De Marez, Lieven and Van Hoecke, Sofie and Ongenae, Femke. Design of a social chatbot with gamification for user profiling and smoking trigger detection. In *Pervasive Computing Technologies for Healthcare, 16th EAI International Conference, PervasiveHealth 2022, Proceedings*, volume 488, pages 636–650. Springer, 2023. ISBN 9783031345869. URL `http://doi.org/10.1007/978-3-031-34586-9_42`. 2

[4] Giovanni Almeida Santos, Guilherme Guy de Andrade, Geovana Ramos Sousa Silva, Francisco Carlos Molina Duarte, Joao Paulo Javidi Da Costa, and Rafael Timoteo de Sousa. A Conversation-Driven Approach for Chatbot Management. *IEEE Access*, 10:8474–8486, 2022. ISSN 2169-3536. doi: 10.1109/ACCESS.2022.3143323. URL `https://ieeexplore.ieee.org/document/9681834/`. 2, 10, 11, 49

[5] Ahmed Fadhil and Adolfo Villafiorita. An Adaptive Learning with Gamification & Conversational UIs: The Rise of CiboPoliBot. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 408–412, Bratislava Slovakia, July 2017. ACM. ISBN 978-1-4503-5067-9. doi: 10.1145/3099023.3099112. URL `https://dl.acm.org/doi/10.1145/3099023.3099112`. 3, 31, 106

[6] Joseph Ollier, Simon Neff, Christine Dworschak, Arber Sejdiji, Prabhakaran Santhanam, Roman Keller, Grace Xiao, Alina Asisof, Dominik Rüegger, Caterina Bérubé, Lena Hilfiker Tomas, Joël Neff, Jiali Yao, Aishah Alattas, Veronica Varela-Mato, Amanda Pitkethly, M^{a} Dolores Vara, Rocío Herrero, Rosa M^{a} Baños, Carolina Parada, Rajashree Sundaram Agatheswaran, Victor Villalobos, Olivia Clare Keller, Wai Sze Chan, Varun Mishra, Nicholas Jacobson, Catherine Stanger, Xinming He, Viktor von Wyl, Steffi Weidt, Severin Haug, Michael Schaub, Birgit Kleim, Jürgen Barth, Claudia Witt, Urte Scholz, Elgar Fleisch, Florian von Wangenheim, Lorainne Tudor Car, Falk Müller-Riemenschneider, Sandra Hauser-Ulrich, Alejandra Núñez Asomoza, Alicia Salamanca-Sanabria, Jacqueline Louise Mair, and Tobias Kowatsch. Corrigendum: Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol. *Frontiers in Public Health*, 9: 809278, November 2021. ISSN 2296-2565. doi: 10.3389/fpubh.2021.809278. URL

https://www.frontiersin.org/articles/10.3389/fpubh.2021.809278/full. 3, 41, 43, 104, 106

[7] Rini Van Solingen and Egon Berghout. *Goal/question/metric method*. McGraw-Hill, New York, NY, January 1999. 5

[8] A Turing. Computing machinery and intelligence. (Mind 49: 433-460), 1950. 7

[9] Donn Emmanuel Gonda, Jing Luo, Yiu-Lun Wong, and Chi-Un Lei. Evaluation of Developing Educational Chatbots Based on the Seven Principles for Good Teaching. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pages 446–453, Wollongong, NSW, December 2018. IEEE. ISBN 978-1-5386-6522-0. doi: 10.1109/TALE.2018.8615175. URL https://ieeexplore.ieee.org/document/8615175/. 7

[10] Joseph Weizenbaum. ELIZA — a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 26(1):23–28, January 1983. ISSN 0001-0782, 1557-7317. doi: 10.1145/357980.357991. URL https://dl.acm.org/doi/10.1145/357980.357991. 7

[11] Eleni Adamopoulou and Lefteris Moussiades. Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2:100006, December 2020. ISSN 26668270. doi: 10.1016/j.mlwa.2020.100006. URL https://linkinghub.elsevier.com/retrieve/pii/S2666827020300062. 7

[12] Sameera A. and Dr. John. Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 2015. ISSN 21565570, 2158107X. doi: 10.14569/IJACSA.2015.060712. URL http://thesai.org/Publications/ViewPaper?Volume=6&Issue=7&Code=ijacsa&SerialNo=12. 8

[13] Charles Callaway and Khalil Sima'an. Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. *Computational Linguistics*, 32(3):451–452, September 2006. ISSN 0891-2017, 1530-9312. doi: 10.1162/coli.2006.32.3.451. URL https://direct.mit.edu/coli/article/32/3/451-452/1920. 8

[14] Gustavo Marques Mostaço, Leonardo Barreto Campos, and Carlos Eduardo Cugnasca. AgronomoBot: a smart answering Chatbot applied to agricultural sensor networks. page 14, 2018. 8

[15] Saurabh Dutta, Ger Joyce, and Jay Brewer. Utilizing Chatbots to Increase the Efficacy of Information Security Practitioners. In Denise Nicholson, editor, *Advances in Human Factors in Cybersecurity*, pages 237–243, Cham, 2018. Springer International Publishing. ISBN 978-3-319-60585-2. 8

[16] Arthur R. T. de Lacerda and Carla S. R. Aguiar. Floss faq chatbot project reuse: How to allow nonexperts to develop a chatbot. In *Proceedings of the 15th International Symposium on Open Collaboration*, OpenSym '19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450363198. doi:

10.1145/3306446.3340823. URL https://doi.org/10.1145/3306446.3340823. 8, 13, 14, 15

[17] Alan Nichol. Conversation-driven development, Oct 2021. URL https://rasa.com/blog/conversation-driven-development-a-better-approach-to-building-ai-assistants/ 11

[18] Teddy Surya Gunawan, Asaad Balla Falelmula Babiker, Nanang Ismail, and Mufid Ridlo Effendi. Development of Intelligent Telegram Chatbot Using Natural Language Processing. page 5. 12, 13

[19] Patrice Bouvier, Karim Sehaba, and Élise Lavoué. A trace-based approach to identifying users' engagement and qualifying their engaged-behaviours in interactive systems: application to a social game. *User Modeling and User-Adapted Interaction*, 24(5):413–451, December 2014. ISSN 0924-1868, 1573-1391. doi: 10.1007/s11257-014-9150-2. URL http://link.springer.com/10.1007/s11257-014-9150-2. 14

[20] Juho Hamari, Lobna Hassan, and Antonio Dias. Gamification, quantified-self or social networking? Matching users' goals with motivational technology. *User Modeling and User-Adapted Interaction*, 28(1):35–74, March 2018. ISSN 0924-1868, 1573-1391. doi: 10.1007/s11257-018-9200-2. URL http://link.springer.com/10.1007/s11257-018-9200-2. 14

[21] Ana Carolina Tomé Klock, Isabela Gasparini, Marcelo Soares Pimenta, and Juho Hamari. Tailored gamification: A review of literature. *International Journal of Human-Computer Studies*, 144:102495, December 2020. ISSN 10715819. doi: 10.1016/j.ijhcs.2020.102495. URL https://linkinghub.elsevier.com/retrieve/pii/S1071581920300975. 14

[22] Sebastian Deterding, Miguel Sicart, Lennart Nacke, Kenton O'Hara, and Dan Dixon. Gamification. using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, page 2425–2428, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450302685. doi: 10.1145/1979742.1979575. URL https://doi.org/10.1145/1979742.1979575. 15

[23] Sebastian Deterding. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human–Computer Interaction*, 30(3-4):294–335, May 2015. ISSN 0737-0024, 1532-7051. doi: 10.1080/07370024.2014.993471. URL http://www.tandfonline.com/doi/full/10.1080/07370024.2014.993471. 15

[24] Benedikt Morschheuser, Lobna Hassan, Karl Werder, and Juho Hamari. How to design gamification? A method for engineering gamified software. *Information and Software Technology*, 95:219–237, March 2018. ISSN 09505849. doi: 10.1016/j.infsof.2017.10.015. URL https://linkinghub.elsevier.com/retrieve/pii/S095058491730349X. 15

[25] Yu kai Chou. *Actionable Gamification: Beyond Points, Badges, and Leaderboards*. Createspace Independent Publishing Platform, 2015. 16, 67

[26] S. Freitas, A. T. Lacerda, P. O. Calado, T. S. Lima, and E. Dias Canedo. Gamification in education: A methodology to identify student's profile. In *2017 IEEE Frontiers in Education Conference (FIE)*, pages 1–8, Los Alamitos, CA, USA, oct 2017. IEEE Computer Society. doi: 10.1109/FIE.2017.8190499. URL https://doi.ieeecomputersociety.org/10.1109/FIE.2017.8190499. 16, 25, 26

[27] Armando Toda, Ricardo Carmo, Alan Silva, and Seiji Isotani. *GAMIFY-SN: A Metamodel for Planning and Deploying Gamification Concepts Within Social Networks - A Case Study*, pages 1357–1366. 03 2018. ISBN 978-3-319-77711-5. doi: 10.1007/978-3-319-77712-2_130. 18

[28] Gustavo F. Tondello, Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E. Nacke. The Gamification User Types Hexad Scale. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pages 229–243, Austin Texas USA, October 2016. ACM. ISBN 978-1-4503-4456-2. doi: 10.1145/2967934.2968082. URL https://dl.acm.org/doi/10.1145/2967934.2968082. 19, 20

[29] Ana Carolina Tomé Klock, Isabela Gasparini, and Marcelo Soares Pimenta. 5W2H Framework: a guide to design, develop and evaluate the user-centered gamification. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems*, pages 1–10, São Paulo Brazil, October 2016. ACM. ISBN 978-1-4503-5235-2. doi: 10.1145/3033701.3033715. URL https://dl.acm.org/doi/10.1145/3033701.3033715. 20, 21

[30] Antonio Bucchiarone, Antonio Cicchetti, and Annapaola Marconi. GDF: A Gamification Design Framework Powered by Model-Driven Engineering. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pages 753–758, Munich, Germany, September 2019. IEEE. ISBN 978-1-72815-125-0. doi: 10.1109/MODELS-C.2019.00117. URL https://ieeexplore.ieee.org/document/8904523/. 22, 23

[31] Rujianto Eko Saputro, Sazilah Salam, Mohd. Hafiz Zakaria, and Toni Anwar. A gamification framework to enhance students' intrinsic motivation on MOOC. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 17(1):170, February 2019. ISSN 2302-9293, 1693-6930. doi: 10.12928/telkomnika.v17i1.10090. URL http://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/10090. 23

[32] Fitri Marisa, Sharifah Sakinah, Zeratul Izzah, Anastasia L, Ronald David, and Anang Aris. Evaluation of Student Core Drives on e-Learning during the Covid-19 with Octalysis Gamification Framework. *International Journal of Advanced Computer Science and Applications*, 11(11), 2020. ISSN 21565570, 2158107X. doi: 10.14569/IJACSA.2020.0111114. URL http://thesai.org/Publications/ViewPaper?Volume=11&Issue=11&Code=IJACSA&SerialNo=14. 25, 27

[33] Susana de Freitas, Farshid Mahmoodi, and Lennart E Nacke. The games people play: Engineering engagement through human-centered design. In *The 15th International Conference on Entertainment Computing*, pages 489–500. Springer, 2013. 27

[34] Michael Sailer, Kristina Kauper, Christoph Hoffmann, and Patrick Prokop. User-experience-focused gamification design: A framework for supporting playful user experiences. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, pages 1757–1766. IEEE, 2017. 27

[35] Richard Hunicke, Marc LeBlanc, and Robert Zubek. Mda: A framework for analyzing game design elements. In *Proceedings of the Level Up! Conference on Game Design*, pages 117–130. EEE, 2004. 27

[36] Victor R. Basili, Gianluigi Caldiera, and H. Dieter Rombach. The goal question metric approach. URL `https://api.semanticscholar.org/CorpusID:13884048`. 28, 47

[37] Barbara Kitchenham, O. Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. Systematic literature reviews in software engineering – a systematic literature review. 51(1):7–15. ISSN 09505849. doi: 10.1016/j.infsof.2008.09.009. URL `https://linkinghub.elsevier.com/retrieve/pii/S0950584908001390`. 28

[38] Ying Xu and Mark Warschauer. A content analysis of voice-based apps on the market for early literacy development. In *Proceedings of the Interaction Design and Children Conference*, pages 361–371, London United Kingdom, June 2020. ACM. ISBN 978-1-4503-7981-6. doi: 10.1145/3392063.3394418. URL `https://dl.acm.org/doi/10.1145/3392063.3394418`. 31

[39] Kantinee Katchapakirin and Chutiporn Anutariya. An Architectural Design of ScratchThAI: A conversational agent for Computational Thinking Development using Scratch. In *Proceedings of the 10th International Conference on Advances in Information Technology - IAIT 2018*, pages 1–7, Bangkok, Thailand, 2018. ACM Press. ISBN 978-1-4503-6568-0. doi: 10.1145/3291280.3291787. URL `http://dl.acm.org/citation.cfm?doid=3291280.3291787`. 32, 33, 105

[40] Ines Hungerbuehler, Kate Daley, Kate Cavanagh, Heloísa Garcia Claro, and Michael Kapps. Chatbot-Based Assessment of Employees' Mental Health: Design Process and Pilot Implementation. *JMIR Formative Research*, 5(4):e21678, April 2021. ISSN 2561-326X. doi: 10.2196/21678. URL `https://formative.jmir.org/2021/4/e21678`. 34, 35, 104, 105

[41] Stefano Valtolina and Liliana Hu. Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. In *CHItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter*, pages 1–5, Bolzano Italy, July 2021. ACM. ISBN 978-1-4503-8977-8. doi: 10.1145/3464385.3464726. URL `https://dl.acm.org/doi/10.1145/3464385.3464726`. 34, 36, 42, 104, 105

[42] Indra Hidayatulloh, Sigit Pambudi, Herman Dwi Surjono, and Totok Sukardiyono. Gamification on Chatbot-Based Learning Media: a Review and Challenges. *Elinvo (Electronics, Informatics, and Vocational Education)*, 6(1):71–80, October 2021. ISSN 2477-2399, 2580-6424. doi: 10.21831/elinvo.v6i1.43705. URL `https://journal.uny.ac.id/index.php/elinvo/article/view/43705`. 36, 37

[43] Oihab Allal-Chérif, Alba Yela Aránega, and Rafael Castaño Sánchez. Intelligent recruitment: How to identify, select, and retain talents from around the world using artificial intelligence. *Technological Forecasting and Social Change*, 169:120822, August 2021. ISSN 00401625. doi: 10.1016/j.techfore.2021.120822. URL https://linkinghub.elsevier.com/retrieve/pii/S0040162521002547. 36, 106

[44] Aigul Syzdykbayeva, Aigerim Baikulova, and Rysty Kerimbayeva. Introduction of Artificial Intelligence as the Basis of Modern Online Education on the Example of Higher Education. In *2021 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, pages 1–8, Nur-Sultan, Kazakhstan, April 2021. IEEE. ISBN 978-1-72817-470-9. doi: 10.1109/SIST50301.2021.9465974. URL https://ieeexplore.ieee.org/document/9465974/. 38, 103

[45] Simone Sandri, Matthias Hemmje, Huiru Zheng, Felix Engel, Anne Moorhead, Haiying Wang, Raymond Bond, Mike McTear, Andrea Molinari, and Paolo Bouquet. Is there an Optimal Technology to Provide Personal Supportive Feedback in Prevention of Obesity? In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1–6, San Diego, CA, USA, November 2019. IEEE. ISBN 978-1-72811-867-3. doi: 10.1109/BIBM47256.2019.8983405. URL https://ieeexplore.ieee.org/document/8983405/. 38, 39, 103, 106

[46] Han Shi Jocelyn Chew. The Use of Artificial Intelligence–Based Conversational Agents (Chatbots) for Weight Loss: Scoping Review and Practical Recommendations. *JMIR Medical Informatics*, 10(4):e32578, April 2022. ISSN 2291-9694. doi: 10.2196/32578. URL https://medinform.jmir.org/2022/4/e32578. 38, 104, 105

[47] Anton Nijholt. Playable cities: A short survey (Keynote paper). In *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)*, pages 1–6, Himeji, Japan, September 2017. IEEE. ISBN 978-1-5386-1023-7. doi: 10.1109/ICIEV.2017.8338610. URL http://ieeexplore.ieee.org/document/8338610/. 39, 103

[48] Lucy Bray, Ashley Sharpe, Phillip Gichuru, Peter-Marc Fortune, Lucy Blake, and Victoria Appleton. The Acceptability and Impact of the Xploro Digital Therapeutic Platform to Inform and Prepare Children for Planned Procedures in a Hospital: Before and After Evaluation Study. *Journal of Medical Internet Research*, 22(8): e17367, August 2020. ISSN 1438-8871. doi: 10.2196/17367. URL http://www.jmir.org/2020/8/e17367/. 40, 107

[49] Reshawn Ramjattan, Patrick Hosein, and Nigel Henry. Using Chatbot Technologies to Help Individuals Make Sound Personalized Financial Decisions. In *2021 IEEE International Humanitarian Technology Conference (IHTC)*, pages 1–4, United Kingdom, December 2021. IEEE. ISBN 978-1-66544-925-0. doi: 10.1109/IHTC53077.2021.9698928. URL https://ieeexplore.ieee.org/document/9698928/. 40, 41, 42, 103, 106

[50] Jeffrey Dalton, Victor Ajayi, and Richard Main. Vote Goat: Conversational Movie Recommendation. In *The 41st International ACM SIGIR Conference on Research*

& Development in Information Retrieval, pages 1285–1288, Ann Arbor MI USA, June 2018. ACM. ISBN 978-1-4503-5657-2. doi: 10.1145/3209978.3210168. URL https://dl.acm.org/doi/10.1145/3209978.3210168. 41, 42, 105

[51] Alan Nichol. Conversational ai platform | superior customer experiences start here, Oct 2021. URL https://rasa.com/. 67

[52] Barbara Ann Kitchenham and Stuart Charters. Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, 07 2007. URL https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf. 102

[53] Barbara Ann Kitchenham, David Budgen, and Pearl Brereton. *Evidence-based software engineering and systematic reviews*. Chapman & Hall/CRC Innovations in Software Engineering and Software Development Series. CRC Press, London, England, 06 2020. 102

[54] Binh Vu, Sebastian Bruchhaus, Anne Moorhead, Huiru Zheng, Luigi D'Arco, Louise Lynch, Luigia Simona Sica, Michela Ponticorvo, Federico Diano, Haithem Afli, Praveen Joshi, Andrea Molinari, and Matthias Hemmje. Towards continuous professional monitoring of health status based on energetic balancing. In *2022 IEEE International Workshop on Sport, Technology and Research (STAR)*, pages 72–77, 2022. doi: 10.1109/STAR53492.2022.9859599. 103, 106

[55] Sophie Henne, Vanessa Mehlin, Elena Schmid, and Sigurd Schacht. Components of digital assistants in higher education environments. In *2022 IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) & 31st International Association For Management of Technology (IAMOT) Joint Conference*, pages 1–8, 2022. doi: 10.1109/ICE/ITMC-IAMOT55089.2022.10033236. 103, 106

[56] Jaeho Jeon. Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective. *Education and Information Technologies*, 27(4):5767–5787, May 2022. ISSN 1360-2357, 1573-7608. doi: 10.1007/s10639-021-10839-y. URL https://link.springer.com/10.1007/s10639-021-10839-y. 103, 104, 106

[57] Denis Benner, Sofia Schöbel, Christoph Suess, Vera Baechle, and Andreas Janson. Level-up your learning – introducing a framework for gamified educational conversational agents. 2022. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85172008559&partnerID=40&md5=4ce1af738dea5d7f9dd91ce12932acf3. Cited by: 0. 104

[58] C. Potts, E. Ennis, R. B. Bond, M. D. Mulvenna, M. F. McTear, K. Boyd, T. Broderick, M. Malcolm, L. Kuosmanen, H. Nieminen, A. K. Vartiainen, C. Kostenius, B. Cahill, A. Vakaloudis, G. McConvey, and S. O'Neill. Chatbots to Support Mental Wellbeing of People Living in Rural Areas: Can User Groups Contribute to Co-design? *Journal of Technology in Behavioral Science*, 6(4):652–665, December 2021. ISSN 2366-5963. doi: 10.1007/s41347-021-00222-6. URL https://link.springer.com/10.1007/s41347-021-00222-6. 104

[59] Joseph Ollier, Simon Neff, Christine Dworschak, Arber Sejdiji, Prabhakaran Santhanam, Roman Keller, Grace Xiao, Alina Asisof, Dominik Rüegger, Caterina Bérubé, Lena Hilfiker Tomas, Joël Neff, Jiali Yao, Aishah Alattas, Veronica Varela-Mato, Amanda Pitkethly, M^{a} Dolores Vara, Rocío Herrero, Rosa M^{a} Baños, Carolina Parada, Rajashree Sundaram Agatheswaran, Victor Villalobos, Olivia Clare Keller, Wai Sze Chan, Varun Mishra, Nicholas Jacobson, Catherine Stanger, Xinming He, Viktor von Wyl, Steffi Weidt, Severin Haug, Michael Schaub, Birgit Kleim, Jürgen Barth, Claudia Witt, Urte Scholz, Elgar Fleisch, Florian von Wangenheim, Lorainne Tudor Car, Falk Müller-Riemenschneider, Sandra Hauser-Ulrich, Alejandra Núñez Asomoza, Alicia Salamanca-Sanabria, Jacqueline Louise Mair, and Tobias Kowatsch. Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol. *Frontiers in Public Health*, 9:625640, October 2021. ISSN 2296-2565. doi: 10.3389/fpubh.2021.625640. URL `https://www.frontiersin.org/articles/10.3389/fpubh.2021.625640/full`. 104

[60] Robbert Jan Beun, Claire Luiten, Chris Verbeek, and Maartje P. Poelman. A rationale for a gamified e-coach application to decrease the consumption of sugar sweetened beverages. *Frontiers in Digital Health*, 2, 2021. ISSN 2673-253X. doi: 10.3389/fdgth.2020.564529. URL `https://www.frontiersin.org/articles/10.3389/fdgth.2020.564529`. 104, 106

[61] 8th international conference on learning and collaboration technologies, lct 2021, held as part of the 23rd international conference, hci international 2021. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12784 LNCS, 2021. URL `https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112153207&partnerID=40&md5=593a801dd39119b4b980531ba9921ec7`. Cited by: 0. 104

[62] 8th international conference on learning and collaboration technologies, lct 2021, held as part of the 23rd international conference, hci international 2021. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12785 LNCS, 2021. URL `https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112045861&partnerID=40&md5=c1022ad8fc315b2fe25a8db8c59c3b55`. Cited by: 0. 104

[63] Stefano Valtolina and Mattia Marchionna. Design of a chatbot to assist the elderly. *End-User Development*, page 153–168, 2021. doi: 10.1007/978-3-030-79840-6_10. URL `http://dx.doi.org/10.1007/978-3-030-79840-6_10`. 104

[64] Jonas Grund, Moritz Umfahrer, Lea Buchweitz, James Gay, Arthur Theil, and Oliver Korn. A gamified and adaptive learning system for neurodivergent workers in electronic assembling tasks. *Proceedings of Mensch und Computer 2020*, Sep 2020. doi: 10.1145/3404983.3410420. URL `http://dx.doi.org/10.1145/3404983.3410420`. 104

[65] Yigal Rosen, Kristin Stoeffler, Michael Yudelson, and Vanessa Simmering. Towards scalable gamified assessment in support of collaborative problem-solving competency development in online and blended learning. *Proceedings of the Seventh ACM*

Conference on Learning @ Scale, Aug 2020. doi: 10.1145/3386527.3405946. URL http://dx.doi.org/10.1145/3386527.3405946. 104

[66] XueYan (Sue) Chen, Meghna Nayak, Tiffany C Wong, Saba Kawas, and Julie A Kientz. Interaction design & children toolkit. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*, IDC '20, page 256–259, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450380201. doi: 10.1145/3397617.3402031. URL https://doi.org/10.1145/3397617.3402031. 104

[67] 6th ibero-american conference on human-computer interaction, hci-collab 2020. *Communications in Computer and Information Science*, 1334, 2020. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85101567240&partnerID=40&md5=ea0cc10bb1f755063544047671c853d1. Cited by: 0. 104

[68] Evropi Stefanidi, Dimitrios Arampatzis, Asterios Leonidis, Maria Korozi, Margherita Antona, and George Papagiannakis. Magiplay: An augmented reality serious game allowing children to program intelligent environments. *Transactions on Computational Science XXXVII*, page 144–169, 2020. doi: 10.1007/978-3-662-61983-4_9. URL http://dx.doi.org/10.1007/978-3-662-61983-4_9. 105

[69] 6th international conference on electronic governance and open society: Challenges in eurasia, egose 2019. *Communications in Computer and Information Science*, 1135 CCIS, 2020. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85082559489&partnerID=40&md5=e20ab766bdf8f4d3f435a3c4c8ba29f9. Cited by: 0. 105

[70] 12th eai international conference on pervasive computing technologies for healthcare, pervasivehealth 2018. 2018. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85116321310&partnerID=40&md5=3c5134faf351777d5db4d82985704478. Cited by: 0. 105

[71] D. Griol and Z. Callejas. Chapter 15 - conversational agents as learning facilitators: Experiences with a mobile multimodal dialogue system architecture. In Santi Caballé and Robert Clarisó, editors, *Formative Assessment, Learning Data Analytics and Gamification*, Intelligent Data-Centric Systems, pages 313–331. Academic Press, Boston, 2016. ISBN 978-0-12-803637-2. doi: https://doi.org/10.1016/B978-0-12-803637-2.00015-4. URL https://www.sciencedirect.com/science/article/pii/B9780128036372000154. 105

[72] 5th international symposium on ambient intelligence, isami 2014. *Advances in Intelligent Systems and Computing*, 291:1 – 253, 2014. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84928975509&partnerID=40&md5=adcc5ff62588e1a3d678ce4055a941cf. Cited by: 0. 105

[73] Icaart 2014 - proceedings of the 6th international conference on agents and artificial intelligence. volume 2, 2014. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84902350021&partnerID=40&md5=8d5d51e0c04516e80db01f871b9b5c4b. Cited by: 0. 105

[74] Michal Mazur, Rafal Rzepka, and Kenji Araki. Proposal for a conversational english tutoring system that encourages user engagement. pages 10–12, 01 2011. 105

[75] Rodrigo Bavaresco, Diorgenes Silveira, Eduardo Reis, Jorge Barbosa, Rodrigo Righi, Cristiano Costa, Rodolfo Antunes, Marcio Gomes, Clauter Gatti, Mariangela Vanzin, Saint Clair Junior, Elton Silva, and Carlos Moreira. Conversational agents in business: A systematic literature review and future research directions. *COMPUTER SCIENCE REVIEW*, 36, MAY 2020. ISSN 1574-0137. doi: 10.1016/j.cosrev.2020.100239. 105

[76] Truong An Bui, Megan Pohl, Cory Rosenfelt, Tatiana Ogourtsova, Mahdieh Yousef, Kerri Whitlock, Annette Majnemer, David Nicholas, Carrie Demmans Epp, Osmar Zaiane, and Francois Bolduc, V. Identifying potential gamification elements for a new chatbot for families with neurodevelopmental disorders: User-centered design approach. *JMIR HUMAN FACTORS*, 9(3), JUL-SEP 2022. ISSN 2292-9495. doi: 10.2196/31991. 105

[77] David Griol, Jose Manuel Molina, and Zoraida Callejas. Incorporating android conversational agents in m-learning apps. *EXPERT SYSTEMS*, 34(4, SI), AUG 2017. ISSN 0266-4720. doi: 10.1111/exsy.12156. 105

[78] Esteban Vazquez-Cano, Santiago Mengual-Andres, and Eloy Lopez-Meneses. Chatbot to improve learning punctuation in spanish and to enhance open and flexible learning environments. *INTERNATIONAL JOURNAL OF EDUCATIONAL TECHNOLOGY IN HIGHER EDUCATION*, 18(1), JUN 28 2021. ISSN 2365-9440. doi: 10.1186/s41239-021-00269-8. 105

[79] Daphne Economou, Ioannis Doumanis, Frands Pedersen, Paresh Kathrani, Markos Mentzelopoulos, and Vassiliki Bouki. Evaluation of a dynamic role-playing platform for simulations based on octalysis gamification framework. In D Preuveneers, editor, *WORKSHOP PROCEEDINGS OF THE 11TH INTERNATIONAL CONFERENCE ON INTELLIGENT ENVIRONMENTS*, volume 19 of *Ambient Intelligence and Smart Environments*, pages 388–395, 2015. ISBN 978-1-61499-530-2; 978-1-61499-529-6. doi: 10.3233/978-1-61499-530-2-388. 11th International Conference on Intelligent Environments (IE), Prague, CZECH REPUBLIC, JUL 13-14, 2015. 106

[80] Sabina Asensio-Cuesta, Vicent Blanes-Selva, Alberto Conejero, Manuel Portoles, and Miguel Garcia-Gomez. A user-centered chatbot to identify and interconnect individual, social and environmental risk factors related to overweight and obesity. *INFORMATICS FOR HEALTH & SOCIAL CARE*, 47(1):38–52, JAN 2 2022. ISSN 1753-8157. doi: 10.1080/17538157.2021.1923501. 106

[81] Mark S. Dworkin, Sangyoon Lee, Apurba Chakraborty, Colleen Monahan, Lisa Hightow-Weidman, Robert Garofalo, Dima M. Qato, Li Liu, and Antonio Jimenez. Acceptability, feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote hiv medication adherence in young hiv-positive african american msm. *AIDS EDUCATION AND PREVENTION*, 31(1):17–37, FEB 2019. ISSN 0899-9546. doi: 10.1521/aeap.2019.31.1.17. 106

[82] Christian Johnson, Maksim Urazov, and Emma Zanoli. Escapeling: A gamified, ai-supported chatbot for collaborative language practice. In D Guralnick, ME Auer, and A Poce, editors, *INNOVATIONS IN LEARNING AND TECHNOLOGY FOR THE WORKPLACE AND HIGHER EDUCATION*, volume 349 of *Lecture Notes in Networks and Systems*, pages 141–148, 2022. ISBN 978-3-030-90677-1; 978-3-030-90676-4. doi: 10.1007/978-3-030-90677-1_14. 14th Annual Conference on Learning Ideas - Innovations in Learning and Technology for the Workplace and Higher Education, ELECTR NETWORK, JUN 14-18, 2021. 106

[83] Serge Bibauw, Wim Van den Noortgate, Thomas Francois, and Piet Desmet. Dialogue systems for language learning: A meta-analysis. *LANGUAGE LEARNING & TECHNOLOGY*, 26(1), AUG 2022. ISSN 1094-3501. 106

[84] Klinta Epalte, Signe Tomsone, Aivars Vetra, and Guna Berzina. Patient experience using digital therapy "vigo" for stroke patient recovery: a qualitative descriptive study. *DISABILITY AND REHABILITATION-ASSISTIVE TECHNOLOGY*, 18 (2):175–184, FEB 17 2023. ISSN 1748-3107. doi: 10.1080/17483107.2020.1839794. 106

[85] Sahar Sajadieh and Hannah Wolfe. Come hither to me performance of a seductive robot. In *CHI EA '19 EXTENDED ABSTRACTS: EXTENDED ABSTRACTS OF THE 2019 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS*. Assoc Comp Machinery; ACM SIGCHI, 2019. ISBN 978-1-4503-5971-9. doi: 10.1145/3290607.3313287. CHI Conference on Human Factors in Computing Systems (CHI), Glasgow, SCOTLAND, MAY 04-09, 2019. 106

[86] Farbod Aprin, Irene-Angelica Chounta, and H. Hoppe. "see the image in different contexts": Using reverse image search to support the identification of fake news in instagram-like social media. 06 2022. ISBN 978-3-031-09679-2. doi: 10.1007/978-3-031-09680-8_25. 106

Appendix I

Literature review

According to [52], a Systematic Literature Review (SLR) is a method for identifying, analyzing, and interpreting available evidence related to a specific research question, field, or area of interest. An SLR comprises a set of nine activities organized into three phases: Review Planning, Review Execution, and Review Documentation.

Phase 1, Review Planning, addresses how the study should be conducted, and it involves three activities: Specification of research questions, Development of the review protocol, and Validation of the protocol. Phase 2, Review Execution, encompasses the following activities: Identification of relevant articles, Selection of primary studies, Quality assessment, Data extraction, and Data synthesis. The definitions outlined in the protocol are applied, and any discrepancies necessitating protocol adjustments must be documented. Phase 3, Review Documentation, involves recording the processes and outcomes of the review [53]).

In this document, we describe the application of a literature review based on the systematic review outlined above, albeit in a more simplified format. This section will present the state-of-the-art relevant studies on the chatbot and gamification intersection. To understand what is already known from this Intersection, the following search string was used:

Search string: **(CHATBOT OR CONVERSATIONAL AGENT) AND (GAMIFICATION)**.

This string was applied in the following bases:

1. IEEE
2. SCOPUS
3. Web of Science

Below is described the selection criteria for the researched articles:

1. **Gamification not presented**: This criterion indicates that the article should not be selected if it does not mention or discuss the gamification concept. In other words, the article should have some content related to gamification to be considered.

2. **Chatbot not presented**: This criterion specifies that the article should be excluded from the selection if it lacks any reference to chatbots. The article must feature some discussion or content related to chatbot technology. **Gamification and Chatbot not presented**: This criterion highlights that an article should be disregarded if it lacks gamification and chatbot-related content. In other words, the article should contain at least one of these elements to be considered relevant for the study.

3. **Not an article**: This criterion filters out materials not qualifying as traditional articles, such as blog posts, forum discussions, or other non-academic sources.

4. **Not accessible**: This criterion is employed when the article is inaccessible or unavailable for review, potentially due to access permissions, publication status, or unavailability of the full text.

5. **Duplicated**: When multiple instances of the same article are encountered, this criterion excludes duplicate entries and streamlines the selection process, ensuring each unique source is considered only once.

Table I.1 containing the complete list of articles:

Id	Title	Reference
A1	Using Chatbot Technologies to Help Individuals Make Sound Personalized Financial Decisions	[49]
A2	Introduction of Artificial Intelligence as the Basis of Modern Online Education on the Example of Higher Education	[44]
A3	Is there an Optimal Technology to Provide Personal Supportive Feedback in Prevention of Obesity?	[45]
A4	Playable cities: A short survey (Keynote paper)	[47]
A5	Towards Continuous Professional Monitoring of Health Status Based on Energetic Balancing	[54]
A6	Components of digital assistants in higher education environments	[55]
A7	Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective	[56]

A8	Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective	[56]
A9	The Use of Artificial Intelligence–Based Conversational Agents (Chatbots) for Weight Loss: Scoping Review and Practical Recommendations	[46]
A10	Level-Up your Learning – Introducing a Framework for Gamified Educational Conversational Agents	[57]
A11	Chatbots to Support Mental Wellbeing of People Living in Rural Areas: Can User Groups Contribute to Co-design?	[58]
A12	Corrigendum: Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol (Front. Public Health, (2021), 9, (625640), 10.3389/fpubh.2021.625640)	[6]
A13	Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol	[59]
A14	Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness	[41]
A15	Chatbot-based assessment of employees' mental health: Design process and pilot implementation	[40]
A16	A Rationale for a Gamified E-Coach Application to Decrease the Consumption of Sugar Sweetened Beverages	[60]
A17	8th International Conference on Learning and Collaboration Technologies, LCT 2021, held as Part of the 23rd International Conference, HCI International 2021	[61]
A18	8th International Conference on Learning and Collaboration Technologies, LCT 2021, held as Part of the 23rd International Conference, HCI International 2021	[62]
A19	Design of a Chatbot to Assist the Elderly	[63]
A20	A gamified and adaptive learning system for neurodivergent workers in electronic assembling tasks	[64]
A21	Towards Scalable Gamified Assessment in Support of Collaborative Problem-Solving Competency Development in Online and Blended Learning	[65]
A22	Extended Abstracts - Proceedings of the 2020 ACM Interaction Design and Children Conference, IDC 2020	[66]
A23	6th Ibero-American Conference on Human-Computer Interaction, HCI-Collab 2020	[67]

A24	MagiPlay: An Augmented Reality Serious Game Allowing Children to Program Intelligent Environments	[68]
A25	6th International Conference on Electronic Governance and Open Society: Challenges in Eurasia, EGOSE 2019	[69]
A25	An Architectural Design of ScratchThAI A conversational agent for Computational Thinking Development using Scratch	[39]
A26	Vote goat: Conversational movie recommendation	[50]
A27	12th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2018	[70]
A28	Conversational Agents as Learning Facilitators: Experiences With a Mobile Multimodal Dialogue System Architecture	[71]
A29	5th International Symposium on Ambient Intelligence, ISAmI 2014	[72]
A30	ICAART 2014 - Proceedings of the 6th International Conference on Agents and Artificial Intelligence	[73]
A31	ICAART 2014 - Proceedings of the 6th International Conference on Agents and Artificial Intelligence	[73]
A32	Proposal for a conversational English tutoring system that encourages user engagement	[74]
A33	Conversational agents in business: A systematic literature review and future research directions	[75]
A34	Identifying Potential Gamification Elements for A New Chatbot for Families With Neurodevelopmental Disorders: User-Centered Design Approach	[76]
A35	Chatbot-Based Assessment of Employees' Mental Health: Design Process and Pilot Implementation	[40]
A36	The Use of Artificial Intelligence-Based Conversational Agents (Chatbots) for Weight Loss: Scoping Review and Practical Recommendations	[46]
A37	Incorporating android conversational agents in m-learning apps	[77]
A38	Vote Goat: Conversational Movie Recommendation	[50]
A39	Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments	[78]
A40	An Architectural Design of ScratchThAI A conversational agent for Computational Thinking Development using Scratch	[39]
A41	Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness	[41]

A42	An Adaptive Learning with Gamification & Conversational UIs: The Rise of CiboPoliBot	[5]
A43	Using Chatbot Technologies to help Individuals make Sound Personalized Financial Decisions	[49]
A44	Evaluation of a dynamic role-playing platform for simulations based on Octalysis gamification framework	[79]
A45	A user-centered chatbot to identify and interconnect individual, social and environmental risk factors related to overweight and obesity	[80]
A46	Is there an Optimal Technology to Provide Personal Supportive Feedback in Prevention of Obesity?	[45]
A47	acceptability, feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote HIV medication adherence in young HIV-positive African American MSM	[81]
A48	Elena plus Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol	[6]
A49	Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective	[56]
A50	Escapeling: A Gamified, AI-Supported Chatbot for Collaborative Language Practice	[82]
A51	Dialogue systems for language learning: A meta-analysis	[83]
A52	A Rationale for a Gamified E-Coach Application to Decrease the Consumption of Sugar Sweetened Beverages	[60]
A53	Components of digital assistants in higher education environments	[55]
A54	Elena+ Care for COVID-19, a Pandemic Lifestyle Care Intervention: Intervention Design and Study Protocol (vol 9, 625640,2021)	[6]
A55	Towards Continuous Professional Monitoring of Health Status Based on Energetic Balancing	[54]
A56	Intelligent recruitment: How to identify, select, and retain talents from around the world using artificial intelligence	[43]
A57	Patient experience using digital therapy "Vigo" for stroke patient recovery: a qualitative descriptive study	[84]
A58	Come Hither to Me Performance of a Seductive Robot	[85]
A59	"See the Image in Different Contexts": Using Reverse Image Search to Support the Identification of Fake News in Instagram-Like Social Media	[86]

A60	The Acceptability and Impact of the Xploro Digital Therapeutic Platform to Inform and Prepare Children for Planned Procedures in a Hospital: Before and After Evaluation Study	[48]

Table I.1: Searched articles

Table I.2 delineates the articles upon which the research papers were predicated, indicating whether each article was chosen for inclusion and, if not, elucidate the rationale behind its exclusion.

Id	Article base	Selected	Reason not selected
A1	IEEE	Yes	-
A2	IEEE	Yes	-
A3	IEEE	Yes	-
A4	IEEE	Yes	-
A5	IEEE	Yes	-
A6	IEEE	Yes	-
A7	Scopus	No	Gamification not presented
A8	Scopus	Yes	-
A9	Scopus	Yes	-
A10	Scopus	Yes	-
A11	Scopus	Yes	-
A12	Scopus	No	- Duplicated
A13	Scopus	Yes	-
A14	Scopus	Yes	-
A15	Scopus	No	Not an article
A16	Scopus	No	Not an article
A17	Scopus	No	Not an article
A18	Scopus	Yes	-
A19	Scopus	No	Not accessible
A20	Scopus	No	Not accessible
A21	Scopus	No	Not an article
A22	Scopus	No	Not an article
A23	Scopus	Yes	-
A24	Scopus	No	Not an article
A25	Scopus	Yes	-
A26	Scopus	Yes	-

A27	Scopus	No	Not an article
A28	Scopus	Yes	-
A29	Scopus	No	Not an article
A30	Scopus	No	Not an article
A31	Scopus	No	Not an article
A32	Scopus	Yes	-
A33	Web of Science	Yes	-
A34	Web of Science	Yes	-
A35	Web of Science	Yes	-
A36	Web of Science	Yes	-
A37	Web of Science	Yes	-
A38	Web of Science	No	Duplicated
A39	Web of Science	No	Gamification not presented
A40	Web of Science	No	Duplicated
A41	Web of Science	No	Duplicated
A42	Web of Science	Yes	-
A43	Web of Science	No	Duplicated
A44	Web of Science	Yes	-
A45	Web of Science	Yes	-
A46	Web of Science	No	Duplicated
A47	Web of Science	No	Not accessible
A48	Web of Science	No	Duplicated
A49	Web of Science	Yes	-
A50	Web of Science	No	Not accessible
A51	Web of Science	No	Not accessible
A52	Web of Science	Yes	-
A53	Web of Science	Yes	-
A54	Web of Science	No	Duplicated
A55	Web of Science	Yes	-
A56	Web of Science	Yes	-
A57	Web of Science	Yes	-
A58	Web of Science	Yes	-
A59	Web of Science	No	Not accessible
A60	Web of Science	Yes	-

Table I.2: Selected articles

Appendix II

GCMP manual

The Gamified Chatbot Management Process (GCMP) is a development process designed to guide the implementation of gamified chatbot solutions. This process is based on the Chatbot Management Process (CMP). For more details, refer to the scientific article at the following link: `https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9681834`.

The visual organization of the GCMP process is shown in the image II.1.

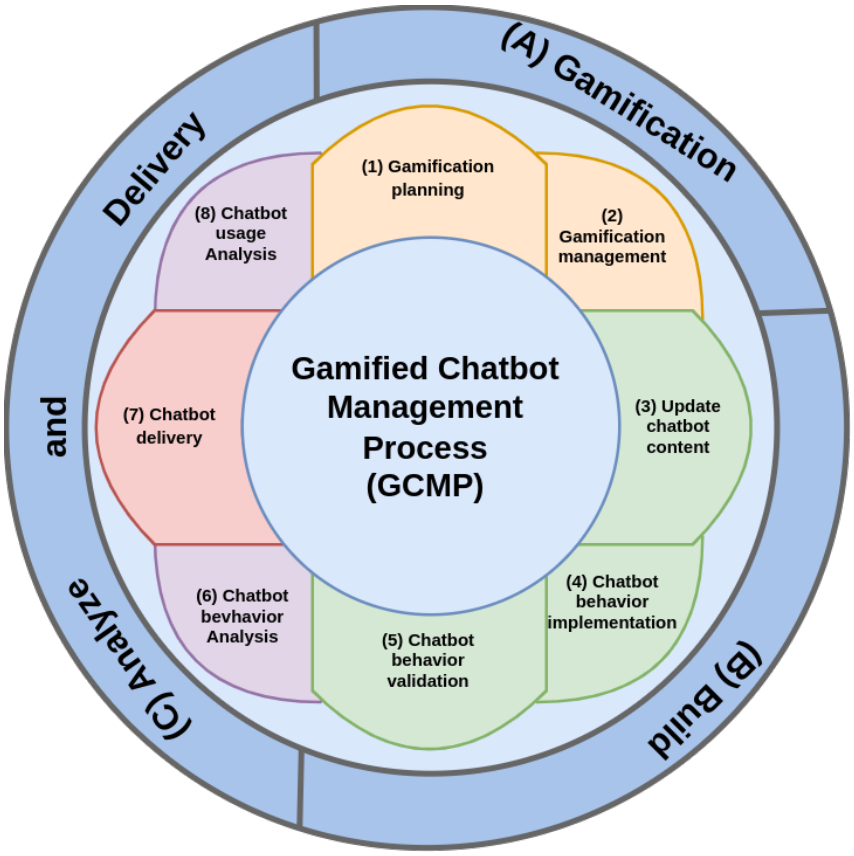


Figure II.1: GCMP

II.1 Phases and steps description

II.1.1 (A) Gamification

The gamification phase aims to address the development and evolution of the application of gamification in the chatbot, ensuring that gamification techniques are implemented in a manner that aligns with the intended format and purpose.

(1) Gamification planning

The Gamification planning step is applied in each GCMP iteration. Still, the first iteration will bring new information about the context, which should be updated in future iterations as needed. It is important to manage the following information:

1. **Business Metrics which lead to Game Objective**: many chatbots have a FAQ behavior. Stakeholders may want the user to find the solution and avoid the actuation of a human, so increasing chatbot retention could be a game objective.

2. **Users, which lead to Players**: Defining the target users and raising the user gamification profile is significant to ensure that the designed gamification will impact each user.

3. **Desired Actions, which lead to win-states**: To define a win-state, the desired actions that the player should achieve are needed. All chatbot messages are essential to guide the user to the desired point.

4. **Feedback Mechanics, which lead to triggers**: feedback mechanics must be designed to tell the user that their conversation is meaningful. All feedback mechanics should become triggers to promote the desired actions.

5. **Incentives, which lead to Rewards**: Finding what the chatbot can give users when they commit the desired actions and arrive at the Win-state is the objective of this item.

Each one of the gamification techniques that most impact the defined core drives should be listed, and the application planned to be applied in the chatbot solution. The plan should be revised and adapted in the subsequent iterations according to the gamification data and insights that emerge from the solution development.

(2) Gamification management

With the initial or rolling gamification plan as income, managing the user's gamification profile to ensure the gamification will have the desired result. Updates on it should impact the usage and add or remove game techniques. Gamification management is affected by the gamification framework chosen, changing how gamification techniques are applied.

All the phases, target player profile, and game techniques should be clear to drive the chatbot conversation behavior on the following GCMP activities.

II.1.2 (B) Build

The build phase primarily focuses on constructing and implementing content within the chatbot, whether it is gamified or not. Additionally, adjustments are made in line with the hypotheses generated during the analysis activity, along with the inclusion of relevant content tailored to the chatbot's context.

(3) Update chatbot content

The chatbot's content update involves adding, removing, or editing content in the chatbot, both in the natural language understanding (NLU) and natural language generation (NLG) components. This step varies according to the chatbot framework chosen. This step results in the new content being added to the chatbot, ready for training, or saved by the chatbot framework.

(4) Chatbot behavior implementation

Following the content update, it is necessary to implement these contents into the chatbot solution. Implementing code responsible for integrating with external services or triggering behaviors may be needed to achieve the desired results of gamification or dialogue. This step varies according to the chatbot framework. In machine-learning chatbots, the dialogue model is trained in this step.

(5) Chatbot behavior validation

This step ensures that the new content, behaviors, and messages are correctly presented to the users. It is also essential to verify that the gamification techniques are behaving as expected to maintain a positive user experience. This step varies according to the chatbot framework and may be needed to run automatized dialog test routines.

It's essential to run conversations to the chatbot, simulating user's behaviors to ensure the desired solution works correctly in the chatbot. If any issue is found with the dialog

related to the chatbot behavior or the gamification application, it is necessary to rerun the steps.

II.1.3 (C) Analyze and Delivery

This phase is responsible for delivering the gamified chatbot solution to the user through the delivery and analysis of the chatbot. This analysis includes examining the conversations to understand the chatbot's behavior and evaluating the chatbot's model metrics.

(6) Chatbot behavior analysis

Prior to the execution of the delivery phase, it is imperative to conduct a verification of the accuracy of the chatbot model metrics. Identification of any issues demands an assessment of their severity, and if deemed necessary, prompt corrective measures should be implemented before the final delivery. This process diverges from a mere analysis of model metrics in cases where the chatbot incorporates such metrics, extending to the consideration of framework adjustments and feedback integration.

In essence, a more holistic approach is warranted to ensure the seamless functionality of all behaviors of the chatbot, encompassing integrations, code structure, gamification, and other pertinent elements, guaranteeing the chatbot solution's comprehensive readiness for delivery.

(7) Chatbot delivery

It is deploying the chatbot solution to its intended users. The nature of this deployment varies depending on the selected chatbot framework. In certain frameworks, the delivery process is simplified to a mere button press, whereas in alternative frameworks, it necessitates the comprehensive configuration of the entire solution infrastructure.

Regardless of the specificities inherent to the chosen framework, the paramount consideration is ascertaining that the solution is updated and operational for the end users.

(8) Chatbot usage analysis

The analysis step is focused on examining user behavior when using the chatbot. Therefore, if possible, read all the conversations, but it is common to read randomly selected conversations or seek specific behaviors, in addition to monitoring metrics, data, and charts such as:

1. Evaluation of user interactions with the chatbot;
2. Assessment of specific user behaviors and patterns;

3. Tracking relevant metrics, data, and charts.

The goal is to gain insights into how users interact with the chatbot, identify any areas for improvement, and assess the solution's overall effectiveness, not only to the chatbot behavior but the effectiveness of the gamification application, too.

Appendix III

GQM data

III.1 Planning

To evaluate the use of GCMP as a software development process, we utilized the GQM framework to establish goals, identify questions, and define metrics for the process. Initially, the general objectives were determined using GCMP in gamified chatbot development. Figure **??** is the GQM diagram used.

III.1.1 GQM Plan

The following objectives were established:

- **G1:** Analyze the chatbot generated by GCMP for the researcher's understanding and assertiveness in the context of the research laboratory.
- **G2:** Analyze the chatbot generated by GCMP for the researcher's understanding of gamification in the context of the research laboratory.
- **G3:** Analyze GCMP for quality from the researcher's perspective in the context of the research laboratory.

Next, we formulated questions to aid in understanding different aspects of the software development process. The questions raised were as follows:

- **Q1.** How assertive is the chatbot?
- **Q2.** What is the evolution of the chatbot's content?
- **Q3.** What is the user satisfaction level with the chatbot?
- **Q4.** What is the impact of gamification on user engagement?

- **Q5.** How is gamification evolving in the chatbot?
- **Q6.** What is the adherence to the process?
- **Q7.** How is the chatbot evolving using GCMP?
- **Q8.** What is the developer's focus level regarding scope, chatbot, and gamification each week?

Subsequently, we identified metrics, i.e., selected metrics capable of providing quantitative or qualitative data to answer the raised questions. These metrics should be meaningful, objective, and measurable. The selected metrics were as follows:

- **M1.** Natural Language Understanding (NLU) Score for each version.
- **M2.** Overall score for each version.
- **M3.** Number of utterances for each version.
- **M4.** Number of intentions for each version.
- **M5.** Utterance usage for each version.
- **M6.** Number of intent examples for each version.
- **M7.** Net Promoter Score (NPS).
- **M8.** Number of messages per user.
- **M9.** Value of each principal driver for gamification per version.
- **M10.** Number of gamification techniques per principal driver per version.
- **M11.** Developer and researcher feedback.
- **M12.** Comparison of expected vs. actual activity execution.
- **M13.** Development time per version.
- **M14.** Number of releases.
- **M15.** Scope focus each week.
- **M16.** Chatbot focus each week.
- **M17.** Gamification focus each week.

These metrics were selected to provide an objective analysis of the gamified chatbot development process using GCMP, allowing the evaluation of various aspects, from process efficiency to user engagement and satisfaction.

Next, we established a measurement procedure for collecting the data defined by GQM. This procedure involves experimenting with the participation of two software engineering students. Initially, we created a questionnaire and administered it to the participants to identify their profiles and knowledge. Then, we established a schedule of follow-up meetings, including a meeting to discuss the project's progress and evolution and an interview to assess specific aspects related to scope, gamification, and chatbot.

In addition to interviews, we extracted project progress metrics from the version control tool, such as GitHub. We also collected metrics from the generated artificial intelligence model, engagement metrics (such as NPS), and chatbot usage metrics.

III.1.2 Measurement Plan

The table below provides the necessary information for the measurement plan of the experiment:

III.2 Analysis Plan

Based on the analysis of the results obtained, we will take measures to address challenges or identify improvement opportunities. These actions may involve adjustments to the development process or team training. Finally, we apply iterative refinement regularly, reviewing and enhancing the established goals, questions, and metrics. The flexibility of the goals and metrics analysis process allows us to make adaptations as needed to project-specific requirements and priorities, ensuring a more effective and appropriate approach to obtaining the necessary data.

Graphs and visualizations are created and evaluated for each version to track project and team evolution. These analyses generate insights and new hypotheses to adjust GQM and the evaluated process, GCMP.

Metric	Formal Description	Data Collection Responsible	Value Range	Collection Interval	Collection Method
M1.	NLU Score per version	Researcher	0 to 10	Per version	Manual Extraction
M2.	Overall score per version	Researcher	0 to 10	Per version	Manual Extraction
M3.	Number of utterances per version	Researcher	0 to N	Per version	Manual Extraction
M4.	Number of intentions per version	Researcher	0 to N	Per version	Manual Extraction
M5.	Utterance usage per version	Researcher	0 to N	Per version	Manual Extraction
M6.	Number of intent examples per version	Researcher	0 to N	Per version	Manual Extraction
M7.	User interview	Developers	1 to 5	Per phase	Forms
M8.	Number of messages per user	Developers	0 to N	Per phase	Forms
M9.	Core driver value by each version	Researcher	0 to 10	Per version	Manual Extraction
M10.	Number of gamification techniques for core driver by each version	Researcher	0 to N	Per version	Manual Extraction
M11.	Developer and researcher feedback	Researcher	Qualitative	Weekly	Forms
M12.	Comparison of expected vs. actual activity execution	Researcher	0 to N	Per version	Manual Extraction
M13.	Development time per version	Researcher	N days	Per version	Manual Extraction
M14.	Number of releases	Researcher	0 to 100	Final Count	Manual Extraction
M15.	Scope focus each week	Researcher	0 to 10	Weekly	Forms
M16.	Chatbot focus each week	Researcher	0 to 10	Weekly	Forms
M17.	Gamification focus each week	Researcher	0 to 10	Weekly	Forms

Table III.1: Measurement table

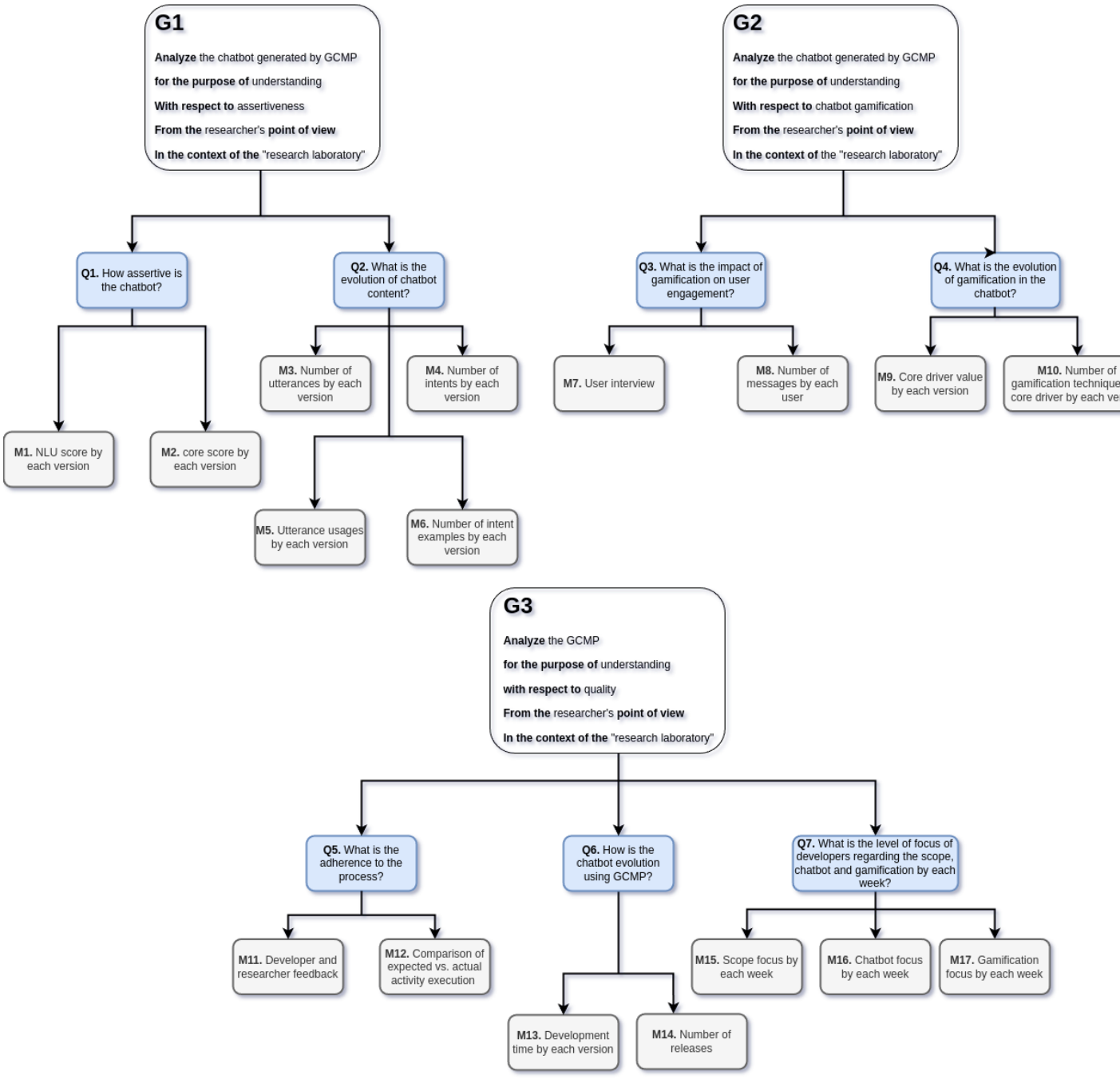


Figure III.1: GQM

Appendix IV

Questionaries

In the appendix, the questionnaires used during the course of the master's program are incorporated. These questionnaires play are important in the systematic collection of data relevant to the research at hand. The first instrument, titled Team Forms was designed to identify the profile and obtain information about the development team responsible for implementing the Gamified Collaborative Modeling Process (GCMP) during the experiment. This questionnaire was distributed on two separate occasions: one before the start of the project and another after the completion of the experiment. The repetition of the questionnaire allows for a longitudinal analysis of the responses, enabling the evaluation of the team's evolution over the experimental period.

The second questionnaire, called User Profile was developed by the development team itself to discern the gamification profile of the users involved in the experiment. Responses from individuals previously identified as part of the target audience provide valuable insights for the adaptation or updating of the employed gamification profile, using the Octalysis framework.

The third questionnaire, titled Users Feedback was implemented in two distinct phases. In the first phase, five individuals previously categorized as belonging to the target audience participated, who were tasked with achieving a specific goal through interaction with the Latte chatbot. Data were meticulously extracted regarding the use of the chatbot by these users. In the second phase, the same five participants from the previous phase repeated the interaction with the chatbot, now accompanied by five new users, all subjected to the most updated version of the chatbot. The purpose of this approach is to obtain information from individuals who have already interacted with the chatbot and, simultaneously, capture data from new users in the latest version of the system. This data collection strategy aims to enrich the overall understanding of the impact and effectiveness of the chatbot in different contexts and phases of the experiment.

IV.1 Team forms

IV.1.1 Questionary

Perfil do Desenvolvedor

Este formulário tem como objetivo levantar o perfil do desenvolvedor que vai participar do desenvolvimento de um chatbot gamificado utilizando o Gamified Chatbot Management Process (GCMP).

Os dados desde questionário são anônimos.

* Indica uma pergunta obrigatória

1. E-mail *

Questões demográficas

Nesta seção informções demográficas são coletadas.

2. Qual a sua idade? *

Marcar apenas uma oval.

- ◯ menos de 16 anos
- ◯ 16 - 20 anos
- ◯ 21 - 30 anos
- ◯ 31 - 40 anos
- ◯ acima de 40 anos

3. Qual o seu grau de instrução? *

Marcar apenas uma oval.

- ◯ Ensino fundamental incompleto
- ◯ Ensino fundamental completo
- ◯ Ensino médio completo
- ◯ Ensino superior completo
- ◯ Pós graduação, mestrado, doutorado, etc

Conhecimento técnico

Nesta seção informções de conhecimento técnico são coletadas.

4. Quantos anos de experiência como desenvolvedor? *

Marcar apenas uma oval.

- ◯ Nenhuma experiência
- ◯ 2 a 5 anos
- ◯ 5 a 10 anos
- ◯ mais de 10 anos

5. Com quais linguagens de programção você já trabalhou ou trabalha? *

Marque todas que se aplicam.

- ☐ Java
- ☐ Python
- ☐ Javascript
- ☐ PHP
- ☐ Ruby
- ☐ Go
- ☐ C++
- ☐ Nenhuma
- ☐ Outro: _____

6. Como você classifica seu conhecimento em relação ao desenvolvimento de chatbots? *

Marcar apenas uma oval.

- ◯ Nenhum
- ◯ Muito baixo
- ◯ Baixo
- ◯ Neutro
- ◯ Alto
- ◯ Muito alto

7. Com quais frameworks de chatbot você já trabalhou? *

Marque todas que se aplicam.

- ☐ DialogFlow
- ☐ IBM Watson
- ☐ Rasa
- ☐ Take Blip
- ☐ Nenhum
- ☐ Outro: _____

Conhecimento de Gamificação

Nesta seção informções de conhecimento sobre gamificação são coletadas.

8. Quantos anos de experiência com projetos de gamificação? *

Marcar apenas uma oval.

- ◯ Nenhuma experiência
- ◯ 2 a 5 anos
- ◯ 5 a 10 anos
- ◯ mais de 10 anos

9. Como você classifica seu conhecimento em relação a gamificação? *

Marcar apenas uma oval.

- ◯ Nenhum
- ◯ Muito baixo
- ◯ Baixo
- ◯ Neutro
- ◯ Alto
- ◯ Muito alto

10. Com quais frameworks de gamificação você já trabalhou? *

Marque todas que se aplicam.

- ☐ Hexad
- ☐ Octalysis
- ☐ Gamification Design Framework (GDF)
- ☐ Nenhum
- ☐ Outro: _____

IV.1.2 Questionaries responses

Perfil do Desenvolvedor

4 respostas

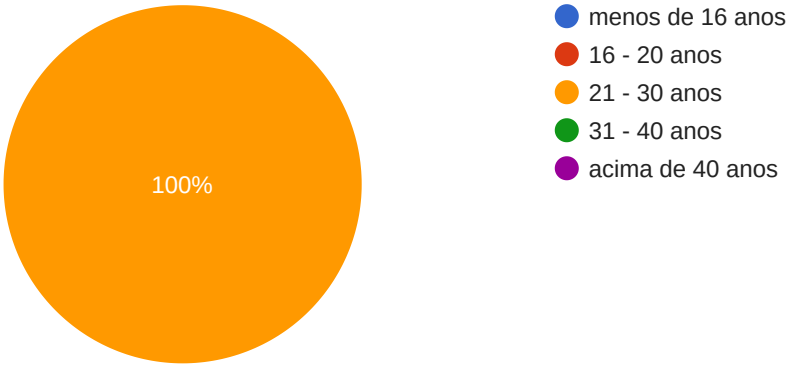
Publicar análise

Questões demográficas

Qual a sua idade?

Copiar

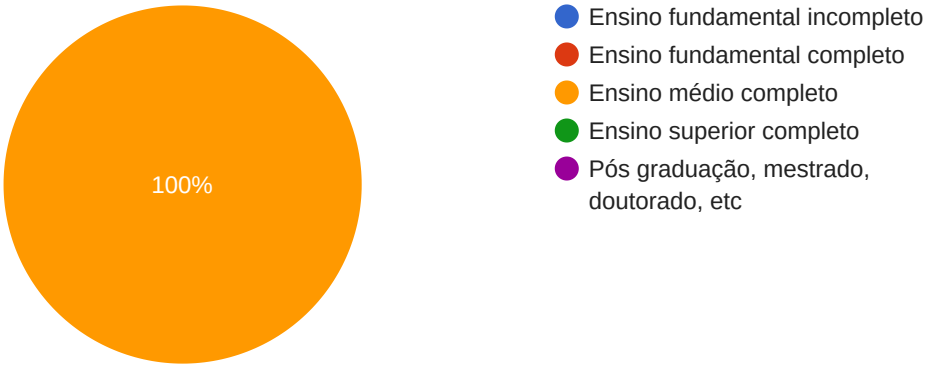
4 respostas



Qual o seu grau de instrução?

Copiar

4 respostas

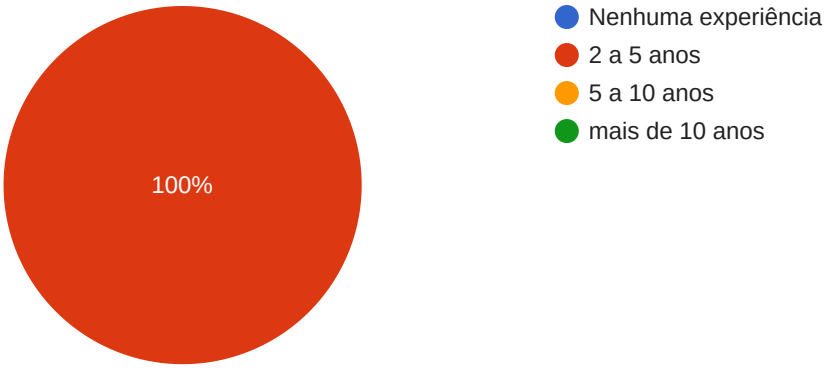


Conhecimento técnico

Quantos anos de experiência como desenvolvedor?

Copiar

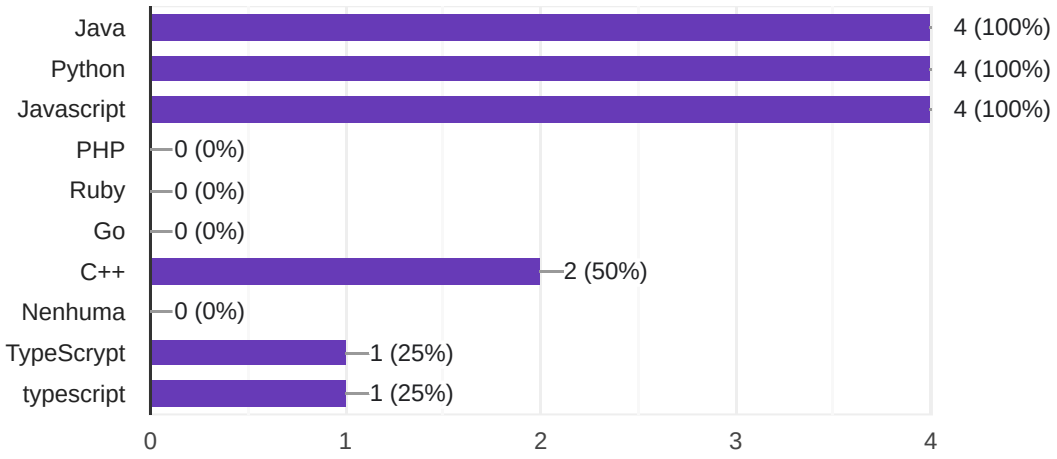
4 respostas



Com quais linguagens de programção você já trabalhou ou trabalha?

Copiar

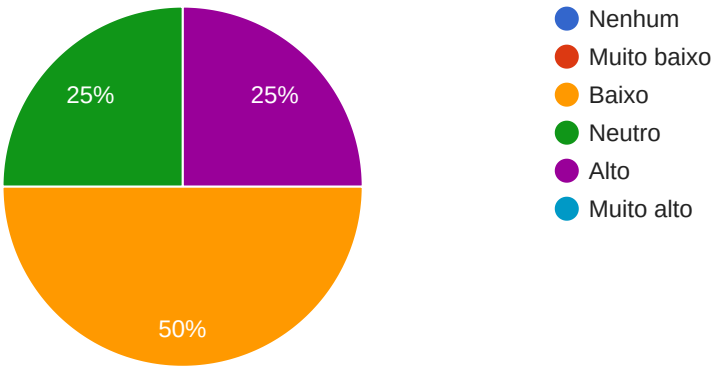
4 respostas



Como você classifica seu conhecimento em relação ao desenvolvimento de chatbots?

Copiar

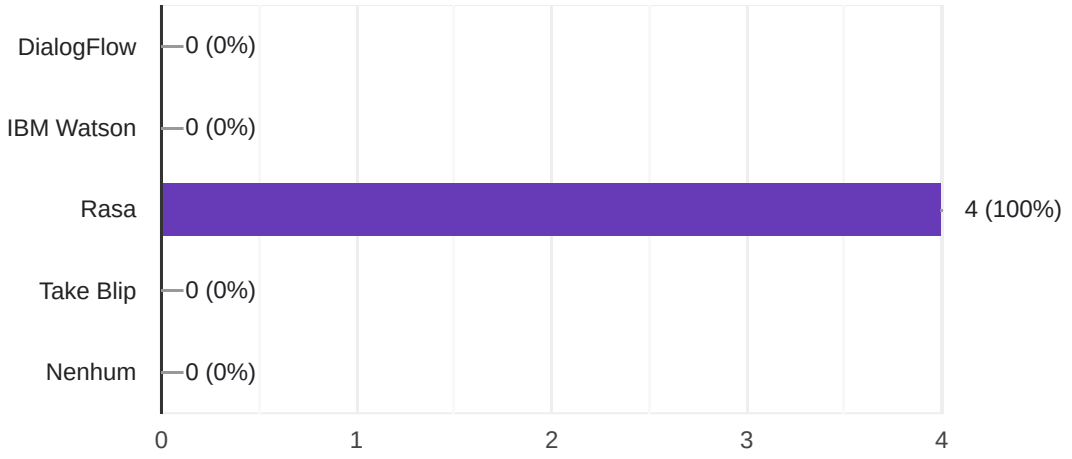
4 respostas



Com quais frameworks de chatbot você já trabalhou?

Copiar

4 respostas

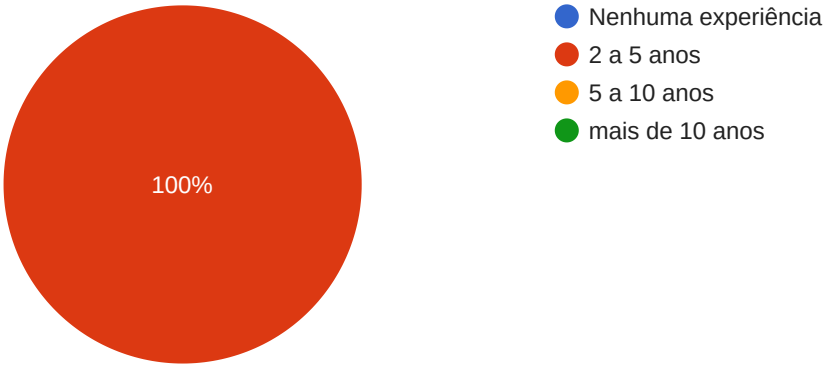


Conhecimento de Gamificação

Quantos anos de experiência com projetos de gamificação?

Copiar

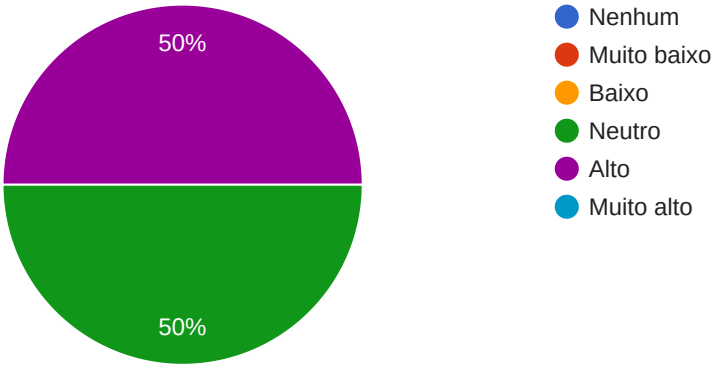
4 respostas



Como você classifica seu conhecimento em relação a gamificação?

Copiar

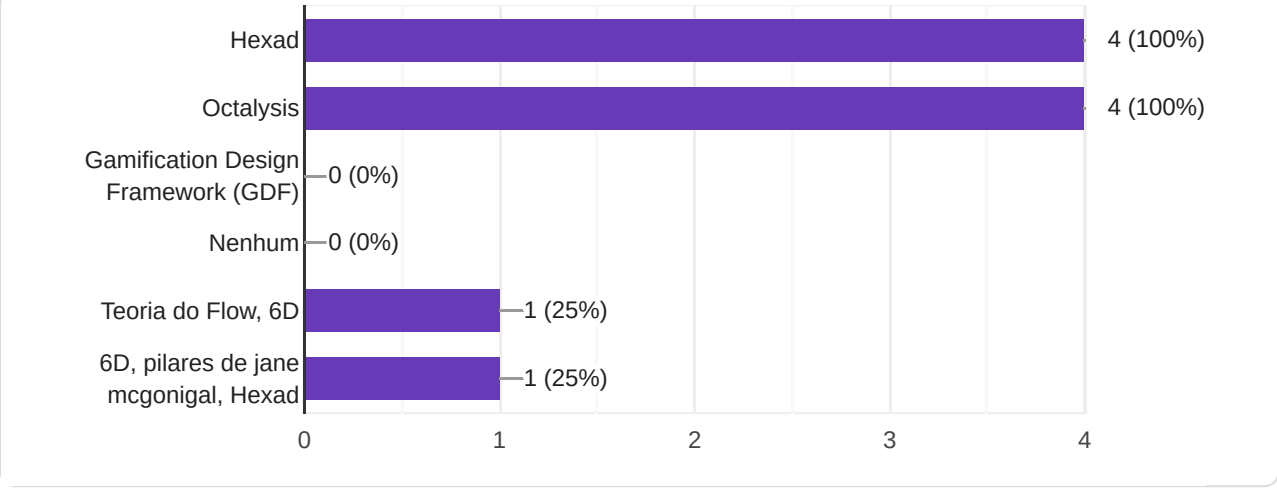
4 respostas



Com quais frameworks de gamificação você já trabalhou?

Copiar

4 respostas



IV.2 Users gamification profile

IV.2.1 Questionary

Latte Chatbot: Uma ferramenta para te auxiliar na escrita de Artigos Científicos

O Latte Chatbot é uma iniciativa de estudantes do curso de Engenharia de Software da Universidade de Brasília para auxiliar pesquisadores na escrita de artigos científicos. Como parte do processo de desenvolvimentos estamos construindo nossa base de conhecimento de maneira a abranger as principais dores da comunidade.

Agradecemos as sua participação!

* Obrigatória

Informações gerais

Nesta seção colocaremos algumas informações gerais para entender o público contemplado na pesquisa.

1. Qual a sua idade? *

- ◯ Menos de 18 anos
- ◯ De 18 a 21 anos
- ◯ De 22 a 30 anos
- ◯ De 31 a 44 anos
- ◯ De 45 a 60 anos
- ◯ Mais de 60 anos

2. Qual seu nível de formação acadêmica? *

- ◯ Ensino Médio Incompleto
- ◯ Ensino Médio Completo
- ◯ Graduação Incompleta
- ◯ Graduação Completa
- ◯ Pós-Graduação Incompleta
- ◯ Pós-Graduação Completa
- ◯ Mestrado incompleto
- ◯ Mestrado completo
- ◯ Doutorado incompleto
- ◯ Doutorado completo
- ◯ Outra

3. Qual sua principal área de pesquisa/estudo? *

Solucionando suas dores

Nesta seção coletaremos as principais dores do nosso publico em relação a escrita de artigos científicos, sua ajuda é de suma importância para o desenvolvimento da ferramenta!

4. Em sua experiência na escrita de trabalhos acadêmicos, em em quais temas se concentram suas duvidas mais frequentes? *

- [ ] Estruturação: como organizar as seções do artigo, como introdução, métodos, resultados e discussão.
- [ ] Redação: como escrever de forma clara e objetiva, sem ambiguidades ou erros gramaticais.
- [ ] Referências bibliográficas: como citar corretamente as fontes utilizadas no trabalho, seguindo as normas da instituição ou periódico onde o artigo será submetido.
- [ ] Seleção de fontes: como escolher quais referências serão utilizadas para embasar o trabalho, de forma a garantir que sejam relevantes e atualizadas.
- [ ] Análise de dados: como apresentar os resultados de forma clara e concisa, utilizando tabelas, gráficos e estatísticas adequados.
- [ ] Estilo de escrita: como adequar a linguagem e o estilo de escrita ao público-alvo e ao tipo de periódico ou evento onde o artigo será submetido.
- [ ] Formatação: como seguir as normas de formatação do periódico ou evento onde o artigo será submetido, incluindo margens, espaçamento, tamanho da fonte e outras especificações.
- [ ] Revisão: como revisar o texto para garantir que não haja erros de ortografia, gramática ou digitação, bem como para verificar se o conteúdo está claro e coeso.
- [ ] Plágio: como evitar a prática de plágio, citando corretamente as fontes e evitando o uso indevido de trechos de textos de outros autores.
- [ ] Originalidade: como garantir que o trabalho seja original e inovador, trazendo contribuições relevantes para a área de estudo.
- [ ] Normas técnicas: utilização adequada das normas técnicas como APA, IEEE, MLA e outros
- [ ] Outra

5. Qual é o seu processo atual para escrever um artigo científico? Nos conte um pouco mais sobre. *

6. Você tem dificuldades em organizar suas ideias e pensamentos antes de começar a escrever? Nos conte um pouco mais sobre. (Caso não, responda: não) *

7. Você possui alguma dificuldade na escrita de artigos científicos que não foi comtemplada na presente pesquisa? Nos conte um pouco mais sobre.

8. Você já utilizou um Chatbot? (Ex: ChatGPT, Siri, Alexa) *

- ◯ Sim
- ◯ Não

Um chatbot Gamificado para escrita de artigos

Com objetivo de melhorar a experiência de nossos usuários estaremos agregando elementos gamificação na aplicação, gostaríamos de entender um pouco do que te motiva!

9. Avalie as afirmações considerando os aspectos que te motivam: *

	Concordo Totalmente	Concordo	Não estou decidido	Desconcordo	Desconcordo Totalmente
Eu me sentiria motivado se o chatbot me oferecesse oportunidades para desenvolver minhas habilidades e conhecimentos.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot me fornecesse feedbacks quanto as minhas ações.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot me oferecesse missões ou objetivos desafiadores que me fizessem descobrir novas informações ou segredos no chatbot.	◯	◯	◯	◯	◯
Eu me sentiria motivado se o chatbot me fizesse sair da minha zona de conforto e experimentar coisas novas.	◯	◯	◯	◯	◯
Eu me					

sentiria motivado(a) se este chatbot gamificado me permitisse tomar decisões e ter controle sobre a situação.	○	○	○	○	○
Eu me sentiria motivado se o chatbot me ajudasse a enxergar o meu desenvolvimento.	○	○	○	○	○
Eu me sentiria motivado(a) se o chatbot oferecesse recompensas exclusivas e raras, que eu possa colecionar.	○	○	○	○	○
Eu me sentiria motivado(a) se o chatbot oferecesse oportunidades para eu me conectar com outros usuários.	○	○	○	○	○

10. Avalie as afirmações considerando os aspectos que te motivam: *

	Concordo Totalmente	Concordo	Não estou decidido	Desconcordo	Desconcordo Totalmente
Eu me sentiria motivado(a) se o chatbot gamificado me alertasse sobre as possíveis perdas de não completar tarefas ou objetivos propostos.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot trabalhasse com lançamentos de novidades.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot me oferecesse desafios e recompensas temporárias e limitadas.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot me desse a opção de poder entrar em grupos exclusivos.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot apresentasse surpresas e	◯	◯	◯	◯	◯

eventos aleatórios durante o jogo.					
Eu me sentiria mais motivado(a) se o chatbot me mostrasse que estou contribuindo para um propósito maior.	◯	◯	◯	◯	◯
Eu me sentiria motivado(a) se o chatbot me permitisse compartilhar meu progresso e resultados com outros usuários.	◯	◯	◯	◯	◯

Muito obrigado pela sua contribuição!

Suas respostas serão de sua importância para o projeto Latte Chatbot! 🤖

11. Você acredita que um chatbot para escrita de artigos científicos poderia ajudá-lo a superar suas dificuldades e otimizar seu tempo? (Considere chatbot como uma ferramenta capaz de tirar duvidas e realizar consultas) *

- [ ] Sim
- [ ] Não

12. Deseja receber em primeira mão a plataforma quando for lançada? Deixe seu melhor e-mail

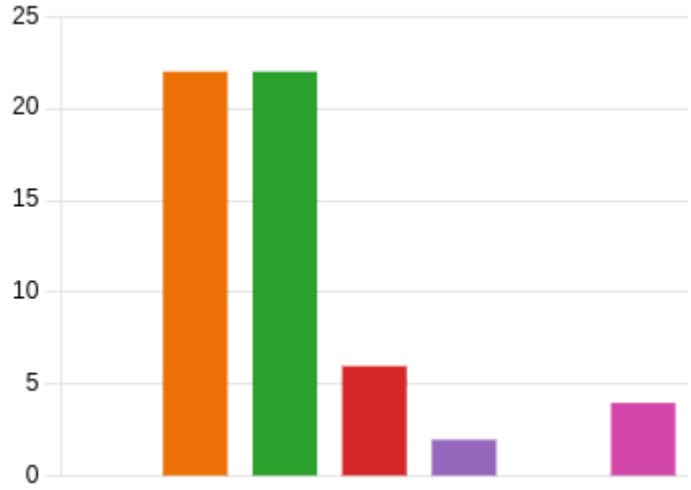
IV.2.2 Questionaries responses

Latte Chatbot: Uma ferramenta para te auxiliar na escrita de Artigos Científicos

56	07:22	Fechado
Respostas	Tempo médio para concluir	Status

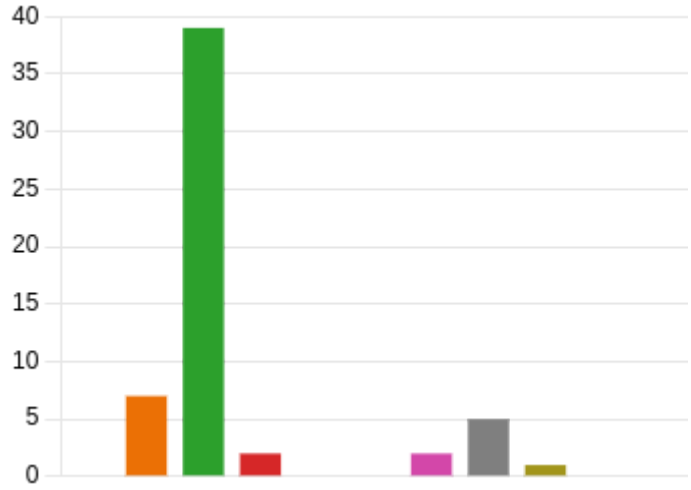
1. Qual a sua idade?

Menos de 18 anos	0
De 18 a 21 anos	22
De 22 a 30 anos	22
De 31 a 44 anos	6
De 45 a 60 anos	2
Mais de 60 anos	0
Outra	4



2. Qual seu nível de formação acadêmica?

Ensino Médio Incompleto	0
Ensino Médio Completo	7
Graduação Incompleta	39
Graduação Completa	2
Pós-Graduação Incompleta	0
Pós-Graduação Completa	0
Mestrado incompleto	2
Mestrado completo	5
Doutorado incompleto	1
Doutorado completo	0
Outra	0



3. Qual sua principal área de pesquisa/estudo?

56
Respostas

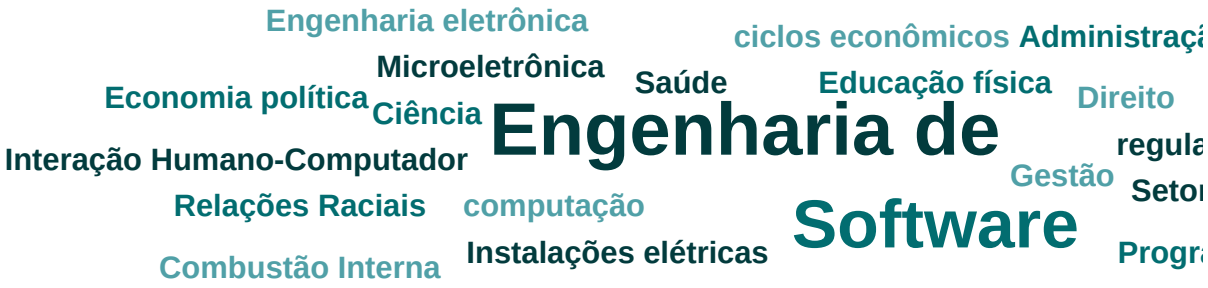
Respostas Mais Recentes

"Engenharia de software "

"Software"

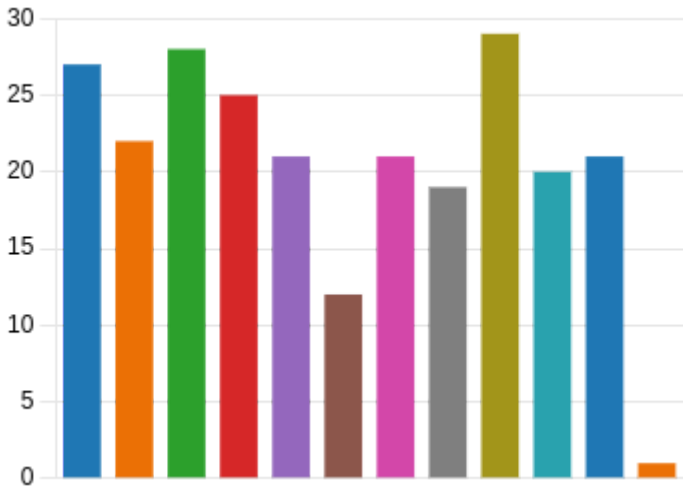
"Engenharia de software"

16 respondentes (**29**%) responderam **Engenharia de** para esta pergunta.



4. Em sua experiência na escrita de trabalhos acadêmicos, em em quais temas se concentram suas duvidas mais frequentes?

Estruturação: como organizar as...	27
Redação: como escrever de for...	22
Referências bibliográficas: como...	28
Seleção de fontes: como escolh...	25
Análise de dados: como apresen...	21
Estilo de escrita: como adequar ...	12
Formatação: como seguir as nor...	21
Revisão: como revisar o texto pa...	19
Plágio: como evitar a prática de ...	29
Originalidade: como garantir qu...	20
Normas técnicas: utilização ade...	21
Outra	1



5. Qual é o seu processo atual para escrever um artigo científico? Nos conte um pouco mais sobre.

56
Respostas

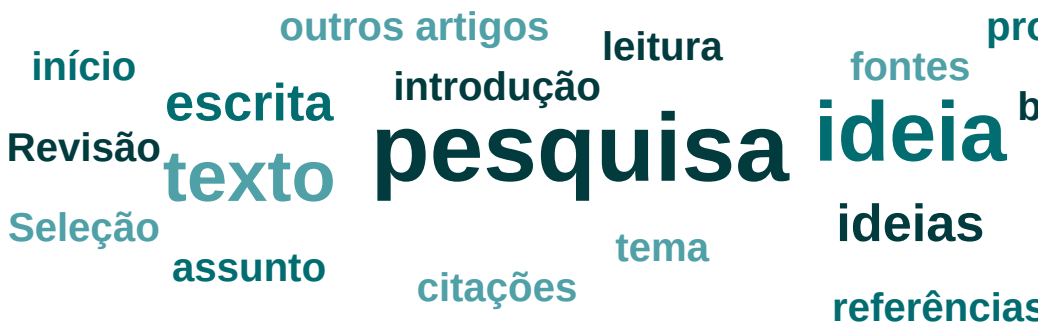
Respostas Mais Recentes

"Não estou escrevendo artigos científicos atualmente."

"1. Especificação do tema escolhido 2. Pesquisa para levanta...

"ideia > pesquisa de outros artigos > escrita > revisão > vol...

9 respondentes (**16**%) responderam **pesquisa** para esta pergunta.



6. Você tem dificuldades em organizar suas ideias e pensamentos antes de começar a escrever? Nos conte um pouco mais sobre. (Caso não, responda: não)

56
Respostas

Respostas Mais Recentes

"Minha maior dificuldade estava em ligar uma ideia a outra ...

"Não'

"nao'

7 respondentes (**13**%) responderam **pouco** para esta pergunta.

7. Você possui alguma dificuldade na escrita de artigos científicos que não foi comtemplada na presente pesquisa? Nos conte um pouco mais sobre.

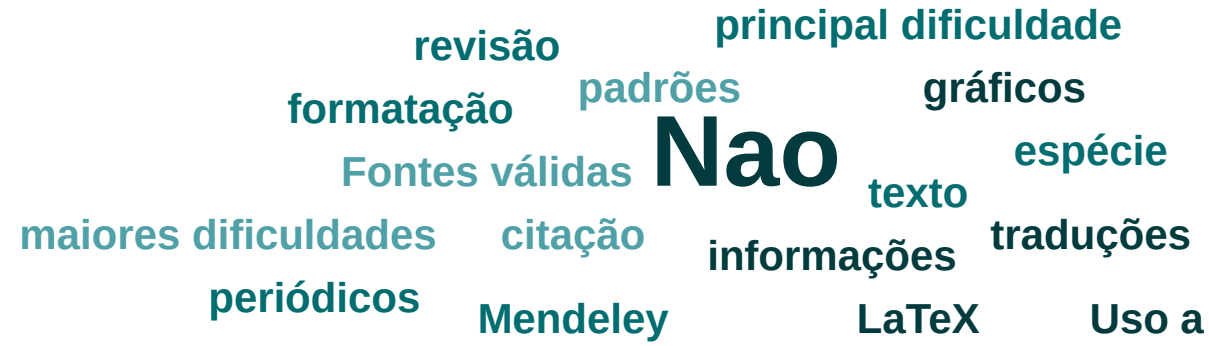
36
Respostas

Respostas Mais Recentes

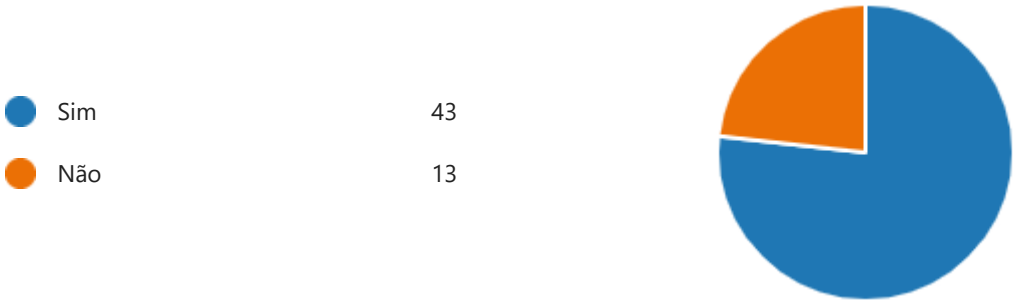
"Não.'

"Não'

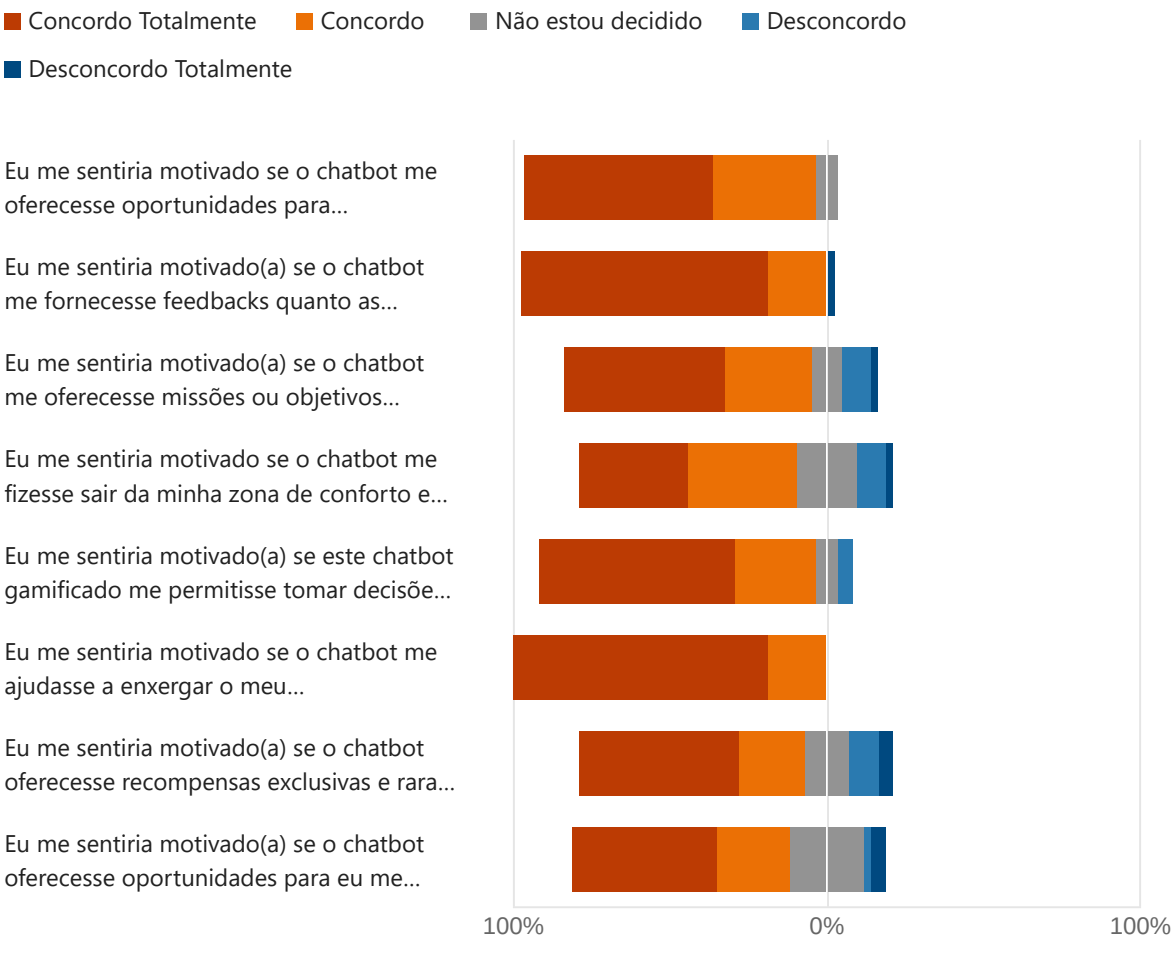
5 respondentes (**14**%) responderam **Nao** para esta pergunta.



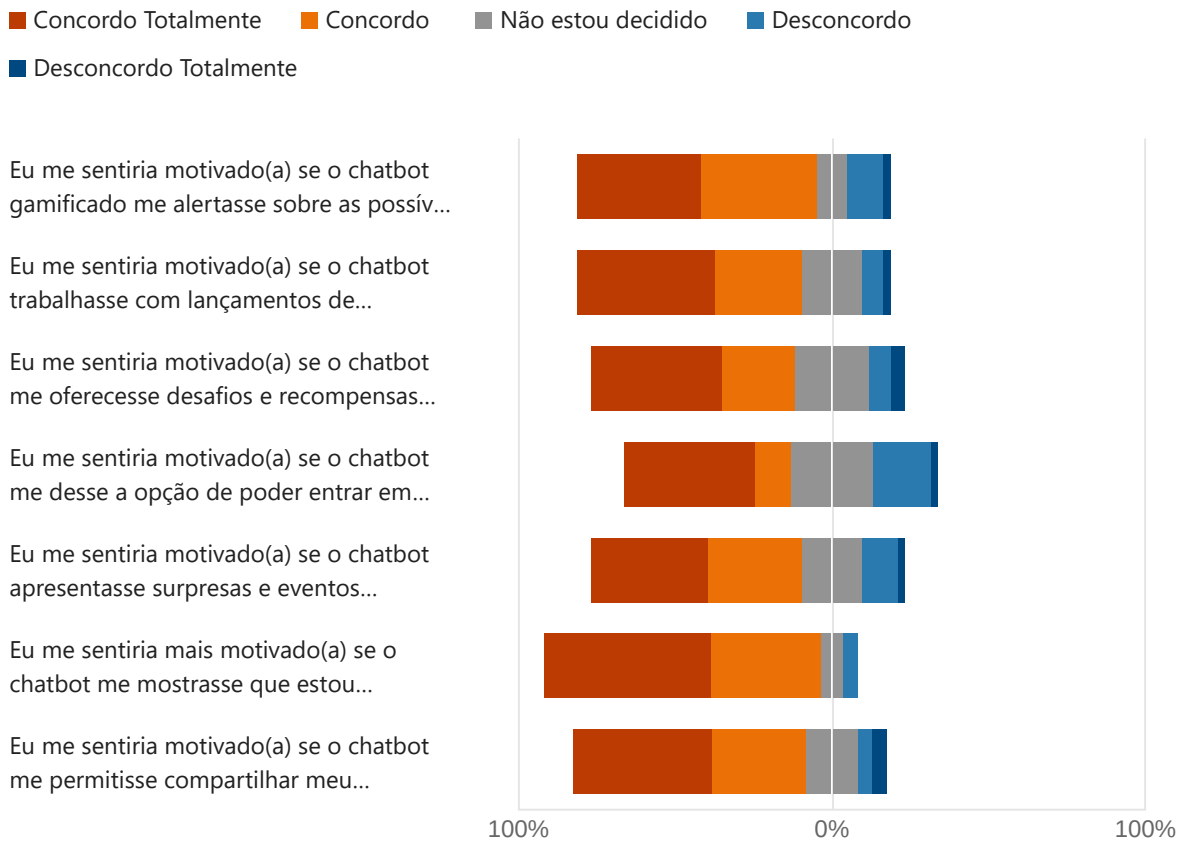
8. Você já utilizou um Chatbot? (Ex: ChatGPT, Siri, Alexa)



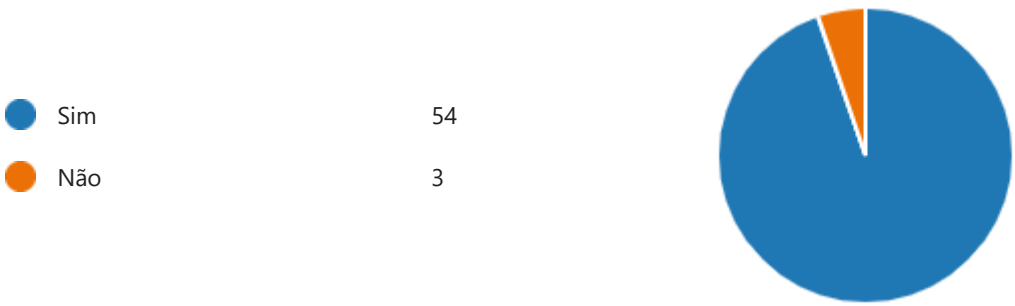
9. Avalie as afirmações considerando os aspectos que te motivam:



10. Avalie as afirmações considerando os aspectos que te motivam:



11. Você acredita que um chatbot para escrita de artigos científicos poderia ajudá-lo a superar suas dificuldades e otimizar seu tempo? (Considere chatbot como uma ferramenta capaz de tirar duvidas e realizar consultas)



IV.3 Chatbot users usage

IV.3.1 First phase

Teste 1.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	1 minuto 3 segundos	3	4
2	Sim	1 minuto 15 segundos	3	4
3	Sim	58 segundos	3	4
4	Sim	1 minuto 27 segundos	4	5
5	Sim	1 minuto 24 segundos	3	3
Media	**100% Sim**	**1 minuto 14 segundos**	**3 Mensaguens**	**3.8**

Teste 2.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	2 minutos e 15 segundos	3	3
2	Sim	2 minutos 10 segundos	4	3
3	Sim	2 minutos 7 segundos	3	4
4	Sim	2 minutos 20 segundos	3	3
5	Sim	2 minutos 42 segundos	4	4
Media	**100% Sim**	**2 minutos 19 segundos**	**3 Mensaguens**	**3.4**

Teste 3.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	1 minuto 43 segundos	4	5
2	Sim	1 minuto 32 segundos	4	4
3	Sim	1 minuto 38 segundos	4	5
4	Sim	1 minuto 49 segundos	4	4
5	Sim	1 minuto 51 segundos	4	4
Media	**100% Sim**	**1 minutos 43 segundos**	**4 Mensaguens**	**4.4**

IV.3.2 Second phase

Teste 1.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	54 segundos	3	4
2	Sim	1 minuto e 11 segundos	3	4
3	Sim	1 minuto e 2 segundos	3	5
4	Sim	58 segundos	3	5
5	Sim	1 minuto e 5 segundos	3	4
Media	**100% Sim**	**1 minuto 2 segundos**	**3**	**4.4**
6	Sim	1 minuto e 17 segundos	3	4
7	Sim	1 minuto e 3 segundos	3	4
8	Sim	1 minuto e 4 segundos	4	3
9	Sim	51 segundos	2	5
10	Sim	1 minuto e 14 segundos	4	5
Media	**100% Sim**	**1 minuto e 6 segundos**	**3**	**4.2**
M. Final	**100% Sim**	**1 minuto 4 segundos**	**3**	**4.3**

Teste 2.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	2 minutos 10 segundos	3	4
2	Sim	1 minutos 57 segundos	3	5
3	Sim	2 minutos 22 segundos	3	4
4	Sim	2 minutos 7 segundos	3	3
5	Sim	2 minutos 11 segundos	3	4
Media	**100% Sim**	**2 minutos e 9 segundos**	**3**	**4.0**
6	Sim	2 minutos 8 segundos	3	3
7	Sim	2 minutos 14 segundo	4	4
8	Sim	2 minuto 3 segundos	3	3
9	Sim	2 minutos 34 segundos	5	3
10	Sim	2 minutos 12 segundos	4	5
Media	**100% Sim**	**2 minutos e 14 segundos**	**4**	**3.6**

M. Final	**100% Sim**	**2 minutos e 12 segundos**	**3**	**3.8**

Teste 3.1				
ID	**Conseguiu concluir o objetivo?**	**Tempo (min/seg)**	**Quantidade de Mensagens**	**Avaliação do Usuário (1 a 5)**
1	Sim	1 minuto 31 segundos	4	5
2	Sim	1 minuto 18 segundos	4	5
3	Sim	1 minuto 23 segundos	4	4
4	Sim	1 minuto 28 segundos	4	5
5	Sim	1 minuto 45 segundos	4	4
Media	**100% Sim**	**1 minuto 29 segundos**	**4**	**4.6**
6	Sim	1 min 21 segundos	4	4
7	Sim	1 min 40 segundos	4	4
8	Sim	1 minuto 45 segundos	5	4
9	Sim	1 minuto 54 segundos	4	5
10	Sim	1 minuto e 41 segundos	4	4
Media	**100% Sim**	**1 minuto 41 segundos**	**4**	**4.2**
M. Final	**100% Sim**	**1 munuto e 35 segundos**	**4**	**4.4**