informatics

MDPI

Article

Toolkit for Inclusion of User Experience Design Guidelines in the Development of Assistants Based on Generative Artificial Intelligence

Carlos Alberto Peláez [1], Andrés Solano [1], Johann A. Ospina [1], Juan C. Espinosa [1], Ana S. Montaño [1], Paola A. Castillo [1], Juan Sebastián Duque [2], David A. Castro [2], Juan M. Nuñez Velasco [3,*] and Fernando De la Prieta [3]

[1] Grupo de Investigación GITI, Universidad Autónoma de Occidente, Cali 760030, Colombia; capa@uao.edu.co (C.A.P.); afsolano@uao.edu.co (A.S.); jaospina@uao.edu.co (J.A.O.); juan_cam.espinosa@uao.edu.co (J.C.E.); ana_sofia.montano@uao.edu.co (A.S.M.); pacastillo@uao.edu.co (P.A.C.)
[2] Corporación Talentum, Cali 760042, Colombia; juan.duque@talentum.edu.co (J.S.D.); david.castro@talentum.edu.co (D.A.C.)
[3] Grupo de Investigación BISITE, Universidad de Salamanca, 37008 Salamanca, Spain; fer@usal.es
* Correspondence: jmnunez@usal.es

Abstract: This study addresses the need to integrate ethical, human-centered principles into user experience (UX) design for generative AI (GenAI)-based assistants. Acknowledging the ethical and societal challenges posed by the democratization of GenAI, this study developed a set of six UX design guidelines and 37 recommendations to guide development teams in creating GenAI assistants. A card-based toolkit was designed to encapsulate these guidelines, applying color theory and Gestalt principles to enhance usability and understanding. The design science research methodology (DSRM) was followed, and the toolkit was validated through a hands-on workshop with software and UX professionals, assessing usability, user experience, and utility. The quantitative results indicated the high internal consistency and effectiveness of the toolkit, while the qualitative analysis highlighted its capacity to foster collaboration and address GenAI-specific challenges. This study concludes that the toolkit improves usability and utility in UX design for GenAI-based assistants, though it identifies areas for future enhancement and the need for further validation across varied contexts.

Keywords: user experience; generative artificial intelligence; human-centered artificial intelligence

check for updates

Academic Editors: Da Tao, Tingru Zhang and Hailiang Wang

Received: 19 November 2024
Revised: 29 December 2024
Accepted: 7 January 2025
Published: 24 January 2025

Citation: Peláez, C.A.; Solano, A.; Ospina, J.A.; Espinosa, J.C.; Montaño, A.S.; Castillo, P.A.; Duque, J.S.; Castro, D.A.; Nuñez Velasco, J.M.; De la Prieta, F. Toolkit for Inclusion of User Experience Design Guidelines in the Development of Assistants Based on Generative Artificial Intelligence. *Informatics* **2025**, *12*, 10. https://doi.org/10.3390/informatics12010010

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A new era has begun for the discipline of artificial intelligence (AI), driven by its democratization. Sectors such as healthcare [1] and education [2] serve as examples of non-IT niches that are beginning to experience profound transformations because of the use and appropriation of AI. The evolution of generative artificial intelligence (GenAI) has proven to be the key enabler of democratization [3]. Neural networks based on transformer architectures have undoubtedly expanded the range of possibilities that these technologies offer [4].

Society must inevitably confront new challenges and risks; like a child who grows up and leaves home, AI has embarked on a journey beyond its origins in computer science. Responsibility for the use, appropriation, and potential future evolution of AI is now

rapidly extending to other knowledge areas and disciplines. These fields are tasked with the challenge of educating society and building a culture capable of making informed, ethical, and moral decisions regarding the use of this technology. In this context, human-centered artificial intelligence (HCAI) is poised to play a key role [5].

Previous experiences have provided us with prior knowledge that enables us to better face this new challenge. For example, we are aware of the role played by the Human–Computer Interaction (HCI) discipline in bridging the gap between non-experts and technology. HCI has facilitated an interaction with technology during the era of computational democratization, which is marked by the emergence of personal computers and the internet [5].

The transition from HCI to HCAI can be understood as a process of conceptual expansion and disciplinary scope. While HCI emerged to ensure that digital technologies were accessible, comprehensible, and useful to people, HCAI takes these fundamental principles further by incorporating the unique capabilities and characteristics of AI [6]. Thus, HCAI inherits from HCI its focus on usability, user satisfaction, and the reduction in cognitive and emotional barriers in the interaction while introducing new challenges, such as the need to design for the transparency of complex models, ethical responsibility in automated decision-making, and fostering trust in autonomous systems [7].

In essence, HCAI not only seeks to make interfaces understandable to users but also to ensure that the algorithms and underlying logics of AI align with human values, moral principles, and sociocultural contexts [8]. In this way, HCAI expands the scope of HCI, taking on not only the design of surface-level user experiences but also the responsible configuration of the technological core that underpins them.

The above presents a new scenario with emerging challenges in light of the rise in solutions based on GenAI-mediated assistants that aims to mitigate the implicit risks posed by the widespread use of this technology [9]. The current literature largely focuses on studies that highlight the potential benefits offered by GenAI-based solutions for user experience (UX) design in interactive systems [10,11]. However, there has been little discussion regarding what UX design should consider in the development of GenAI-based assistant solutions [12]. This research specifically aims to contribute to this discussion by providing a toolkit that compiles a set of guidelines and recommendations for the UX design of GenAI-based assistants. These guidelines are aimed at designers and developers to increase the perception of trust, usefulness, and UX in human–AI interactions. To this end, this paper first reviews the related works, noting their purpose, contributions, and limitations. Then, it provides a description of the methodology used, detailing its implementation throughout the course of this study. The last sections present the results obtained from the experiments conducted with the stakeholders, as well as a final discussion of the findings.

2. Background

HCAI is a field that strongly emphasizes human experiences, satisfaction, and needs. Its goal is to improve, augment, and optimize human performance while ensuring that AI systems are dependable, secure, and trustworthy. This approach supports human self-efficacy, encourages creativity, defines responsibilities more clearly, and fosters greater social engagement [13]. These aspects are considered to be empirically significant relative to their influence on the UX of GenAI-based systems. However, some researchers have underscored the need for a further investigation into the UX of AI-assisted chatbots [14]. Concerns have also been raised by various authors about GenAI-based tools like ChatGPT, particularly regarding their tendency to generate inaccurate or fabricated information and

their ability to circumvent systems designed to detect duplicate content, which is a critical issue in domains where originality is paramount [15].

A comprehensive review of 223 documents sourced from various electronic databases revealed the key factors influencing UX design for the development of GenAI-based systems. This state-of-the-art analysis highlighted the substantial impact of the principles of the HCAI discipline [12].

Multiple studies have highlighted how GenAI can reshape a variety of domains, such as healthcare, education, finance, tourism, and cultural heritage, by enhancing personalization and improving overall UX [16–19]. Central to this evolution is the shift from HCI principles to HCAI approaches [20], emphasizing ethics, inclusivity, fairness, and cultural sensitivity in GenAI-driven solutions [21].

Human feedback is a crucial factor in maintaining the quality and relevance of GenAI outputs [16,22] and ensuring trust, reliability, transparency, and explainability in interactive systems [23]. Researchers underscore that effective GenAI solutions must align with user needs, expectations, and values while incorporating the psychological, social, and cognitive dimensions of interactions [24].

Studies on GenAI's multimodal capabilities have indicated progress in enhancing UX, especially when bridging the gap between algorithmic complexity and user-friendly interfaces [25]. This includes enabling more natural communication through conversational systems, improving personalization, and ensuring that designs remain accessible to diverse user groups [26].

The collective insights from the reviewed works underscore the need to incorporate a human-centered, ethically grounded UX design that integrates transparency, fairness, accountability, and explainability into GenAI solutions [27–30]. This approach entails inclusive design, user education, and continuous feedback loops, ensuring that GenAI solutions are not only technologically advanced but also respectful of human values, social contexts, and long-term well-being [31,32]. Studies have suggested that user trust in GenAI-based technologies is largely dependent on how these ethical aspects are addressed, particularly in sensitive sectors such as healthcare and education.

Likewise, inclusivity and accessibility are key priorities. Studies have advocated for designing GenAI solutions that are accessible to a wide range of users, regardless of their abilities or socioeconomic context [21,33]. This trend represents an effort to democratize access to advanced technologies and ensure that the benefits of GenAI are available to everyone.

A final discussion, resulting from the review of these studies, highlights the concerns raised in the literature regarding how to address sensitive factors, such as ethics, transparency, inclusivity, accessibility, and fairness, among others, in GenAI-mediated solutions. However, while it is common for these studies to emphasize HCAI as an essential discipline for providing such a foundation, no studies explicitly offer precise contributions to incorporate these considerations, for instance, at the design process level for such solutions.

Based on the above elements, this study makes a contribution through a toolkit that defines a set of UX design guidelines for GenAI-based assistants. This toolkit serves as a resource for generating spaces for analysis, discussion, and conceptualization among members of a multidisciplinary team during the design process of GenAI-based assistant solutions.

3. Materials and Methods

3.1. Methodology

The initial challenge posed by the need to address the scope and implications of UX design in the development of GenAI-based assistants led to the creation of an arti-

fact as a proposed solution. In this case, the solution comprises a set of guidelines and recommendations for the UX design of GenAI-based assistants. Additionally, a toolkit was designed to facilitate the application of these guidelines and recommendations and improve the experience of teams developing a GenAI assistant. In addressing the challenge of artifact development, the design science research methodology (DSRM) [34] offered the best approach.

Following the processing guided by the DSRM, an adaptation was made based on the framework for UX design and the assessments of Adikari et al. [35]. To address the challenges faced by teams developing software solutions, we initiated a research process focused on reviewing existing, validated experiences and guidelines from the literature on UX design, particularly in the context of designing GenAI-based assistant solutions [12].

The design process for the UX guidelines involved a team of expert UX researchers and experienced industry designers. This process emphasized the need to develop a set of recommendations for integrating UX design principles, particularly when creating GenAI-based software solutions. The first part of the methodology applied consisted of two phases: an expert panel phase followed by a descriptive and inferential analysis of the conducted survey. Each guideline was characterized using the Mann–Whitney test [36] to evaluate potential significant differences between experts' perceptions regarding utility and clarity for each UX design guideline.

The knowledge base of UX design guidelines, which was reviewed and discussed by experts, helped structure the content for developing the toolkit. A group of designers participated in discussions about the most effective colors, illustrations, and formats for presenting the guidelines and their recommendations, both digitally and on printed cards.

The definition of the elements that comprise the cards in the toolkit is based on the IDEO method card set [37] and the Value-Sensitive Design Toolkit [38]. Both references provided a foundation for structuring the information in an accessible, practical, and easy-to-implement way. Additionally, the specific context of GenAI requires consideration of the unique challenges in this field, such as the need to explain complex AI concepts to non-technical users and the importance of managing expectations regarding system capabilities.

The printed card set was used to test the toolkit in the workplace of a UX design team at a software development company. The goal was to assess the toolkit's contribution in terms of user experience, usability, and utility in identifying factors associated with UX design for GenAI-based assistants. The process is outlined in Figure 1.

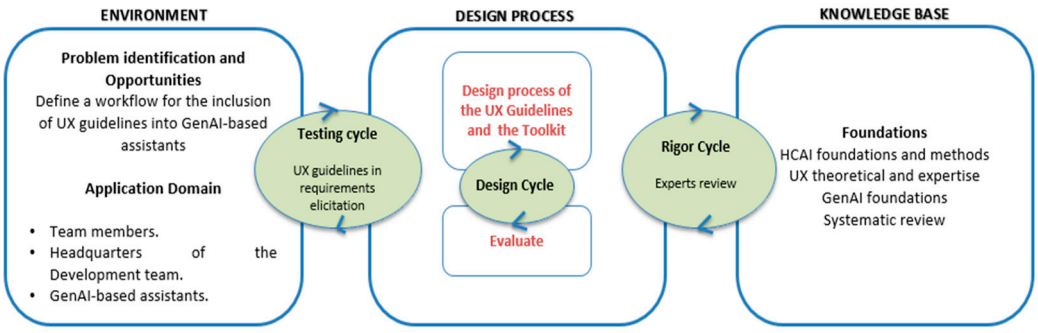


Figure 1. The DSRM process guided the development of the UX design guidelines toolkit for GenAI-based assistants. Adapted from [35].

3.2. UX Design Guidelines and Recommendations for GenAI-Based Assistants

The general findings identified in the literature led to the recognition of several elements where emphasis was placed on various factors with a potential influence on how people perceive a good UX as a result of using and adopting a GenAI-based assistant. These factors encompass ethical and responsible design considerations, inclusivity for appropriate

AI usage, fostering effective human–AI collaboration, personalization of the human–AI experience, trust, and reliability in the human–AI relationship, and addressing issues of inaccuracy and variability.

A key factor in the discussion of these elements is the relationship between innovation and necessity. Undoubtedly, topics such as the discussion of ethical and moral factors in the development of interactive systems [39], as well as accessibility as an essential element for inclusive solutions [40], are widely addressed in the literature of disciplines such as HCI. These critical considerations of morality, ethics, and inclusion should not be overlooked; on the contrary, they highlight the urgent need for reassessment considering the unique characteristics of GenAI technologies and the challenges they pose. Consistent with the present analysis, this study identified six UX design guidelines that should be considered in the development of a GenAI-based assistant.

A set of 37 recommendations was identified to be associated with the UX design guidelines for the development of GenAI-based assistants. These recommendations allow for a more detailed and precise interpretation of the scope of each UX design guideline in the indicated context. In addition, they provide a reference for the foundation of an analysis and discussion process among the team members responsible for designing a GenAI-based assistant solution according to the context and stakeholder needs.

Table 1 presents the six UX design guidelines for the development of GenAI-based assistants, each accompanied by a brief description. A specific nomenclature is used to facilitate the inclusion of those guidelines in both the description of recommendations and the toolkit that is introduced further in this paper.

Table 1. UX design guidelines for the development of GenAI-based assistants.

Code	UX Guideline	Description
G1	Design Based on Ethical and Responsible Principles	This guideline should be considered a cross-cutting factor in the development of any solution based on GenAI, regardless of its application context [41]. Incorporating a GenAI-based solution or evolving an existing one should prompt us to question how that solution or its evolution will contribute to creating value for the UX [42]. A design grounded in ethical and responsible principles contributes to creating value in UX by enhancing the user's perception of transparency in GenAI-based assistants. It enables mechanisms that clearly communicate the solution's capabilities, limitations, and potential risks, prevents misunderstandings, and includes bias mitigation as a key element that requires traceability and control throughout the human–AI interaction process [43].
G2	Design Based on Principles of Inclusion	A design based on the principles of inclusion ensures that GenAI-based assistants can be used and understood by a diverse and representative audience. This is achieved by creating solutions adaptable to various accessibility needs, addressing several disabilities within the population. Furthermore, it emphasizes the importance of considering cultural and linguistic variations to ensure their utility for a global and diverse audience and mitigating potential biases and prejudices [44]. This approach promotes the reduction in barriers to the adoption of GenAI in disadvantaged contexts, particularly in emerging economies, recognizing the need to advance toward models whose development is not constrained by data derived predominantly from customs, cultures, and contexts belonging to industrialized countries [21].

Table 1. *Cont.*

Code	UX Guideline	Description
G3	Design that Promotes AI–Human Collaboration	This guideline responds to the significant variety of human tasks that can be supported by GenAI-based solutions at various stages, tailored to specific needs throughout their execution [45]. Grounded in the principle of HCAI, this guideline adheres to the premise that AI solutions should exist to augment capabilities, not replace them. In this context, various types of human activities can be greatly enhanced through the mediation of GenAI-based solutions in their processes.
G4	Design Oriented Toward Personalization of the AI–Human Experience	GenAI-based solutions offer the potential to usher in a new era of information technologies, with the capacity to scale toward a more natural and personalized human–computer experience. A key factor enabling this evolution is the anthropomorphism resulting from human–GenAI interaction [46]. It is crucial, however, to ensure that users remain aware that they are interacting with technology, not another human (a conscious being). This guideline highlights the challenge of value tension [38], which must be considered in UX design for such solutions.
G5	Design Based on Trust and Reliability in GenAI–Human Interaction	In the context of emerging technologies, such as GenAI-based solutions, trust and reliability are critical factors for their adoption and evolution. Trust arises when a system delivers useful results, a high quality, and is verifiable [29]. It is essential for users to identify issues such as hallucinations, biases, and errors and evaluate whether these problems are manageable or acceptable. This capability allows users to adjust their expectations, maintain control, and ensure informed use of the system. In contrast, reliability refers to the system's consistency and predictability under different conditions. A reliable system not only reduces errors but also strengthens user perception by demonstrating safe and consistent performance. Both concepts—trust and reliability—are essential pillars for ensuring effective and sustainable interaction in GenAI-based solutions. In this regard, the adoption of GenAI follows a cycle similar to Gartner's Hype Cycle for emerging technologies. Currently, many GenAI solutions are at a peak of inflated expectations [47], where the initial enthusiasm of users contrasts with their difficulties in recognizing the limitations and potential risks of these technologies. However, as these solutions progress toward maturity, trust needs to be established in human–AI interactions to determine their long-term success and adoption.
G6	Design with Consideration of Inaccuracy and Variability	The results generated by GenAI-based solutions are far from infallible. This may significantly deviate from the originally estimated or expected results. On the one hand, responses may be subject to inaccuracies due to the inherent limitations of the algorithms, data, and architectures underlying these solutions. On the other hand, users should be aware that GenAI systems can provide responses that vary in quality and form, even if the input provided by the user remains unchanged [45].

Associated with each of the previously discussed UX design guidelines, a set of recommendations is provided to guide the UX design team during the analysis and creation phases of the GenAI-based assistant. These recommendations are intended to streamline the incorporation of the guidelines into the design process and to ensure their effective

implementation in the final specification of the solution. Table 2 provides a detailed overview of the recommendation associated with each guideline.

Table 2. UX design guidelines and their associated recommendations.

UX Guideline	Associated Recommendations
G1	G1R1. Communicate the System's Scope and Limitations: Ensure that stakeholders are aware of the system's scope, limitations, and primary intended uses. Clearly specify the purposes of the GenAI system. The system should be prepared to assist users when they inquire about its capabilities. G1R2. Suggest Limits and Offer Assistance: The system should suggest its own limitations and deficiencies when inaccuracies are detected in the generated results. This can foster trust and reassurance in users. If the system fails, it should provide users with the option to be transferred to human customer support. G1R3. Communicate Possibilities of Emergent Behaviors: If your system allows for emergent behaviors, ensure that users understand that the GenAI-based system may provide feedback on topics for which they have not been explicitly trained. G1R4. Monitor and Mitigate Biases and Inappropriate Content: Test, monitor, and track the results generated from user interactions with your GenAI solution, paying attention to outcomes that may exhibit potential biases, prejudices, or even toxic content or misinformation. If such issues are identified, various mechanisms should be explored to mitigate their impact, both at the system level and through communication with stakeholders. G1R5. Document Limitations and Known Issues of the Model: Ensure that you identify and document the potential limitations and known issues associated with the model used for the development of your GenAI-based solution. Investigate whether there is verifiable information about similar solutions that have exhibited biases, prejudices, or negative stereotypes when using the same model. G1R6. Establish Ethical Boundaries and Detect Bias in Responses: Identify biased response patterns related to racial, gender, ethnic, or political identities. Set clear boundaries for the behavior of the AI model, particularly concerning sensitive content and language. Prioritize ethical guidelines that promote more holistic decision-making, incorporating ethics, morality, and fairness. G1R7. Raise Stakeholder Awareness of the Real Capabilities of AI: Educate stakeholders throughout the lifecycle, guiding the development and eventual launch of the GenAI-based solution, providing a realistic and informed perspective on how the system functions. This will help them avoid misinformation, false expectations, and exaggerations about the capabilities of these types of solutions. G1R8. Clearly Communicate the Scope of the Solution: Clearly articulate the solution's scope at the outset of the interaction to prevent user frustration when the system encounters requests that are beyond its capabilities. Ensure that the organization transparently informs users that they are engaged with a GenAI-based system.
G2	G2R1. Adapt the GenAI Solution to the Context and Communities: Acknowledge communities and their context by understanding their cultural, economic, social, and governance limitations. Consider how to tailor your GenAI-based solution to be useful in this context and strive to mitigate any potential features associated with this model and pre-training that could undermine its primary purpose. G2R2. Assessing the Need for Multimodality to Ensure Accessibility: Although incorporating multimodality into GenAI solutions naturally entails a significantly higher cost, its necessity within a specific context and communities should be evaluated. Determine the potential modalities required to ensure accessibility for most stakeholders. If multimodality is not feasible in the initial versions due to various factors, ensure that the solution's architecture is sufficiently modular, flexible, and economically viable to accommodate future evolutions of GenAI systems. G2R3. Use Open and Adaptable Models: Encourage the use of open models that allow for greater adaptability to a knowledge base aligned with the context and communities targeted by your GenAI solution. G2R4. Communicate Evidence of Biases and Prejudices in GenAI Responses: Upon identifying potential biases and prejudices in your GenAI solution's responses that could affect or harm users, such as those related to gender ideology, sexual preferences, religious beliefs, or ethnic and social discrimination, promptly communicate these issues to stakeholders. If possible, implement corrective actions to evolve the solution to minimize this type of response. G2R5. Promote Inclusion and Consideration of GenAI Developers: Raise awareness among stakeholders to adopt an attitude of inclusion and consideration toward those responsible for developing the GenAI solution. Encourage an understanding of the limitations and risks inherent in the current use and adoption of such technologies. This approach fosters relationships that support the system's evolution through a more inclusive design process that involves all stakeholders. G2R6. Provide Accessible Alternatives to Generated Content: Offer alternatives to generated content, such as text descriptions of images. These alternatives are essential to ensure that all users, regardless of their abilities, can access and understand the information provided by the AI. Adopt recognized accessibility standards, such as the Web Content Accessibility Guidelines (WCAGs). G2R7. Ensure Accessibility Technologies and Navigation Options: Consider assistive technologies and alternative navigation options to ensure that the experience is accessible to users with functional differences. Ensure that multimodal information is accessible to all users by providing options such as subtitles for videos, audio descriptions for images, contrast and readable font settings, adjustable speed of the information being presented, and flexible interaction methods with different resources.

Table 2. *Cont.*

UX Guideline	Associated Recommendations
G3	G3R1. Enable Co-Editing of Work Outputs: Ensure that your GenAI assistant promotes the co-editing of work outputs, allowing users to enhance them with the system's assistance. G3R2. Identify Opportunities for Human–AI Collaboration: Adapt workflows and cooperative dynamics by identifying points of the value chain where the use of a GenAI assistant can provide significant advantages. Define work outputs amenable to generative processes, ensuring that the responsible use of the solution remains a core principle throughout its implementation. G3R3. Provide Guidance for Crafting Effective Prompts: Offer users guidance on how to improve the creation of prompts to generate work outputs that are more consistent and aligned with users' needs. The solution should provide options and support to help users understand how to define a clear prompt, identify relevant keywords, and establish an appropriate overall structure for writing prompts, tailored to their needs and the solution's capabilities. G3R4. Implement Real-Time Bidirectional Feedback: Incorporate functionalities in your solution that enable real-time bidirectional feedback during human–AI interactions, allowing for the continuous improvement in work outputs generated by the GenAI assistant. G3R5. Provide Contextual Help for New Features: This feature provides users contextual assistance when they are about to use a feature for the first time.
G4	G4R1. Enable Context and Process Flow Configuration: Provide users with the ability to configure factors related to the application context and process flow within the GenAI assistant, tailoring the generation of work outputs to meet their expectations. G4R2. Enable Interaction History Management: Allow users to retain their history of previous interactions with the GenAI assistant, providing them with the ability to manage these records by deleting or preserving those deemed relevant and necessary for their needs and goals. G4R3. Provide Control Over Work Output Parameters: Allow users to control parameters related to generative work outputs according to their needs. For example, users should be able to specify the number of images they wish to produce or the number of alternatives they wish to generate for a specific planning strategy, among other needs, depending on the context of the GenAI assistant solution. G4R4. Personalize the Experience Through Individual Parameters and Preferences: Facilitate human–AI interaction by promoting the personalization of the experience through a set of user-specific information and basic configuration parameters (e.g., personal information, colors, interface layout, images, audio, etc.). Additionally, enable the modeling of a structure for more personalized and user-centered generative response styles tailored to the needs of each user. G4R5. Adapt the Communication Style to Context and User Profile: Ensure that the GenAI assistant offers a communication style in human–GenAI interactions that is adaptable and consistent with both the application context and the specific characteristics of the user profile for whom the solution is intended. G4R6. Control and Monitor Anthropomorphization and Clarify the Nature of AI: Regulate and monitor the anthropomorphization in human–GenAI interactions, reminding users that they are not interacting with a human being (raising their awareness). As such, AI should not be attributed with intelligence, sensitivity, or empathy, and its primary purpose is to assist users more efficiently.
G5	G5R1. Guide Users in Identifying Expected Benefits: Support the user in defining the specific benefits anticipated to arise from using a GenAI assistant on their processes, workflows, and work styles. In addition, it helps them identify the key stages where the solution should be applied to maximize its impact. G5R2. Disclose the Sources of Information Used: Inform stakeholders of the resources (sources) of the original information, which enables the generation of responses that the GenAI system provides to the user. G5R3. Set Realistic Expectations About Capabilities and Limitations: Communicate both the capabilities and limitations of GenAI-based solutions to stakeholders to align user expectations with the system's actual performance. This approach helps build consistent trust in the system's abilities and lays the foundation for future developments that can progressively enhance user confidence in the GenAI system. G5R4. Explain the Generation Processes and Mechanisms: Clearly and simply communicate to users the processes and mechanisms used by AI models that enable the generation of responses or work outputs. G5R5. Promote Safe Practices and Protect Sensitive Data: Inform users about the importance of not using sensitive data during interactions with the GenAI system to ensure data security and prevent practices such as phishing, privacy violations, confidential information leaks, and other security breaches. It is crucial that users are aware of how to interact safely with a GenAI-based assistant and understand the risks associated with the inadvertent introduction of sensitive data. G5R6. Enable Persistence of Information in Interaction Sessions: The solution should be capable of maintaining persistence in the information and elements generated during human–GenAI interaction sessions, thereby enhancing the optimization of prompt engineering. G5R7. Provide Users Control Over Personal Data: Offer intuitive interfaces within the GenAI assistant that allow users to manage their personal information. This includes options to opt out of personal data collection and mechanisms to ensure continuous and direct control over personal data. G5R8. Scale the Solution While Preserving Privacy and Quality: Implement mechanisms that allow the GenAI assistant to scale to a larger user base without compromising data privacy or the quality of the generated responses.
G6	G6R1. Encourage Critical Evaluation of Generated Results: Raise user awareness about the importance of critically analyzing the work outputs generated by GenAI mediation and assessing their accuracy and validity. G6R2. Provide Multiple Generative Response Options: Whenever possible, offer users a set of generative response options produced by the solution, allowing them to select the one that best aligns with their interests and needs. G6R3. Provide Predefined Questions Aligned with the Scope: Consider offering users a predefined list of questions for using a solution that aligns with its scope, allowing users to quickly obtain information. However, note that predefined questions may limit the ability to explore all the system's options.

3.3. Toolkit for UX Design Guidelines for GenAI-Based Assistants

While the guidelines and recommendations provided for the UX design of GenAI-based assistants can offer a solid foundational basis to a team responsible for creating such solutions, they are not sufficient on their own to ensure activities that genuinely enrich the collaborative and creative work dynamics necessary for determining how to integrate these guidelines. This is particularly true in processes that require ideation and collaboration dynamics with a creative focus for UX design [48], which can significantly influence the considerations to be considered in a requirement engineering process within the context of developing GenAI-based solutions.

Card-based design tools have significantly enhanced ideation and creativity processes, fostering high levels of collaboration and interaction among participants [28]. Challenges that require diverse approaches—such as addressing ethical and moral issues in the mediation of information technologies—are widely supported by theoretical approaches, including Value-Sensitive Design (VSD), through the use of card-based toolkits [25].

The use of card-based tools has also been extended to other disciplines, including software engineering. In this field, standards such as Essence apply card-based tools to design artifacts, including Alphas, to specify states and verification criteria [49]. This strategy aims to facilitate communication among team members regarding project evolution, thereby strengthening collaborative work.

Given the importance of collaborative work in the ideation and creativity processes associated with UX design [50], a card-based toolkit was developed. This toolkit incorporates UX design guidelines and recommendations for developing GenAI-based assistants. The structure for the toolkit's design was primarily based on previous experiences, such as the IDEO method card set [37] and the VSD toolkit [38].

Figure 2 presents the digital version of a set of cards from the UX design guidelines toolkit for GenAI-based assistants, specifically corresponding to guidelines G1 "Design Based on Ethical and Responsible Principles" and G2 "Design Based on Principles of Inclusion".

In designing the toolkit, color theory was used to facilitate the understanding and categorization of the guidelines. Each guideline is associated with a specific color that evokes psychological responses aligned with its content: red represents ethics and responsibility; blue conveys urgency [32], associated with inclusion, and suggests trust [33]; orange, for AI–human collaboration, evokes dynamism; purple is linked to personalization and connotes innovation; green, representing reliability, communicates stability; and light blue, for precision, conveys clarity. This color-coding enhances the tool's visual esthetics and improves the cognitive assimilation of the presented concepts.

From a Gestalt psychology perspective, the card design implements principles that optimize the organization and perception of information. Proximity and similarity are applied to the grouping of textual elements and the consistent use of styles, creating visual coherence and facilitating categorization. The principle of closure is evident in how color blocks frame content, forming complete perceptual units. The figure-ground principle, which uses contrasting text over solid color backgrounds, ensures a clear visual hierarchy and optimal legibility. This content structure, grounded in Gestalt principles, seeks to improve the toolkit's usability and enhance the effective communication of guidelines and recommendations for UX design in GenAI solutions [51].

Figure 3 shows the printed version of the toolkit, which presents the deck associated with the six UX guidelines for GenAI-based assistants.



Figure 2. A sample of digital cards from the GenAI toolkit for guidelines G1 and G2 and recommendations G1R1 and G2R1. Created by the author.



Figure 3. A sample of printed version cards of the GenAI toolkit for the six UX guidelines. Created by the author.

3.4. Evaluation

A validation process was conducted to evaluate the contribution of the UX design guidelines toolkit in the design of GenAI-based assistants, focusing on three key components: usability, user experience, and utility. This process was conducted through a hands-on workshop with an experienced software development company of 10 participants organized into three interdisciplinary teams. These teams included UX designers, developers, requirements analysts, and other relevant roles critical to this project.

The initial hypothesis to be validated was that using the toolkit would enhance the perceived usability, utility, and user experience during the UX design process, particularly in the requirements analysis phase of GenAI-based assistant development. This premise was evaluated from the perspective of the interdisciplinary development teams responsible for its implementation.

The toolkit was applied to a set of 11 predefined functional and non-functional requirements for developing a GenAI-based assistant as part of a project to establish an educational assistant for the Valle del Cauca, Colombia, school system. These requirements were defined based on their representativeness in incorporating the UX design guidelines for the development of the GenAI assistant. This selection was conducted in adherence to the principles of the DSRM [52], focusing on requirements that provide representativeness for the validation of the artifacts under evaluation. This project was led by a software development company with the participation of the University Autónoma de Occidente in Cali, Colombia, and the University of Salamanca in Salamanca, Spain. Following the workshop, the participants completed a questionnaire adapted to the MEEGA+ instruments [37] and meCUE 2.0 [38] for a usability assessment in systems. The questionnaire included 23 statements related to the defined evaluation components, rated on a 5-point Likert scale, along with three open-ended questions about positive aspects, negative aspects, and additional comments on the instrument. Table 3 provides a detailed overview of the questions answered by the users.

Table 3. Evaluation questionnaire items for the UX guidelines toolkit for GenAI system.

Factor	Code	Question
Usability	P1	The visual design of the toolkit facilitates the rapid identification of guidelines for GenAI solutions.
Usability	P2	The instructions and examples provided in the toolkit are clear and easily understandable.
Usability	P3	Learning to use the toolkit was straightforward and intuitive for me.
Usability	P4	I believe that most individuals can quickly learn to use this toolkit effectively.
Usability	P5	The recommendations and guidelines provided by the toolkit are clear and comprehensible within the specific context of UX design for GenAI.
Usability	P6	The fonts (size and style) used in the toolkit are easily readable.
Usability	P7	The structure of the toolkit allows for easy navigation across different UX aspects specific to GenAI.
User Experience	P8	I am satisfied with how the toolkit assisted me in addressing specific UX challenges.
User Experience	P9	The use of the toolkit was a pleasant experience, free from frustration.
User Experience	P10	I find the content of the toolkit to be relevant and useful for my work.
User Experience	P11	The toolkit provides concrete and applicable examples of how to implement UX guidelines across various types of GenAI solutions.

Table 3. *Cont.*

Factor	Code	Question
User Experience	P12	Using the toolkit enabled me to enhance the accuracy of the defined requirements.
User Experience	P13	The guidelines and examples provided by the toolkit are useful and applicable to my specific needs.
User Experience	P14	The toolkit promotes collaboration and discussion within the work team.
User Experience	P15	If possible, I would use this toolkit on a daily basis to tackle UX design projects for GenAI.
User Experience	P16	I would choose the toolkit for use in future projects related to GenAI.
User Experience	P17	I prefer using the physical version of the toolkit over the digital version.
Utility	P18	The toolkit helped me identify critical UX aspects that are unique to GenAI solutions.
Utility	P19	The toolkit's recommendations were directly applicable to the UX challenges encountered in the projects.
Utility	P20	The use of the toolkit allowed me to save time in identifying and resolving GenAI-specific UX problems.
Utility	P21	The toolkit provides valuable guidance on how to effectively communicate the capabilities and limitations of GenAI to end-users.
Utility	P22	The toolkit was useful in anticipating and mitigating potential ethical issues in UX design for GenAI.
Utility	P23	The toolkit is a valuable resource for enhancing UX in GenAI projects.

The analysis of the results was conducted in two phases. The first phase involved an analysis of the quantitative data, using non-parametric statistics and the statistical software R-3.4.2 [53] to identify significant conclusions from the participants' evaluations. The second phase involved a thematic analysis of the qualitative information obtained through observations and discussions.

To quantitatively evaluate the toolkit, the analysis incorporates several statistical techniques. The internal consistency of the instruments is assessed through Cronbach's Alpha coefficient, which was calculated for three primary dimensions: usability, user experience, and utility. This coefficient measures the reliability of the item set, with values above 0.70 indicating strong internal consistency [54]. Cronbach's Alpha is formally defined as follows:

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum_{i=1}^{N} \sigma_{Y_i}^2}{\sigma_X^2}\right) \quad (1)$$

where N denotes the number of items, $\sigma_{Y_i}^2$ represents the variance of each item Y_i, and σ_X^2 indicates the total variance of the test.

A Principal Component Analysis (PCA) was applied to reduce data dimensionality and identify the variables explaining the majority of variability within each dimension. The PCA included supplementary variables (role, experience, and frequency of use) to enhance

the interpretation of the results. Formally, PCA attempts to identify a matrix P of principal components, such that the following is true [55]:

$$Z = XP \tag{2}$$

where X is the standardized data matrix, P is the eigenvector matrix defining the directions of maximum variance, and Z is the transformed matrix of the principal components.

To visualize and interpret the results, contribution plots and biplots were generated. Additionally, a sentiment analysis was conducted using natural language processing (NLP) techniques on the qualitative observations collected from the questionnaire.

4. Results and Discussion

4.1. Quantitative Results

The reliability analysis results using Cronbach's Alpha coefficient demonstrated high internal consistency across the three evaluated dimensions. For the usability section, a raw Cronbach's Alpha of 0.82 and a standardized Alpha of 0.88 were obtained, reflecting high reliability with an average inter-item correlation of 0.5. In the user experience dimension, the raw Alpha was 0.82 and the standardized Alpha was 0.81, with an average inter-item correlation of 0.3, indicating moderate consistency. Finally, in the utility dimension, a raw Alpha of 0.84 and a standardized Alpha of 0.91 were achieved, indicating excellent internal consistency with an average inter-item correlation of 0.63.

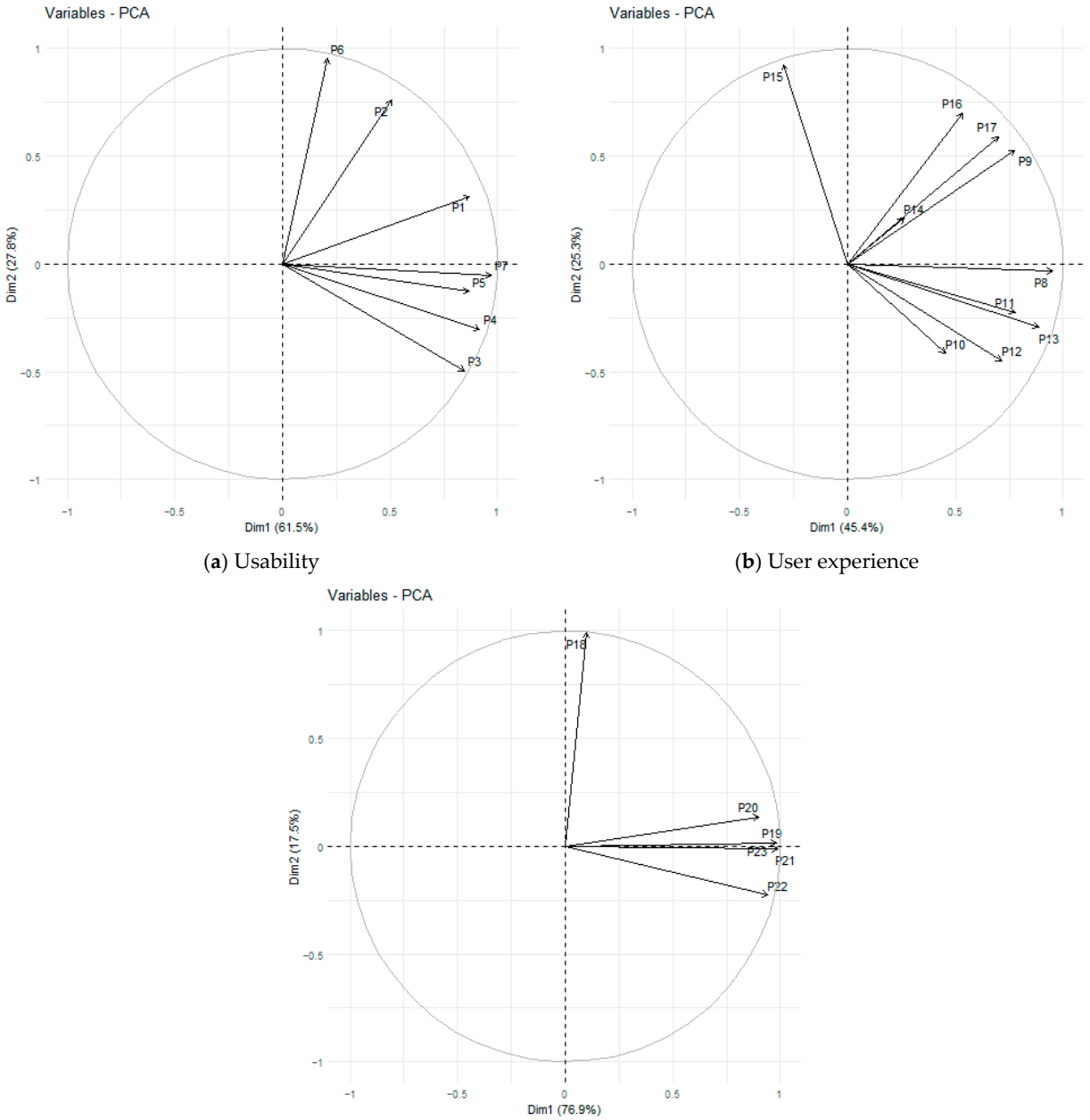
Table 4 presents the cumulative percentage of variance explained by each component across the three categories. The results suggest that retaining two components per category is sufficient to explain at least 70% of the variability, indicating that only the first factorial plane of each category can be interpreted.

Table 4. Cumulative percentage of variance by category.

Component	Cumulative Percentage of Variance		
	Usability	**User Experience**	**Utility**
1	61.5	45.4	76.9
2	89.3	70.7	94.5
3	97.4	83.8	98.1
4	99.6	92.3	99.4
5	100.0	99.0	100.0
6	100.0	100.0	100.0

Figure 4 illustrates the behavior of the variables in the first factorial plane for each category.

The results of the Principal Component Analysis (PCA) reveal distinct patterns across the three evaluated dimensions: usability, user experience, and utility. In the usability dimension, the first principal component explains 61.54% of the variance, while the second component accounts for 27.79%, totaling 89.33% of the explained variance. The first component is dominated by variables related to the clarity and ease of use of the toolkit (P1, P3, P4, P5, P7), which cluster closely on the positive side of the horizontal axis. The second component reflects aspects of accessibility and structure, primarily represented by P6 (font readability) and P2 (operability).



(**a**) Usability (**b**) User experience

(**c**) Utility

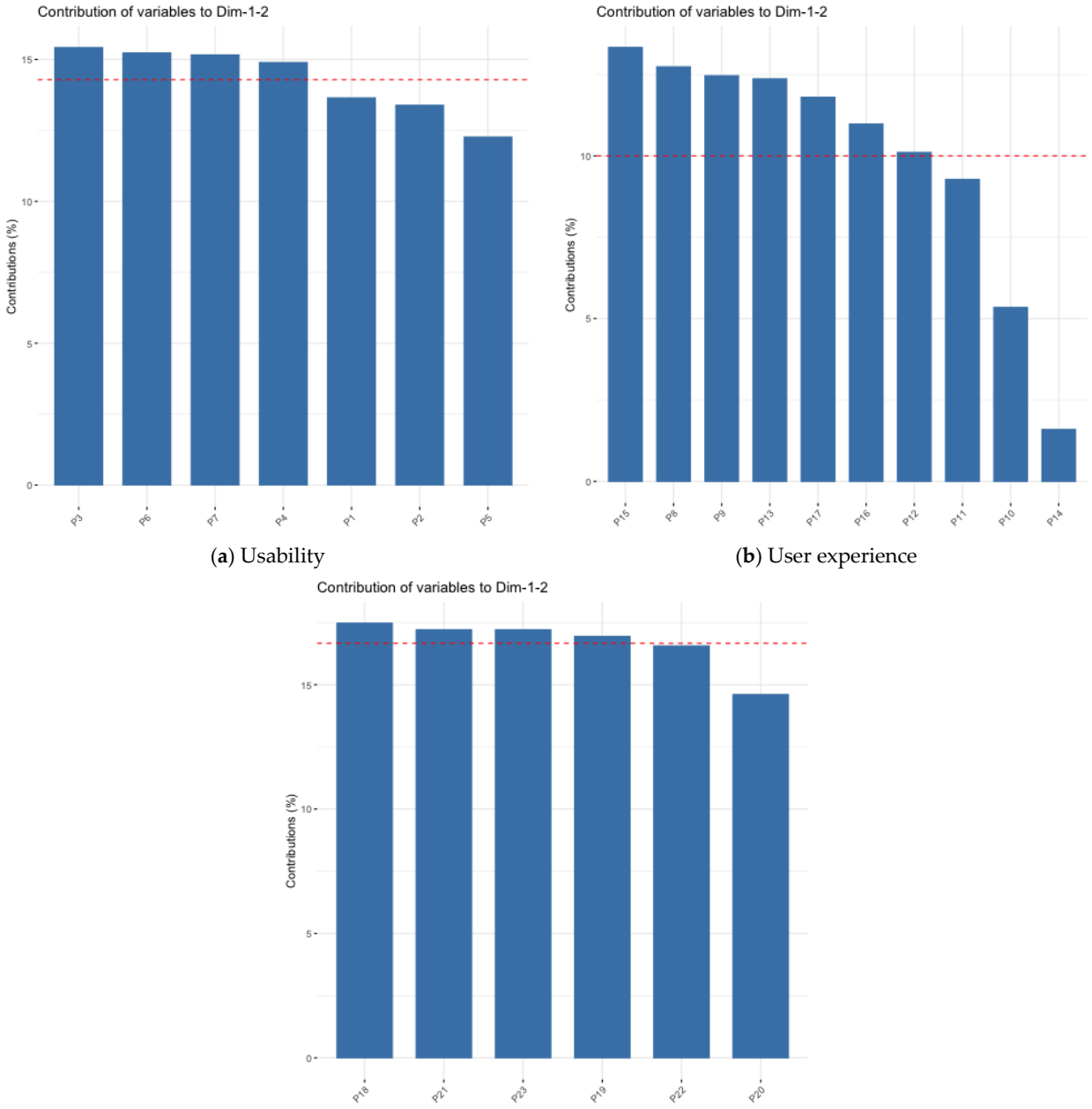
Figure 4. Biplot by category: (**a**) Usability, (**b**) User experience, (**c**) Utility.

In the user experience category, the first component explains 45.42% of the variance, while the second component accounts for 25.29%, totaling 70.71%. A clear separation is observed between two groups of variables: one capturing overall satisfaction and the perceived utility of the toolkit (P8, P10, P11, P12, P13) and another related more closely to future usage intent and format preferences (P9, P15, P16, P17).

In the utility dimension, the first component explains a very high proportion of the variance (76.95%), with all variables (P19 to P23) showing strong correlations and clustering toward the positive side of this component. This suggests that all the assessed aspects of utility contribute similarly to the overall perception of the toolkit's utility. However, P18 (identification of critical UX aspects) stands out due to its strong alignment with the second component, which indicates that it may capture a unique dimension of the toolkit's utility.

Figure 5 presents the contribution analysis of each component. In the usability category, variables P3, P6, P7, and P4 exhibit the highest contributions. These variables correspond

to aspects such as visual design, toolkit structure, ease of learning, and the perception that others can quickly learn to use the toolkit. This result suggests that ease of use and the toolkit's structure are fundamental aspects in the perception of its usability.



(**a**) Usability (**b**) User experience

(**c**) Utility

Figure 5. Contribution of the first factorial plane by category: (**a**) Usability, (**b**) User Experience, (**c**) Utility. The red line represents 1/number of variables × 100%.

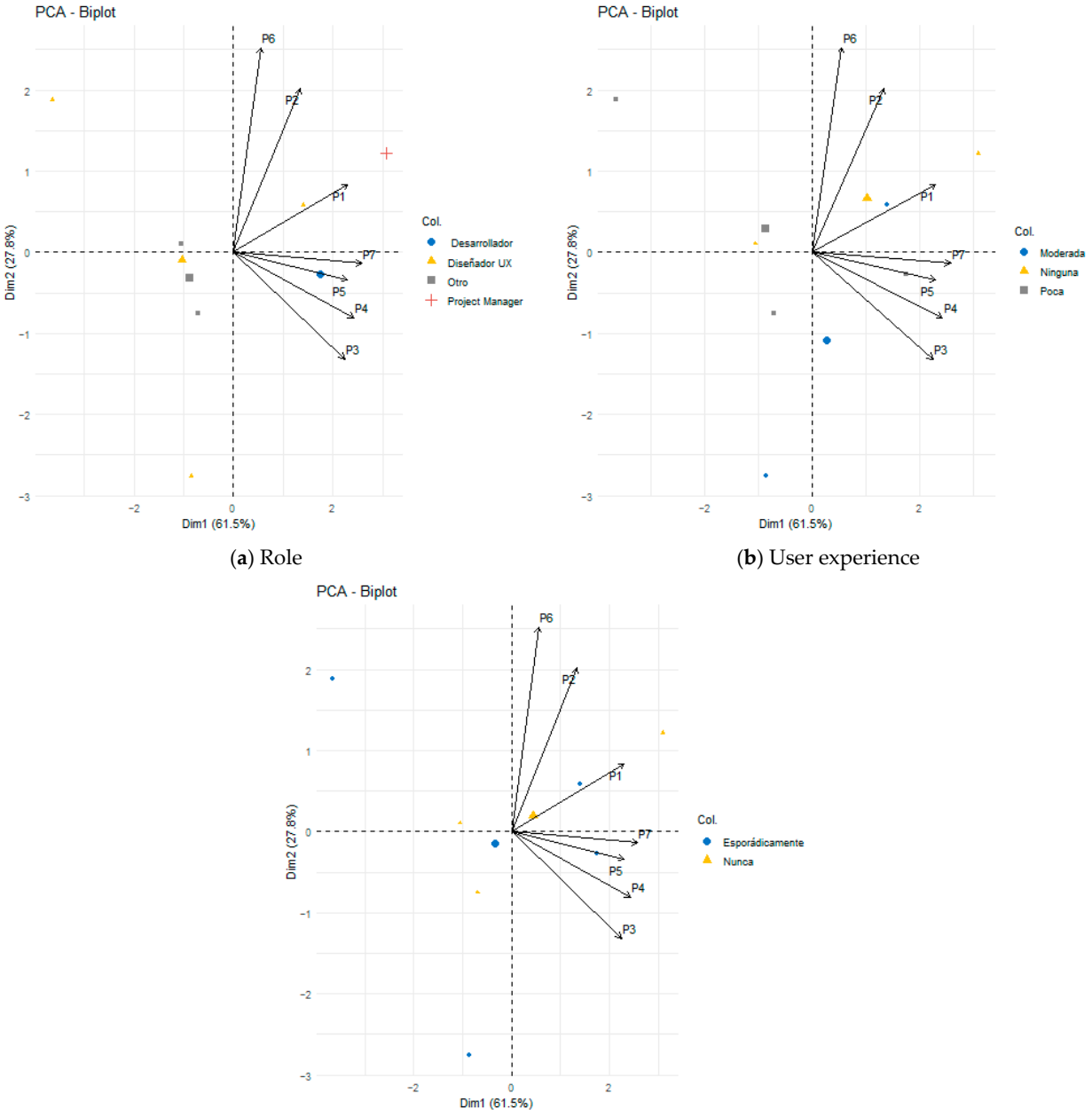
In the user experience category, variables P15, P8, P9, and P13 show the highest contributions. These variables are associated with the willingness to use the toolkit daily, the pleasantness of the user experience, satisfaction with how the toolkit aids in addressing specific UX challenges, and the provision of concrete and applicable examples. This indicates that overall satisfaction, perceived utility, and practical applicability are key aspects of the UX with the toolkit.

In the utility dimension, a more uniform distribution of contributions is observed. Variables P18 and P21 stand out slightly, closely followed by P23, P19, and P22. These variables assess the toolkit's capacity to identify critical UX aspects in GenAI solutions,

provide useful guidance on communicating GenAI capabilities and limitations, and assess its overall value for enhancing user experience in GenAI projects. The high contribution of all the variables in this category indicates that the toolkit is perceived as a comprehensive and valuable tool for addressing various UX aspects in GenAI projects [56].

It is noteworthy that, across all categories, the variables with the highest contributions are related to the practical and applicable aspects of the toolkit, reinforcing its perception as a useful and relevant tool for professionals working on GenAI projects [57,58].

Figure 6 presents the relationship between the collected demographic variables (project role, experience with similar tools, and frequency of use of similar tools) and the usability component.



(**a**) Role

(**b**) User experience

(**c**) Frequency of use

Figure 6. Relationship between demographic variables and usability component: (**a**) Role, (**b**) User Experience, (**c**) Frequency of use.

UX designers and project managers align with variables P1, P2, and P7. This suggests that these roles, often involved in project supervision and planning, place high value on the toolkit's ability to provide clear and easily accessible information. In contrast, developers and other roles display a weaker association with these variables, which may indicate that they prioritize different aspects of the toolkit or interact with it in a distinct manner.

Users with little or no prior experience cluster more closely with variables P3 and P4, which correspond to the ease of learning the toolkit. This indicates that novice users find the toolkit's learning curve manageable, which is a crucial factor for its adoption in diverse teams. In contrast, highly experienced users demonstrated a stronger association with P2 and P6, suggesting that they focused more on the presentation details and structure of the toolkit.

Frequent users of design tools tend to value the same variables—P1, P2, and P6—as users in higher-responsibility roles. This suggests that greater familiarity with the toolkit enhances appreciation of its design elements and clear instructions. In contrast, those who never used design tools showed a stronger association with questions related to ease of learning, implying that their lack of use may be linked to perceived learning barriers.

Figure 7 illustrates the relationship between the collected demographic variables (project role, experience with similar tools, and frequency of use of similar tools) and the UX component.

Project managers and UX designers align more closely with variables P9 (pleasant experience), P16 (choice for future projects), and P17 (preference for the physical version). This alignment indicates that these roles find the toolkit satisfactory and are likely to integrate it into their workflow. Developers, on the other hand, show stronger associations with P11 (concrete examples) and P12 (accuracy improvement), suggesting that they value the toolkit's practical applications and its ability to enhance their work output.

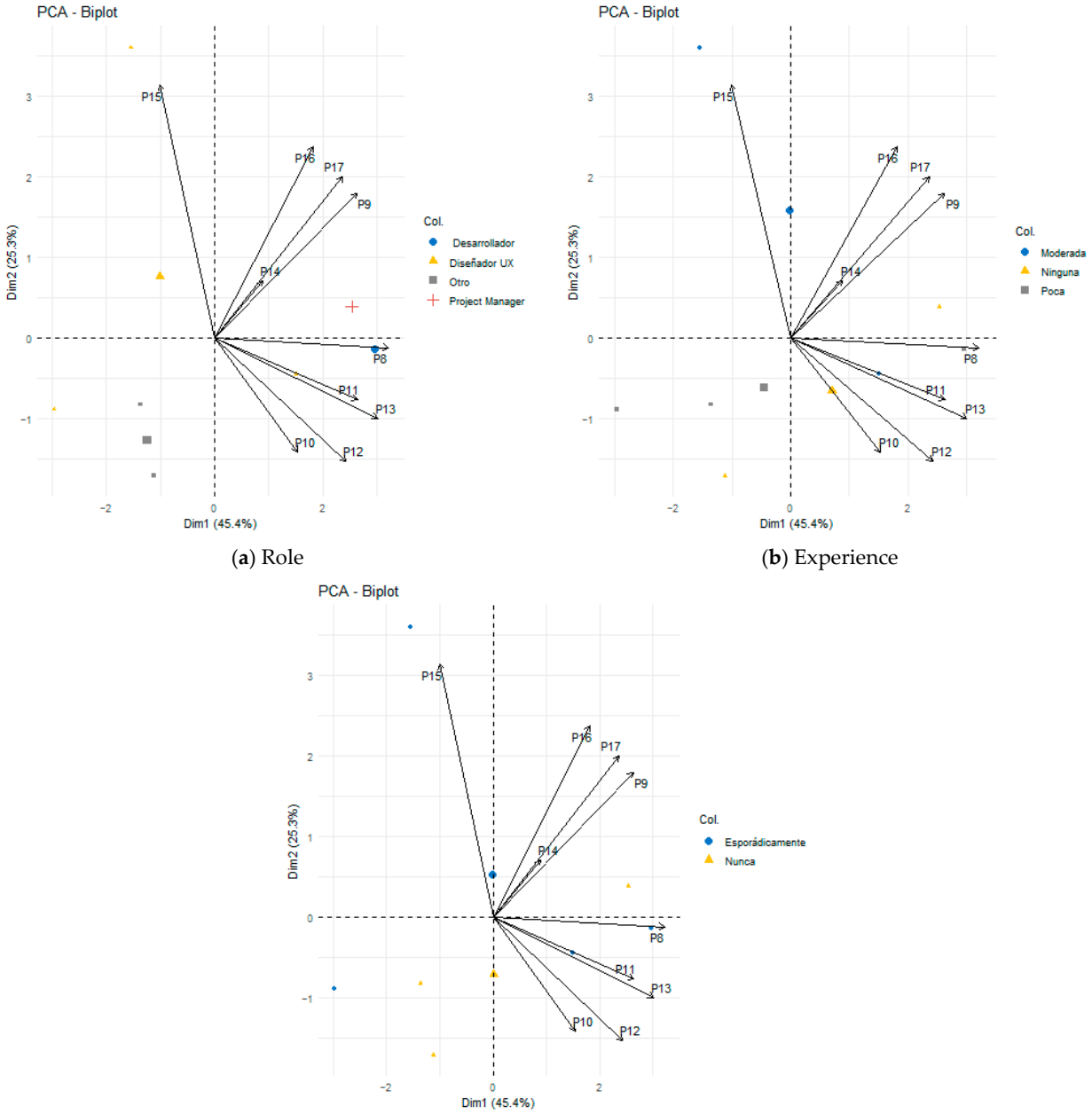
Experienced users exhibit a strong alignment with variables P15 and P16, which indicates that those with more field experience recognize the long-term value of the toolkit and are more likely to integrate it into their regular practices. In contrast, users with little or no experience are more likely to be associated with P12 and P13, suggesting that they appreciate the toolkit's ability to provide concrete guidance and enhance the quality of their work.

Occasional users show a connection with P15 and P16, indicating that even infrequent use fosters appreciation for the toolkit's value in daily work and future projects. Those who never used the toolkit were more likely to have questions about the relevance of the examples (P10, P11), suggesting that demonstrating the toolkit's applicability to their specific tasks could increase adoption rates.

Figure 8 presents the relationship between the collected demographic variables (project role, experience with similar tools, and frequency of use of similar tools) and the utility component.

Project managers and UX designers align with P18 (identification of critical UX aspects), highlighting their focus on high-level UX strategies. Developers, on the other hand, are more associated with variables related to time-saving and the communication of GenAI capabilities, reflecting their interest in practical efficiency and the technical communication aspects of the toolkit.

Highly experienced users align with P18, suggesting that they find the toolkit particularly valuable for identifying critical UX issues in GenAI projects. Users with little or no experience are more likely to be associated with P21 and P22 (ethical issue mitigation), indicating that the toolkit provides essential guidance on complex GenAI aspects for those with less field experience.



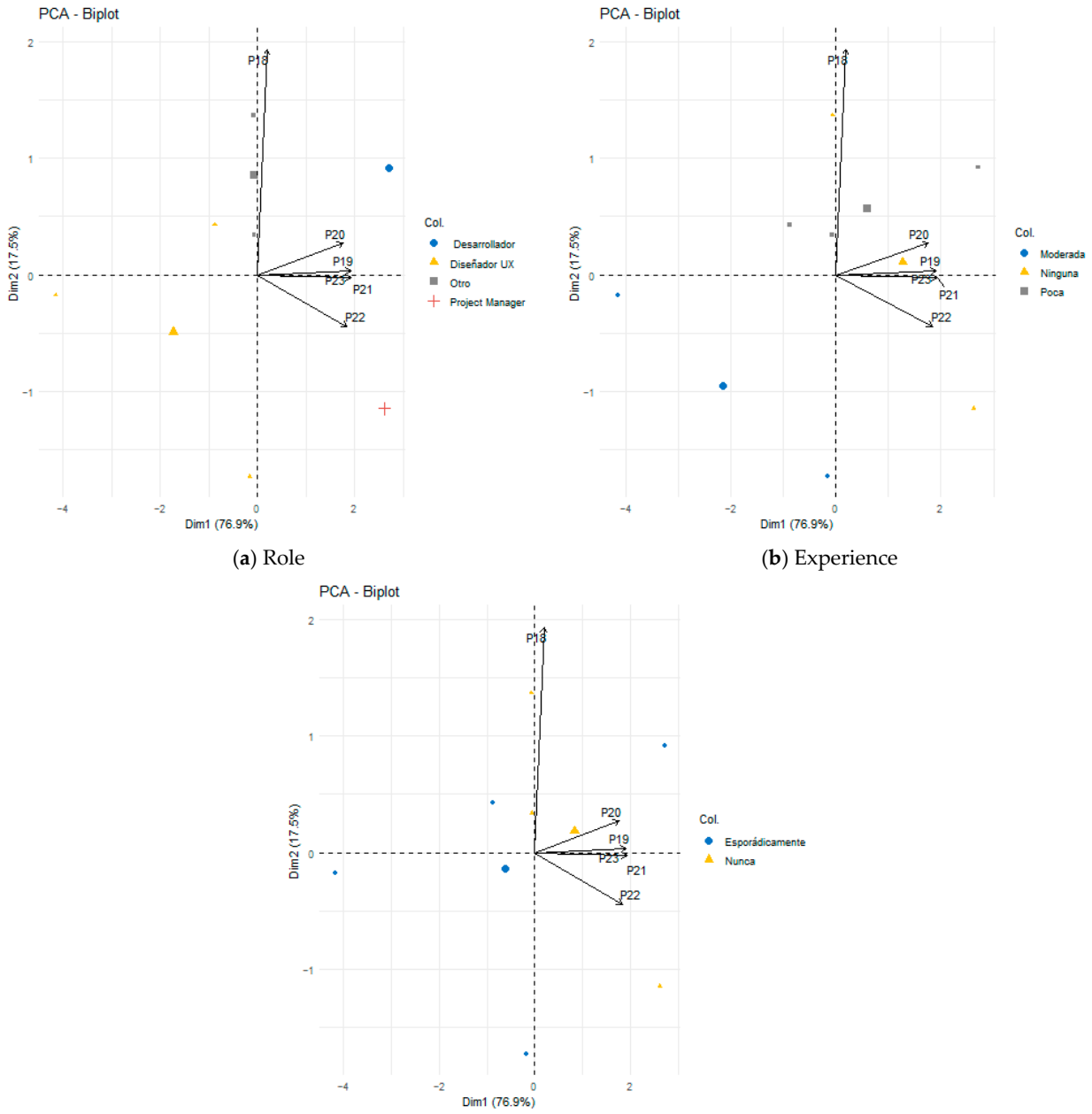
(**a**) Role (**b**) Experience

(**c**) Frequency of use

Figure 7. Relationship between demographic variables and UX component: (**a**) Role, (**b**) Experience, (**c**) Frequency of use.

Occasional users value P18 and P19 (applicability to UX challenges), suggesting that even infrequent use aids in identifying and addressing critical UX issues. Users who have never used design tools align more closely with P20, implying that the toolkit's capacity to address ethical challenges could be a compelling feature to increase adoption within this group.

This analysis reveals that the toolkit provides value across different roles, experience levels, and usage frequencies, with each group deriving unique benefits. The toolkit is particularly effective in offering clear guidance, facilitating learning, and addressing critical UX and ethical issues in GenAI projects. These findings will inform future iterations of the toolkit, ensuring that it continues to meet the diverse needs of its user base and potentially increases its adoption and effectiveness across all groups.



(**a**) Role (**b**) Experience

Figure 8. Relationship between demographic variables and utility component: (**a**) Role, (**b**) Experience, (**c**) Frequency of use.

The quantitative analysis supports the effectiveness and perceived value of the UX guidelines toolkit for the design of GenAI assistants, thereby enhancing the perceived usability, utility, and user experience during the UX design process. The results indicate that the toolkit is viewed as clear, applicable, and highly useful for enhancing UX and addressing specific challenges in GenAI projects. The observed differences across roles, experience levels, and usage frequency suggest that the toolkit is versatile and adaptable to diverse needs and usage contexts, while also highlighting potential areas for improvement in future iterations of the toolkit.

4.2. Qualitative Results

Figure 9 presents the results of the sentiment analysis conducted on the qualitative responses from the questionnaire.

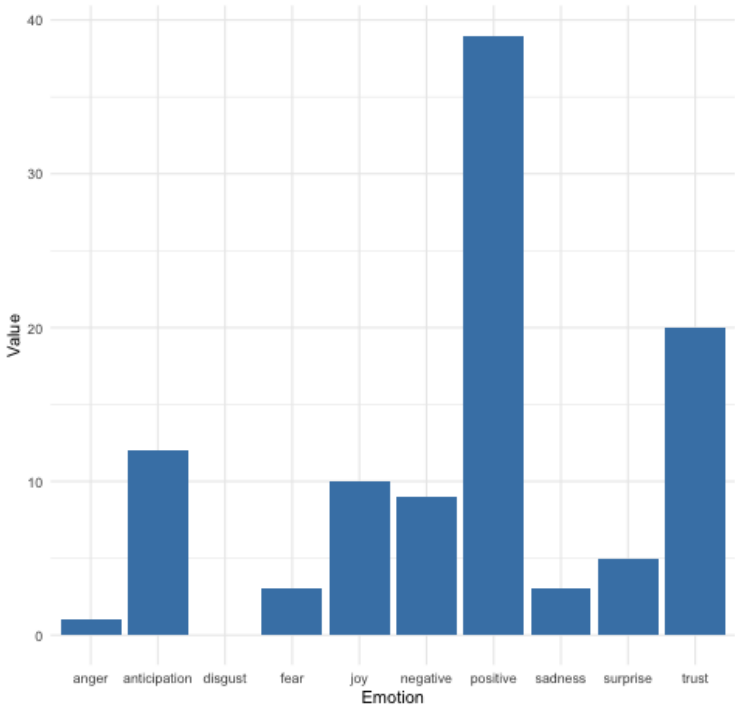


Figure 9. Sentiment analysis of qualitative questionnaire responses.

The results reveals that positive sentiment is predominant, followed by expressions of trust, suggesting that participants generally had a favorable experience with the toolkit and felt confident in its capabilities and applicability. Finally, Figure 10 presents a word cloud generated from the participants' observations.



Figure 10. Word cloud of terms from qualitative responses.

This visualization highlights key terms, such as "participants", "toolkit", "cards", "instructions", and "guidelines", reflecting the most relevant aspects of user interaction with the toolkit.

The qualitative information gathered following the use of the toolkit during the discussion session and observational process was analyzed using a thematic analysis approach, identifying patterns and recurring themes in participants' responses. This analysis provided an in-depth understanding of how users interacted with the toolkit, the challenges they encountered, and the areas in which they perceived the greatest value.

Regarding usability and learning curve, participants demonstrated a general ability to use the toolkit without the need for extensive additional instructions. An initial learning curve was observed, followed by an adaptation phase that allowed users to explore the content more efficiently. Although the toolkit has an intuitive design, it could benefit from more detailed introductory instructions to facilitate an initial understanding of the instrument.

Preferences in terms of format and interaction were evident, with a clear preference for handling the physical cards of the toolkit; this facilitated collective visualization and collaboration in group discussions [59]. Although the participants interacted with the digital version, they found the physical version more useful for collaborative activities. These results suggest that maintaining visual and practical elements in the future design of the instrument is important for enhancing its use in team-based environments.

The toolkit effectively fosters idea exchange and active collaboration among participants. Insightful discussions emerged around topics such as usability and utility, indicating that the proposed instrument is well suited to promote reflection on relevant UX principles in the design of GenAI assistants [60].

Participants required more time to complete the proposed tasks than initially planned. This suggests a need to re-evaluate and adjust the recommended time allocations for effective toolkit application in real productive contexts. Such adjustments are essential to ensure that users can thoroughly explore and integrate the toolkit's guidelines and recommendations without compromising efficiency in UX design processes [61].

The toolkit demonstrated adaptability to various work styles and group dynamics. Participants employed different approaches to organizing and using the cards, indicating that the tool is sufficiently flexible to accommodate diverse methodologies and team collaboration preferences.

5. Conclusions and Future Work

The results in Section 4 are consistent with the initial hypothesis, indicating that the use of the toolkit indeed enhances perceived usability, utility, and user experience in UX design within the development process of GenAI-based assistants, particularly during the requirements analysis phase, where experimentation was conducted. The study participants highlighted aspects such as the clarity of the recommendations, the organized structure of the toolkit, and its utility in addressing specific GenAI design challenges, including human–AI collaboration, ethical considerations in design, and adaptability to various contexts.

The structure of the toolkit, which is grounded in principles of ethical design, customization, and reliability, facilitated the addressing of complex, specific aspects of GenAI-based assistants, such as the need for transparency and alignment with user expectations. In addition, the recommendations on managing the communicability of system limitations and potential emergent behaviors in GenAI were perceived as valuable for mitigating misaligned expectations and fostering responsible human–AI interaction as a foundation for high-quality UX.

For users who require more time to become familiar with the toolkit, several recommendations are suggested to facilitate their engagement. First, providing more detailed introductory instructions can help reduce the initial learning curve by providing a clear guide on the structure and purpose of each card and guideline before the first application in context. Additionally, implementing guided exercises or practical examples can be considered, allowing users to explore the toolkit in a structured and gradual manner, focusing on step-by-step familiarization. In line with this approach, the development of a website is currently being considered to enable stakeholders to access this foundational information before using the toolkit.

Despite these encouraging results, the toolkit has some limitations. The findings suggest the need to validate the toolkit with a broader range of design teams and end-users, particularly those operating in varied organizational and cultural contexts. Expanding this validation would ensure that the toolkit maintains its effectiveness and applicability

across diverse scenarios, thereby contributing to its robustness and adaptability to different design needs.

In future work, we will suggest investigating the integration of the toolkit into collaborative methodologies that involve end-users throughout the entire lifecycle of the GenAI-based assistant development process, from the early design phases. This includes human-centered design methodologies, in which users actively participate from the ideation phase to the final implementation, ensuring that development aligns with users' needs and expectations. Value-Sensitive Design is also applicable, as it integrates ethical and moral principles directly into the design process. In addition, employing these approaches in iterative design frameworks and agile methodologies would support the continuous adaptation of UX guidelines to respond to changing user needs and contexts. This research could be extended to study user satisfaction, the efficiency of human-centered design processes, and, ultimately, its validation within human–GenAI assistant interactions.

Author Contributions: Conceptualization, C.A.P., A.S. and J.C.E.; methodology, C.A.P., J.A.O., P.A.C. and J.C.E.; validation, C.A.P., J.C.E., J.A.O., J.S.D., D.A.C. and A.S.M.; formal analysis, J.A.O., J.C.E., J.M.N.V., C.A.P. and F.D.l.P.; investigation, C.A.P., A.S., J.C.E., A.S.M. and P.A.C.; resources J.S.D. and D.A.C.; data curation, J.A.O. and J.C.E.; writing—original draft preparation C.A.P., A.S., J.C.E. and J.A.O.; writing—review and editing J.M.N.V., F.D.l.P. and C.A.P.; supervision, C.A.P. and A.S.; visualization, J.A.O. and J.C.E.; project administration, C.A.P., J.S.D. and F.D.l.P.; funding acquisition J.M.N.V. and F.D.l.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of the International Chair Project on Trustworthy Artificial Intelligence and Demographic Challenge within the National Strategy for Artificial Intelligence (ENIA), in the framework of the European Recovery, Transformation and Resilience Plan. Reference: TSI-100933-2023-0001. This project is funded by the Secretary of State for Digitalization and Artificial Intelligence and by the European Union (Next Generation).

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Talentum Corporation with internal code 23INTER-461 (approved on January 2024) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: All the original datasets corresponding to the three case studies are available in the following repository: https://github.com/Juankidd/toolkitgeai (accessed on 1 November 2024).

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

1. Chen, A.; Liu, L.; Zhu, T. Advancing the democratization of generative artificial intelligence in healthcare: A narrative review. *J. Hosp. Manag. Health Policy* **2024**, *8*, 12. [CrossRef]
2. Dessimoz, C.; Thomas, P.D. AI and the democratization of knowledge. *Sci. Data* **2024**, *11*, 268. [CrossRef] [PubMed]
3. Rajaram, K.; Tinguely, P.N. Generative artificial intelligence in small and medium enterprises: Navigating its promises and challenges. *Bus. Horiz.* **2024**, *67*, 629–648. [CrossRef]
4. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoriet, J.; Jones, L.; Gomez, A.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 11.
5. Dix, A.; Finlay, J.; Abowd, G.D.; Beale, R. *Human-Computer Interaction*; Taylor & Francis, Inc.: Philadelphia, PA, USA, 2004.
6. Xu, W.; Dainoff, M. Enabling human-centered AI: A new junction and shared journey between AI and HCI communities. *Interactions* **2023**, *30*, 42–47. [CrossRef]
7. Xu, W.; Gao, Z. Enabling human-centered AI: A methodological perspective. In Proceedings of the 2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS), Ontario, ON, Canada, 15–17 May 2024.

8. Capel, T.; Brereton, M. What is Human-Centered about Human-Centered AI? A Map of the Research Landscape. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023.
9. Sison, A.J.G.; Daza, M.T.; Gozalo-Brizuela, R.; Garrido-Merchán, E.C. ChatGPT: More Than a "Weapon of Mass Deception" Ethical Challenges and Responses from the Human-Centered Artificial Intelligence (HCAI) Perspective. *Int. J. Hum. Comput. Interact.* **2023**, *40*, 4853–4872. [CrossRef]
10. Casteleiro-Pitrez, J. Generative artificial intelligence image tools among future designers: A usability, user experience, and emotional analysis. *Digital* **2024**, *4*, 316–332. [CrossRef]
11. Takaffoli, M.; Li, S.; Mäkelä, V. Generative AI in User Experience Design and Research: How Do UX Practitioners, Teams, and Companies Use GenAI in Industry? In Proceedings of the 2024 ACM Designing Interactive Systems Conference, Copenhagen, Denmark, 1–5 July 2024.
12. Peláez, C.; Solano, A.; Nuñez, J.; Castro, D.; Cardona, J.; Duque, J.; Espinosa, J.; Montaño, A.; De la Prieta, F. Designing User Experience in the Context of Human-Centered AI and Generative Artificial Intelligence: A Systematic Review. In Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence—DCAI 2024, Salamanca, Spain, 26–29 June 2024.
13. Bingley, W.J.; Curtis, C.; Lockey, S.; Bialkowski, A.; Gillespie, N.; Haslam, S.A.; Worthy, P. Where is the human in human-centered AI? Insights from developer priorities and user experiences. *Comput. Hum. Behav.* **2023**, *141*, 107617. [CrossRef]
14. Baabdullah, A.M.; Alalwan, A.A.; Algharabat, R.S.; Metri, B.; Rana, N.P. Virtual agents and flow experience: An empirical examination of AI-powered chatbots. *Technol. Forecast. Soc. Chang.* **2022**, *181*, 121772. [CrossRef]
15. Gill, S.S.; Xu, M.; Patros, P.; Wu, H.; Kaur, R.; Kaur, K.; Buyya, R. Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet Things Cyber-Phys. Syst.* **2024**, *4*, 19–23. [CrossRef]
16. Banh, L.; Strobel, G.; Markets, E. Generative artificial intelligence. *Electron. Mark.* **2023**, *33*, 63. [CrossRef]
17. Peruchini, M.; da Silva, G.M.; Teixeira, J.M. Between artificial intelligence and customer experience: A literature review on the intersection. *Discov. Artif. Intell.* **2024**, *4*, 4. [CrossRef]
18. York, E. Evaluating ChatGPT: Generative AI in UX Design and Web Development Pedagogy. In Proceedings of the 41st ACM International Conference on Design of Communication, Orlando, FL, USA, 26–28 October 2023.
19. Liu, Y.; Siau, K. Generative Artificial Intelligence and Metaverse: Future of Work, Future of Society, and Future of Humanity. In Proceedings of the AI-generated Content, Proceedings of the first International Conference, AIGC 2023, Shanghai, China, 25–26 August 2023; Springer: Singapore, 2023.
20. Liu, F.; Zhang, M.; Budiu, R. AI as a UX Assistant. 27 October 2023. Available online: https://www.nngroup.com/articles/ai-roles-ux/ (accessed on 21 February 2024).
21. Qadri, R.; Shelby, R.; Bennett, C.L.; Denton, E. AI's Regimes of Representation: A Community-centered Study of Text-to-Image Models in South Asia. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, Chicago, IL USA, 12–15 June 2023.
22. Wang, X.; Attal, M.I.; Rafiq, U.; Hubner-Benz, S. Turning Large Language Models into AI Assistants for Startups Using Prompt Patterns. In Proceedings of the International Conference on Agile Software Development, Copenhagen, Denmark, 13–17 June 2022.
23. Shah, C.S.; Mathur, S.; Vishnoi, S.K. Continuance Intention of ChatGPT Use by Students. In Proceedings of the International Working Conference on Transfer and Diffusion of IT, Nagpur, India, 15–16 December 2023.
24. Li, J.; Cao, H.; Lin, L.; Hou, Y.; Zhu, R.; El Ali, A. User experience design professionals' perceptions of generative artificial intelligence. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024.
25. Kim, P.W. A Framework to Overcome the Dark Side of Generative Artificial Intelligence (GAI) Like ChatGPT in Social Media and Education. *IEEE Trans. Comput. Soc. Syst.* **2023**, *11*, 5266–5274. [CrossRef]
26. Asadi, A.R. LLMs in Design Thinking: Autoethnographic Insights and Design Implications. In Proceedings of the 2023 5th World Symposium on Software Engineering, Tokyo, Japan, 22–24 September 2023.
27. Sun, J.; Liao, Q.V.; Agarwal, M.M.M.; Houde, S.; Talamadupula, K.; Weisz, J.D. Investigating explainability of generative AI for code through scenario-based design. In Proceedings of the 27th International Conference on Intelligent User Interfaces, Helsinki, Finland, 22–25 March 2022.
28. Oniani, D.; Hilsman, J.; Peng, Y.; Poropatich, R.K.; Pamplin, J.C.; Legault, G.L.; Wang, Y. Adopting and expanding ethical principles for generative artificial intelligence from military to healthcare. *NPJ Digit. Med.* **2023**, *6*, 225. [CrossRef] [PubMed]
29. Shin, D.; Ahmad, N. Algorithmic nudge: An approach to designing human-centered generative artificial intelligence. *Computer* **2023**, *56*, 95–99. [CrossRef]
30. Tlili, A.; Shehata, B.; Adarkwah, M.A.; Bozkurt, A.; Hickey, D.T.; Huang, R.; Agyemang, B. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learn. Environ.* **2023**, *10*, 15. [CrossRef]
31. Fisher, J. Centering the Human: Digital Humanism and the Practice of Using Generative AI in the Authoring of Interactive Digital Narratives. In Proceedings of the International Conference on Interactive Digital Storytelling, Kobe, Japan, 11–15 November 2023.

32. Mao, Y.; Rafner, J.; Wang, Y.; Sherson, J. A Hybrid Intelligence Approach to Training Generative Design Assistants: Partnership Between Human Experts and AI Enhanced Co-Creative Tools. *Front. Artif. Intell. Appl.* **2023**, *368*, 108–123.
33. Guo, M.; Zhang, X.; Zhuang, Y.; Chen, J.; Wang, P.; Gao, Z. Exploring the Intersection of Complex Aesthetics and Generative AI for Promoting Cultural Creativity in Rural China after the Post-Pandemic. In *AI-Generated Content, Proceedings of the First International Conference, AIGC 2023, Shanghai, China, 25–26 August 2023*; Springer Nature: Singapore, 2023; pp. 313–331.
34. Geerts, G. A design science research methodology and its application to accounting information systems research. *Int. J. Account. Inf. Syst.* **2011**, *12*, 142–151. [CrossRef]
35. Adikari, S.; McDonald, C.; Campbell, J. A design science framework for designing and assessing user experience. In Proceedings of the Human-Computer Interaction, Design and Development Approaches: 14th International Conference, HCI International, Orlando, FL, USA, 9–14 July 2011.
36. Hollander, M.; Wolfe, D. *Nonparametric Statistical Methods: Solutions Manual to Accompany*; Wiley-Interscience: Hoboken, NJ, USA, 1999.
37. IDEO. *IDEO Method Cards: 51 Ways to Inspire Design*; William Stout: San Francisco, CA, USA, 2003.
38. Friedman, B.; Hendry, D. The envisioning cards: A toolkit for catalyzing humanistic and technical imaginations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 30 April–5 May 2012.
39. Friedman, B.; Hendry, D. *Value Sensitive Design: Shaping Technology with Moral Imagination*; MIT Press: Boston, MA, USA, 2019.
40. Sarsenbayeva, Z.; van Berkel, N.; Hettiachchi, D.; Tag, B.; Velloso, E.; Goncalves, J.; Kostakos, V. Mapping 20 years of accessibility research in HCI: A co-word analysis. *Int. J. Hum.-Comput. Stud.* **2023**, *175*, 103018. [CrossRef]
41. Hassan, A. Factors Affecting the Use of ChatGPT in Mass Communication. In *Emerging Trends and Innovation in Business and Finance*; Springer Nature: Singapore, 2023; pp. 671–685.
42. Brandtzaeg, P.B.; You, Y.; Wang, X.; Lao, Y. "Good" and "Bad" Machine Agency in the Context of Human-AI Communication: The Case of ChatGPT. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 23–28 July 2023.
43. Rana, N.P.; Pillai, R.; Sivathanu, B.; Malik, N. Assessing the nexus of Generative AI adoption, ethical considerations and organizational performance. *Technovation* **2024**, *135*, 103064. [CrossRef]
44. Creely, E. The possibilities; limitations, and dangers of generative AI in language learning and literacy practices. In Proceedings of the International Graduate Research Symposium 2023, Hanoi, Vietnam, 27–29 October 2023.
45. Weisz, J.D.; Muller, M.; He, J.; Houde, S. Toward General Design Principles for Generative AI Applications. 13 January 2023. Available online: https://arxiv.org/abs/2301.05578 (accessed on 17 February 2024).
46. Sarraf, S.; Kar, A.K.; Janssen, M. How do system and user characteristics, along with anthropomorphism, impact cognitive absorption of chatbots–Introducing SUCCAST through a mixed methods study. *Decis. Support Syst.* **2024**, *178*, 114132. [CrossRef]
47. Perri, L. What's New in the 2023 Gartner Hype Cycle for Emerging Technologies. 23 August 2023. Available online: https://www.gartner.com/en/articles/what-s-new-in-the-2023-gartner-hype-cycle-for-emerging-technologies (accessed on 1 July 2024).
48. Harley, H. Ideation in Practice: How Effective UX Teams Generate Ideas. 29 October 2017. Available online: https://www.nngroup.com/articles/ideation-in-practice/ (accessed on 19 November 2024).
49. Object Management Group. *Essence–Kernel and Language for Software Engineering Methods*; OMG: Needham, MA, USA, 2018.
50. Feng, K.K.; Li, T.W.; Zhang, A.X. Understanding collaborative practices and tools of professional UX practitioners in software organizations. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023.
51. Wagemans, J.; Elder, J.H.; Kubovy, M.; Palmer, S.E.; Peterson, M.A.; Singh, M.; von der Heydt, R. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychol. Bull.* **2012**, *138*, 1172–1217. [CrossRef] [PubMed]
52. Johannesson, P.; Perjons, E. *An Introduction to Design Science*; Springer: Stockholm, Sweden, 2014.
53. Team, R.C. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022.
54. Nunnally, J. *Psychometric Theory*, 2nd ed.; McGraw-Hill: New York, NY, USA, 1978.
55. Jolliffe, I.T. Principal component analysis for special types of data. In *Principal Component Analysis*; Springer: New York, NY, USA, 2002; pp. 338–372.
56. Shneiderman, B. *Human-Centered AI*; Oxford University Press: Oxford, UK, 2022.
57. Elliot, A.J.; Maier, M.A. Color psychology: Effects of perceiving color on psychological functioning in humans. *Annu. Rev. Psychol.* **2014**, *65*, 95–120. [CrossRef] [PubMed]
58. Kress, G.; van Leeuwen, T. Colour as a semiotic mode: Notes for a grammar of colour. *Vis. Commun.* **2002**, *1*, 343–368. [CrossRef]
59. Roy, R.; Warren, J.P. Card-based design tools: A review and analysis of 155 card decks for designers and designing. *Des. Stud.* **2019**, *63*, 125–154. [CrossRef]

60. Petri, G.; von Wangenheim, C.G.; Borgatto, A.F. MEEGA+, systematic model to evaluate educational games. In *Encyclopedia of Computer Graphics and Games*; Springer International Publishing: Cham, Switzerland, 2024; pp. 1112–1119.
61. Minge, M.; Thüring, M. The meCUE questionnaire (2.0): Meeting five basic requirements for lean and standardized UX assessment. In Proceedings of the Design, User Experience, and Usability: Theory and Practice: 7th International Conference, DUXU 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, 15–20 July 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.