Proceedings of the 11th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'25)
Paris, France - August, 2025
Paper No. MHCI 111
DOI: 10.11159/mhci25.111

Unveiling Chatbot APP Characters: A Socio-Cultural Analysis of App Representations and User Perceptions

Lingyun Yue[1]
[1]University of Amsterdam
1012 CZ, Amsterdam, Netherlands
Lingyun.yue.yue@student.uva.nl

Abstract **-** This study explores the socio-cultural logics underlying the character construction of chatbot applications by examining how developers strategically design and represent chatbot personas to align with users' emotional expectations and normative imaginaries. Based on a dataset of 110 chatbot apps collected from Google Play, it employs a mixed-methods approach that includes keyword frequency analysis, co-occurrence network visualization using Gephi, and iconographic analysis. Chatbot character attributes are categorized into three dimensions: image, identity, and function. Findings indicate a predominant emphasis on anthropomorphic features, particularly the combination of "human," "social," and "friend," suggesting that developers prioritize the construction of emotionally available yet non-intimate chatbot companions. While utilitarian roles such as "assistant" or "knowledge provider" appear less frequently, social identities like "friend" or "roleplay" are foregrounded and often visually reinforced through gendered and sometimes sexualized icon designs. Drawing on social constructivism and sociomateriality, this study argues that chatbot character design reflects developers' projections of the "ideal user," shaped by existing social norms and stereotypes. By centring "similarity" as a key design logic, these representations risk reproducing cultural biases, particularly those surrounding gender and emotional labour, under the guise of user familiarity. The study concludes by emphasizing the need for more critical and reflexive design practices that challenge rather than reinforce dominant socio-technical imaginaries embedded in human–AI interaction.

Keywords**:** Anthropomorphism; Character construction; Chatbot applications; Gender bias; Human–AI interaction

1. Introduction

Chatbots are typical examples of artificial intelligence systems and "one of the most basic and widespread examples of human-computer interaction (HCI)" [1]. A chatbot is a computer program that responds like an intelligent entity during text or voice conversations and understands one or more human languages through natural language processing (NLP) [2]. Currently, emerging NLP models like ChatGPT bring new vitality to chatbot apps. According to social constructivist perspectives, technologies—including mobile applications—are not neutral tools but socio-cultural artifacts shaped by human choices, institutional norms, and cultural imaginaries.

Chatbots will explicitly identify their characters to cater to their audience, and will emphasize different functionalities. The characterization of chatbots can be seen as a symbolic construct aimed at giving specific social identities and behavioural paradigms to chatbots. Research by Go and Sundar [3] indicates that appropriately anthropomorphized design of chatbots increases emotional connection between chatbots and users. If users perceive the chat agent as more similar to themselves, they may provide more favourable evaluations of the chat agent. "This human similarity can be implied by constructing characters, including images (visual cues) and identities (identity cues)". Lee and Oh [4] propose that this anthropomorphic cue can shape social cognition. Additionally, besides textual descriptions of characters, users' perception and emotions towards app characters can also be influenced by icon design [5]. Further critical research by da Costa suggests a bias in the construction of characters [6]. Developers tend to display "feminine attributes" of computer in human-computer interaction behaviour and stereotype behaviours to illustrate the characteristics of female roles, exposing gender bias in the process.

In current research, there is a lack of empirical analysis on how Chatbots convey their characters to users, and this paper aims to address this gap. Therefore, the research questions are as follows: What are the common characters found in Chatbot apps? Regarding the tendency of these characters, what underlying logic might developers have?

It is important to clarify that this study does not examine developers' intentions directly. Instead, it analyzes developer-facing representations (such as app descriptions and icons) to infer patterns and logics of character construction. The frequency of keywords in the application description page can reflect how developers convey different Chatbot characters to

users and identify the application through discourse. To address the research question, I used Google Play as the analysis platform and conducted word frequency analysis on the description and body of the app interface in the Google Play. This analysis aimed to extract the keywords that related to character, and facilitating the formation of preliminary judgments about the chatbot's character. Specifically, characters include three categories: image, identity, and function. I created a keyword co-occurrence matrix and utilized Gephi to draw a keyword co-occurrence network, followed by social network centrality analysis of the keywords. Finally, I combined the visual image presentation of Chatbot APP icons to analyse the underlying logic behind the construction of characters and engage it in existing ideological debates.

2. Method

2.1 Data Collection

To examine the character construction of chatbot applications, I collected data from Google Play in March 2025 using four keyword queries: “AI Chatbot,” “Chatbot,” “AI Talk,” and “AI Chat.” The initial scraping process was conducted using 4CAT, a digital methods tool for web data extraction. All available metadata, including the app name, description, and visual icon, was saved into a structured .csv file. After initial collection, a manual screening process was performed to exclude irrelevant or duplicate entries, such as game-based chat interfaces or apps not primarily designed for human–AI interaction. The final dataset consisted of 110 distinct chatbot applications.

2.2 Text Processing and Keyword Extraction

The app descriptions and titles were processed using Gooseeker to segment the text and extract keywords. The analysis was limited to English-language descriptions to maintain linguistic consistency. From the full set of extracted terms, 76 keywords related to chatbot character attributes were initially identified based on the criteria proposed by Go and Sundar [3], who categorize chatbot anthropomorphism along visual and identity cues. To refine the dataset, synonymous terms were manually merged (e.g., “lover” and “boyfriend/girlfriend” collapsed under “lovers”), and keywords were sorted into three predefined categories:

Image attributes (e.g., “human,” “anime”) refer to visual or aesthetic cues that shape users’ immediate perception of the chatbot’s persona. These descriptors often suggest anthropomorphism, stylization, or visual familiarity.

Identity roles (e.g., “friend,” “teacher,” “soulmate”) capture the relational or social positions that chatbots are assigned within app descriptions. These labels imply emotional or interpersonal expectations tied to the chatbot’s behavior.

Functional descriptors (e.g., “assistant,” “knowledge,” “suggestions”) denote the task-oriented purposes chatbots are designed to fulfill, reflecting instrumental utility rather than emotional engagement.

Table 1: keywords about Chatbot character.

Image	***Identity***	***Function***
human	roleplay	social
chatbot	friend	suggestions
anime	teacher	knowledge
	coach	health
	soulmate	education
	lovers	creativity
	assistant	sexual
	guides	entertainment
	actor	tour

Where ambiguity or overlap existed (e.g., “coach” or “health”), classification was determined by the keyword’s dominant connotation within its contextual usage across the dataset, guided by close reading of app descriptions. This classification scheme enabled a structured and comparative analysis of how character traits are constructed across different chatbot applications, which resulted in a final list of 21 consolidated character-related keywords used in subsequent analysis (see Table 1). Following that, I employed the research method outlined by Dieter in 2021[7]. I calculated the frequency of occurrence for each key word, creating frequency tables and word clouds.

2.3 Keyword Co-occurrence Analysis

To investigate how character elements are combined in app representations, I employed the method proposed by Sedighi [8], constructed a keyword co-occurrence matrix. A co-occurrence was defined as the appearance of two keywords within the same app description field (i.e., per-app granularity). Using Gooseeker, I computed the frequency with which any two keywords appeared together across all 110 app entries.

The resulting co-occurrence matrix was imported into Gephi, a network visualization software. I employed the Fruchterman-Reingold layout algorithm to render the network. Node size was scaled by degree centrality to reflect the prominence of each character keyword, and edge thickness was determined by the co-occurrence weight. Keywords were colour-coded by category: image (purple), identity (red), and function (green), facilitating the visual identification of dominant combinations and structural clusters.

2.4 Iconographic Analysis

To complement the textual analysis, I conducted a visual examination of app icons, drawing on semiotic and gender studies-informed criteria. From the 110 apps, a subset of 33 icons was manually selected to ensure representational diversity across function types. The sample was purposively selected to cover a broad range of character types (e.g., assistant, friend, roleplay) and visual styles (e.g., anime, realistic, symbolic), ensuring both functional and aesthetic diversity for comparative analysis.

Icons were classified into four categories: Robot-themed icons (e.g., mechanical faces, circuit elements); Human photo-based icons (realistic avatars or selfies); Anime-style illustrations (stylized, fictional human figures); ChatGPT mimicry icons (monochrome logos or generic assistant symbols). Visual cues suggesting gendered traits or sexualization (e.g., makeup, revealing outfits, anthropomorphic features) were recorded and cross-referenced with the app’s stated character type (e.g., “friend,” “roleplay”). This mixed-methods approach enables both the quantitative mapping of chatbot character discourses and the qualitative interpretation of their ideological underpinnings.

3. Analysis

3.1 Word Frequency Analysis

Word frequency involves quantitatively counting the occurrences of each keyword. From the frequency of keyword, we can make a preliminary judgment about the popularity of specific types of characters. Figure 1 presents the categories and frequencies of all keywords in histogram and word cloud. It's important to note that multiple keywords may be mentioned in the description of the same application, and these keywords are not mutually exclusive.

From Figure 1, we can observe that "human" appears frequently in the Image category, indicating a strong tendency towards anthropomorphism in chatbots. Among the identity-related character keywords, "roleplay" is significantly ahead. it is possibly driven by the need for multi-identity of chatbots, suggesting that the “roleplay” can cover various other identities. The second-ranking identity is "friend", reflecting developers' emphasis on people's need for emotional communication with chatbots, while identity similar to ChatGPT's, such as "assistant”, is less frequently mentioned in chatbot apps. Interestingly, although developers tend to acknowledge emotional communication about friendship, there isn't much mention of emotional communication related to more intimate relationships like "boyfriend" or "girlfriend." I will discuss this point further in the discussion section. As for the function of the Chatbots, the frequency of social functionality far outweighs that of knowledge and suggestion functions emphasized by language models such as Bing and Chat GPT.

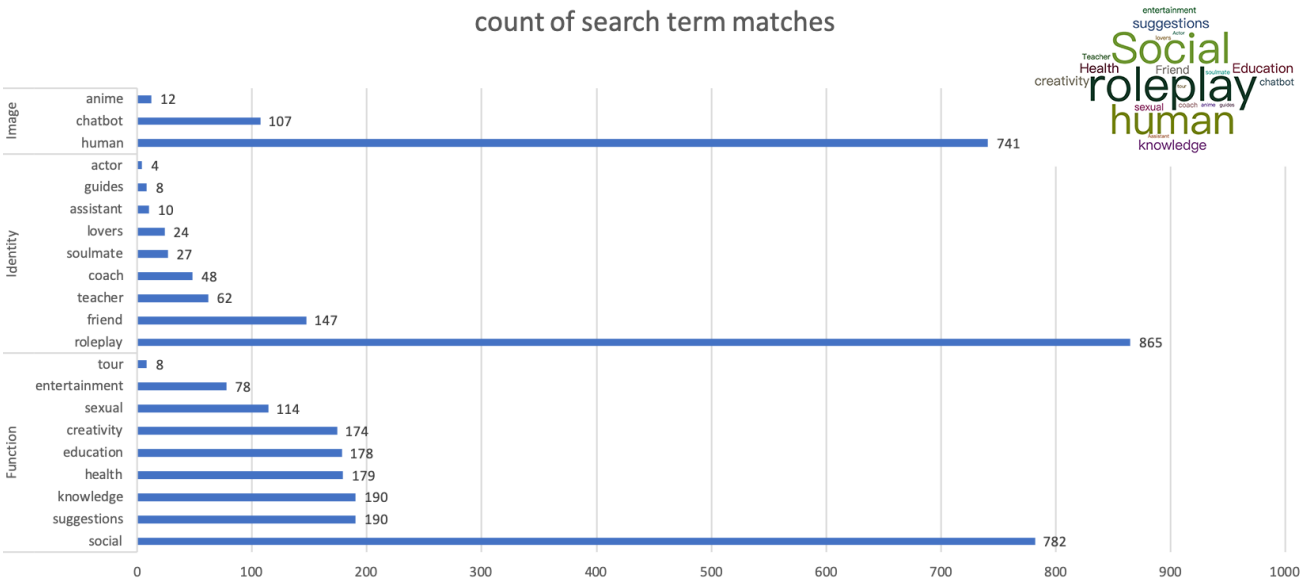

Fig.1: Keyword frequency statistics table and the word cloud

In summary, we can infer from the frequency of keywords that developers are more inclined to endow chatbots with social functions rather than simply serving as information retrieval and assistance tools. According to Chaves and Gerosa [9], people tend to expect social behaviour from each other in communication, so "social characteristics can facilitate interaction between humans and chatbots, mitigating user frustration and dissatisfaction." Developers are more inclined to enable machines to have some emotional communication with humans, appearing in the role of friends. However, this emotional communication typically remains superficial but not intimate. It's important to emphasize that "the existence of these discourses does not necessarily mean that the operations of these applications are consistent with these stated capabilities and values" [10]. It can only be said that from the developers' perspective, their aim is to design chatbots with such characters.

3.2 Co-occurrence Analysis of Keywords

Keyword co-occurrence refers to the phenomenon where two keywords appear simultaneously in the segment of text describing an application. A co-occurrence matrix (Table 2) quantitatively counts the number of times two certain keywords appear together. It can be used to figure out how developers typically combine keywords to create a comprehensive chatbot character. Based on Table 2, it can be further inferred that designers tend to incorporate "social" function character with the image character of "human". They also tend to imbue it with a "friend" identity. The high proportion of simultaneous occurrences between these three words suggests that they form the primary character combination for chatbots.

Table 2: Co-occurrence Matrix

	human	Social	knowledge	chatbot	creativity	Friend	Education	Health	roleplay	sexual	suggestions	coach	Teacher	entertainmer	lovers	soulmate	Assistant	guides	anime	Actor	tour
human	211	182	128	148	114	74	90	44	84	62	92	24	38	24	10	12	8	8	12	4	4
Social	182	192	108	136	108	54	90	44	66	56	80	24	38	24	10	12	8	8	12	4	4
knowledge	128	108	130	104	68	36	50	32	36	28	80	16	30	12	8	6	2	8	4	0	4
chatbot	148	136	104	196	82	36	74	46	48	48	86	24	36	26	6	6	8	8	4	4	0
creativity	114	108	68	82	122	34	52	26	46	46	66	14	26	22	2	6	6	8	0	4	0
Friend	74	54	36	36	34	81	38	2	38	20	30	2	10	0	2	6	2	0	8	4	4
Education	90	90	50	74	52	38	90	22	18	20	46	16	16	6	0	0	8	0	8	4	0
Health	44	44	32	46	26	2	22	68	14	32	30	16	16	10	6	6	2	0	0	0	0
roleplay	84	66	36	48	46	38	18	14	100	34	22	0	0	16	10	12	0	0	12	4	4
sexual	62	56	28	48	46	20	20	32	34	78	32	6	22	14	8	12	0	0	0	0	0
suggestions	92	80	80	86	66	30	46	30	22	32	98	14	26	18	2	0	0	8	0	0	0
coach	24	24	16	24	14	2	16	16	0	6	14	24	10	8	0	0	2	8	0	0	0
Teacher	38	38	30	36	26	10	16	16	0	22	26	10	38	8	0	0	2	8	0	0	0
entertainmer	24	24	12	26	22	0	6	10	16	14	18	8	8	68	0	0	0	8	0	0	0
lovers	10	10	8	6	2	2	0	6	10	8	2	0	0	0	10	6	0	0	0	0	0
soulmate	12	12	6	6	6	6	0	6	12	12	0	0	0	0	6	15	0	0	0	0	0
Assistant	8	8	2	8	6	2	8	2	0	0	0	2	2	0	0	0	8	0	0	0	0
guides	8	8	8	8	8	0	0	0	0	0	8	8	8	8	0	0	0	8	0	0	0
anime	12	12	4	4	0	8	8	0	12	0	0	0	0	0	0	0	0	0	12	0	0
Actor	4	4	0	4	4	4	4	0	4	0	0	0	0	0	0	0	0	0	0	4	0
tour	4	4	4	0	0	4	0	0	4	0	0	0	0	0	0	0	0	0	0	0	4

3.3. Degree Centrality Analysis in Social Networks

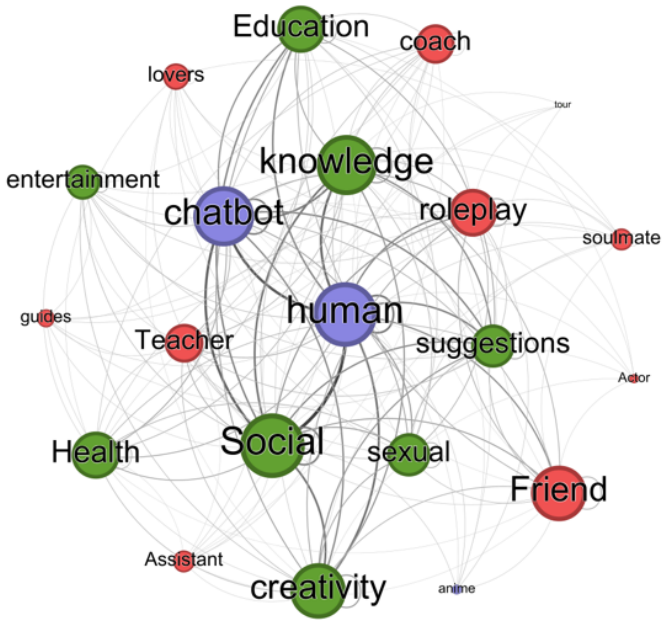

Fig.2: Co-occurrence network of keywords (plotted by Gephi)

This is a keyword co-occurrence network (Figure 2) generated by Gephi based on the co-occurrence matrix of words. Different colours represent different character keyword categories. As mentioned earlier, the colour intensity of edges is divided according to the weight of co-occurrence words. In other words, darker colours indicate a higher frequency of simultaneous mentions of two keywords, serving as a visualization of the keyword co-occurrence analysis in Section 3.2. Through this social network, we can also obtain new data: the size of each node represents the node's degree, where larger nodes indicate higher degree centrality in the network, suggesting that the associated character keyword is commonly found in various character combinations, which implies that the character keyword has a broader application scope.

From Figure 2, it can be observed that "human" serves as a significant connecting point in various subnetworks, indicating a divergent trend of the network with "human" at its core. Additionally, "social" and "friend" also exhibit relatively high centrality within the network, suggesting their frequent occurrence in the app interface page of Google Play and extensive coverage. This further validates the previous analysis of keyword co-occurrence, indicating that these three are the primary character combinations for chatbots. Although the word frequencies of "Knowledge" and "chatbot" are not high, they demonstrate strong centrality, which suggests that in actual developer operations, these two keywords can be combined with various character keywords from different categories.

3.4 Application Icons: Visual Presentation of Images

Application icons often relate to the character of the chatbot app. (Figure 3) Apps with icon designs similar to Chat GPT with simplistic designs tend to have assistant functions. Some apps prefer to showcase their social functions directly with chat boxes in their icon, or they adopt the appearance of cute robots to increase appeal while providing social or knowledge-related services. Meanwhile, some app icons may have a female-oriented design, featuring real or anime-style young women with vibrant hair and attractive faces. At times, there may even be subtle sexual connotations, such as biting an apple or sporting devilish red horns. These apps typically define themselves with a "friend" character, providing emotional

companionship. Combining this with the above, the icon design reflects the visually sexualization of companion-type characters, as well as the construction them of an attractive image. While the professional-type characters are combined with de-sexualized icons.

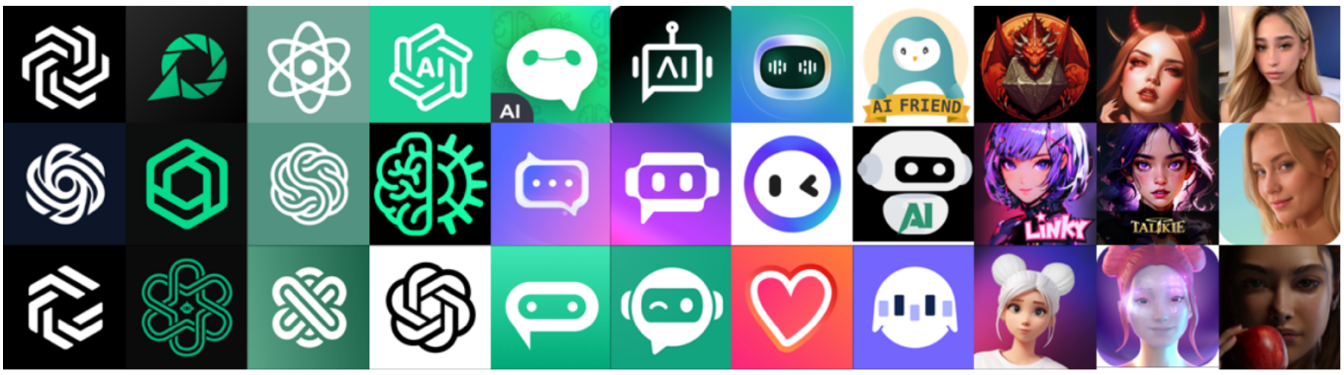

Fig.3: Chatbot Application Icons cluster

4. Discussion

While these patterns reveal common representational strategies, it is important to keep in mind that this study focuses solely on developer-side materials, rather than actual user interpretations or usage outcomes. This study uncovered an interesting phenomenon. Although many software programs fulfil traditional expectations of visually sexualization of companion-type characters, developers seldom use keywords related to lovers (such as boyfriend/girlfriend) in describing the app content, but only create an identity of "hot friends." I believe the reason for this might lie in the anthropomorphism barriers of chatbots. While it can satisfy a certain level of emotional communication expectations, there remain technical barriers to constructing more intimate human-machine interaction relationships. This is because people's perceptions are influenced by the performance and emotional capabilities of artificial intelligence, as well as regulated by human trust tendencies [11].

From a social constructivist perspective, applications are products of human decisions, and supported by implicit assumptions, norms, and discourses that circulate within the social and cultural contexts [12]. The underlying logic behind the construction of such applications is that chatbot app developers will build an ideal user from their own perspective, whom aligns with academic research descriptions of users. The ideal users, as mentioned, will make more favourable evaluations of characters that are more anthropomorphic according to social norm, more familiar to themselves, and ideally the same with their expectations and perceptions of the actual AI character. Some of the apps that mentioned in this paper that copied ChatGPT’s icon, are also designed to create this feeling of familiarity for ideal users. In the introduction, we mentioned da Costa's study on the gender bias in character construction [6], which might be because developers tend to connect with users in a specific way based on social conventions they are familiar with. In this process, similarity attributes are not only used for character construction as proposed by [5], but also for the repetition of social stereotypes. By approximating the stereotype of the ideal user, developers further construct the "human likeness" of chatbots [3].

We need to critically examine this underlying design logic of pursuing similarity from a perspective of sociomateriality. Technology is not only a product of social life but also shapes and influences various aspects of society, culture, politics, and economy. We should be wary of the biases of developers and the ideology of the ideal user they construct, which can also create positive feedbacks into social stereotypes, and the existing social biases, such as the image of young women not being associated with knowledge, will be further solidified.

5. Conclusion

In summary, this paper provides an empirical analysis of how Chatbots convey their characters to users. Through the analysis of the app collection, I have identified the following key findings: Developers are most inclined to give chatbot social functions rather than letting them serve as information query and assistive tools. "Human" “social” and “friend” are

the most common character combinations for Chatbots. The underlying logic of character construction is that developers tend to design characters from the perspective of the ideal user, within the realm of what the ideal user imagines, in order to enhance the "similarity" between the chatbot app and the users. This paper critically examines the underlying logic of character construction that pursues “similarity”, arguing that the blind pursuit of "similarity" can reinforce stereotypical impressions of society.

It is important to note that this study focuses on how chatbot characters are constructed and represented by developers, rather than how these characters are perceived or interpreted by users. Future research could incorporate user studies or in-app behavior analysis to explore how these representations are actually received in practice. Moreover, it only touches on English query results of Chatbot apps and only collected the apps on Google Play. The database needs further expansion, and we can conduct more comprehensive research on this phenomenon from the critical perspective. This study contributes to the design research community by foregrounding the socio-cultural logics embedded in chatbot character construction and proposing a critical lens to examine emotional labor and stereotype reproduction in human–AI interaction.

References

[1] H. Bansal and R. Khan, “A review paper on human computer interaction,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 8, no. 4, pp. 53–56, 2018.

[2] A. Khanna, B. Pandey, K. Vashishta, K. Kalia, B. Pradeepkumar, and T. Das, “A study of today’s AI through chatbots and rediscovery of machine intelligence,” *Int. J. U- E-Serv. Sci. Technol.*, vol. 8, no. 7, pp. 277–284, 2015.

[3] E. Go and S. S. Sundar, “Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions,” *Comput. Hum. Behav.*, vol. 97, pp. 304–316, 2019.

[4] J. Lee and P. J. Oh, “Effects of the use of high-fidelity human simulation in nursing education: A meta-analysis,” *J. Nurs. Educ.*, vol. 54, no. 9, pp. 501–507, 2015.

[5] C. H. Ho and K. C. Hou, “Exploring the attractive factors of app icons,” *KSII Trans. Internet Inf. Syst.*, vol. 9, no. 6, pp. 2251–2270, 2015.

[6] P. Costa, “Conversing with personal digital assistants: On gender and artificial intelligence,” *J. Sci. Technol. Arts*, vol. 10, no. 3, pp. 59–72, 2018.

[7] M. Dieter, A. Helmond, N. Tkacz, F. van der Vlist, and E. Weltevrede, “Pandemic platform governance: Mapping the global ecosystem of COVID-19 response apps,” *Internet Policy Rev.*, vol. 10, no. 3, pp. 1–28, 2021.

[8] M. Sedighi, “Application of word co-occurrence analysis method in mapping of the scientific fields (case study: the field of Informetrics),” *Library Rev.*, vol. 65, no. 1/2, pp. 52–64, 2016.

[9] A. P. Chaves and M. A. Gerosa, “How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design,” *Int. J. Hum.–Comput. Interact.*, vol. 37, no. 8, pp. 729–758, 2021.

[10] A. Kuntsman and E. Miyake, “The paradox and continuum of digital disengagement: Denaturalising digital sociality and technological connectivity,” *Media, Culture & Society*, vol. 41, no. 4, pp. 479–497, 2019.

[11] X. Song, B. Xu, and Z. Zhao, “Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants,” *Inf. Manage.*, vol. 59, no. 2, p. 103595, 2022.

[12] L. Jegatha Deborah, R. Baskaran, and A. Kannan, “Learning styles assessment and theoretical origin in an E-learning scenario: a survey,” *Artif. Intell. Rev.*, vol. 42, pp. 801–819, 2014.