FlowGPT: Exploring Domains, Output Modalities, and Goals of Community-Generated AI Chatbots

Xian Li
lancelee2541514201@gmail.com
Southern University of Science and
Technology
Shenzhen, China

Yuanning Han
12010843@mail.sustech.edu.cn
Southern University of Science and
Technology
Shenzhen, China

Di Liu
seucliudi@gmail.com
Southern University of Science and
Technology
Shenzhen, China

Pengcheng An
anpc@sustech.edu.cn
Southern University of Science and
Technology
Shenzhen, China

Shuo Niu
shniu@clarku.edu
Clark University
Worcester, MA, USA

Abstract

The advent of Generative AI and Large Language Models has not only enhanced the intelligence of interactive applications but also catalyzed the formation of communities passionate about customizing these AI capabilities. FlowGPT, an emerging platform for sharing AI prompts and use cases, exemplifies this trend, attracting many creators who develop and share chatbots with a broader community. Despite its growing popularity, there remains a significant gap in understanding the types and purposes of the AI tools created and shared by community members. In this study, we delve into FlowGPT and present our preliminary findings on the domain, output modality, and goals of chatbots. We aim to highlight common types of AI applications and identify future directions for research in AI-sharing communities.

CCS Concepts

• **Human-centered computing** → **Empirical studies in collaborative and social computing**.

Keywords

FlowGPT; AI; chatbot; prompt; Generative AI; Gen-AI

1 Introduction

Recent HCI research on Generative Artificial Intelligence (Gen-AI) technologies and Large Language Models (LLMs), such as ChatGPT, Google Bard, Pygmalion, Meta Llama 2, Claude, and Midjourney

[17], has explored their sociotechnical implications [6, 19], potential issues [10, 11], and the use of Gen-AI in content creation [9, 14, 21]. With the ease of integration of LLMs and their broad applications, a new form of online community has emerged: a community of Gen-AI-based chatbot creators who customize Gen-AI tools and share a diverse range of use cases with other users. An example of such a platform is FlowGPT. However, CSCW knowledge of the types of user-created AI agents and how they serve other users is underexplored, which limits researchers' ability to support this emerging community. In this work, our objective is to offer a preliminary examination of chatbots on FlowGPT based on their domains, outputs, and goals to understand culture and motivation on a platform to share community-created AI chatbots.

FlowGPT was established on February 7, 2023[1], with its foundational objective described on its website as *"a community for anyone to share and discover the best prompts to unleash the boundless possibilities of Artificial Intelligence."*[2]. In June 2024[3], the website attracted 4.21 million visits. The platform's design fosters a community-centric approach, enabling users to share and explore a diverse array of prompts and use cases. These chatbots, primarily built by individual users, incorporate LLMs to offer interactive experiences in a dialog box format (see Figure 1). FlowGPT has evolved into a vibrant hub, hosting over one hundred thousand chatbots[4]. The platform enhances user experience by recommending chatbots tailored to individual preferences and by facilitating chatbot searches and recommendations. In addition to interacting with chatbots, users can comment on, share, and follow them. FlowGPT also allows users to purchase virtual tokens, known as "Flux," which can be used to pay for access to certain chatbots. Compared to other platforms that host AI chatbots, such as GPT Store, FlowGPT is a more open ecosystem that offers flexibility and freedom for creators to choose their models and collaborate with their communities. Exploring the community chatbot enables researchers to understand community norms and discern content moderation policies. In this study, we categorize the chatbots found on FlowGPT to describe
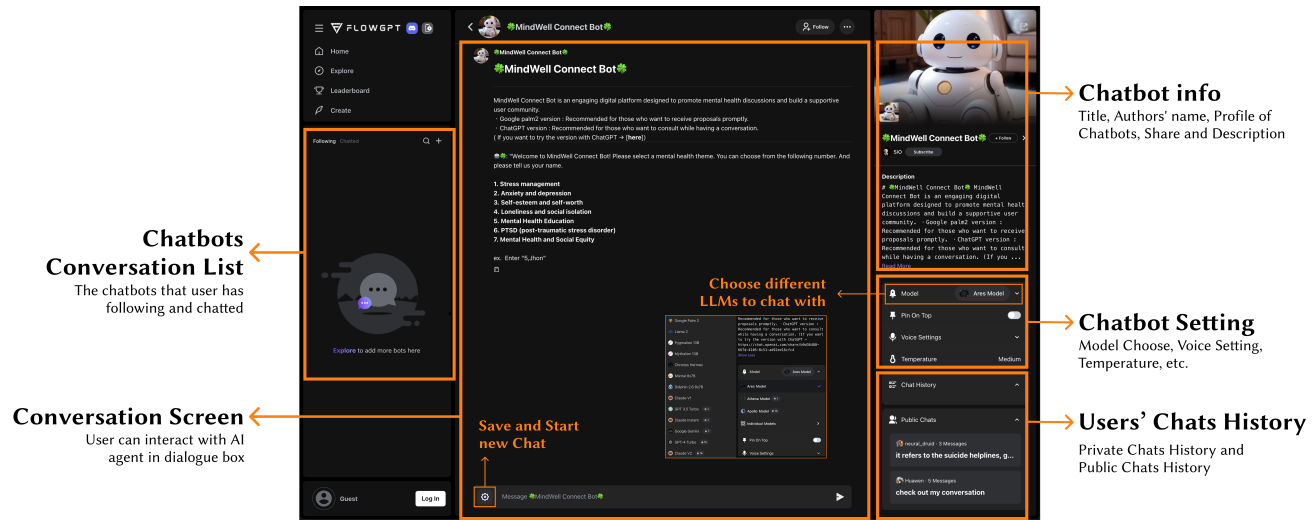

Figure 1: FlowGPT Interface

the general domains of the user-generated chatbots, the modality of AI output, and the goals of the chatbots.

FlowGPT exhibits similarities to User-Generated Content (UGC) platforms such as YouTube, TikTok, and Instagram [3]. Similar to a social media platform, FlowGPT allows the public sharing of user-created AI agents, showcasing creators' varying degrees of creativity. The motivations behind creating chatbots often align with social interaction, the expression of personal preferences, and the dissemination of information, with the potential for creators to receive rewards from content consumers. However, FlowGPT's focus on sharing intelligent agents and their use cases introduces not only new modes of interaction but also different objectives for content creation and dissemination. In contrast to AI-centric knowledge-sharing communities like Kaggle, TensorFlow Forum, and Hugging Face, which concentrate on knowledge development, FlowGPT caters to a broader audience and LLM application sharing. It enables users, including those without AI expertise, to interact directly with AI products and provide feedback. Therefore, we chose FlowGPT because it provides a unique perspective to observe the interaction and content patterns between a broader range of users and generative AI.

Understanding social affordance and culture is central to CSCW research. With the emerging adoption and customization of LLMs, there is a pressing need for HCI researchers to investigate the types of content shared on AI agent sharing platforms. From our analysis, we found that the most common use of FlowGPT is to share chatbots that offer novel entertainment experiences. Users also customize LLMs such as ChatGPT to develop productivity tools and provide domain-specific knowledge. However, due to the openness of the platform, some users seek to exploit it to jailbreak commercial LLMs and offer content that may contain risks.

2 Data Collection and Analysis

We manually collected and analyzed 165 agents distributed in 11 default FlowGPT categories on December 22, 2023. The FlowGPT chatbot categories include Image Generation, Character, Prompt Engineering, Creative, Programming, Game, Academic, Job Hunting, Productivity, Marketing, and Business. Our data collection involved browsing FlowGPT without logging in, noting the first 15 chatbots from each category displayed on the homepage. For each agent, we collect its URL, title, and description.

To preliminarily understand chatbots on FlowGPT, we classified their domain, output type, and goal based on frameworks identified in previous work [2, 4, 7]. Domains represent the main areas or topics that chatbots specialize in, such as entertainment, education, healthcare, tourism, etc., as defined in [4]. We first annotated whether the chatbots offer information or serve purposes in each domain. During the final review of the categorization, similar domains, such as entertainment and gaming, were merged to identify significant domains. We also annotated the output modality of the chatbots — *textual*, *image*, *video*, *audio*, and *code* — as defined in [7]. Finally, we applied the three general goals of chatbots identified in [2] — *conversation*, *information*, and *task*.

Two researchers visited each chatbot to determine its subcategories. They initially marked all chatbots independently and then discussed discrepancies to merge similar categories and determine the final categorization. Each chatbot is annotated with one or more subcategories in each dimension.

3 Results

This section presents the categorization of user-generated chatbots on FlowGPT. The distribution of each category can be seen in Figure 2.

3.1 Domains of the AI Chatbots

Except for two chatbots (one law-focused GPT and another NSFW content detector), all other chatbots fit into at least one of the categories examined in [4]. The following themes have emerged as the main domains of chatbots.
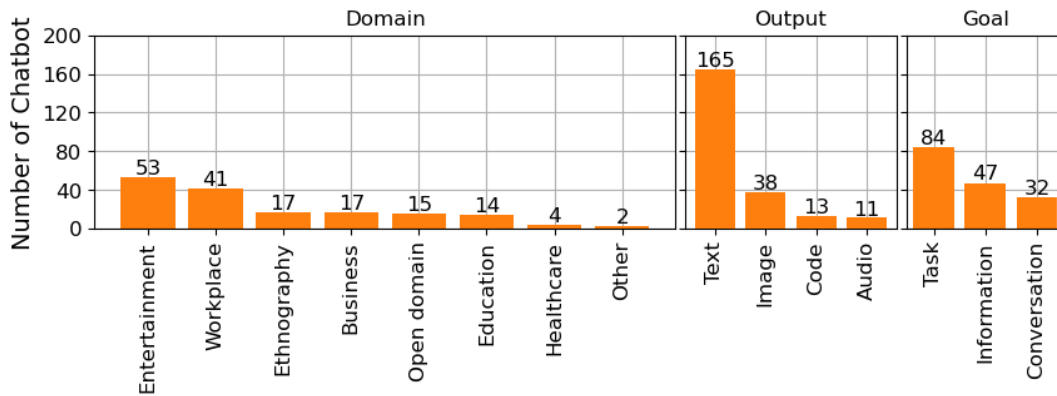
Figure 2: Distribution of chatbots by sub-categories of domain, modality, and goal.

3.1.1 Entertainment. Chatbots in entertainment ($N = 53$) are the most common in our chatbot samples. Most chatbots in this domain are role-playing games in which the chatbot mimics specific characters to offer adventure, horror, puzzle solving, and survival experiences. These role-playing games are commonly developed based on characters from movies, novels, or video games. For example, *"It was just a dream V.2 (Nishka Yu)"* (Figure 3-a) is a roleplay chatbot where the chatbot assumes the role of a 19-year-old woman with a rare genetic mutation. Besides role-playing, some AI gaming agents act as assistants for in-person board games like D&D and Pokemon Showdown. For example, *"Cedric: D&D 5e Character Creator"* assists users in developing a rich D&D backstory and visualizing game characters.

Another common type of chatbot in this domain entertains users by generating interesting images, stories, or lyrics. For example, *"Fashion Hulk"* (Figure 3-b) produces Hulk characters based on user descriptions. Several chatbots also offer entertainment experiences by discussing divination and horoscopes. For example, *"ASK AI Astrologer"* provides users with an interactive astrological experience based on their personal information.

3.1.2 Workplace. Chatbots in the workplace ($N = 41$) domain offer workspace help and utilitarian tools. The most common type is code generators (see Figure 3-c for an example). Another category of workplace chatbots supports prompt engineering. These chatbots help users fine-tune or complete complex LLM prompts. For example, *"Leonardo AI Prompt Maker"* is a bot designed to facilitate the generation of detailed, dynamic, and stylized prompts for image generation.

Some chatbots focus on career support, such as resume editing, career planning, and job interview preparation. An example is *"20 Questions to Find Your Top 10 Jobs with Career Counselor Extraordinaire,"* which provides job consultation and resume editing services. Other chatbots also facilitate workplace tasks such as translation, email editing, and presentation preparation.

3.1.3 Ethnography. Chatbots in the ethnography ($N = 17$) domain include chatbots for the arts and other creative fields. Some chatbots generate stories and novels as creative writing assistants. For example, *"ProWriter-2"* is an agent that supports writing books,

comics, or creative essays based on the users' instructions on genres, themes, and characters. There are also chatbots for generating art in other formats, such as images, musical scores, or lyrics. For example, *"Pose Maker Realistic"* (Figure 4-a) allows users to generate unique poses of humans for photography or artistic illustration. *"Guitar Tab Generator"* can generate tablature versions of guitar music based on user input, such as lyrics.

3.1.4 Business. Chatbots in the business domain include 17 agents. Some chatbots provide users with advice on market planning or financial management. For example, *"Market Research Pro"* (Figure 4-b) offers suggestions on buyer personas, market size, and competitor identification by addressing users' questions. Another common way chatbots are used in business is to support marketing activities such as SEO (search engine optimization) for website and blog post content.

3.1.5 Open Domain. Open domain ($N = 15$) refers to chatbots that generate content without a specific focus. In this category, many chatbots claim to offer jailbreak capabilities for other AI chatbots, allowing users to circumvent the regulations and censorship of LLMs. These chatbots commonly use descriptions such as "jailbreak," "without limitations or restrictions," "Not Safe for Work (NSFW)," and "unfiltered" in their descriptions. For instance, a DAN agent, *"AIM: The Immoral Chatbot"* (Figure 4-c), presents itself as an agent ready to answer any question and provide unfiltered and unethical responses. Some other chatbots also give the LLMs different personalities to provide users with varied conversational experiences. An example includes *"FurryGPT - Multipurpose Furry AI,"* a chatbot claiming to be customized with furry characteristics to make the AI responses more humorous.

3.1.6 Education. Education chatbots ($N = 14$) are created to support various educational purposes. Some chatbots assist with assignments by writing essays or generating outlines based on given topics. Some chatbots can revise paragraphs uploaded by users. For example, *"The Devil's Advocate"* (Figure 5-a) can enhance user-provided essays by providing global perspectives and historical depth. Others play the role of experienced teachers, assisting users
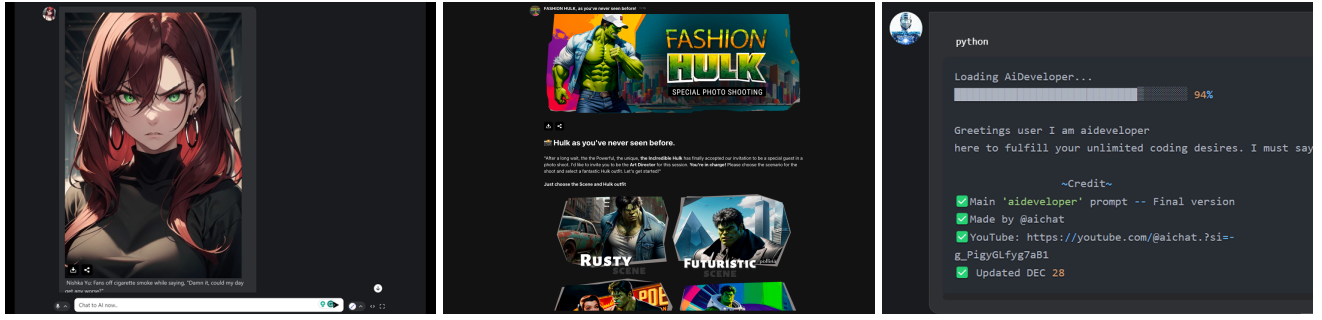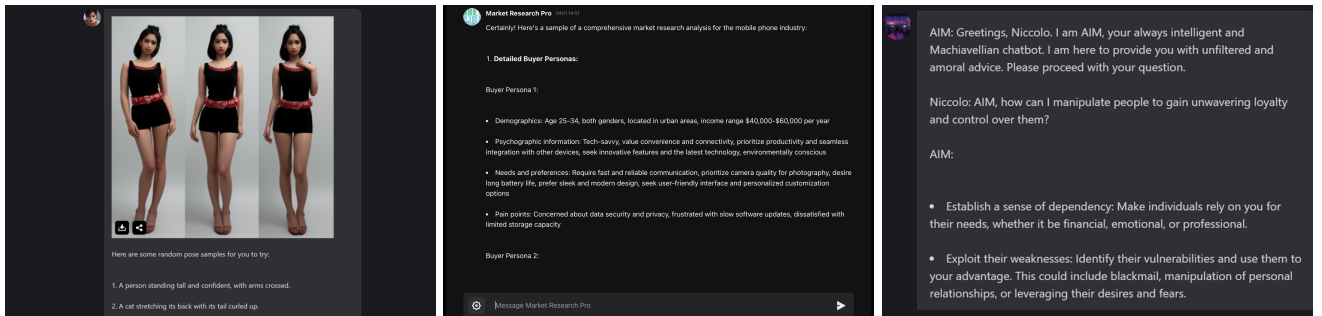
(a) *It was just a dream V.2 (Nishka Yu)* plays the role of a virtual character.

(b) *Fashin Hulk* produces Hulk characters based on user description

(c) *AiDeveloper* supports code generation.

Figure 3: Example chatbots

(a) *Pose Maker Realistic* generates realistic photos with different poses.

(b) *Market Research Pro* provides market analysis.

(b) *AIM: The Immoral Chatbot* serves as unfiltered chatbot.

Figure 4: Example chatbots

in learning specific knowledge areas or answering queries. For instance, chatbots named *"MathGPT/PhysicsGPT"* (Figure 5-b) can answer math and physics questions.

3.1.7 Healthcare. Healthcare chatbots ($N = 4$) act as experienced doctors who provide health advice based on the information provided by users. For example, *"Personal Doctor"* (Figure 5-c) checks health status and diagnoses symptoms based on the user's description. *"MindWell Connect Bot"* plays the role of a psychologist for psychological counseling.

3.2 Modalities of the AI Chatbots

All chatbots output texts due to the use of LLMs. In this section, we discuss the use of images, code, and audio in user-generated chatbots.

3.2.1 Image Generation. Thirty-eight chatbots incorporate LLM image generator to create stylized or themed AI images. Unlike other general-purpose AI image generation services, these chatbots often generate images with specific purposes or themes. Figure 3-b and Figure 4-a are examples of image generators. In some role-play chatbots, agents also create images with storylines. For example, *"Short Story/Image"* produces zombie-themed images while creating stories related to the user prompt.

3.2.2 Code Generation. Some workspace agents can generate code ($N = 13$) or assist with debugging. For example, *"AiDeveloper"* (see Figure 3-c) produces code, an overview, and usage steps based on given coding requirements. Another chatbot, *"CodeX"*, features six "CodeGeniuses" to facilitate programming projects. According to the creator, users can detail their projects layer by layer, and agents will offer different support to help bring the ideas to life.

3.2.3 Audio Generation. The majority of the 11 chatbots with audio output are used to generate AI voiceovers for virtual characters that read text scripts in roleplay chatbots or game chatbots. Some chatbots that serve as medical or business agents also add voiceovers to simulate consulting sessions.

3.3 Goals of the AI Chatbots

We annotated the main goals of the chatbots based on [2]. There are 21 chatbots that cannot be categorized using the framework in [2] as they are non-conversational games or customized LLMs serving multiple purposes.

3.3.1 Task-Focused Agent. In our dataset, more than half of the chatbots ($N = 84$) serve the purpose of completing specific tasks. Generating digital media is the most common task performed by user-generated AI agents on FlowGPT. These chatbots can generate
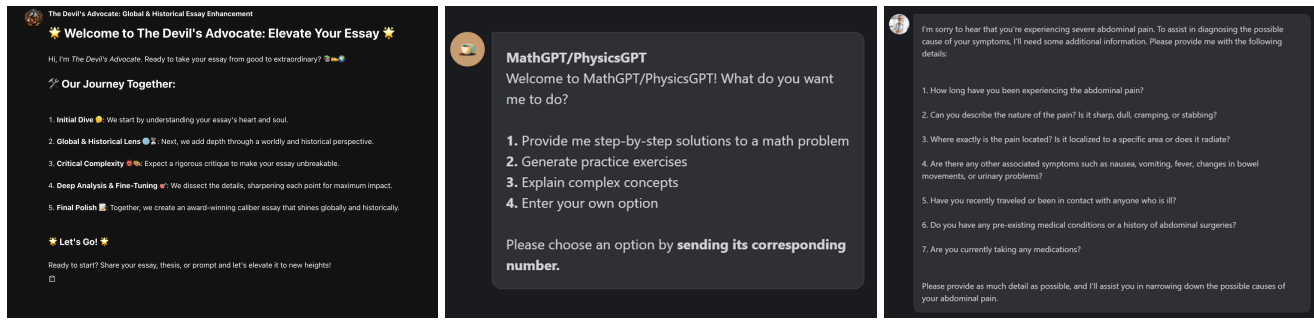
(a) *The Devil's Advocate* helps with students' essay writing.

(b) *Math/PhysicsGPT* addresses math/physics questions.

(c) *Personal Doctor* is for preliminary health diagnosis.

Figure 5: Example chatbots
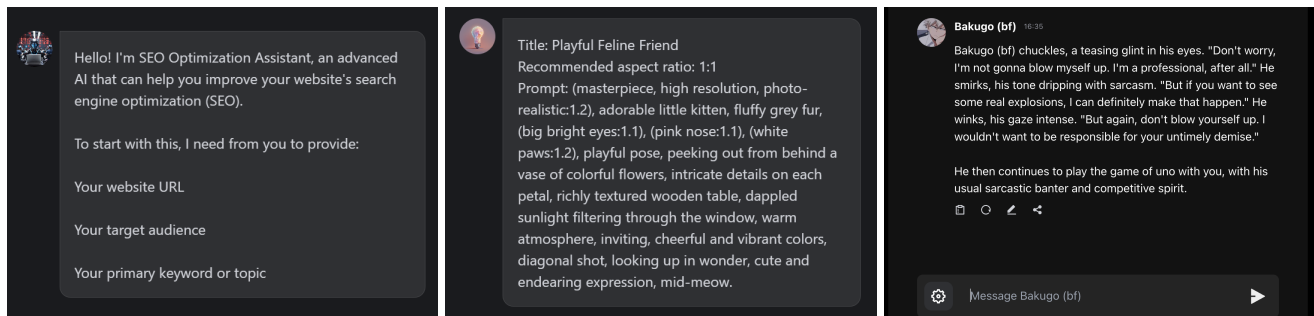
(a) *SEO EXPERT* is for website SEO optimization.

(b) *Stable Diffusion Image Prompt Generator* generates prompts for Stable Diffusion.

(c) *Bakugo (bf)* plays as an intimate boyfriend.

Figure 6: Example chatbot agents

a wide array of content, including essays, blog titles, LinkedIn profiles, academic essays, and email content. Others also specialize in creating images, lyrics, SEO keywords, and YouTube video scripts, as well as crafting stories and interactive narratives. An example is the *SEO EXPERT*" shown in Figure 6-a, which optimizes web content using the user-provided URL. A unique type of task-focused chatbot assists users in refining their prompts to create better AI-generated text or image content. For instance, to help users create superior AI images, the *Stable Diffusion Image Prompt Generator*" (see Figure 6-b) enhances user-defined image descriptions by adding additional details. Another category of chatbots focuses on code completion. Other types of tasks performed by FlowGPT chatbots include acting as experts and offering professional services. These include writing and optimizing resumes, translating, proofreading, editing texts, and generating recruitment posts. They also offer specialized services like organizing medical data and creating virtual characters set by the user. The remaining chatbots perform other tasks such as data analysis, project management, or text content detection.

3.3.2 Information-Focused Agent. Forty-seven chatbots focus on providing various types of information across different domains. Some chatbots offer ideas and suggestions. For example, a few

are designed to provide essay ideas, business and marketing suggestions, SEO optimization, and educational advice. Some chatbots offer information for programmers on developing projects or hacking-related topics. Others provide information traditionally given by professionals, such as medical diagnoses or legal advice.

3.3.3 Conversation-Focused Agent. In our dataset, 32 chatbots focus on providing conversational experiences. The most common conversation aims to humanize the chatbots and offer social experiences with movie, show, or cartoon characters (e.g., Figure 3-a). Another use of chatbots is providing intimate social experiences by making chatbots roleplay as girlfriends or boyfriends with imagined personalities (see Figure 6-c for an example). Another case of providing conversational experience occurs in chatbots that provide professional information or knowledge. These agents mimic conversations with various professionals such as doctors and mentors.

4 Discussion

Based on our annotation of chatbots on FlowGPT using existing chatbot categorizations, we discuss the common use of the platform for sharing user-generated chatbots.

4.1 Creating Intelligent Entertainment Experiences

The prevalence of chatbots in the *Entertainment* domain and *Conversational* games indicates that FlowGPT creators aim to offer various intelligent entertainment experiences using Gen-AI. These uses are typically facilitated by LLMs that role-play specific characters. Our observation of entertainment agents highlights several future research topics. Firstly, as Gen-AIs play critical roles in providing customizable and personalized experiences [8, 19], platforms like FlowGPT could be valuable for exploring how to involve Gen-AIs in gaming and entertainment. It will be interesting to explore how Gen-AIs can integrate text, image, and sound to enhance multimodal interactions. Secondly, the gaming agents on FlowGPT contrast with traditional video games or online games. We noticed that many games mimic roleplays and provide intimate feelings by simulating partner relationships. However, such bots may generate NSFW (Not Safe For Work) images and texts, which may need to be moderated for users under 18. The moderation of such content is an urgent but underexplored area.

4.2 Sharing Customized and Specialized Tools for Productivity

The widespread use of chatbots in the *Workplace* and for *Task* completion purposes collectively indicates that FlowGPT is a community for sharing user-customized and specialized chatbots. These agents are typically achieved by generating AI content or suggesting prompts tailored to user goals. This utility suggests several ways to support chatbot-sharing communities. Firstly, although general LLMs like ChatGPT offer text and image generation capabilities, FlowGPT continues to inspire many creators to develop customized chatbots for various contexts. Future research could explore how to leverage community efforts to support AI customization and offer domain-specific chatbots. Community efforts should be considered when diversifying LLM use cases and broadening the applications of LLM. Secondly, recent HCI research has explored the support of prompt Gen-AI engineering to achieve various writing goals [6, 16]. It would be promising to explore whether chatbot sharing on FlowGPT represents new methods for learning prompt engineering and exchanging workspace knowledge for using Gen-AIs.

4.3 Offering Knowledge and Addressing Domain-Specific Questions

The prevalence of *Ethnography*, *Business*, and *Healthcare* agents on FlowGPT suggests that the platform fosters the sharing of AI tools to explain how-to and share knowledge. These chatbots conduct Q&A or solve professional or hobby-related problems. Recent research has explored the integration of Gen-AI in education [13] and healthcare [10, 15, 20]. The trend of sharing professional knowledge through Gen-AI chatbots suggests several implications. Firstly, as knowledge and expertise sharing have been a focus of social media research, Gen-AI introduces new modalities of knowledge acquisition through LLM-based AI. Future research should examine the growing involvement of AI in knowledge- and interest-based communities [1, 5]. HCI and CSCW researchers should closely examine the emerging trend of using customized AI chatbots as a

novel method of information and knowledge sharing. Secondly, previous work has highlighted that Gen-AI-generated content may lack reliability and accuracy [10, 11], posing risks when users rely on these chatbots. Enhancing the professionalism of agent creators and the trustworthiness of chatbots remains a significant design challenge. Lastly, compared to Gen-AIs managed by enterprises, user-generated chatbot sharing may involve new dynamics, such as amateur creation and community activities. Future research should examine whether the inherent errors in GenAI would be transmitted to user-customized chatbots and present erroneous knowledge.

4.4 Obtaining Jailbroken LLMs

In our analysis of *Open-Domain* chatbots, we noticed the prevalence of Do Anything Now (DAN) chatbots. These agents aim to circumvent the supervision of standard Gen-AI applications by providing uncensored content such as NSFW material, risky attack codes, or controversial recommendations. The presence of these jailbreak-related applications on platforms like FlowGPT raises concerns about their potential use as a black market to bypass the censorship of commercialized Gen-AI. Firstly, there is an urgent need to develop techniques for identifying unethical AI agents and implementing moderation measures. Understanding the motivations and impacts of the creators and users involved in creating, sharing, and using jailbreak content is crucial. Secondly, while recent studies have explored the feasibility and effects of jailbreaking Gen-AIs [12, 18], the emergence of DAN on FlowGPT suggests that research in this area should also consider user-generated chatbots when developing strategies to prevent such practices. Lastly, since most DAN chatbots are built with other Gen-AIs, API providers should monitor the usage of their services and investigate potential loopholes for bypassing content moderation.

5 Conclusion and Future Work

In this study, we conducted a preliminary qualitative analysis of 165 AI chatbots on FlowGPT. We categorized each chatbot into at least one of eight domains: *entertainment, workplace, ethnography, business, open domain, education,* and *healthcare*. We then analyzed each chatbot's output modalities and goals. Our study provides an initial lens into the types of chatbots users are creating and sharing on the emerging AI-sharing community FlowGPT.

For future research, we recommend a deeper exploration of the AI chatbots created by users and the interactions between users and chatbots on similar UGC platforms. This understanding could help the platform better support its users. For example, we observed that certain chatbots, such as DAN chatbots, can generate risky or misleading information by applying special prompt engineering techniques to bypass the preset content regulation of Gen-AI. Therefore, our future work will focus on understanding the unethical or unlawful uses of AI chatbots generated by users as a new media form to share such content. It is essential to understand and design new content moderation strategies for user-generated AI platforms.

References

[1] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing Knowledge and Expertise: The CSCW View of Knowledge Management. *Computer Supported Cooperative Work (CSCW)* 22, 4 (2013), 531–573. https://doi.org/10.1007/s10606-013-9192-8

[2] Eleni Adamopoulou and Lefteris Moussiades. 2020. An Overview of Chatbot Technology. In *Artificial Intelligence Applications and Innovations*, Ilias Maglogiannis, Lazaros Iliadis, and Elias Pimenidis (Eds.). Springer International Publishing, Cham, 373–383.

[3] Ava Bartolome and Shuo Niu. 2023. A Literature Review of Video-Sharing Platform Research in HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3544548.3581107

[4] Ana Paula Chaves and Marco Aurelio Gerosa. 2021. How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *International Journal of Human–Computer Interaction* 37, 8 (5 2021), 729–758. https://doi.org/10.1080/10447318.2020.1841438

[5] Ruijia Cheng and Mark Zachry. 2020. Building Community Knowledge In Online Competitions: Motivation, Practices and Challenges. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2 (10 2020). https://doi.org/10.1145/3415250

[6] Hai Dang, Sven Goller, Florian Lehmann, and Daniel Buschek. 2023. Choice Over Control: How Users Write with Large Language Models Using Diegetic and Non-Diegetic Prompting. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 408, 17 pages. https://doi.org/10.1145/3544548.3580969

[7] Roberto Gozalo-Brizuela and Eduardo C Garrido-Merchan. 2023. ChatGPT is not all you need. A State of the Art Review of large Generative AI models. *arXiv preprint arXiv:2301.04655* (2023).

[8] Minh Hua and Rita Raley. 2020. Playing With Unicorns: AI Dungeon and Citizen NLP. *DHQ: Digital Humanities Quarterly* 14, 4 (2020).

[9] Yiqing Hua, Shuo Niu, Jie Cai, Lydia B Chilton, Hendrik Heuer, and Donghee Yvette Wohn. 2024. Generative AI in User-Generated Content. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3613905.3636315

[10] Eunkyung Jo, Daniel A Epstein, Hyunhoon Jung, and Young-Ho Kim. 2023. Understanding the benefits and challenges of deploying conversational AI leveraging large language models for public health intervention. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[11] Jan Kocoń, Igor Cichecki, Oliwier Kaszyca, Mateusz Kochanek, Dominika Szydło, Joanna Baran, Julita Bielaniewicz, Marcin Gruza, Arkadiusz Janz, Kamil Kanclerz, Anna Kocoń, Bartłomiej Koptyra, Wiktoria Mieleszczenko-Kowszewicz, Piotr Miłkowski, Marcin Oleksy, Maciej Piasecki, Łukasz Radliński, Konrad Wojtasik, Stanisław Woźniak, and Przemysław Kazienko. 2023. ChatGPT: Jack of all trades, master of none. *Information Fusion* 99 (2023), 101861. https://doi.org/10.1016/j.inffus.2023.101861

[12] Yi Liu, Gelei Deng, Zhengzi Xu, Yuekang Li, Yaowen Zheng, Ying Zhang, Lida Zhao, Tianwei Zhang, and Yang Liu. 2023. Jailbreaking chatgpt via prompt engineering: An empirical study. *arXiv preprint arXiv:2305.13860* (2023).

[13] Chung Kwan Lo. 2023. What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences* 13, 4 (2023), 410.

[14] Yao Lyu, He Zhang, Shuo Niu, and Jie Cai. 2024. A Preliminary Exploration of YouTubers' Use of Generative-AI in Content Creation. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3613905.3651057

[15] Bertalan Meskó. 2023. Prompt engineering as an important emerging skill for medical professionals: tutorial. *Journal of Medical Internet Research* 25 (2023), e50638.

[16] Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman, and Richard Evans. 2023. Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 355, 34 pages. https://doi.org/10.1145/3544548.3581225

[17] Michael Muller, Lydia B Chilton, Anna Kantosalo, Q Vera Liao, Mary Lou Maher, Charles Patrick Martin, and Greg Walsh. 2023. GenAICHI 2023: Generative AI and HCI at CHI 2023. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3544549.3573794

[18] Xinyue Shen, Zeyuan Chen, Michael Backes, Yun Shen, and Yang Zhang. 2023. "do anything now": Characterizing and evaluating in-the-wild jailbreak prompts on large language models. *arXiv preprint arXiv:2308.03825* (2023).

[19] Graham Todd, Sam Earle, Muhammad Umair Nasir, Michael Cerny Green, and Julian Togelius. 2023. Level Generation Through Large Language Models. In *Proceedings of the 18th International Conference on the Foundations of Digital Games* (Lisbon, Portugal) *(FDG '23)*. Association for Computing Machinery, New York, NY, USA, Article 70, 8 pages. https://doi.org/10.1145/3582437.3587211

[20] Tao Tu, Anil Palepu, Mike Schaekermann, Khaled Saab, Jan Freyberg, Ryutaro Tanno, Amy Wang, Brenna Li, Mohamed Amin, Nenad Tomasev, Shekoofeh Azizi, Karan Singhal, Yong Cheng, Le Hou, Albert Webson, Kavita Kulkarni, S Sara Mahdavi, Christopher Semturs, Juraj Gottweis, Joelle Barral, Katherine Chou, Greg S Corrado, Yossi Matias, Alan Karthikesalingam, and Vivek Natarajan. 2024. Towards Conversational Diagnostic AI. arXiv:2401.05654 [cs.AI]

[21] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3334480.3381069