



Generating personas using LLMs and assessing their viability

Andreas Schuller
andreas.schuller@iao.fraunhofer.de
Human-Technology Interaction
Fraunhofer Institute for Industrial
Engineering IAO
Stuttgart, Germany

Theresa Maria Probst
theresa.maria.probst@iao.fraunhofer.de
Human-Technology Interaction
Fraunhofer Institute for Industrial
Engineering IAO
Stuttgart, Germany

Doris Janssen
doris.janssen@iao.fraunhofer.de
Human-Technology Interaction
Fraunhofer Institute for Industrial
Engineering IAO
Stuttgart, Germany

Michael Schmidt
ms@code2design.de
code2design
Stuttgart, Germany

Julian Blumenröther
julian.blumenroether@iao.fraunhofer.de
Institute of Human Factors and
Technology Management
University of Stuttgart
Stuttgart, Germany

Chandan Kumar
chandan.kumar@iao.fraunhofer.de
Human-Technology Interaction
Fraunhofer Institute for Industrial
Engineering IAO
Stuttgart, Germany

ABSTRACT

User personas, reflecting human characteristics, play a crucial role in human-centered design, contributing significantly to ideation and product design processes. However, expressing a diverse range of product-related human characterizations poses a challenging and time-consuming task for UX experts. This paper explores the utilization of Large Language Models (LLMs) to streamline the generation of personas, thereby enhancing the efficiency of UX researchers and providing inspiration for stakeholder discussions. Towards this objective, we devised strategic prompts and guidelines involving stakeholders and potential product features, resulting in the creation of candidate user personas. These personas were then compared with those crafted by human experts in a remote study involving 11 participants assessing 16 personas each. The analysis revealed that LLM-generated personas were indistinguishable from human-written personas, demonstrating similar quality and acceptance.

CCS CONCEPTS

• **Human-centered computing** → **User centered design**; *User studies*.

KEYWORDS

personas, automated persona generation, user centered design, innovation processes, generative AI

ACM Reference Format:

Andreas Schuller, Doris Janssen, Julian Blumenröther, Theresa Maria Probst, Michael Schmidt, and Chandan Kumar. 2024. Generating personas using LLMs and assessing their viability. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3613905.3650860>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
CHI EA '24, May 11–16, 2024, Honolulu, HI, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0331-7/24/05
<https://doi.org/10.1145/3613905.3650860>

1 INTRODUCTION

Human-Centered Design (HCD) plays a crucial role in shaping product design, development, and advancements in human-computer interaction research [11]. An intrinsic element of the HCD process is empathy, requiring UX experts and designers to connect with users, their context, and their aspirations. One widely adopted technique to foster this understanding is the creation of personas [6, 16]. These personas can represent user groups, providing an impression of their personalities to unveil insights into their needs and desires [18]. However, the traditional manual approach to persona creation poses challenges, especially when a multitude of personas are needed for focus group discussions and ideation workshops. In this regard, various tools such as User Persona¹, PersonAI - User Persona Generator², and QoqoAI³, facilitate persona creation with detailed descriptions based on demographics and scenarios. However, these tools often provide a standardized format and lack adaptation options for factors like use case, output format, stakeholder role, or age group. Hence, there's a growing need for more advanced artificial intelligence (AI) technologies to support HCD specifically in persona creation [14, 21].

In the realm of AI technologies for automation and generative tasks, large language models (LLMs) like GPT-3.5⁴, and GPT-4 have gained widespread popularity [3, 5]. LLMs are now integral in various environments, including information retrieval [26], complex writing [4], creativity and innovation [2, 25]. The evolution has also fueled extensive research in unlocking the potential of LLMs for human-centered development [21]. LLMs can be particularly powerful for reshaping HCD methodologies, as they are trained on large user data and encode human experiences. Furthermore, readily available LLMs like ChatGPT offer this advantage without necessitating the creation of specific models, potentially saving time and effort for UX researchers and designers. Therefore, recent research has examined the viability of LLMs in supporting HCD processes, such as requirement analysis, market research, and user data creation [12, 14, 19]. Notably, the work by Hämäläinen et al. [12]

¹<https://userpersona.dev/>

²<https://www.figma.com/community/plugin/1287786847239653675/personai-user-persona-generator>

³<https://www.qoqo.ai>

⁴<https://openai.com/chatgpt>

explores the use of GPT-3 Playground for rapid synthetic user research data generation (generating responses to questionnaires), revealing results similar to real user responses. Similarly, several recent studies have examined the extent to which AI language models can simulate human behavior and capabilities [1, 7–9, 20]. In alignment with this, and recognizing the crucial role of persona creation in HCD, investigating the feasibility and acceptance of LLMs becomes increasingly interesting, forming the central focus of this paper.

Towards this focus, a limited number of works have looked into the utilization of LLMs for persona generation. Notable studies, such as Goel et al.'s [10] and Kocaballi's [14], leverage ChatGPT for interactive persona creation, with a specific focus on novice designers, human-AI collaboration, and direct dialog prompting for specific cases. However, the effective implementation of an automated persona generation process requires flexibility in expected results to accommodate diverse project settings and an easily navigable approach that doesn't require extensive knowledge of prompting techniques. Additionally, it is necessary to validate the LLM-generated personas with human experts, to comprehend their viability in HCD and innovation processes, which is an aspect that has not been covered in previous studies. Thus, the two primary objectives of our work are:

- (1) To examine the characteristics of effective prompting strategies and identify techniques that yield promising results for persona generation.
- (2) To assess the perceived quality and acceptance of AI-generated personas in contrast to those created by humans.

Towards the first objective, we present the details of our proposed AI-assisted persona creation strategy and examples in Section 2. Addressing objective two, we describe the evaluation methodology involving 11 UX experts and designers who assess 16 personas, aiming to distinguish between ChatGPT and human-written personas in Section 3. Finally, in section 4, the results are discussed, offering valuable insights into the perceived quality and authenticity of AI-generated personas, and their future potential.

2 PROMPTING STRATEGIES FOR AI-ASSISTED PERSONA CREATION

The issue of automated persona generation arose from the requirements of the design agency⁵. The agency conducts user-centered innovation workshops to drive long-term brand identity enhancement for its clients. As part of their innovation process, they create personas to assess the success factors of new product concepts from the perspectives of their target user groups. Their approach to pre-create the personas in back office for using them in innovation workshops, is only one means to use personas. More often even, personas are created in a collaborative workshop. However, this would be hard to substitute by an AI (as the creation process leads to better understanding of the personas in all participating people). So for this paper, we focus on the scenarios of desktop-creation of personas.

Seeking to streamline their workshops, they aimed to automate certain steps, leading to our collaboration. Together, we identified

persona creation as a step in the innovation process that could benefit from recent advancements in generative AI technology. To address their specific needs, we conducted a case study using ChatGPT for persona generation. Given the agency's diverse client base, the personas created are highly customized for individual use cases within their innovation workshops. As an example, the most recent scenario involved a supplier of aesthetic and functional door and window fittings. For their innovation workshop, they needed personas that target the user groups of architects, investors, fabricators, and end-users, each having specific output formats. To maintain individuality, we utilized different prompting strategies to fit the results to any scenario or use case. By editing the variable parts of our final prompts, the resulting personas can be tailored to any use case or scenario. Our approach for the creation of prompts to generate personas within ChatGPT is discussed in the following.

2.1 Prompt formulation Experiments

We performed comparative analyses on the generation capabilities of GPT 3.5 and GPT 4.0, revealing negligible disparities in output quality. Consequently, we elected to retain GPT 3.5 along with its default configurations. Our initial experimentation entailed employing rudimentary free text prompts for persona generation, resulting in outputs lacking requisite structural and stylistic coherence. Subsequent iterative refinement facilitated the formulation of effective strategies for persona generation, which are discussed in the following.

2.1.1 General Guidelines. Since the input prompt for ChatGPT is free text, it leaves a lot of room for creativity and different approaches for inputs to arrive at the desired results. We have combined various strategies and principles in the development of our prompt for persona creation. Some general principles we used were inspired by Bsharat et al. [4].

Idea: Our prompts follow proven prompting principles. These include: (1) Providing context; (2) Specification of requirements; (3) Use of delimiters;

Application: (1) To provide ChatGPT with background information, our prompt contains a company description and the context of use. (2) To be precise in our statements, we clearly specify how many personas, which stakeholders and attributes are required, and what format the output should have (e.g. tables). We provide information on how long a paragraph should be and from which perspective it should be formulated. We require the personas to be consistent, so we emphasize this point. (3) Our prompt works for various persona formats by replacing specific information with variables. The variables have to be defined depending on the existing scenario and persona format. To keep the different parts distinguished, we use delimiters.

Strategy Examples:

- (1) *Company description* = "" "" "" ""
- (2) *Make sure that the personas are consistent; The output is in the form of tables.*
- (3) *Prompt* = "" "" "You are acting on behalf of company XYZ. "" {Company description} "" "" ""

Thus we have already restricted (parts of) ChatGPT's output and so we have already approximated the structure of the persona to

⁵<https://www.code2design.de>

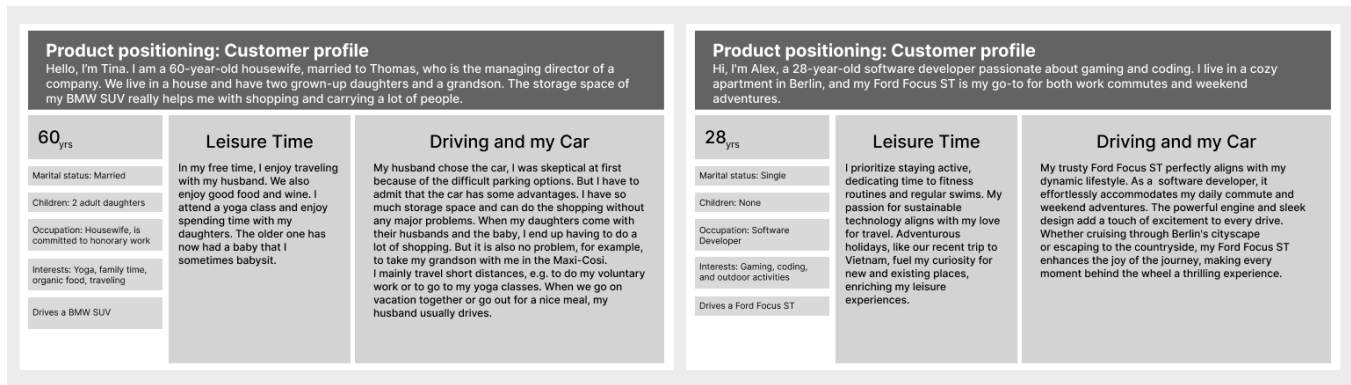


Figure 1: Personas of Scenario 2 used for the Survey. Left human-written, right AI-generated.

the prevailing ideas. However, we have refined our prompt with more specific strategies.

2.1.2 Role-play prompting, Persona Pattern. *Idea:* The so-called persona pattern [23] or role-play prompting [15], consists of instructing ChatGPT to take on a specific role to formulate the output from a corresponding perspective.

Application: To outsource the creation of personas for ideation purposes, design agencies can be consulted. Therefore, we gave ChatGPT the role of a consultant who acts on behalf of the company in question.

Strategy Example:

You are a consultant. You are acting on behalf of company XYZ.

By applying the role of a consultant to ChatGPT, we expect to receive more precise and sophisticated results.

2.1.3 One-Shot prompting. *Idea:* Provide an example of the desired output for ChatGPT to understand the underlying structure, format, and language used. This strategy is also known as one-shot prompting [3].

Application: To adapt the resulting personas to the style of the already existing human-expert written personas, we provided an example of them to ChatGPT.

Strategy Example

Example="""Target group: Architect Name: Ann-Marie Bien Age: 32 Profession: Interior designer, employed Company: young architectural office with 20 employees Gross income per Month: 3500 EUR, gross, Living situation: Vienna, co-living 45 sqm, rent Hobby + interest: loves architecture, design and art, theatre, rides an e-bike, is a visual person and is interested in new trends""

By providing this example to ChatGPT within our prompt, we aim to receive similar results in terms of structure and form. In general, the style could also be imitated by providing a large set of descriptions and instructions to restrict the output to the desired results. However, we opted for the more concise and comprehensible version of "One-Shot prompting".

2.1.4 Incremental Prompting. *Idea:* Although ChatGPT is proficient at solving simple problems almost effortlessly. However, once the prompt becomes too large, usually some parts of the desired results might be missing or inconsistent. Therefore, a larger task can be broken down into multiple simpler subtasks [4].

Application: Depending on the context, format, and scenario of the persona, the generation might be too complex to be processed by ChatGPT in a single prompt. It can therefore be beneficial to divide up the creation of personas for this more complex task to improve the quality of the output. The larger task of creating a fully decked-out persona can be divided into multiple smaller subtasks like demographic information generation, attribute selection or story writing, etc. To keep the consistency of the results, the prompts are subsequently used in the same chat to query ChatGPT.

Strategy Example:

- (1) *Prompt: Create Personas*
- (2) *Prompt: Add the following attributes for each previously created persona: <keywords>*
- (3) *Prompt: ...*

Originally, we tried to generate the personas within one prompt. Nevertheless, when splitting up the original prompt into multiple sub-tasks, we received a noticeably better result.

2.2 Experiment results

During our prompt formulation experiments with ChatGPT, we made some interesting observations.

- ChatGPT does not always solve the tasks in the same way (e.g. it creates different types of tables).
- Creating several personas at once, increased the variation in the set of personas created.
- Writing the company description in the language of the company's stakeholders, despite an otherwise English prompt, resulted in more credible personas.

By utilizing the previously discussed strategies, we were able to formulate prompts for ChatGPT with which we could create personas in an identical format and with a similar content style, which were almost indistinguishable from the human expert written personas. Even though the results look promising, for the usage of the personas in a work-related context, the results would still have



Figure 2: Examples of the two persona formats used. Left scenario 1 human-written persona for an appliance company, right scenario 2 AI-generated persona in an automotive context.

to be manually checked to ensure sufficient quality and consistency. Even so, the amount of effort used can be reduced to a minimum compared to manually creating the personas.

Applying these strategies, we were able to create relevant and consistent personas. Figure 1 illustrates the similarity between a human-created and an artificially generated persona using ChatGPT. The next section discusses our survey setup to evaluate the quality of the results of our prompting strategies.

3 EVALUATION

The evaluation methodology for assessing the viability of generated personas draws inspiration from Turing test settings commonly used to evaluate responses from conversational agents like LLMs and Chatbots [13, 22]. Its goal is to determine whether humans can distinguish between an AI solution and a human-built one. Following this, we conducted an evaluation in the form of a remote online study (created using the Limesurvey-based organizational platform). The following subsections describe the survey setup, methodology as well as the results gathered.

3.1 Methodology and Procedure

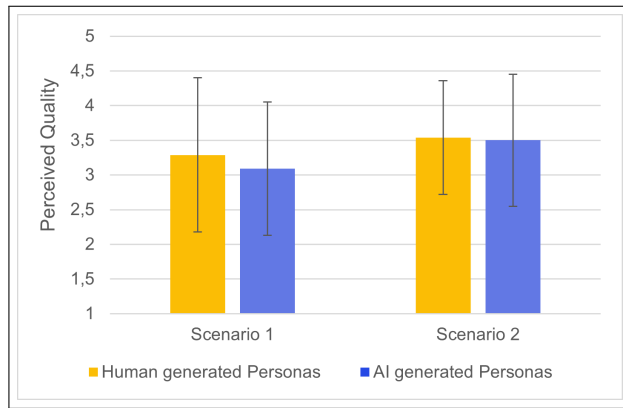
For evaluation purposes, we created personas in two distinct application scenarios, both derived from previous industrial use cases with varied stakeholders requiring personas for innovation and ideation. Scenario 1 depicted an appliance company specializing in door and window fittings, with specified stakeholder groups like "architects or investors". Scenario 2 focused on a luxury automotive company, targeting only end users. As illustrated in Figure 2, the persona formats differed; Scenario 1 personas featured individual attributes with bullet points, while Scenario 2 personas were predominantly continuous text with some bullet points. Each scenario consisted of 8 personas, where half of the shown persons were AI-generated and half were created by human experts. For both scenarios, The AI-generated personas were explicitly generated

with the prompting strategies as shown in Section 2. GPT-3.5. was used with the default settings as the model for generation. The generated output was transferred to the same template that was used by the human experts. No changes were made to the content created by ChatGPT, except for the company name for anonymization purposes.

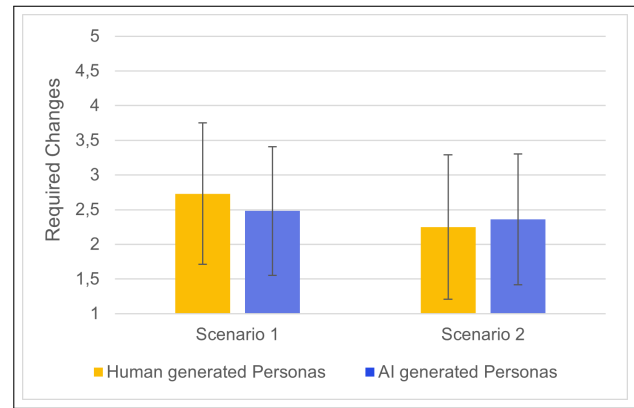
The online questionnaire was split into four different parts. In the first part, we gathered demographic information about our participants. The second part assessed Scenario 1, starting with an explanation and introduction of the scenario. Subsequently, the users were presented with eight different personas in randomized sequence and answered questions about their quality, need for changes, and an estimation of the creation method (manual or AI) for this specific persona. In the same format, Scenario 2 was assessed in the survey's third part. The survey concluded with general questions about utility, acceptance, and novelty of such a concept in the follow-up survey. On average, the survey took about 30-45 minutes per participant for completion.

Scenario	# correct	# incorrect	Accuracy
<i>-Human written personas-</i>			
Scenario 1	17	27	38,64%
Scenario 2	34	10	75,00%
<i>-AI generated personas-</i>			
Scenario 1	26	18	59,09%
Scenario 2	28	16	63,64%
<i>-Scenarios overall-</i>			
Scenario 1	43	45	49%
Scenario 2	62	26	70%

Table 1: Overview of allocation results: Correctly or Incorrectly classified



(a) Perceived quality of the shown persona (1=poor quality, 5=very high quality)



(b) Perceived necessary changes to the shown persona (1=no changes, 5=complete redesign)

Figure 3: Comparison between AI-generated and Human-written Personas based on average participant scores, with error bars representing Standard Deviation

3.2 Participants

To conduct this study, we recruited participants from our research department who have a UX, software development, or interaction design background through email and ms-teams channels. 11 participants (7 male, 4 female) took part in the study. The age of participants ranged from 28 to 44 ($M = 37.18$, $SD = 4.77$). On average the years of experience of participants in the user experience field was about nine working years ($M = 9.09$, $SD = 5.92$). 4 participants stated that they use personas frequently (several times per year) in their professional context, while 4 of the remaining participants stated an occasional (1-3 times per year) usage and 3 rare (once every few years) usage of personas. 5 participants chose "little experience", 2 "medium experience" and 4 "high experience" when asked for their self-rated experience with generative text AI tools such as ChatGPT.

3.3 Results

By conducting the survey, we gathered some results that cover the identification and the perceived quality of human-created compared to AI-generated personas. This section covers the results we gathered during the survey in detail.

3.3.1 Correct identification of personas. Table 1 illustrates the classification of evaluated personas, revealing that participants could not distinguish between AI and human-created personas significantly. However, slight differences emerged based on the scenarios. In the first scenario, accuracy slightly fell below the chance level, while in the second scenario, participants demonstrated an advantage in discerning human-created from AI-generated personas. A potential explanation for this discrepancy lies in the distinct persona formats—bullet-point style in the first scenario and continuous text with short sentences in the second, offering varied cues for classification based on language formulation.

When asked about strategies how to decide about the authorship on a persona, people mostly mentioned language and style (6 mentions, e.g. "language not linear", "too marketing bloomy" or "spelling

errors"), while only some people also argued about the content (3 mentions, e.g. "facts that were hallucinated", "inconsistencies: AI are more generic wrong, humans are more specific wrong").

3.3.2 Perceived quality. Figure 3a displays the participants' average quality assessments, revealing that the mean perceived quality was slightly above medium for both AI and human-created personas. This suggests that AI-generated personas were rated as favorably as those created by humans. The non-parametric Wilcoxon rank-sum test revealed that there were no statistically significant differences between the scores given by participants for human-written and AI-generated personas ($p > .05$).

3.3.3 Further changes to the personas deemed necessary. In evaluating the personas for real-life use in a user experience context, we inquired about the necessary changes. Figure 3b illustrates a consistent pattern, with a consensus favoring partial customization or adjustments to specific details before actual usage. The non-parametric Wilcoxon rank-sum test revealed that there were no statistically significant differences between the scores for required changes given by participants for human-written and AI-generated personas ($p > .05$).

3.3.4 General assessment of AI-supported persona creation. At the study's conclusion, participants shared their overall opinion on the concept of AI persona generation. The idea was generally deemed useful (mean 4.18 on a scale from 1 to 5) and applicable in daily work (mean 4.09 on the same scale). Regarding novelty, the presented concepts were perceived as only slightly novel (3.09 on the same scale), potentially influenced by the AI-savvy nature of the surveyed user group. A nuanced examination of usefulness revealed a correlation between experience level and task requirements. More experienced UX designers expressed a preference for enhanced interaction with an AI Persona rather than a simple document creation process.

4 DISCUSSIONS AND LIMITATIONS

The study suggests that personas generated by ChatGPT are challenging to distinguish and judge similarly to human-created ones, indicating their viability with experts. However, the findings are subjective to study limitations. For example, as the original personas created by hand were only available in German, there needed to be an additional translation step. This might have introduced some weak points to the descriptions as some of our evaluators mentioned, the exact phrasing and flow of language as an important differentiator when giving estimations about the creator of the personas. Additionally, the study recognizes limitations in assessing the representativeness and diversity of target user groups, emphasizing the need for actual end-user involvement. The study suggests that, like human-created personas, automated personas may also contribute to stereotypes and biases, potentially offering an opportunity for promoting more equality in persona creation. As personas are said to be a source of stereotypes and biases also in human-created personas [17], this could even be an opportunity towards more equality.

Furthermore, automated persona generation may lack empathy compared to manual creation. In discussions on the approach, some stressed the importance of empathizing through the persona's usage and perspective. Our method might involve less psychological connection, potentially leading to reduced empathy with automatically generated personas. This was acknowledged by a survey participant emphasizing the importance of the creative process in building a persona's perspective.

Nevertheless, in our approach, when we engaged in discussions regarding the initial requirements and usage context with our collaborative partners (i.e., prospective users of the generated personas), the emphasis leaned more towards saving time and effort during group moderation sessions. It also focused on devising practical templates and prompts to facilitate workshops and serve as a foundation for group activities. Particularly during instances when group activities stall or lose momentum, having pre-generated templates could aid in swiftly reigniting the conversation flow.

Furthermore, it's essential to underscore that the automated persona results were not selectively chosen (i.e., cherry-picked). In practical scenarios, UX experts can effortlessly generate numerous personas automatically, selecting the most suitable ones as a foundation for further refinement. This approach has the potential to expedite iteration, enhance quality, and achieve better alignment with the specific context through a more iterative and collaborative creation process.

5 CONCLUSION AND FUTURE WORK

In this study, we developed a ChatGPT-driven prompting strategy for persona generation, validated through a small-scale expert study. This lays a crucial foundation for deeper analyses and applications of generative AI in persona creation and Human-Centered Design (HCD) processes. However, further research is necessary to broaden the scope of the reported study by involving other potential stakeholders and obtaining a larger sample size to enhance the robustness of the findings. Our future plans involve creating an application that leverages these strategies to expedite persona

creation for designers, developers, and UX practitioners. The application, integrated with models like GPT-4.0, will allow users to customize usage descriptions and persona formats while concealing underlying structures and steps. From our discussions with industry partners, they have shown keen interest in incorporating the generated persona into future live workshops with their clients.

A compelling future concept entails the impersonation of both generated and hand-made personas by generative AI. This may materialize as a persona-bot, assimilating the traits and background information of a designated persona. This concept holds promise for investigating user acceptance and empathy towards such personas. Further exploration into multi-agent, autonomously impersonated personas interacting directly with designers and developers, or among themselves [24], presents exciting possibilities. Future studies should delve into the practical applications, interactive dynamics, and user acceptance within this evolving landscape, paving the way for innovative advancements in persona-driven design processes.

ACKNOWLEDGMENTS

This research received support from the Ministry of Economic Affairs, Labour, and Tourism of Baden-Wuerttemberg within the project "KIFortschrittszentrum Lernende Systeme und Kognitive Robotik". Additionally, we extend our gratitude to all participants at Fraunhofer IAO Stuttgart for generously contributing their time and valuable feedback to the study.

REFERENCES

- [1] Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. 2023. Out of one, many: Using language models to simulate human samples. *Political Analysis* 31, 3 (2023), 337–351.
- [2] Volker Bilgram and Felix Laarmann. 2023. Accelerating innovation with generative AI: AI-augmented digital prototyping and innovation methods. *IEEE Engineering Management Review* (2023).
- [3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
- [4] Sondos Mahmoud Bsharat, Aidar Myrzakhan, and Zhiqiang Shen. 2023. Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4. *arXiv preprint arXiv:2312.16171* (2023).
- [5] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. 2023. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712* (2023).
- [6] Alan Cooper. 1999. *The inmates are running the asylum*. Springer.
- [7] Molly Crockett and Lisa Messeri. 2023. Should large language models replace human participants? (2023).
- [8] Danica Dillion, Niket Tandon, Yuling Gu, and Kurt Gray. 2023. Can AI language models replace human participants? *Trends in Cognitive Sciences* (2023).
- [9] Christoph Engel, Max RP Grossmann, and Axel Ockenfels. 2024. Integrating machine behavior into human subject experiments: A user-friendly toolkit and illustrations. *MPI Collective Goods Discussion Paper* 2024/1 (2024).
- [10] Toshali Goel, Orit Shaer, Catherine Delcourt, Quan Gu, and Angel Cooper. 2023. Preparing Future Designers for Human-AI Collaboration in Persona Creation. In *Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. 1–14.
- [11] Jan Gulliksen, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson, and Åsa Cajander. 2003. Key principles for user-centred systems design. *Behaviour and Information Technology* 22, 6 (2003), 397–409.
- [12] Perttu Hämmäläinen, Mikke Tavast, and Anton Kunnari. 2023. Evaluating large language models in generating synthetic hci research data: a case study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [13] Daniel Jannai, Amos Meron, Barak Lenz, Yoav Levine, and Yoav Shoham. 2023. Human or Not? A Gamified Approach to the Turing Test. *arXiv:2305.20010 [cs.AI]*
- [14] A. Baki Kocaballi. 2023. Conversational AI-Powered Design: ChatGPT as Designer, User, and Product. *arXiv:2302.07406 [cs.HC]*

- [15] Aobo Kong, Shiwan Zhao, Hao Chen, Qicheng Li, Yong Qin, Ruiqi Sun, and Xin Zhou. 2023. Better zero-shot reasoning with role-play prompting. *arXiv preprint arXiv:2308.07702* (2023).
- [16] Anne Elisabeth Krueger. 2022. Two methods for Experience Design based on the Needs Empathy Map: Persona with Needs and Needs Persona. *Mensch und Computer 2022-Workshopband* (2022).
- [17] Nicola Marsden, Jasmin Link, and Elisabeth Bülesfeld. 2014. *Personas und stereotype Geschlechterrollen*. De Gruyter Oldenbourg, Berlin, München, Boston, 91–104. <https://doi.org/doi:10.1515/9783110363227.91>
- [18] John Pruitt and Tamara Adlin. 2010. *The persona lifecycle: keeping people in mind throughout product design*. Elsevier.
- [19] Liying Qiu, Param Vir Singh, and Kannan Srinivasan. 2023. How Much Should We Trust LLM Results for Marketing Research? *Available at SSRN 4526072* (2023).
- [20] Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. 2023. Whose opinions do language models reflect?. In *International Conference on Machine Learning*. PMLR, 29971–30004.
- [21] Albrecht Schmidt, Passant Elagroudy, Fiona Draxler, Frauke Kreuter, and Robin Welsch. 2024. Simulating the Human in HCD with ChatGPT: Redesigning Interaction Design with AI. *Interactions* 31, 1 (2024), 24–31.
- [22] AM Turing. 1950. Computing Machinery and Intelligence in "Mind" 59 (236) 433–460. *Cerca con Google* (1950).
- [23] Jules White, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C Schmidt. 2023. A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382* (2023).
- [24] Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Shaokun Zhang, Erkang Zhu, Beibin Li, Li Jiang, Xiaoyun Zhang, and Chi Wang. 2023. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155* (2023).
- [25] Jingfeng Yang, Hongye Jin, Ruixiang Tang, Xiaotian Han, Qizhang Feng, Haoming Jiang, Shaochen Zhong, Bing Yin, and Xia Hu. 2023. Harnessing the power of llms in practice: A survey on chatgpt and beyond. *ACM Transactions on Knowledge Discovery from Data* (2023).
- [26] Yutao Zhu, Huaying Yuan, Shuting Wang, Jiongnan Liu, Wenhan Liu, Chenlong Deng, Zhicheng Dou, and Ji-Rong Wen. 2023. Large language models for information retrieval: A survey. *arXiv preprint arXiv:2308.07107* (2023).