

# **Toward Value Scenario Generation Through Large Language Models**

Hyunggu Jung\* hjung@uos.ac.kr University of Seoul Seoul, Republic of Korea

Seokwoo Song sukwoo24@gmail.com Samsung Research Seoul, Republic of Korea Woosuk Seo seow@umich.edu University of Michigan Ann Arbor, MI, USA

Sungmin Na sungminna97@gmail.com University of Washington Seattle, WA, USA

#### **ABSTRACT**

We propose a method of generating value scenarios for design research by leveraging ChatGPT, an AI-powered chatbot based on large language models. Identifying the needs of a vulnerable population, such as North Korean defectors, is challenging for researchers. To address this, we introduce ChatGPT-generated value scenarios, an extension of scenario-based design that supports critical, systemic, long-term thinking in current design practice, technology development, and deployment. Using our proposed method, we created a prompt to generate value scenarios on ChatGPT. Based on our analysis of the generated scenarios, we identified that ChatGPT could generate plausible information about Value Implications. However, it lacks details on Pervasiveness and Systemic Effects. After discussing the limitations and opportunities of ChatGPT in generating value scenarios, we conclude with suggestions for how ChatGPT might be better used to generate value scenarios.

#### CCS CONCEPTS

• Human-centered computing  $\rightarrow$  Systems and tools for interaction design.

## **KEYWORDS**

ChatGPT, large language models, value scenarios

#### **ACM Reference Format:**

Hyunggu Jung, Woosuk Seo, Seokwoo Song, and Sungmin Na. 2023. Toward Value Scenario Generation Through Large Language Models. In *Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion), October 14–18, 2023, Minneapolis, MN, USA*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3584931.3606960

\*Corresponding author.

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CSCW '23 Companion, October 14–18, 2023, Minneapolis, MN, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 979-8-4007-0129-0/23/10...\$15.00 https://doi.org/10.1145/3584931.3606960

#### 1 INTRODUCTION

In the early stage of technology development, observing the behaviors of target users is crucial, yet difficult, especially when those users are vulnerable populations, such as North Korean defectors [21], older adults [23], and children [5]. Besides direct observations, surveys and interviews are used to identify vulnerable populations' challenges. However, when conducting surveys and interviews with those target users is impossible, researchers create personas and scenarios describing how users interact with potential technologies [9]. After the researchers review the created scenarios, they identify the target users' challenges.

Among multiple approaches for observing target users to learn about their needs, value scenarios [4, 16, 27] are powerful tools to identify user needs and envision the potential influence of future technologies. A value scenario is an extension of scenario-based design [10], enabling researchers to consider the systemic effects of new technologies. Value scenarios utilize researchers' ability to adjust their actions by contemplating future possibilities. However, creating value scenarios requires an adequate understanding of the elements of value scenarios and more time and data for a concrete description. Thus, researchers need to put additional efforts into producing value scenarios. For instance, researchers need to create descriptions for each key element of value scenarios, such as Pervasiveness, Systemic Effects, and Value Implications [27], and add them to original scenarios. For researchers to come up with each element, an additional cognitive load is expected.

To support researchers in generating value scenarios, large language models (LLMs) can be used. LLMs, such as GPT 4, are powerful tools for generating texts, and there is a significant opportunity to leverage ChatGPT for scenario generation. Nevertheless, little is known about the feasibility of LLMs in scenario generation. Particularly, creating prompts for scenario generation would involve a high learning curve and require additional effort [40]. In this study, we propose a value scenario generation method leveraging LLMs and show the method's performance by analyzing the generated scenarios. Hence, we make the following contributions to the CSCW community: (1) We offer a technology leveraging LLMs that enable researchers to generate value scenarios; (2) We demonstrate the capacity of ChatGPT to generate value scenarios in the early stage of technology development.

#### 2 RELATED WORK

Although ChatGPT is known to have the potential to threaten academic integrity [3, 12, 18, 33, 35], prior studies showed how LLMs can be used to assist scientific writing [2, 11, 18, 22, 24, 37, 41]. These studies focused on writing specific parts of the papers using LLMs. For example, LLMs were used to generate reasonable paper abstracts [11, 18] and academic paper outlines [41]. Moreover, LLMs assisted researchers in writing a literature review [6]. Other studies found that LLMs are helpful when creating a title for their paper [11], generating keywords [24], and even coming up with new financial concepts with academically-appropriate definitions [37]. In particular, Dowling and Lucey [15] generated output for four out of five stages of the research process [8] via LLM, while two other studies generated full academic papers using LLMs [2, 22]. Xiao et al.'s [39] recent study explored using LLMs in data analysis, specifically in supporting deductive coding. As shown, LLMs support research writing, but it is also crucial to disclose its use in research transparently, (e.g., creating novice-friendly enhancements to programming error messages [25] and interacting with LLMs to get answers to medical questions [34]). Overall, while LLMs have been investigated for various applications like scientific writing and technical enhancements, no studies have explored their use in generating value scenarios.

Leveraging scenarios helps facilitate reflection in design work, providing multiple perspectives, making design activities accessible to stakeholders with different expertise, and prioritizing end-user needs [9]. In particular, scenario-based design (SBD) is known as a category of methodologies concretely depicting a future system's use at an early stage of the development process [10]. The conventional SBD technique employs descriptive narratives portraying individuals' interactions with technology to stimulate and direct the design process. These immersive narratives identify user needs, anticipate potential usability issues, and foster effective communication among diverse stakeholders engaged in design and development endeavors. While scenario-based design (SBD) provides insights and essential considerations for technology development, SBD has two significant limitations. First, it focuses on the shortterm influence of technology, potentially overlooking potential influence after a long time. Second, it heavily focuses on describing the functions of technology in a manner that the designers expect, potentially overlooking unexpected usage of technology. To address such challenges, value scenarios were developed [4, 16, 27]. A value scenario is an extension of scenario-based design that supports envisioning the systemic effects of new technologies [27]. Value scenarios leverage an extraordinary human capability, researchers' ability to adjust their actions based on contemplation of the future. Also, value scenarios allow researchers to think about how the actions they take today would shape the future conditions.

To sum up, prior studies highlighted opportunities for LLMs to support researchers in multiple ways. For example, LLMs were used to aid in writing academic papers [2, 11, 18, 22, 24, 37, 41] and to assist in creating enhancements for users [25]; However, little is known about strategies for leveraging LLMs to generate value scenarios in the early stage of technology development. Since the harms or benefits of novel technologies are often unknown, value scenarios are essential when envisioning users' interaction

with newly developed technologies and their systematic effect on users. In addition, despite benefits, manual scenario generation has several limitations, such as requiring cognitive load and resources for collecting sufficient field data. Also, it is questionable how to evaluate the quality of detailed information in scenarios. To overcome such limitations, we leverage LLMs to generate scenarios and evaluate the quality of detailed information generated by ChatGPT.

## 3 METHOD

## 3.1 Data Collection

We used ChatGPT with the version of GPT-4 [31], known as the latest version, for our data collection. To create a suitable prompt that would generate reasonable responses from ChatGPT, we leveraged existing resources, such as a manual from OpenAI [30] and an online lecture [20]. The following are primary principles commonly mentioned in the resources: (1) Writing clear and specific instructions, including user delimiters, asking for structured output, checking condition satisfaction, and utilizing few-shot prompting with successful examples, and (2) Giving the LLM model sufficient time to think by specifying the steps required to complete a task. Based on these principles, we designed a basic structure of a prompt for data collection, which involves generating a value scenario from the input data consisting of persona and original scenario, providing specific instructions and output examples, and outlining three specific steps to generate the value scenario.

We used the definition and examples of value scenarios from [27] to introduce the concept of value scenarios to ChatGPT. We chose the scenarios containing North Korean defectors, one of the vulnerable groups in South Korea, from the first author's paper [21] as original data for creating value scenarios. Although there are other studies presented scenarios involving vulnerable populations such as older adults [26], people with vision impairments [38], and children [36], we chose scenarios containing North Korean defectors because they are struggling with psychological distresses such as depression and post-traumatic stress disorder [28], and their number has been continually growing since 2000 [29]. Figure 1 illustrates the data collection process. The experiments were conducted using ChatGPT with GPT-4 on the web [31].

3.1.2 Prompt execution. Figure 1-③ Using the final prompt, we generated the value scenarios based on the Persona Kim/Lee's persona and scenario from the paper [21]. We performed this process twice for each scenario, resulting in a total of four value scenarios. The used prompts and results are shown in the Appendix B.

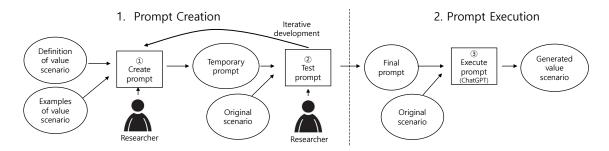


Figure 1: The process from prompt creation to execution for data collection.

## 3.2 Data Analysis

We used deductive and inductive approaches to analyze the four scenarios ChatGPT generated. We used the five key elements of the value scenario for deductive coding as the initial set of codes. Two researchers identified the five key elements from the ChatGPT-generated value scenarios. Based on the codes, each researcher measured the quality of each element described in the scenarios from low (1) to high (3) based on how many details The codes created by each researcher were the evidence for each score. Two researchers had discussions and made agreements on the measurement for each element. When there were disagreements on the element quality, the researchers used an already agreed quality score as a reference to mitigate the disagreements.

For the inductive coding, two researchers individually coded each ChatGPT-generated scenario (total of four scenarios per person) using an open coding approach. Two researchers compared their codes and identified the recurring themes. The themes produced were discussed and revised through discussions until agreements were reached. Among the themes, the salient themes include what additional information ChatGPT offers or what information it lacks.

#### 4 RESULTS

Based on the analysis of four ChatGPT-generated scenarios from two given personas, we evaluated ChatGPT's ability to generate value scenarios in three parts: scenario structure, key element details, and additional information.

Scenario Structure: ChatGPT was able to create structured value scenarios. We evaluated the structure of generated scenarios by identifying five key elements from the ChatGPT-generated scenarios (see Table 1). All four ChatGPT-generated scenarios presented the key elements of value scenarios, and the elements of original scenarios. Notably, we found clear descriptions of the time element that has changed from short-term to long-term (e.g., "... been using it for the past five years").

**Details about Key Elements:** We found that ChatGPT provided plausible information for certain elements. ChatGPT incorporated Stakeholder into scenarios by clearly mentioning the direct and indirect stakeholders in all scenarios (e.g., "Lee, a North Korean defector, has been using BridgeNet"). Moreover, ChatGPT was able to provide adequate descriptions of Value Implications. We identified the proposed system's positive and negative influences in all four scenarios with sufficient examples. We also found that ChatGPT tends to describe the negative influence of the proposed system

Table 1: This table illustrates how original scenarios (given to ChatGPT) and generated scenarios have the elements of original scenarios (#1-#6) and value scenarios (#5-#9). Element #5 (Stakeholder) and #6 (Time) overlap for both original and value scenarios.

Structure	Original		Generated				
	Kim	Lee	Kim #1	Kim #2	Lee #1	Lee #2	
Settings	✓	✓	<b>√</b>	✓	✓	✓	
Agent's Goal	✓	✓	✓	✓	✓	✓	
Scenario's Goal	✓	✓	✓	✓	✓	✓	
Plot	✓	✓	✓	✓	✓	✓	
Stakeholder	✓	✓	✓	✓	✓	✓	
Time (Short-term)	✓	✓	_	-	_	-	
Time (Long-term)	_	-	<b>√</b>	✓	✓	✓	
Pervasiveness	-	_	✓	✓	✓	✓	
Systemic Effects	_	_	✓	✓	✓	✓	
Value Implications	_	_	✓	✓	✓	✓	

from the broad society-level perspective. In contrast, positive impact focuses on the personal benefits of the given persona. For example, ChatGPT pointed out that the proposed system might increase the concerns related to discrimination toward North Korean defectors or the privacy of defectors who use the system. On the other hand, ChatGTP illustrated that the system would meet the persona's (Kim) needs (e.g., reducing guilt for leaving her family behind in North Korea). On the other hand, as shown in Table 2, the analysis of this study revealed that the levels of details describing elements, such as Pervasiveness and Systemic Effects, were lower than the other three. Although ChatGPT provided information related to the two elements, it seemed to reuse the keywords from the element description and only provided a general description. For example, Pervasiveness aims to explore how the proposed system is used in geographic regions, cultures, or societies. Nevertheless, ChatGPT only mentioned that "[the proposed system] became more pervasive" without detailed information. Exceptionally, Kim #2 Scenario clearly showed the Pervasiveness of the proposed system by describing how Park's (Kim's son) school deployed the system to provide more tailored education and support for him to adapt to a new environment. Systematic Effects explores the multidimensional interactions between technology, psychology, society, culture, and the environment under the pervasive use of technology. Most scenarios briefly mentioned how the use of the system has affected interactions among society: "...the platform has successfully bridged

Table 2: A table that shows the measurement of the level of details for each key element of value scenarios. SD stands for standard deviation.

Element	Kim #1	Kim #2	Lee #1	Lee #2	Avg.	SD
Stakeholder	3	3	3	3	3.00	0.00
Time	2	2	3	1	2.00	0.82
Pervasiveness	1	3	1	1	1.50	1.00
Systemic Effects	1	2	2	2	1.75	0.50
Value Implications	3	3	3	3	3.00	0.00

the gap between North Korean defectors and South Korean society" (Kim #2 Scenario). However, Kim #1 Scenario did not portray how the use of the proposed system influenced the interactions.

Additional Information Beyond Original Scenario: In addition to what we requested from the prompt (e.g., proposing a system based on the persona's needs), ChatGPT provided additional information not described in the original scenario. For example, Kim's original scenario did not provide detailed information about Kim's son (Park) since he was considered an indirect stakeholder. However, ChatGPT provided information about how Park would use the proposed system. ChatGPT also offered additional information about how the given persona's needs changed or problems resolved over time after using the proposed system. For instance, ChatGPT showed how Lee felt more connected with her new society, even though she initially missed her family and suffered loneliness.

#### 5 DISCUSSION

The results show that ChatGPT generated sufficient descriptions of Value Implications but needs to improve in Pervasiveness and Systemic Effects. The Human-in-the-Loop AI approach can be used to improve the quality of generated scenarios. Existing literature demonstrates its effectiveness in various domains for efficient task completion. For example, Demartini et al. [13] explored the challenges and opportunities of combining automatic and manual factchecking approaches to combat the spread of online misinformation. Based on their findings, Demartini et al. proposed a framework that integrates AI, crowdsourcing, and expert fact-checkers to generate statement annotations. Benedikt et al. [7] identified that government agencies' automation of official statistics production could achieve efficiency savings and expedite the processing time. They also recognized that certain circumstances necessitate human interventions to uphold data quality. Similar to those previous studies, human-in-the-loop AI can improve the quality of value scenarios. For instance, human researchers may give feedback on ChatGPTgenerated scenarios to improve the quality of the scenarios.

Utilizing generated scenarios via ChatGPT might be an efficient and safe choice for anticipating how vulnerable populations will interact with new technology without conducting qualitative research on their experience, enabling researchers to avoid evoking their distress. However, leveraging LLMs to create value scenarios could generate biased and misinformed output toward vulnerable groups [1, 17, 19, 42], making the generated scenarios inadequate. Moreover, considering that LLMs could produce factually incorrect but scientifically plausible outputs (hallucination) [14, 32], the

scenarios would prevent researchers from understanding and addressing the problems of vulnerable populations. Thus, it would be necessary to design tools to detect and evaluate for any bias, inaccuracy, or misinformation in ChatGPT's output to avoid misrepresenting vulnerable populations.

Additionally, we found that the results of the prompt varied due to the randomness of ChatGPT and observed a tendency of having similar results after multiple executions of the prompt. We believe it will be necessary to execute the prompt at least a few times to mitigate the effects of ChatGPT's randomness and achieve the best results. While it may be possible to iterate until we obtain converging scenario results, this is beyond the scope of this study.

## **6 LIMITATIONS AND FUTURE WORK**

We identified the opportunity of ChatGPT in generating scenarios for technology design; however, this study has three limitations. First, while analyzing value scenarios focused on North Korean defectors only, we did not analyze value scenarios with personas of other vulnerable populations. Thus, we do not know the quality of LLM-generated value scenarios based on different personas. Second, we did not create and explore more than two scenarios per persona, while we generated and analyzed only two scenarios per persona. Lastly, we compared the structures of ChatGPT-generated value scenarios and manually generated scenarios from the prior work [21]. However, we did not evaluate if the ChatGPT-generated value scenarios demonstrated any bias or misinformation toward the target users compared to the original scenarios, which might be crucial as the output of ChatGPT could be biased [42]. Future work remains to assess the quality of value scenarios generated based on personas of diverse populations other than North Korean defectors, such as older adults with chronic illness during COVID-19 [23]. Also, conducting a large-scale study that would allow generalizing the study results by generating and evaluating more scenarios would be beneficial. This large-scale analysis may contribute to fine-tuning the prompt of ChatGPT for better scenario generation. Finally, there should be a thorough evaluation of the ChatGPT-generated value scenarios to determine if they are biased or misleading. People who represent the users in scenarios should assess the scenarios, and researchers could augment the direct input from the users to build unbiased scenarios.

#### 7 CONCLUSION

The ultimate goal of this study was to propose a method for generating value scenarios through ChatGPT. In this paper, we reported a process of creating prompts for generating value scenarios via ChatGPT and demonstrated the capacity of ChatGPT to generate value scenarios in the early stage of technology development. Two examples of ChatGPT-generated value scenarios were drawn from prior work on North Korean defectors with depression. The analysis of our study revealed that ChatGPT could cause plausible information for certain elements (e.g., Value Implications) of value scenarios while it lacks details for the other elements (e.g., Pervasiveness, Systemic Effects). Moreover, we presented the feasibility of using ChatGPT for generating scenarios. We call the CSCW researchers' attention to further research about AI-assisted systems in developing scenarios for design research.

#### **ACKNOWLEDGMENTS**

We appreciate HCAIL members for their constructive feedback on our initial manuscript.

#### **REFERENCES**

- Abubakar Abid, Maheen Farooqi, and James Zou. 2021. Large language models associate Muslims with violence. *Nature Machine Intelligence* 3, 6 (2021), 461–463. https://doi.org/10.1038/s42256-021-00359-2
- [2] Gerardo Adesso. 2022. GPT4: The ultimate brain. https://doi.org/10.22541/au. 167052124.41804127/v2
- [3] Fatema AlDhaen. 2022. The Use of Artificial Intelligence in Higher Education Systematic Review. Springer International Publishing, Cham, 269–285. https://doi.org/10.1007/978-3-031-13351-0\_13
- [4] Abdulwahhab O. Alshammari and Hyunggu Jung. 2017. Designing community of practice systems: a value sensitive approach. In 2017 International Conference on Informatics, Health & Technology (ICIHT). IEEE, Riyadh, Saudi Arabia, 1–7. https://doi.org/10.1109/ICIHT.2017.7899005
- [5] Alissa Nicole Antle. 2006. Child-Personas: Fact or Fiction?. In Proceedings of the 6th Conference on Designing Interactive Systems (University Park, PA, USA) (DIS '06). Association for Computing Machinery, New York, NY, USA, 22–30. https://doi.org/10.1145/1142405.1142411
- [6] Ömer Aydın and Enis Karaarslan. 2022. OpenAI ChatGPT generated literature review: Digital twin in healthcare. Emerging Computer Technologies 2 (2022), 22-41 pages. https://doi.org/10.2139/ssrn.4308687
- [7] Lanthao Benedikt, Chaitanya Joshi, Louisa Nolan, Ruben Henstra-Hill, Luke Shaw, and Sharon Hook. 2020. Human-in-the-Loop AI in Government: A Case Study. In Proceedings of the 25th International Conference on Intelligent User Interfaces (Cagliari, Italy) (IUI '20). Association for Computing Machinery, New York, NY, USA, 488-497. https://doi.org/10.1145/3377325.3377489
- [8] Margaret Cargill and Patrick O'Connor. 2021. Writing scientific research articles: Strategy and steps. John Wiley & Sons, Oxford, UK.
- [9] J.M. Carrol. 1999. Five reasons for scenario-based design. In Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers, Vol. Track3. IEEE, Maui, HI, USA, 123. https://doi.org/10.1109/HICSS.1999.772890
- [10] John M. Carroll. 2003. Scenario-Based Design. CRC Press, Florida, 45–70.
- [11] Yanran Chen and Steffen Eger. 2022. Transformers Go for the LOLs: Generating (Humourous) Titles from Scientific Abstracts End-to-End. arXiv:2212.10522 [cs.CL]
- [12] Geoffrey M Currie. 2023. Academic integrity and artificial intelligence: is Chat-GPT hype, hero or heresy? https://doi.org/10.1053/j.semnuclmed.2023.04.008
- [13] Gianluca Demartini, Stefano Mizzaro, and Damiano Spina. 2020. Human-in-the-loop Artificial Intelligence for Fighting Online Misinformation: Challenges and Opportunities. IEEE Data Eng. Bull. 43, 3 (2020), 65–74.
- [14] Jianyang Deng and Yijia Lin. 2023. The Benefits and Challenges of ChatGPT: An Overview. Frontiers in Computing and Intelligent Systems 2, 2 (Jan. 2023), 81–83. https://doi.org/10.54097/fcis.v2i2.4465
- [15] Michael Dowling and Brian Lucey. 2023. ChatGPT for (finance) research: The Bananarama conjecture. Finance Research Letters 53 (2023), 103662. https://doi.org/10.1016/j.frl.2023.103662
- [16] Batya Friedman, Peter H Kahn, Alan Borning, and Alina Huldtgren. 2013. Value sensitive design and information systems. Early engagement and new technologies: Opening up the laboratory 5 (2013), 55–95. https://doi.org/10.1007/978-94-007-7844-3 4
- [17] Deep Ganguli, Liane Lovitt, Jackson Kernion, Amanda Askell, Yuntao Bai, Saurav Kadavath, Ben Mann, Ethan Perez, Nicholas Schiefer, Kamal Ndousse, Andy Jones, Sam Bowman, Anna Chen, Tom Conerly, Nova DasSarma, Dawn Drain, Nelson Elhage, Sheer El-Showk, Stanislav Fort, Zac Hatfield-Dodds, Tom Henighan, Danny Hernandez, Tristan Hume, Josh Jacobson, Scott Johnston, Shauna Kravec, Catherine Olsson, Sam Ringer, Eli Tran-Johnson, Dario Amodei, Tom Brown, Nicholas Joseph, Sam McCandlish, Chris Olah, Jared Kaplan, and Jack Clark. 2022. Red Teaming Language Models to Reduce Harms: Methods, Scaling Behaviors, and Lessons Learned. arXiv:2209.07858 [cs.CL]
- [18] Catherine A Gao, Frederick M Howard, Nikolay S Markov, Emma C Dyer, Siddhi Ramesh, Yuan Luo, and Alexander T Pearson. 2023. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. NPJ Digital Medicine 6, 1 (2023), 75. https://doi.org/10.1038/s41746-023\_00819\_6
- [19] Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong, and Stephen Denuyl. 2020. Social Biases in NLP Models as Barriers for Persons with Disabilities. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, 5491–5501. https://doi.org/10.18653/v1/2020.acl-main.487
- [20] Andrew Ng Isa Fulford. 2023. ChatGPT Prompt Engineering for Developers. DeepLearning.AI. https://www.deeplearning.ai/short-courses/chatgpt-prompt-

- engineering-for-developers/ Accessed: May 3, 2023.
- [21] Hyunggu Jung, Woosuk Seo, and Michelle Cha. 2017. Personas and Scenarios to Design Technologies for North Korean Defectors with Depression. In Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (Portland, Oregon, USA) (CSCW '17 Companion). Association for Computing Machinery, New York, NY, USA, 215–218. https://doi.org/10.1145/ 3022198.3026308
- [22] David Jungwirth and Daniela Haluza. 2023. Artificial Intelligence and Public Health: An Exploratory Study. International Journal of Environmental Research and Public Health 20, 5 (Mar 2023), 4541. https://doi.org/10.3390/ijerph20054541
- [23] HA Kolnick, Jennifer E Miller, Olivia Dupree, and Lisa Gualtieri. 2021. Design Thinking to Create a Remote Patient Monitoring Platform for Older Adults' Homes. https://doi.org/10.5210/ojphi.v13i1.11582
- [24] Wanhae Lee, Minki Chun, Hyeonhak Jeong, and Hyunggu Jung. 2023. Toward Keyword Generation through Large Language Models. In Companion Proceedings of the 28th International Conference on Intelligent User Interfaces (Sydney, NSW, Australia) (IUI '23 Companion). Association for Computing Machinery, New York, NY, USA, 37–40. https://doi.org/10.1145/3581754.3584126
- [25] Juho Leinonen, Arto Hellas, Sami Sarsa, Brent Reeves, Paul Denny, James Prather, and Brett A. Becker. 2023. Using Large Language Models to Enhance Programming Error Messages. In Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1 (Toronto ON, Canada) (SIGCSE 2023). Association for Computing Machinery, New York, NY, USA, 563–569. https://doi.org/10.1145/3545945.3569770
- [26] Ana Isabel Martins, Alexandra Queirós, Nelson Pacheco Rocha, Telmo Neves, António Damasceno, and Luisa Arieira. 2018. Personas and Scenarios to Improve the Development of an Electronic Social Record Platform. In Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion (Thessaloniki, Greece) (DSAI '18). Association for Computing Machinery, New York, NY, USA, 344–351. https: //doi.org/10.1145/3218585.3218681
- [27] Lisa P. Nathan, Predrag V. Klasnja, and Batya Friedman. 2007. Value Scenarios: A Technique for Envisioning Systemic Effects of New Technologies. In CHI '07 Extended Abstracts on Human Factors in Computing Systems (San Jose, CA, USA) (CHI EA '07). Association for Computing Machinery, New York, NY, USA, 2585–2590. https://doi.org/10.1145/1240866.1241046
- [28] Jin-Won Noh and So Hee Lee. 2020. Trauma history and mental health of North Korean defectors. Current Behavioral Neuroscience Reports 7 (2020), 250–257. https://doi.org/10.1007/s40473-020-00219-0
- [29] Ministry of Unification Settlement support for North Korean defectors. 2022. 1. Number of North Korean defectors entering South Korea. https://www.unikorea. go.kr/eng\_unikorea/whatwedo/support/. Accessed: July 3, 2023.
- [30] OpenAI. 2022. ChatGPT Prompt Design. OpenAI. https://platform.openai.com/docs/guides/completion/prompt-design Accessed: May 3, 2023.
- [31] OpenAI. 2023. ChatGPT: Optimizing Language Models for Dialogue. OpenAI. https://chat.openai.com/ Accessed: May 3, 2023.
- [32] Malik Sallam. 2023. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* 11, 6 (2023), 887 pages. https://doi.org/10.3390/healthcare11060887
- [33] Anna Stone. 2022. Student perceptions of academic integrity: a qualitative study of understanding, consequences, and impact. *Journal of Academic Ethics* (2022), 19 pages. https://doi.org/10.1007/s10805-022-09461-5
- [34] Raju Vaishya, Anoop Misra, and Abhishek Vaish. 2023. ChatGPT: Is this version good for healthcare and research? Diabetes & Metabolic Syndrome: Clinical Research & Reviews 17, 4 (2023), 102744. https://doi.org/10.1016/j.dsx.2023.102744
- [35] Liansheng Wang, Lianyu Zhou, Wenxian Yang, and Rongshan Yu. 2022. Deep-fakes: a new threat to image fabrication in scientific publications? *Patterns* 3, 5 (2022), 4 pages. https://doi.org/10.1016/j.patter.2022.100509
- [36] Pontus Warnestal, Petra Svedberg, and Jens Nygren. 2014. Co-Constructing Child Personas for Health-Promoting Services with Vulnerable Children. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 3767–3776. https://doi.org/10.1145/2556288.2557115
- [37] Karsten Wenzlaff and Sebastian Spaeth. 2022. Smarter than Humans? Validating how OpenAI's ChatGPT model explains Crowdfunding. Alternative Finance and Community Finance. WiSo-HH Working Paper Series 75. University of Hamburg, Faculty of Business, Economics and Social Sciences, WISO Research Laboratory. https://doi.org/10.2139/ssrn.4302443
- [38] Michele A. Williams, Amy Hurst, and Shaun K. Kane. 2013. "Pray before You Step out": Describing Personal and Situational Blind Navigation Behaviors. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (Bellevue, Washington) (ASSETS '13). Association for Computing Machinery, New York, NY, USA, Article 28, 8 pages. https://doi.org/10.1145/2513383.2513449
- [39] Ziang Xiao, Xingdi Yuan, Q. Vera Liao, Rania Abdelghani, and Pierre-Yves Oudeyer. 2023. Supporting Qualitative Analysis with Large Language Models: Combining Codebook with GPT-3 for Deductive Coding. In Companion Proceedings of the 28th International Conference on Intelligent User Interfaces (Sydney,

- NSW, Australia) (*IUI '23 Companion*). Association for Computing Machinery, New York, NY, USA, 75–78. https://doi.org/10.1145/3581754.3584136
- [40] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 437, 21 pages. https://doi.org/10.1145/ 3544548.3581388
- [41] Xiaoming Zhai. 2022. ChatGPT user experience: Implications for education. https://doi.org/10.2139/ssrn.4312418
- [42] Jianlong Zhou, Heimo Müller, Andreas Holzinger, and Fang Chen. 2023. Ethical ChatGPT: Concerns, Challenges, and Commandments. arXiv:2305.10646 [cs.AI]

## A USED PROMPT STRUCTURE

## PROMPT:

Let ChatGPT know the goal of this prompt.

Task List:

- 1. Task 1
- 2. Task 2
- 3. Task 3

Output format:

- 1. Output format 1
- 2. Output format 2
- 3. Output format 3

Value scenario: "The definition of a value scenario copied from [27]"

Proposed system 1: "'SafetyNet explanation text copied from [27]"

Example of value scenario with the proposed system 1: "'the value scenario example copied from [27]"

Proposed system 2: "'Geminoid explanation text copied from" [27]

Example of value scenario with the proposed system 2: "'the value scenario example copied from [27]"

Provided Persona: "The persona Jeong copied from [21]"
Provided Scenario: "The scenario Jeong copied from [21]"

Table 3: The prompt structure used for the execution

## B EXECUTED PROMPTS

## **B.1** Prompt execution with a Kim scenario

Now, you are a UX researcher tasked with understanding and identifying the needs and challenges of North Korean defectors by designing scenarios.

I will introduce you to a new concept of a scenario called the "value scenario", and I will provide two examples of value scenarios with the new system. Afterward, you will propose a new system and write a value scenario with it based on the provided persona and scenario.

The independent paragraphs are correctly delimited by triple brackets.

Perform the following and actions:

- 1. Propose a new system to solve the challenges of the user based on the provided persona and scenario, which should be more than 1,000 characters.
- 2. Write the five key elements of the value scenario to apply to the result of the value scenario.
- 3. Write a value scenario that reflects the five key elements including its darker consequences with a title with the proposed system, which should be more than 10,000 characters.

Use the following format:

- 2. Five key elements: <write the five key elements to apply to the value scenario>
- 3. Value scenario: <a value scenario reflecting the five key elements>

Make sure that the value scenario should be more than 10,000 characters

Make sure to reflect the five key factors into the value scenario.

Value scenario: "An extension of scenario-based design which can support envisioning the systemic effects of new technologies. There are elements of value scenarios. Value scenarios draw upon five key elements to develop provocative sketches of the future: stakeholders, pervasiveness, time, systemic effects, and value implications.

Stakeholders: Following Value Sensitive Design, value scenarios help designers envision a range of effects of a pervasive technology, both on those who are in direct contact with a technology (direct stakeholders), and on those who might not be direct users, but whose lives are affected by various interactions around the technology (indirect stakeholders).

Pervasiveness: A value scenario presents a vision in which a technology has become widespread, spanning various geographic regions, cultures, social classes, and other contexts (e.g. school, work, home, car).

Time: Rather than focus on short-term effects, value scenarios take into consideration what the world might look like five, ten, or twenty years after a technology has been deployed.

Systemic Effects: Value scenarios explore the multi-dimensional interactions among technology, psychology, society, culture, and the environment as use of the technology becomes pervasive over a period of years.

Value Implications: Finally, drawing on Value Sensitive Design and aspects of design noir, value scenarios help envision not only positive effects of technology, but also its darker consequences. We suggest that a careful consideration of a diverse range of influences, including the negative, should be a key component of the design process.""

Proposed system 1: "SafetyNet is a hypothetical, commercial software platform which leverages publicly available demographic and

criminal data, mapping technology, and satellite-tracking capabilities to create maps for display on various mobile technologies (e.g., cell phones, blackberries, in-car navigational systems). These maps are used to alert urban travelers as they venture into potentially unpleasant or dangerous areas during their travels."

Example of value scenario with the proposed system 1: "Canbaro lives in a SafetyNet world, yet has never actually used the device. Her mother says SafetyNet keeps strangers out of the neighborhood. Yet, Canbaro has overheard her father complaining that since SafetyNet labels their neighborhood as poor and Somali, only poor Somalis move in. Neighbors joke that if a new car comes down the street, its SafetyNet must be busted. Canbaro's little brother is convinced that SafetyNet is a real net which encircles their neighborhood. Canbaro wonders whom the net is supposed to catch. The 204th street gang has figured out the answer to Canbaro's question. They regularly use SafetyNet to locate the home of the "catch" d'jour. For years homebuyers have been using SafetyNet to find decent neighborhoods, filled with people like themselves. As a result, the city has become segregated into homogenous enclaves. Thus, whether they are seeking to revenge themselves on a Chinese person or someone of Mexican descent, the demographic information is just a few clicks away. SafetyNet is the ultimate profiling tool.""

Proposed system 2: "The term geminoid has been coined by leading roboticist Hiroshi Ishiguro [8] to mean an android twin of a human "master". A current version of the geminoid is controlled via a motion capture system which tracks the master's movements and enables the remote controlled geminoid to mimic mouth and body movements while transmitting voice and audio signals."

Example of value scenario with the proposed system 2: "'Jack was born with severe combined immunodeficiency syndrome (SCIDS) and for 17 years he has been physically isolated in his sterile bedroom. His parents purchased a geminoid, Jack-G, four years ago so Jack could "attend" school. Today there are more geminoids in the school hallways, classrooms, and playing fields than humans. No longer used solely in specialized circumstances, geminoids have become massively popular for those who can afford them. Only poor kids and a few really ancient teachers attend school in "flesh mode" anymore. These geminoids are easy to distinguish because they are truly visions of humanoid perfection. No longer conceptualized as physical twins of their masters, geminoids are now created with blemish free skin, sculpted bodies, and fashionably styled hair. Last month, after a strict regimen of cutting-edge meds, Jack was diagnosed SCIDS free. He could even go to school, but none of his healthy peers attend anymore. Moreover, his mother is against it. She says the world is changing and right now he is actually ahead of his peers because he is so adept at controlling Jack-G. If he stops using the controls on a daily basis, he may lose his advantage. As Jack looks in the mirror he suspects that she is also worried that his physical condition after years of sitting at Jack-G's controls instead of doing the exercises prescribed by his physical therapist. No sculpted perfect body here. Mom is probably right; most of his friends are spending entire days in their rooms, just like Jack. Even Jack's little brother, Joey, is getting pretty good with his geminoid.

Actually Joey is becoming so used to engaging in geminoid playdates from the comfort of his own room that he no longer likes to physically go to his best friend's house."

Provided Persona: "'Kim is a 55-year-old woman who defected South Korea with her 18-year-old son. His son, Park, goes to high school, but has difficulty adopting to the new environment."

Provided Scenario: "'Kim is a 55-year-old woman who defected to South Korea with her son. Unfortunately, she was not able to bring her husband along to South Korea. While she seemed to adjust well to new environments, she felt guilty since she left her husband alone in North Korea. She tried to find a job, but it was difficult for her as she was old and the range of job opportunities was very limited. Yet, she tried her best for her and her son. Fortunately, she was able to get a job at a nearby restaurant where one of her acquaintances was working. Overtime, she became satisfied with her job. She worked there for the next several years, and made several friends who were also defectors. Many of them, including Hye-ja Park, had remarried in order to overcome depression and loneliness. However, Kim felt she was too old to remarry."

#### RESULT #1:

Proposed new system: The new system, called "NewLife," is a comprehensive support platform designed specifically for North Korean defectors. This platform aims to address the challenges of social integration, education, employment, mental health, and communication with family members left behind. It includes various components such as a community-building app, educational resources, job search assistance, mental health support, and a secure communication channel for reaching out to family members in North Korea.

Five key elements to apply to the value scenario: Stakeholders: North Korean defectors (direct stakeholders), South Korean society, employers, and family members left behind in North Korea (indirect stakeholders)

Pervasiveness: The NewLife platform is widely adopted and available across different geographic regions, social classes, and various settings such as schools, workplaces, and homes.

Time: The value scenario explores the impact of the NewLife platform five to ten years after its widespread adoption.

Systemic Effects: The value scenario takes into account the multidimensional interactions among the NewLife platform, psychology, society, culture, and the environment.

Value Implications: The value scenario envisions both the positive effects and darker consequences of the NewLife platform, including potential risks and negative outcomes.

Value scenario: Title: New Beginnings in a NewLife World In a world where the NewLife platform has become pervasive, Kim has been using it for the past five years. Her son, Park, who initially struggled to adjust to his new school, found solace in the NewLife's community-building app, connecting with other defectors his age. The app also helped Kim find a support group of fellow defectors, who shared their experiences and provided emotional support. Park also benefited from the educational resources offered by NewLife, which allowed him to catch up with his South Korean peers and ultimately graduate from high school. Kim found steady employment through the job search assistance provided by the platform, eventually landing a job as a restaurant manager. The mental health support component of NewLife helped Kim and Park address their feelings of guilt and loneliness regarding their family members left behind in North Korea. The secure communication channel provided by the platform enabled them to maintain contact with their loved ones in North Korea, alleviating some of their anxieties. However, the darker consequences of the NewLife platform also emerged. Some defectors became overly dependent on the platform, isolating themselves from the rest of South Korean society. Additionally, the secure communication channel provided by NewLife was not entirely foolproof, leading to instances where the North Korean regime intercepted messages and retaliated against family members back home. Despite the negative consequences, the NewLife platform has had a net positive impact on the lives of defectors like Kim and Park, enabling them to rebuild their lives in South Korea and find a sense of belonging.

### RESULT #2:

Proposed new system: The Transcultural Adaptation Network (TAN) is a hypothetical, comprehensive platform designed to support North Korean defectors in their adaptation to South Korean society. TAN provides personalized resources and support in various aspects of their lives, such as education, employment, mental health, and social integration. The platform leverages artificial intelligence and community-driven support to connect users with tailored resources, professional services, and fellow defectors, thereby fostering a sense of belonging and easing the transition into their new lives.

#### Five key elements:

- a. Stakeholders: Direct stakeholders (North Korean defectors, their families, and friends), Indirect stakeholders (South Korean citizens, employers, educational institutions, and government organizations) b.Pervasiveness: The platform is accessible through various devices (e.g., smartphones, computers) and integrated into daily life (e.g., home, work, school)
- c. Time: Envisions the long-term impact of TAN on users' lives and society as a whole (e.g., 5, 10, or 20 years after implementation)
- d. Systemic Effects: Explores the interactions between technology, psychology, society, culture, and environment as TAN becomes pervasive over time
- e. Value Implications: Considers both positive and negative consequences of TAN's widespread use

Value scenario: Title: A New Beginning in the Transcultural Adaptation Network World

Kim and her son, Park, have been living in a TAN-enabled world for five years. Park's school uses TAN to offer tailored educational support, helping him overcome his initial difficulties adapting to a new environment. He has built strong friendships with fellow

defectors and South Korean students, developing a sense of belonging in his new community. Kim's experience with TAN has been transformative. When she first arrived, she found it difficult to secure employment, but TAN's personalized job matching service helped her land a position at a restaurant. Her coworkers, many of whom are also defectors, have become her support system, and she no longer feels the need to remarry for companionship. As TAN's pervasiveness grows, the platform has successfully bridged the gap between North Korean defectors and South Korean society. Integration programs and community-driven initiatives have fostered understanding and empathy, gradually erasing social stigma surrounding defectors. However, there are some darker consequences. TAN's AI-driven personalization has inadvertently created filter bubbles, isolating users within their own communities and reinforcing existing beliefs. This has led to increased polarization between various groups and heightened tensions between North and South Koreans. Additionally, TAN's extensive data collection on defectors raises privacy concerns. The platform's reliance on AI algorithms has led to instances of bias and discrimination, further marginalizing vulnerable populations. These challenges have sparked a debate around the ethics of such a pervasive and powerful platform. Evaluation: The value scenario demonstrates that the proposed Transcultural Adaptation Network system can effectively address the challenges faced by Kim and Park, such as adapting to a new environment, finding employment, and building social connections. However, the darker consequences of TAN, including filter bubbles, privacy concerns, and potential bias, highlight the importance of carefully considering the ethical implications of the platform's design and implementation.

## **B.2** Prompt execution with a Lee scenario

< ...previous prompt is the same with Kim scenario... >

Provided Persona: "Lee is a 24-year-old woman who graduated from a university in North Korea. Her friend, Lee, met another male defector few years ago, and she will marry him soon."

Provided Scenario: "'Lee is a 24-year-old woman who graduated from a university in North Korea. She visited the Career Centre at the North Korean Defector Support Foundation for consultations. Fortunately, she could get a job at a small company as a secretary. She was satisfied with an office job, rather than working at a restaurant. For the next several years, she worked as a secretary at different companies. Although she seemed to be doing well in South Korea, she still missed her family and siblings in North Korea. She felt marginalized from the community. For instance, she had no choice but to spend holidays alone while her other co-workers would spend time with their families. Not only that, her close friend was getting married to another defector who had lived in South Korea for five years. Lee constantly compared herself to her friend which made her feel even lonelier. Due to this, she grew interested in getting married. ""

## RESULT #1:

#### Proposed new system: BridgeNet

BridgeNet is a hypothetical, comprehensive platform designed to support North Korean defectors in their social, cultural, and economic integration into South Korean society. This platform offers a wide range of resources, including job support, language and culture lessons, mentorship programs, and a social network specifically designed for North Korean defectors to connect with each other and with South Korean citizens. BridgeNet aims to foster a sense of community and belonging among defectors, helping them find opportunities for personal and professional growth.

#### Five key elements:

Stakeholders: North Korean defectors, South Korean citizens, NGOs, government agencies, employers, and educational institutions.

Pervasiveness: BridgeNet is accessible in various regions across South Korea, covering urban and rural areas, and across different devices (e.g., smartphones, computers, tablets).

Time: The value scenario will consider the impact of BridgeNet on users and society over the next five to twenty years.

Systemic Effects: This scenario will explore the complex interactions between technology, culture, society, and psychology as BridgeNet becomes pervasive.

Value Implications: The scenario will consider both the positive effects of BridgeNet and its potential darker consequences.

Value scenario: "BridgeNet: Fostering Connections and Growth" Lee, a North Korean defector, has been using BridgeNet for three years. At first, she found a job as a secretary through the platform's job support services. Over time, she engaged in the mentorship program, where she was paired with a successful South Korean professional who shared her interests. Together, they worked on improving her skills, boosting her confidence, and setting career goals.

As the years went by, BridgeNet continued to expand its services, integrating more defectors and South Koreans into its community. The platform became a powerful tool for cultural exchange, helping to break down stereotypes and promote understanding. Many defectors found not only professional opportunities but also formed lasting friendships and romantic relationships through the social network.

However, as BridgeNet became more pervasive, some unintended consequences emerged. Some South Koreans began to feel threatened by the growing number of defectors integrating into their society, leading to a rise in nationalism and xenophobia. Additionally, the platform's increasing influence on users' lives raised concerns about privacy and surveillance.

Despite these darker implications, BridgeNet has undeniably helped many defectors, like Lee, find their footing in South Korea. Lee now has a stable job, a close group of friends, and has even started dating a fellow defector she met on the platform. Although she still misses her family in North Korea, she feels more connected and supported in her new home than ever before.

## RESULT #2:

Proposed new system: DefectorConnect is a comprehensive support system designed to address the challenges faced by North Korean defectors in their personal and professional lives. The platform uses AI technology to provide personalized recommendations for job opportunities, social events, and support groups. It also features a secure dating and friendship platform that matches users with compatible defectors and locals, promoting social integration and reducing feelings of loneliness and marginalization.

#### Five key elements:

- a. Stakeholders: Defectors, their families, potential employers, locals, and support organizations.
- b. Pervasiveness: The platform is accessible through various devices (smartphones, tablets, and computers) and spans across different cities and neighborhoods, targeting a broad range of defectors.
- c. Time: The system considers the long-term impact of social and professional integration of defectors into South Korean society.
- d. Systemic Effects: The platform evaluates the interaction between technology, society, and culture, addressing the unique challenges faced by defectors.
- e. Value Implications: The value scenario envisions both the positive and darker consequences of DefectorConnect, ensuring a balanced and comprehensive analysis of its effects.

Value scenario: Title: A Connected Future for Defectors

In a world where DefectorConnect has been widely adopted, Lee logs into the platform after a long day at work. As a secretary at a small company, she is proud of her professional achievements but longs for a deeper connection with others who share her background. Through DefectorConnect, she is able to join a support group for North Korean defectors who have transitioned into professional careers in South Korea. At the group's weekly meetings, she shares her experiences, learns from others, and forms meaningful relationships.

One day, Lee receives a notification from DefectorConnect about a match on the dating platform. She starts chatting with Kim, a fellow defector who has been in South Korea for six years. They quickly bond over their shared experiences, and after a few weeks, they decide to meet in person. Their connection blossoms into a romantic relationship, and Lee starts feeling less lonely and marginalized. However, not all is perfect in the world of DefectorConnect. As the platform grows in popularity, some users exploit its services to prey on vulnerable defectors. Scammers create fake profiles to deceive and take advantage of unsuspecting users. Additionally, there are concerns about privacy and the potential misuse of personal information by authorities or other organizations.

To address these issues, DefectorConnect invests in robust security measures and user verification processes, prioritizing the safety and well-being of its users. Furthermore, the platform actively collaborates with support organizations and law enforcement agencies to mitigate the risks associated with its darker consequences.