Breaking Barriers to Creative Expression: Co-Designing and Implementing an Accessible Text-to-Image Interface

Atieh Taheri[1], Mohammad Izadi[2], Gururaj Shriram[2], Negar Rostamzadeh[2], Shaun Kane[2]
[1]UC Santa Barbara, [2] Google
[1]a_taheri@ucsb.edu, [2]{izadi, gshriram, nrostamzadeh, shaunkane}@google.com

Abstract

Text-to-image generation models have grown in popularity due to their ability to produce high-quality images from a text prompt. One use for this technology is to enable the creation of more accessible art creation software. In this paper, we document the development of an alternative user interface that reduces the typing effort needed to enter image prompts by providing suggestions from a large language model, developed through iterative design and testing within the project team. The results of this testing demonstrate how generative text models can support the accessibility of text-to-image models, enabling users with a range of abilities to create visual art.

1 Introduction

Recent advancements in text-to-image (T2I) models have allowed users, including non-artists, to produce high-quality images from textual prompts [1, 2]. Notable models like DALL-E 2 [3], Midjourney [4], and Stable Diffusion [5] interpret text or text+image prompts to generate unique images. These models are significant for individuals with motor disabilities, offering an accessible image creation alternative to traditional methods [6]. Text input, adaptable through various means such as keyboards or voice, is central to their accessibility [7]. However, for optimal user experience, especially for those with motor disabilities, T2I tools must be designed with interactive features that cater to their specific needs [6]. In this context, we introduce *PromptAssist*, an accessible web-based interface for T2I prompts. It incorporates a large language model to provide prompt suggestions based on user input. PromptAssist's design stems from the first-hand experiences of its diverse team, most of them are individuals with motor disabilities, addressing interface accessibility issues. Its main features are: (1) auto-suggestions for prompts or additional details; (2) an inclusive text entry interface supporting text and pointer-based interactions; (3) a guided workflow for prompt creation.

2 PromptAssist: an accessible prompting interface for text-to-image models

PromptAssist was conceived after extensive testing of existing T2I models by researchers Taheri, Rostamzadeh, and Kane. Its objective is to simplify text-to-image prompt composition for users who find prolonged typing challenging. Drawing inspiration from tools like Promptomania's Prompt Builder [8], which aids users in generating creative prompts by suggesting image attributes, PromptAssist refines this approach (Related Work are included in AppendixA.) It emphasizes 1) diverse input methods for prompt creation and 2) minimized user input for prompt generation. Using generative text models, PromptAssist aids in text-to-image prompt creation. Unlike Promptomania's fixed prompt list, PromptAssist enhances user input. This system employs an internal transformer-based language model akin to GPT-3 [9]. The interface guides users through a structured prompt creation process. Users can: (1) select an image environment; (2) pick objects for the image; (3) detail scene

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

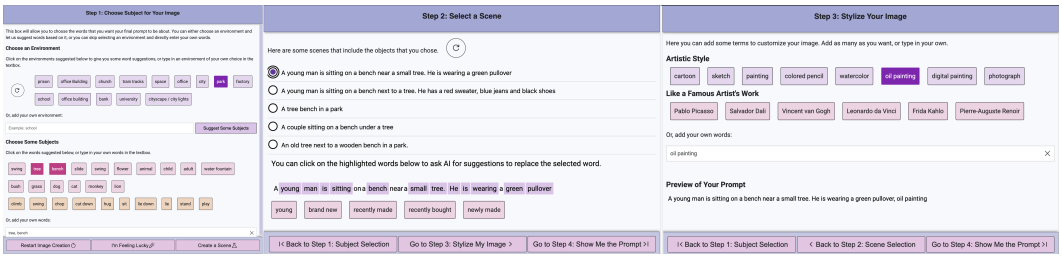

Figure 1: PromptAssist's interface. (Left) Step 1: users select an environment, subjects, and optionally actions. (Middle) Step 2: users choose and edit a scene. (Right) Step 3: users can add artistic styles.

activities; (4) select a scene that describes the intended prompt and optionally modify or replace words in the scene; (5) optionally specify the visual style. Accessibility is a priority; at each step, users can either type or select from a list of suggestions. Figure 1 provides an interaction demo. The language model prompts used in our prototype, as well as the resulting images generated by copying the resulting prompt into DALL-E, are provided in Appendix B.

PromptAssist was co-designed over several months with a team including researchers with motor disabilities. Led by Taheri, Rostamzadeh, and Kane, our diverse team spanned fields like computer science, human-computer interaction, and accessibility. Taheri's personal challenges, living with Spinal Muscular Atrophy, significantly influenced the project's direction. This lived experience enriched our mission to develop a more accessible T2I platform. The trio—Taheri, Rostamzadeh, and Kane—initially conceptualized PromptAssist, and its growth was marked by continuous testing and feedback. As the prototype matured, Izadi and Shriram, although not from a research background, joined to aid in its refinement through study sessions (more details in Appendix C) and to learn about accessibility research. Testing was primarily done using Parti [10], a T2I model akin to DALL-E.

Session 1: collaborative planning. Taheri outlined the project's scope and goals, emphasizing collaborative testing and feedback. A dedicated group chat room was established for seamless team communication.

Session 2: familiarization with existing T2I tools. Team members explored existing T2I tools, experimenting with different prompts and understanding their capabilities. The need for usability and accessibility enhancements was evident.

Session 3: testing PromptAssist Taheri introduced PromptAssist abd how it tackled previously identified issues. Team members interactively tested the prototype, which then demanded prompts to follow a specific sequence. Feedback centered on improving user interface, accessibility, and prompt flexibility. The addition of autocomplete and more prompt suggestions were also emphasized.

Session 4: testing the revised PromptAssist Taheri presented the revamped PromptAssist, addressing prior feedback. The team appreciated its enhanced navigation and the flexibility it offered.

3 Discussion and concluding remarks

PromptAssist arose from the need for a more accessible T2I tool, addressing challenges faced by individuals with disabilities in typing longer prompts on conventional platforms. Enhancing T2I accessibility, PromptAssist supports varied input methods, aligning with WCAG's *operability* principle. Instead of limiting users to predefined categories, it employs an LLM to provide contextual suggestions, enhancing both accessibility and broadening creative horizons. T2I models unlock unique creative potential, especially for those hindered by traditional tools' accessibility constraints. Despite their promise, there is room for refining their accessibility and usability. A challenge in PromptAssist's design was ensuring both creative freedom and user-friendliness. We found that while automated suggestions can assist, they risk overshadowing user creativity. Thus, PromptAssist empowers users by allowing them to modify or reject system suggestions, ensuring their creative intent remains paramount. PromptAssist blends accessibility with creative freedom. Its features, rooted in an LLM, offer potential for broader creative platforms. Future versions could integrate varied inputs, such as speech, and adapt based on prior outputs. Its features could benefit a range of creative platforms, supporting users from various backgrounds. Generative AI models, with user-centric interfaces, can revolutionize media creation. Future work should prioritize user collaboration, sharing, and combined human-AI artistic endeavors. For further detail please refer to Appendix D.

4 Ethical Implications

The development of more accessible text-to-image interfaces raises important ethical considerations. On one hand, enhancing accessibility of creative tools is an important goal that can promote inclusion for people with disabilities. However, reliance on imperfect AI systems also raises valid concerns around loss of autonomy and authorship. We aimed to mitigate this risk in PromptAssist's design - it provides prompt suggestions to users, but allows them to accept, modify or reject these ideas, reinforcing creative independence. Preserving user agency was a priority. Still, biases in the training data could lead to limiting diversity or promoting harmful stereotypes. As designers utilizing these powerful but imperfect AI systems, we must remain cognizant of these pitfalls, maximizing creative freedom while minimizing biased outcomes. Overall, we believe the benefits of increased access enabled by PromptAssist outweigh its potential risks. But we acknowledge our duty to proactively monitor its real-world impacts, making adjustments to address any issues around bias that arise. Our goal is to responsibly co-develop technologies that expand accessibility and empowerment for all. It is also important to note that the rights and credits of the artists should be in mind when images from specific styles or artists are generated [11, 12].

References

[1] Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, and Honglak Lee. Generative adversarial text to image synthesis. In *International conference on machine learning*, pages 1060–1069. PMLR, 2016.

[2] Minsuk Chang, Stefania Druga, Alex Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie Cai, and Michael Terry. The prompt artists. *arXiv preprint arXiv:2303.12253*, 2023.

[3] DALL·E 2 — openai.com. `https://openai.com/dall-e-2`, 2022. [Accessed 23-09-2023].

[4] Midjourney — midjourney.com. `https://www.midjourney.com`, 2022. [Accessed 23-09-2023].

[5] Stability AI — stability.ai. `https://stability.ai`, 2022. [Accessed 23-09-2023].

[6] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)*, 3(3):1–27, 2011.

[7] Heidi Horstmann Koester and Sajay Arthanat. Text entry rate of access interfaces used by people with physical disabilities: A systematic review. *Assistive Technology*, 30(3):151–163, 2018.

[8] Promptomania. Promptomania, 2023.

[9] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[10] Jiahui Yu, Yuanzhong Xu, Jing Yu Koh, Thang Luong, Gunjan Baid, Zirui Wang, Vijay Vasudevan, Alexander Ku, Yinfei Yang, Burcu Karagol Ayan, et al. Scaling autoregressive models for content-rich text-to-image generation. *arXiv preprint arXiv:2206.10789*, 2022.

[11] Harry H Jiang, Lauren Brown, Jessica Cheng, Mehtab Khan, Abhishek Gupta, Deja Workman, Alex Hanna, Johnathan Flowers, and Timnit Gebru. Ai art and its impact on artists. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 363–374, 2023.

[12] Negar Rostamzadeh, Emily Denton, and Linda Petrini. Ethics and creativity in computer vision. *arXiv preprint arXiv:2112.03111*, 2021.

[13] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S Sara Mahdavi, Rapha Gontijo Lopes, et al. Photorealistic text-to-image diffusion models with deep language understanding. *arXiv preprint arXiv:2205.11487*, 2022.

[14] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, 2022.

[15] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.

[16] Sri Sunarti, Ferry Fadzlul Rahman, Muhammad Naufal, Muhammad Risky, Kresna Febriyanto, and Rusni Masnina. Artificial intelligence in healthcare: opportunities and risk for future. *Gaceta Sanitaria*, 35:S67–S70, 2021.

[17] Xinge Guo, Tianyiyi He, Zixuan Zhang, Anxin Luo, Fei Wang, Eldwin J Ng, Yao Zhu, Huicong Liu, and Chengkuo Lee. Artificial intelligence-enabled caregiving walking stick powered by ultra-low-frequency human motion. *ACS nano*, 15(12):19054–19069, 2021.

[18] Francesco Aggogeri, Tadeusz Mikolajczyk, and James O'Kane. Robotics for rehabilitation of hand movement in stroke survivors. *Advances in Mechanical Engineering*, 11(4):1687814019841921, 2019.

[19] Shaun K. Kane, Anhong Guo, and Meredith Ringel Morris. Sense and accessibility: Understanding people with physical disabilities' experiences with sensing systems. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20, New York, NY, USA, 2020. Association for Computing Machinery.

[20] Katerina Zdravkova, Venera Krasniqi, Fisnik Dalipi, and Mexhid Ferati. Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective. *Frontiers in Artificial Intelligence*, page 240, 2022.

[21] Rohit Ranchal, Teresa Taber-Doughty, Yiren Guo, Keith Bain, Heather Martin, J Paul Robinson, and Bradley S Duerstock. Using speech recognition for real-time captioning and lecture transcription in the classroom. *IEEE Transactions on Learning Technologies*, 6(4):299–311, 2013.

[22] Jim Fruchterman and Joan Mellea. Expanding employment success for people with disabilities. *Benetech, November*, 2018.

[23] Maarten Buyl, Christina Cociancig, Cristina Frattone, and Nele Roekens. Tackling algorithmic disability discrimination in the hiring process: An ethical, legal and technical analysis. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, page 1071–1082, New York, NY, USA, 2022. Association for Computing Machinery.

[24] Meredith Whittaker, Meryl Alper, Cynthia L Bennett, Sara Hendren, Liz Kaziunas, Mara Mills, Meredith Ringel Morris, Joy Rankin, Emily Rogers, Marcel Salas, et al. Disability, bias, and ai. *AI Now Institute*, page 8, 2019.

[25] Alan Foley and Beth A Ferri. Technology for people, not disabilities: Ensuring access and inclusion. *Journal of Research in Special Educational Needs*, 12(4):192–200, 2012.

[26] Pilar Orero and Anna Matamala. Accessible opera: Overcoming linguistic and sensorial barriers. *Perspectives: Studies in Translatology*, 15(4):262–277, 2007.

[27] Atieh Taheri, Ziv Weissman, and Misha Sra. Exploratory design of a hands-free video game controller for a quadriplegic individual. In *Augmented Humans Conference 2021*, pages 131–140, 2021.

[28] Domen Novak and Robert Riener. A survey of sensor fusion methods in wearable robotics. *Robotics and Autonomous Systems*, 73:155–170, 2015.

[29] Joan Lobo-Prat, Peter N Kooren, Arno HA Stienen, Just L Herder, Bart FJM Koopman, and Peter H Veltink. Non-invasive control interfaces for intention detection in active movement-assistive devices. *Journal of neuroengineering and rehabilitation*, 11(1):1–22, 2014.

[30] Abdulhamit Subasi, Saeed Mian Qaisar, et al. The ensemble machine learning-based classification of motor imagery tasks in brain-computer interface. *Journal of Healthcare Engineering*, 2021, 2021.

[31] Ricardo Chavarriaga, Melanie Fried-Oken, Sonja Kleih, Fabien Lotte, and Reinhold Scherer. Heading for new shores! overcoming pitfalls in bci design. *Brain-Computer Interfaces*, 4(1-2):60–73, 2017.

[32] Zhang Yi, Dai Lingling, Luo Yuan, and Huosheng Hu. Design of a surface emg based human-machine interface for an intelligent wheelchair. In *IEEE 2011 10th International Conference on Electronic Measurement & Instruments*, volume 3, pages 132–136. IEEE, 2011.

[33] Atieh Taheri, Ziv Weissman, and Misha Sra. Design and evaluation of a hands-free video game controller for individuals with motor impairments. *Frontiers in Computer Science*, 3:751455, 2021.

[34] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun K. Kane. "the less i type, the better": How ai language models can enhance or impede communication for aac users. CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.

[35] Monica Murero, Salvatore Vita, Andrea Mennitto, and Giuseppe D'Ancona. Artificial intelligence for severe speech impairment: innovative approaches to aac and communication. In *PSYCHOBIT*, 2020.

[36] Satoshi Ikeda, Hiroko Fukuda Siddiqi, Mayuko Mori, Hiromitsu Kawajiri, Misa Hirasawa, Takashi Kawaguchi, and Kaoru Yasuda. An online art project based on the affirmative model of disability in japan. *International Journal of Art & Design Education*, 41(4):532–546, 2022.

[37] Adrian Zbiciak and Tymon Markiewicz. A new extraordinary means of appeal in the polish criminal procedure: the basic principles of a fair trial and a complaint against a cassatory judgment. *Access to Justice in Eastern Europe*, 6(2):1–18, March 2023.

[38] Adrian Zbiciak and Tymon Markiewicz. A new extraordinary means of appeal in the polish criminal procedure: the basic principles of a fair trial and a complaint against a cassatory judgment. *Access to Justice in Eastern Europe*, 6(2):1–18, March 2023.

[39] Audrey Aldridge and Cindy L Bethel. A systematic review of the use of art in virtual reality. *Electronics*, 10(18):2314, 2021.

[40] Dazed. How AI could increase art world accessibility for disabled artists, April 2023.

[41] Chris Creed. Assistive technology for disabled visual artists: exploring the impact of digital technologies on artistic practice. *Disability & Society*, 33(7):1103–1119, 2018.

[42] Faye Ginsburg. Disability in the digital age. In *Digital anthropology*, pages 101–126. Routledge, 2020.

[43] Evan D Rapoport, Erin M Nishimura, Jonathan R Zadra, Peter M Wubbels, Dennis R Proffitt, Traci H Downs, and J Hunter Downs III. Engaging, non-invasive brain-computer interfaces (bcis) for improving training effectiveness & enabling creative expression. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 591–594. SAGE Publications Sage CA: Los Angeles, CA, 2008.

[44] Rosie Flewitt, Natalia Kucirkova, and David Messer. Touching the virtual, touching the real: ipads and enabling literacy for students experiencing disability. *Australian Journal of Language and Literacy, The*, 37(2):107–116, 2014.

[45] Dominique Burger. Improved access to computers for the visually handicapped: New prospects and principles. *IEEE transactions on rehabilitation engineering*, 2(3):111–118, 1994.

[46] George H Williams. Disability, universal design, and the digital humanities. *Debates in the digital humanities*, pages 202–212, 2012.

[47] Davide Mulfari, Antonio Celesti, and Massimo Villari. A computer system architecture providing a user-friendly man machine interface for accessing assistive technology in cloud computing. *Journal of Systems and Software*, 100:129–138, 2015.

[48] Pauline Oliveros Foundation et al. Adaptive use musical instruments (aumi), 2007. https://aumiapp.com/.

[49] William Payne, Ann Paradiso, and Shaun K. Kane. Cyclops: Designing an eye-controlled instrument for accessibility and flexible use. In *NIME'20*, July 2020.

[50] Zacharias Vamvakousis and Rafael Ramirez. The eyeharp: A gaze-controlled digital musical instrument. *Frontiers in psychology*, 7:906, 2016.

[51] Microsoft Research. Microsoft hands-free music - microsoft research — microsoft.com, March 2018.

[52] Anthony Hornof, Anna Cavender, and Rob Hoselton. Eyedraw: a system for drawing pictures with eye movements. *ACM SIGACCESS Accessibility and Computing*, (77-78):86–93, 2003.

[53] Microsoft Research. Eyewriter — eyewriter.org, 2010.

[54] Susumu Harada, Jacob O Wobbrock, and James A Landay. Voicedraw: a hands-free voice-driven drawing application for people with motor impairments. In *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, pages 27–34, 2007.

[55] Ellen Jiang, Kristen Olson, Edwin Toh, Alejandra Molina, Aaron Donsbach, Michael Terry, and Carrie J Cai. Promptmaker: Prompt-based prototyping with large language models. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–8, 2022.

A Related Work

A.1 Text-to-image generation models

Recent years have seen a great expansion in the capabilities of T2I models, resulting in more detailed and realistic images. Key models include Reed et al.'s Generative Adversarial T2I Synthesis model [1] which employs conditional generative adversarial networks (cGANs) to create images from text, the Imagen [13], a T2I diffusion model and Parti [10], which uses sequence-to-sequence techniques and ViT-VQGAN tokenization to generate photorealistic images.

Recently, much attention has been drawn to models that are available to users, often through public betas, include OpenAI's DALL-E 2 [14], which creates high-resolution images from text using 650 million image-text pairs, Midjourney, an open beta tool popular among artists available as a Discord chatbot, and Stable Diffusion [15], an open-source model operable on users' own devices. Research, like that by Chang et al. [2], explored how users create and share prompts within online communities. Our work provides some early insights into the potential use of T2I models by individuals with disabilities.

A.2 Improving accessibility with AI

Significant strides in AI have enhanced assistive technologies for individuals with motor disabilities, streamlining healthcare [16], rehabilitation [17, 18, 19], education [20, 21], and employment [22, 23, 24], novel engagement modes in digital entertainment and video gaming [25, 26, 27] through machine learning, computer vision, and natural language processing [19, 28, 29, 30, 31, 32, 33]. Notably, language model advances have revolutionized augmentative and alternative communication (AAC) options for non-verbal individuals [34], and systems like Voiceitt have amplified speech capabilities for those with severe speech impairments [35].

The art world has also seen a growing integration of AI to facilitate creation for artists with disabilities, such as AI working in tandem with prosthetics and bionics to improve accessibility [36, 37, 38, 39, 40]. In this work, we focus on T2I tools—a groundbreaking AI-driven medium for artistic visual generation. By leveraging LLMs, we aim to boost interface accessibility, hence extending their utility to a wider audience. To our knowledge, the use of LLMs in visual creation via T2I models has not yet been explored.

A.3 Accessibility and creative tools

Computers, even general-purpose software can enable art creation for people with disabilities [41]. Researchers have explored innovative interfaces to facilitate creative tasks [42, 43, 44]. Accessible creative tools emphasize simplified interfaces and designs catering to users with diverse abilities [45, 46, 47].

Notably, work has been done to make music composition accessible with tools like AUMI [48], Cyclops [49], EyeHarp [50], and Hands-Free Music [51] using diverse inputs. Several projects have facilitated visual art creation for those with motor disabilities. EyeDraw [52], EyeWriter [53] uses eye movements, and VoiceDraw [54], uses non-speech vocalizations for drawing. Our work introduces an accessible visual art creation approach utilizing generative AI models. Unlike previous methods requiring learning, our system uses LLMs to generate creative text with reduced click or keystroke efforts, relying on existing interaction methods.

B LLM prompts and T2I output images

This section includes the prompts that were used to generate suggestions in the PromptAssist prototype. These prompts evolved as the prototype developed; this appendix includes the prompts used in the final round of testing and Figure 2 shows the resulting images after copying the resulting prompt from PromptAssist into DALL-E at the end of this section.

These prompts use a system inspired by PromptMaker [55]; the prompt is passed to a large language model, and the output is returned to PromptAssist. In these examples, variable input to the prompt is represented by the symbol **<input>** and output from the language model is represented by the symbol **<output>**.

B.1 Suggest environments

```
Name: environment
Suggestion: university
Name: environment
Suggestion: ocean
Name: environment
Suggestion: hospital
Name: environment
Suggestion: <output>
```

B.2 Suggest subjects for an environment

```
Environment: school
Suggestions: blackboard, teacher, chair, book, student, class, eraser, whiteboard,
             notebook, pen, pencil, eraser, paper

Environment: work office
Suggestions: desk, computer, pen, paper sheet, folder, fax, phone, pencil, paper
             shredder, light

Environment: forest
Suggestions: tree, animal, bird, monkey, lion, fox, eagle, plant, flower, insect,
             tiger, horse, wolf

Environment: home
Suggestions: TV, bed, table, sofa, couch, light, console table, remote control,
             carpet, rug, room, kitchen

Environment: sea
Suggestions: fish, jelly fish, star fish, shark, dolphin, whale, island, boat,
             ship, coral, crab
```

```
Environment: <input>
Suggestions: <output>
```

B.3 Suggest actions for subjects

```
word: tv
verbs: watch, work, fix, turn on, turn off, turn up, turn down, put, pick up

word: cat
verbs: play, eat, sit, run jump, scratch, pet, sleep, feed, jump, meow, brush,
       groom, bathe, cuddle, love

word: pool
verbs: fill, swim, drawn, go down, dive, jump, splash, play, go down, drink,
       eat, throw, throw up, spit, pee, pee in

word: paper
verbs: write, read, draw, cut, tear, color, crumple, make, throw, pick up,
       put down, fold, take, give, put away, color, spin

word: <input>
verbs: <output>
```

B.4 Generate prompts from scenes and subjects

```
words: dog
scene: A small dachshund doing a kickflip on a skateboard

words: cat, tree
scene: A young DSH cat sitting on a small tree branch with leaves in a garden

words: space, rocket
scene: A black and white space ship with a rocket attached to the engine.
       There is a trail of smoke following the rocket flying through
       the space full of planets and stars

words: chair, cup, parrot
scene: A parrot with a blue feather and a black/grey beak sitting on a
       high-backed chair. There is a small table next to the chair and a red
       cup of tea on the table

words: fish, waves
scene: A red snapper fish with a white body and black eyes and mouth is swimming
       through the waves at the ocean

words: flying car, fly, bird
scene: a flying car is driving in the sky and a bird is flying next to the car

words: paper sheet, fold, folder
scene: a paper sheet is being folded on a blue folder with a white paperclip on
       it

words: sofa, sleep
scene: A man is sleeping on the sofa. He has a red t-shirt, blue shorts and brown
       hair,

words: monkey, swing, tree
scene: A small monkey is swinging on a tree branch. There is a red banana next
       to the branch
```

```
words: take blood from
scene: A doctor is taking blood from a syringe and putting it in a small tube,

words: eiffel tower, old man, old woman
scene: An old man and an old woman are drinking wine next to the eiffel tower.,

words: cat, mouse
scene: An outdoor cat patiently waiting by a mouse hole for its next meal.

words: <input>
scene: <output>
```

B.5 Generate synonyms for words

```
word: blue
replacements red, pink, orange, yellow, purple, green, brown
word: small
replacements big, tiny, giant, medium, large, huge, miniature
word: young
replacements old, adult, child, teenager, infant, baby, middle-aged
word: <input>
replacements <output>
```

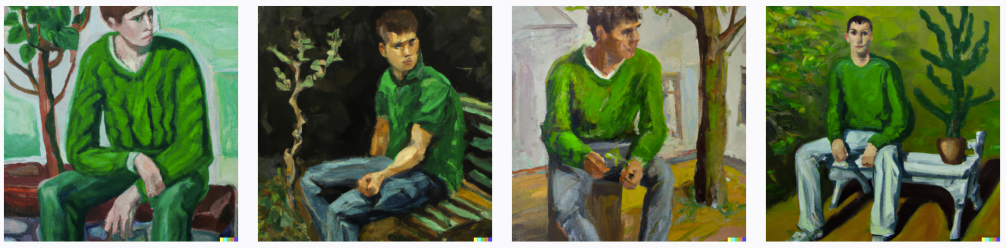

Figure 2: DALL-E 2 output for prompt "A young man is sitting on a bench near a small tree. He is wearing a green pullover, oil painting." This prompt was created by the authors using suggestions from PromptAssist.

C Iterative development of PromptAssist

Taheri, Rostamzadeh, and Kane conceived the initial concept for PromptAssist, and developed its early iterations. They refined this prototype through testing and recurring group discussion. Both Taheri and Kane used their own personal accessibility tools during testing. Once the PromptAssist prototype matured that could be more easily tested, they recruited Izadi and Shriram via an internal disability interest group within their organization. While Izadi and Shriram did not have a background in research, they joined the project because of their interest in shaping development of more accessible T2I technologies and to gain experience in accessibility research.

Taheri, Izadi, Shriram, and Kane were present at every test session, while Taheri generally served as the session facilitator. The research team tested prototype of PromptAssist on their own devices, but generally tested their prompts using an internal version of Parti [10], an autoregressive text-to-image model that produces images comparable to models such as DALL-E (although it is worth noting that prompts generated by PromptAssist can be copied and pasted into any T2I model).

C.1 Test session 1: collaborative planning and discussion

This session functioned as an informative gathering. Taheri detailed the research goals and future plans, particularly, guiding less experienced members, Izadi and Shriram, about their roles in iterative

testing and feedback. A group chat room was created for the research duration to facilitate questions and image sharing among the team.

C.2 Test session 2: experimenting with existing T2I tools

In this session, the team lead familiarized the team with existing T2I tools, encouraging them to experiment with the current Parti user interface, which consisted of only a text box and submit button. The team explored the tools' potential and boundaries by generating images based on varying concepts, styles, and detail levels. Izadi, for example, focused on an idea of a bird with a corgi's face flying over San Francisco. Shriram experimented with multiple prompts to understand the safety filter's operation. Kane used different color terms to observe their impact on cat portrait images. The team members communicated their experiences throughout the session in a "think-aloud" format, and Taheri and Kane documented the proceedings.

C.2.1 Feedback on current T2I systems

The team members, particularly those new to T2I systems, were impressed by the output quality but agreed on the need for usability and accessibility improvements. Issues identified included the absence of access keys or keyboard shortcuts, the inability to cancel the slow image generation process, and the lack of autocomplete or grammar correction. The possibility of benefiting from pre-made prompt categories or ideas was also discussed. One theme that arose during testing was difficulty in generating longer prompts or iterating on prompts. Shriram, in particular, found it difficult to type longer or more prompts. The team discussed how to refine prompts using natural language instructions like "make this more descriptive." Izadi expressed interest in using natural language to edit or merge images generated by the model.

C.3 Test session 3: experimenting with PromptAssist

In this collaborative session led by Taheri, the team tested the PromptAssist prototype. Taheri explained how PromptAssist addressed the usability and accessibility issues identified in the previous T2I systems session. Team members worked together and shared feedback via a "think-aloud".

This version of PromptAssist, while similar to the one described above, required users to compose their prompt in a specific order, without skipping steps.

C.3.1 Feedback from the session

Shriram spent much of his time generating images related to the ocean. He was particularly interested in what subjects were added to the image, noting that requesting ocean images tended to include plants but not animals. Izadi noted that the user interface did not quite match their goals; while they appreciated that PromptAssist offered suggestions, they wanted to be able to start with their own specific prompt and build upon it. While Izadi noted that selecting parts of prompts using the mouse reduces the need for typing, they desired the ability to do both.

The team suggested a variety of user interface changes, including: reducing white space in the interface; increasing color contrast; and making it easier to go back or restart a prompt. This early version usually provided 3-5 suggestions in each category; participants requested that ten or more suggestions would be useful, especially when coming up with their initial idea for the prompt. As discussed in the first session, adding autocomplete would be helpful so that users could enter a partial prompt and let the model finish it.

C.4 Test Session 4: revised PromptAssist

Following the previous session, Taheri updated PromptAssist per user feedback. Enhancements included: adding a multi-page layout, improving color contrast, writing human-readable error messages, increasing the number of suggestions, allowing users to generate additional suggestions, adding the option to skip steps like artistic style, and enabling keyboard-only navigation of all buttons and menus.

As in the previous session, Taheri walked the team through the updates made in response to their feedback. The team members tested the enhanced PromptAssist, sharing their thoughts. They all

found that the changes made it easier to navigate through the process and to follow their own creative ideas. The revised prototype's notable advancement was its flexibility in allowing user-defined creative processes over prescribed sequences. In contrast to the prior version, which enforced a specific prompt format, the allowed users to enter their own prompts, accessing assistance only when needed.

D Detailed discussion

D.1 Improving usability and accessibility of T2I tools

This paper documents the co-design and evolution of an accessible prompt creation tool for T2I systems, which was designed in response to feedback from disabled team members who found typing prompts, particularly longer ones, challenging in the existing text-based interface, which was essentially just an HTML text input control.

The creation of PromptAssist provides ways to enhance T2I interface accessibility. It facilitates prompt creation through typing, pointing, clicking, or a combination, adhering to the WCAG's concept of *operability*—allowing any input device to perform all actions. Unlike existing tools that limit input to predefined categories, PromptAssist uses an LLM to offer contextual suggestions, allowing users to expand their own ideas, rather than choosing from a preset list of image ideas. In addition to enabling more accessible forms of input, these contextual suggestions could be beneficial for all users regardless of their abilities.

D.2 Unleashing creativity with T2I tools

It is clear that T2I models can enhance the creative abilities of people, enabling them to create images that they would be unable to create by other means. T2I systems may be especially empowering to individuals who are unable to effectively use other image creation tools because of accessibility barriers. The focus on text input in current T2I models is beneficial in some ways, as many people with disabilities have already found accessible ways to input text. However, as shown in this work, there remains the opportunity to increase accessibility and ease of use of these systems.

D.3 Creativity *vs.* ease of use

One tension that arose in designing PromptAssist's user interface was balancing creative flexibility and ease of use. While pre-generated contextual prompt ideas can help users who find typing challenging, it risks limiting creativity and user autonomy. Interviews with experienced T2I users revealed that prompt crafting is viewed as part of their creative work [2]. Over-reliance on language generation could also diminish the perceived independence of a user with disability, giving the impression that the system, rather than the user, is doing the work [34].

To mitigate concerns about creativity and autonomy, PromptAssist enables users to view system suggestions, which they can then accept, modify, or reject. This approach amplifies their original ideas without making creative choices for them, an aspect that remains crucial in the development of accessible creative tools.