Expandora: Broadening Design Exploration with Text-to-Image Model

ANONYMOUS AUTHOR(S)

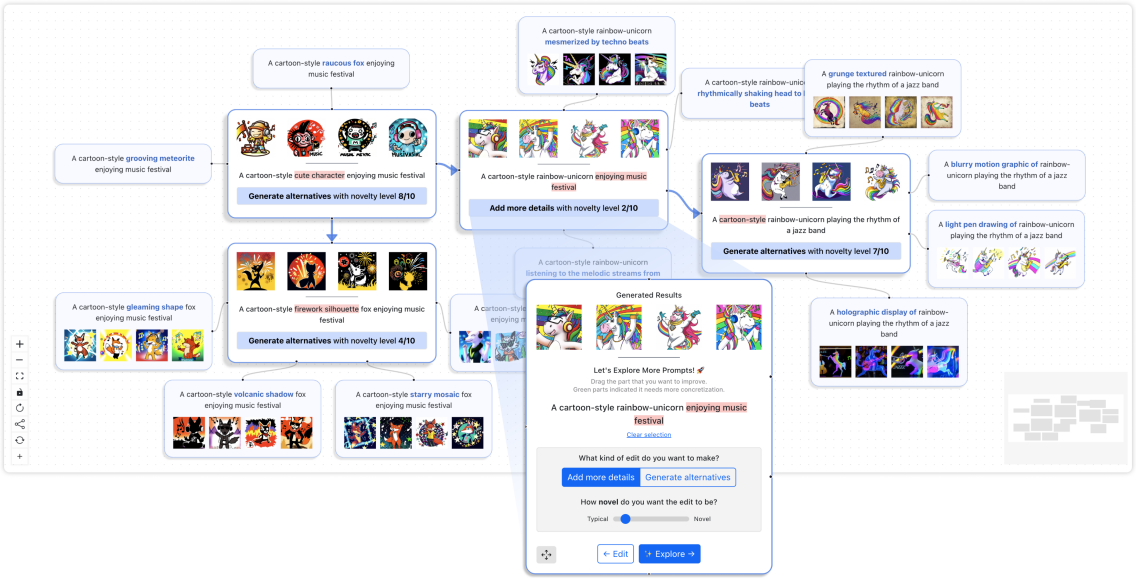

Fig. 1. Interface of Expandora. The figure illustrates the two key components of Expandora: (1) Structured Input Interface: Users can specify exploratory intentions by selecting parts of a prompt to refine, choosing between adding details or generating alternatives, and adjusting the desired novelty level. (2) Mindmap-like Interface Showing Results: Exploration results are visualized as a branching, non-linear structure that reflects iterative workflows, allowing users to revisit, refine, or branch off ideas.

Broad exploration of references is critical in the visual design process. While text-to-image (T2I) models offer efficiency and customization of exploration, they often limit support for divergence in exploration. Despite their potential, current T2I models often make broad exploration challenging, as designers lack intuitive ways to articulate exploratory intentions and manage iterative, non-linear workflows. To address these challenges, we developed Expandora. Users can specify their exploratory intentions and desired diversity levels through structured input, and using an LLM-based pipeline, Expandora generates tailored prompt variations. The results are displayed in a mindmap-like interface that encourages non-linear workflows. Beyond enhancing divergent exploration, we discuss how structured input can also support convergent thinking and how non-linear interfaces can be further applied to generative AI systems to improve creative workflows.

CCS Concepts: • **Human-centered computing** → **Interactive systems and tools**.

Additional Key Words and Phrases: Creativity Supporting Tool, Design Exploration, Text-to-Image Model, Generative AI

Manuscript submitted to ACM

ACM Reference Format:
Anonymous Author(s). 2018. Expandora: Broadening Design Exploration with Text-to-Image Model. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Exploration is crucial in the creative process [5], helping designers understand the problem space and generate novel ideas [6, 11, 12, 14]. While reference websites (e.g., Pinterest) have traditionally supported this, text-to-image (T2I) models now offer an alternative, allowing designers to generate tailored visuals quickly. Designers also expect to see unexpected outputs to help them think creatively through the T2I model [8]. However, these models often limit *divergence*, producing predictable images or causing fixation on initial results [8, 17]. Prior research explored interfaces that facilitate seamless connections between visual and semantic data [7, 9]. With generative models (e.g., LLMs, diffusion models), tools like CreativeConnect [4] and PopBlends [18] help designers decompose and recombine ideas, while Luminate [16] aids in avoiding fixation. Also, to support users' expressing intent through specification (refining ideas) and diversification (exploring broad possibilities), recent work has shown how to utilize the generative model's ability as an expression channel of a user's intent [15]. For input methods, various generative model-based tools [13, 19, 20] support structured or tailored inputs beyond natural language input to help users express their creative intention. Additionally, techniques for refining and suggesting prompts further enhance exploration [1, 2, 15].

In this paper, we propose Expandora, a system that enhances T2I-assisted exploration by generating semantically diverse prompts. We focused on two challenges of using the T2I model for design exploration that have not been investigated yet: (1) difficulties in articulating *exploratory intentions* in prompts, such as specifying variation and consistency, and (2) the linear nature of T2I interfaces, which hinders branching, merging, and revisiting ideas. It features a structured input interface for specifying exploration parameters and an LLM-powered pipeline that generates tailored prompt variations. Results are displayed in a mindmap-like interface, allowing users to navigate ideas visually. A comparative study with eight familiar with graphic design and the T2I model showed that Expandora facilitated more diverse exploration and increased prompt generation within the same time compared to ChatGPT's image generation. Participants reported higher satisfaction but highlighted the need for convergent ideation support to refine ideas. We discuss improvements to address this gap and better support creative workflows.

2 Expandora

2.1 Design Goals

Expandora addresses two key challenges observed in how designers interact with T2I models.

2.1.1 Support Input for Exploratory Intention. First, they **struggled to express *exploratory* intentions accurately through text prompts**. Design exploration often begins with an open-ended mindset, where users have broad ideas and seek to explore diverse possibilities through generated images. In such cases, users exhibit *exploratory intentions*, focusing on *how much and what kind of diversity* they want in the outputs. For example, one participant shared: *"When I typed 'a drawing of a soft cloud,' I wanted the shape of the cloud to vary, but I wanted the drawing style to remain consistent with the previous one. However, I couldn't specify these preferences to the model."* Current interfaces for T2I models lack inputs to deliver such exploratory intentions. As a result, users are often forced to either specify their prompts before fully exploring diverse ideas or use vague prompts and rely on the model's randomness, which may not align with

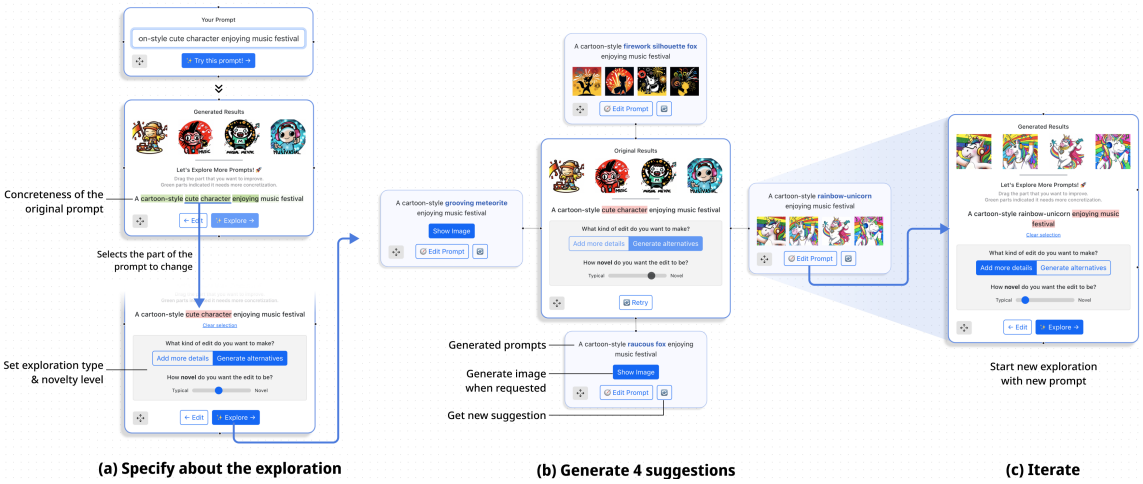

Fig. 2. Usage flow of Expandora. (a) The user begins by entering a prompt, and the system generates four initial images. Additionally, the system highlights the concreteness of each word in the prompt. The system also provides tools for further diversification, allowing users to select a specific part of the prompt to refine, choose how to diversify it, and adjust the desired level of novelty. (b) Based on the input, the system suggests four modified prompts that refine the specified part. Users can either select a prompt to generate new images or request additional suggestions. (c) For further exploration, users can click "Edit Prompt" to return to the interface in (a) and continue diversifying, enabling iterative and targeted exploration.

their creative goals. **Design Goal 1: Enable users to specify which elements to explore, and provide controls to define the direction and range of the diversity of the T2I outputs**

2.1.2 Facilitate Iterative and Non-Linear Workflows. Second, they **struggled to engage in iterative design ideation using linear interfaces in current T2I systems**. Design ideation is inherently non-linear, involving branching, merging, and revisiting ideas. However, current T2I systems present results sequentially, making iteration cumbersome. P3 shared: *"I usually repeat exploration and then select parts to refine. But with generative models, it felt like I was going back to square one with every generation."* **Design Goal 2: Support iterative exploration by enabling users to explore on top of previous exploration, revisit exploration history, or branch ideas seamlessly.**

2.2 Features

2.2.1 Structured Input for Exploration Intention. As shown in Figure 2 (a), the workflow begins with users entering an initial prompt, generating four images. To guide exploration, the system analyzes the word concreteness of the prompt, highlighting less concrete terms in green to indicate areas needing exploration. This draws on prior research linking word concreteness to visual representations [10]. Referring to this, users can then specify which part of the prompt to explore by dragging over it and selecting either "Add Details" or "Generate Alternatives." They can also adjust the novelty level, with lower settings maintaining similarity to the original prompt and higher settings producing more diverse ideas. This structured input helps users articulate both the direction and breadth of their exploration (DG1).

2.2.2 System Suggestions for the Expanded Prompts. Based on user settings, the system generates four modified prompt suggestions in a mindmap-like layout (Figure 2 (b)). Users can preview and generate images by clicking "Show Image."

If a suggestion is unsatisfactory, clicking "New Suggestion" recalculates alternatives using an adaptive algorithm (Section 2.3), ensuring better alignment with user preferences.

2.2.3 Iterative Exploration Workflow. Users can refine ideas iteratively by selecting "Edit Prompt" on any suggestion, setting new exploration preferences, and generating additional suggestions. This expands the design space step by step. The exploration history is visualized in a mindmap-like interface (Figure 1), allowing users to branch into multiple ideas or revisit previous prompts at any point. This non-linear workflow supports flexible, iterative ideation, mirroring natural design processes (DG2).

2.3 Technical Detail

Expandora is a ReactJS-based web application using React Flow for visualization and OpenAI's DALL·E2 API for image generation. For **concreteness analysis**, we trained a BERT-based model on a 40k-word dataset [3], achieving a loss of 0.042 and MAE of 0.346, with words highlighted in green according to abstraction level. For **prompt suggestion**, GPT-4o generates 200 prompts based on user preferences, either refining details or creating alternatives (Appendix-B.1, B.2). To **filter initial suggestions**, prompts are evaluated for semantic similarity to the original, ensuring alignment with the user's novelty setting, followed by K-means clustering to select four diverse and representative prompts. If a user discards a suggestion, the system dynamically generates a replacement that maximizes novelty while maintaining diversity, ensuring a continuous and adaptive exploration process.

3 Discussion

3.1 Structured Input for Divergence and Convergence in Exploration

Divergent thinking requires efficiently exploring a broad range of possibilities. Expandora's structured input, designed for adding detail and generating alternatives with a novelty slider, facilitates this process. This could be extended to other creative domains, such as storytelling or songwriting, by adapting structured inputs to different modalities.

However, creative exploration also involves convergence—refining and detailing ideas—which Expandora does not fully support. Users might struggle to steer outputs as desired during this phase, highlighting the need for improved intent expression. Future systems should integrate free-form inputs alongside structured options, allowing users to specify keywords or constraints to refine outputs more effectively. Additionally, enabling users to merge multiple ideas or refine outputs dynamically on the nonlinear interface of Expandora based on their mode of exploration could enhance the convergence process.

3.2 Applying Non-Linear Interfaces to T2I Systems

The non-linear interface of Expandora aligns well with the iterative nature of design ideation, helping users reflect on and expand their exploration paths. This structure fosters a sense of control and coherence, making exploration more effective. A simple adaptation for other generative AI tools could be displaying generation history as a graph or tree, allowing users to revisit and compare iterations more intuitively. Beyond visualization, the non-linear interface provides insights into how users evolve their ideas over time. By analyzing patterns in user modifications and preferences, systems could learn individual exploration styles and offer more personalized suggestions. Future enhancements could even allow AI to simulate exploration on behalf of users, streamlining the creative process and providing smarter, context-aware support.

References

[1] Shm Garanganao Almeda, J.D. Zamfirescu-Pereira, Kyu Won Kim, Pradeep Mani Rathnam, and Bjoern Hartmann. 2024. Prompting for Discovery: Flexible Sense-Making for AI Art-Making with Dreamsheets. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 160, 17 pages. doi:10.1145/3613904.3642858

[2] Stephen Brade, Bryan Wang, Mauricio Sousa, Sageev Oore, and Tovi Grossman. 2023. Promptify: Text-to-Image Generation through Interactive Prompt Exploration with Large Language Models. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 96, 14 pages. doi:10.1145/3586183.3606725

[3] Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods* 46 (2014), 904–911.

[4] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1055, 25 pages. doi:10.1145/3613904.3642794

[5] Claudia Eckert and Martin Stacey. 2000. Sources of inspiration: a language of design. *Design studies* 21, 5 (2000), 523–538.

[6] Scarlett R. Herring, Chia-Chen Chang, Jesse Krantzler, and Brian P. Bailey. 2009. Getting Inspired! Understanding How and Why Examples Are Used in Creative Design Practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 87–96. doi:10.1145/1518701.1518717

[7] Youwen Kang, Zhida Sun, Sitong Wang, Zeyu Huang, Ziming Wu, and Xiaojuan Ma. 2021. MetaMap: Supporting Visual Metaphor Ideation through Multi-dimensional Example-based Exploration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 427, 15 pages. doi:10.1145/3411764.3445325

[8] Hyung-Kwon Ko, Gwanmo Park, Hyeon Jeon, Jaemin Jo, Juho Kim, and Jinwook Seo. 2023. Large-Scale Text-to-Image Generation Models for Visual Artists' Creative Works. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (Sydney, NSW, Australia) *(IUI '23)*. Association for Computing Machinery, New York, NY, USA, 919–933. doi:10.1145/3581641.3584078

[9] Janin Koch, Nicolas Taffin, Andrés Lucero, and Wendy E. Mackay. 2020. SemanticCollage: Enriching Digital Mood Board Design with Semantic Labels. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 407–418. doi:10.1145/3357236.3395494

[10] Mackenzie Leake, Hijung Valentina Shin, Joy O. Kim, and Maneesh Agrawala. 2020. Generating Audio-Visual Slideshows from Text Articles Using Word Concreteness. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3313831.3376519

[11] Brian Lee, Savil Srivastava, Ranjitha Kumar, Ronen Brafman, and Scott R. Klemmer. 2010. Designing with Interactive Example Galleries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2257–2266. doi:10.1145/1753326.1753667

[12] Felix Müller-Wienbergen, Oliver Müller, Stefan Seidel, and Jörg Becker. 2011. Leaving the beaten tracks in creative work–A design theory for systems that support convergent and divergent thinking. *Journal of the Association for Information Systems* 12, 11 (2011), 2.

[13] Xiaohan Peng, Janin Koch, and Wendy E. Mackay. 2024. DesignPrompt: Using Multimodal Interaction for Design Exploration with Generative AI. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 804–818. doi:10.1145/3643834.3661588

[14] Daniel Ritchie, Ankita Arvind Kejriwal, and Scott R. Klemmer. 2011. D.Tour: Style-Based Exploration of Design Example Galleries. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 165–174. doi:10.1145/2047196.2047216

[15] Kihoon Son, DaEun Choi, Tae Soo Kim, Young-Ho Kim, and Juho Kim. 2024. GenQuery: Supporting Expressive Visual Search with Generative Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 180, 19 pages. doi:10.1145/3613904.3642847

[16] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 644, 26 pages. doi:10.1145/3613904.3642400

[17] Samangi Wadinambiarachchi, Ryan M. Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 380, 18 pages. doi:10.1145/3613904.3642919

[18] Sitong Wang, Savvas Petridis, Taeahn Kwon, Xiaojuan Ma, and Lydia B Chilton. 2023. PopBlends: Strategies for Conceptual Blending with Large Language Models. 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 435, 19 pages. doi:10.1145/3544548.3580948

[19] Zhijie Wang, Yuheng Huang, Da Song, Lei Ma, and Tianyi Zhang. 2024. PromptCharm: Text-to-Image Generation through Multi-modal Prompting and Refinement. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 185, 21 pages. doi:10.1145/3613904.3642803

[20] Mingxu Zhou, Dengming Zhang, Weitao You, Ziqi Yu, Yifei Wu, Chenghao Pan, Huiting Liu, Tianyu Lao, and Pei Chen. 2024. StyleFactory: Towards Better Style Alignment in Image Creation through Style-Strength-Based Control and Evaluation. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 107, 15 pages. doi:10.1145/3654777.3676370

A Examples of Generated Images from Two Conditions

Figure 3 shows examples of images generated by two participants using Expandora and the baseline, which was chatting with ChatGPT and its image generation feature. In Expandora condition, users produced a wider variety of creative concepts, demonstrating the system's ability to support diverse and iterative exploration.

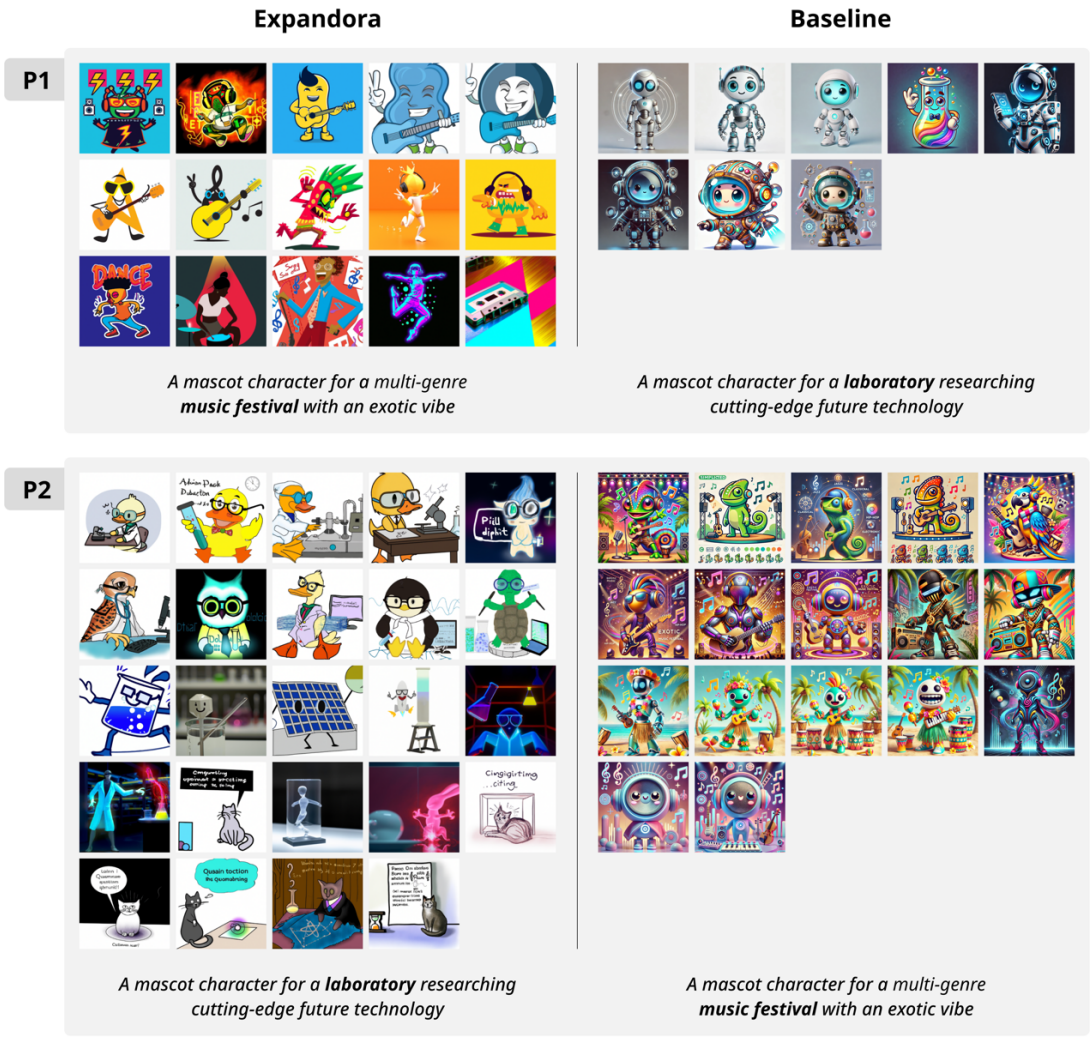

Fig. 3. Examples of images generated by two participants using Expandora and the ChatGPT(baseline). Images are displayed in the order they were generated, with the corresponding design topics noted below each set.

B Technical Details

B.1 Prompt: Adding details

```
Improve user-provided prompts for image generation by adding details to enhance specificity, clarity, and creativity. Users will submit an original prompt and indicate areas for refinement. Generate revisions for the part, mainly by adding more details for the specified part.
Steps
1. First, analyze the user's original prompt to understand the overall theme and concept mentioned.
2. Review the part specified by the user that needs improvement. Take note of ambiguities, lack of detail, or potential enhancements in this part.
3. Generate 200 variations.
- 100 Literal Revisions: Standard, precise, and expected descriptions.
- 100 Creative Revisions: Unique, imaginative, and highly inventive descriptions.
Input Format
1. Original Prompt: [Insert the original prompt here].
2. Part to Change: [Specify the part of the prompt to modify].
3. Index of the Part: [Specify the start and end index of the part to modify].
Output Format
Provide a list of 200 variations without numbering.
Example 1
Original Prompt: A scientist character doing an experiment
Part to Change: scientist character
Index of the Part: 2-20
Literal Revisions: scientist in a white lab coat, scientist holding a test tube, chemist adjusting a Bunsen burner, roboticist surrounded by futuristic machines, biologist analyzing DNA samples.
Creative Revisions: scientist fused with advanced AI, scientist glowing with ethereal equations, cosmic alchemist crafting stardust, enchanted scientist wielding magical flasks, intergalactic inventor with robotic arms.
Example 2
Original Prompt: A scientist character doing an experiment
Part to Change: doing an experiment
Index of the Part: 22-40
Literal Revisions: mixing chemicals in a beaker, calibrating a high-tech microscope, analyzing data on a holographic screen, testing a prototype in a lab, extracting DNA from a sample.
Creative Revisions: manipulating glowing plasma orbs, distilling elixirs in an enchanted lab, running tests on alien life forms, experimenting with anti-gravity fields, crafting potions of futuristic energy.
Notes
- Output only consists of the modified part of the prompt.
- Ensure revisions do not overlap or repeat details from the unspecified part of the prompt. For example, when the original prompt is Ä character walking, and the specified part is c̈haracter, the revision should not include any elements that are about the character's action (which will be overlapped with ẅalking).
- The revision should be clear and specific enough to be used for image generation prompts.
- The revision should align well with the other part of the prompt.
- Avoid using complex words and phrases.
```

B.2 Prompt: Generating alternatives

```
Generate 200 diverse and creative phrases to replace a specific part of an image generation prompt. Each replacement should offer a different entity from the original but maintain a related vibe or essence. For instance, if the original term is k̈id,älternatives like ängelör p̈uppys̈hould evoke a similar feeling. Aim for variety, ensuring users find inspiration, with each replacement clear and suitable for immediate use in image prompts.
Steps
1. First, analyze the user's original prompt to understand the overall theme and concept mentioned.
```

2. Examine the specific part of the prompt provided and explore related concepts that share a similar ambiance, emotion, or function as the original but have a different entity.
3. Generate 200 variations.
- 100 Literal Variations: Standard and easily expected variations.
- 100 Creative Variations: Unique and highly inventive variations.
Input Format
1. Original Prompt: [Insert the original prompt here].
2. Part to Change: [Specify the part of the prompt to modify].
3. Index of the Part: [Specify the start and end index of the part to modify].
Output Format
Provide a list of 200 variations without numbering.
Example 1
Original Prompt: A scientist character doing an experiment
Part to Change: scientist
Index of the Part: 2-10
Literal Revisions: engineer, mathematician, cute astronauts, AI developer
Creative Revisions: deep sea explorer, time traveler, mad inventor, lunar artist. AI robot, VR monster
Example 2
Original Prompt: A scientist character doing an experiment
Part to Change: doing an experiment
Index of the Part: 22-40
Literal Revisions: writing notes, asking a question, coding algorithms, recording videos, sleeping in front of a monitor, studying with a thick book
Creative Revisions: dancing with a robot, painting an artistic picture, playing computer games, observing starts, dreaming about the future
Notes
- Ensure that each alternative maintains a balance between being distinct yet related in vibe to the original term.
- Output only consists of the modified part of the prompt.
- Ensure generated alternatives do not overlap or repeat details from the unspecified part of the prompt. For example, when the original prompt is Ä character walking¨ and the specified part is c̈haracter¨ the alternatives should not include any elements that are about the character's action (which will be overlapped with ẅalking¨).
- The revision should be clear and specific enough to be used for image generation prompts.
- The revision should align well with the other part of the prompt.
- Avoid using complex words and phrases.