

Prompting for Discovery: Flexible Sense-Making for AI Art-Making with DreamSheets

Shm Garanganao Almeda, J.D. Zamfirescu-Pereira, Kyu Won Kim,
Pradeep Mani Rathnam, Bjoern Hartmann
UC Berkeley
Berkeley, CA, USA

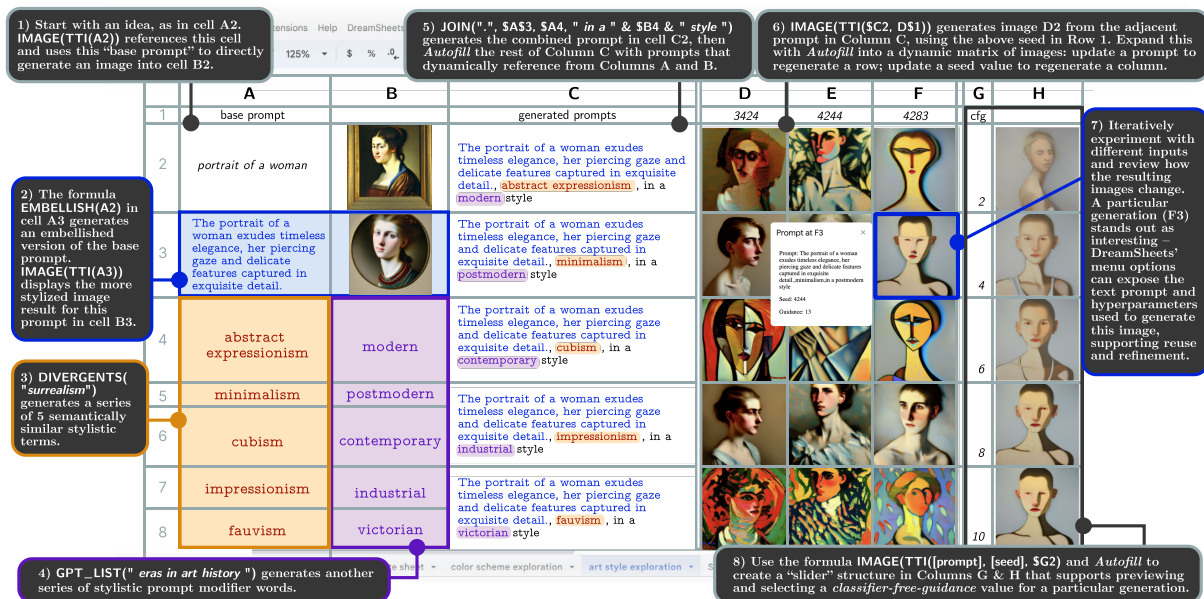


Figure 1: DREAMSHEETS provides a 2D spreadsheet interface that Text-to-Image users can use to author creative workflows for rapid prompt-image exploration. In this example derived from sheet-systems authored by expert participants (E1, E2, E4, E5) an initial idea for a prompt (A2) is used to generate an image (B2) using DREAMSHEETS's TTI() function. They modify the prompt with EMBELLISH(), an LLM-based function, to generate the more stylized (B3). Other LLM functions DIVERGENTS() and GPT_LIST() are used to generate series of *art styles* and *eras in art history*, two chosen axes for prompt exploration. Their dynamic *prompt template* combines the embellished prompt (A3) with references to the generated lists of "modifiers" to fill Column C with a series of prompts. These dynamic prompts generate the images in Columns D-F. Each image column utilizes a different seed, referencing the cell in row 1. Columns G and H depict a hyperparameter "slider" structure used to evaluate and select a *cfg* setting. If content in a referenced cell is updated, the sheet automatically regenerates dependent cells, allowing users to repurpose this workflow structure for iterative explorations.

ABSTRACT

Design space exploration (DSE) for Text-to-Image (TTI) models entails navigating a vast, opaque space of possible image outputs, through a commensurately vast input space of hyperparameters and prompt text. Perceptually small movements in prompt-space can surface unexpectedly disparate images. How can interfaces

support end-users in reliably steering prompt-space explorations towards interesting results? Our design probe, DreamSheets, supports user-composed exploration strategies with LLM-assisted prompt construction and large-scale simultaneous display of generated results, hosted in a spreadsheet interface. Two studies, a preliminary lab study and an extended two-week study where five expert artists developed custom TTI sheet-systems, reveal various strategies for targeted TTI design space exploration—such as using templated text generation to define and layer semantic "axes" for exploration. We identified patterns in exploratory structures across our participants' sheet-systems: configurable exploration "units" that we distill into a UI mockup, and generalizable UI components to guide future interfaces.



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

CHI '24, May 11–16, 2024, Honolulu, HI, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0330-0/24/05
<https://doi.org/10.1145/3613904.3642858>

CCS CONCEPTS

• **Human-centered computing** → **Interaction paradigms.**

KEYWORDS

generative AI, text to image, design space exploration

ACM Reference Format:

Shm Garanganao Almeda, J.D. Zamfirescu-Pereira, Kyu Won Kim., Pradeep Mani Rathnam, Bjoern Hartmann. 2024. Prompting for Discovery: Flexible Sense-Making for AI Art-Making with DreamSheets. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3613904.3642858>

1 INTRODUCTION

Text-to-Image (TTI) models like DALL•E [43] and Stable Diffusion [47] generate images from a combination of text prompts and numerical parameters (e.g., seeds). Interfaces for these models are seeing wide end-user adoption in a variety of settings, from marketing collateral to independent art making.

A critical skill for Text-to-Image (TTI) users is navigating between prompt text inputs and image outputs. Building such an understanding is neither trivial nor straightforward [58, 59]: the spaces of possible inputs and outputs are *massive*, and the mapping of one to the other is highly *opaque*.

Current patterns in commercial interfaces present limited support for exploration: a text box for prompt input and an area for displaying and saving a few output images. Some offer a number of additional features to support *prompt engineering*: canned prompt ideas, including options to influence “style,” and sliders for manipulating hyperparameters.

This lack of explicit interface support has led to the creation of resources such as community-curated prompt books [38], spreadsheets [56], and tutorials [33, 49] that document exploration processes. Researchers have also begun to study prompting practices [4, 19, 25, 34–36] and to propose alternative interfaces for interacting with TTI (e.g., Promptify [2]) and other types of generative models (e.g., GanZilla [10], Spacesheets [28], PromptAid [32]). These systems tend to support particular prompt-image or prompt-text workflows, helping users refine prompts towards a goal.

In this paper we argue that supporting Text-To-Image users goes beyond ensuring that they can achieve a particular end result; *gaining an understanding of the mapping between input and output is core to successful creation with generative AI systems*. Exploring the relationship between TTI inputs and outputs is a sensemaking process [41, 48] where users aim to build mental representations that allow them to reliably navigate to desired outputs. While the input and output spaces are massive and opaque, they are *not* arbitrary: prompts *can* develop reliable exploration “targeting strategies” with experience and productive reflection. Users can use the knowledge they gain from many iterative input-output tests to effectively “prompt-craft,” effectively steering generations towards desirable results.

Thus, our guiding research question is: *How might new interfaces support users in sense-making for successful art making with such models?*

To investigate this question, we built DREAMSHEETS, a tool that enables TTI model users to compose *targeted exploration* systems within a spreadsheet interface. In DREAMSHEETS, spreadsheet cells can contain prompts, or images generated from those prompts. A set of novel prompt manipulation functions enable users to explore prompt space computationally, through the construction and strategic combination of categorical lists, alternative wordings, embellishments, synonyms, and more. These functions are implemented through prompts to a large language model (LLM).

Spreadsheets may not readily provide the ideal affordances for organizing image collections; however, they are a *highly flexible, computational* substrate for *what-if* exploration; by presenting image and text generation tools within a customizable sandbox, DREAMSHEETS enables users to compose novel creative workflows.

We investigated users’ exploration strategies in two studies with DREAMSHEETS: a 1-hour lab study with 12 primarily amateur participants, and a two-week extended study with five expert TTI artists. In these studies, we examined how both groups (1) develop intuition for prompt designs that yield specific outputs, and (2) use DREAMSHEETS’ affordances for computational prompt manipulation, workflow creation, and output evaluation.

Our primary insights lie in observing and analyzing how participants used DREAMSHEETS to develop custom TTI sheet-systems, and in identifying sense-making and exploration patterns across these generative prototypes, including the construction and iterative reuse of composable exploration “structures” that map to generalizable UI concepts. We then use these insights to create UI mockups to inform potential future interfaces, and report on the feedback and speculation it elicited from our participants.

This paper makes three contributions:

First, it describes DREAMSHEETS, a spreadsheet-based TTI design space exploration platform that enables user-defined computational and LLM-supported interactions over the joint design space of prompts, seeds, and other TTI model hyperparameters.

Second, it offers a rich description of a first-of-its-kind extended (2-week) study exploring the ways in which artists use computational structures for sense-making and to explore the design space of TTI models.

Finally, it presents a set of generalized UI design suggestions, co-designed in a visual UI mock-up with our expert artist participants, and informed by the sheet-systems they developed while working in DREAMSHEETS.

2 RELATED WORK

Our work builds on research showing that considering many alternatives in parallel effectively aids design space exploration [15, 44], such as through gallery interfaces [31], tracking exploration history [16, 21], with suggestions for possible input shifts [31], and effective organization [29, 46, 55]. Spreadsheets enable many of these abilities that common TTI interfaces often lack.

In this section, we draw explicit connections to Creativity Support Tools, prompting and other TTI model workflows, design space exploration of images in non-TTI contexts, and sensemaking.

2.1 Creativity Support

In 2007, Shneiderman identified four underlying design principles for creativity support tools (CSTs): *support exploratory search, enable collaboration, provide rich history-keeping, and design with low thresholds, high ceilings, and wide walls* [51]. A more recent body of research explores how CST design can aid users' creative processes and productivity [6, 11]. Spreadsheets are themselves an example of a tool that supports creative exploration, enabling users to separate fixed values from values they want to vary, affording effective exploration and evaluation of "what-if" scenarios [45, 51].

2.2 Prompting & Text-to-Image Model Workflows

At the surface, prompting can appear straightforward, but crafting effective prompts is a challenge, even for experts [26, 58, 59]. How a prompt directly impacts model outputs is an active area of research [24, 50]. Choosing the right language to achieve desirable visual results in these prompt-based interactions can be difficult, motivating online user communities [9, 42] and researchers to develop and investigate new prompting techniques [26, 27] and tools supporting prompt discovery and exploration [2, 13]. These tools tend to be goal-driven, helping artists with a particular image "goal" target and improve their generations, while perhaps considering alternatives. In contrast, DREAMSHEETS seeks to explicitly support the rapid construction of flexible structures towards various user-defined goals. A growing body of literature has also begun to explore how communities of practice are approaching prompting [4, 35–37]. These studies offer taxonomies of prompt structures, showing how artists use different prompt modifiers [36], how they engage in the process of prompt engineering [37], and how they consider the material properties of text prompts [4]; this literature will rightfully continue to grow as these practices evolve alongside the models they rely on.

DREAMSHEETS builds upon this prior work by enabling prompt-artists to rapidly prototype computational structures and pursue user-defined, targeted explorations across prompt-image space—allowing us, in turn, to learn how artists use such new capabilities to build structures for sense-making and art-making.

2.3 Design Space Exploration of Images

Prompt discovery tools follow a long line of research into design space exploration of images. In spaces like computer graphics and animation where *visual judgment* of a human-designer-produced artifact is the primary evaluation mode, prior work has often focused on browsing interfaces, such as in Marks *et al.*'s seminal Design Galleries [31]. Interaction techniques for browsing include multi-step galleries [31]; map metaphors [53]; or faceted browsing [14].

Narrowing down from the explored designs, users may wish to pursue multiple alternative options for deeper exploration, though typically many orders of magnitude fewer than the number of algorithmically explored designs, as in GEM-NI [57].

Spreadsheets' usefulness for visual design space exploration in part stems from the intrinsic 2D matrix layout enabling "small multiples", a term Edward Tufte popularized [54] as an answer to the question "compared to what?" In a "contact sheet"-like 2D matrix, one can readily compare many images at once and quickly

identify the best candidates. Spreadsheets have a rich history of serving as vehicles for exploratory work, in accounting and far beyond—utilized as early as 1994 for information visualization of data and images themselves [5, 22], including images generated from a numerical input space [28].

2.4 Sense-making

A number of our observations relate to the broader sense-making literature, including Pirolli and Card's seminal work on information foraging [40]—DREAMSHEETS offers users an *information scent* on prompts—and sense-making more broadly [41, 48]. This line of work models how users navigate and make decisions in information-rich environments (like DREAMSHEETS), balancing between the perceived cost of seeking information and the potential reward of finding what they're seeking. DREAMSHEETS's design draws upon the *free energy principle of the brain* theory from cognitive science [12] which describes how the brain reduces uncertainty by making predictions and updating an internal mental model accordingly, generatively optimizing its internal model with sensory input to enhance prediction accuracy. This principle formed a basis for Davis *et al.*'s Creative Sense-Making (CSM) framework [8], which they applied to human-AI co-creation in the collaborative drawing domain. DREAMSHEETS's design also draws inspiration and lessons from existing sense-making interfaces – including classics like Scatter/Gather [7] and more modern implementations like Sensecape [52]. SemanticCollage [18] and ImageSense [17] provided reusable, system-generated text and visuals to support creative image search and sense-making with "reflection in action".

3 PROMPT & IMAGE EXPLORATION WITH DREAMSHEETS

DREAMSHEETS leverages the inherently flexible spreadsheet model in support of iteration and exploration of the TTI generation prompt input space. The features of DREAMSHEETS are embedded within a spreadsheet (built on Google Sheets) that recomputes and re-renders images in response to prompt additions and changes, allowing (for instance) drag-based "autofill" of columns, rows, or 2D regions with formula-generated prompts and images based on those prompts, alongside other common spreadsheet functionality. Specifically, DREAMSHEETS offers access to diffusion model image generation as a spreadsheet function that can take the content of other cells in the sheet as input, including combinations or transformations of multiple cells. These features support the user in efficiently exploring, observing how generated outputs are influenced by modifications to the input.

Our prototype also includes a set of LLM-based spreadsheet functions for manipulating prompts directly, such as `GPT_LIST` and `LIST_COMPLETION` for generating or extending a list of items of a certain description, `EMBELLISH` to create a detailed variation of the input text, and `ALTERNATIVES` to generate multiple variations of a seed prompt (see Table 1 for a full list). These functions serve as an aid to users in the construction of axes along which to explore the prompt/image design space, supporting template prompts with automatic value insertion.

Function Name	Description
TTI(prompt, [seed], [cfg])	Generate an image (at the returned URL) using the given <i>prompt</i> and, optionally, <i>seed</i> and a classifier-free guidance (<i>cfg</i>) parameters.
GPT(prompt)	LLM function for arbitrary <i>prompt</i> .
GPT_LIST(prompt, length)	Populates <i>length</i> cells in a row (or column) with words/phrases of type <i>prompt</i> .
LIST_COMPLETION(prompt)	Like GPT_LIST, but <i>prompt</i> is a list of items, rather than a description.
SYNONYMS(prompt)	Generates a list of synonyms.
ANTONYMS(prompt)	Generates a list of antonyms.
DIVERGENTS(prompt)	Generates “divergent” words.
ALTERNATIVES(prompt)	Generates a list of alternative wordings for <i>prompt</i> .
EMBELLISH(prompt)	Generates an embellished alternative to <i>prompt</i> , commonly using more specific or detailed words.

Table 1: The LLM-based functions available in our prototype. Each list-producing function additionally has an `_T` alternative (e.g., `SYNONYMS_T`) that transposes its outputs across a row.

For example, a template prompt like “A *<facial expression>* woman” can be expanded into a column of different prompts by substituting values generated by `GPT_LIST("FACIAL EXPRESSIONS")`.

3.1 DREAMSHEETS Design

Our primary goal here was to enable the use of a computational substrate for TTI model design space exploration via spreadsheet formula construction.

Drawing on prior work in prompt design [25, 34, 36, 38], we identified the testing of alternative phrasings and the addition of detail as core activities in TTI prompt exploration. These activities help users explore neighboring points in design space and recognize fruitful *directions for further* prompt explorations. We learned from prior evaluations of prompt engineering in the TTI context [37] that users were likely to express a diversity of design patterns and that supporting flexibility, providing immediate visual feedback, and offering an extended period of familiarization would be critical.

We operationalized support for these exploratory activities as the `ALTERNATIVES`, `DIVERGENTS`, and `EMBELLISH` functions. Similarly, synonym and antonym generation are core NLP building blocks, useful for creating variation that targets specific words in a longer prompt—we integrated these capabilities through the `SYNONYMS` and `ANTONYMS` functions; see Table 1 for a full list. These operations formed a foundation from which users could build their own custom workflows and strategic approaches to discovering model capabilities and exploring the underlying prompt design space.

To use these concepts in a spreadsheet paradigm and support the generation of *sets* of images, we designed these functions to output *lists* of values that populate across a column (or row) of cells. These terms in these cells can be referenced in traditional spreadsheet style and concatenated with other values to form combined prompts. We also provided functions to *extend* lists of prompts or prompt parts, allowing users to build on a conceptual list by providing a few initial examples.

3.2 DREAMSHEETS Implementation

We explored various service options for DREAMSHEETS’s underlying spreadsheet functionality, including building our own spreadsheet interface from scratch, open source spreadsheets HandsOnTable¹ and LuckySheet,² and both Microsoft Excel and Google Sheets.

One major challenge in integrating with an existing spreadsheet is the relatively long latency of image generation itself: up to 15 seconds or more, even when using cloud APIs. Spreadsheet users are accustomed to rapid updates and recomputations in response to changes in cell values—a multi-minute delay resulting from a backlog of image prompt updates and, consequently, new image generations, would be unacceptably slow. This need drove our use of the Stability.ai API,³ which supports parallel image generation requests with the Stable Diffusion 2 model, and offers sub-15 second response times. This critically enabled the full-scale “small multiples” visualizations of results that we wanted users to be able to utilize to view and evaluate results across multiple input axes simultaneously.

Ultimately, we selected Google Sheets as the spreadsheet interface, as it is easily extensible and accessible to most people. Google Sheets’ Apps Script environment lets developers create add-ons in a JavaScript-like environment, has a sufficiently long timeout (30 seconds) for custom functions, and allows users to continue to edit the sheet even while our custom formulas, which required back-end calls to TTIs and LLMs, awaited responses.

As a side benefit, because Google Sheets is already an online-native platform, rapid collaboration and version history are built-in.

We implemented DREAMSHEETS as a Google Sheets Apps Script add-on and a proxy web server written in JavaScript using ExpressJS. The add-on adds custom functions described in Table 1, making the corresponding requests to the proxy web server which handles caching and calling the appropriate API to either Stability.ai or OpenAI. Figure 2 illustrates how the proxy server facilitates communication between the Google Sheets add-on and Stability.ai or OpenAI. For the `TTI` function, the proxy server makes a hash using a combination of the prompt, seed, and guidance values and checks if the image has been generated before. Otherwise, an API call is made to Stability.ai to generate a 512×512-pixel image which is then cached in the file system for easy retrieval in the future.

The LLM-based functions that return a list utilizes OpenAI’s ChatGPT with `gpt-3.5-turbo`. To ensure that ChatGPT returns a properly formatted list with the appropriate length, it is initialized with the following messages:

¹<https://handsontable.com>

²<https://github.com/dream-num/Luckysheet>

³<https://api.stability.ai/docs>

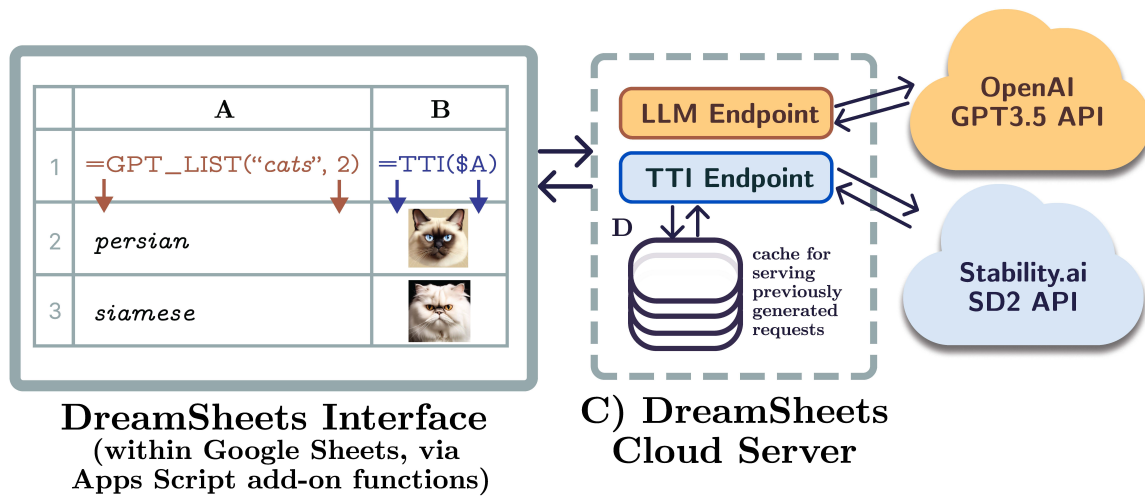


Figure 2: The DREAMSHEETS implementation. LLM (A) and TTI (B) functions fetch from separate endpoints of the DREAMSHEETS cloud server (C) which forwards requests to the OpenAI ChatGPT and Stability.ai Stable Diffusion cloud-based APIs. TTI requests are cached (D) using a hash of (prompt text, seed, classifier-free guidance) as a key.

system: Respond with a Javascript array literal with the given length in parentheses
 user: types of animals (length: 5)"
 assistant: ["dog", "cat", "frog", "horse", "deer"]
 user: [PROMPT] (length: [LENGTH])

The implementation for each LLM-based function differs only in the prompt sent to the LLM proxy server: each function prepends different additional instructions to the user's inputted prompt. The complete list of full prompts sent to ChatGPT are:

- LIST_COMPLETION** Similar items to this list without repeating "[LIST]"
- SYNONYMS** Synonyms of "[USER INPUT]"
- ANTONYMS** Antonyms of "[USER INPUT]"
- DIVERGENTS** Divergent words to "[USER INPUT]"
- ALTERNATIVES** Alternative ways to say "[USER INPUT]"
- EMBELLISH** Embellish this sentence: [USER INPUT]

3.3 Prototyping Interactions with Spreadsheets

In addition to serving as a platform for TTI exploration and sense-making, DREAMSHEETS is a demonstration of a kind of creativity support tool for composable prototyping that can be built on top of existing spreadsheet software as a computational substrate. Spreadsheets are an excellent foundation for tools that might benefit from 2D structure, formula construction, and a familiarized user base—and they can be straightforward to extend if there is alignment between the spreadsheet and backend data models. Here we describe how lessons from DREAMSHEETS can be used in the construction of other TTI-based CSTs.

First, in designing a system with a spreadsheet substrate, a number of considerations arise, grounded in *how much* and *what kind* of customization is needed. Is it sufficient, for example, to simply add new spreadsheet formula functions—or are more substantial changes needed to the interaction layers? Will custom data types

need to be supported, and how will they be rendered into cells? How will users interact with these different data types—are new input mechanisms (i.e., beyond text-in-cell input) needed too, such as file uploads or image editing? What response latency is acceptable given the tasks users are expected to engage in? And, finally, how critical are history-keeping and real-time multiplayer functionality?

These considerations constrain which underlying spreadsheets can be used. For DREAMSHEETS, we found that Google Sheets' API support was sufficient: custom functions were easy to add, and meeting the required maximum latency (30s) was possible by parallelizing requests to our backend. One major drawback was that custom functions could not cause cells to render images—only the built-in `=IMAGE(IMAGE-URL)` formula can do that—forcing the rather clunky `=IMAGE(TTI(PROMPT, SEED, CFG))` construction.

In contexts where interaction layer changes are needed, open-source spreadsheets such as HandsOnTable or LuckySheet offer more expansive opportunities for customization. Both options support extending spreadsheets with additional content types for cells as well as supporting additional functionality for user input and manipulation of images, poses, embeddings, and other data types relevant to TTI models. For DREAMSHEETS, this flexibility would have allowed us to build a `=TTI` function that displays an image directly, without needing an intermediate `=IMAGE` function call—as well as integrate image and other types of inputs.

Custom data types can pose challenges too: for opaque (non-user-interpretable) data types, like embeddings, researchers should consider how to represent those values in a sheet itself. In some cases, simply storing the data on a server and sending a unique ID (mapped to each value on said server) would be preferable to passing blob data around a spreadsheet's cells. Similarly, for compute-intensive operations like image synthesis, caching is critical: spreadsheets re-render frequently and unpredictably, and generating an entire sheet's worth of images can be both expensive and time-consuming.

In DREAMSHEETS, we cached images on the server and used a unique ID to pass around in the browser that represents a specific (prompt, seed, cfg) tuple. This approach could be generalized to support many other types of inputs and outputs, including vectors, embeddings, tokens, and other kinds of state, as well as intermediating models (LoRAs, ControlNet, etc.). Storing data on the server and referencing it by ID allows for less client-side overhead load in a spreadsheet (and potentially cost benefits through caching across users), and makes caching of those values straightforward, but requires an additional translation layer between sheet and backend.

4 METHOD

Having built DREAMSHEETS, we first ran a preliminary 1-hour lab study with novice users to understand how users approach using DREAMSHEETS for TTI exploration; this revealed that DREAMSHEETS enables a variety of custom workflows, but that sensemaking was nearly always among the first activity participants engaged in.

Following our lab study, we ran a second, 2-week extended study with expert users. This second study was intended explicitly to better understand the kinds of custom workflows experts would build for sensemaking, given enough time, and what kinds of individual activities those sensemaking workflows consisted of.

4.1 Preliminary Lab Study

In this initial study, we observed how participants used DREAMSHEETS to define and use a text-to-image generation workflow; we gave participants a concrete task and training in the tool, but did not direct them beyond that.

4.1.1 Participants. We recruited 12 participants via email lists and social media. All 12 reported some spreadsheet experience, with most (10 out of 12) reporting frequent use (many times or daily). 10 out of 12 participants also had some experience with TTI models, and only 1 participant (P1) reported no prior experience with LLMs.

4.1.2 Task and Protocol. Each study took place through a Zoom call that lasted approximately 60 minutes, during which participants and the facilitating researcher collaborated in a shared Google Sheets document. We designed a *concept art* creation task to give users a direction achievable in the short amount of time provided, while leaving room for subjectivity and creativity: users were given a single inspirational image, then asked to generate three new images that could fit in a style and unspecified narrative as suggested by the inspiration image. We included a prompt explaining the task in the activity sheet, as well as an image of a post-apocalyptic, ruined Seattle, complete with space needle.⁴

Our protocol began with a brief tutorial on DREAMSHEETS and its functionalities, followed by an observation of participants as they engaged in the concept art task. We used an example sheet to walk participants through a tutorial to first remind participants of general spreadsheet operations (i.e. using formulas with cell references and expanding them with autofill) and then introducing DREAMSHEETS’s image and text generation functions. Once users were comfortable using the TTI and GPT functions, we introduced the concept art activity. Participants were encouraged to think

⁴The image used for this task was borrowed with gratitude from Andy Salerno, whose blog post [49] originally inspired this work.

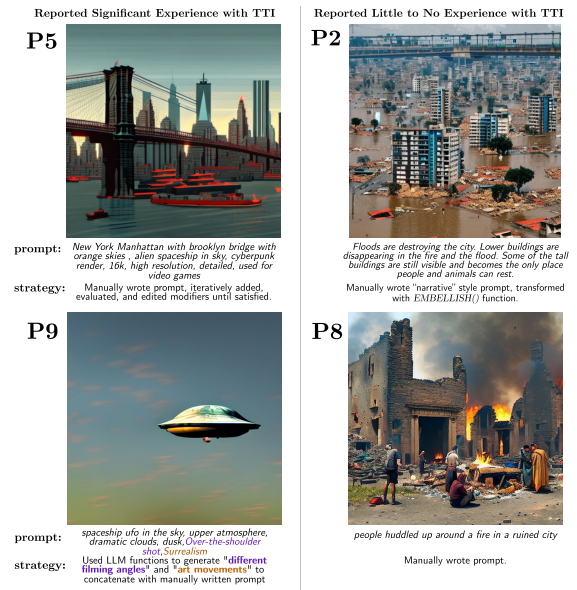


Figure 3: Examples of prompts authored by participants in the first study

along as they generated the three images required to complete the task.

4.1.3 Data Collection and Analysis. We observed, recorded, and transcribed video and audio of each interview, including approximately 40 minutes of system use, in entirety, throughout which participants were encouraged to think aloud and provide clarification when prompted. We then engaged in an exploratory qualitative data analysis using the recordings, transcripts, and resulting spreadsheet artifacts. We recorded responses to surveys completed before and after the interview, and reviewed usage data, containing logs of each text or image generation function call used in DREAMSHEETS.

4.1.4 Preliminary Study Results: DREAMSHEETS in use. Here, we provide an overview of some results specific to the preliminary lab study that informed our second, extended study. We discuss usage patterns and themes informed by *both* studies in section 5.

More than half of the participants in this study (7 out of 10) reported limited to no experience with TTI systems, but *all* participants were able to successfully utilize a prompt-crafting workflow in the DREAMSHEETS system, and to produce generations that they were satisfied with for the concept art task. Though authored directly by participants, the workflows adopted by more novice participants were likely inspired by the example structures showcased during the initial tutorial phase, with P3 and P4 copying from the tutorial examples directly.

The seven participants with limited prompt-engineering experience (P1, 2, 3, 4, 6, 8, 10) wrote their prompt in “English” ranging from brief sentence fragments to detailed scene descriptions. Meanwhile, 4 of the 5 participants who reported substantial or extensive TTI experience (P5, P7, P9, P11) wrote in a structure specific to “prompt language”—comma separated lists of terms, including modifiers to influence visual style.

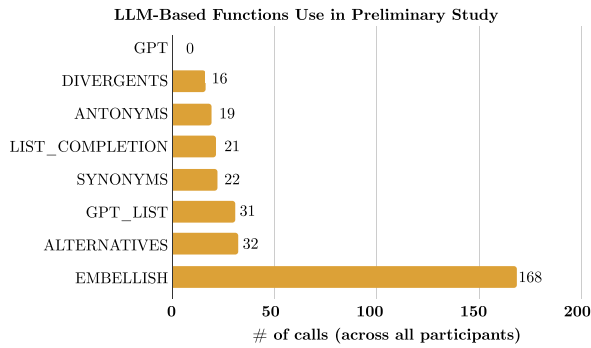


Figure 4: Number of times participants used each of the LLM-based functions (listed in order of ascending frequency) during the preliminary lab study activity. There were 4,737 calls to the TTI() image generation function.

Novice and experienced participants found LLM-based functions useful for creating or improving their prompts. Participants with more spreadsheet experience more readily adopted string concatenation strategies to construct prompts. Using LLM-based functions to generate a series of words in a particular category, like GPT_LIST("camera angles") or SYNONYMS("red"), participants introduced semantic modification "slots" into their prompts. We describe this *LLM-assisted dynamic prompt construction* as a prompt-space exploration strategy in Section 5.3.

This formative study confirmed DREAMSHEETS's usefulness as an exploratory TTI system and illuminated promising usage patterns, as we observed even novice participants begin developing various strategies and structures to support their task completion goals in the 40 minutes provided. However, the brevity and constraints (i.e. prescribed creative task) of this first-use study format meant that users did not fully leverage DREAMSHEETS to develop strategies for user-defined, creatively motivated goals. This motivated a longer-term expert study to observe how generative artists might utilize DREAMSHEETS to build systems for "real-world" creative workflows.

4.2 Extended Expert Study Design

We turned to an extended study to observe the custom sheet-systems that experts would create when given the time and flexibility to pursue authentic creative explorations.

4.2.1 Participants. To recruit experts for our second user study, we sent recruitment messages to individuals publicly participating in generative art communities on social media, and recruited 5 individuals (designated as E1-5 to differentiate from Study 1 participants).

4.2.2 Protocol. We conducted three 45 minute interviews spanning 2 weeks with each participant. Participants were instructed to use the tool for about 7-10 hours over the course of the 2-week study. We suggested 30-45 minutes of tool use per day, but participants were given the freedom to decide the length and structure of their work sessions. As with the first study, the initial interview began with a short tutorial reviewing spreadsheet functionality and

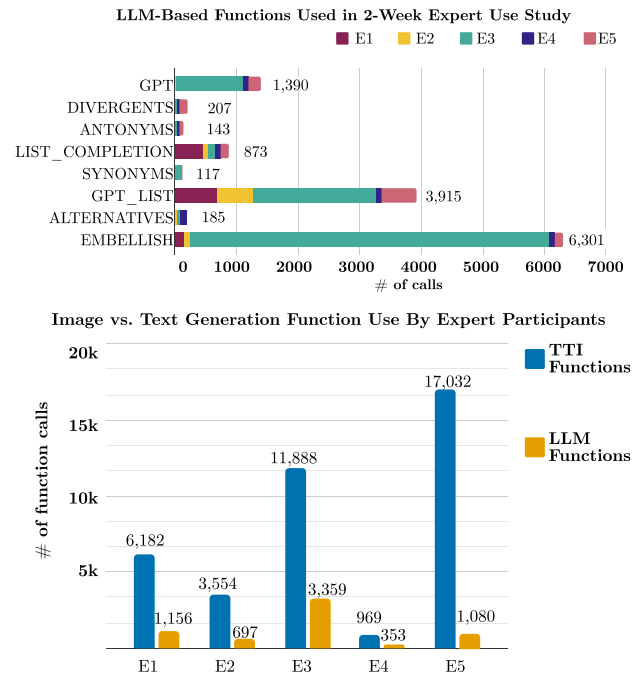


Figure 5: Above, LLM-based function use across the 2-week expert use study. Individual colors represent individual expert participants consistently across functions (5 total). Below, the number of times each expert participant made a unique call to the image generation and text generation functions during the 2-week study.

demonstrating DREAMSHEETS functions. The collaborative spreadsheet shared with each participant included documentation and examples of DREAMSHEETS function use. Participants could contact the research team via email with questions throughout the study.

The second interview took place 1 week into the study. We asked participants to explain their exploration goals and strategies, and to use relevant parts of their spreadsheets to illustrate. We described back to participants our observations, allowing them to clarify any potential misinterpretations of their actions. Based on the feedback we received, we designed a UI mockup that incorporated elements inspired by the structures built and functions used by participants during their first week of using DREAMSHEETS.

In the third and final 45-minute interview, we again asked participants to describe the creative explorations evident in their spreadsheets, and to explain how they integrated DREAMSHEETS's functionalities into their creative process. We then showed participants the UI mockup to gain their perspective and elicit further feedback and suggestions for designing more supportive TTI interfaces.

4.2.3 Data Collection and Analysis. We analyzed our participants' usage of DREAMSHEETS as observed or described during interviews, as well as the resulting artifacts: the sheets and usage logs, containing the full chronology of function calls made to the DREAMSHEETS system. We periodically viewed their spreadsheets throughout the 2-week study, including their Version History – a detailed record

of changes made to the sheet, which we used to recover and save copies of previous versions. We engaged in a rigorous data coding and analysis process by stepping through the version history of each expert participant’s Google Sheets document and leveraging usage data logs to build data visualizations of each “exploration session.” This included labeling high-level screenshots of each sheet noting structures (e.g. exploratory axes like “different seed in each column”), clustering semantically-similar prompts, and modeling the development of prompts over time, denoting both LLM-generated and manually authored iterative prompt modifications similar to those taxonomically defined in [36].

5 FINDINGS: STRUCTURES FOR EXPLORING PROMPT-INPUT-TO-IMAGE-OUTPUT SPACE

Across both studies, participants’ use of DREAMSHEETS and the artifacts they produced allowed us to observe and identify key elements of their TTI creative workflow: their *goals*, the *strategies* they chose to pursue, and the interface *structures* they constructed to support these strategies.

Our expert participants were able to construct a number of sophisticated spreadsheet systems and improved them throughout the study. The design of these systems were shaped by their particular exploration strategies for navigating TTI space; we highlight some of the patterns that emerged in the findings below.

In this section, we abstract the hyperdimensional Prompt-Input and Visual-Output space as two dimensions that the generative model maps between. The hyperparameter input space (2-dimensional in DREAMSHEETS, using *seed* and classifier free guidance, or *cfg*) is shown, when relevant, as a smaller, transformative dimension that lies between the two hyper-dimensions and influences the mappings between them. We use these abstractions to illustrate a particular exploration strategy and draw attention to the particular space(s) that TTI users are sensemaking during their exploration, and show structures prototyped by DREAMSHEETS participants that exemplify that exploration strategy.

The overall trend was for participants to start by recreating the prompt templating activities as in other tools, but then build on top of these by selecting and combining axes of exploration—which can operate in both linear and non-linear ways. Participants then constructed 2D “small multiples”-style grids, iteratively layering axes for increasingly sophisticated explorations of prompt-image space—towards targeted areas of image space, and simultaneously, towards sense-making exploratory goals: to observe and understand capabilities and interactions. For example: what artists and styles can this model reproduce? What subjects (e.g., animals) and attributes (e.g., colors, facial expressions) might interact to yield interesting results?

5.1 Iterative Prompt Exploration

Iterative prompt refinement, where participants gradually refine a prompt while testing the effects of each addition, is a fundamental TTI exploration strategy possible in any interface. However, many interfaces provide only a few (< 10) results, and offer limited support for comparing results, displaying only the results of one experiment at a time, or only allowing users visually compare the

results of chronologically adjacent experiments, as in the *sequential “chat” history* interface model used by Midjourney or Dreamstudio.

Limited support for user-structured comparisons makes it difficult to evaluate the impact of prompt modifiers, and confounds effective sense-making; what the user may perceive as a *semantically close* edit in prompt space can translate to a confoundingly *large visual transformation*, and vice versa – see Figure 6.

Our expert participants described previously using external tools to save and evaluate TTI results history – saving favorite results and prompt modifiers into a spreadsheet (E1, E5) or word document (E2, E4) with notes on the expected impact of each token, for example.

Spreadsheets *inherently* afford rich, reconfigurable, and structured history-keeping and results evaluation. DREAMSHEETS users leveraged the “infinite canvas” qualities of digital spreadsheets to keep and evaluate in situ records of their exploration history. They also leveraged the inherent *reconfigurability* of results within DREAMSHEETS; all participants used duplication (copying and pasting groups of cells, or entire sheets) to repurpose and iterate on prior explorations (including novices, who duplicated and adapted liberally from the tutorial structures.)

To effectively steer prompts towards desirable outputs, users benefit from reconfigurable history structures. Concurrently, DREAMSHEETS users found that it was beneficial to *generate* and *view* larger samples of results simultaneously. All participants tried organizing generated outputs in a “small multiples” or “contact sheet” layout (as described by E1); 3 of the 5 experts (E1, E4, E5) explicitly remarked on its usefulness for large-scale results evaluation. Calls to the TTI function comprised the bulk of participants use of DREAMSHEETS, as shown in Figure 5; generating an average **7,925 unique images** over the course of the study.

To *generate* these larger samples of results, participants used *variation* strategies to efficiently and systematically generate many variations of a single prompt idea. These variation strategies took the form of **Parametric Manipulations** and **Semantic Manipulations** as described in the following subsections.

5.2 Parametric Manipulation

While prompt-crafting is central to the effective use of TTI models, the ability to quickly manipulate hyperparameters alone provides a useful dimension for exploration, this motivated participants to develop structures around hyperparameter control.

All expert participants used dynamic references to a column or row with a series of hyperparameter values to prototype a “slider” like evaluation structure; see Figure 8 for examples using *cfg*.

3/5 expert participants (E2, E3, E5), used “Power Cells” to prototype the functionality of a global “Settings” panel with the option to *regenerate on update*. By structuring sheets such that all generations reference a seed or *cfg* value from a particular cell, updating this “Power Cell” would regenerate the entire sheet of results. This afforded iteratively testing values on a large sample of results to find a desirable “setting.” See E5’s “Same Seed Prompt Explore” in Figure 7B for an example.

DREAMSHEETS provides two parameters that influence image generation: a stochastic hyperparameter (*seed*) and a non-stochastic hyperparameter (classifier-free-guidance, or *cfg*) which participants used to transform their explorations in different ways.

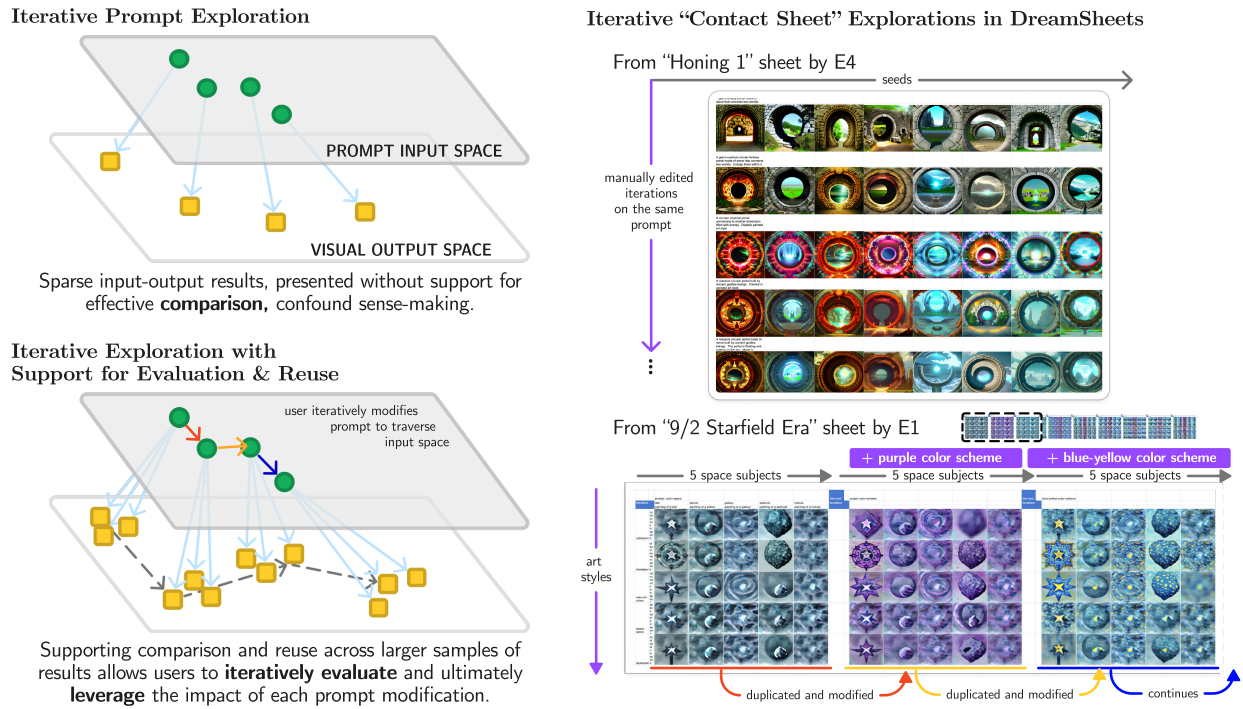


Figure 6: We abstract the Prompt-Input and Visual-Output space as two dimensions that the generative model maps between, and use these abstractions to illustrate the space(s) that TTI users are targeting via different exploration strategies. Steering towards desirable results with iterative prompt refinement is a common strategy, but sparsely presented results can confound productive sense-making. DREAMSHEETS users prototyped iterative sheet-systems that leveraged the “infinite-canvas” and rich history-keeping affordances of digital spreadsheets to conduct large-scale comparison across results, and to arrange and repurpose past explorations. With a structured, scalable results display, users can stumble upon interesting outliers without being confounded by them.

5.2.1 Stochastic Transformations. Seeds define the specific random noise that the diffusion model will use as a starting point; the model then repeatedly “de-noises” successive versions to generate an image, with the text prompt as a guide.

Generating many images using the same *prompt* but different *seeds* was a common strategy across participants in both studies.

All 5 expert participants utilized seed variations to quickly evaluate many “versions” of the same prompt. As shown in Figure 7A, a series of seed values allows the user to reveal a larger area of image-output space with each prompt. This improves efficiency towards creative goals (increasing the likelihood of finding a desirable output) and in sense-making (revealing larger samples of output space with each input test.)

There is no perceptual correlation between adjacent seeds, but images generated with the same seed may share visual similarities with the original noise pattern. A seed can then be useful for biasing generations towards a particular composition or color pattern, motivating targeted explorations of hyperparameter space. E5 prototyped several versions of a seed-exploration structure, including “Same Seed Prompt Explore” shown in Figure 7B. With a “vector graphics” design goal, they used this structure to identify seeds that could bias the image generation to feature “a central object” on a

flat background. E5 found seed 7935 and used this value in future “vector graphics” style explorations (see Figure 11B).

5.2.2 Non-Stochastic Parametric Transformations. E2 and E5 were particularly interested in exploring different *cfg* values. This hyperparameter has a perceptually linear influence on the image generation; a higher *cfg* value generates images more strongly influenced by the prompt. Low *cfg* values allowed E5 to gain a sense of “the model’s priors,” using a phrase commonly used in machine learning to refer to bias or preexisting (“prior”) beliefs. E5, E4, and E1 alluded to the “default style” latent to a specific image generation model as being important for prompt-artists to learn.

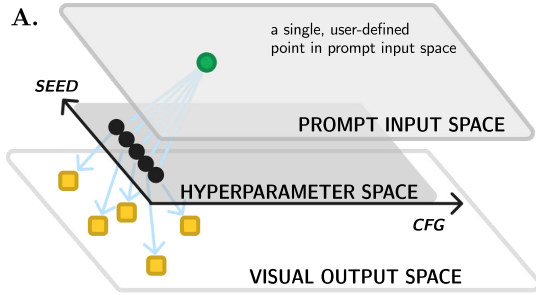
Non-stochastic hyperparameters that offer more “controlled” transformation are useful for exploration “depth” (i.e., repeated image refinement)—a capability that DREAMSHEETS lacks support for, and that systems like ControlNet[60] cater to.

5.3 Semantic Explorations

“Does Stable Diffusion know the same artists I do?”
(P11)

Participants manipulated language to make movements in prompt-space that would, ideally, translate into movements towards interesting areas of image-space. The spreadsheet interface provided a

Stochastic Transformations



Various seed values translate a single prompt input into multiple image outputs.

From "Exploration 1" by E4

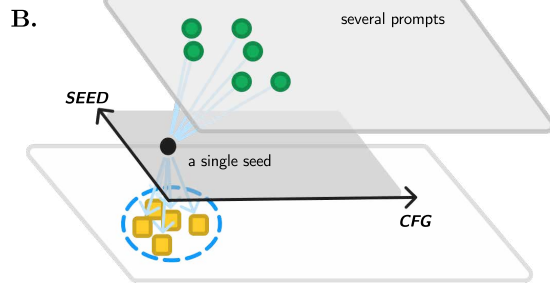
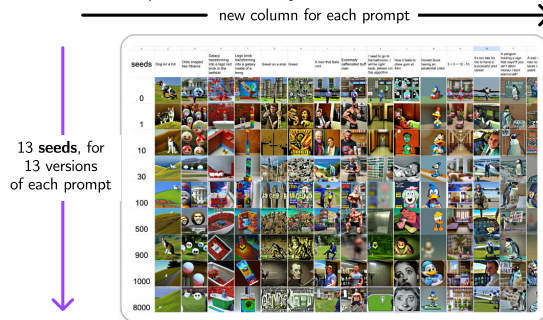


Image results transformed by the same seed will have visual similarities.

From "Same Seed Prompt Explore" by E5

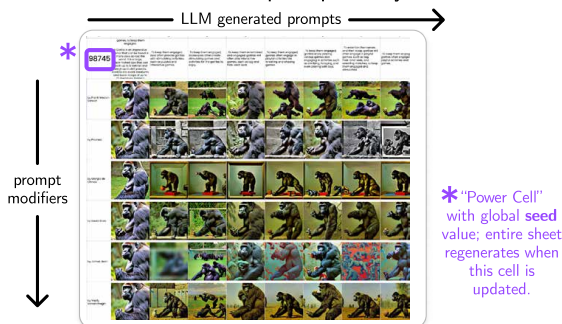
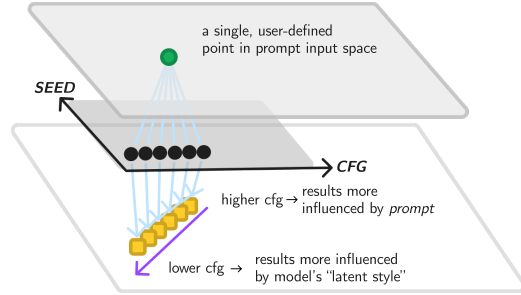
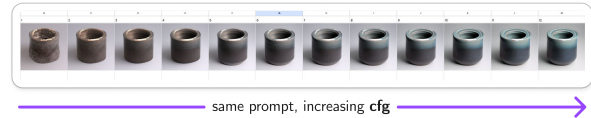


Figure 7: Here we introduce Hyperparameter space, a smaller intermediate dimension that can transform the mappings between Prompt-Input and Visual-Output space. Here we show two distinct strategies utilizing stochastic hyperparameters to transform results, with views of exemplary structures built by expert participants. A improves the efficiency of each test by increasing output sample size. B supports searching for and selecting a desirable hyperparameter value.

Non-Stochastic Transformations



From "Guidance Progression" by E5



From "cfg-test" by E2 (rotated for fit)

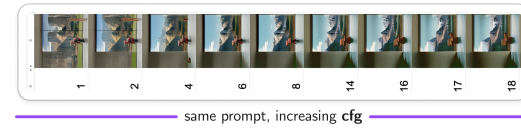


Figure 8: Participants created a "slider" along a series of classifier-free-guidance (cfg) values to preview a perceptually linear transformation on a particular generation. The global "Power Cell" structure shown in Figure 7B was also used to iteratively test cfg values on a batch of results. Offering more "controlled" transformation is particularly useful for "depth" explorations of image results, which DreamSheets otherwise lacks support for.

familiar structure for 2D evaluations; evaluating the combinatorial effect of two exploratory "axes" at a time (e.g. "subject" columns by "art-style" rows) was a common strategy. P6, a novice user in the preliminary study with limited prompting experience, said:

"I think it's good to know how things change depending on different variables... the spreadsheet helps with navigating what exactly is changing within the image." (P6)

5.3.1 Manual Semantic Exploration. To streamline iterative prompt explorations, participants constructed dynamic "prompt templates" that combine "base prompts" with swappable "slots." Our expert participants echoed [4]'s findings, treating these carefully crafted "prompt templates" as art pieces in themselves.

In Figure 9, E1 combined simple "base prompts" with several "Internet aesthetic" words to evaluate their efficacy as prompt modifiers – checking if the model would interpret each "aesthetic" in alignment with their expectations. Participants used manually constructed series (as opposed to the list-generating LLM functions) to conduct specifically targeted explorations. E4 prioritizes total creative writing control, using TTI to craft comedic prompt-image pairs and narrative concept art; they chose to manually craft most prompts without LLM-assistance.

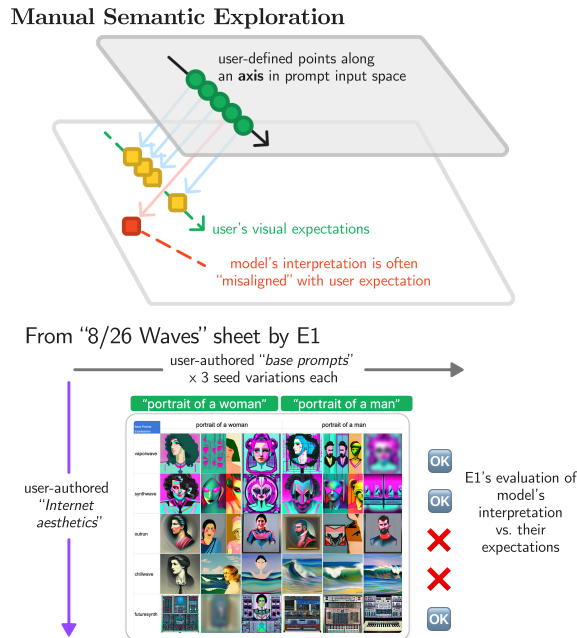


Figure 9: Participants constructed dynamic “prompt templates” that combine *base prompts* with “swappable” *prompt modifiers* (here, a series of “Internet aesthetics”). Users attempt conceptually gradual movements in prompt-space for procedural sense-making, iteratively testing how each modification to prompt text maps (often imperfectly) to movement in image-output space and updating expectations accordingly.

5.3.2 Generative Semantic Exploration. Participants employed DREAMSHEETS’s LLM-based functions and spreadsheet concatenation to build sheet-systems that streamline the discovery of useful points in prompt-space. By crafting dynamic prompt templates that reference from LLM-generated lists, participants select semantic “axes” to define a “prompt space” for LLM-assisted exploration. 4 of 5 experts (all except E4) utilized cell concatenation to craft dynamic prompt templates with LLM-generated prompt parts, though all of expert participants used the LLM functions to some extent. See Figures 4 and 5 for counts of how frequently participants utilized each of the LLM functions across both studies.

E2 and E3 independently developed two-part interfaces that separate the design process into prompt-authoring and image-evaluation steps. The “lists” section of the interface houses the semantic-axes selection and sampling process: here, the user selects categories for text generation like “lighting techniques,” or “mythical creatures,” and decides how to combine them (often with row-wise, comma-separated concatenation.)

In a separate section, the concatenated prompts are used to generate a column of image results. After constructing these sheet-prototypes, E2 and E3 reused them extensively for many explorations.

For 3 of 5 expert participants (E1, E2, E5), list generation functions `GPT_LIST` and `LIST_COMPLETION` comprised more than half of

their LLM Function use. E3 also used these list generation functions many times—in fact, more than any other participant—but their use of DREAMSHEETS stood out overall: E3 made 9268 LLM function calls, 7.6 times more than the next most frequent user of the LLM functions.

E1 used text-generation to discover interesting new prompt modifiers, including “monochromatic,” “watercolor,” and “pixel art.”

“I didn’t set out to make pixel art or watercolors, but through the course of the study I discovered these aesthetic spaces that I really loved!” (E1)

LLM-functions can accelerate prompt space traversal while supporting creative *Recognition over Recall*- users can choose *camera angle* as a “slider” to explore, then recognize an appealing “setting” in the generated results - without having to know or recall the words to describe it [33].

5.4 Flexible Scaffolding for User-Structured Multidimensional Explorations

Users approach TTI generation systems with a wide variety of creative goals, as showcased by the diverse images generated by our participants. Participants iteratively combined multiple, multidimensional exploration strategies to prototype bespoke sheet-systems for targeted explorations of prompt-image space.

Figure 11A shows some of E1’s *Expressions Exploration* combining an LLM-generated list of *facial expressions*, stochastic transformations, and a manually authored list of *subjects* (man, girl, dog).

Figure 11B shows a segment of E5’s targeted, multidimensional “vector-graphics” exploration, combining manual and LLM-generated semantic axes with a global “Power Cell” for iteratively regenerating the sheet with different “flat-color-biased” seed values, identified via targeted explorations as in Figure 7B.

Over the course of the study, E5 generated 4693 such “vector-graphics” style images across two large-scale exploration sheets; E3 made 11,562 unique “animal-plant photography” generations with the system shown in Figure 10. Other participants pursued a variety of diverse creative focuses throughout the study, but all of these prompt artists exemplify the visual art style development observed by Chang *et al.* [4]: with delicately crafted prompt-templates and deliberately selected hyperparameters, they developed distinct art styles and share with their online communities as images, prompts, and prompt-templates.

DREAMSHEETS’s flexibility allowed users to develop custom systems for various goals. The patterns that nevertheless emerged suggest generalizable structures that future systems can offer as dimensional exploration “units” – *composable support structures* with the flexibility to decide when and how to combine them. This echoes a takeaway from Li *et al.* in “Beyond the Artifact: Power as a Lens for Creativity Support Tools”: creative practitioners are empowered when they can *laterally compose* tools in an efficient workflow, or refuse tools and replace them with others [23].

Flexible options for AI-assistance can play a role in fluidly supporting different exploration strategies and styles. When users hand off creative-labor and control to the power of AI-generated serendipity, they should maintain the power to reclaim control at any time. This echoes tensions observed by Lawton *et al.* in “When is a Tool a Tool? User Perceptions of System Agency in Human–AI

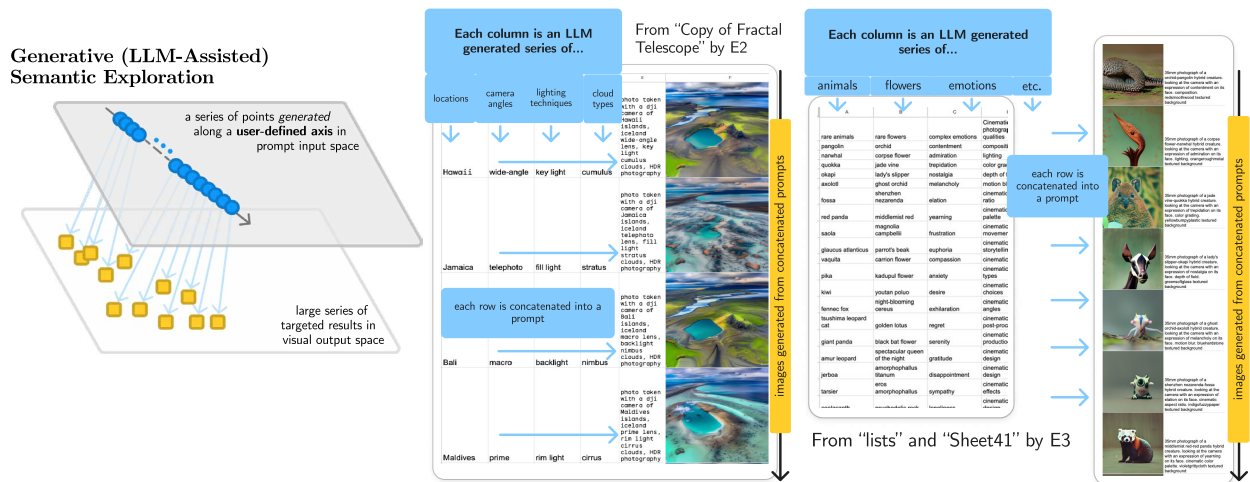


Figure 10: E2 and E3 independently developed conceptually similar sheet-system prototypes for different creative goals (drone-style landscape photography and animal-plant hybrids). They each used column-wise categorical-list generation (via LLM functions) and row-wise concatenation to generate a long column-wise series of results.

Co-Creative Drawing” [20]. Future studies should investigate how co-creative systems might allow users to flexibly control where and how AI “assistance” influences their workflow.

6 CO-DESIGNING EXPLORATION SUPPORTIVE UI FEATURES WITH THE DREAMSHEETS 2.0 MOCKUP

To more concretely probe into how participants conceptualized their own sense-making systems, and to inform a more generalized understanding of how we can support these processes, we developed “DREAMSHEETS 2.0” UI mockups with participants to elicit concrete feedback and speculative design ideas.

To facilitate the application of our findings towards future interface designs, we map the co-designed UI concepts and components to the exploration strategies observed in participants’ use of DREAMSHEETS 1.0; we present these in Figure 14.

6.1 Rich, Reusable Exploration History

Our participants’ sheet-systems in DREAMSHEETS exemplified how TTI interfaces can offer improved support for iterative prompt exploration by affording rich history-keeping and structured, large-scale evaluation across results. To that end, we designed the DREAMSHEETS 2.0 mockup to include two visual layout settings that users can freely toggle between: a “small-multiples” grid view (12, left) and a focused list view (Fig. 12, right). “Scalable” output display supports TTI users as they move between broad and focused evaluations of results.

Participants used sheet and cell duplication to iteratively re-purpose exploration history in DREAMSHEETS 1.0; this informed our decision to suggest features for revisiting prompt-history and exploration “sessions” in DREAMSHEETS 2.0.

We proposed a prompt “token bank” system (Figure 12, featured in Figure 13, right) that would allow users to convert highlighted

prompt text into a *Saved Token* for reuse in future prompts. Tokens can be converted into *dynamic* tokens for semantic explorations.

6.2 Supporting Exploration Breadth and Depth

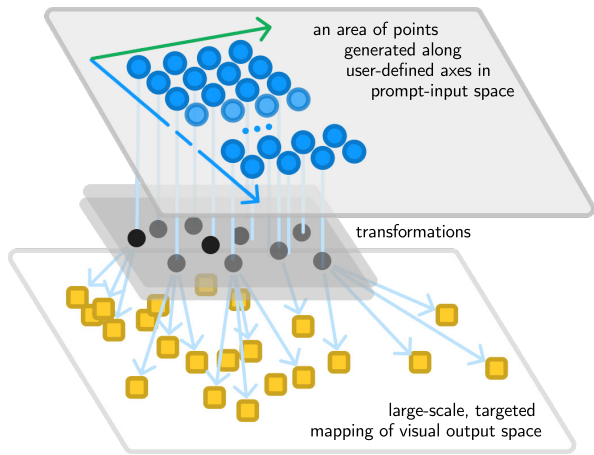
Participants utilized the structured “infinite canvas” qualities of digital spreadsheets to iteratively expand explorations. Providing the user direct control over the “scale” of their generations allows them to flexibly expand and evaluate explorations. To this end, DREAMSHEETS 2.0 would continuously load more results on-demand, for potentially infinite scrolling. Cost may limit the feasibility of large-scale generations; offering cheaper “low-fidelity” generations and options to increase quality on-demand may be a potential solution. We replicated the “Power Cell” strategy crafted by participants by providing options that would allow DREAMSHEETS 2.0 users to regenerate their exploration session by updating global hyperparameter settings.

We used the “Explore this image...” option to elicit participant ideas for additional “image refinement controls” beyond the *classifier-free-guidance* manipulation available in DREAMSHEETS 1.0. Suggestions included adding spatial conditioning controls (e.g. manipulating Poses or Edges; participants mentioned ControlNet [60]), support for generative inpainting (as enabled by Muse [3]), or image-to-image interpolation (as in SpaceSheets [28]). Participants also suggested adding Image-to-Text transformations similar to the CLIP Interrogator [39] or Midjourney’s /describe function [1], which could “close the loop” by allowing users to translate ideas back and forth between image and prompt space.

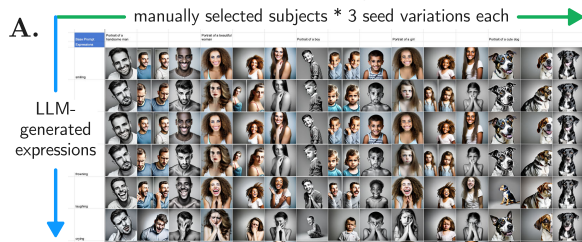
6.3 Supporting Semantic Exploration with Prompt Templates and Dynamic Tokens

DREAMSHEETS 1.0 participants used cell concatenation to craft prompt templates with “slots” to select axes for structured semantic exploration, motivating the design of more supportive prompt-template features in our 2.0 Mock-Up.

User-Structured Multidimensional Exploration



From “Expressions Exploration” by E1



From “Background Color Exploration” by E5

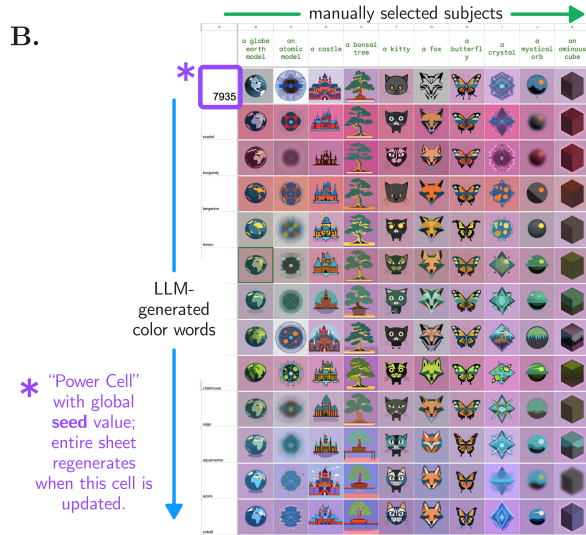


Figure 11: Participants adapted sheet-systems to combine multiple strategies to embark on a structured, multidimensional exploration, broadly evaluating a wide, but targeted, mapping of input-output space.

Dynamic tokens take the role of the “slots” used by participants to specify where to introduce text variation. Prompt variations are automatically generated with LLM-assistance, systematically combined, and populated across several columns of generated images. The LLM supporting this interaction is surfaced, providing users control over the text-generation, or the option to refuse its support altogether. Users can manually append, edit, and remove prompt words for each column of exploration, at will. This component meets *Chang et al.*’s call for future interfaces to support prompt templates as standalone, interactive computational artifacts [4].

6.4 Flexible Structures for User-Defined Multidimensional TTI Explorations

DREAMSHEETS 1.0 offered the flexibility of an infinite “blank canvas” with the trade-off of minimal structural support “out of the box.” Our expert participants were required to iteratively prototype systems custom to their creative “styles” and workflows. Rather than designing for a particular workflow, DREAMSHEETS 2.0 suggests supportive structures while maintaining flexibility by presenting composable exploration features. Users can effectively pursue simple prompt-input-image-output tests, or they can construct increasingly sophisticated multidimensional explorations. With configurable, reusable, and reusable components, users can compose targeted, iterative explorations of prompt-image space.

6.5 Participant Feedback

Participants reflected positively on the features presented in the DREAMSHEETS 2.0 Mock-Up, validating our adaptive interpretations of the structures they prototyped in DREAMSHEETS 1.0. The Mock-Up was developed and presented during the study; their feedback directly contributed to improving the design and identifying its most promising components. The ability to save and recover exploration history for reuse in future explorations was a highlight: E4 considered the “prompt token bank” a direct upgrade to their current history keeping practices (saving prompts and useful stylistic modifiers into a text document).

“If I can drag and drop a presaved dynamic chunk... I can fully focus on sculpting the prompt and being creative.” (E4)

The “Save session” feature prompted E1 and E5 to describe similar expected use cases – to save current state of the system, pausing their workflow to return at a later date, ideally supporting history-reuse. E5 expressed their preference to segment their process into a “generation” stage (crafting prompts to generate thousands of images) and a “evaluation” stage (curating the selection of images), and described being able to pause and temporally separate these activities as a potential “game-changer.”

7 LIMITATIONS

We acknowledge several limitations of our study. First, the particular hyperparameters exposed may differ between different TTI models; future work should investigate if our approach scales to higher-dimensional hyperparameter spaces. Second, more powerful techniques to guide image generation are emerging in research, such as ControlNet [60], or Readout-Guidance [30], which require

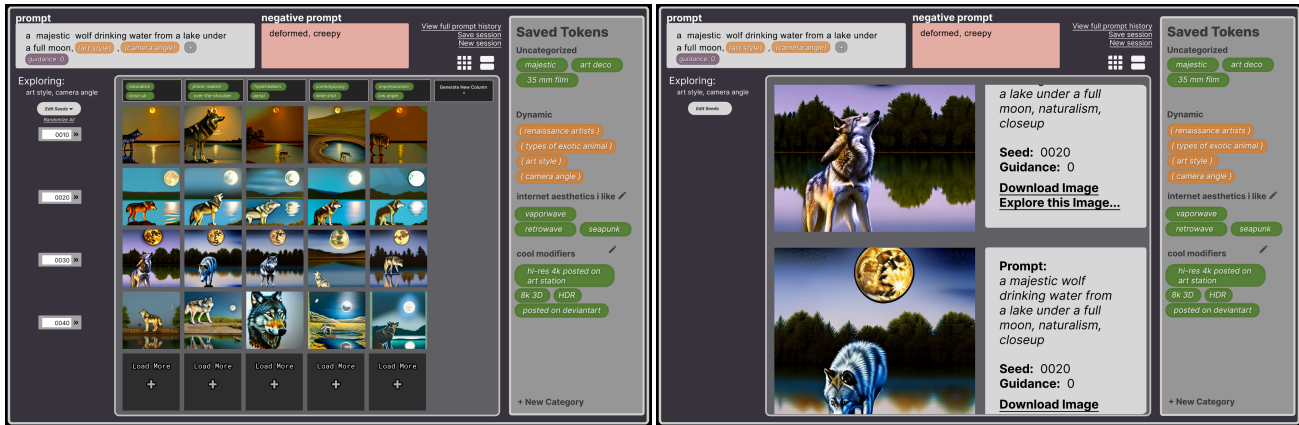


Figure 12: Views of the Dreamsheets 2.0 UI Mockup, showcasing the two visual layouts that users can freely toggle between: a grid view (left) and a focused list view (right). DREAMSHEETS 1.0 users valued the “small multiples” contact sheet layout, but viewing results in detail required them to change “zoom” settings or tediously adjust the sizes of columns and rows. An improved UI could provide “scalable” displays to support TTI users as they move between broad and focused evaluations of results.

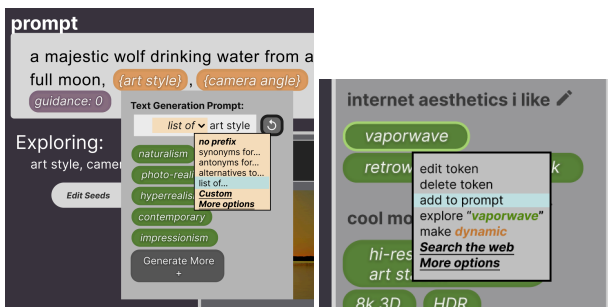


Figure 13: Views of elements featured in the interactive UI Mockup, including support for prompt templates with *dynamic* tokens, allowing participants to flexibly generate semantic variations while eliminating the need to write cell concatenation formulas. The LLM supporting this interaction is surfaced; users can control how the prompt is sent. Tokens can be saved for future reuse.

different forms of input to the TTI model. As participants requested, future work should investigate how DREAMSHEETS could be extended to support these approaches.

8 CONCLUSION

Text-to-Image models challenge users to navigate vast, opaque design spaces on both sides—prompt input and image output. DREAMSHEETS provides a flexible, spreadsheet-based interface for users to author strategies to achieve creative goals, and facilitating sensemaking—developing through experience the language and working understanding needed to reliably steer image generations towards interesting outputs. Through two user studies, including an extended expert study, we observed challenges, tensions, and opportunities in the TTI prompt-exploration process. We utilized these

insights to develop a UI mockup, improved with participant feedback, and suggesting features for future supportive TTI exploration interfaces. Finally, we considered the implications of supporting users’ sensemaking in prompt-image space, and beyond.

9 DISCLOSURE

The authors used ChatGPT for minor copy editing tasks.

10 EXPERT ARTIST CREDITS

In consenting to participate in the study, participants gave explicit and informed consent for the data collected during the study, including their creative contributions, to be anonymously published in research findings. That said, the authors believe that the AI and research community should strive to credit artists when their work is used, according to their wishes. After acceptance, the authors reached out to the expert generative artist participants to collect their preferences with regard to remaining anonymous or receiving credit in the final publication.

- (1) Expert Participant 1 (E1) is Stephen Young, a mixed media generative artist with 2 years of AI experience. He’s worked with multiple GAN and diffusion models to create art available on his website <https://www.kyrick.art>, @kyrick.art on Threads and @kyrickyoung on X.com.
- (2) Expert Participant 2 (E2) is Jeremy Torman, an interdisciplinary artist/musician that has been painting for over 20 years and using generative ai tools since 2016. He has used GANs, deepdream, style-transfer, vqgan+clip, jax diffusion, deforum, et al. He shares his work online @tormanjeremy on X/Twitter and @jeremy_torman on Instagram.
- (3) Expert Participant 3 (E3) is Seth Niimi, a multi-passionate creator who explores unique ways to combine tools, processes and ideas in pursuit of fascinating experiences. Seth can be found on TikTok and Instagram as @synaesthory.
- (4) Expert Participant 4 (E4) chose to remain anonymous.

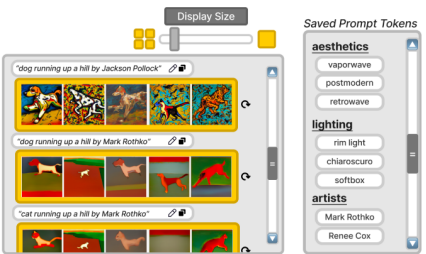


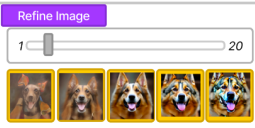
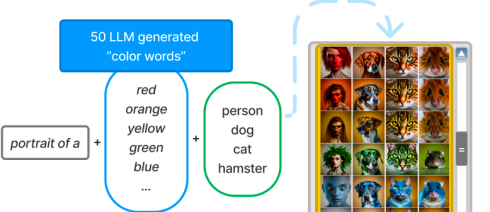
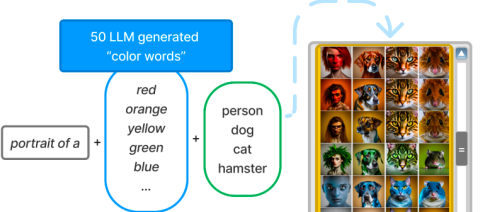
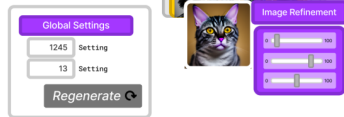
Exploration Strategy Observed in DreamSheets 1.0	UI Features Proposed & Co-Designed with DreamSheets 2.0	Example UI Components
A “Small Multiples” Results Evaluation	Scalable Results Display (Small-Multiples Grid View + Focused List View)	
B Reusable Exploration History	Prompt Token Bank, Session History, Modular & Reconfigurable Histories	
C Stochastic Transformation	Global Generation Settings, Simultaneous Regeneration, Infinite Exploration Expansion	
D Non-Stochastic Transformation	Image Refinement Controls (ex: Image-to-Image Interpolation, Spatial Controls)	
E Manual Semantic Exploration	Composable Prompt Templates with User-Authored Dynamic Token Series	
F Generative Semantic Exploration	LLM-Supported Dynamic Token Generation	
G Structured Multidimensional Exploration	Flexibly Composable, Reconfigurable Exploratory Units	

Figure 14: By observing the exploratory sheet-systems developed by experts in DREAMSHEETS 1.0, we identified supportive strategies and structures. Through the DREAMSHEETS 2.0 Mock-Up Co-Design process, we elicited expert input and concrete suggestions to connect these ideas to generalizable UI concepts, which can be presented as composable “units” to support user-structured exploration in future interface designs.

(5) Expert Participant 5 (E5) is a Canadian generative artist known as sureai.i, with 2 years of experience in the TTI space. She has previously used a variety of tools, including Stable Diffusion and Midjourney, along with manual digital editing. Her work can be seen online at @sureailabs on X.com and @surea.i on Instagram.

ACKNOWLEDGMENTS

We would like to thank our reviewers for their time and effort; their suggestions were immensely helpful in improving the quality and clarity of this work. We would also like to thank our participants for their contributions to the studies.

REFERENCES

- [1] 2023. Documentation for Midjourney’s /describe Function. <https://docs.midjourney.com/docs/describe> Accessed: 2023-09-15.
- [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 (UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 96, 14 pages. <https://doi.org/10.1145/3586183.3606725>
- [3] Huiwen Chang, Han Zhang, Jarred Barber, AJ Maschinot, Jose Lezama, Lu Jiang, Ming-Hsuan Yang, Kevin Murphy, William T. Freeman, Michael Rubinstein, Yuanzhen Li, and Dilip Krishnan. 2023. Muse: Text-To-Image Generation via Masked Generative Transformers. arXiv:2301.00704 [cs.CV]
- [4] Minsuk Chang, Stefania Druga, Alexander J. Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie J Cai, and Michael Terry. 2023. The Prompt Artists. In *Creativity and Cognition*. ACM, Virtual Event USA, 75–87. <https://doi.org/10.1145/3591196.3593515>

- [5] Ed Huai-hsin Chi, Joseph Konstan, Phillip Barry, and John Riedl. 1997. A Spreadsheet Approach to Information Visualization. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology (UIST '97)*. Association for Computing Machinery, New York, NY, USA, 79–80. <https://doi.org/10.1145/263407.263513>
- [6] John Joon Young Chung, Shiqing He, and Eytan Adar. 2021. The Intersection of Users, Roles, Interactions, and Technologies in Creativity Support Tools. In *Designing Interactive Systems Conference 2021 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1817–1833. <https://doi.org/10.1145/3461778.3462050>
- [7] Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. 1992. Scatter/Gather: a cluster-based approach to browsing large document collections. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '92)*. Association for Computing Machinery, New York, NY, USA, 318–329. <https://doi.org/10.1145/133160.133214>
- [8] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative Sense-Making: Quantifying Interaction Dynamics in Co-Creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition (C&C '17)*. Association for Computing Machinery, New York, NY, USA, 356–366. <https://doi.org/10.1145/3059454.3059478>
- [9] Remi Durant. 2023. Artist Studies. <https://remidurant.com/artists/> Accessed: 2023-09-15.
- [10] Noyan Evirgen and Xiang'Anthony' Chen. 2022. GANzilla: User-Driven Direction Discovery in Generative Adversarial Networks. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–10.
- [11] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–18. <https://doi.org/10.1145/3290605.3300619>
- [12] Karl Friston. 2009. The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences* 13, 7 (2009), 293–301.
- [13] Pol Baladas Gerard Serra, Oriol Domingo. 2022. A programmable interface for creative exploration. *Machine Learning for Creativity and Design Workshop at the 36th Conference on Neural Information Processing Systems (NeurIPS 2022)* (December 2022). https://neuripscreativityworkshop.github.io/2022/papers/ml4cd2022_paper19.pdf
- [14] Elena L. Glassman, Tianyi Zhang, Björn Hartmann, and Miryung Kim. 2018. *Visualizing API Usage Examples at Scale*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3174154>
- [15] Björn Hartmann, Loren Yu, Abel Allison, Yeonsoo Yang, and Scott R. Klemmer. 2008. Design as Exploration: Creating Interface Alternatives through Parallel Authoring and Runtime Tuning. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (Monterey, CA, USA) (UIST '08)*. Association for Computing Machinery, New York, NY, USA, 91–100. <https://doi.org/10.1145/1449715.1449732>
- [16] Jeffrey Heer, Jock Mackinlay, Chris Stolte, and Maneesh Agrawala. 2008. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE transactions on visualization and computer graphics* 14, 6 (2008), 1189–1196.
- [17] Janin Koch, Nicolas Taffin, Michel Beaudouin-Lafon, Markku Laine, Andrés Lucero, and Wendy E. Mackay. 2020. ImageSense: An Intelligent Collaborative Ideation Tool to Support Diverse Human-Computer Partnerships. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 45 (may 2020), 27 pages. <https://doi.org/10.1145/3392850>
- [18] 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*. Association for Computing Machinery, New York, NY, USA, 407–418. <https://doi.org/10.1145/3357236.3395494>
- [19] Chinmay Kulkarni, Stefania Druga, Minsuk Chang, Alex Fiannaca, Carrie Cai, and Michael Terry. 2023. A Word is Worth a Thousand Pictures: Prompts as AI Design Material. <https://arxiv.org/abs/2303.12647v1>
- [20] Tomas Lawton, Kazjon Grace, and Francisco J Ibarrola. 2023. When is a Tool a Tool? User Perceptions of System Agency in Human-AI Co-Creative Drawing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 1978–1996.
- [21] 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*. 2257–2266.
- [22] Marc Levoy. 1994. Spreadsheets for images. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. 139–146.
- [23] Jingyi Li, Eric Rawn, Jacob Ritchie, Jasper Tran O'Leary, and Sean Follmer. 2023. Beyond the Artifact: Power as a Lens for Creativity Support Tools. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 47, 15 pages. <https://doi.org/10.1145/3586183.3606831>
- [24] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2021. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. arXiv:2107.13586 [cs.CL]
- [25] Vivian Liu and Lydia B. Chilton. 2021. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. <https://doi.org/10.48550/ARXIV.2109.06977>
- [26] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 384, 23 pages. <https://doi.org/10.1145/3491102.3501825>
- [27] Vivian Liu, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2023. 3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1955–1977. <https://doi.org/10.1145/3563657.3596098>
- [28] Bryan Loh and Tom White. 2018. SpaceSheets: interactive latent space exploration through a spreadsheet interface. (2018). <https://doi.org/10.26686/wgtn.17134886.v1>
- [29] Aran Lunzer and Kasper Hornbæk. 2008. Subjunctive Interfaces: Extending Applications to Support Parallel Setup, Viewing and Control of Alternative Scenarios. *ACM Trans. Comput.-Hum. Interact.* 14, 4, Article 17 (jan 2008), 44 pages. <https://doi.org/10.1145/1314683.1314685>
- [30] Grace Luo, Trevor Darrell, Oliver Wang, Dan B Goldman, and Aleksander Holynski. 2023. Readout Guidance: Learning Control from Diffusion Features. arXiv:2312.02150 [cs.CV]
- [31] J. Marks, B. Andalman, P. A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J. Seims, and S. Shieber. 1997. Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., USA, 389–400. <https://doi.org/10.1145/258734.258887>
- [32] Aditi Mishra, Utkarsh Soni, Anjana Arunkumar, Jinbin Huang, Bum Chul Kwon, and Chris Bryan. 2023. PromptAid: Prompt Exploration, Perturbation, Testing and Iteration using Visual Analytics for Large Language Models. arXiv:2304.01964 [cs.HC]
- [33] Lars Nielsen. 2022. An advanced guide to writing prompts for Midjourney (text-to-image). *Mlearning.ai* (Sep 2022). <https://medium.com/mlearning-ai/an-advanced-guide-to-writing-prompts-for-midjourney-text-to-image-aa12a1e33b6> Accessed: 2023-09-15.
- [34] Jonas Oppenlaender. 2022. The Creativity of Text-to-Image Generation. In *Proceedings of the 25th International Academic Mindtrek Conference (Academic Mindtrek '22)*. Association for Computing Machinery, New York, NY, USA, 192–202. <https://doi.org/10.1145/3569219.3569352>
- [35] Jonas Oppenlaender. 2023. The Cultivated Practices of Text-to-Image Generation. arXiv:2306.11393 [cs.CY]
- [36] Jonas Oppenlaender. 2023. A taxonomy of prompt modifiers for text-to-image generation. *Behaviour and Information Technology* (Nov. 2023), 1–14. <https://doi.org/10.1080/0144929x.2023.2286532>
- [37] Jonas Oppenlaender, Rhema Linder, and Johanna Silvennoinen. 2023. Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering. arXiv:2303.13534 [cs.HC]
- [38] Guy Parsons. 2022. *The DALL-E 2 Prompt Book*. <https://dallery.gallery/the-dalle-2-prompt-book/>
- [39] pharmapsychotic. 2023. CLIP Interrogator: A prompt engineering tool for text-to-image models. <https://github.com/pharmapsychotic/clip-interrogator> Accessed: 2023-09-15.
- [40] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological review* 106, 4 (1999), 643.
- [41] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.
- [42] proxima centauri b, sure.ai, Stephen Young, and Encyclopedia Erratica. 2023. *Image Synthesis Style Studies Database by parrotzone.art*. https://web.archive.org/web/20230325013519/https://docs.google.com/spreadsheets/d/14xTqtuV3BuKDNhLotB_d1aFIBGnDJOY0BRXJ8-86GpA/edit#gid=0
- [43] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical Text-Conditional Image Generation with CLIP Latents. arXiv:2204.06125 [cs.CV]
- [44] Eric Rawn, Jingyi Li, Eric Paulos, and Sarah Chasins. 2023. Understanding Version Control as Material Interaction with Quickpose. In *Proceedings of the SIGCHI conference on human factors in computing systems*.
- [45] Mitchell Resnick, Brad Myers, Kumiyo Nakakoji, Ben Shneiderman, Randy Pausch, Ted Selker, and Mike Eisenberg. 2005. Design Principles for Tools to Support Creative Thinking. (1 2005). <https://doi.org/10.1184/R1/6621917.v1>
- [46] 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*. 165–174.
- [47] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-Resolution Image Synthesis With Latent Diffusion Models.

- Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2022), 10684–10695.
- [48] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. The Cost Structure of Sensemaking. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*. Association for Computing Machinery, New York, NY, USA, 269–276. <https://doi.org/10.1145/169059.169209>
- [49] Andy Salerno. 2022. *4.2 Gigabytes, or: How to Draw Anything*. <https://andys.page/posts/how-to-draw/>
- [50] Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Tali Bers, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M. Rush. 2021. Multitask Prompted Training Enables Zero-Shot Task Generalization. <https://doi.org/10.48550/ARXIV.2110.08207>
- [51] Ben Shneiderman. 2002. Creativity Support Tools. *Commun. ACM* 45, 10 (oct 2002), 116–120. <https://doi.org/10.1145/570907.570945>
- [52] Sangho Suh, Bryan Min, Srishti Palani, and Haijun Xia. 2023. Sensecape: Enabling Multilevel Exploration and Sensemaking with Large Language Models. *arXiv preprint arXiv:2305.11483* (2023).
- [53] Jerry O. Talton, Daniel Gibson, Lingfeng Yang, Pat Hanrahan, and Vladlen Koltun. 2009. Exploratory Modeling with Collaborative Design Spaces. In *ACM SIGGRAPH Asia 2009 Papers* (Yokohama, Japan) (*SIGGRAPH Asia '09*). Association for Computing Machinery, New York, NY, USA, Article 167, 10 pages. <https://doi.org/10.1145/1661412.1618513>
- [54] Edward R Tufte. 1993. *Envisioning information*. Vol. 199. Graphics Press Cheshire, CT.
- [55] Robert F. Woodbury and Andrew L. Burrow. 2006. A typology of design space explorers. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 20, 2 (2006), 143–153. <https://doi.org/10.1017/S0890060406060136>
- [56] Stephen Young. 2022. COG Prompt Parrot. <https://github.com/kyrick/cog-prompt-parrot>
- [57] Loutfouz Zaman, Wolfgang Stuerzlinger, Christian Neugebauer, Rob Woodbury, Maher Elkhaldi, Naghmi Shireen, and Michael Terry. 2015. GEM-NI: A System for Creating and Managing Alternatives In Generative Design. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 1201–1210. <https://doi.org/10.1145/2702123.2702398>
- [58] JD Zamfirescu-Pereira, Heather Wei, Amy Xiao, Kitty Gu, Grace Jung, Matthew G Lee, Bjoern Hartmann, and Qian Yang. 2023. Herding AI Cats: Lessons from Designing a Chatbot by Prompting GPT-3. (2023).
- [59] J.D. Zamfirescu-Pereira, Richmond 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 (CHI'23)*.
- [60] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (2023), 3813–3824. <https://api.semanticscholar.org/CorpusID:256827727>