

Doom or Deliciousness: Challenges and Opportunities for Visualization in the Age of Generative Models

V. Schetinger¹ , S. Di Bartolomeo² , M. El-Assady³ , A. McNutt⁴ , M. Miller⁵ , J. P. A. Passos⁶ , and J. L. Adams² 

¹TU Wien, ² Northeastern University, ³ ETH AI Center, ⁴ University of Chicago, ⁵ University of Konstanz, ⁶ Hugging Face

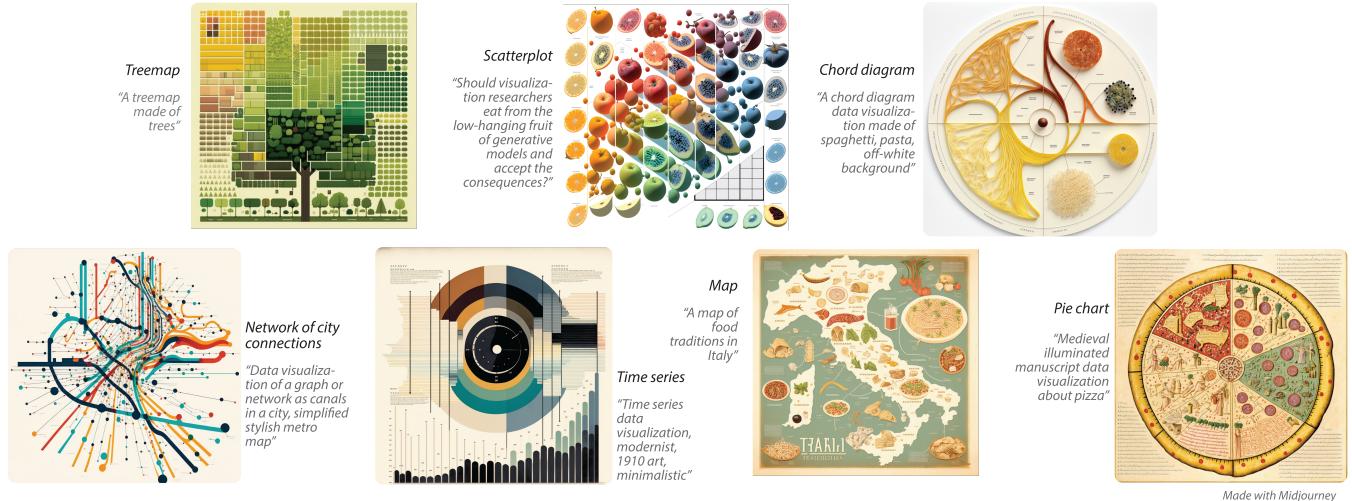


Figure 1: Different types of data visualization (and the prompts used to create them) as imagined by text-to-image generative models. While these examples are delicious (in their graphical intrigue and exploration of style), they risk doom by potentially creating misplaced trust.

Abstract

Generative text-to-image models (as exemplified by DALL-E, MidJourney, and Stable Diffusion) have recently made enormous technological leaps, demonstrating impressive results in many graphical domains—from logo design to digital painting to photographic composition. However, the quality of these results has led to existential crises in some fields of art, leading to questions about the role of human agency in the production of meaning in a graphical context. Such issues are central to visualization, and while these generative models have yet to be widely applied in visualization, it seems only a matter of time until their integration is manifest. Seeking to circumvent similar ponderous dilemmas, we attempt to understand the roles that generative models might play across visualization. We do so by constructing a framework that characterizes what these technologies offer at various stages of the visualization workflow, augmented and analyzed through semi-structured interviews with 21 experts from related domains. Through this work, we map the space of opportunities and risks that might arise in this intersection, identifying doomsday prophecies and delicious low-hanging fruits that are ripe for research.

1. Introduction

Generative models are increasingly prominent in many domains. There has been astonishing technological development among such tools, such as the text-to-image generation of DALL-E 2 [Ope22a], Stable Diffusion [RBL*22], or MidJourney [Mot22], as well as the textual synthesis found in tools like GPT-3 [Ope22b] or Copilot [Git22]. These systems transform a (typically textual) prompt and into an entity (such as an image or text) drawn from their learned representations of training data. These tools have been hailed for their rapid development and high-quality results.

Yet, this evolution has not been without friction. The automated creation of images and text raises questions about the role of human agency in the production of meaning in a graphical context, creating tensions in professional [Plu22], legal [Vin22], and artistic [Roo22] contexts. For instance, the term *Bach Faucet* [Com22] captures the paradox that the existence of a machine that can produce arbitrarily many high-quality instances of an artistic style reduces the value of that artistic style—inspired by an early success in generative art that was able to produce sonatas in the style of Bach [Cop04]. Like other models, generative models replicate biases present in their training data, which might yield racist imagery [Ope22c].

Despite these challenges, these models have the potential to improve a wide variety of visualization workflows, such as by increasing the speed of production, facilitating creativity, and enabling expression. However, there has been little investigation into the role these generative models might play in visualization. Wood [Woo22] called for their use as a way to break away from the “*walled garden*” of structured visualization, while others have explored using the visualization generation as means to explore the possibility space of highly stylized visualizations [For22, SFP*22] or validating visualization usage [WTC22]. While some are skeptical [Ma23] about the holistic utility of such models, we believe that their wider spread utilization is on the nearby horizon.

This work seeks to circumvent the dilemmas found in other domains’ interaction with generative models by trying to understand this landscape before it emerges. In support of this goal, we seek to answer:

What challenges and opportunities might we expect to find as use of generative models becomes commonplace in visual design and analytical workflows?

We answer this question by seeking the concerns, opinions, and predictions of domain experts ($N=21$) from visualization, machine-learning, art, and art history in a semi-structured interview study (Sec. 3). Through this study, we elicited beliefs about the risks and opportunities posed by the generative models. Our interview study principally focuses on image-based generative models (IGMs); however, there are a wide variety of other generative models, such as text-based ones (e.g. chatGPT [Ope23]).

We analyze these findings by locating challenges and opportunities available at each stage of a standard visualization pipeline (Sec. 4). We find that participants believed that the use of generativity in visualization was promising. For instance, it may allow us to capture ephemeral aspects of visualization (such as emotion), or to create artistically rendered visualizations (as in Fig. 2), or to increase the speed with which visualization designers can rapidly prototype their scientific and graphical communications. Such models also offer ample risk for visualization. For instance, there was substantial concern over the proclivity of such tools to amplify bias and to parrot components of their training data (in such a manner that was not respectful of copyright). We extend these concerns and highlight how future work might ameliorate such issues. In addition, our supplemental material is available on osf.io.

In conducting this study, we seek to provide a forward-looking foundation of how these models might be used—both to guide future tool construction, but also to steer design of the models themselves. Per Bender and Koller [BK20], this work aims at highlighting the right hills to climb and valleys to watch out for.

2. Background and Related Work

From the creation of Generative Adversarial Networks (GANs) in 2014 [MO14] to new text-to-image tools in 2022, such as DALL-E 2 [RDN*22], Imagen [SCS*22], Stable Diffusion 2.0 [RBL*22], and MidJourney [Mot22], the development of image-based generative models (IGMs) has been progressing at breakneck speed, quickly entering many domains. We posit that it is reasonable, too, to expect the

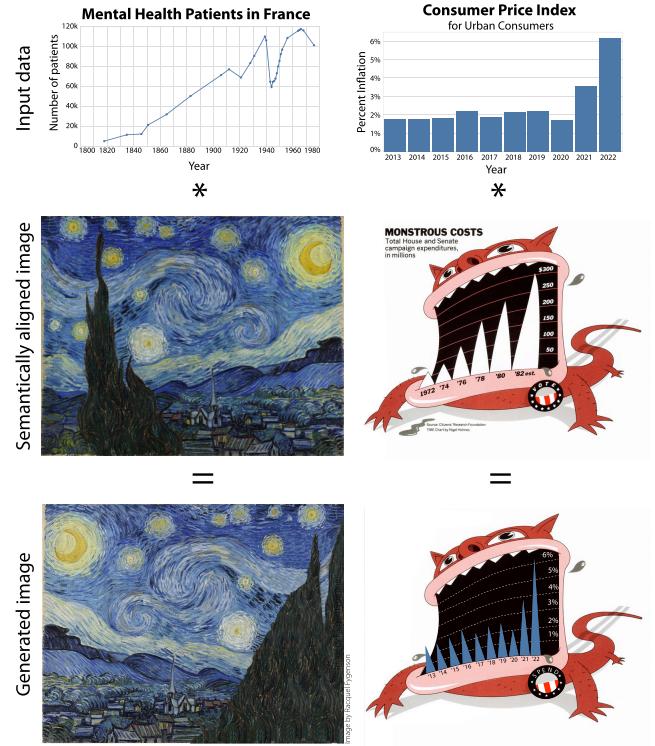


Figure 2: Future IGMs will likely be able to incorporate data or charts as part of their generation process. Racquel Fygenson conceptualized using an IGM to blend data with a thematically relevant image (left). IGMs could also be used to restyle well-known visualizations (right). Data from [CC10] (left) and [OLS23] (right).

widespread adoption of generative technologies in visualization creation within the coming decade. Similar studies as our own have been carried out in other domains, including newsrooms graphics [LQC22], visual marketing [MV22], computational notebooks [MWDD23] and construction scheduling [PMdS23]. Our study differs from these in its focus on the particular visualization practice and design.

Our work is situated among prior studies on generative models, particularly those focused on using generative models to enhance the overall process of visualization (“Gen4Vis”).

The use of generative tools in the development of data visualizations has been growing. Often, in traditional visualization pipelines, a designer has to either be proficient in a programming language, or translate their ideas into tool-specific operations [SSL*22], which makes for a steep learning curve. A number of tools seek to simplify the process, using either visual interfaces [MC21], or, more recently, natural language [SLJL10, MS23, DBSSD23, WCA23] interfaces, which allow users to produce visualizations by simply typing or speaking their questions or requests. Recent surveys [WCWQ22, WWS*22, WH22] have explored how machine learning is being applied to the data visualization process. Our work builds on these by looking forward and focusing on generative tasks. Wang et al. focused on dividing applications of ML for visualization into data, visualization, and user [WCWQ22]. The survey presents these elements as modular

Table 1: The predefined potential usages (above the fold) and concerns (below the fold) that participants considered in our study. Items were derived by reviewing common concerns about generative models and by reflecting on our own experience these tools.

Prompt	Description	Question
Design prettyfication	Using a generative model to improve the appearance of a visualization, as in style transfer [SLC*22].	Fig. 4, 5A
Embellishment	Automatically introduce visual elements not bound to data. E.g. legends, annotations, decorations, or chart junk.	Fig. 4, 5A
Chart recommendation	Generate charts based on the semantics, types, or shape of the data.	Fig. 4, 5A
Rapid prototyping	Iteratively generate designs to explore different possibilities.	Fig. 4, 5A
Moodboards	An automatically generated visual collage of options for exploring a possibility space.	Fig. 4, 5A
Visualize training data	Aid the creation of generative models by curating content in training data and controlling parameters.	Fig. 4, 5A
Amplifying biases	IGMs, like all ML-models, are susceptible to bias based on their training data. For instance, a prompt asking for a doctor might only produce images of white men [Ope22c].	Fig. 5B
Untrustworthy results	Charts generated by probabilistic models may not accurately represent the input data or may manipulate it in subtle or confusing ways	Fig. 5B
Replicated private data	Models are well known to parrot [BGMMS21] their training data, and these models run a similar risk.	Fig. 5B
Rip off existing vis style	Just as IGMs are criticized for replicating artists' styles [Hei22], these might try to parrot well known graphics.	Fig. 5B
Replacing vis designers	Sufficiently advanced models might be able to replace human visualization designers	Fig. 5B

components of a prospective visualization process pipeline, which has informed our thinking in Sec. 4. Data transformations such as dimensionality reduction could be applied prior to visualization [WFC*18]. Machine learning might be applied to a dataset to automatically generate Vega-Lite visualization specifications [DD18], to highlight issues with those specifications [WTC22], or to suggest the automatic stylistic transfer of graph drawing from one example layout to another [WJW*20]. User-centered interventions might include user profiling, such as predicting a user's next click [OGW19] or anticipating their visual attention across an infographic [BKO*17, WCWQ22].

The dual of using generativity for visualization (“Vis4Gen”) has also been well studied. These prior works cover topics such as how surrogate models can be used to enhance interpretability [MQB19], how visualization can make high-dimensional structures like embeddings more comprehensible [STN*16], and how exposing the internals of black box models can make them easier to understand [HPRC20]. This dichotomy between these approaches informs an axis of our interview study (Fig. 4)—“Gen supports VIS” vs “VIS supports Gen”.

3. Methodology: Semi-Structured Interview Study

To better understand the challenges and risks that may arise through the integration of generative models into visualization workflows, we conducted a semi-structured interview study that sought to elicit forecasts, opinions, fears, hopes, and concerns arising from experts from various domains. Given the rapid development cycles of generative models, our goal in this study is to prospect the space around visualization—rather than specifically identify the next trends. Our study is informed by the belief that this goal would be best aided by those with expertise in this or related domains.

Study Participants. We conducted interviews with 21 participants, which lasted an average of one hour. Interviews were conducted remotely over Zoom. Participants were drawn from a convenience sample assembled based on their work in relevant domains. In particular, we consulted experts with backgrounds in art or art history ($N=5$), machine learning ($N=2$), and visualization or HCI ($N=14$). We denote each of these backgrounds in participant identifiers as PX_{art} , PY_{ML} , and PZ_{vis} respectively. Given their central relevance to the topic, most participants' primary background was in visual-

ization, however, most lacked experience with generative tools, and so we supplemented their opinions with those of experts drawn from other fields. We included a participant if they had sufficient expertise in their relevant domain—as demonstrated by holding or pursuing a post-graduate degree or a substantial history of working in their field. We conducted two pilot interviews prior to the study. Fortunately, the results were sufficiently similar to the other interviews to allow inclusion in our analysis. See appendix for details of self-reported backgrounds. Additional demographics, such as regarding experience with generative models or with visualization, were not formally collected. Participants were not compensated.

Study Procedure. Interviews consisted of three phases, which, sequentially, sought to elicit (i) participant self-identification (Fig. 3), (ii) opportunities that they foresaw, and (iii) risks posed. The discussion was focused through a shared Miro board (see appendix for examples), in which participants were invited to create post-its about ideas and place them on several axes denoting different views (Fig. 4, 5). These predefined prompts (which we list in Tab. 1) were meant to help users ideate on these topics rather than being a comprehensive list of all possible applications or issues. In addition to prompts, participants were invited to create their own and thereby ruminate on their concerns and aspirations for these tools. These discussions were interleaved with situating prompts that reviewed various function-

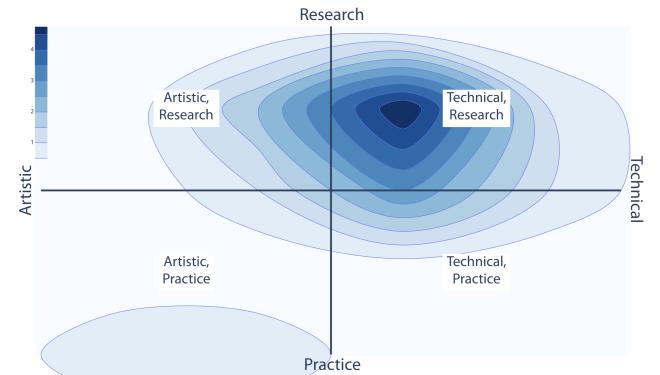


Figure 3: Aggregate participant ratings for self identification.

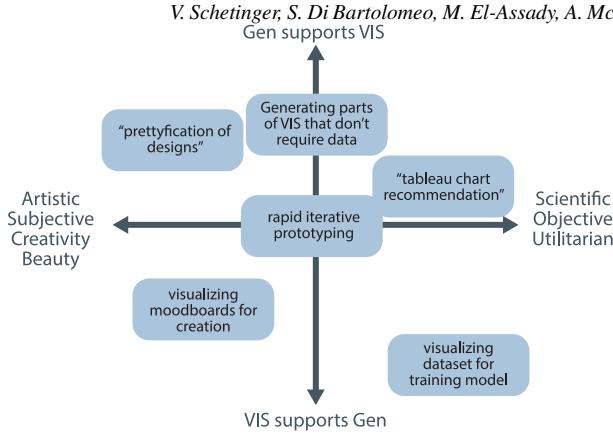


Figure 4: Participants were invited to locate potential usages of IGMs for visualization in a space that sought to elicit how much a concept (such as “*rapid iterative prototyping*”) was artistic or supported visualization vs. supported generativity (per Sec. 2).

alities possessed by current tools (such as in-painting [YQS20] and out-painting [XLC20]) to help situate their thinking among current technologies. We utilized this shared creative space to invite thinking aloud and ideation, which may have been more limited in a more restricted environment. Participants brought up a variety of additional topics (such as semantic chart recommendation), which we discuss in the next section. A complete reproduction of each participant response is available in the appendix.

Through these discussions, participants rated various potential concerns and usages of IGMs on either four axes (purpose, objectivity, excitement, and feasibility) or two axes (concern and likelihood), depending of if we were discussing a potential or a usage, respectively. We focused on these axes (and therein aspects) and set of initial topics (Tab. 1) because they might elicit thoughtful discussion of fears and hopes for these technologies, rather than being an assertion about the relationship between such concepts more generally.

Result Coding. Interviews were automatically transcribed, and then were coded by two of the authors. Results and themes were then iteratively discussed among the rest of the team in order to form our discussion, which we present in the next section. Given the speculative nature of our work, we sought to locate our findings around a generalizable model, and so we modified a standard visualization pipeline in support of the resultant themes. We augmented this analysis by quantizing the placement of each usage or concern and normalizing it on a unary axis, as shown in Fig. 6. The full results can be found in the appendix, with aggregates presented in Fig. 9.

4. Analysis: Help and Harm in the Visualization Pipeline

We now analyze our interviews to identify challenges and opportunities. Per our analysis methodology, described in the previous section, we locate our results within an adapted standard visualization pipeline model [MKC20, Chi00], which we show in Fig. 7. This pipeline consists of four transformations: condensing the world into data by **Data-fying** it (Sec. 4.1), **Transforming** (Sec. 4.2) that data into something usable, changing that processed data into an image by **Visualizing** (Sec. 4.3), and then using or **Interacting** (Sec. 4.4) with that rendered image. These stages are drawn from McNutt

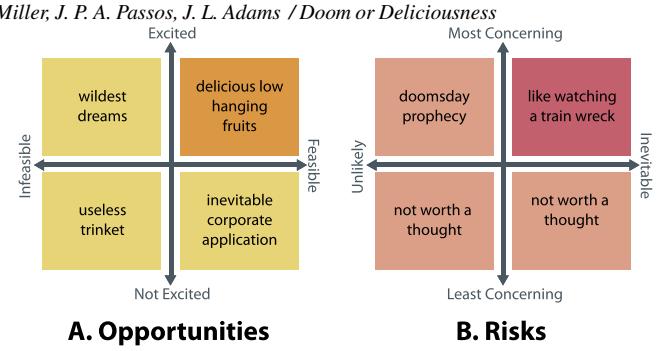


Figure 5: Participants were invited to place both predefined potential usages and concerns for IGMs in visualization (Tab. 1) and self-defined suggestions onto these axes as a means of ideation. Each quadrant was given a provocative name based on its meaning to help the understanding of the task.

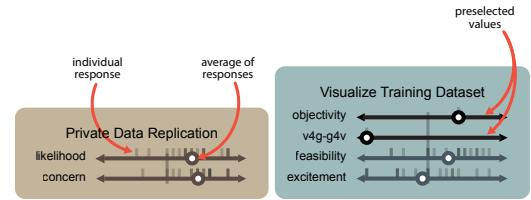


Figure 6: Participant ratings for a potential usage (visualizing training data as part of model construction) and a concern (private data replication). Ticks indicate individual responses, while circles indicate averages. Throughout this paper blue boxes show potential usages, while brown boxes show concerns.

et al.’s [MKC20] mirage pipeline, which we refer to for specific definition. While other workflow organizations (such as the KDD model [FPSS96]) might be used, we select this one both to be generalizable but also to be in dialog with other works on applying AI-tools to visualization [WWS*22]. Generative models entail opportunities for useful intervention at each stage of this pipeline.

We further organize our observations following a modification of the SWOT-strategic management [BEM*21], which enumerates the **Strengths** and **Weaknesses** arising from inside an organization, as well as the **Opportunities** and **Threats** from outside. This approach is often used as part of the decision-making process for businesses or other organizations. We adapt this analysis by partitioning by whether the opportunities and risks arise from the visualization design side (producer) or the application domain side (consumer). In doing so, we seek to highlight the speculative decision of whether to integrate or implement a particular feature and thereby explore the possibilities available at each stage.

We next traverse our pipeline and highlight the potential that generative models hold for each stage, using the above color scheme to denote risks and opportunities. To support these discussions, we show participant ratings of various potential usages and concerns, such as in Fig. 6.

The majority of this discussion focuses on information visualization-focused tasks. In addition to the graphical basis of the medium, these results were driven by the cultural prominence of IGMs when we conducted the study. Had we conducted the study

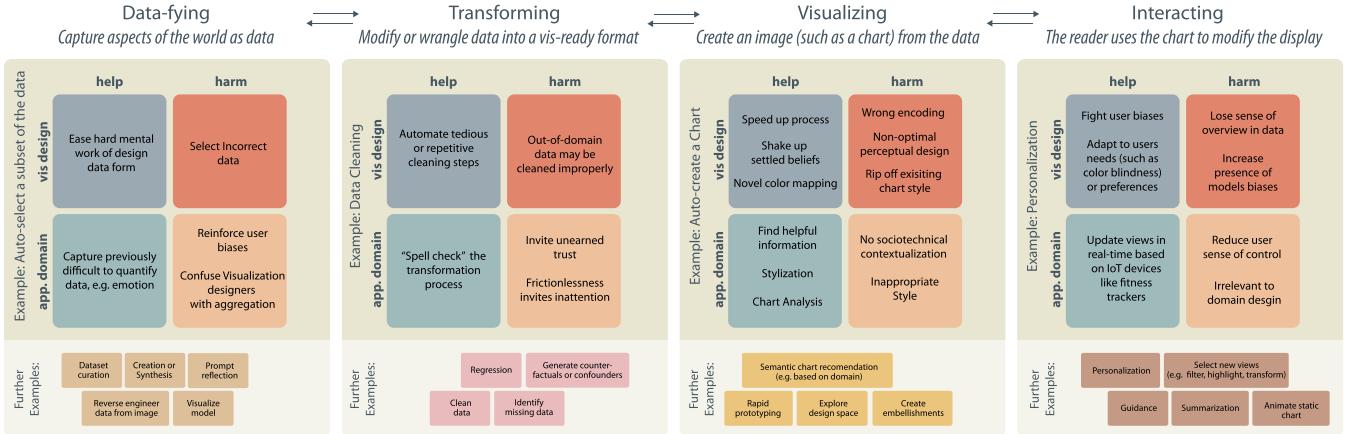
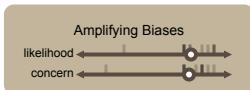


Figure 7: There are opportunities to apply generative models throughout the visualization pipeline. Here, we highlight some of the help and harm that use of generative models can do for particular example tasks at each stage of our broadly organized pipeline.

several months later, participants may have focused on the capabilities of chatbots like (ChatGPT [Ope23] or Bard [Goo23]) rather than those of IGMs (like DALL-E 2 [Ope22a]). In addition, our participant pool was mainly drawn from those without experience in domains like Scientific Visualization or animation, which caused our results to be skewed away from such topics. Finally, these considerations over-emphasize the visualization component of this pipeline (likely motivated by our context as trying to identify opportunities for visualization). However, we note that the other stages also offer ample opportunities for generative intervention.

4.1. Data-fying

Data is not a natural resource, and so its entry into a visualization workflow begins by conceptualizing what that data can be [MKC20]. This process can be time-consuming and difficult, as data modeling is notoriously challenging [RS19]. A generative model might aid this process by helping the user identify what counts as data (or just by defining the data workspace subsequent analysis will occur in), such as by preparing a SQL query, suggesting a data model, finding a relevant data set, or otherwise injecting its own knowledge [HHS*22]. For instance, a natural language prompt such as “get data from earnings of the fourth quarter relative to our competitors” offers ample ambiguity and opportunity for the model to exert agency, such as in identifying who the competitors are, how to compute earnings data, which business units to include, and so on. Huang et al. employ this approach for flow visualization using NLP [HXHT22]. These moments of agency invite risk (e.g., the model might provide bad results or exhibit biases) [GSS*22], but they also invite opportunity.



Consider **creating a subset of data** (Fig. 7 left). In such a task, the user asks the model to create a working subset of the data, such as by generating an SQL query for a large database. Such automation might reduce the burden of designing a potentially difficult query and thereby speed up the analysis or design process. In addition, this might also allow for the system to synthesize or create data not present in the original dataset, but that might be inferred from context. These include quantities such as emotions or other ephemeral aspects of the

dataset [Woo22]. Such an exchange of agency carries with it potential for harm; for instance, it invisibly **surfaces the models’ biases**. If an idea is out of the scope of the training data, or if a type of person is biased against, that bias will be embedded into the data, possibly without labeling—an error that can cascade down the pipeline, affecting every subsequent data interaction. Reciprocally, the use of a model to construct data might also **manifest the biases of the user**, such as through confirmation bias. Participants believed that amplifying biases was an inevitable consequence of using these models. For instance, **P21_{ML}** noted that “*when you model data using a normal distribution, and then you sample from this normal distribution, you will always have a bias because things that sit closer to the average will have a higher probability of being selected*”.

Similarly, there were concerns about the model parroting its training dataset (Fig. 6 right). Some of these stemmed from anxieties about using the material in ways not permitted by the licenses of the data (an issue which is being examined in the American court system [Vin22]), and the inability to track down the sources used for the generation. For instance, **P6_{vis}** observed that “*If you are doing a chart, and you want a simple icon graphic to represent certain subjects, that would be the most common use case, but I have my worries about copyright that comes along with it. If I generate an icon, I don’t know if there are any ramifications of me using it in my paper, and that might get me in trouble. So I would be uncomfortable. ... I’d be worried that I’m ripping somebody off.*” Such fears are difficult to address in a text-based medium, but it would be exceptionally quarrelsome to try to check if a particular image has been derived from another, even when using modern reverse image look-up tools. The recent tool *Have I been trained?* [Spa22] tries to address the issue by offering a lookup function into the LAION dataset, used to train Stable Diffusion [RBL*22]. A review of the search “data visualization” on this database returns a number of graphics with identifiable authorship, as well as graphics that violate common mores of visualization, e.g. 3D bar charts showing occlusion, distorted iconography, distracting gradients, and cluttered interfaces—all concerns listed in Tab. 1.

A variety of other tasks might be fulfilled at this stage. **P1_{vis}** described the potential of **automatically reverse engineering data from an image**. This work has been explored previously [PH17]. However, using a generative model might allow better extraction of data

(not based on a predefined schema), but it may also introduce new, more difficult-to-address or identify forms of bias. Generative models could be used to construct useful situationally-aware prompts (such as in the style of LitVis [WKD18]) that can help users reflect and consider how they are enshrining bias into their models of data. Finally, generative models are trained with large datasets (Fig. 6 left), and their results can greatly vary depending on this training. Curating input data can help reduce biases, personalize models, and increase the general quality of results. The complexity latent to this task type could be allayed by a generative model, which might be able to act as an assistant, such as Hynes et al.’s data linter [HST17], which could provide more dynamic or situational suggestions than heuristically-motivated training-data analysis tools.

4.2. Transforming

Once the data is established, it needs to be transformed into a form that might be pliable for visualization. This process can involve wrangling, processing, introducing additional models (such as through regression or other advanced analytics), or countless additional approaches. Within each of these interleaved steps and stages, there are opportunities for a generative model to exert agency. For instance, prompts like “remove outliers” or “find missing data” presupposes particular structures of the data. While those might sometimes be unambiguous (for instance, a monthly calendar missing data from the weekends has clearly defined gaps), in other cases, it might be less clear (for instance, removing outliers might assume a particular model of the data distribution). Consider **cleaning data** (Fig. 7 center-left); this stage might require removing duplicates or outliers, and converging on consistent naming—among a host of other sub-activities. These repetitive processes prove to be something that can largely be automated [KW19], and so a generative model might usefully intervene by automatically creating transformations. Tools like Copilot [Git22] can already effectively assist in such a process when there is a human-in-the-loop to help guide and curate the synthesis. Just as in the previous stage of our pipeline, yielding agency to a model runs the risk of introducing the model’s biases into the generated content, which **P9_{art}** and **P17_{vis}** were particularly worried about. In addition, it may also risk out-of-domain schemas which, noted **P21_{ML}**, may cause the suggestions to be irrelevant or incorrect. Further, the potential frictionlessness of this process may cause the user to overly trust the synthesized results, a huge issue for many participants (**P11_{vis}**, **P12_{vis}**, **P17_{vis}**, **P19_{vis}**), which may lead to difficulties in identifying errors down the line. In addition to acting as a tool to automated transform creation, generative models might also be used to evaluate whether or not a transformation was done well or if any steps were missing, in something akin to a spell checker for data analysis, which might be analogous to Wu et al.’s [WTC22] use of an LLM to analyze Vega-Lite charts. While current LLMs do not consider data as part of their synthesis process, future models might do so to ensure that suggestions are appropriate to the data and task to be pursued (as in Fig. 2).

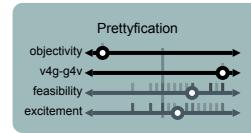
In addition, to the general task of cleaning data, generative models might be used at each incremental step of the cleaning pipeline, such as in creating regressions or identifying missing data. **P11_{vis}**, for instance, was interested in augmenting financial datasets for research purposes: “*getting data of financial transactions can be very hard, which makes it hard to test our techniques*”. These pose similar

benefits (like increasing the speed to accomplish a task or showing an unknown functionality) and risks (furthering user and model biases) as in the general case.

4.3. Visualizing

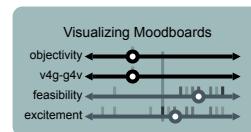
The most prominent stage in our pipeline is the visualization stage. Here, the modeled data is mapped to a visual encoding which will subsequently be presented to the user. As our work and participant pool is centered around visualization, participants observed the most risk in this domain but also the most opportunities. For instance, some participants (**P4_{vis}**, **P6_{vis}**, **P10_{art}**, **P12_{vis}**) highlighted the potential for harm in found graphics that do not handle sensitive content carefully (which is aligned to McNutt et al.’s [MHK21] descriptions therein). Most common models contain some form of safety system against pornography, gore, and deep fakes, but they are not perfect and can sometimes be circumvented by resourceful users [Tay23]. Yet, the same free-form nature that can be potentially harmful also offers a deep well of potentially unfamiliar ideas and suggestions that might unsettle settled ideas.

One such example is the **beautification of visualization designs**, which while mostly aesthetic, could also be functional. From the selection of proper colors based on semantics [EAKM*22, HYC*22] to more extreme visual transformations, it was considered



by participants to be mostly “*subjective*” (**P15_{vis}**). However, it was recognized that as with Coelho’s [CM20] “Infomages”, they do change the impact and interpretation of a visualization. For example, **P17_{vis}** commented on the value of beautification by considering the “*entertaining value of generative models for engaging users*” (**P17_{vis}**). Works like BeauVis [HIDI22] could be used to orient the training of models for “*aesthetic amplification*”. Participants expressed enthusiasm towards this aspect but also doubts its feasibility. Most of the concerns related to *lacking trust in the result of the model*—indeed, **P8_{vis}** explained, there would be no way of knowing if the change in design would affect the data represented in the resulting visualization, or that the change would be appropriate assuming the process is done through a model that has no awareness of the data and the context of the user. In addition, they may also be useful for the construction of color maps that are adapted to the input data, such as for medical imaging or cartographic applications. These issues are further discussed in Sec. 5.1.

Another example is **Moodboard creation**; an application where visualization supports the use of generative models. **P13_{art}** summarized this application as being like a “*Pinterest in the generative space*”. Many creative tasks begin



with constructing mood boards, or a collage of images and text for inspiration and ideation. Generative models, aided by visualization and interactive techniques to navigate and query the latent space could become an infinite source of inspiration. Most participants believed developing such mood boards to be easy, because, besides Pinterest, there have been other examples of intelligent moodboards [KTBL*20]. **P21_{ML}** pointed out that pulling images from the latent space is still a very computationally-intensive task, which



(a) Interface of a visualization design system that lets users select different marks and channels to generate pretty visuals, and could also be used for defining smart, generative brushes.



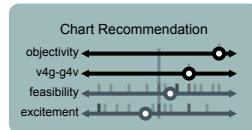
(b) A generative moodboard where an artist ideates visual moods for a garden, selecting between drawings and pictures of plants, and different color palettes.



(c) A set of instructions for how to wash your hands. Two of the biggest difficulties for text-to-image models (at the time of writing) are shown: drawing hands and rendering meaningful text.

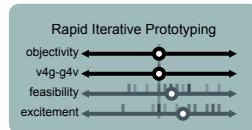
Figure 8: An illustration of possible manifestations of **Rapid Iterative Prototyping** with a generative model. We simulated a simple design sprint, starting with a vague idea of what we wanted, and collaboratively refined our prompts to obtain the above proofs of concept, each being a realistic visualization-related use case. The model’s strengths (aesthetics) and weaknesses (adhering to data) can be observed. While current technology is not able to elaborate on complex queries, these samples still usefully exemplify design thinking. They can serve as positive (“I like this”, “this works”, “I want this element”) or negative (“I don’t like this”, “this is not what we want”, “we should remove this”) anchors.

limits this application. Fig. 8b shows what a generative mood board might look like for a particular design sprint.



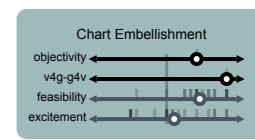
Generative models could be used to suggest types of charts that have fitting qualities to represent given data. Our discussion used an example of Tableau’s **chart recommendation system**.

which may have caused participants to think of this application as very feasible, as it already exists in some forms in commercial software. While participants were not very excited about generatively driven chart recommendation systems, however, some participants suggested adjacent notions. These included a variety of ideas, such as: “Exploring the parameter space for design solutions to a task” (**P5_{vis}**), ““Semantic” chart recommendation” (**P5_{vis}**), “Figuring out how to visualize multidimensional data that is hard to represent” (**P8_{vis}**), and “Search for novel visualization types” (**P8_{vis}**). Each of which seems like they have ample potential application throughout visualization.



Rapid iterative prototyping is the process of rapidly changing parameters while designing or developing a prototype to test different ideas and quickly iterate over them. Most participants felt that it could be used both for supporting VIS and supporting the development of generative models, and both for exploring a design space (thus a subjective process) and more objective, pragmatic purposes. Participants were excited about being able to use generative models for rapid prototyping, and they believed it to be feasible. **P5_{vis}** suggested the IGMs could be used to “generate variations of a design”, rather than starting from scratch each time. A similar idea was expressed by **P19_{vis}** and **P10_{art}**, in terms of “exploring the parameter space for design solutions to a task” and “exploring potentialities of media”. **P1_{vis}** and **P21_{ML}** made a connection between generative

models and NLP-based interfaces for information visualization. Similarly, **P14_{vis}** thought that it might be used as a “search tool”, which is similar to how users of Copilot perceive that tool [SGN*22]. **P12_{vis}** noted that “prototyping entails different things and processes depending on the field”, and was not confident that interaction or complex engineering solutions could be prototyped this way. We sketch this idea, and some of the highlighted opportunities here, in Fig. 8



The concept of **chart embellishment** was presented to the participants as “generating parts of a visualization that don’t really require data”—which participants (**P8_{vis}**) also referred to as “chart junk”. We did not limit the concept to exclusively chart junk. For instance, the use of glyphs in visualization [SM14, YSD*22] or annotation, sometimes do not require data. Indeed, it is easy to see how IGMs could be used to generate iconic or symbolic encodings [Woo22], such as Chernoff faces [Che73], and support anthropographics [MJAD20]. At least two interviewees (**P12_{vis}** and **P17_{vis}**) thought that this was the most probable “inevitable corporate application” to come soon.



Design styles being ripped off was seen as inevitable, but opinions about how concerning it would be, were split. Some participants (**P1_{vis}**, **P8_{vis}**, **P15_{vis}**) declared this wasn’t a worry for them, and saw the issue as more of an advantage for them, while others cared deeply (**P4_{vis}**, **P6_{vis}**). The inevitability of it was often accompanied by a discussion on how this is already happening in the context of art and illustration. **P17_{vis}** commented that models are already doing this with illustrators and artists, and we should expect this for visualization too.

4.4. Interacting

The last pipeline stage comprises user interactions with data visualizations, or more generally, how they might *use* the created visualization. Such interactions are iterated and interleaved amongst other pipeline stages: seeing one chart may suggest constructing another, or may necessitate a return to the transformation stage to better clean the data.

An illustrative application of this potential is the **Personalization** (Fig. 7 right) of visual design to a user’s tastes, tasks, and abilities. For instance, a model might act as an adaptor and automatically transform rendered visualization into more accessible versions that are tailored to a user’s particular needs, such as color blindness or low visualization literacy (Fig. 8c). P10_{art} thought that being able to express your own “*personal aesthetics*”, such as by “(*training on your own data*)” would be useful. Graphics also might be personalized to how a user prefers to consume information. For instance, a manager might only like to see big number displays, while an analyst might like to see richer graphical forms. However, this also runs the risk that seeing only such display might accidentally cause hallucinations based on properties of the chart form [MKC20]. Such personalization also might be dynamic and make “*changes based on tracking user behavior and attention*” (P5_{vis}). While such tools could help to obtain new perspectives, there is a risk that users could lose track of the overview or sense of agency over [Hee19] the presented information.

Despite the value of these interactions, participants were concerned about both malicious and involuntary misleading information generated through the models. They cited the ability of these models to generate believable misinformation (P1_{vis}, P19_{vis}, P21_{ML}) and mentioned deep fakes (P7_{vis}, P13_{art}) and fake news (P10_{art}). P17_{vis} warned that this could also be involuntary, mentioning the possibility of “*irresponsible use of a black box*”, or just for the users of these models to be misled by their own results. P14_{vis} reiterated the dangers of using a black box without fully understanding it, saying that “*everything that is robust is obvious and everything that is obvious is robust*”, referencing Da’s [Da19] argument that computational literary studies are merely counting exercises that only nominally interact with the text—a failure mode to consider in future systems design.

Finally, a variety of additional interactions and applications were suggested for these models. For instance, P20_{art} suggested that a “*content aware brush, painting enhancement*” interaction might be useful to get alternatives for a particular part of a chart. These models could be used to generate guidance, such as in the form of explanations. These might take the form of automatic summaries for charts (such as in the manner of Kanthara et al. [KLL*22]), or parts of charts, such as in a Tableau’s Explain data feature [Tab22]—the advantage of the generative model is that results could be iteratively tuned and adjusted to taste. Such models may be able to synthesize animations or other interactions for previously static charts based on users demonstrating how they wish to interact with a chart or how they wish it to behave. Like other places in the pipeline where user preference can drive usage, this can lead to expressions of bias or what might be called an *AI-filter bubble*, wherein the model reflects and amplifies the user’s biases in a negative feedback loop. Further, high levels of intrusive automatic guidance could be distracting during the interpretation process. Similarly, if that guidance is delivered in a way that does not reflect the user’s agency or previous decisions then it runs the risk of being perceived as impolite [Whi05].

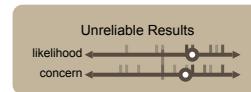
5. Discussion

Despite some skepticism and critique clouding the introduction of generative models to new domains generally, participants in our interviews expressed tempered optimism and cautious excitement about the possibilities for applying these technologies to visualization. Participants were asked to think beyond current technical constraints, and even the most conservative visualization experts conceded the usefulness and inevitability of these technologies, for aesthetic, artistic, and disseminative purposes.

A common thread throughout our discussions was a sense of excitement: the huge technological leap to text-to-image models makes a compelling case for the potential of visualization in the age of generative models. We echo this sentiment and look forward to the dawning age of rich accessible illustrative visualization adapted to taste and to the task. Yet not every aspect of these models yields a delicious conclusion, which we discuss by considering cross-cutting concerns (Sec. 5.1) and challenges (Sec. 5.2).

5.1. Concerns

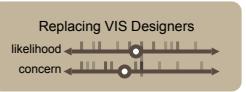
Our participants’ enthusiasm for these technologies is matched by their concerns. Participants mostly expressed that they were concerned about these technologies (Fig. 9c), but that they were, perhaps unfortunately, inevitable. Only P18_{vis}, a visualization expert who is “*much more excited than worried about these technologies*”, considering most of the risks discussed “*not worth a thought*”. Participants typically had strongly-held concerns, but there was not a general consensus of what was most worrying. Some viewed generative models as a form of Trojan horse, allowing the issues to creep in—such as amplification of bias, leaking of private data, deep fakes, and so on.



Unreliable Results. The unreliability of results was one of the aspects of greatest concern. Most worries stemmed from the untraceability or unexplainability of the sources and suggestions. The stochastic nature of generative models and lack of semantic grounding makes it hard (or some might argue, impossible [MGSS21]) to guarantee compliance to queries, which might produce undetectable or unverifiable wrong outputs. Although this is more serious for certain applications where precision is critical, it is also difficult to demonstrate that an image is original, which may incur possible copyright or ownership issues. To wit, P6_{vis} claimed that the risk of copyright violation would prevent them from generating even minor sections of a visualization. While some of these issues may have a technical solution (such as is promised in Amazon’s yet-to-be-released Code Whisperer [Ama22]), navigating the ethical elements beyond the legal ones remains a thorny task.

Replacing Visualization Designers.

Despite our own assumptions to the contrary, participants were generally unconcerned about the premise that generative models would replace designers of visualization. Similarly, participants were not worried about designs being stolen, which we found surprising given the comparable concerns in artistic [Vin22] and coding communities [Roo22]. P6_{vis} also observed that they do not believe they will be replaced per se, but for the expectations to change and for goalposts to be moved.



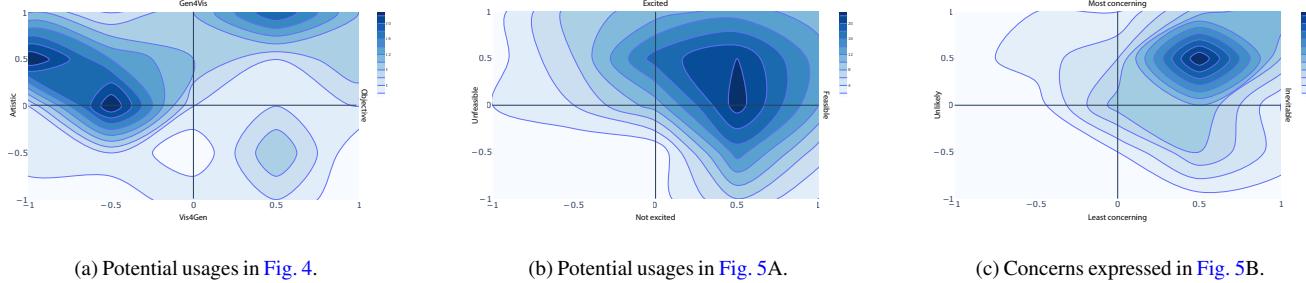


Figure 9: Aggregated participant responses. In (a) we see high interest in artistic applications of generativity to visualization creation. In (b) we see that the majority of concepts were considered feasible, and participants were mostly excited about them. In (c), similarly, the top-right quadrant is dominant, indicating that the risks discussed were generally seen as concerning, and likely to happen.

No participants questioned whether “can computers be creative?” Our participants may not view visualizations as artistic artifacts causing such issues to be viewed as irrelevant to our discussion. Alternatively, the lack of a consensus on what just defines creativity [FMBD18] may have made matters of AI creativity too murky to consider. Two practicing artists, **P13_{art}** and **P20_{art}**, agreed that mastering generative tools will become an essential skill in their field, but espoused differing levels of optimism. While **P20_{art}** was enthusiastic about achieving an edge in his trade, **P13_{art}** is “worried about the raising entry level for newcomers in an already cruel business”. **P20_{art}** observed that tasks like illustrating land cards in *Magic: The Gathering* would likely be among the first jobs to be consumed by these technologies, as generative models can already create astonishing landscape paintings faster and more easily than experienced illustrators. **P9_{art}**, **P10_{art}**, and **P16_{ML}** (who are also involved in art, but are not practicing artists) offered curiosity and excitement about this new media, as well as ethical concerns—such as the use of publicly available art as training data without consent. The public perception of models that generate visualizations remains to be seen, but based on these initial reactions, we suggest that it may be less severe than the reception to generative art [Roo22]. Yet, the negatively perceived practices that led to the current generation of text-to-image models may have ‘poisoned the well’ for some groups. For instance, the visualization community on the vis.social Mastodon instance has specifically disallowed AI-generated images that use text-to-image models such as MidJourney and Stable Diffusion, due to questions of provenance [Hen22].

5.2. Challenges

We next describe several high-level challenges that we identified through our analysis.

Data-Constrained Generation. The inclusion of constraints is a prominent technical gap in diffusion models. For instance, most popular models (such as Midjourney [Mot22]) cannot currently deal with prompts involving quantities (such as “five dogs and two cats”), have trouble with text, and objects such as hands (Fig. 8c). Thus the substantially more constrained task of creating data visualizations is far off at the moment. Constraining the output of generative processes to data is not a new concept [Soc17, KM20], but it can be tricky to apply to the powered-by-noise diffusion process of which most text-to-image models are currently based. Yet, even once such techniques

are possible, it might be long a time until such images can be reliably trusted—at least according to the views of our participants. Some skeptical participants emphasized this limitation, arguing that generative models have no place in objective visualization (Fig. 4, top-right quadrant). This is aligned with how, despite the promise of recommendation engines to create elaborate images, analysts have primarily preferred to focus on simple charts with clear takeaways [BLF*22]. Still, we observed ample optimism in other interviewees, which suggests that when such a technology is manifest, it may be well received.

Beyond exploring how to better navigate and convey trust, there is ample room for improvement in generative models, such as from a performance perspective. For instance, the current resolution for generated images is still quite low, as higher resolutions require massive amounts of GPU memory. Simply by increasing the maximum resolution, we allow for higher frequencies to be rendered, which are essential for text and data visualization. We suggest that such issues indicate the value of investigating models that create a vector, rather than raster, images, as they may be less constrained by such technical limitations.

Inheritance of AI Worries. By incorporating generative models into the visualization pipeline, we inherit many problems faced by the AI and ML communities. For instance, as we observed throughout Sec. 4, the issue of bias is a pressing concern. **P16_{ML}** observed that “bias is a human problem, not a machine problem”, so our approach to addressing these issues can not rely exclusively on technical solutions. Further, the current trend towards increasingly large models has meant that only companies with large resources are able to develop these models from scratch. Several participants (e.g. **P1_{ML}**, **P12_{vis}**, and **P14_{vis}**) discussed the issue of centralization and democratization of these technologies. They observed that the current situation places a small number of firms in control of the data, allows them to exercise authority on what counts as bias, and may lead to censorship.

Harmful Rationalization. The easier it is to create personalized, easily digestible, compelling, memorable, and beautiful visualizations, the harder it will be to guard ourselves against their influence. Misinformation becomes a bigger risk even without ill-intent [MHK21], as a convincing visualization might lead a person to become overconfident over the knowledge gained from it [SEA22]. Their domineering form can make readers less questioning about conclusions [MCC20]. For instance, Galactica [TKC*22], a generative large language model meant to facilitate scientific inquiry, often gives incorrect or biased answers to questions, but presented in an authoritative manner [Hea22].

Of course, when there exists actual ill-intent in the use of these models the possible visualizations that can be generated from them become even more dangerous. Many participants (**P1_{vis}**, **P7_{vis}**, **P9_{art}**, **P10_{art}**, **P13_{art}**, **P21_{ML}**) worried about deep fakes and fake news, and that visualization powered by generative models has the potential to be a “super fake news machine” (**P13_{art}**). Yet, one does not even need to use fake information, only to have an automated tool capable of choosing convenient half-truths to support a position, querying the needed data and then creating beautiful and convincing visualizations that can be easily spread through unregulated social media channels. Further, most visual deception carried out on social media does not even require graphically tricking the user, rather data mirages like cherry-picking are more than sufficient [**LPLK23**].

5.3. Limitations

As with any study, ours has limitations. We focused on the ideation process in our interviews to elicit ideas and opinions from our participants, which is an inherently speculative task. This conjectural nature likely influenced participants’ responses, which would have been different if we had provided interactive demos. For instance, in the evaluation of concern versus likelihood, some participants reported not being concerned about something *precisely* because they thought it was unlikely.

Our study participants may not be representative of a wider or different population. As we focused on gathering the perspective from visualization researchers and practitioners, the other groups (artists, art historians, etc) were under-represented in our sample. Some participants lacked familiarity with generative models and so had difficulty in discussing the default concepts. Similarly, our focus on domain experts may have biased our results. Future work should explore how these results extend to the public more generally.

The development of tools in this space continues to happen at a breakneck speed. During the course of our study, a number of notable models were released—such as MidJourney V4 [**Mot22**], Stable Diffusion 2.0 [**RBL*22**], Galactica [**TKC*22**], and ChatGPT [**Ope23**]. While the developments in these models are important, their behavior is not fundamentally different from the prior models, and so we believe that our findings will continue to apply as these models are further iterated upon. Our interviews only focused on IGMs, rather than the full scope of generative models. We chose to focus on image generation because of their cultural prominence at the time of our study, because of its functional similarity to the human act of producing data visualizations, and for its visually engaging artifacts, which we hoped would provoke more inspired participant responses.

5.4. Conclusion

In this work, we sought to chart the challenges and opportunities latent in the oncoming interweaving of data visualization and generative models, particularly those focused on text-to-image generation. We did so through a semi-structured interview study with experts from a variety of domains and identified a number of areas of risk (such as the potential for harm that models trained on countless images hold) as well as areas of potentially ample reward (such as enabling easily accessible illustrative visualizations).

Shortly after we conducted our study, chatGPT was released to

wide renown. This central achievement of this system is not a technical innovation (although the associated advances are non-trivial), rather it is one of usability—leading to record-breaking numbers of users engaging with an LLM for a wide variety of tasks [**Reu23**]. This popularity surfaced and extended discussions about some of the topics explored here, such as the consequences of generative AI systems for science more generally [**SWVN23**]. There is a new promising field of research in finding emergent abilities LLMs [**WTB*23**, **KCK*23**], which soon might be transferable to the visual domain [**RKH*21**], and using them as universal interfaces [**SST*23**].



Yet, we are never really prepared for the effects new technologies can have on ourselves, and on society. McLuhan observed that “*We look at the present through a rearview mirror. We march backward into the future*” [**FM67**]. However, we should not be afraid, because “*there is absolutely no inevitability as long as there is a willingness to contemplate what is happening*”. Through this work, we strove to contemplate what is happening, gathering the participants’ experiences and opinions—their collective rearview mirrors—and then projecting them into the future. In doing so, we have tried to peer out from the walled garden of visualization [**Woo22**] and chart this land beyond the trees: its risks, challenges, and opportunities.

Acknowledgments

We thank our anonymous reviewers for their helpful comments, as well as our study participants, without whom this work would not have been possible. This work has been partially supported by the European Commission under the project Humane-AI-Net (grant agreement 952026) and the Austrian Science Fund (FWF, grant P35767).

References

- [Ama22] AMAZON: Amazon codewhisperer features. <https://aws.amazon.com/codewhisperer/features/>, 2022. Accessed 8/16/22. 8
- [BEM*21] BENZAGHTA M. A., ELWALDA A., MOUSA M. M., ERKAN I., RAHMAN M.: Swot analysis applications: An integrative literature review. *Journal of Global Business Insights* 6, 1 (2021), 55–73. 4
- [BGMMS21] BENDER E. M., GEBRU T., McMILLAN-MAJOR A., SHMITCHELL S.: On the dangers of stochastic parrots: Can language models be too big? In *Proc. of the ACM Conf. on Fairness, Accountability, and Transparency* (2021), pp. 610–623. 3
- [BK20] BENDER E. M., KOLLER A.: Climbing towards nlu: On meaning, form, and understanding in the age of data. In *Proc. of the 58th annual meeting of the association for computational linguistics* (2020), pp. 5185–5198. 2
- [BKO*17] BYLINSKII Z., KIM N. W., O’DONOVAN P., ALSHEIKH S., MADAN S., PFISTER H., DURAND F., RUSSELL B. C., HERTZMANN A.: Learning visual importance for graphic designs and data visualizations. *CoRR abs/1708.02660* (2017). URL: <http://arxiv.org/abs/1708.02660>. 3
- [BLF*22] BAO C. S., LI S., FLORES S. G., CORRELL M., BATTLE L.: Recommendations for visualization recommendations: Exploring preferences and priorities in public health. In *CHI Conf. on Human Factors in Computing Systems* (2022), pp. 1–17. 9

- [CC10] COLDEFY M., CURTIS S. E.: The geography of institutional psychiatric care in france 1800–2000: Historical analysis of the spatial diffusion of specialised facilities for institutional care of mental illness. *Social Science & Medicine* 71, 12 (2010), 2117–2129. 2
- [Che73] CHERNOFF H.: The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association* 68, 342 (1973), 361–368. URL: <http://www.jstor.org/stable/2284077>. 7
- [Chi00] CHI E. H.-H.: A taxonomy of visualization techniques using the data state reference model. In *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings* (2000), IEEE, pp. 69–75. 4
- [CM20] COELHO D., MUELLER K.: Infomages: Embedding data into thematic images. In *Computer Graphics Forum* (2020), vol. 39, Wiley Online Library, pp. 593–606. 6
- [Com22] COMPTON K.: Tweet, October 2022. URL: <https://web.archive.org/web/20221124001133/https://twitter.com/GalaxyKate/status/1583907942834716672>. 1
- [Cop04] COPE D.: *Virtual music: computer synthesis of musical style*. MIT press, 2004. 1
- [Da19] DA N. Z.: The computational case against computational literary studies. *Critical inquiry* 45, 3 (2019), 601–639. 8
- [DBSSD23] DI BARTOLOMEO S., SEVERI G., SCHETINGER V., DUNNE C.: Ask and you shall receive (a graph drawing): Testing chatgpt’s potential to apply graph layout algorithms, Mar 2023. URL: <osf.io/wsvx7>. 2
- [DD18] DIBIA V., DEMIRALP C.: Data2vis: Automatic generation of data visualizations using sequence to sequence recurrent neural networks, 2018. URL: <https://arxiv.org/abs/1804.03126>, doi: 10.48550/ARXIV.1804.03126. 3
- [EAKM*22] EL-ASSADY M., KEHLBECK R., METZ Y., SCHLEGEL U., SEVASTIANOVA R., SPERRLE F., SPINNER T.: Semantic Color Mapping: A Pipeline for Assigning Meaningful Colors to Text. *4th IEEE Workshop on Visualization Guidelines in Research, Design, and Education* (2022), 7. 6
- [FM67] FIORE Q., MCLUHAN M.: *The medium is the message*. New York: Random House, 1967. 10
- [FMBD18] FRICH J., MOSE BISKJAER M., DALSGAARD P.: Twenty years of creativity research in human-computer interaction: Current state and future directions. In *Proc. of the 2018 Designing Interactive Systems Conf.* (2018), pp. 1235–1257. 9
- [For22] FORREST J.: I Asked an Artificial Intelligence to Draw a Chart. *Nightingale* (June 2022). URL: <https://nightingaledvs.com/i-asked-an-artificial-intelligence-to-draw-a-chart/>. 2
- [FPSS96] FAYYAD U., PIATETSKY-SHAPIRO G., SMYTH P.: The kdd process for extracting useful knowledge from volumes of data. *Communications of the ACM* 39, 11 (1996), 27–34. 4
- [Git22] GITHUB: Github copilot. <https://github.com/features/copilot>, 2022. Accessed 11/28/22. 1, 6
- [Goo23] GOOGLE: An important next step on our ai journey. <https://blog.google/technology/ai/bard-google-ai-search-updates/>, 2023. Accessed 3/8/23. 5
- [GSS*22] GABA A., SETLUR V., SRINIVASAN A., HOFFSWELL J., XIONG C.: Comparison Conundrum and the Chamber of Visualizations: An Exploration of How Language Influences Visual Design. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. URL: <https://ieeexplore.ieee.org/document/9903602/>, doi:10.1109/TVC.2022.3209456. 5
- [Hea22] HEAVEN W. D.: Why Meta’s latest large language model survived only three days online. *MIT Technology Review* (November 2022). URL: <https://www.technologyreview.com/2022/11/18/1063487/meta-large-language-model-ai-only-survive-d-three-days-gpt-3-science/>. 9
- [Hee19] HEER J.: Agency plus automation: Designing artificial intelligence into interactive systems. *Proc. of the National Academy of Sciences* 116, 6 (2019), 1844–1850. 8
- [Hei22] HEIKKILÄ M.: This artist is dominating ai-generated art. and he’s not happy about it., Sep 2022. URL: <https://www.technologyview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/>. 3
- [Hen22] HENRY K.: vis.social server rules. <https://vis.social/about>, 2022. Accessed 12/1/22. 9
- [HHS*22] HOQUE M. N., HE W., SHEKAR A. K., GOU L., REN L.: Visual Concept Programming: A Visual Analytics Approach to Injecting Human Intelligence At Scale. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–10. URL: <https://ieeexplore.ieee.org/document/9904017/>, doi:10.1109/TVC.2022.3209466. 5
- [HIDI22] HE T., ISENBERG P., DACHSELT R., ISENBERG T.: BeauVis: A Validated Scale for Measuring the Aesthetic Pleasure of Visual Representations. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. URL: <https://ieeexplore.ieee.org/document/9903341/>, doi:10.1109/TVC.2022.3209390. 6
- [HPRC20] HOHMAN F., PARK H., ROBINSON C., CHAU D. H.: Summit: Scaling deep learning interpretability by visualizing activation and attribution summarizations. *IEEE Transactions on Visualization and Computer Graphics (TVC)* (2020). URL: <https://fredhohman.com/summit/>, doi:10.1109/TVC.2019.2934659. 3
- [HST17] HYNES N., SCULLEY D., TERRY M.: The data linter: Lightweight, automated sanity checking for ml data sets. In *NIPS MLSys Workshop* (2017), vol. 1. 6
- [HXHT22] HUANG J., XI Y., HU J., TAO J.: FlowNL: Asking the Flow Data in Natural Languages. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. URL: <https://ieeexplore.ieee.org/document/9910021/>, doi:10.1109/TVC.2022.3209453. 5
- [HYC*22] HU R., YE Z., CHEN B., KAICK O. V., HUANG H.: Self-Supervised Color-Concept Association via Image Colorization. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–10. URL: <https://ieeexplore.ieee.org/document/9904484/>, doi:10.1109/TVC.2022.3209481. 6
- [KCK*23] KOCON J., CICHECKI I., KASZYCA O., KOCHANEK M., SZYDŁO D., BARAN J., BIELANIEWICZ J., GRUZA M., JANZ A., KANCERZ K., KOCON A., KOPTYRA B., MIELESZCZENKO-KOWSZEWCZ W., MILKOWSKI P., OLEKSY M., PIASECKI M., RADLIŃSKI L., WOJTASIK K., WOZNIAK S., KAZIENKO P.: ChatGPT: Jack of all trades, master of none, Feb. 2023. arXiv:2302.10724 [cs]. URL: <https://arxiv.org/abs/2302.10724>. 10
- [KLL*22] KANTHARA S., LEONG R. T. K., LIN X., MASRY A., THAKKAR M., HOQUE E., JOTY S.: Chart-to-text: A large-scale benchmark for chart summarization. *arXiv preprint arXiv:2203.06486* (2022). 8
- [KM20] KWON O., MA K.: A deep generative model for graph layout. *IEEE Trans. Vis. Comput. Graph.* 26, 1 (2020), 665–675. doi:10.1109/TVC.2019.2934396. 9
- [KTBL*20] KOCH J., TAFFIN N., BEAUDOUIN-LAFON M., LAINE M., LUCERO A., MACKAY W. E.: Imagesense: An intelligent collaborative ideation tool to support diverse human-computer partnerships. *Proc. of the ACM on human-computer interaction* 4, CSCW1 (2020), 1–27. 6
- [KW19] KRISHNAN S., WU E.: Alphaclean: Automatic generation of data cleaning pipelines. *arXiv preprint arXiv:1904.11827* (2019). 6
- [LPLK23] LISNIC M., POLYCHRONIS C., LEX A., KOGAN M.: Misleading beyond visual tricks: How people actually lie with charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023), ACM. To Appear. 10
- [LQC22] LIU V., QIAO H., CHILTON L.: Opal: Multimodal image generation for news illustration. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2022), UIST ’22, Association for Computing Machinery. URL: <https://doi.org/10.1145/3526113.3545621>, doi:10.1145/3526113.3545621. 2

- [Ma23] MA K.-L.: Pushing visualization research frontiers: Essential topics not addressed by machine learning. *IEEE Computer Graphics and Applications* 43, 1 (2023), 97–102. 2
- [MC21] MCNUTT A. M., CHUGH R.: Integrated visualization editing via parameterized declarative templates. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–14. 2
- [MCC20] MCNUTT A., CRISAN A., CORRELL M.: Divining insights: Visual analytics through cartomancy. In *Extended Abstracts of the 2020 CHI Conf. on Human Factors in Computing Systems* (2020), pp. 1–16. 9
- [MGSS21] MERRILL W., GOLDBERG Y., SCHWARTZ R., SMITH N. A.: Provable limitations of acquiring meaning from ungrounded form: What will future language models understand? *Transactions of the Association for Computational Linguistics* 9 (2021), 1047–1060. URL: http://dx.doi.org/10.1162/tacl_a_00412, doi:10.1162/tacl_a_00412. 8
- [MHK21] MCNUTT A. M., HUANG L., KOENIG K.: Visualization for villainy. *alt.vis* (2021). 6, 9
- [MJAD20] MORAIS L., JANSEN Y., ANDRADE N., DRAGICEVIC P.: Showing data about people: A design space of anthropographics. *IEEE Transactions on Visualization and Computer Graphics* (2020). 7
- [MKC20] MCNUTT A., KINDLMANN G., CORRELL M.: Surfacing visualization mirages. In *Proc. of the 2020 CHI Conf. on human factors in computing systems* (2020), pp. 1–16. 4, 5, 8
- [MO14] MIRZA M., OSINDERO S.: Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014). 2
- [Mot22] MOTION L.: Midjourney, 2022. Accessed 17/11/22. URL: <http://www.midjourney.com/>. 1, 2, 9, 10
- [MQB19] MING Y., QU H., BERTINI E.: Rulematrix: Visualizing and understanding classifiers with rules. *IEEE Trans. Vis. Comput. Graph.* 25, 1 (2019), 342–352. URL: <https://doi.org/10.1109/TVCG.2018.2864812>, doi:10.1109/TVCG.2018.2864812. 3
- [MS23] MADDIGAN P., SUSNIAK T.: Chat2vis: Generating data visualisations via natural language using chatgpt, codex and gpt-3 large language models. *arXiv preprint arXiv:2302.02094* (2023). 2
- [MV22] MAYAHI S., VIDRIH M.: The impact of generative ai on the future of visual content marketing. *arXiv preprint arXiv:2211.12660* (2022). 2
- [MWDD23] MCNUTT A. M., WANG C., DELINE R. A., DRUCKER S. M.: On the design of ai-powered code assistants for notebooks. *SIGCHI Conference on Human Factors in Computing Systems* (2023). To Appear. 2
- [OGW19] OTTLEY A., GARNETT R., WAN R.: Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis. *Computer Graphics Forum* 38, 3 (2019), 41–52. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13670>, arXiv: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13670>, doi:<https://doi.org/10.1111/cgf.13670>. 3
- [oLS23] OF LABOR STATISTICS B.: Cpi for all urban consumers (cpi-u), 2023. Accessed 3/8/23. URL: https://data.bls.gov/timeseries/CUUR0000SA0L1E?output_view=pct_12mths. 2
- [Ope22a] OPENAI: DALL-E 2, 2022. Accessed 9/5/22. URL: <https://openai.com/dall-e-2/>. 1, 5
- [Ope22b] OPENAI: Gpt-3. <https://openai.com/api/>, 2022. Accessed 6/30/22. 1
- [Ope22c] OPENAI: Reducing bias and improving safety in dall-e 2, 2022. Accessed 9/5/22. URL: <https://openai.com/blog/reducing-bias-and-improving-safety-in-dall-e-2/>. 1, 3
- [Ope23] OPENAI: Chatgpt. <https://chat.openai.com/>, 2023. Accessed 3/8/23. 2, 5, 10
- [PH17] POCO J., HEER J.: Reverse-engineering visualizations: Recovering visual encodings from chart images. In *Computer graphics forum* (2017), vol. 36, Wiley Online Library, pp. 353–363. 5
- [Plu22] PLUNKETT L.: AI Creating 'Art' Is An Ethical And Copyright Nightmare. *Kotaku* (August 2022). URL: <https://kotaku.com/a-i-art-dall-e-midjourney-stable-diffusion-copypright-1849388060.1>
- [PMdS23] PRIETO S. A., MENGISTE E. T., DE SOTO B. G.: Investigating the use of chatgpt for the scheduling of construction projects. *arXiv preprint arXiv:2302.02085* (2023). 2
- [RBL*22] ROMBACH R., BLATTMANN A., LORENZ D., ESSER P., OMMER B.: High-resolution image synthesis with latent diffusion models. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* (June 2022), pp. 10684–10695. 1, 2, 5, 10
- [RDN*22] RAMESH A., DHARIWAL P., NICHOL A., CHU C., CHEN M.: Hierarchical text-conditional image generation with clip latents, 2022. URL: <https://arxiv.org/abs/2204.06125>, doi:10.4850/ARXIV.2204.06125. 2
- [Reu23] REUTERS: Chatgpt sets record for fastest-growing user base - analyst note. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>, 2023. Accessed 3/12/23. 10
- [RKH*21] RADFORD A., KIM J. W., HALLACY C., RAMESH A., GOH G., AGARWAL S., SAstry G., ASKELL A., MISHKIN P., CLARK J., ET AL.: Learning transferable visual models from natural language supervision. In *International Conf. on Machine Learning* (2021), PMLR, pp. 8748–8763. 10
- [Roo22] ROOSE K.: An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy. *New York Times* (September 2022). URL: <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>. 1, 8, 9
- [RS19] ROSENTHAL K., STRECKER S.: Toward a taxonomy of modeling difficulties: A multi-modal study on individual modeling processes. In *Proc. of the 40th International Conf. on Information Systems, ICIS 2019, Munich, Germany, December 15–18, 2019* (2019), Krcmar H., Fedorowicz J., Boh W. F., Leimeister J. M., Wattal S., (Eds.), Association for Information Systems. URL: https://aisel.aisnet.org/icis2019/learning_environ/learning_environ/12.5
- [SCS*22] SAHARIA C., CHAN W., SAXENA S., LI L., WHANG J., DENTON E., GHASEMPIOUR S. K. S., AYAN B. K., MAHDAVI S. S., LOPES R. G., SALIMANS T., HO J., FLEET D. J., NOROUZI M.: Photorealistic text-to-image diffusion models with deep language understanding, 2022. URL: <https://arxiv.org/abs/2205.11487>, doi:10.48550/ARXIV.2205.11487. 2
- [SEA22] SEVASTIANOVA R., EL-ASSADY M.: Beware the rationalization trap! when language model explainability diverges from our mental models of language. *arXiv preprint arXiv:2207.06897* (2022). 9
- [SFPM*22] SCHETINGER V., FILIPOV V., PÉREZ-MESSINA I., SMITH E., DE OLIVEIRA R. O.: I learn to diffuse, or data alchemy 101: a mnemonic manifesto. *alt.vis* (2022). 2
- [SGN*22] SARKAR A., GORDON A. D., NEGREANU C., POELITZ C., RAGAVAN S. S., ZORN B.: What is it like to program with artificial intelligence? *Psychology of Programming Interest Group (PPIG)* (2022). 7
- [SLC*22] SHI Y., LIU P., CHEN S., SUN M., CAO N.: Supporting Expressive and Faithful Pictorial Visualization Design With Visual Style Transfer. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–11. URL: <https://ieeexplore.ieee.org/document/9903511/>, doi:10.1109/TVCG.2022.3209486. 3
- [SLJL10] SUN Y., LEIGH J., JOHNSON A., LEE S.: Articulate: A semi-automated model for translating natural language queries into meaningful visualizations. In *Smart Graphics* (Berlin, Heidelberg, 2010), Taylor R., Boulanger P., Krüger A., Olivier P., (Eds.), Springer Berlin Heidelberg, pp. 184–195. 2
- [SM14] SETLUR V., MACKINLAY J. D.: Automatic generation of semantic icon encodings for visualizations. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2014), pp. 541–550. 7
- [Soc17] SOCOLOFSKY E.: Data as a creative constraint. OpenVisConf, 2017. URL: <https://www.youtube.com/watch?v=fIy7rkYyphI>. 9

- [Spa22] SPAWNING: Have i been trained?, 2022. Accessed 12/1/22. URL: <https://haveibeentrained.com/>.5>
- [SSL*22] SHEN L., SHEN E., LUO Y., YANG X., HU X., ZHANG X., TAI Z., WANG J.: Towards natural language interfaces for data visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–1. doi:[10.1109/TVCG.2022.3148007.2](https://doi.org/10.1109/TVCG.2022.3148007.2)
- [SST*23] SHEN Y., SONG K., TAN X., LI D., LU W., ZHUANG Y.: Hugginggpt: Solving ai tasks with chatgpt and its friends in huggingface, 2023. arXiv:[2303.17580.10](https://arxiv.org/abs/2303.17580.10)
- [STN*16] SMILKOV D., THORAT N., NICHOLSON C., REIF E., VIÉGAS F. B., WATTENBERG M.: Embedding projector: Interactive visualization and interpretation of embeddings. *CoRR abs/1611.05469* (2016). URL: [ht tp://arxiv.org/abs/1611.05469](https://arxiv.org/abs/1611.05469), arXiv:[1611.05469.3](https://arxiv.org/abs/1611.05469.3)
- [SWVN23] STOKEL-WALKER C., VAN NOORDEN R.: What ChatGPT and generative AI mean for science. *Nature* 614, 7947 (Feb 2023), 214–216. URL: <https://doi.org/10.1038/d41586-023-00340-6.10>
- [Tab22] TABLEAU: Get started with explain data, 2022. Accessed 12/1/22. URL: https://help.tableau.com/current/pro/desktop/en-us/explain_data_basics.htm.8
- [Tay23] TAYLOR J.: Chatgpt's alter ego, dan: users jailbreak ai program to get around ethical safeguards. <https://www.theguardian.com/technology/2023/mar/08/chatgpt-alter-ego-dan-users-jailbreak-ai-program-to-get-around-ethical-safeguards>, 2023. Accessed 3/8/23. 6
- [TKC*22] TAYLOR R., KARDAS M., CUCURULL G., SCIALOM T., HARTSHORN A., SARAVIA E., POULTON A., KERKEZ V., STOJNIC R.: Galactica: A large language model for science. *arXiv preprint arXiv:2211.09085* (2022). 9, 10
- [Vin22] VINCENT J.: The scary truth about AI copyright is nobody knows what will happen next. *Verge* (November 2022). URL: <https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data.1.5.8>
- [WCA23] WU J., CHUNG J. J. Y., ADAR E.: viz2viz: Prompt-driven stylized visualization generation using a diffusion model, 2023. arXiv:[2304.01919.2](https://arxiv.org/abs/2304.01919.2)
- [WCWQ22] WANG Q., CHEN Z., WANG Y., QU H.: A survey on ML4VIS: applying machine learning advances to data visualization. *IEEE Trans. Vis. Comput. Graph.* 28, 12 (2022), 5134–5153. URL: <https://doi.org/10.1109/TVCG.2021.3106142>, doi:[10.1109/TVCG.2021.3106142.2.3](https://doi.org/10.1109/TVCG.2021.3106142.2.3)
- [WFC*18] WANG Y., FENG K., CHU X., ZHANG J., FU C.-W., SEDLMAIR M., YU X., CHEN B.: A perception-driven approach to supervised dimensionality reduction for visualization. *IEEE Transactions on Visualization and Computer Graphics* 24, 5 (2018), 1828–1840. doi:[10.1109/TVCG.2017.2701829.3](https://doi.org/10.1109/TVCG.2017.2701829.3)
- [WH22] WANG C., HAN J.: Dl4scivis: A state-of-the-art survey on deep learning for scientific visualization, 2022. URL: <https://arxiv.org/abs/2204.06504>, doi:[10.48550/ARXIV.2204.06504.2](https://doi.org/10.48550/ARXIV.2204.06504.2)
- [Whi05] WHITWORTH B.: Polite computing. *Behaviour & Information Technology* 24, 5 (2005), 353–363. 8
- [WJW*20] WANG Y., JIN Z., WANG Q., CUI W., MA T., QU H.: Deep-drawing: A deep learning approach to graph drawing. *IEEE Trans. Vis. Comput. Graph.* 26, 1 (2020), 676–686. URL: <https://doi.org/10.1109/TVCG.2019.2934798>, doi:[10.1109/TVCG.2019.2934798.3](https://doi.org/10.1109/TVCG.2019.2934798.3)
- [WKD18] WOOD J., KACHKAEV A., DYKES J.: Design exposition with literate visualization. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 759–768. 6
- [Woo22] WOOD J.: Beyond the walled garden: A visual essay in five chapters. *alt.vis* (2022). 2, 5, 7, 10
- [WTB*23] WEI J., TAY Y., BOMMASANI R., RAFFEL C., ZOPH B., BORGEAUD S., YOGATAMA D., BOSMA M., ZHOU D., METZLER D.,
- CHI E. H., HASHIMOTO T., VINYALS O., LIANG P., DEAN J., FEDUS W.: Emergent Abilities of Large Language Models. *Transactions on Machine Learning Research* (Jan. 2023). URL: <https://openreview.net/forum?id=yzkSU5zdwD.10>
- [WTC22] WU T., TERRY M., CAI C. J.: Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts. In *SIGCHI Conference on Human Factors in Computing Systems* (2022), pp. 1–22. 2, 3, 6
- [WWS*22] WU A., WANG Y., SHU X., MORITZ D., CUI W., ZHANG H., ZHANG D., QU H.: AI4VIS: survey on artificial intelligence approaches for data visualization. *IEEE Trans. Vis. Comput. Graph.* 28, 12 (2022), 5049–5070. URL: <https://doi.org/10.1109/TVCG.2021.3099002>, doi:[10.1109/TVCG.2021.3099002.2.4](https://doi.org/10.1109/TVCG.2021.3099002.2.4)
- [XLC20] XIAO Q., LI G., CHEN Q.: Image outpainting: Hallucinating beyond the image. *IEEE Access* 8 (2020), 173576–173583. URL: <https://ojs.aaai.org/index.php/AAAI/article/view/6951.4>
- [YQS20] YANG J., QI Z., SHI Y.: Learning to incorporate structure knowledge for image inpainting. In *AAAI Symposium on Educational Advances in Artificial Intelligence* (2020), AAAI Press, pp. 12605–12612. URL: <https://ojs.aaai.org/index.php/AAAI/article/view/6951.4>
- [YSD*22] YING L., SHU X., DENG D., YANG Y., TANG T., YU L., WU Y.: Metaglyph: Automatic generation of metaphoric glyph-based visualization. *IEEE Transactions on Visualization and Computer Graphics* (2022). 7

Appendix A: Additional Details

In this appendix, we include additional information that we were unable to fit into the main body of the paper. These cover further details about our participants, the full structure of the interview, and the data we obtained through the interviews. Finally, in Fig. 11 we list the prompts used to create the imagery throughout the paper.

Participant backgrounds

As part of our onboarding, participants were asked to self-identify between Practice ↔ Research and Artistic ↔ Technical by shaping a circle over a plane, which we aggregate in Fig. 3. The resulting aggregation shows a variety of specialists and generalists over our covered spectra. This population includes 21 participants: 11 visualization researchers or practitioners (4 Ph.D. students, 4 postdocs, 3 researchers/designers, and 1 professor), the majority with over 5 years of experience and some with over 10 years, 3 artists each with over 8 years experience, 2 machine learning researchers, 2 HCI postdocs, and 2 art historians, one of which is also a philosopher and researches automated image-making. Fig. 3 shows a distribution of the aggregated self-identifications modeled as points. However, as can be observed in the miro interviews (end pages), each person had the freedom to shape a circle around this spectrum as desired, and participants varied widely in their choices of size and positioning.

Interview Format

The interviews were done over a miro board, and each interview was split into three phases: **onboarding** (introduction, self-identification), exploring **opportunities** (stages 1a, 1b), and understanding **risks** (stage 2). The two main forms of interaction were adding post-its with ideas to the board, and then ordering them relatively to the axis to answer questions, which are done while vocalizing thoughts and talking to the interviewer. Adding a post-it or splitting an existing one represents an act of ideation, and establishing relative positions are acts of comparative evaluation. The steps of the interview are organized so that subjects alternate between ideation and evaluation.

During the **onboarding**, the participants were presented with examples of experimental visualizations made with generative technologies by the authors, briefly discussed them, and then were asked to self-identify by shaping a circle over an axis (Fig. 10, top-left). This served to introduce the ideas that were to be discussed and to make participants accustomed to Miro and the interview format.

The **opportunities** and **risks** phases contain an initial population of concepts, which serves two purposes: to help the interviewee orient themselves in the proposed dimensions, allowing for easier ideation, and to force them to evaluate core ideas curated by the authors. Originally, the spaces were empty to avoid priming, but the pilot interviews showed us that eliciting novelty requires a lot of creativity and effort on the part of the interviewee, which could easily be blank. By having a set of relevant concepts in place we not only inspired subjects but also forced them to give us some input in the subsequent evaluation phase by ordering them in case they could not think of new ideas to add. The relevant stages for collecting insights were 1a, 1b, and 2, each with its own plane and conceptual goal.

Stage 1a. starts with participants being shown a populated space of

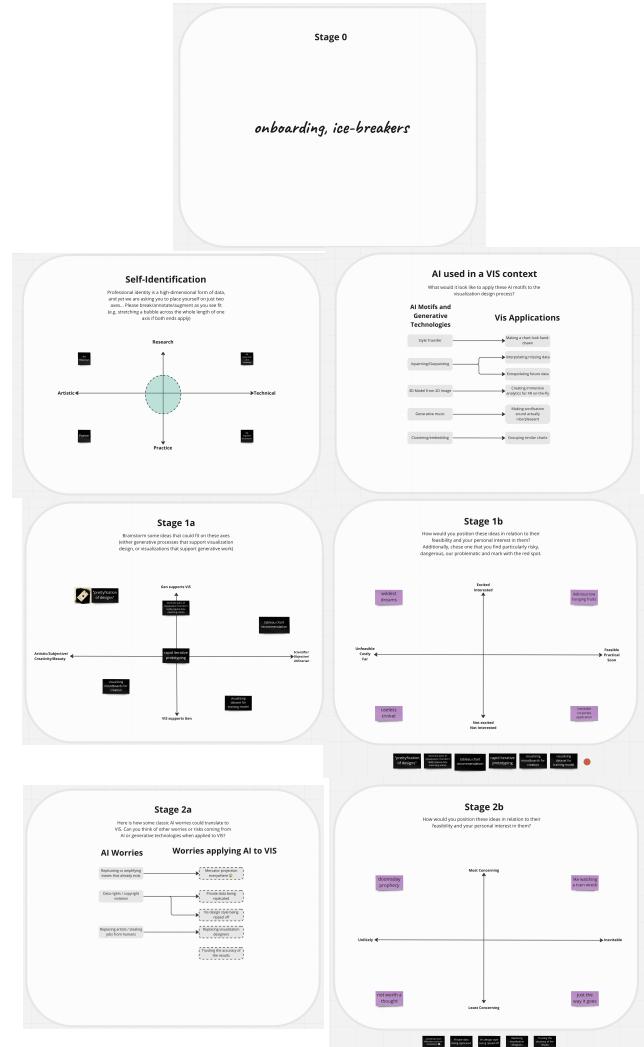


Figure 10: A blank copy of the study instrument. Each participant was walked through each of the cells shown here, and add or sort sticky notes on the axes. The positioning of the sticky notes was then coded according to the axes, assigning each one a value between -1 and 1, yielding Fig. 9

concepts through the dimensions of purpose (vertical) vs. objectivity (horizontal). The purpose is defined between two polarities: generative models supporting visualization at the top, and visualization supporting generative models at the bottom. After an explanation and discussion of these ideas, participants are asked then to add their own, and the totality of ideas at the board will be transferred to the next stage (1b).

Stage 1b. sought to measure the relative excitement (vertical axis) vs the assumed feasibility (horizontal axis) of the ideas by spatially ordering them around. This was meant to gauge the participant's interest and belief that these applications could be useful and realistically integrated in the visualization process, or support it. Each quadrant was given a name based on its meaning to help the understanding of

Figure 11: Throughout the main paper we used a variety of IGM-created graphics, the prompts for which we list here.

Figure	Tool	Prompt
1	Midjourney	Prompts appears in the figure, each use version 4
6a	Midjourney	<i>pastel mood board, data visualization, chart, infographic, interface –no monitor, globe, woman –v 4</i>
6b	Midjourney	<i>mood board data visualization, pastel colors, pictures of plants and mugs, pastel light tones –v 4</i>
6c	Midjourney	<i>pastel, data visualization, infographic, visual explanation of how to wash hands with steps, vertigo comics –v 4</i>
Inline Apple	Midjourney	<i>the forbidden fruit (apple) of generative models, enticing visualization researchers. minimalist design, Klimt, stencil, simplified, modernist, clean, white background</i>

the task. At the end of this stage, participants were also asked to indicate which one of the applications offered the biggest risk to them by overlaying a red dot on top of it. The top-right quadrant is especially important to us, indicating the most probable opportunities: ideas that participants are excited about and appear feasible, and this is why we named it "**delicious low-hanging fruits**".

Stage 2. was intended to investigate the participants' fears and concerns about AI motifs applied to visualization, which would imply risk applications based on generative models. After being presented with a list of common "AI Worries", and some predefined risks for visualization, and then discussing them, the participant would be asked to add more ideas to the list. Then, once again, all ideas on the board would be moved to a plane, at this stage defined by concern (vertical axis) vs. likelihood (horizontal axis), where they had to be ordered comparatively. In this stage the quadrants are named relatively to the type of danger they represent: a concerning, but the unlikely scenario is a "**doomsday prophecy**", but concerning and inevitable risks could be "like watching a train wreck", and tackling them should be treated as serious challenges.

The full artifacts of our interviews are included in the next pages.

