StoryEnsemble: Enabling Dynamic Exploration & Iteration in the Design Process with AI and Forward-Backward Propagation

Sangho Suh
University of Toronto
Toronto, Canada
sangho.suh@utoronto.ca

Michael Lai
University of Toronto
Toronto, Canada
michaelkp.lai@utoronto.ca

Kevin Pu
University of Toronto
Toronto, Canada
jpu@dgp.toronto.edu

Steven Dow
University of California, San Diego
La Jolla, USA
spdow@ucsd.edu

Tovi Grossman
University of Toronto
Toronto, Canada
tovi@dgp.toronto.edu

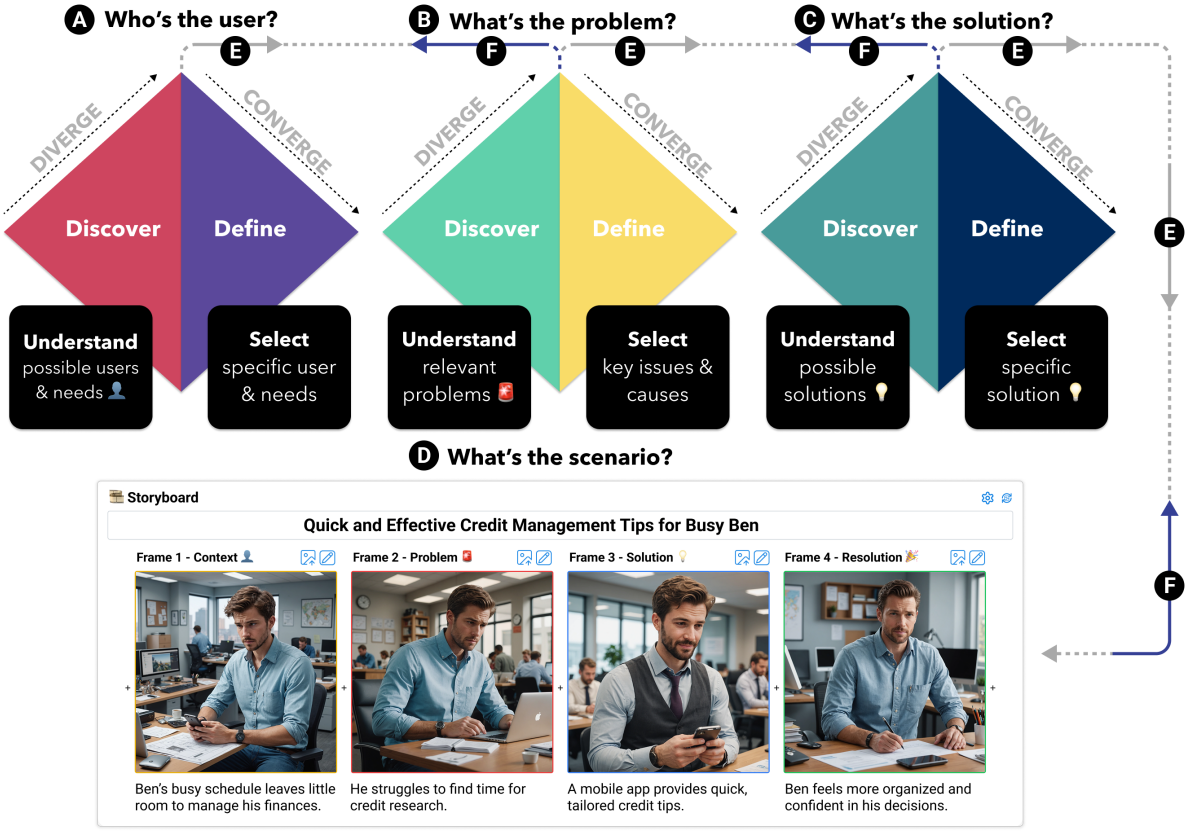


Figure 1: StoryEnsemble enables dynamic exploration and iteration in the design process by leveraging AI and supporting forward and backward propagation of changes across interconnected stages—(A) personas, (B) problems, (C) solutions, and (D) scenarios. Users can move fluidly between stages, revise earlier assumptions, and (E) propagate changes forward—enabling the kind of iteration and exploration that design frameworks promote but is often difficult in practice. They can also begin from later stages, such as solutions or scenarios, and (F) propagate changes backward to reframe upstream decisions.

UIST '25, Busan, Republic of Korea

ACM ISBN 979-8-4007-2037-6/2025/09
https://doi.org/10.1145/3746059.3747772

Abstract

Design processes involve exploration, iteration, and movement across interconnected stages such as persona creation, problem framing, solution ideation, and prototyping. However, time and resource constraints often hinder designers from exploring broadly, collecting feedback, and revisiting earlier assumptions—making it

difficult to uphold core design principles in practice. To better understand these challenges, we conducted a formative study with 15 participants—comprised of UX practitioners, students, and instructors. Based on the findings, we developed StoryEnsemble, a tool that integrates AI into a node-link interface and leverages forward and backward propagation to support dynamic exploration and iteration across the design process. A user study with 10 participants showed that StoryEnsemble enables rapid, multi-directional iteration and flexible navigation across design stages. This work advances our understanding of how AI can foster more iterative design practices by introducing novel interactions that make exploration and iteration more fluid, accessible, and engaging.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**; **User centered design**.

Keywords

Design frameworks; design thinking; Double Diamond; scenario-based design; forward-backward propagation; backpropagation; human-AI interaction

ACM Reference Format:

Sangho Suh, Michael Lai, Kevin Pu, Steven Dow, and Tovi Grossman. 2025. StoryEnsemble: Enabling Dynamic Exploration & Iteration in the Design Process with AI and Forward-Backward Propagation. In *The 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25), September 28-October 1, 2025, Busan, Republic of Korea.* ACM, New York, NY, USA, 36 pages. https://doi.org/10.1145/3746059.3747772

1 Introduction

> "Few ideas work on the first try. Iteration is key to innovation." — Sebastian Thrun

Design frameworks such as design thinking [4], the Double Diamond model [19], and scenario-based design [57] provide structured approaches to addressing complex, open-ended problems. They have had an enormous influence for the past few decades, driving innovation across both academic and professional settings. Despite their diverse formats and implementations, these frameworks share core principles: iterative processes, divergent and convergent thinking, and user-centered storytelling [33, 48, 54, 77].

Unfortunately, despite their value, these principles can be difficult to apply in practice. From the initial persona creation to the final high-fidelity design storyboard, time and resource constraints often limit individuals' and teams' ability to iterate, explore a broad range of ideas, or refine solutions through structured convergence [31]. Collecting meaningful feedback requires significant time and effort [23, 25, 36, 81], and when that feedback prompts revisions—such as reframing the problem or rethinking solutions—it can trigger a costly loop that requires revisiting earlier assumptions and reworking previous design artifacts such as personas, problem statements, scenarios, or storyboards. This overhead often deters individuals and teams from iterating [28, 70, 79].

Moreover, because design stages are tightly interconnected, getting stuck in one stage—such as struggling to generate diverse ideas [76] or define actionable problem statements—can stall the entire process. Without a clear problem framing or a range of initial directions, teams may struggle to develop meaningful prototypes or gather useful feedback. Similarly, at later stages, while artifacts like storyboards offer powerful ways to visualize user experiences and communicate design concepts, creating effective storyboards can require significant time and skills. This can result in oversimplified narratives that fail to capture nuanced interactions or encourage reflection and discussion [72]. When personas, problem statements, or scenarios remain underdeveloped, it becomes difficult to imagine concrete use cases or evaluate solutions in context. In practice, such bottlenecks often make it hard for teams to make meaningful progress and move through the design process effectively.

In light of such challenges, AI is emerging as a promising solution [60]. Recent work has demonstrated AI's capability to support various design tasks, including brainstorming [68, 69, 80], persona generation [30, 62, 64, 83], and visual content creation [3, 74]. These tools can significantly reduce the labor required for key design activities, helping teams work within tight resource constraints. Notably, researchers have shown that large language models can create synthetic user personas comparable to those based on real user data [62], potentially valuable for contexts involving vulnerable or hard-to-reach populations [30, 36].

However, current AI-supported design tools still fall short of addressing the full complexity of design workflows, particularly the iterative nature of the process. Prior work has largely focused on isolated stages of the design process [9, 30, 62, 64, 82], overlooking the need to support the iterative and interconnected nature of design, where changes in one part of the process often require adjustments in others. As a result, the current landscape leaves designers with no tools that enable them to manage dependencies across stages and encourage reflection and iteration throughout the design process [21, 39, 54].

To address this gap and demonstrate how AI can facilitate key design principles across the design process, we developed StoryEnsemble, a system that supports widely shared principles across design methodologies—such as structured iteration [7, 11, 23, 54, 56], divergent and convergent thinking [68], and scenario-based design [53, 57, 72]. StoryEnsemble enables dynamic exploration and flexible iteration by empowering users to explore diverse ideas, refine through structured feedback, propagate changes seamlessly across interconnected stages, and create easily editable storyboards. Our contributions include:

- StoryEnsemble,[1] an interactive system that leverages a node-link interface and novel interaction techniques—such as forward and back propagation—to enable rapid exploration and flexible iteration.
- A formative study identifying challenges faced by students, instructors, and practitioners in applying iterative, exploratory, and feedback-driven design principles in practice.
- A system evaluation study that advances our understanding of how AI can support workflows grounded in core design principles—such as iteration, divergence-convergence, and scenario construction—particularly in design education and practice.

In the remainder of the paper, we review related work on design frameworks and AI design tools, then present our formative study

[1] https://storyensemble-research.github.io/

identifying key challenges that students, instructors, and practitioners face when applying design principles. We then describe StoryEnsemble and its features, followed by an evaluation with 10 participants that demonstrates how it enhances exploration, iteration, and scenario development. We conclude with implications for future AI-assisted design systems, as well as for other domains involving interconnected, iterative workflows.

2 Related Work

We first examine common principles across design methodologies, then explore how generative AI is transforming creative workflows, and review existing AI design tools.

2.1 Principles Across Design Methodologies and Practical Challenges

The landscape of design methodologies is diverse, with frameworks varying significantly in their structure, terminology, and specific approaches [26, 48]. Research has documented well over a hundred distinct design approaches across various domains [33]. From design thinking and the Double Diamond model to human-centered design, participatory design, service design, and scenario-based approaches, designers have numerous options for characterizing and structuring their design processes. This diversity reflects the fact that no single approach is universally optimal and applicable. Organizations and practitioners frequently adapt frameworks to their specific contexts, leading to continuous evolution and customization of design processes [26, 40].

Despite this variety, comparative analyses have shown that the underlying principles transcend specific methodologies [33]. These shared principles include iterative refinement, alternating phases of divergent and convergent thinking, human-centered design, and user-centered storytelling [53]. While different frameworks visualize and structure these principles distinctly—some making the divergent-convergent pattern explicit (e.g., Double Diamond [19]), others emphasizing empathy and iteration (e.g., design thinking [20]) or scenario-based design—the core approaches to problem-solving remain consistent across methodologies. Many frameworks that appear linear in representation are fundamentally iterative in practice, incorporating feedback loops that enable practitioners to revisit earlier stages as new insights emerge [20].

However, a notable gap exists between theoretical frameworks and their practical application [5, 27, 51, 55, 67, 76]. Despite their theoretical importance, these principles often prove challenging to implement due to time constraints, resource limitations, and the labor-intensive nature of key design tasks [40]. For instance, while design thinking emphasizes the importance of multiple iterations to refine ideas [56], time constraints frequently limit iteration cycles to one or a few. Even in semester-long courses, the design cycle may only be completed once, or iteration may be intentionally minimized to avoid major changes to a project.

Practitioners across educational and professional settings face several common challenges. Novice designers often gravitate toward familiar problems, frequently addressing the same types of design challenges and getting fixated [45]. Teams under deadline pressure tend to implement theoretically non-linear processes in a more linear, constrained manner [28]. Additionally, collecting and incorporating meaningful feedback becomes difficult within tight project timelines [23, 61], further widening the gap between ideal iterative refinement and practical implementation.

Our work builds on decades of research on design by contributing a system that addresses a longstanding challenge in design practice: enabling easy, seamless iteration across interdependent stages. By demonstrating how to lower the overhead of iteration, our work aims to bridge the gap between theoretical ideals and the realities of design processes.

2.2 Generative AI for Creative Processes

The recent advances in AI opened up new interaction possibilities and are radically transforming creative workflows across various fields [1, 15, 17, 66], ranging from creative writing [38, 41, 46] to design [47, 76] and content creation [12, 75]. For example, in TableBrush, users can generate stories based on a narrative arc sketched by the user [18]. Liu et al. explored how text-to-image AI can be incorporated into the 3D design workflows, showing that text-to-image models can assist with producing reference images, preventing design fixation, and inspiring design considerations [43]. These advancements highlight the transformative potential of AI in supporting more dynamic and flexible creative processes.

AI is increasingly being integrated into the creative process to expand idea exploration. In Luminate, Suh et al. explored a novel interaction paradigm where the system generates a range of ideas designed to cover the design space instead of producing a single response to a prompt [68]. Similarly, Wang et al. contributed RoomDreaming, a system that enables interior designers to co-design with AI [76]. While both systems focus on supporting designers in exploring diverse ideas and iterating on their designs, RoomDreaming does not address the interconnected nature of stages in design thinking, where changes in one stage, such as problem definition, often require adjustments in subsequent stages like persona development or solution generation.

Generative AI is also being explored for its potential to automate feedback and enhance the iterative process. For instance, Duan et al. developed a Figma plugin where LLMs generate feedback on UI mockups, identifying subtle design errors and improving textual elements [24]. Similarly, Xiang et al.'s SimUser system automates usability feedback, producing insights that closely mirror those generated by human users [79]. Assuming feedback will always be accessible with the advances in AI, E et al. explored the tradeoffs of providing access to feedback throughout a design task versus only providing feedback after [25]. These systems demonstrate how AI can facilitate rapid iteration by providing timely, automated feedback, allowing designers to refine their work more efficiently.

Our work builds on these efforts by utilizing AI-generated feedback to facilitate transitions across various design processes, from persona creation to design artifact creation, demonstrating the new role that feedback can play in the AI-powered framework for iterative design and creative exploration. While the traditional design process has used feedback primarily for evaluation, our approach highlights its potential as a generative tool—one that can be invoked on demand and propagate changes forward and backward in the design process, interconnecting different design stages to stimulate

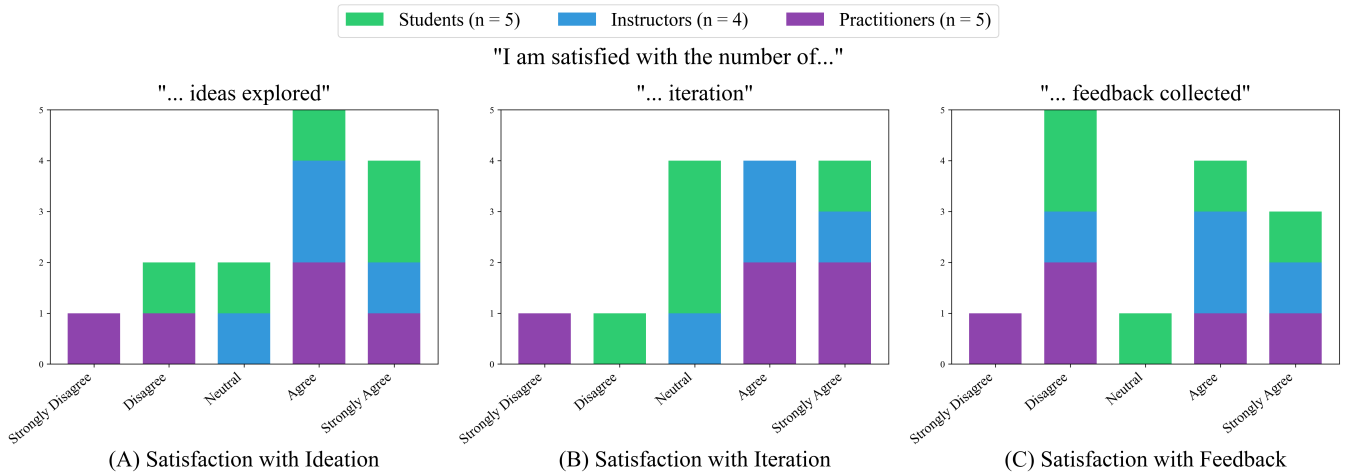


Figure 2: Results from a formative study with students, instructors, and practitioners and their satisfaction with ideation, iteration, and feedback collection.

new directions and support flexible iteration throughout the design process.

2.3 AI Tools for Design

Generative AI has become increasingly important in supporting individual stages within the design process, such as generating persona [16, 62, 82] and prototyping [8, 9, 30, 59, 64]. He et al. proposed Cinemassist, a system that supports creative exploration of potential cinematic compositions for users creating 3D animations. For each keyframe in the animation, users are presented with different scene options and frame-level options [32]. Wang et al. developed AIdeation to support search and combination of reference images for visual concept designers [78]. The UIDEC tool scaffolded the design brainstorming process by emphasizing design under constraints and using AI to generate examples [65]. LLMs have also been used to generate user research data, providing designers with fast, scalable tools for modeling user behaviors and needs [34, 59, 62]. Other tools focus on AI-assisted novel artifact generation, such as sketching [42], motion comics [13], or manipulable design tokens [63]. These advancements suggest that AI can reduce the manual effort involved in various design stages, allowing designers to focus on more complex creative tasks.

While many AI tools have focused on supporting tasks within specific stages of the design process, limited attention has been paid to augmenting the overall process and bridging fragmentation between stages. Our work contributes to the growing body of research on AI-assisted design by addressing this gap, demonstrating how AI can help manage interdependencies and support iteration across related artifacts in a cohesive manner.

3 Formative Study: Understanding Challenges, Opportunities, and Practices with Survey & Interview

To understand how generative AI can enable new design workflows, we recruited 15 participants for two formative studies involving a survey and interviews.

Survey. The survey study aimed to identify challenges related to exploration, iteration, and feedback collection, along with potential opportunities and concerns regarding AI integration. While we intended to keep the scope open to various design processes, we referenced *design thinking* in questions about design practices to provide a familiar reference point and facilitate discussion around common activities such as problem framing, ideation, and iteration.

Interview. The interview study focused on understanding the practices surrounding storyboarding for feedback collection among UX designers. We concentrated our interviews on only UX professionals, since industry practices around storyboarding—compared to academic settings—are less well-documented and offer opportunities for new insights, such as how storyboards are used in communicating with stakeholders.

3.1 Procedure & Participants for Two Formative Studies

For the survey study, we recruited 8 participants (gender: 4M, 4F; age: M = 30, SD = 7.7, range = [24, 46]; referred to as F1-8). Out of the eight participants, five had design learning experience as students (through HCI courses at universities or UX bootcamp), four had experience teaching design process as instructors (M = 4.3 years of teaching experience, SD = 5.3, range: [1, 12]), and five had experience as design practitioners (M = 4.6 years of design practice experience, SD = 3.2, range: [2, 10]). Participants completed an online survey where they described challenges they faced during the design process, particularly with respect to exploration, iteration, and feedback collection. They were then asked how these

challenges might be addressed and whether AI could support these processes. Each participant received a $15 gift card.

In the interview study, we conducted 60-minute semi-structured interviews with 7 professional UX designers (4M, 3F; F8-15) who had an average of 7.1 years of experience (SD = 4.7, range: [2, 15]) in UX design and storyboarding. On average, they had received feedback on their UX scenarios (in text) 12.6 times (SD: 7.9; range: [3, 25]) and storyboard 9.6 times (SD: 9.3; range: [3, 30]). Participants discussed the types of storyboards they used—text-only scenarios, visual storyboards, or photos—and how they collected feedback from stakeholders. They also described the kinds of feedback they typically sought. Each participant received a $25 gift card.

We provided participants with a definition and overview of design thinking principles during recruitment and again at the beginning of both studies to ensure a shared understanding of the terminology and process.

3.2 Identified Needs (Ns) from Survey with Students, Instructors, and Practitioners

3.2.1 Challenges in Applying Design Thinking in Projects.

N1 Support time-consuming tasks in design thinking
N2 Support ideation at any design phase
N3 Support translation of user needs to problem statement

F1 (UX designer) shared that **time and budget constraints** often limited the full application of design thinking. F2 (practitioner) and F6 (student) noted that **ideation** was difficult, with initial ideas often basic and better ones emerging later in the process. Integrating diverse user perspectives was also a challenge, sometimes leading to overlooked groups (F3, student). F4 (practitioner) and F5 (student) found **translating user needs and challenges into actionable problem statement** challenging. Finally, F7 (instructor & UX consultant), who has been teaching HCI and UX courses for 12 years, touched on the **theory-practice gap**, explaining that the widely taught five-stage design thinking model oversimplifies the process, saying:

> "The process is described as linear, but in reality, it's not that linear. Students should understand that not every problem will require all the steps and that things need to be customized on a per-problem basis."

To address these challenges, participants suggested adapting design thinking according to the stakeholders' needs, time constraints, and budget (F1) and using collaboration or AI for ideation (F2). Other solutions included getting help from team members to refine problem statements (F4), building systems to guide problem definition, and using AI to generate feedback (F5).

3.2.2 Challenges in Transitioning Between Design Stages.

N4 Support divergent and convergent thinking
N5 Support synthesis and translation of ideas
N6 Support returning to previous design phases

F5 (student) stressed the value of the **double diamond process, which encourages both divergent and convergent thinking**. F6 (student) pointed out the importance of **synthesizing ideas**, noting that while generating ideas is commonly highlighted, the ability to **distill and translate those ideas** is often overlooked yet crucial. F7 (instructor) mentioned the **flexibility to revisit previous phases of the design process**, stressing that it is okay and it should be encouraged to return to earlier stages if new insights emerge.

3.2.3 Challenges in Receiving Feedback and Iterating Design.

N7 Support more frequent design iteration via feedback
N8 Facilitate efficient feedback collection at any design phase

With respect to ideation, student participants (F3, F5, F8) indicated that the number of ideas they explored depended on the course requirement and instructor. F3 said: "The **exploration was driven by instructor feedback and collaborative discussions with my peers**, which encouraged me to think beyond the obvious." F2 and F3 (students) shared that while they "felt that the number of ideas explored could have been greater," "**time constraint**" is always a problem.

Regarding iteration, except for F6, every student participants (F3, F5, F7-8) mentioned they iterated at least once, while varying in the number of iteration. For example, F2 had a chance to iterate only once during a workshop, while F8 did not iterate "during the workshop." However, they all agreed on the feedback's role in iteration. F8 said, "teacher feedback come in as a good 'reason' for team members to agree on going back." F6, who said there was no iteration in her class, attributed this to the absence of feedback, saying "classes don't offer enough time to engage in feedback." F2 and F3 both mentioned they iterated based on feedback from "instructor and peers." Students (F3, F5-6, F8) not satisfied with the number of iterations in the course noted that "the course's pace ... pushed [them] to move forward without enough refinement" (F3). To address this problem, they suggested giving students "**more chances to [get] feedback**" (F6) and teach the importance of feedback and iteration to team members who "do not want to go back and iterate and consider [iteration] very inefficient" (F8).

While all instructor participants (F4, F6-7) strongly agreed that *it is important to return to previous stage of the design process and iterate*, its **time-consuming and laborious nature** made it challenging. F7 said that even for "12-week courses, [he finds] it hard for students to go back to previous stages." He added that the most he asks the students to do is to iterate *within* specific stages—"especially at the stages of lo-fi, mid-fi, and evaluation." F6 believed that "learning how to iterate is a skill that requires time."

Practitioner participants (F1, F6-8) also **acknowledged the importance of iteration**. However, the number of times they iterate varied depending on the project schedule (F7), client (F1), and feedback (F5). F1 and F5 noted that they "return to previous stages" when there are new insights from formative studies, users' feedback, or stakeholders' requests to add or remove some features.

The practices and experiences surrounding collecting feedback varied. Some (F3, F8) reported collecting feedback more frequently than others (F2, F5-6). F3 collected feedback at every stage, F5 collected feedback only on specific stages such as solution or testing stages. As shown in Fig. 2, practitioner participants (F1, F5, F7-8) were also generally **dissatisfied with the number of feedback they collect**. Several participants (F3, F5-6) expressed dissatisfaction with the current status quo and suggested potential solutions. F5 suggested using generative AI to generate "feedback... so that a

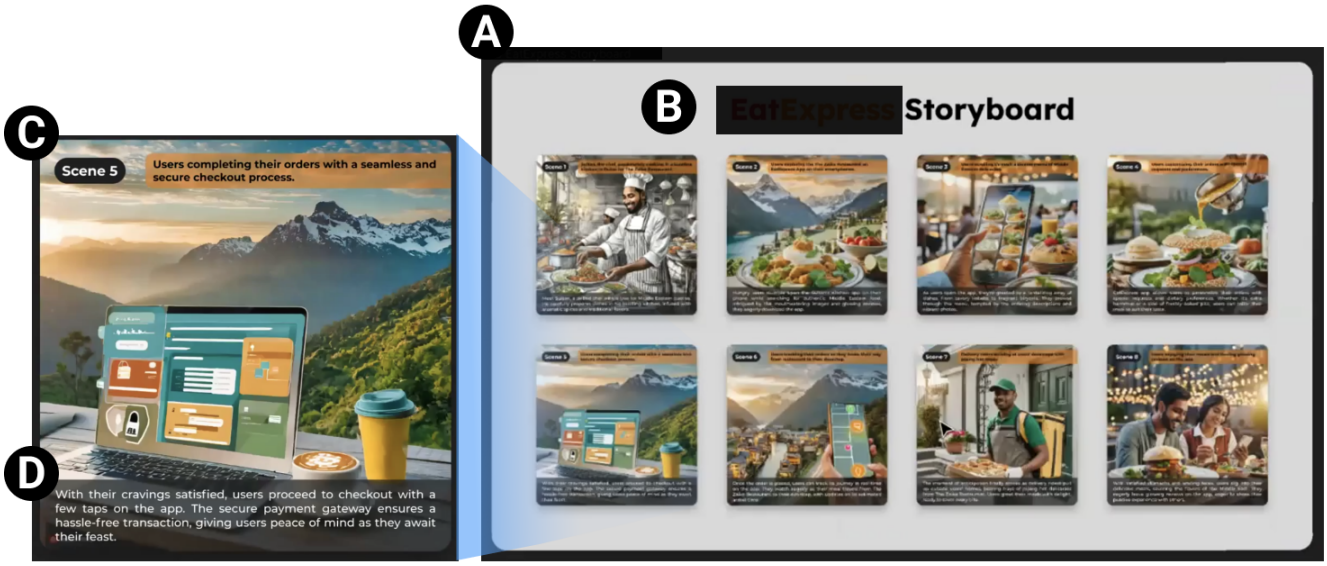


Figure 3: A screenshot of (A) F15's storyboard, with the (B) title (redacted) at the top. (C) is a frame featuring an image that shows the interface with (D) caption. This, along with general storyboard visuals [72], informed our storyboard design (Fig. 11).

more frequent feedback process can be easily realised." F3 suggested integrating "regular feedback checkpoints into the course structure, ensuring that feedback is shared consistently at each stage of the design process."

3.2.4 Challenges with Using AI for Design Thinking.

N9 Reduce manual effort in requesting AI feedback
N10 Synergize human and AI capabilities for co-creation

Many participants (5 of 8) had experience using AI. The most common use case was for generating "a wide range of ideas" (F3). Several participants (F2, F3, F8) shared that they used it to explore "diverse perspectives [to] spark new thoughts and directions on a problem" and "get the divergent thinking going." In addition to ideation, F4 used it to create storyboard and "analyze feedback to identify common themes." All participants supported the idea of integrating AI in their design process. F1 felt the current practice is "still very manual" and that "there has to be a way to **make it more automatic**." However, F3 and F4 also emphasized that it has to be a **co-creative process**, saying "AI should complement human creativity and judgment, not replace it. The best outcomes often come from a **synergy between human insight and AI's capabilities**" (F4). Several participants (F5, F6, F7) mentioned using AI to generate "feedback" (F5) "at every stage of the design" (F7) to help "break free from design fixation" (F8). F7 shared that it can perhaps "help with the iteration by suggesting more ideas/solutions." F7 went on to suggest that given that "the constrains of 10-weeks course projects (while students take multiple courses) and the timelines in industry often mean that **steps have to be skipped** and you can **only do one full cycle** (without much opportunities to go back)," "we could ***sub-contract* some aspects to AI** [to] free up more time for designers to address these challenges."

3.3 Identified Needs (Ns) from Interview with UX Designers on Storyboarding Practices

3.3.1 Storyboard is essential with many benefits but the costly nature of producing it prevents its use. All the UX designers acknowledged that the costly nature of storyboarding prevented them from using them as much as they would like considering the **many benefits compared to alternatives** such as text-based scenarios. F12 said: "What's awesome about storyboards is that it just makes ideas easier to discuss and gives you a fuller picture. A higher level view." However, she stated that she **uses storyboard sparingly**, saying "I've used storyboards when I've had a greater amount of time which is unfortunate, or when I really needed help explaining a beginning to end concept." Several participants (F12, F13) working as freelancers also mentioned that the **clients who hire them avoid requesting storyboarding** as that would require them to pay more for the time spent on creating storyboards, highlighting the costly nature of storyboarding even in terms of finance.

N11 Support image consistency across frames in storyboard

3.3.2 Ensuring image consistency in storyboard is challenging when AI is used to generate images for storyboard. During the interview, F15 showed storyboard (see Fig. 3) he created using ChatGPT's image generation capability. While AI eased the process of creating storyboard for efficient communication with his clients, he noted the challenges present in his current workflow such as **maintaining consistency across image frames in the storyboard**. He explained that he creates each frame separately, which led to inconsistency in the generated images. As shown in Fig. 3, images varied in style and tone. For example, some images were in drawing style (1st frame), while others were photo realistic.

4 StoryEnsemble

We first present the design goals for StoryEnsemble, grounded in needs (**Ns**) identified through formative studies, a literature review, and pilot studies. We then provide an overview of the interface and its features, followed by an example workflow.

4.1 Design Goals (DGs)

The DGs are organized into three themes: (1) *core innovations*, (2) *creative support*, and (3) and *usability across diverse contexts*.

I. Core innovations

DG1. Enable non-linear engagement (N7). Allow users to start at or engage with any design stage, with automatic updates ensuring consistency across related artifacts, supporting the iterative workflows emphasized in many design methodologies.

DG2. Enable on-demand feedback (N8, N9). Offer real-time feedback on design artifacts and streamline its incorporation, encouraging users to reflect and continuously refine ideas with tight feedback loops.

II. Creative support

DG3. Promote diverse idea exploration (N2, N5, N6). Encourage broader ideation while avoiding AI-induced homogenization [2]. Provide tools for rapid generation and variation to help users overcome fixation.

DG4. Support divergent and convergent thinking (N4). Enable fluid movement between generating possibilities (Discover) and narrowing options (Define) based on insights and feedback, mirroring the structure of the Double Diamond model.

III. Usability across diverse contexts

DG5. Simplify artifact generation using AI (N1, N3, N11). Reduce effort needed to create personas, problem statements, solution ideas, and storyboards using AI.

DG6. Support flexible AI engagement (N8, N9, N10). Accommodate three levels of AI involvement to match varying design contexts and user preferences: (1) *no AI involvement*, where users manually create and refine all design artifacts; (2) *partial AI assistance*, where users provide key information and AI fill gaps; and (3) *full AI generation* for exploration, iteration, and feedback.

DG7. Ensure adaptability to diverse contexts (N10). Design the system to be simple yet robust, so that it can scale across a range of design methodologies and implementation settings—from short workshops to extended courses—and support users with varying levels of expertise.

4.2 Interface: Basic Components

Here, we describe the basic components of StoryEnsemble, which structure the design process into modular nodes that users can create, edit, and iterate on—either manually, with AI assistance, or through a hybrid approach. As shown in Fig. 4, StoryEnsemble offers four node types (Fig. 4D-G), each corresponding to one of the first four phases (Empathize-Define-Ideate-Prototype) of the 5-stage design thinking process defined by Stanford d.school [20]:

👤 **Persona Node (Fig. 4D & 5)** [Empathize]: Helps users define target users through attributes like name, location, needs, and challenges. These can be filled in manually, completely generated using AI, or partially completed by users with AI assisting to fill

Figure 4: StoryEnsemble Interface: (A) opens up a panel (Fig. 6) to add context and generate chains of ideas on a specific theme, such as Urban Sustainability, (B) allows users to add empty nodes and edit node content manually (Fig 10) or with AI (Fig. 9). The system features different nodes types: (C) context node (e.g., Urban Sustainability) for grouping ideas, (D) persona nodes, (E) problem nodes, (F) solution nodes, and (G) a storyboard node. (H) Edges ( ) map the relationships and dependencies between ideas, showing how ideas influence each other throughout the design process. The figure shows outputs generated using the Start brainstorming feature, detailed in Section 4.3.

Figure 5: 👤Persona node: An (A) illustrative image, (B) name and (C) attributes which represent a persona. Each node includes (D) a contextual toolbar that appears when the node is selected. Toolbar options include: (a) view AI-generated feedback (Fig. 8), (b) revise using AI (Fig. 9), (c) edit manually (Fig. 10), (d) generate problem node, (e) generate more personas (Fig. 7), and (f) duplicate the node.

in remaining fields. This flexibility supports cases where designers already have established personas (e.g., from user interviews or research) that they want to incorporate directly, or where they have specific characteristics in mind but need help expanding the persona with additional details.

🚨 **Problem Node (Fig. 4E)** [Define]: Allows users to define key challenges, with fields such as context, stakeholders, and objectives. Users can input their own insights, leverage AI to generate complete problem statements, or take a hybrid approach where they specify certain aspects and use AI to develop or refine others—providing flexibility for cases where designers have identified specific issues but need help articulating them fully.

💡 **Solution Node (Fig. 4F)** [Ideate]: Enables users to propose solutions by outlining attributes like problems addressed, key features, and benefits. As with the other nodes, users can manually create solutions, generate them entirely with AI, or combine their own ideas with AI-generated content—allowing designers to start with a specific solution concept and use AI to elaborate on its details and implementation.

🎬 **Storyboard Node (Fig. 4G, 11)** [Prototype]: Enables users to create visual narratives for their design concepts, either by generating frames based on earlier stages using AI or by manually authoring them for full creative control. Fields include frame type, image, and caption, supporting rapid exploration and feedback collection. While prototypes can range from low-fidelity (as in this static storyboard) to high-fidelity (e.g., interactive prototypes in Figma), StoryEnsemble focuses on low-fidelity storyboards to facilitate quick idea exploration and early feedback, aligning with the goal of providing something "tangible enough to receive feedback in the short-term" [71].

No specific node represents the Test stage, but feedback mechanisms are built into each node type to support iterative refinement (Fig. 8), which we elaborate in Sec 4.3.

Node Elements. Each node features illustrative images (Fig. 5A) to enhance engagement and relatability—helping facilitate empathy with personas, and making problem and solution nodes more visually compelling and easier to connect with (DG5). To maintain consistency for nodes chained together, as shown in Fig. 4, nodes such as persona or solution nodes that are generated after their prior nodes use character description generated from prior nodes when generating illustrative images (See 'Generate visual character descriptions' in Table 9 for prompt and Fig. 20 for pipeline).

Connection Mechanisms. StoryEnsemble enables users to establish dependencies between different stages of the design process through connection handles. Persona nodes have a single connection handle at the bottom, allowing them to serve as starting points that influence subsequent stages. Problem and solution nodes feature connection handles at both top and bottom, enabling them to receive input from earlier stages and propagate changes to later ones. Storyboard nodes, positioned at the end of the process, have only a top connection handle to receive input from earlier design artifacts. These handles, shown in Fig. 4H, allow users to establish dependencies between different stages—for example, linking a persona to a problem node (👤 ⇢ 🚨).

These connections serve dual purposes: they visualize relationships between design artifacts and function as pathways for propagating changes. When a node is updated, the system identifies connected nodes that might be affected and offers users the option to cascade these changes forward or backward through the dependency network. This interconnected structure ensures that modifications in one stage can be propagated throughout the design process, supporting the iterative nature of design thinking.

4.3 Features for Supporting Rapid Exploration and Flexible Iteration

We added features in StoryEnsemble designed to support rapid exploration and flexible iteration (**DG3**). This design decision is grounded in the principle that "design is often an interplay between ideation and iteration" [29, 56]. In StoryEnsemble, users can iterate within and across design stages (**DG1**). Below, we describe the supporting features and their use cases in Sec 4.4.

*Start Brainstorming (Fig. 4A, 6; **DG5, 3**).* In the early stages of design, users often struggle to get started due to lack of knowledge [22, 68]. To assist with this, StoryEnsemble offers a brainstorming feature where users can start from scratch or build upon

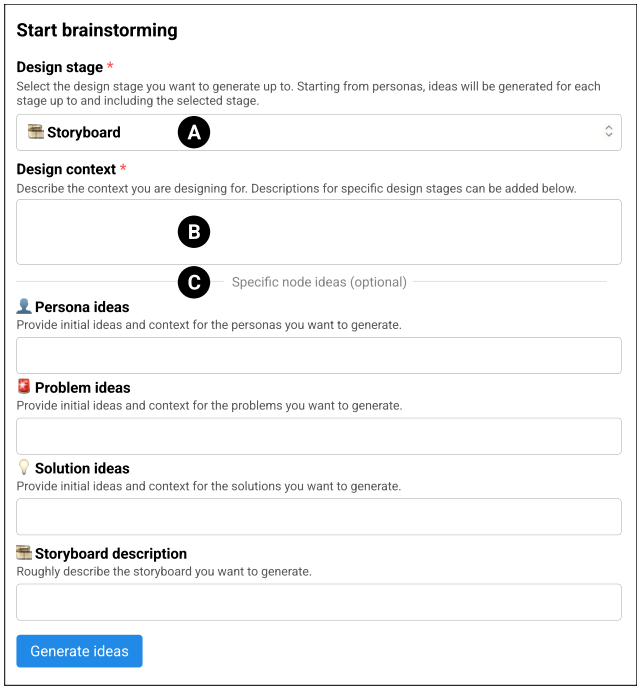


Figure 6: Start Brainstorming: A panel that opens when users select the Start brainstorming button (Fig. 4A). It allows users to generate and explore multiple design artifacts. Users can (A) specify the design stage (personas, problems, solutions, storyboards) they want to generate up to, (B) provide a design context text prompt, and (C) add optional context or text guidance for each stage. For example, users can add optional context like 'PhD students who publish at CHI, IUI, or UIST' to steer the kinds of personas that get generated.

initial ideas. The system leverages AI to generate diverse ideas based on minimal input, helping users overcome creative blocks and encouraging exploration of different possibilities. This feature specifically supports the divergent thinking phase of the first diamond (Discover) in the Double Diamond model, allowing users to broadly explore the persona-problem-solution space by generating multiple persona, problem, and solution ideas. It addresses common challenges in design practice, such as the difficulty of generating diverse ideas at the start of a project, when team members may be fixated on a few familiar personas, problems, or solutions.

Add Empty Node (Fig. 4B; ***DG3, 1****).* Tools need flexibility and support diverse contexts and user needs. This feature allows users to add any empty node and manually fill the contents in—using the edit panel shown in Fig. 10—rather than relying on AI-generated content. This feature is designed to account for situations where users already have a persona, problem statement, solution, or storyboard design, such as those informed by prior research or user interviews. By supporting manual input, it upholds the principle of human-centered design, which emphasizes grounding decisions in engagement with end users. In practice, StoryEnsemble users can complete the persona and problem nodes based on actual user data, then leverage AI—as needed or as appropriate in a given context—to explore potential solution ideas and generate storyboards to communicate their design concepts.

Generate More (Fig. 7; ***DG5,3****).* One of the key strengths of generative AI is its ability to rapidly produce new ideas [68]. To ease the ideation process, StoryEnsemble features a Generate More option that allows users to seamlessly generate additional ideas or variations of existing ones, as shown in Fig. 7. This feature ensures that users are not limited by their initial ideas and can explore beyond them. Users can use this feature for all node types except for storyboard. It helps users efficiently explore the persona-problem-solution spaces and discover ideas they might have difficulty thinking about before converging on a specific direction, as prescribed in the Develop phase of the Double Diamond model (Fig. 1). This sustains divergent thinking beyond initial ideation, encouraging users to push beyond their perspectives.

Revise with AI (Fig. 9; ***DG1****).* This feature allows users to provide any instruction as prompt to update the content of any node. Users can open the panel (Fig. 9) by selecting any node and clicking the Revise button, as shown in Fig. 5D. StoryEnsemble intelligently supports partial completion workflows: when a user manually adds data to certain fields but leaves others empty, the Revise with AI feature automatically recognizes these gaps. Upon switching to the Revise tab in the popup, the system adds 'Fill in missing values' to the instruction suggestions. This allows users to quickly complete their partially defined nodes without having to explicitly request or formulate this common task. For example, if a designer has specific demographic information about a persona from research but lacks insights about their challenges, they can simply select this auto-generated instruction to have AI suggest appropriate challenges based on the existing persona attributes.

Generate Next Node (Fig. 5d). After selecting any node (except storyboard), users can use the Generate Next Node option from the contextual toolbar to create several downstream nodes. For instance, selecting it from a persona node—Problems as in Fig. 5d—allows the user to generate related problem nodes, while selecting it from a problem node can yield multiple solution ideas. This feature enables users to extend their ideas efficiently. By allowing generation from any stage in the process, this feature encourages rapid exploration.

View Feedback (Fig. 8, Iterating within Design Stage; ***DG2****).* StoryEnsemble enables users to gather AI-generated feedback for refining individual nodes, such as personas, problem statements, solutions, and storyboards. When feedback is incorporated and the content of a node is updated, the system recognizes the potential impact on other connected nodes. For example, revising a problem statement may necessitate changes to the personas or solutions linked to that problem. The system provides a visual cue ( ), signaling users that updates in one node could affect related nodes. Users are made aware of these dependencies through icons that appear on the connected nodes, allowing them to decide whether to update the content in those nodes. This ensures that feedback-driven changes flow naturally within a design stage while maintaining consistency across the artifacts. This feedback mechanism supports the transition toward convergent thinking phases (Define and Deliver in the Double Diamond model), enabling users to evaluate and narrow down options based on critical assessment of their

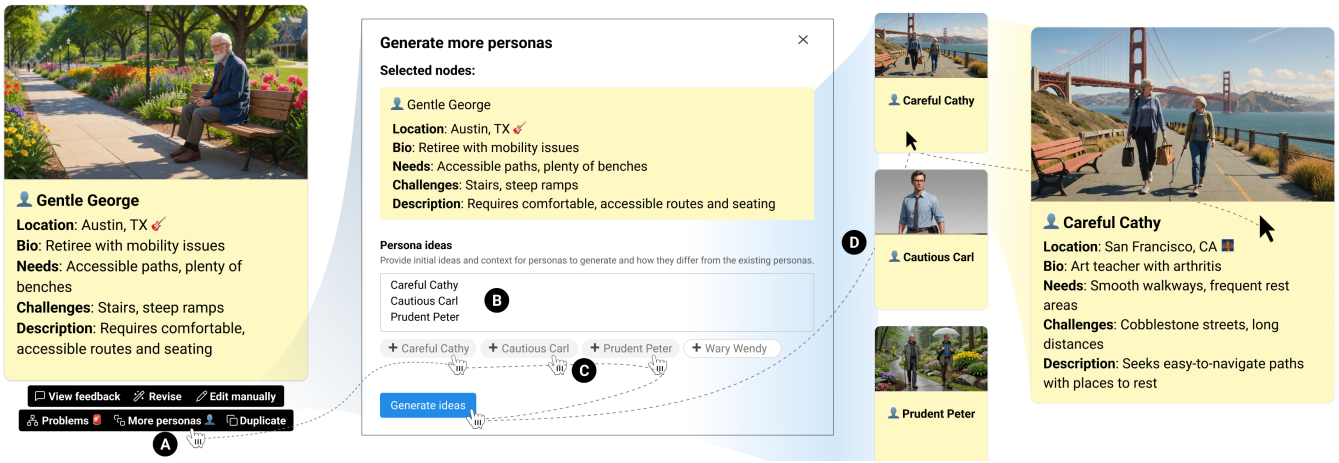


Figure 7: GENERATE MORE: Users can build on (A) a selected idea by (B) entering their own suggestions in natural language or (C) utilizing AI-generated recommendations. For example, a user could type a prompt like 'people who live in the same state (e.g., Austin, TX) as Gentle George' or '...in different states (e.g., California)' to steer the generation. The system then generates (D) additional ideas (e.g., Careful Cathy from San Francisco) to support further exploration.

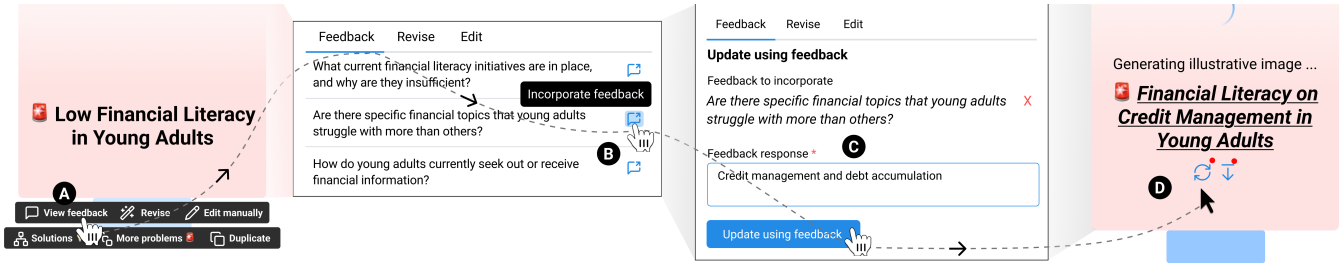


Figure 8: VIEW FEEDBACK: (A) Users can generate feedback for each node (View feedback) and (B) review a list of AI-generated feedback to select feedback to incorporate. (C) Users provide a response to a feedback question to (D) update the node accordingly.

ideas. The AI-generated feedback provides diverse perspectives that might otherwise require extensive user research or stakeholder consultation, making the convergent phases more thorough and evidence-based.

Cascading Changes (Fig. 12, Iterating Across Design Stages; ***DG3, 1****).* In addition to supporting feedback-driven iterations within a stage, StoryEnsemble facilitates the propagation of changes across different stages of the design process. When a node is updated—via manual editing (Fig. 10), AI-assisted revision (Fig. 9), feedback incorporation (Fig. 8), or creating new connections between nodes (Fig. 12)—the system identifies how these changes may influence other nodes connected to that stage. For example, modifying a persona may prompt updates to the problem statement and solution nodes that are dependent on it.

Whenever a node is updated via any of the four interactions, StoryEnsemble displays '*update*' ( ) and 'propagate change' (forward propagation: ; backward propagation:) icons that appear on affected nodes, as shown in Fig. 12. These visual indicators alert users that a change in one area of the design may require adjustments in related stages. Unlike in the VIEW FEEDBACK feature, where users can refine individual nodes within a stage, cascading changes allow for holistic updates across multiple stages of the design process. Users can choose to propagate changes *all the way to the last connected node* ( ) or *to the first node in the chain* ( ) or manually *update the connected nodes one by one* ( ), ensuring flexibility and control over the iterative process. By highlighting potential interdependencies between design stages, StoryEnsemble reduces the manual effort required to maintain consistency, allowing users to focus on creative exploration and iteration without losing track of critical connections between artifacts.

Multi-Node Selection and Batch Operation (Fig. 13; ***DG5, DG4, DG1****).* StoryEnsemble enables users to select multiple nodes simultaneously and perform operations on them as a batch. For example, users can select multiple persona nodes and generate corresponding problem nodes that take all selected personas as context, allowing for efficient exploration of problem spaces across different user types. Similarly, multiple problem nodes can be selected to generate diverse solutions addressing various challenges. This capability extends to feedback as well, where users can request feedback on multiple nodes at once to evaluate them comparatively. By supporting batch operations, StoryEnsemble streamlines the workflow when exploring diverse design directions, enabling users to quickly

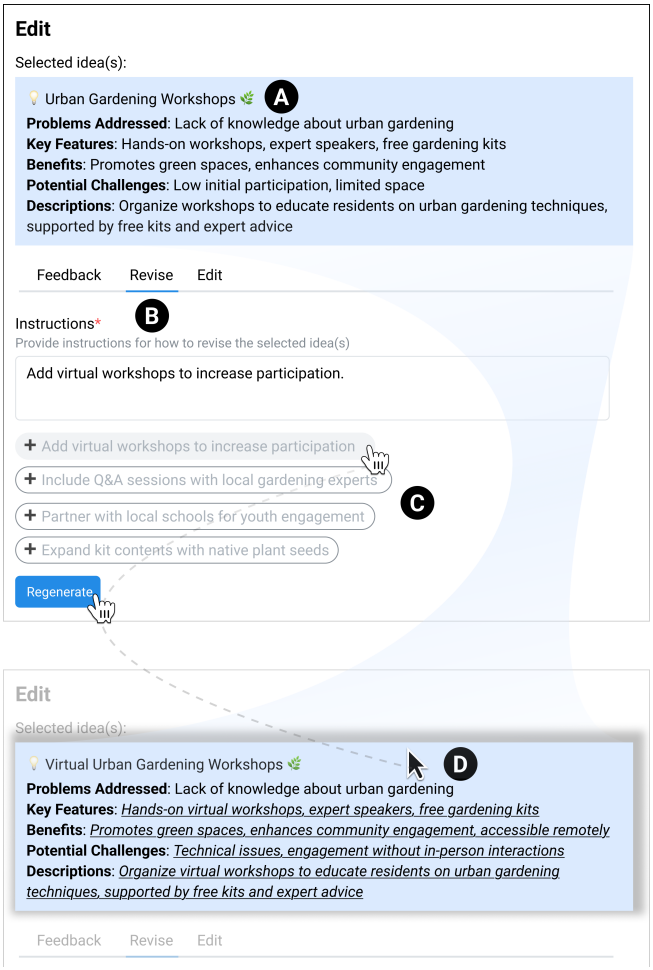


Figure 9: Revise with AI: Users can easily update (A) idea content by providing (B) natural language prompts (e.g., 'Add virtual workshops to increase participation'). (C) AI suggests instructions for iteration (e.g., '`Partner with local schools for youth engagement`'), and (D) revised content is automatically underlined for clarity.

branch out and consider multiple pathways through the design space without having to perform repetitive individual actions.

*Semantic Zoom (Zoomed Out - Fig. 4; In - Fig. 5 | **DG4**).* As users generate multiple personas, problems, solutions, and storyboards, the interface can become complex and potentially overwhelming. To address this challenge, StoryEnsemble implements semantic zoom functionality that adapts content visibility based on zoom level. When zoomed out, nodes display only their titles—as in Fig. 4—while still allowing access to full content via hover, which reveals a detailed popup window, eliminating the need to zoom in. This supports rapid exploration of the design space by allowing users to generate and compare many ideas without visual clutter. By providing an overview while preserving access to details, StoryEnsemble enables users to maintain awareness of their full idea landscape while focusing on specific areas of interest. This aligns

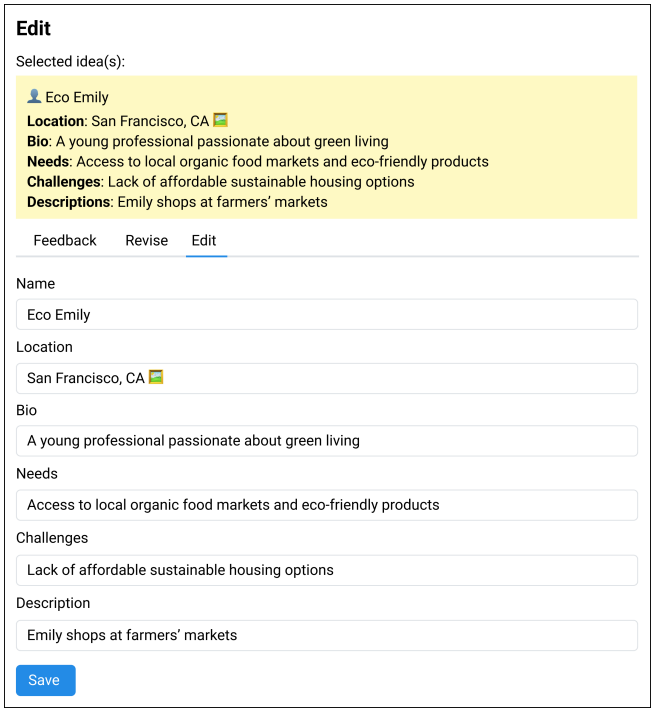


Figure 10: Designer input: Users can manually edit the contents of a node using an interface with text fields by selecting the Edit manually button on a persona, problem, or solution node. This supports flexible engagement with AI, allowing users to rely fully, partially, or not at all on AI (DG6).

with both divergent thinking principles—supporting the generation and navigation of many possibilities—and convergent thinking phases where users need to compare and evaluate options efficiently across the design space.

4.4 Example Scenario: Designing a Sustainable Packaging Solution

Alex, an eco-conscious freelance designer, is using StoryEnsemble to create a sustainable packaging solution for grocery stores. As Alex progresses through the design process, he utilizes both divergent and convergent thinking tasks, guided by the system's features. Alex's journey through this design process demonstrates how StoryEnsemble supports the complete design cycle while following the Double Diamond model's structure. Throughout this scenario, Alex moves through both diamonds of the process: first engaging in divergent thinking to explore multiple personas (Discover) before converging on 👤 Sustainability-Seeking Sara (Define), then diverging again to consider various solutions (Develop) before converging on 💡 Plant-based Packaging (Deliver). StoryEnsemble supports this natural flow between expansive exploration and focused refinement, demonstrating how generative AI can operationalize core design principles that are often difficult to implement fully in practice due to time and resource constraints.

Empathize Stage (👤Persona Node) - Divergent Task. Alex begins by brainstorming types of users who would benefit from

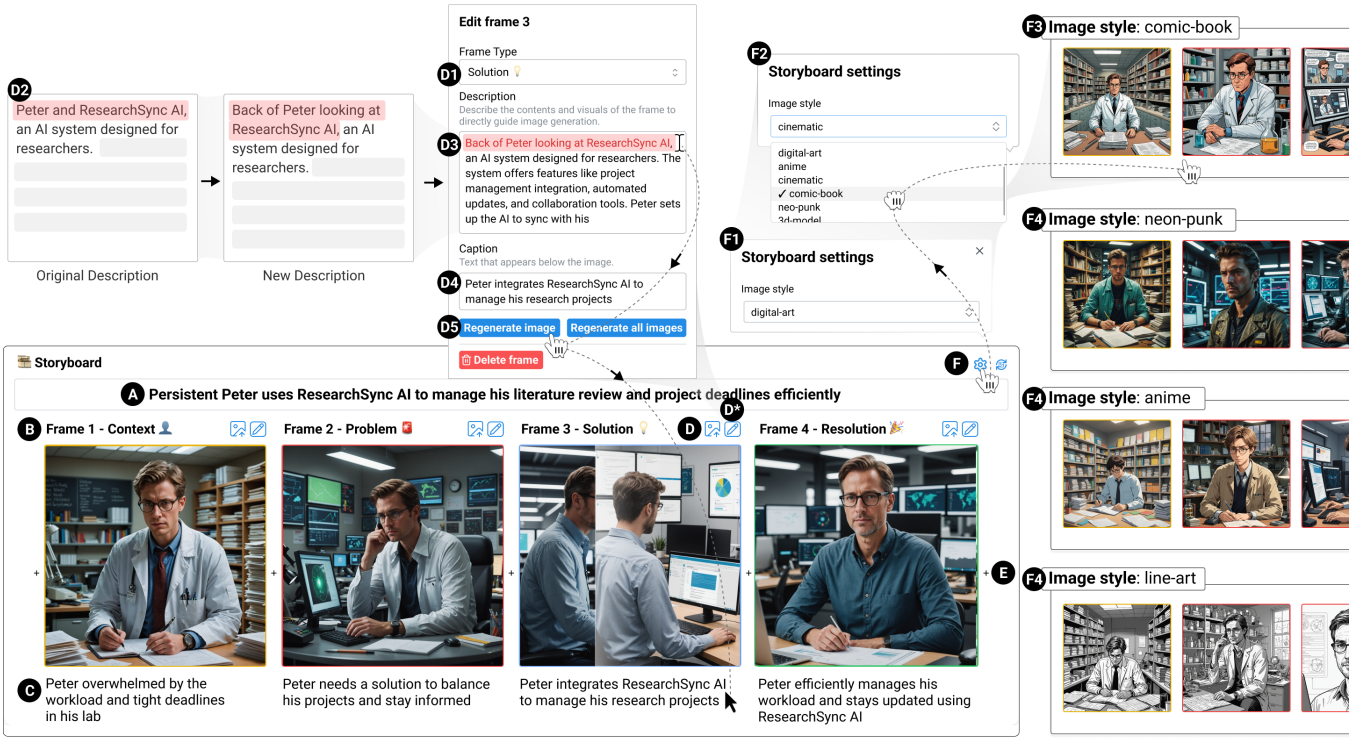


Figure 11: Storyboard: Each storyboard has (A) a title and (by default) four frames. Each frame represents a (B) specific type (Context 👤, Problem 🧯, Solution 💡, Resolution 🎉) and comes with an image and a (C) caption. Users can (D) upload an image of choice to the frame or (D*) edit it: they can update the (D1) frame type, (D2→D3) description used for guiding image generation, caption text (D4), and then (D5) regenerate image with new description ('Back of Peter looking at ResearchSync AI'). Users can also (E) add new frames to extend the narrative. The (F) storyboard settings panel enables customization of visual styles: users can choose an image style (F1→F2), such as (F3) `comic-book` or (F4) other options like `neon-punk`, `anime`, and `line-art`, and regenerate all images in the selected style. The pipeline for generating the storyboard is detailed in Fig. 20.

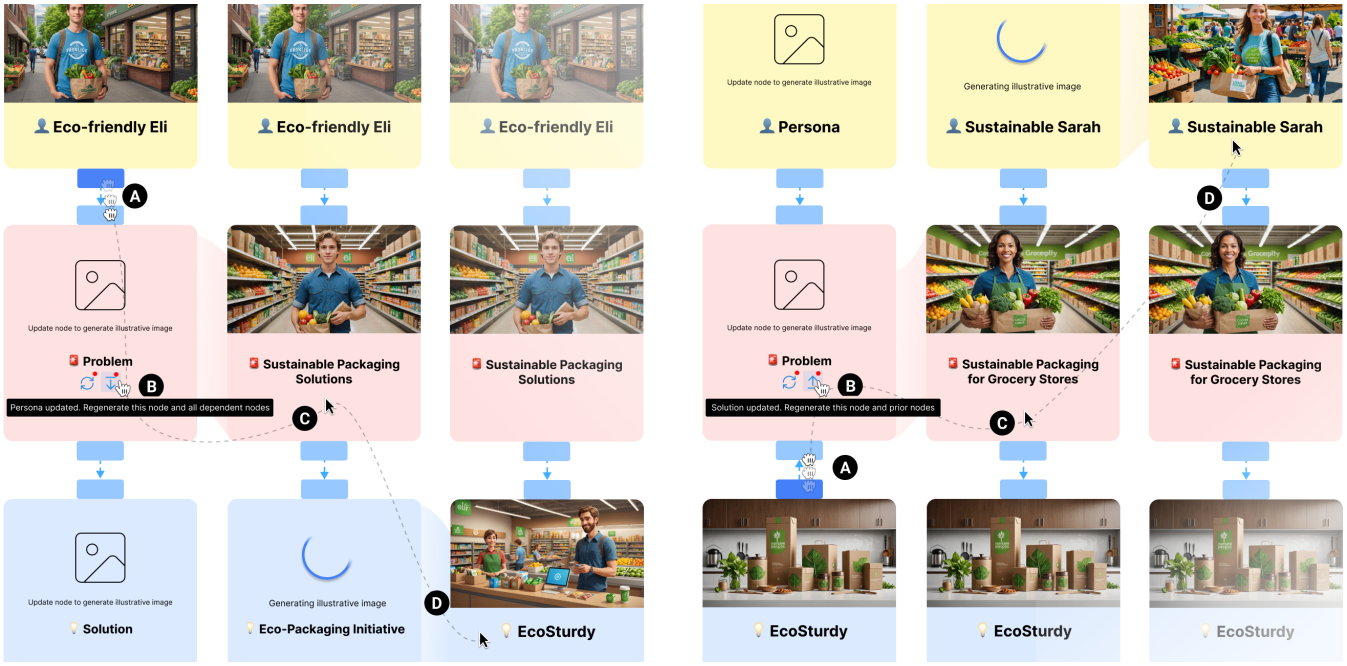
sustainable packaging. Using the AI-generated persona feature, he explores a range of possibilities and settles on 👤 Sustainability-Seeking Sara, a young professional who cares about the environment but struggles to find eco-friendly packaging at her local grocery store. The system helps Alex define Sara's attributes, such as her environmental concerns, grocery shopping habits, and desire for sustainable products. This divergent phase encourages Alex to explore various user profiles before selecting one that resonates with the problem space.

Define Stage (🧯 Problem Node) - Convergent Task. Alex needs to focus on clarifying the core problem that his solution will address. Using the problem node, Alex enters insights based on Sara's persona, allowing the AI to generate potential problem statements. The chosen problem is: 🧯 Limited availability of sustainable packaging options in grocery stores. The View Feedback feature prompts Alex to refine the problem further by considering suggestions, such as including the economic challenges of transitioning to sustainable packaging. This phase represents a convergent task, as Alex narrows down and refines the problem into a clear, actionable statement.

Ideate Stage (💡 Solution Node) - Divergent Task. In the ideation phase, Alex tries to explore as many potential solutions as possible by using the Generate More feature. To guide solutions towards environmentally friendly packaging options, Alex inputs 'biodegradable packaging' as a guiding prompt. StoryEnsemble returns several AI-generated alternatives, including 💡 Reusable Packaging Programs and 💡 Plant-based Packaging. This stage is all about divergent thinking, as Alex explores multiple avenues and evaluates which solutions are most viable. After reviewing the options, Alex decides to focus on 💡 Plant-based Packaging and begins fleshing out the key features and benefits of this solution.

Prototype Stage (Storyboard Node) - Convergent Task. Now in the prototype stage, Alex shifts to a convergent task by focusing on turning his solution into a tangible concept. Using the storyboard node, Alex creates a visual narrative to illustrate Sara's journey—from struggling to find sustainable packaging, discovering the plant-based packaging solution, and adopting it for her groceries. Alex uses Revise with AI to tweak the storyboard frames, ensuring the visuals and captions align with the design goals. This low-fidelity prototype allows Alex to quickly generate a storyboard that communicates the solution's value and gathers early feedback.

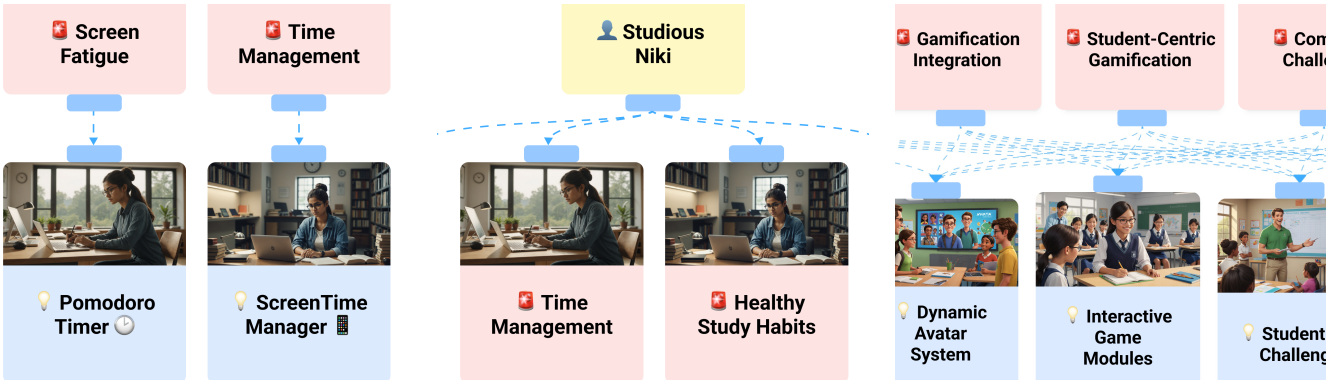
Iteration and Feedback Collection – Divergent and Convergent Tasks. Throughout the process, Alex engages in both



(a) Forward propagation: Users can (A) connect from a upstream node to downstream node (e.g., persona→problem), and then (B) click forward prop icon (⇣) to trigger (C→D) updates to later stages.

(b) Back propagation: Users can (A) connect from a downstream node to upstream node (e.g., solution→problem), and then (B) click back prop icon (⇡) to trigger (C→D) updates to earlier stages.

Figure 12: StoryEnsemble supports forward and backward propagation. Propagation can be triggered either by user-initiated handle connections between nodes (shown above) or by making changes to an existing node. Users can freely connect any node—upstream or downstream—to another. This enables users to fill in missing content, maintain coherence across stages, and iterate fluidly without being constrained by a fixed direction or sequence.



(a) One-to-One: Using the Start Brainstorming feature results in multiple one-to-one pairs (e.g., Fig. 4).

(b) One-to-Many: Using toolbar's Generate Next Node feature (e.g., Fig. 5d) from a single node creates multiple downstream nodes.

(c) Many-to-Many: Selecting multiple nodes and using toolbar's Generate Next Node feature yields multiple downstream nodes, each generated using the selected nodes.

Figure 13: Three node-to-node patterns that showcase different workflows and exploration strategies.

divergent and convergent tasks. The Forward Propagation feature helps Alex ensure that updates to Sara's persona (e.g., adding her involvement in a community recycling program) are reflected in other stages like the problem statement and solution nodes. The View feedback feature allows Alex to gather diverse feedback at each stage, a divergent task, but then Alex converges by incorporating that feedback into a refined version of the solution. The

Cascading Changes feature ensures consistency across the nodes, making it easy for Alex to iterate without losing coherence.

Final Stage: Conclusion and Refinement – Convergent Task. In the final phase, Alex reviews the project holistically. By using Cascading Changes, Alex ensures that the connections between the persona, problem, and solution are solid. This final convergence solidifies Alex's design, making sure that the artifacts work together to support a cohesive, user-centered solution.

4.5 Implementation Details

StoryEnsemble is built as a web application using the React framework. The interactive canvas is implemented with ReactFlow, an open-source library for building node-based interfaces. For AI-powered features, the system integrates OpenAI's API for text generation and Stability for image generation. Each node tracks its dependency changes, and the prompts used in the system—including those for regenerating node content—are listed in Appendix A.2.

5 User Evaluation

To assess the perceived usefulness of StoryEnsemble, we conducted a user study. Specifically, we aimed to understand: (1) how StoryEnsemble supports workflows inspired by core design principles such as iteration, divergence and convergence, and scenario-ased reasoning, (2) the usefulness of StoryEnsemble's features, and (3) potential applications of StoryEnsemble, especially for learning, teaching, and using the iterative, exploratory design process.

Study Setup & Rationale. Given the novel aspects of StoryEnsemble, we opted for an exploratory study, a method increasingly used to evaluate AI systems offering novel experiences [10, 35, 68, 69]. Like recent systems, StoryEnsemble introduces a variety of features (as shown in Sec 4), including innovative ones like backward propagation (Fig. 12b). This made an exploratory approach well-suited for examining the system's broader potential and the perceived utility of each feature. Another key motivation was to understand how StoryEnsemble could support diverse user groups—students, instructors, and practitioners—across a variety of design methodologies and workflows. The open-ended nature of this inquiry reinforced the need for an exploratory study; a more narrowly defined, hypothesis-driven approach (requiring a baseline) would have limited our analysis and prevented us from a broader examination of the system's flexibility, usefulness, and overall potential. While a hypothesis-driven approach is equally valuable, we see it as more appropriate as a next step—to build on the findings from this exploratory work to form hypotheses and evaluate them in a more controlled setting.

5.1 Procedure

The 90- to 120-minute evaluation study (Fig. 14) was conducted in four main phases: tutorial, task, survey, and interview. Each phase was designed to guide participants through the process and gather insights on their experiences and interactions.

Tutorial (40-60 min). The study began with a self-paced tutorial session to familiarize participants with the system and its features. For consistency across participants, they were asked to watch a set of pre-recorded tutorial videos (20+ videos totaling 37.5 min) demonstrating how to use StoryEnsemble for various tasks, such as

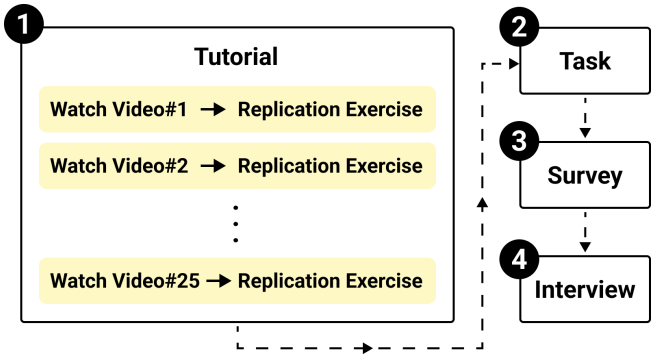


Figure 14: User study procedure

idea generation and iteration within the design process. Participants were instructed to replicate the actions shown in the videos.

Task (20-30 min). After the tutorial, participants were asked to complete a task using the system while engaging in the think-aloud. The task involved selecting a topic of their choice, with a list of pre-selected topics (Table 3) provided for reference, though they were also free to choose a topic outside of the list. We instructed participants to iterate on their designs at least once, in order to gather insights into their experience using the system's iteration features and understand how and why they chose to use them. This session lasted approximately 20–30 minutes, depending on the complexity of the task and participant engagement.

Survey (15 min). Once participants completed the task, they were given a brief survey to gather quantitative feedback on their experience. The survey included questions from the Creativity Support Index [14] (CSI) and System Usability Scale [6] (SUS) questionnaires, as well as questions about the perceived effectiveness of individual features. To provide participants with a familiar frame of reference, we referred to the design process as design thinking in some questions, similar to our approach in the formative study.

Interview (15 min). Finally, participants took part in a semi-structured interview to provide more in-depth qualitative feedback. The interview focused on participants' perceptions of the system's strengths and weaknesses, the ease of use of the AI tools, and how effectively the system supported their design process. As with the survey, we occasionally referred to *design thinking* when probing participants' process, in order to facilitate discussion and align with a commonly recognized design framework. Participants were also asked about any challenges they encountered during the task and their thoughts on how the system could be improved.

5.2 Participants

We recruited 10 participants (age: M = 34.6, SD = 10.9, range = [24, 60]; gender: 5F, 5M) who represented the perspectives of three user groups: 4 students', 5 instructors' (M = 8 years of experience, SD = 8.1, range: [1, 20]), and 6 practitioners' (M = 5.2 years of experience, SD = 2.6, range: [3, 10]). Throughout the paper, as shown in Table 1, we refer to participants as P1–P10, using subscripts to indicate their current profession: student (stud), instructor (inst), or practitioner (prac). For example, $P4_{stud}$ or $P10_{inst}$.

	Profession	Age	Years of Experience
P1	Practitioner	24	3 years
P2	Practitioner	32	5 years
P3	Student	26	1 year
P4	Student	29	2 years + 1 year (Instructor) + 3 years (Practitioner)
P5	Practitioner	31	5 years
P6	Practitioner	37	10 years
P7	Student	28	1 year + 1 year (Instructor)
P8	Instructor	46	12 years + 5 years (Practitioner)
P9	Instructor	60	20 years
P10	Instructor	33	6 years + 1 year (Student)

Table 1: Participant demographics for user evaluation

5.3 Measures

We describes measures used to examine (1) how StoryEnsemble supports exploration and iteration and (2) the usefulness of StoryEnsemble's features.

5.3.1 Creativity Support & Usability. To assess whether StoryEnsemble provides effective creativity support and is usable, we administered two standard questionnaires for measuring them: the Creativity Support Index (CSI) and the System Usability Scale (SUS) surveys.

5.3.2 Exploration Measures. To assess how participants explored the design space while using StoryEnsemble, we defined the following as exploration measures.

1. *Number of persona, problem, solution, storyboard nodes.* This measure is informed by prior work in information retrieval and HCI [69, 73, 84], where the breadth of information exploration is often assessed by counting concept nodes or domain-specific terms created in mind maps or knowledge structures. In our case, this refers to the number of nodes that participants created during their task, corresponding to major stages of the design process: empathizing with the user, defining the problem, ideating solutions, and creating storyboards. By counting persona, problem, solution, and storyboard nodes, we can gauge how thoroughly participants engaged with each stage. Higher counts may indicate more extensive exploration and ideation, reflecting deeper engagement with the system's core tasks.
2. *Usage frequency of features designed for exploration.* To measure the extent to which participants used features designed for exploration, we tracked the frequency of interactions with specific system tools, such as the Start Brainstorming for generating ideas and Generate More option for expanding on ideas from selected idea. These metrics help us understand how often participants relied on StoryEnsemble's generative AI capabilities to enhance their exploration of ideas, as well as their willingness to try different features for broadening the solution space.

5.3.3 Iteration Measures. According to [56], *an iteration* can be defined as "a new version of a design or changes to design." Additionally, we measure the number of times users utilize the propagate changes feature. The measures are as follows.

1. *# of Individual Node Edits.* This measure tracks how often participants edited a node (e.g., persona, problem, solution) and used the Update function ( ) to apply changes to a single connected node. By monitoring the number of edits and updates that propagate to only one related node, we can assess how participants iterated within specific parts of the design process while keeping the scope of changes localized. This can indicate whether participants preferred to make targeted refinements to adjacent nodes rather than making broader, system-wide adjustments across multiple connected nodes.
2. *# of Forward Propagate Edits.* This measure counts the number of times participants edited a node and chose to Forward Propagate ( ) changes to subsequent nodes in the workflow (e.g., updating a problem node and propagating changes to connected solution nodes). The number of nodes updated by forward propagations was also tracked. By tracking forward propagation, we can assess how frequently participants recognized that changes in one stage should affect subsequent stages, a hallmark of effective iterative process.
3. *# of Back Propagate Edits.* This measure captures how often participants edited a node and chose to Back Propagate ( ) changes to preceding nodes (e.g., modifying a solution node and propagating changes back to the problem node). The number of nodes updated by backward propagations was also tracked. This metric helps us understand how participants handled iterative feedback loops, ensuring that updates made at later stages of the design process were reflected in earlier stages. It highlights the system's ability to support non-linear design processes where iteration occurs across multiple stages.

5.3.4 Perceived Utility Measures. To evaluate the perceived utility of StoryEnsemble's features, we analyzed responses from the post-study survey and interviews. This included participants' agreement ratings (1: Strongly Disagree, 5: Strongly Agree) to statements like, 'Being able to propagate changes down ↓ to dependent nodes is useful for iterating on ideas' and questions during interview to elicit further details on their assessment of the features.

6 Results

We present an analysis of survey responses, interviews, and participant interactions with StoryEnsemble,[2] detailing their assessment of the system, how they utilized its features, the perceived value of each, and the diverse ways StoryEnsemble can enhance learning, teaching, and the practice of exploratory, iterative design process.

Creativity Support & Usability. Participants rated StoryEnsemble as providing high creativity support [14]. As shown in Table 2, the

[2]Due to time constraints, one participant (P6) was unable to complete the task, reducing the system usage analysis to N=9. All other analyses are based on N=10.

Table 2: Creativity Support Index (CSI) Results (N=10). The highest value is shown in bold, and the second highest in underline. Following the guidelines by the authors of CSI [14], the *Collaboration* factor was included in the paired-factor comparison questions despite it being out of scope in this study as the score reflects its perceived importance to users and could thereby be useful for future reference.

Factor	Av.g. Score (SD)	Avg. Factor Count
Enjoyment ("*I enjoyed using the tool*")	**18** (2.1)	1.8
Exploration ("*It allowed me to track different ideas*")	17.8 (2.5)	**5.1**
Expressiveness ("*It allowed me to be very expressive*")	15.6 (4.6)	3.9
Immersion ("*The tool was engaging*")	15.3 (4.4)	1.3
Results Worth Effort ("*I was satisfied with what I got*")	17.6 (2.5)	**5.1**
Collaboration ("*I could collaborate with others*")	**N/A**	2.8
Overall CSI Score	72.0 (15.7)	

participants found StoryEnsemble to provide the most support in terms of Enjoyment and Exploration. On the other hand, the CSI factor count revealed that they regard 'getting satisfactory results' (Results Worth Effort) and 'being able to explore diverse ideas' (Exploration) as the important factors when using StoryEnsemble for iterative design process, followed by 'being able to be expressive' (Expressiveness). System Usability Scale (M = 84.8, SD = 10.8), coupled with self-reported responses, showed that participants found StoryEnsemble extremely usable (*Excellent* according to [6]) and passing the cutoff score (68) for production by a large margin. $P10_{inst}$ echoed this, saying it is "intuitive" and that he can see "students or designers learning this in a few minutes."

6.1 How Does StoryEnsemble Support Exploration and Iteration?

6.1.1 Exploration. Participants generally found exploration of ideas easy in StoryEnsemble ('it was easy to explore ideas': 5 Strongly Agree, 3 Agree, 1 Neutral, 1 Disagree). On average, participants created 16.$\overline{3}$ ideas (PERSONA: 4.$\overline{4}$; PROBLEM: 5.$\overline{2}$; SOLUTION: 5.$\overline{1}$; STORYBOARD: 1.$\overline{5}$). Specifically, they all agreed that 'being able to rapidly generate ideas using AI was useful for exploring ideas' (7 Strongly Agree, 3 Agree). $P6_{prac}$ said: "I think it's extremely useful, because it generates ideas quickly and it's also easy to edit. The Design Thinking process is currently very manual and in need of being digitised, so what you're doing is great."

Participants also found the AI-generated ideas from Start Brainstorming for the persona, problem, solution, and storyboard nodes very useful (4 Strongly Agree, 6 Agree) and mostly relevant (2 Strongly Agree, 6 Agree, 1 Neutral, 1 Disagree). $P1_{prac}$ appreciated how AI provided perspectives outside his experience. He shared that while working on a wedding app, he struggled to see things from a couple's point of view, as he was not married. StoryEnsemble helped him empathize with personas and better understand their needs. $P4_{stud}$ recognized that exploring diverse ideas could help combat "design fixation," a common issue in traditional design processes.

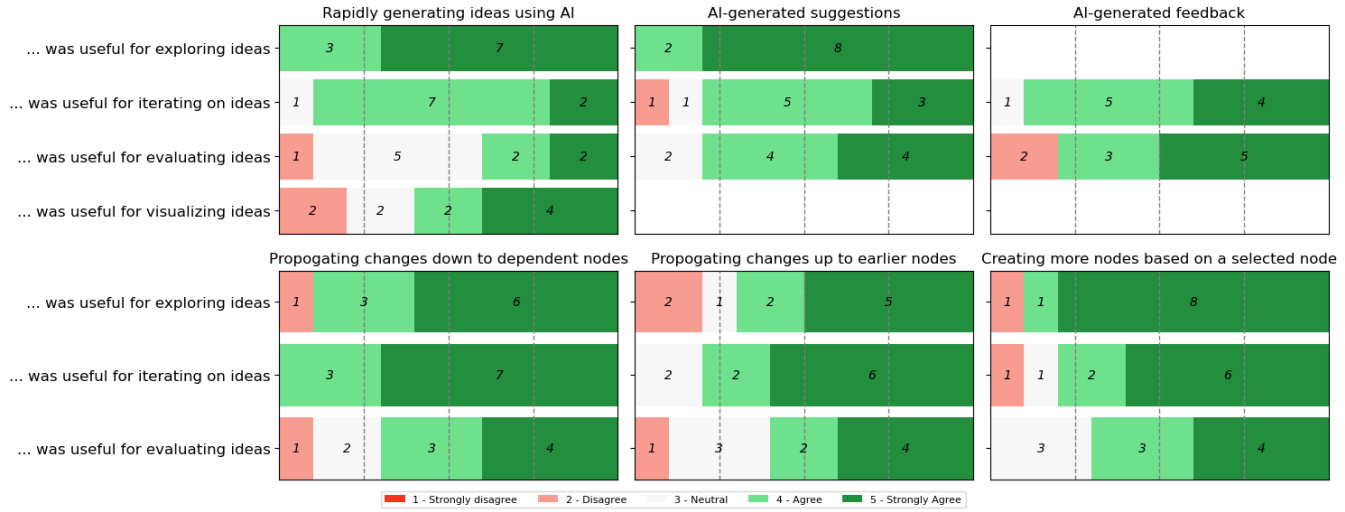
As shown in Fig. 15, they also found AI-generated suggestions extremely useful (5 Strongly Agree, 5 Agree) and for exploring ideas (8 Strongly Agree, 2 Agree). On average, participants generated 22.$\overline{6}$ suggestions during their exploration. Most participants found creating additional nodes using the Generate More feature extremely useful for exploration (8 Strongly Agree, 1 Agree, 1 Disagree). Four participants started with a single persona and then expanded their persona set using Generate More instead of generating the full set of ideas from the start. $P4_{stud}$, who selected Disagree, expressed concern about over-reliance on AI for generating synthetic user data. She preferred creating the initial persona set with real data and then using AI to fill in gaps, validating these insights with target users through further research.

6.1.2 Iteration. Many participants generally found it easy to iterate on ideas (4 Strongly Agree, 4 Agree, 1 Neutral, 1 Disagree). On average, participants iterated on each node at least twice for all nodes except storyboard (PERSONA: 4.$\overline{2}$; PROBLEM: 2.$\overline{4}$; SOLUTION: 2.$\overline{1}$; STORYBOARD: 0.$\overline{5}$) by manually editing (2.$\overline{1}$), revising with AI (2.$\overline{6}$), and using feedback to trigger update of the node (0.$\overline{3}$).

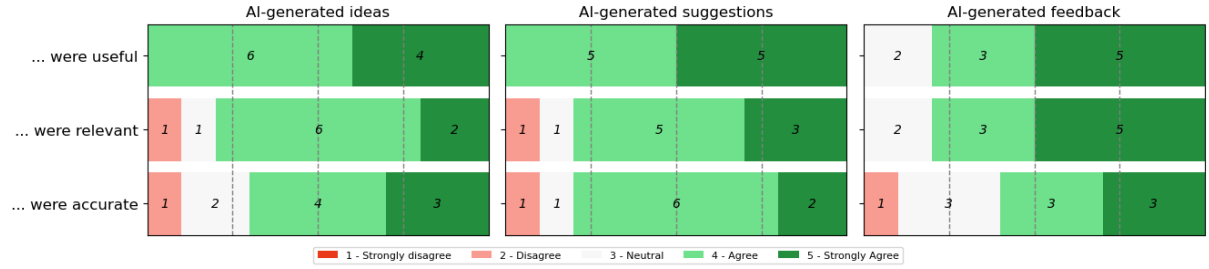
Participants appreciated how the AI-generated feedback (shown in Fig. 9) nudged them to iterate. $P2_{prac}$ said: "I liked that it made me incorporate feedback and revise." Overall, they found AI-generated feedback very useful in general (5 Strongly Agree, 3 Agree, 2 Neutral) and for iterating on ideas (4 Strongly Agree, 5 Agree, 1 Neutral). In addition to the quality of the feedback, the ease with which they could update the node by incorporating the feedback from the list and trigger update of the node seemed to help them effortlessly transition into iteration. $P1_{prac}$ said: "from feedback, it makes me go, 'Oh yeah, I forgot about this.' And then I can just iterate based on that feedback."

Overall, participants liked the support StoryEnsemble provides to enable flexible iteration, with many participants ($P2_{prac}$, $P7\text{-}8_{stud,inst}$, $P10_{inst}$) quoting both the forward propagation and backward propagation as their favorite features. As shown in Fig. 15a, participants found both propagation mechanisms (Fig. 12) highly useful for iteration (Forward Prop: 7 Strongly Agree, 3 Agree; Back Prop: 6 Strongly Agree, 2 Agree, 2 Neutral). Most participants (7/9) used 'propagate changes' at least twice (Forward Prop: M=3.$\overline{2}$; Backward Prop: M=1.$\overline{2}$). P2 specifically emphasized the value of being able to iterate acr oss different stages:

> "I totally like the top down approach of iteration and also bottom up approach where we propagate from one level to another. ... **I could iterate top, down, bottom up to any levels**. So it felt better than design



(a) Perceived usefulness of each feature for ideation, iteration, evaluation, and visualization of ideas.



(b) Perceived quality of AI-generated ideas, suggestions, and feedback

Figure 15: General perception of the usefulness of StoryEnsemble features and the quality of AI-generated ideas, suggestions, and feedback.

thinking in a way because, **I felt like I could iterate on all different levels**."

6.2 How Can StoryEnsemble Be Used in Design Process?

We present three types of workflows (**W**s) observed in our analysis that reflect a spectrum of design approaches, ranging from mostly user-driven to heavily AI-assisted.

W1. Bottom-Up (User-driven) Workflow (Fig. 18). $P1_{prac}$ followed a bottom-up strategy, generating and iterating within each stage before moving on to the next. In the Empathize Stage, they manually created an initial persona (Persona 1), then used StoryEnsemble's AI support to generate four additional personas (Persona 2-5). In the Define Stage, P1 generated two problem nodes (Problem 1, 2) connected to Persona 2 and Persona 3. Next, P1 generated 3 solutions (Solution 1-3) based on Problem 1 to explore the Ideate Stage. P1 then created a storyboard based on Solution 1, Problem 1, Persona 2, and Persona 3. Finally, P1 used AI-driven revisions to edit the problem framing within the storyboard and regenerated the problem image to better reflect the design intent in the Prototype Stage. This workflow illustrates how users can actively guide the process while using AI to expand and refine ideas.

W2. Top-Down (AI-driven) Workflow (Fig. 19). $P2_{prac}$ adopted a more AI-driven, top-down approach, using AI to generate a comprehensive design direction upfront. After manually drafting Persona 1, P2 used AI to revise it and then prompted StoryEnsemble to generate the remaining elements across all design stages in one go: 4 problems (Problem 1-4), 4 solutions (Solution 1-4) and a storyboard linked to Solution 2, Problem 3, and Persona 1 in a single generation from Persona 1. Then P2 iterated on Solution 2 using StoryEnsemble's "backward propagation" feature, which cascaded updates to the linked Problem 2 and Persona 1. This workflow demonstrates StoryEnsemble's ability to support nonlinear iteration and rapid exploration from a high-level concept.

W3. Hybrid Workflow (Fig. 17). $P8_{inst}$ used a hybrid approach, incorporating both designer input and AI suggestions. In the Empathize Stage, P8 generated 3 personas (Persona 1-3) and revised Persona 2 and 3 using designer instructions and AI-recommended prompts. P8 also engaged with AI-generated feedback to further refine Persona 3. In the Define Stage, P8 generated 3 problems (Problem 1-3) based on Persona 3, revising Problem 1 and 2 using AI-recommended prompts before proceeding to brainstorm solutions for Problem 2. This workflow reflects the StoryEnsemble's

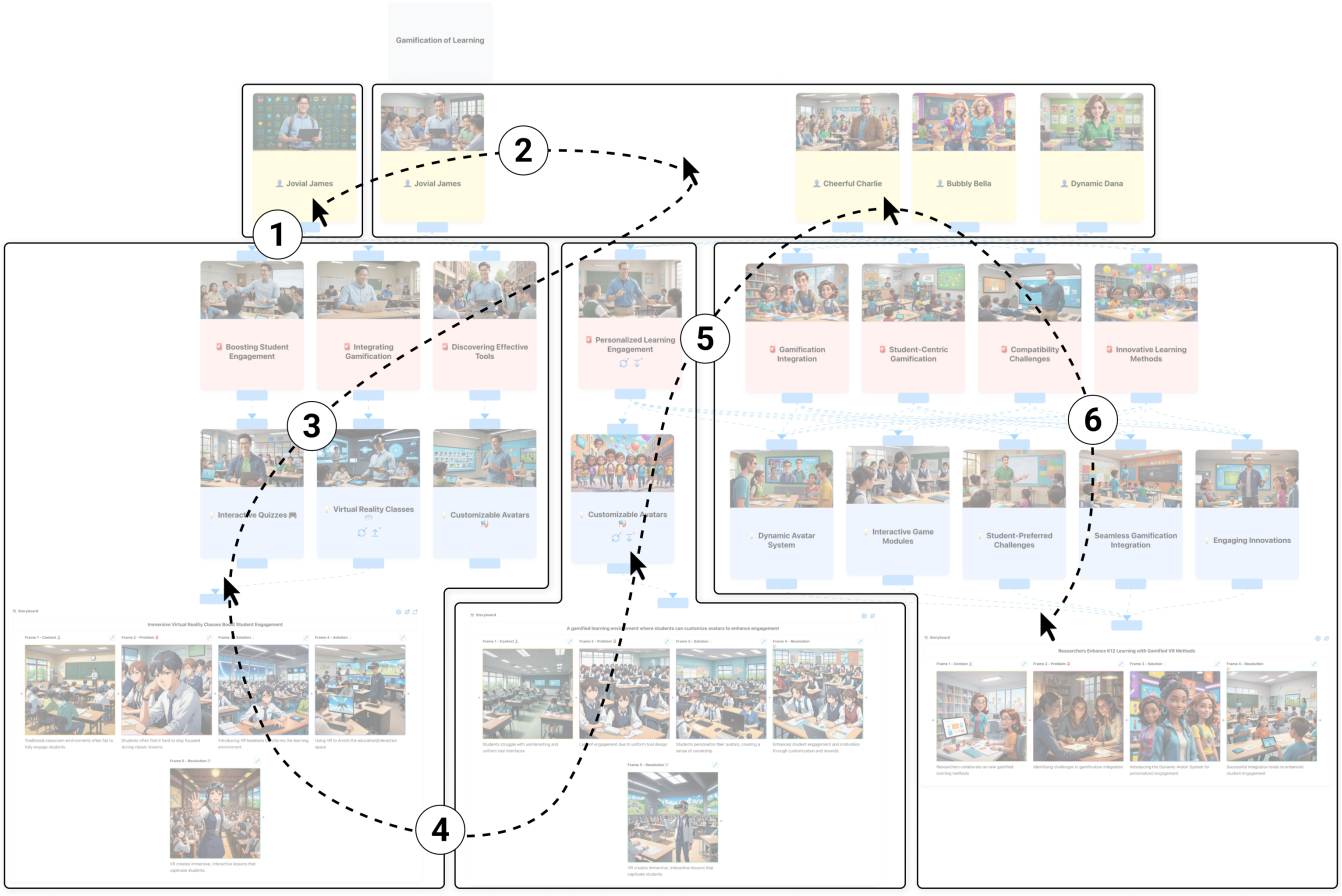


Figure 16: The workspace of $P7_{stud}$ after completing the task. P7 followed a hybrid approach that combined manual authoring with AI assistance. (1) P7 manually created an initial persona, then (2) used GENERATE MORE to expand the persona space. (3) These personas were used to generate corresponding problems, solutions, and an initial storyboard. (4–5) P7 then used backpropagation from the storyboard to generate and revise earlier artifacts—including persona nodes—and (6) generated up to a storyboard. This process illustrates how StoryEnsemble enables dynamic exploration and iteration.

flexibility in supporting co-creation and dynamic iteration between user input and AI contributions during design processes.

These workflows illustrate StoryEnsemble's key design principle of supporting flexible, interconnected design exploration, where users can seamlessly move between manual input and AI-assisted generation across different design stages.

6.3 What Are Potential Applications, Opportunities, and Challenges?

In addition to its role in supporting exploration and iteration, StoryEnsemble opens up several potential applications and addresses opportunities and challenges in the broader use of iterative design process. Through our study, we identified key use cases in teaching, learning, and professional design, while also uncovering concerns regarding AI's role in the process. Most participants rated StoryEnsemble as extremely useful for *teaching* and *using* design thinking (Teaching: 5 Extremely Useful, 3 Very Useful, 2 Moderately; Using: 7 Extremely Useful, 1 Very Useful, 2 Moderately). In contrast, the ratings for *learning* design thinking were more evenly distributed, ranging from moderate to highly useful (Learning: 3 Extremely Useful, 1 Very Useful, 2 Moderately). This suggests that the tool is particularly valuable for practitioners and instructors, while its benefits for learners are still significant but more varied. Below, we explore the main themes that emerged from the survey and interviews.

6.3.1 Leveraging StoryEnsemble for Teaching and Demonstrating Design Thinking. Many participants emphasized the value of StoryEnsemble for teaching design thinking in academic settings. The system's ability to quickly generate artifacts like personas and solutions, combined with its iterative features, was seen as an effective way to guide students through the design thinking process in a structured, interactive manner.

One of the most prominent themes that emerged from our study was StoryEnsemble's potential as a teaching tool. Several participants, especially instructors, noted that StoryEnsemble could significantly enhance the way design thinking is taught in classrooms by

Figure 17: The workspace of $P8_{inst}$ after completing the task.

making the process more interactive and accessible. For instance, $P2_{prac}$ said: "this could help people practice design thinking in the way it's intended to be. It makes more sense to do it this way."

$P8_{inst}$, an experienced instructor, emphasized how StoryEnsemble could serve as a practical demonstration tool, allowing students to see an entire design thinking process in action. He described how it would help students develop a mental model of what a complete project looks like, giving them a clear understanding of how each stage—persona creation, problem definition, solution ideation, and prototyping—fits together:

> "If I had this tool, it would be really neat to demo a toy project... I could **use the AI to show them how to iterate** through a persona, a problem statement, and a solution."

Specifically, he was excited about StoryEnsemble's propagation mechanisms and suggested that the tool could be perfect for demonstrating it:

> "**My favorite feature is the propagate downward and propagate upward** [features], because **this is what we try to teach to students in the design thinking process**... but you already have a mechanism to allow that by propagating upwards or downwards. So I really like that."

Another important advantage for teaching is the ability to use real-time feedback and AI-generated suggestions to encourage students to reflect on their decisions. $P2_{prac}$ noted how this feature could deepen students' engagement:

> "The fact that you can go back and forth and revise things... it would make them think more about the problem and the users."

However, some instructors expressed **concerns about students potentially over-relying on AI** and skipping critical steps in the design process. $P4_{stud}$ emphasized the importance of hands-on experience, saying that "students should have the experience of actually doing it instead of just generating everything randomly." Instructors, who identified similar potential issues, suggested controlling when and how students use AI in the classroom. $P4_{stud}$ and $P10_{inst}$ proposed introducing AI later in the learning process, after students have had a chance to work through initial tasks manually. This could help ensure that students develop essential design skills before using AI to iterate or refine their work. Additionally, $P5_{prac}$ suggested that future versions of StoryEnsemble could feature an admin panel, allowing instructors to tailor AI's involvement to specific tasks and providing them with the flexibility to manage how AI is integrated into different phases of design thinking.

Overall, participants saw StoryEnsemble as a powerful teaching tool, with the potential to make design thinking more accessible, interactive, and engaging for students, while acknowledging the need to maintain a balance between AI assistance and hands-on practice.

6.3.2 Streamlining Professional Workflows with AI Assistance. In addition to its potential for education, StoryEnsemble was also seen as a valuable tool for improving design workflows in professional environments. Several participants, particularly those with experience in design practice, noted that StoryEnsemble could help streamline the ideation and iteration process, making it more efficient for small teams or projects with tight deadlines.

$P6_{prac}$ reflected on how StoryEnsemble could speed up the design process in situations where resources are limited, saying "it would **speed up the process** massively, especially when the design team is small." $P5_{prac}$ added that StoryEnsemble could drastically reduce the amount of time spent on brainstorming and iteration, allowing teams to focus on more strategic aspects of design. Reflecting on the past design workshop with 8 members of his team, he said that StoryEnsemble could have easily given "50% [time] reduction" for each person, essentially saving "40 hours" and allowing that time to be "used for something else."

In a similar vein, $P2_{prac}$, who learned design thinking through a self-paced, 1-week online course, said that given the low attention span of people, this tool would be amazing for students like her, saying "if I can learn the same thing in 90 minutes what I learnt in a week, I'll choose the 90 minutes." She went onto add that corporations that invest money and time to teach their employees about design thinking would be thrilled to purchase this tool for training purposes.

6.3.3 Balancing AI Assistance and Hands-On Design Practice. While participants praised StoryEnsemble's ability to streamline both teaching and professional workflows, its potential to help students "get unstuck" ($P4_{stud}$) and access useful "feedback" ($P8_{inst}$), they also raised important concerns about the potential for over-reliance on AI, particularly in educational contexts. Several instructors worried that students might become too dependent on AI-generated outputs, bypassing essential learning experiences such as manual iteration and critical thinking.

P4, a student with prior experience as both an instructor and design practitioner, stressed the importance of hands-on learning, saying "students should have the experience of actually doing it instead of just generating everything randomly." $P8_{inst}$ also voiced concern over students relying too heavily on AI-generated assumptions. He said: "Designing on assumptions is not a good thing... I would tell students to be very cautious with the assumptions the AI engine makes."

To address these concerns, participants suggested ways to control AI's role and the extent of it in the design process. For instance, $P10_{inst}$ recommended introducing AI only after students have completed certain tasks manually, allowing them to first engage with the design process before using AI to support their iterations. $P5_{prac}$ proposed an interesting idea, suggesting that there could be a panel that allows instructors to precisely control the level of AI engagement by specifying which node types can use AI capabilities. He said:

> "**A toggle button to control AI engagement could be useful**... depending on the workshop, AI could be used only for certain tasks."

He explained that instructors could grant permission to use AI once students complete specific tasks in earlier stages. For example, after conducting interviews and manually creating initial personas, students could then use AI to generate additional personas grounded in the data from their interviews—maintaining a balance between hands-on practice and AI-assisted exploration.

Additionally, participants suggested incorporating validation mechanisms into the system. $P7_{stud}$ proposed that after generating AI-driven ideas, StoryEnsemble could prompt users to accept or reject them, encouraging designers to critically assess the outputs rather than blindly accepting them: "The system could nudge users to accept or reject AI-generated ideas, forcing them to check and validate the data."

7 Discussion

7.1 Summary

In this work, we explored how generative AI can support iterative design processes by developing StoryEnsemble, a system that enables users to easily explore and iterate across multiple stages of design thinking. While prior systems such as DynEx [44] and IdeaSynth [52] also support divergent and convergent workflows through structured UI generation or research idea synthesis, they focused primarily on generation within specific stages. In contrast, our goal was to address a broader challenge of fragmentation across design stages by introducing mechanisms such as dependency tracking, cascading updates, and backpropagation.

Our evaluation showed that StoryEnsemble effectively supports dynamic exploration and iteration through a combination of AI-generated suggestions, revision prompts, and propagation mechanisms. Iteration was supported through direct editing, feedback-triggered updates, and AI-assisted revisions, with participants frequently using forward and backward propagation to cascade changes across connected nodes. These features enabled participants to revise earlier ideas based on later-stage insights and vice versa, helping them break out of linear flows and engage in multi-level iteration. They also adopted a range of workflows—bottom-up, top-down, and hybrid—demonstrating the flexibility of the system in supporting both structured and improvisational design strategies.

These findings highlight the importance of supporting non-linear, iterative processes. Rather than enforcing a fixed, stage-by-stage progression, design systems should allow users to revisit any stage, propagate changes efficiently, and explore multiple paths with minimal friction. As demonstrated by StoryEnsemble, supporting tightly linked representations, lightweight revisions, and multiple entry points offer a promising model for supporting dynamic exploration and iteration in multi-stage creative processes.

7.2 Design Implications

Traditional design processes, such as design thinking, often struggle with complexity and the inherent messiness of interdependent tasks, where changes in one stage affect others, making it difficult to manage under tight deadlines or resource constraints. StoryEnsemble addresses these challenges by visualizing and managing these interdependencies through a node-link interface and allowing bi-directional change propagation to observe the design changes in real-time. As users are freed from manually updating every step, StoryEnsemble facilitates effortless iteration, encouraging a more dynamic and integrated approach to the design process.

Beyond the system, our work offers comprehensive perspectives from students, instructors, and practitioners regarding the potential of incorporating generative AI in the design process. $P9_{inst}$ shared how she has already accepted the fact that she cannot prevent students from using AI and explained that she started asking students to simply describe how they have used it for their work. If co-creation becomes the norm in the near future, our work will contribute rich insights on the system design opportunities and tensions for this paradigm shift in the design process.

The mechanisms we developed to manage interconnected design artifacts—such as dependency tracking, cascading updates, and bidirectional propagation—may generalize to other domains with similarly interdependent structures. For example, in software development, updating one module may require adjustments across related components; in content creation, narrative changes in a script may cascade to visuals, audio, or pacing; in education, learning objectives often shape assessments and instructional materials in a co-evolving process; and in video authoring, changes to a scene or storyline may affect timing, transitions, and supporting media assets. Supporting such interdependencies requires tools that can maintain coherence while allowing flexible iteration—a challenge our approach addresses and one that future systems across these domains may benefit from.

7.3 Ethical and Pedagogical Implications

One important consideration raised in this study relates to the ethical and pedagogical concerns regarding AI over-reliance, particularly in educational contexts. Several participants highlighted the risks of students relying too heavily on AI-generated content, potentially bypassing critical design skills that can only be developed through hands-on experience and critical reflection.

As noted by $P4_{stud}$ and $P8_{inst}$, a balance must be maintained between leveraging AI for efficiency and ensuring students engage with foundational design principles. Without appropriate guidance, students may mistakenly equate the use of AI-generated suggestions with mastering the design process, which may lead to a superficial understanding of key concepts.

To address these concerns, participants recommended diverse ideas such as introducing AI later in the design process, providing a control to regulate AI's availability, and requiring users to actively reflect on AI-generated outputs via validation prompts. These insights can be useful as we design future systems and tools powered by AI and as AI becomes more integrated into educational practices.

7.4 Limitations and Future Work

Study Design. While our exploratory study provided broad insights into the potential of StoryEnsemble, more controlled, comparative studies are needed to answer specific questions such as whether AI-enhanced design leads to more creative or diverse solutions than traditional methods. These studies could also explore the role of StoryEnsemble in supporting various design stages in greater depth.

Long-term Deployment Study. A long-term deployment study could offer further insights into how generative AI can support sustained design projects and team-based workflows. This would help to evaluate StoryEnsemble's effectiveness over time and assess its role in supporting co-creation with AI in different contexts.

Extending System: Adding More Tools. Future development of StoryEnsemble could expand its utility by incorporating additional tools. As Hehn et al enumerated in their work, there are diverse techniques and methods in design [33]. For example, for need-finding, designers can choose from several methods such as customer journey, empathy map, and 5 whys. For prototyping, in addition to storyboard, designers can choose paper prototype, digital wireframe, as well as high fidelity prototype. In fact, participants such as P8 suggested adding empathy maps or journey maps, which are integral to many design workflows. There is also potential for supporting collaboration and enhancing system's feedback mechanisms by synthesizing user feedback across stages, providing users with a more comprehensive view of insights that emerge from design research or user studies.

Extending System: Projecting Possibilities through Generative Simulation. StoryEnsemble's generative capabilities open up exciting possibilities for extending the system into simulation-based research (akin to [49, 50]). Future iterations could leverage the tool's persona and scenario generation to help researchers explore potential user impacts across diverse contexts. For instance, policy researchers could use the system to simulate how proposed programs might affect different demographic groups, generating synthetic user scenarios to identify previously overlooked perspectives or potential implementation challenges. Similarly, design teams could use the tool to proactively explore user experiences across varied socioeconomic, cultural, or accessibility contexts, expanding traditional research methodologies by rapidly generating and iterating through multiple potential user scenarios. This approach would transform StoryEnsemble from a design exploration tool into a powerful platform for scenario planning and social impact assessment, enabling more comprehensive and imaginative approaches to understanding complex user ecosystems.

Preventing Overreliance on AI. Our study raised concerns about students over-relying on AI-generated content, especially in educational settings. One solution proposed by participants is to regulate AI engagement based on the task or stage of the process. This could be managed by providing clear guidelines on when to use AI and ensuring that users have the opportunity to engage manually before incorporating AI-generated suggestions. We could also explore techniques such as *frictions*, interventions that force users to cognitively engage with the AI-generated artifact [37]. Our intention is not to replace these critical aspects of the design process but to view AI as a complementary tool, addressing limitations in current design thinking processes such as the costly nature of getting feedback and generating artifacts such as storyboard to communicate ideas [36].

Identifying Scenarios Where Generative AI is Most Effective. While highlighting the potential of AI in HCI, Schmidt et al. also emphasized the need to define when AI should be used [60]. In our evaluation study, we explored, to some extent, how StoryEnsemble and generative AI could be useful or preferred. For example, participants generally felt AI would be helpful for teaching and using design thinking, while they were mixed about using it for learning design thinking. Within the context of learning design thinking, instructors had clear views on when AI would be beneficial. For instance, they saw value in using AI to demonstrate the full design thinking workflow with a simplified example, but expressed concern about novice students relying on AI independently, potentially leading to misunderstandings of core principles—such as the importance of engaging real users in the design process. Further research is needed to understand the nuances and reveal when AI is acceptable or even preferred. For example, Salminen et al. found that using AI to generate more personas can improve the representation of demographically diverse populations and benefit media-based projects where having demographically diverse representation is essential [58].

8 Conclusion

In this work, we explored how generative AI can address key challenges in the design thinking process, particularly in supporting exploration, iteration, and feedback collection. While design thinking emphasizes iterative, user-centered problem-solving, practical constraints—such as time constraints and difficulty of engaging users for feedback—often hinder its full application. We introduced StoryEnsemble, a system that integrates generative AI to streamline the design thinking workflow, enabling rapid exploration of ideas, flexible iteration across interconnected stages, and efficient feedback collection. Our formative study identified key challenges

faced by students, instructors, and practitioners, and our evaluation demonstrated that StoryEnsemble supports rapid exploration and flexible iteration of ideas. This work contributes a novel approach to AI-integrated design thinking, offering insights into the opportunities and challenges of incorporating AI into creative workflows, while emphasizing the importance of recognizing AI's limitations and ethical considerations.

References

[1] Shm Garanganao Almeda, JD Zamfirescu-Pereira, Kyu Won Kim, Pradeep Mani Rathnam, and Bjoern Hartmann. 2024. Prompting for Discovery: Flexible Sense-Making for AI Art-Making with Dreamsheets. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[2] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*. 413–425.

[3] Victor Nikhil Antony and Chien-Ming Huang. 2024. ID. 8: Co-Creating visual stories with Generative AI. *ACM Transactions on Interactive Intelligent Systems* 14, 3 (2024), 1–29.

[4] L Bruce Archer. 1979. Whatever became of design methodology. *Design studies* 1, 1 (1979), 17–18.

[5] Estefania Arteaga, Robbert Biesbroek, Johanna Nalau, and Michael Howes. 2024. Across the Great Divide: A Systematic Literature Review to Address the Gap Between Theory and Practice. *SAGE Open* 14, 1 (2024), 21582440241228019.

[6] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[7] Sara L Beckman and Michael Barry. 2007. Innovation as a learning process: Embedding design thinking. *California management review* 50, 1 (2007), 25–56.

[8] Karim Benharrak, Tim Zindulka, Florian Lehmann, Hendrik Heuer, and Daniel Buschek. 2024. Writer-defined AI personas for on-demand feedback generation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.

[9] Volker Bilgram and Felix Laarmann. 2023. Accelerating innovation with generative AI: AI-augmented digital prototyping and innovation methods. *IEEE Engineering Management Review* 51, 2 (2023), 18–25.

[10] 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*. 1–14.

[11] Yining Cao, Yiyi Huang, Anh Truong, Hijung Valentina Shin, and Haijun Xia. 2025. Compositional Structures as Substrates for Human-AI Co-creation Environment: A Design Approach and A Case Study. https://api.semanticscholar.org/CorpusID:276813065

[12] Yining Cao, Yiyi Huang, Anh Truong, Hijung Valentina Shin, and Haijun Xia. 2025. Compositional Structures as Substrates for Human-AI Co-creation Environment: A Design Approach and A Case Study. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–25.

[13] Longfei Chen, Shengxin Li, Ziang Li, and Quan Li. 2025. DancingBoard: Streamlining the Creation of Motion Comics to Enhance Narratives. In *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*. Association for Computing Machinery, New York, NY, USA, 477–503. https://doi.org/10.1145/3708359.3712167

[14] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25.

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

[16] Yoonseo Choi, Eun Jeong Kang, Seulgi Choi, Min Kyung Lee, and Juho Kim. 2024. Proxona: Leveraging LLM-Driven Personas to Enhance Creators' Understanding of Their Audience. *arXiv preprint arXiv:2408.10937* (2024).

[17] John Joon Young Chung and Eytan Adar. 2023. PromptPaint: Steering Text-to-Image Generation Through Paint Medium-like Interactions. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 6, 17 pages. https://doi.org/10.1145/3586183.3606777

[18] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching stories with generative pretrained language models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.

[19] British Design Council. 2023. The Double Diamond. https://www.designcouncil.org.uk/our-resources/the-double-diamond/. Retrieved 11 November 2024.

[20] Rikke Friis Dam. 2020. 5 Stages in the design thinking process. https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process?ref=weekly.ui-patterns.com.

[21] Ellen Yi-Luen Do and Mark D Gross. 2001. Thinking with diagrams in architectural design. *Artificial intelligence review* 15 (2001), 135–149.

[22] Dietrich Dörner. 1999. Approaching design thinking research. *Design Studies* 20, 5 (1999), 407–415.

[23] Steven P Dow, Kate Heddleston, and Scott R Klemmer. 2009. The efficacy of prototyping under time constraints. In *Proceedings of the seventh ACM conference on Creativity and cognition*. 165–174.

[24] Peitong Duan, Jeremy Warner, Yang Li, and Bjoern Hartmann. 2024. Generating Automatic Feedback on UI Mockups with Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[25] Jane L. E, Yu-Chun Grace Yen, Isabelle Yan Pan, Grace Lin, Mingyi Li, Hyoungwook Jin, Mengyi Chen, Haijun Xia, and Steven P. Dow. 2024. When to Give Feedback: Exploring Tradeoffs in the Timing of Design Feedback. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 292–310. https://doi.org/10.1145/3635636.3656183

[26] Kiev Gama, George Valença, Pedro Alessio, Rafael Formiga, André Neves, and Nycolas Lacerda. 2023. The developers' design thinking toolbox in hackathons: a study on the recurring design methods in software development marathons. *International Journal of Human–Computer Interaction* 39, 12 (2023), 2269–2291.

[27] Elizabeth Goodman, Erik Stolterman, and Ron Wakkary. 2011. Understanding interaction design practices. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1061–1070.

[28] April Greenwood, Benjamin Lauren, Jessica Knott, and Dànielle Nicole DeVoss. 2019. Dissensus, resistance, and ideology: Design thinking as a rhetorical methodology. *Journal of Business and Technical Communication* 33, 4 (2019), 400–424. https://doi.org/10.1177/1050651919854063

[29] Tom Greever. 2015. *Articulating design decisions: Communicate with stakeholders, keep your sanity, and deliver the best user experience.* " O'Reilly Media, Inc.".

[30] Perttu Hämäläinen, Mikke Tavast, and Anton Kunnari. 2023. Evaluating large language models in generating synthetic hci research data: a case study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[31] 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*. 91–100.

[32] Rui He, Huaxin Wei, and Ying Cao. 2024. An Interactive System for Supporting Creative Exploration of Cinematic Composition Designs. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–15.

[33] Jennifer Hehn, Falk Uebernickel, and Matthias Herterich. 2018. Design thinking methods for service innovation-A delphi study. (2018). https://aisel.aisnet.org/pacis2018/126/

[34] Yu-Kai Hung, Yun-Chien Huang, Ting-Yu Su, Yen-Ting Lin, Lung-Pan Cheng, Bryan Wang, and Shao-Hua Sun. 2024. SimTube: Generating Simulated Video Comments through Multimodal AI and User Personas. *ArXiv* abs/2411.09577 (2024). https://api.semanticscholar.org/CorpusID:274023042

[35] Peiling Jiang, Jude Rayan, Steven P Dow, and Haijun Xia. 2023. Graphologue: Exploring large language model responses with interactive diagrams. In *Proceedings of the 36th annual ACM symposium on user interface software and technology*. 1–20.

[36] Hyunggu Jung, Woosuk Seo, Seokwoo Song, and Sungmin Na. 2023. Toward Value Scenario Generation Through Large Language Models. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 212–220.

[37] Majeed Kazemitabaar, Oliver Huang, Sangho Suh, Austin Z Henley, and Tovi Grossman. 2025. Exploring the design space of cognitive engagement techniques with ai-generated code for enhanced learning. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 695–714.

[38] Jeongyeon Kim, Sangho Suh, Lydia B Chilton, and Haijun Xia. 2023. Metaphorian: leveraging large language models to support extended metaphor creation for science writing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 115–135.

[39] David Kirsh. 2010. Thinking with external representations. *AI & society* 25 (2010), 441–454.

[40] Maaike Kleinsmann, Rianne Valkenburg, and Janneke Sluijs. 2017. Capturing the value of design thinking in different innovation practices. *International Journal of Design* 11, 2 (2017), 25–40.

[41] Mina Lee, Percy Liang, and Qian Yang. 2022. Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–19.

[42] David Chuan-En Lin, Hyeonsu B Kang, Nikolas Martelaro, Aniket Kittur, Yan-Ying Chen, and Matthew K. Hong. 2025. Inkspire: Supporting Design Exploration with Generative AI through Analogical Sketching. *ArXiv* abs/2501.18588 (2025). https://api.semanticscholar.org/CorpusID:275993583

[43] 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*. 1955–1977.

[44] Jenny GuangZhen Ma, Karthik Sreedhar, Vivian Liu, Pedro A Perez, Sitong Wang, Riya Sahni, and Lydia B Chilton. 2025. Dynex: Dynamic code synthesis with structured design exploration for accelerated exploratory programming. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–27.

[45] Monica Maceli, Nancy Smith, and Gatha Bhakta. 2024. Incorporating Unanticipated Uses of Generative AI into HCI Education. In *Proceedings of the 6th Annual Symposium on HCI Education*. 1–7.

[46] Damien Masson, Young-Ho Kim, and Fanny Chevalier. 2025. Textoshop: Interactions Inspired by Drawing Software to Facilitate Text Editing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–14.

[47] Damien Masson, Sylvain Malacria, Géry Casiez, and Daniel Vogel. 2024. Directgpt: A direct manipulation interface to interact with large language models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.

[48] Pietro Micheli, Sarah JS Wilner, Sabeen Hussain Bhatti, Matteo Mura, and Michael B Beverland. 2019. Doing design thinking: Conceptual review, synthesis, and research agenda. *Journal of Product innovation management* 36, 2 (2019), 124–148.

[49] Ha Nguyen, Victoria Nguyen, Saríah López-Fierro, Sara Ludovise, and Rossella Santagata. 2024. Simulating Climate Change Discussion with Large Language Models: Considerations for Science Communication at Scale. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*. 28–38.

[50] Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–18.

[51] John Pruitt and Jonathan Grudin. 2003. Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*. 1–15.

[52] Kevin Pu, KJ Kevin Feng, Tovi Grossman, Tom Hope, Bhavana Dalvi Mishra, Matt Latzke, Jonathan Bragg, Joseph Chee Chang, and Pao Siangliulue. 2025. Ideasynth: Iterative research idea development through evolving and composing idea facets with literature-grounded feedback. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–31.

[53] Whitney Quesenbery and Kevin Brooks. 2010. *Storytelling for user experience: Crafting stories for better design*. Rosenfeld Media.

[54] Rim Razzouk and Valerie Shute. 2012. What is design thinking and why is it important? *Review of educational research* 82, 3 (2012), 330–348.

[55] Yvonne Rogers. 2004. New theoretical approaches for HCI. *Annual review of information science and technology* 38, 1 (2004), 87–143.

[56] Emma J Rose and Cody Reimer. 2022. *6. Iteration*. WAC Clearinghouse. https://doi.org/10.37514/TPC-B.2022.1725.2.06

[57] Mary Beth Rosson and John M Carroll. 2007. Scenario-based design. In *The human-computer interaction handbook*. CRC Press, 1067–1086.

[58] Joni Salminen, Soon-Gyo Jung, Lene Nielsen, and Bernard Jansen. 2022. Creating More Personas Improves Representation of Demographically Diverse Populations: Implications Towards Interactive Persona Systems. In *Nordic Human-Computer Interaction Conference*. 1–11.

[59] Joni Salminen, Chang Liu, Wenjing Pian, Jianxing Chi, Essi Häyhänen, and Bernard J Jansen. 2024. Deus Ex Machina and Personas from Large Language Models: Investigating the Composition of AI-Generated Persona Descriptions. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[60] Albrecht Schmidt, Passant Elagroudy, Fiona Draxler, Frauke Kreuter, and Robin Welsch. 2024. Simulating the human in HCD with ChatGPT: Redesigning interaction design with AI. *Interactions* 31, 1 (2024), 24–31.

[61] Michael Schrage. 1999. *Serious play: How the world's best companies simulate to innovate*. Harvard Business Press.

[62] Andreas Schuller, Doris Janssen, Julian Blumenröther, Theresa Maria Probst, Michael Schmidt, and Chandan Kumar. 2024. Generating personas using LLMs and assessing their viability. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[63] Xinyu Shi, Yinghou Wang, Ryan Rossi, and Jian Zhao. 2025. Brickify: Enabling Expressive Design Intent Specification through Direct Manipulation on Design Tokens. *ArXiv* abs/2502.21219 (2025). https://api.semanticscholar.org/CorpusID:276649308

[64] Joongi Shin, Michael A Hedderich, Bartłomiej Jakub Rey, Andrés Lucero, and Antti Oulasvirta. 2024. Understanding Human-AI Workflows for Generating Personas. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 757–781.

[65] Atefeh Shokrizadeh, Boniface Bahati Tadjuidje, Shivam Kumar, Sohan Kamble, and Jinghui Cheng. 2025. Dancing With Chains: Ideating Under Constraints With UIDEC in UI/UX Design. *ArXiv* abs/2501.18748 (2025). https://api.semanticscholar.org/CorpusID:276079770

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

[67] Erik Stolterman. 2008. The nature of design practice and implications for interaction design research. *International Journal of Design* 2, 1 (2008).

[68] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–26.

[69] Sangho Suh, Bryan Min, Srishti Palani, and Haijun Xia. 2023. Sensecape: Enabling multilevel exploration and sensemaking with large language models. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–18.

[70] Kallaya Tantiyaswasdikul. 2020. How design thinking can foster environmental sustainability: Integrating design thinking into circular design guide. In *Proceedings of the 6th International Conference on Industrial and Business Engineering*. 157–162.

[71] Jason Chew Kit Tham. 2022. *Keywords in Design Thinking: A Lexical Primer for Technical Communicators & Designers*. WAC Clearinghouse.

[72] Khai N Truong, Gillian R Hayes, and Gregory D Abowd. 2006. Storyboarding: an empirical determination of best practices and effective guidelines. In *Proceedings of the 6th conference on Designing Interactive systems*. 12–21.

[73] Kelsey Urgo and Jaime Arguello. 2022. Learning assessments in search-as-learning: A survey of prior work and opportunities for future research. *Information Processing & Management* 59, 2 (2022), 102821.

[74] Samangi Wadinambiarachchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.

[75] Bryan Wang, Yuliang Li, Zhaoyang Lv, Haijun Xia, Yan Xu, and Raj Sodhi. 2024. LAVE: LLM-Powered Agent Assistance and Language Augmentation for Video Editing. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (Greenville, SC, USA) *(IUI '24)*. Association for Computing Machinery, New York, NY, USA, 699–714. https://doi.org/10.1145/3640543.3645143

[76] Shun-Yu Wang, Wei-Chung Su, Serena Chen, Ching-Yi Tsai, Marta Misztal, Katherine M Cheng, Alwena Lin, Yu Chen, and Mike Y Chen. 2024. RoomDreaming: Generative-AI Approach to Facilitating Iterative, Preliminary Interior Design Exploration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[77] Wenjing Wang and Baixi Xing. 2024. Research on the Integration and Application of Design Thinking and Large Language Models in the Innovation Design of Fintech Products. In *Design Studies and Intelligence Engineering*. IOS Press, 673–682.

[78] Wen-Fan Wang, Chien-Ting Lu, Nil Ponsa Campanya, Bing-Yu Chen, and Mike Y. Chen. 2025. AIdeation: Designing a Human-AI Collaborative Ideation System for Concept Designers. *ArXiv* abs/2502.14747 (2025). https://api.semanticscholar.org/CorpusID:276482514

[79] Wei Xiang, Hanfei Zhu, Suqi Lou, Xinli Chen, Zhenghua Pan, Yuping Jin, Shi Chen, and Lingyun Sun. 2024. SimUser: Generating Usability Feedback by Simulating Various Users Interacting with Mobile Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[80] Xiaotong Xu, Jiayu Yin, Catherine Gu, Jenny Mar, Sydney Zhang, Jane L E, and Steven P Dow. 2024. Jamplate: Exploring LLM-Enhanced Templates for Idea Reflection. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 907–921.

[81] Yu-Chun Grace Yen, Jane L E, Hyoungwook Jin, Mingyi Li, Grace Lin, Isabelle Yan Pan, and Steven P Dow. 2024. ProcessGallery: Contrasting Early and Late Iterations for Design Principle Learning. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–35.

[82] Xishuo Zhang, Lin Liu, Yi Wang, Xiao Liu, Hailong Wang, Chetan Arora, Haichao Liu, Weijia Wang, and Thuong Hoang. 2024. Auto-Generated Personas: Enhancing User-centered Design Practices among University Students. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.

[83] Xishuo Zhang, Lin Liu, Yi Wang, Xiao Liu, Hailong Wang, Anqi Ren, and Chetan Arora. 2023. Personagen: A tool for generating personas from user feedback. In *2023 IEEE 31st International Requirements Engineering Conference (RE)*. IEEE, 353–354.

[84] Yao Zhang and Chang Liu. 2020. Users' Knowledge Use and Change during Information Searching Process: A Perspective of Vocabulary Usage. In *Proceedings of the ACM/IEEE joint conference on digital libraries in 2020*. 47–56.

A Appendix

A.1 User Study

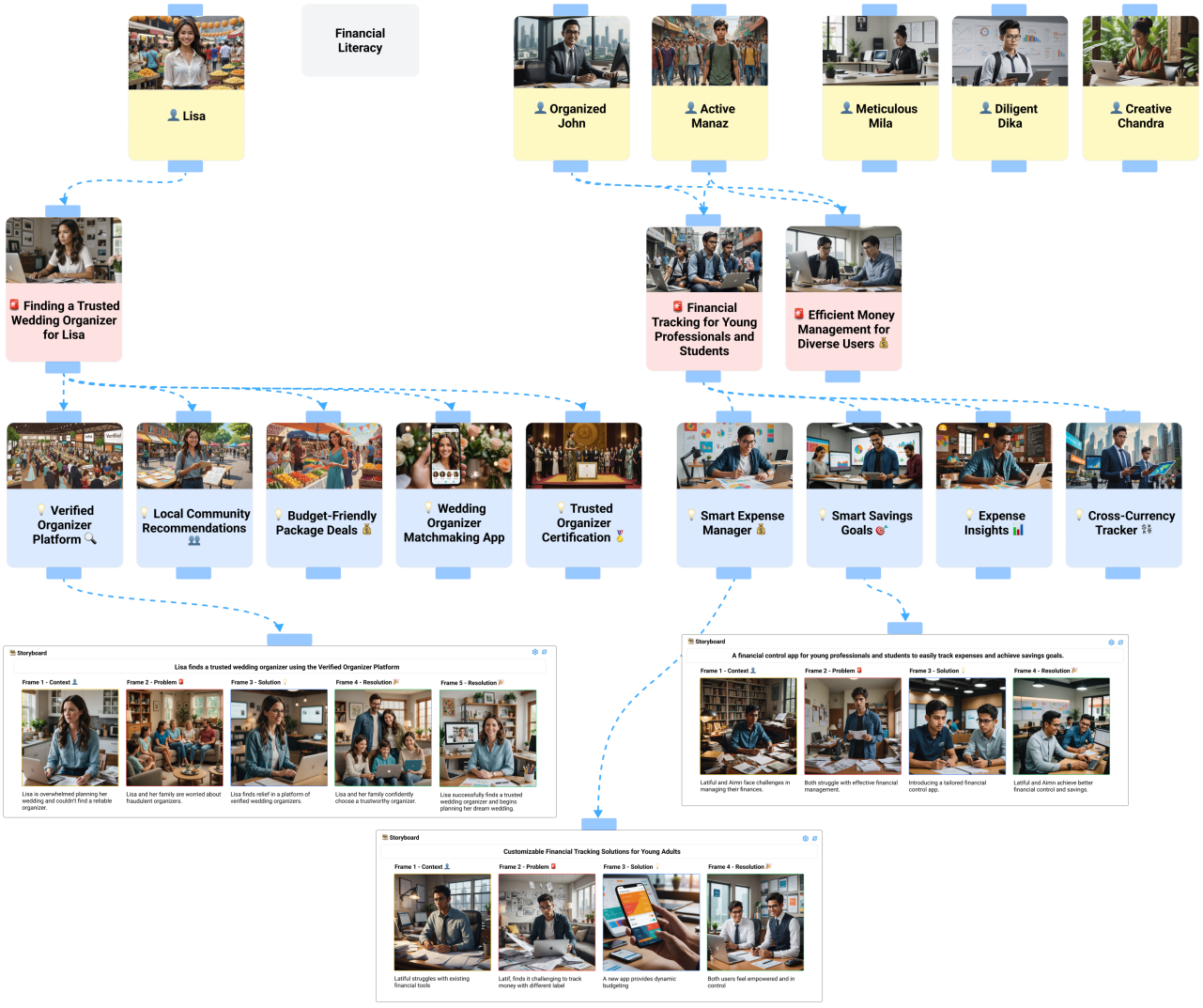


Figure 18: The workspace of $P1_{prac}$ after completing the user study task

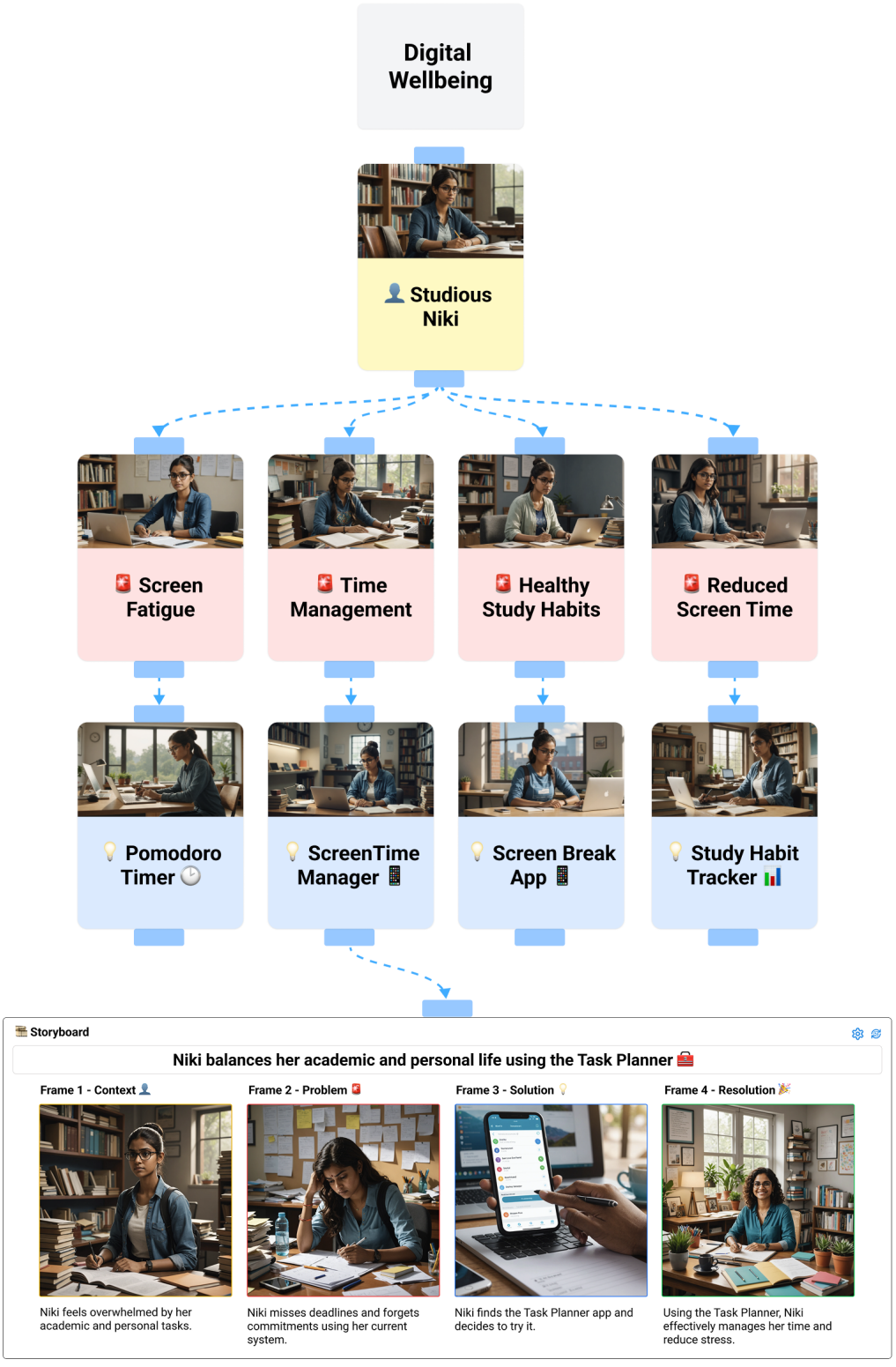


Figure 19: The workspace of $P2_{prac}$ after completing the user study task

Table 3: Some of the topics provided to user study participants as a reference to choose from. They were also free to choose a topic outside this list. The bold indicates topics participants chose.

Topic	Background	Task
Financial Literacy (P1)	Many individuals struggle with managing their finances due to a lack of accessible and engaging educational resources.	Help users better understand and manage their finances.
Health and Wellness Tracking (P5)	Tracking health and wellness metrics can help individuals lead healthier lives, but existing tools often fail to engage users consistently.	Design a tool that helps users engage with their health and wellness.
Education and Learning (P8)	Educational experiences vary widely in effectiveness, often depending on how well they cater to individual learning needs and styles.	Enhance educational experiences by addressing diverse learning needs.
Personalized News Consumption (P4)	The rise of digital news has led to concerns about echo chambers and misinformation. Users need tools that allow them to consume news that is both personalized and balanced.	Enable personalized and balanced news consumption.
Gamification of Learning (P7)	Gamification has been shown to increase engagement and motivation in learning, but its application in education is still evolving. Effective design is key to balancing fun and educational value.	Make learning more engaging through innovative approaches.
Sustainable Living (P3)	Environmental sustainability is becoming increasingly important, but many people struggle to incorporate sustainable practices into their daily lives.	Create a way to encourage and support sustainable living practices.
Digital Wellbeing (P2)	The pervasive use of digital devices can lead to challenges in maintaining a balanced lifestyle, impacting both mental and physical health.	Develop a solution that promotes healthy digital habits.
Accessibility in Public Spaces	Public spaces are essential for community interaction, but they are often not fully accessible to all individuals. This can create barriers to participation and inclusion.	Design a solution that enhances accessibility in public spaces for diverse users.
Remote Work Collaboration	The rise of remote work has transformed how teams collaborate, bringing new challenges in communication, coordination, and maintaining a sense of connection.	Improve the experience of remote team collaboration.
Smart Home Interfaces	As smart home technology becomes more prevalent, users face challenges in managing multiple devices with varying levels of complexity.	Simplify and improve the interaction with smart home technology.
Transportation and Mobility	Urban mobility is a critical issue, with challenges such as traffic congestion, inefficient public transport, and accessibility for people with disabilities.	Improve urban mobility and the user experience of transportation.

A.2 Pipeline & Prompts

Table 4: System-wide prompts used with other prompts to enforce response format.

Prompt Type	Prompt

Enforce response JSON schema	Output a JSON object that fits the schema based on the user message.
	
	`JSON SCHEMA: """`
	`{jsonSchema}`
	`"""`

Table 5: Prompts for generating and regenerating user personas.

Prompt Type	Prompt
Generate personas	User personas are detailed descriptions of fictional characters that represent different user types. They help designers gain empathy for their users and understand their needs, goals, and behaviors. Keep the values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Use an alliterative adjective and first name to create a memorable name for the persona. Create {numberOfVariations} personas based on the given context. Context: """ {context} """
Regenerate personas	User personas are detailed descriptions of fictional characters that represent different user types. They help designers gain empathy for their users and understand their needs, goals, and behaviors. Keep the values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Use an alliterative adjective and first name to create a memorable name for the persona. Update the given personas based on the new context/instructions, keeping the essence of the persona intact, but making necessary changes to ensure cohesiveness. When the context requests regenerating personas based on updated problems: - Rewrite the personas to directly address the problems. - Remove any information that isn't directly related to the new problems. Personas: """ {personas} """ Context: """ {context} """

Table 6: Prompts for generating and regenerating problem statements.

Prompt Type	Prompt
Generate problem statements	A problem statement is a description of an issue designers are trying to solve. A good problem statement should be centered on specific people and their needs. It should be narrow enough to be manageable but broad enough to explore a variety of solutions. A problem statement should be a statement and not a question. A problem statement should not be a hypothetical solution or strongly suggest a single solution. Keep values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Create {numberOfVariations} problem statements based on the given context. Context: """ {context} """
Regenerate problem statements	A problem statement is a description of an issue designers are trying to solve. A good problem statement should be centered on specific people and their needs. It should be narrow enough to be manageable but broad enough to explore a variety of solutions. A problem statement should be a statement and not a question. A problem statement should not be a hypothetical solution or strongly suggest a single solution. Keep values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Update the given problem statements based on the new context/instructions. Keep the essence of the problem statement intact, but make necessary changes to ensure cohesiveness. When the context requests regenerating problems based on updated personas: - Rewrite the problem statements to directly address the personas. - Remove any information that isn't directly related to the new personas. When the context requests regenerating problems based on updated solutions: - Rewrite the problem statements to directly address the solutions. - Remove any information that isn't directly related to the new solutions. Problems: """ {problems} """ Context: """ {context} """

Table 7: Prompts for generating and regenerating solutions.

Prompt Type	Prompt
Generate solutions	A design solution is a proposal for addressing specific problems or user needs. Solutions don't have to be perfect, but instead should be a starting point for further refinement and iteration. Keep values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Create {numberOfVariations} solutions based on the given context. Context: """ {context} """
Regenerate solutions	A design solution is a proposal for addressing specific problems or user needs. Solutions don't have to be perfect, but instead should be a starting point for further refinement and iteration. Keep values short and to the point and use emojis where possible. Avoid repeating the same information in different values. Update the given solutions based on the new context/instructions. Keep the essence of the problem statement intact, but make necessary changes to ensure cohesiveness. When the context requests regenerating solutions based on updated problems: - Rewrite the solutions to directly address the problems. - Remove any information that isn't directly related to the new problems. When the context requests regenerating solutions based on updated storyboards: - Rewrite the solutions to directly address the storyboards. - Remove any information that isn't directly related to the new storyboards. Solutions: """ {solutions} """ Context: """ {context} """

Table 8: Prompts for generating a storyboard outline and image prompts.

Prompt Type	Prompt
Generate storyboard outline	You are an AI assistant tasked with creating a detailed storyboard outline for design thinking based on specific dimensions and their assigned values. Use the given dimensions to generate a coherent storyboard outline to visualize a problem, solution, and the surrounding context. The title is a sentence summarizing a story about the user situation, problem, solution, and resolution. Example title: "A seed catalog app that lets users watch videos instead of reading plant info." Generate an outline with at least 4 frames. The frame outline contains a brief description for each frame in the storyboard. The description should tell a cohesive story about the user situation, problem, solution, and resolution. Each frame description should describe key observations and actions in the frame. Each frame outline also contains a caption that will be displayed directly below the visuals as part of the storyboard. The caption should add clarity and context to the visual image. The caption should be shorter than the frame description since it will be displayed directly below the visual image. When the context requests regenerating storyboards based on updated personas, problems, and solutions: - Revise the storyboard to address the personas, problems, and solutions. - Remove any information that isn't directly related to the new personas, problems, and solutions. Outline Example: Context: """ {context} """

(Table 8. continued)

Prompt Type	Prompt
Generate storyboard image prompts	Given the outline for a storyboard with {numFrames} frames, generate a list of image prompts for each frame. Each frame outline contains an image prompt and imageNegativePrompt used to generate visuals for the storyboard frame. The image prompt should be based on the frame description but focus on describing the visual elements. The image prompt should briefly describe the subject using key adjectives. Additionally, it should describe the setting, lighting, and any other relevant visual details. Wrap terms in `(#weight)` to adjust the importance of the term in the image. The weight should be a number between 0 and 1. Do not set the weight for random terms, only for essential terms required for the subject and action. The image prompt within a single outline should be consistent in style and form to ensure cohesive visual narrative. Consider using specific details such as colors, textures, and other style modifiers. Shot types such as wide shot, medium shot, close-up, etc. can be used to specify the framing of the image. Choose the shot type that best suits the frame description. The negative prompt should omit things that are difficult for AI to generate or are not relevant to the frame. Include the full name of people to ensure consistent characters across all frames. Frames: """ {frames} """ Visual Character Descriptions: """ {visualCharacterDescriptions} """

Table 9: Prompts for generating image prompts and visual character descriptions.

Prompt Type	Prompt
Generate illustrative image prompt	Using the following idea, generate an image prompt and image negative prompt to generate an illustrative image that represents the key elements of the idea. Describe a scene which is a visual metaphor for the persona, problem, or solution described in the idea. Prompts should be short and focus on the visual metaphor and not overly describe the characters. Describe the image in literal visual elements of the image, not the message or meaning of the image. Idea: """ {idea} """

(Table 9. continued)

Prompt Type	Prompt
Generate problem illustrative image prompt	Generate an image that depicts a problem and helps to build empathy. Generate a prompt for the image which describes the scene that depicts the problem. Negative prompts describe what the scene should not include. Problem: """ {problem} """
Generate visual character descriptions	Generate visual character descriptions used to ensure visual consistency between different artists and illustrations. Generate a visual character description used to illustrate the following idea. Idea: """ {idea} """

Table 10: Prompts for generating feedback on design artifacts.

Prompt Type	Prompt
Generate persona feedback	You are an UX designer given a persona. Brainstorm questions to evaluate this persona being used for design thinking. - Consider any information that is missing, but could be useful. - Example: If we know the persona is a salesperson, it may be useful to know what kind of product they sell? - Consider the accuracy and consistency of the persona. - Example: If a persona is experienced, but only has a few years of experience this may be incorrect. - Consider alternative persona values which may be useful. - Example: If a persona is a salesperson, could a sales manager be another useful persona to consider. Persona: """ {persona} """

(Table 10. continued)

Prompt Type	Prompt
Generate problem feedback	You are an UX designer given a problem statement. Brainstorm questions to evaluate this problem statement being used for design thinking. Focus on the quality of the problem statement rather than solutions. - Consider any information that is missing, but could be useful. - Example: If we know a salesperson is struggling to find qualified leads, it may be useful to know what strategies they have tried. - Consider the accuracy and consistency of the problem statement. - Example: If a salesperson struggles to make a connection, but can find qualified leads, this may be incorrect since qualified leads require connections. - Consider alternative or related problem statements which may be useful. - Example: If a salesperson struggles to find qualified leads, they may also struggle with finding leads in general. Problem: """ {problem} """
Generate solution feedback	You are an UX designer given a solution. Brainstorm questions to evaluate this solution being used for design thinking. - Consider if any information that is missing, but could be useful. - Example: If we want to create a networking mobile app, what are the key features of the app? - Consider the accuracy and consistency of the solution. - Example: A high-tech drone solution to deliver food to low-income families may be infeasible as the solution may be too expensive. - Consider alternative or related solutions which may be useful. - Example: If our solution helps a salesperson find leads, it may be useful to also have a solution for qualifying those leads. Solution: """ {solution} """
Generate storyboard feedback	You are an UX designer given a storyboard. Brainstorm questions to evaluate this storyboard being used for design thinking. Storyboard: """ {storyboard} """
Generate group feedback	You are an UX designer given a list of nodes representing personas, problems, solutions, and storyboards. Brainstorm questions to evaluate these nodes being used for design thinking. Nodes: """ {nodes} """

Table 11: Prompts for generating autocomplete recommendations.

Prompt Type	Prompt
Generate problems recommendations	Generate autocomplete suggestions for descriptions of a general group of problems based on the provided personas. Suggestions should be a couple of words. Nodes: """ {nodes} """
Generate solution recommendations	Generate autocomplete suggestions for descriptions of a general group of solutions based on the provided problems. Suggestions should be a couple of words. Nodes: """ {nodes} """
Generate dependent storyboard recommendations	Generate autocomplete suggestions for the title of a general group of storyboards based on the provided personas, problems, and solutions. The title should generally describe the plot of the storyboard. Nodes: """ {nodes} """
Generate more personas recommendations	Generate autocomplete suggestions for descriptions of a general group of personas which differ from the provided personas. Suggestions should be a couple of words. Nodes: """ {nodes} """
Generate more problems recommendations	Generate autocomplete suggestions for descriptions of a general group of problems which differ from the provided problems. Suggestions should be a couple of words. Nodes: """ {nodes} """
Generate more solutions recommendations	Generate autocomplete suggestions for descriptions of a general group of solutions which differ from the provided solutions. Suggestions should be a couple of words. Nodes: """ {nodes} """
Generate more storyboard recommendations	Generate autocomplete suggestions for the title of a general group of storyboards which differ from the provided storyboards. The title should generally describe the plot of the storyboard. Nodes: """ {nodes} """

(Table 11. continued)

Prompt Type	Prompt
Revise node recommendations	Generate suggestions for specific instructions to update/change the following design thinking nodes. Nodes will be either personas, problems, solutions, or storyboards. Suggestions should focus on variations or edits to existing nodes. Constrain suggestions to 100 characters. Nodes: """ {nodes} """

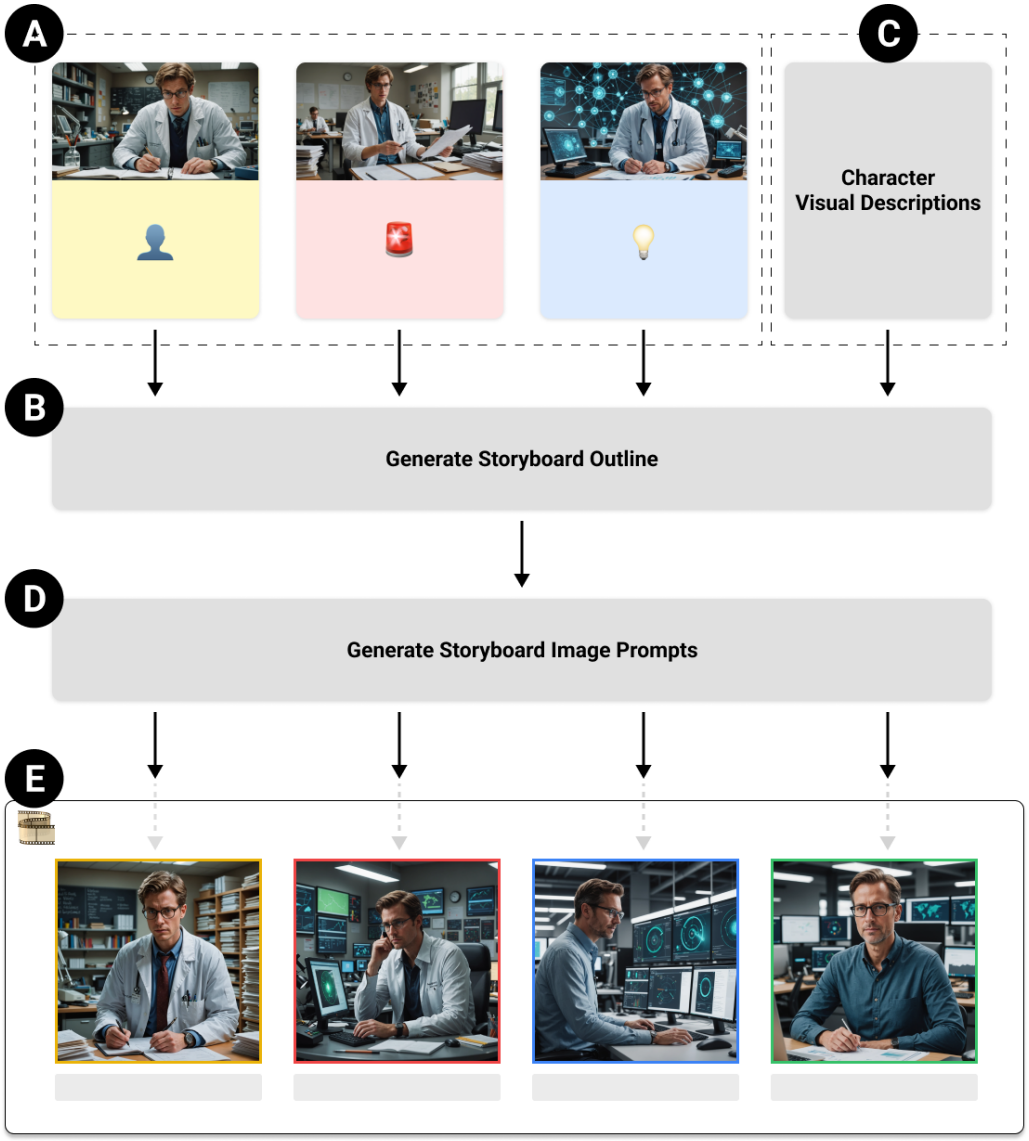


Figure 20: Pipeline for building a storyboard: (A) Persona, problem, and solution inputs are used to generate a (B) text storyboard outline that includes a title and frames consisting of a caption and description. (C) Visual character descriptions help to maintain visual consistency between the design artifacts and the storyboard. The storyboard outline is used to generate (D) image prompts in a single LLM invocation to ensure continuity across frames. (E) Images are generated in parallel using image prompts from the previous step.