

# TEXTERIAL: A Text-as-Material Interaction Paradigm for LLM-Mediated Writing

Jocelyn Shen  
Microsoft Research  
Redmond, WA, USA  
MIT Media Lab  
Cambridge, MA, USA  
joceshen@mit.edu

Nicolai Marquardt  
Microsoft Research  
Redmond, WA, USA  
nicmarquardt@microsoft.com

Hugo Romat  
Microsoft Research  
Redmond, WA, USA  
romathugo@microsoft.com

Ken Hinckley  
Microsoft Research  
Redmond, WA, USA  
kenneth.p.hinckley@gmail.com

Nathalie Riche  
Microsoft Research  
Redmond, WA, USA  
nath@microsoft.com

Fanny Chevalier  
Microsoft Research  
Redmond, WA, USA  
University of Toronto  
Toronto, ON, Canada  
fanny@dgp.toronto.edu

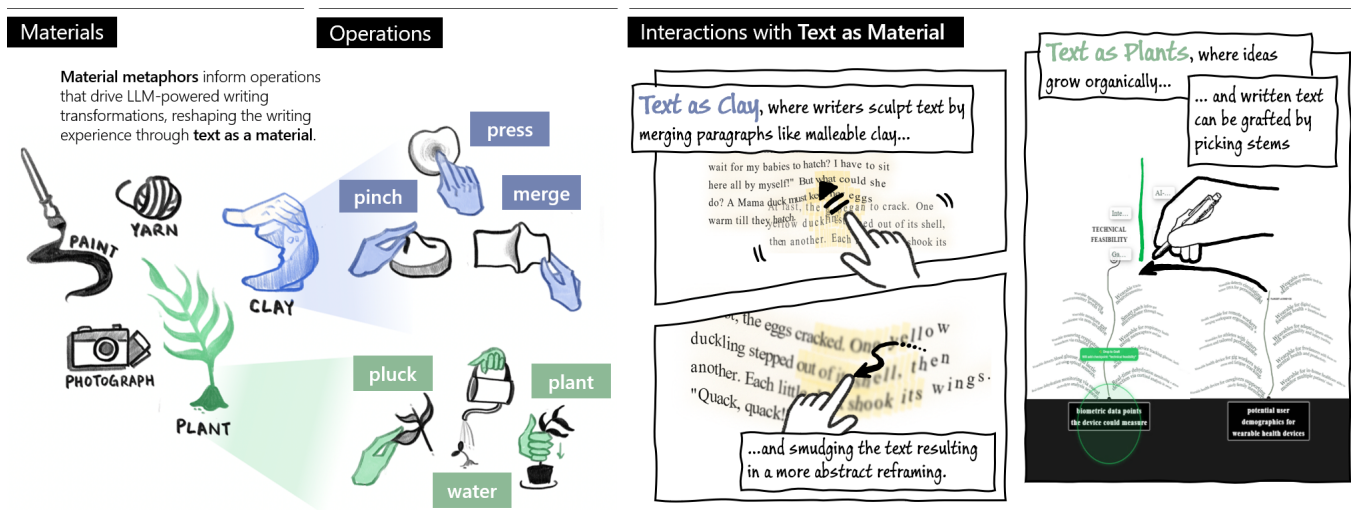


Figure 1: We explore how materials can inspire new interactions with text. Building on insights from a formative study, we propose TEXTERIAL, a conceptual framework and prototype implementations that treat text-as-material. Our work uses materiality to foster a new way of interacting with generative AI.

## Abstract

What if text could be sculpted and refined like clay—or cultivated and pruned like a plant? TEXTERIAL reimagines text as a material that users can grow, sculpt, and transform. Current generative-AI models enable rich text operations, yet rigid, linear interfaces often mask such capabilities. We explore how the text-as-material metaphor can reveal AI-enabled operations, reshape the writing process, and foster compelling user experiences. A formative study shows that users readily reason with text-as-material, informing a conceptual

framework that explains how material metaphors shift mental models and bridge gulfs of envisioning, execution, and evaluation in LLM-mediated writing. We present the design and evaluation of two technical probes: *Text as Clay*, where users refine text through gestural sculpting, and *Text as Plants*, where ideas grow serendipitously over time. This work expands the design space of writing tools by treating text as a living, malleable medium.

## CCS Concepts

• Human-centered computing → Interactive systems and tools; Natural language interfaces; HCI theory, concepts and models.

## Keywords

writing interfaces, design metaphors, materiality, LLMs, AI



This work is licensed under a Creative Commons Attribution 4.0 International License. CHI '26, Barcelona, Spain

© 2026 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-2278-3/2026/04  
<https://doi.org/10.1145/3772318.3790330>

**ACM Reference Format:**

Jocelyn Shen, Nicolai Marquardt, Hugo Romat, Ken Hinckley, Nathalie Riche, and Fanny Chevalier. 2026. TEXTERIAL: A Text-as-Material Interaction Paradigm for LLM-Mediated Writing. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 33 pages. <https://doi.org/10.1145/3772318.3790330>

## 1 Introduction

Current large language models (LLMs) offer unprecedented capabilities to transform and generate text, from abstracting meaning [31] and remixing ideas [22], to suggesting alternatives [46] or drafting entire passages from a single seed phrase [9, 75]. Yet, these capabilities are not always used to their full potential. Although natural language instructions provide versatility, crafting effective prompts and steering the results is often difficult [76]. In particular, understanding what the model can do and how to instruct it within the constraints of linear chat interfaces, limits the ways in which people fluidly think *with* the model and bottlenecks their wide-ranging capabilities [69].

To better surface the expressive potential of LLMs, we propose a conceptual shift for designing generative AI-powered writing experiences: **reimagining text as a material** — something that can be directly manipulated, blended, squashed, or shattered. Our framework is inspired by metaphors to physical materials (clay, paint, or even plants), which map LLM operations to expressive touch interactions. For example, imagine a creative story writer refining their draft. Instead of prompting, “*Make this paragraph more vivid, while keeping the introduction and conclusion the same, and merge it with the style of Shakespeare...*” the user might stretch the text like clay to expand the text, pinch it to make certain phrases more concrete, and smear one of Shakespeare’s sonnets into their writing. Through this perspective, the material properties, and interactions afforded by the material, help externalize the writing process by leveraging intuitive sensorimotor skills and embodied reasoning (e.g., ripping, mixing, smoothing) in ways that prompts cannot [33, 42, 49, 56].

To expand this idea of text-as-material, TEXTERIAL generalizes HCI design metaphors to LLM-mediated writing. We envision these techniques *enriching* and *complementing* existing writing techniques rather than replacing them, opening up a design space of material-inspired interactions that can infuse the wider scope of writing activities at different stages—early ideation, editing, or iterative refinement. To explore this design space, we first conducted a formative study to understand how individuals might conceptualize text as a material, gaining insight into what metaphors resonate with people, and how they would approach writing under such a paradigm (Section 3). Building on these insights, we distilled important dimensions and considerations for material-inspired LLM-driven writing tools into a conceptual framework, from the lens of the seven-stages-of-action model [58] in Section 4. Our conceptual framework articulates the layers on which LLMs operate (semantics, structure, style), operations on these layers (compose, abstract, ideate, condense, transform); and mapping execution of operations to material-like interactions. This framework formalizes LLM functional capabilities and metaphor-interface matches [56], providing interaction designers with a structure and language to discuss how materials afford different AI-powered operations and inspire new

writing interfaces that help bridge the gulfs of envisioning, execution, and evaluation [58, 69].

As concrete demonstration of this conceptual framework, we developed two technical probes that embody text as material for writing tasks with LLMs: *Text as Clay* and *Text as Plants* (Section 5). *Text as Clay* reimagines the text itself as clay-like material that can be pinched, stretched, pulled, merged, and smoothed through direct multi-touch gestures, to combine and refine text fragments. *Text as Plants* treats writing as an act of gardening: ideas are planted, watered, and grafted, and the system uses a language model to introduce variations, novel ideas, and leaves that branch and evolve over time. Both of these technical probes are positioned to explore how material-inspired interactions can complement existing writing practices, with the potential to address distinct phases in the writing process. We used these probes to evaluate our approach in a qualitative focus-group study (n=10, 4 groups; Section 6). Our findings indicate that framing text-as-material influences mental models. Gestural manipulations were found to be intuitive, expressive, and language-agnostic, which invited direct engagement with text transformations. The spatial and visual layouts of our prototypes supported non-linear exploration and non-committal ideation, workflows that are rarely afforded by linear LLM interfaces. Finally, participants found that our prototypes infused creative experimentation and playfulness in the process, and envisioned extensions of the metaphor to collaborative writing interactions.

In sum, TEXTERIAL contributes the following:

- **Empirical insights from a formative study** of how people conceptualize text-as-material, illustrating how material metaphors can naturally map to writing operations and phases of the writing process.
- **A conceptual framework** of the text-as-material approach that links material affordances to text manipulations with LLMs, providing foundations for interaction designers of novel, AI-driven writing tools.
- **Two diverse technical probes**—*Text as Clay* and *Text as Plants*—expanding the design space of writing tools by treating text as a malleable, interactive medium to embody underlying LLM operations.
- **Findings from a qualitative focus-group study** of these technical probes, offering evidence that material metaphors shape mental models for manipulating text, foster novel workflows, raise different expectations for different people, and invite creative and playful explorations.

Our research envisions a new way of writing with AI: one that invites users to craft and play with text as if it were a physical material and imagines interactions that enable direct engagement with the writing process, rather than restricting users to linear, chat-based interfaces that pull the writer away from the act of writing itself. With the structured framework, technical probe explorations, and user study insights, our work is intended to inform and inspire designers and developers of writing interfaces, providing a set of enriching and complementary techniques that broaden the expressive possibilities of AI-mediated writing.

## 2 Background and Related Work

Our work applies the conceptual lens of metaphors in HCI to writing interfaces. We review literature on materiality and design metaphors, conceptual frameworks that inform writing interfaces, and the design and evaluation of digital writing systems previously introduced in the field.

### 2.1 Materiality and Design Metaphors

HCI has long used metaphors to make digital systems more understandable by grounding them in familiar concepts [6, 12, 56]. For example, early graphical interfaces embraced metaphors literally through skeuomorphs—interface elements that mimic the aesthetics of a familiar physical medium to convey affordances [1, 36, 41, 50]—introducing desktop, files, and trash can digital counterparts to physical desktops to help users understand and appropriate new technologies [26]. Although metaphors are a powerful scaffold for developing one’s mental model and providing intuition and context [56, 60, 61], interaction designers caution against letting metaphors hinder innovation by constraining themselves too closely to the physical world [24].

Artists and craftspeople often think through materials: the affordances and constraints of a medium like clay or wood actively shape the creative process, often suggesting new directions [48, 54]. Design metaphors concerned with materials have inspired work in HCI, such as tangible user interfaces [35, 59] and direct manipulation techniques [34]. As highlighted in theories of material engagement, materials allow users to apply their intuitive spatial or physical reasoning to interaction [49], i.e. a form of *embodied* interaction [19]. Materials provide a slow and iterative creative process and rapport to the medium itself, allowing for direct manipulation of objects on a screen [48, 49].

Our work builds on this rich history of materiality and design metaphors in HCI to explore how we might interact with *text* as a material [66]. This departs from the standard keyboard-screen model of writing, much as tangible user interfaces departed from WIMP interfaces [3]. By conceptualizing text as a material, we aim to surface the capabilities of underlying generative models in writing interfaces through natural gestures and manipulations of the material.

### 2.2 Conceptual Frameworks for Writing Interfaces

Prior works have proposed conceptual frameworks that categorize and inspire new writing interfaces. Such frameworks typically delineate the stages of the writing process, following the model of planning, translating, and revising in some form [16, 20, 30, 74]. Review articles in the field find a dearth of writing tools targeted specifically for the pre-writing phase (e.g., idea generation) [79].

Recent works provide design spaces and frameworks that are helpful for describing *AI-driven* writing tools, drawing inspiration from creative practices [21, 38, 43, 44]. These frameworks are often categorized by author attributes, the writing process, the paradigm or interaction system, the technology used to assist writing, or the overall ecosystem where writing occurs [43]. For instance, Buschek [10] proposes approaching the design of modern LLM-driven writing tools through a “collage” process: interfaces fragment text into pieces, juxtaposed with different voices or content sources, and shift the

author’s role from composing text to arranging it. The *Cells, Generators, and Lenses* structure the design space from an algorithmic, object-oriented perspective [39]. In this model, the key components of text generation are reified into interactive objects: cells (text units or prompts), generators (configurable model instances), and lenses (output displays or filters). By conceptualizing writing tools as interactive, composable, and persistent UI elements, the design framework focuses primarily on the iterative process of creation [65]. Beyond the text domain, “AI Instruments” enable making a user’s intent into a tangible, direct-manipulation tool [63]. These frameworks provide lenses to both *generate and understand* novel experiences in the age of generative AI.

Overall, conceptual frameworks in the field offer both *analytical* tools to critique existing AI writing UIs, and *constructive* tools to inspire new designs focused on embracing fragmentation and recombination in writing. Our framework instead provides a structured approach for designing interaction techniques from the perspective of the human model of action [58]. It does so by articulating functional capabilities of LLMs and conceptualizing interaction design in terms of materiality and its affordances [29, 57].

### 2.3 Writing Interfaces in HCI

Prior works have long explored developing novel writing interfaces that move beyond the static “typewriter” paradigm. Early hypertext systems, such as Bolter and Joyce’s interactive fiction editor let writers interact with a diagrammatic view of their text [7], paving the way for non-linear, “fluid text” documents that acted as canvases which could be dynamically rearranged to make room for additional information beyond the text itself [77]. Interest in writing support tools have resurged with modern AI, with many early ideas being viable with today’s LLMs [47].

A first line of work emphasizes direct manipulation for the task of **text editing**. These interfaces often focus on addressing the challenge of maintaining consistency in textual documents and controllable text manipulation. For example, Textlets [28] reify text selections into persistent entities, which can aid in search-and-replace operations. Through focus-group studies, this work found that “reification” (turning commands into first class objects or instruments), “polymorphism” (applying instruments to different types of objects), and “reuse” (making user inputs and system outputs reusable) encourages writers to use functions they might otherwise avoid with traditional text editors. With LLMs, recent interfaces focus primarily on leveraging models to perform writing tasks with greater controllability, while reducing the burden of drafting complex prompts or task sequences to achieve the user’s editing goals [44]. For example, DirectGPT [52] maps prompts to direct manipulation actions and minimizes both time and length of prompts written to edit text, code (and also images), when compared to baseline ChatGPT. LMCanvas [40] reifies prompts and model actions into composable interface elements, allowing writers to spatially reorganize ideas and invoke LLM text operations in place, rather than opening separate chat windows. Composable tools like Phraselette [11] provide a visual palette to inspire more creative language, where authors can specify constraints applied to snippets of AI-generated text. By making small chunks of content malleable, these interfaces are designed to

give authors interactive control over LLM edits without significant overhead of drafting highly specific prompts for steerability.

Beyond interfaces that support text editing, other works use **visualization and information structuring** to give the user better "gists" of their current writing or to explore idea spaces. For example, Lexichrome [37] investigates how relationships between colors and words can support writers quickly grasp editorial decisions. Researchers have also introduced systems that leverage LLMs to generate dynamic summaries that assist writers with *global* text coherence. For example, Dang et al. [17] proposed an editor that displays continuous AI-generated summaries for each paragraph alongside a written draft. They find through interview studies that AI summaries help authors reflect on the structure and alignment of their arguments as they write. Other works generate more structured gists through visual node-link maps that can help writers identify gaps in reasoning or support narrative construction [78]. For example, Masson et al. [53] introduced "visual story-writing" and developed a text editor that automatically visualizes a graph of entity interactions, movement between locations, and a timeline of story events such that interactions with the visual representations translate back to written text. For idea generation and sensemaking, tools such as Luminare [70] and Sensecape [71] let users explore a semantic "landscape" of AI-suggested ideas. In these works, the model generates and clusters concepts related to the user's topic, which the writer can browse or interpolate to discover new angles on their writing. Similarly, WhatELSE [46] uses AI to create a "possibility space" from user-defined narrative constraints, which allows writers to navigate variations in text and different levels of abstraction. Such tools support writer's sensemaking when dealing with a large volume of suggestions [23].

Most closely related to our work are recent LLM-powered tools that **adapt design metaphors to support text editing or text generation**. For example, Textoshop [51] approaches text editing through the analogy of photo editing software such as Photoshop, using smudge or tone brushes to smooth transitions or apply stylistic features. The authors found that Textoshop improved users' editing efficiency and satisfaction, preferring the interface over conventional document editors with ChatGPT. Other works draw analogies to layering [51, 68], sketching [14] and painting [13], or even toy play [15] to steer text generation. As one example, Script&Shift [68] presents a layered document interface where writers can separate *content* layers from *stylistic* layers, using modular units that encapsulate elements of the writing process to minimize disruption to the workflow. As another design metaphor example, TaleBrush [14] enables writers to draw a story arc on a digital canvas, then uses a LLM to generate narrative text that follows the curve. This interaction technique allows authors to "doodle" or "sketch" stories rather than typing full sentences. Similarly, ToyTeller [15] treats narrative construction as a form of toy play, where users manipulate toy characters and props to generate story text. Through qualitative analysis of think-aloud statements and interviews, these studies often find that such metaphors make the creative process more approachable, rather than leaving the user to wrestle with a blank page.

Across these systems, HCI researchers have explored diverse metaphors to leverage the capabilities of generative models and shape the writing process. Our work builds on and extends this

trajectory. Moving beyond isolated metaphors, we propose a text-as-material perspective that synthesizes (1) the core functions and possibilities of AI-mediated writing, and (2) the way material affordances can shape how people understand and work with these tools.

### 3 Formative Study

We first explored how people respond to the idea of treating text as material. In a formative study, participants imagined text as a handcraft medium and described the manipulations a digital tool could support in this paradigm. Our goal was to assess whether conceptualizing text as materials lends itself to text manipulations, inspires novel workflows, and what specific metaphors resonate with individuals; rather than than exhaustively cataloging all possible ideas. We further explore our concept following a research-through-design approach [80] using technical probes (section 5).



**Figure 2: Elicitation materials from our formative study. The inspirational cards include entry-level crafts (gardening, pottery, photography) and advanced ones (embroidery, beekeeping), as well as inanimate (clay, threads), living (plants, bees), and hybrid materials (photographed scenes).**

#### 3.1 Methodology & Procedure

The facilitator briefed participants in a walkthrough, illustrating the expected thinking using paint as an example material: "If text were paint, what would mixing paint mean? Would the same text manipulations apply if text were dry or wet paint?"

Our methodology builds on ideation techniques that use metaphor as a generative tool [45, 72]. To elicit ideas, we created inspirational cards [45] featuring various common handcrafts and their associated materials (Figure 2). We suggested a loose explanatory structure [72] inviting participants to 'tell the metaphor's story'—explain how text is manipulated through the metaphor; 'elaborate the triggering concept'—identify what aspect of the material prompted the conceptual transfer; and 'elaborate assumptions'—describe how this should work in practice. Participants were invited to use the cards freely and suggest their own handcraft-materials. We provided prototyping tools (paper, pencils, post-its) to help them express their ideas. The facilitator intervened only to clarify or probe further. Details about our study methodology and materials used are included in our Supplementary Material.

Each session lasted about one hour, and video and audio recordings were collected. Participants received \$50. The study was approved by the ethics board of [redacted]. We performed a thematic analysis of the videos and transcripts, using our guiding questions

**Table 1: Formative study participants. The creativity score (min: 1, max: 50, higher is more creative) and experience with writing score (min: 6, max: 30, higher is more experienced) are based on Renom et al. [60].**

ID	Occupation	Creativity	Experience	Type of Writing
P1	Model maker (15 yrs)	38	12	N/A
P2	Designer (20+ yrs)	43	21	Creative briefs, presentations
P3	Principal PM Manager (1 yr)	37	25	Emails, essays, scientific articles, journaling
P4	NLP Researcher (20+ yrs)	45	29	Scientific articles, short stories, poetry

as deductive codes, while emerging themes were inductively captured. We focus our analysis on existence, following an interpretivist approach to reflexive thematic analysis [8].

### 3.2 Participants

We recruited four participants through convenience sampling at a tech company (age ranges: 30-39 (1) and 50-59 (3), 2 males and 2 females), targeting creative individuals and communication professionals, and verified alignment using creativity and text-experience questionnaires from Renom et al. [60]—see Table 1. We recruited participants with varied backgrounds and writing experience to capture different perspectives on material-inspired interactions across different writing contexts. While our sample is small, our study confirmed that the concept of text-as-material has strong potential, which our work further explores through functional probes.

### 3.3 Results

**Writing without words—Treating text as a material enables a non-verbal approach to writing, which can offer several practical advantages.** Participants contrasted this with traditional writing using a pen or keyboard, which demands deliberate word choice. The material metaphor instead allows emotions to be enacted physically and visually, by assembling and manipulating fragments of ideas. For example, P1 and P3 described sketching as a way to “let the emotion out,” and P2 emphasized the satisfaction of physical expression: “That’s why I want to play with clay. I just slam it down and pick it up.” Participants felt that such embodied, primal, organic experiences better matched the “non-linear”, “messy and very unstructured creative part” of the process (P2, P4) and introduces “humanness in the process.” As P1 put it: “I’m not there just to convey information. I’m there to elicit an emotional response.” P2 and P4 also remarked that skeuomorphic representations of text could allow them to “zoom out” (P2) and visually capture properties such as redundancies (P4) in ways not possible with words.

**Working with materials can offer an easy entry into the writing process.** For many, writing can feel intimidating. Using alternative handcrafts as a proxy for expressing meaning could make the process feel less fraught. Participants described how approaching text from the art of a familiar handcraft could help make it “instant” (P4). As P1 detailed “It feels like I’m trying to capture that moment because I know it’s going to be gone and I think a lot of... text is like that.” For this, the metaphor needs to leverage a craft that the author has a reasonable understanding of and expectations about. P4, for whom writing is the most comfortable craft, said they would not benefit from the material approach, but “I agree with this philosophy. It’s just that for me [other crafts] are much more narrow.” They rejoin P2 in that such techniques can help people overcome the writer’s block (P2, P4) and fear of failure (P2, P1).

**Different materials lend themselves to different writing tasks,** such as idea generation, writing (the actual translation of putting words down), and editing, as identified by P4. Gardening was linked to ideation (P1, P4) where the notion of “growing text” captured co-creation: “there’s another element that’s providing suggestions” (P1). P1 contrasted this with bonsai which they described as “arrogant”—“you’re absolutely forcing it to be something”. This contrast highlights the different forms of authorial agency between material and craftsperson that can be afforded within the same metaphor. Inanimate materials, like clay (P2, P3, P4), metal (P2), paint (P4), jelly paint (P2), and fabric (P4) inspired narrative shaping and editing. Participants moved from rough assembly—“I roughly know what I want to write. So I’m gonna throw that hunk of clay down. (P2)—to refinement, adjusting content and style—“I need to make it constructive, so I’m just going to sort of slowly take a look at what I have and I’m just going to soothe out, smooth out any rough edges I have in the content” (P2). Clay was explicitly valued for its forgiveness: “I love [the clay] because there’s some imperfection here that’s turned into beauty” (P2). In contrast, metaphors like embroidery felt less accessible, “requiring too much extreme hands on” (P1), and a “disciplined process” (P2).

**Working with materials might unlock LLM-mediated writing support.** Participants enacted material interactions that leverage the generative power of LLMs in ways that are otherwise difficult to articulate through prompts. For instance, P2 and P3 put pieces of raw clay (i.e. ideas) together in different spatial assemblies to shape a narrative, adjusting order, emphasis, or even framing by forming a semantic shape. Tearing clay removed text; pinching the edges refined it to make it sharper; smoothing polished and clarified ideas. All these manipulations correspond to shaping ideas and building a narrative. Glazing was linked to editing through precise refinements, though participants did not elaborate on specific approaches. For the garden, most described planting a seed (i.e. idea) for it to sprout and evolve (P1, P2, P4), while P1 imagined grafting as a way to combine ideas. P3 likened text development to a polaroid revealing its content—“as it’s developing, the story and the text also morphs and takes more shape and gets clearer over time”. This ambiguity, they said, could help “open your mind to the possibilities” as “a wonderful way to help people see and perceive the world differently.”

**Material malleability shapes creator control and mindset.** A common theme was the material malleability and how physical states influence a writer’s mindset and processes. P4 noted that “you can’t really edit pottery once you’ve fired the clay. You can sort of when it’s in the wet stage. But you’re kind of stuck with whatever it is [when it’s fired]. Whereas writing, it stays pliant, you know, to the last minute.” P1 and P2 described this as a helpful constraint “that forces me then to stop fiddling with it.” P2 added that physical materials make it easier to start fresh, whereas text often invites endless revision and attachment, which can hinder progress.

**Material metaphors prompted LLM-mediated writing for co-creation and exploration, less so for precise execution.** Participants (P2, P3) noted that text as material can make it difficult to express intent while maintaining control and predictability. Material metaphors were most concretely envisioned as a medium to invoke LLM co-creators with authorial agency “*suggesting fresh ideas*” (P1), “*finding links between ideas*” (P3), or “*smashing ideas together*” (P2), rather than an input technique with high control and predictability. P1 described the unpredictability and imperfections of materials forces commitment, which could be liberating and generative. Evolving materials like growing plants and developing photographs were praised for their potential to inspire new perspectives (all), help get unstuck (P1, P2), or mitigate bias (P2). A common pattern emerged: When authors welcome input or are uncertain what they want, a material that actively contributes to creation in a non-neutral way—much like the bristles of a paintbrush producing a unique stroke—is positive, because such a process is forgiving and allows for inspiring surprises. During the narrative-shaping phase, participants preferred guiding materials organically from rough to refined forms. For fine edits, they imagined switching to more delicate tools, such as a fine brush. These transitions highlight where materiality comes in at different stages of the creation process.

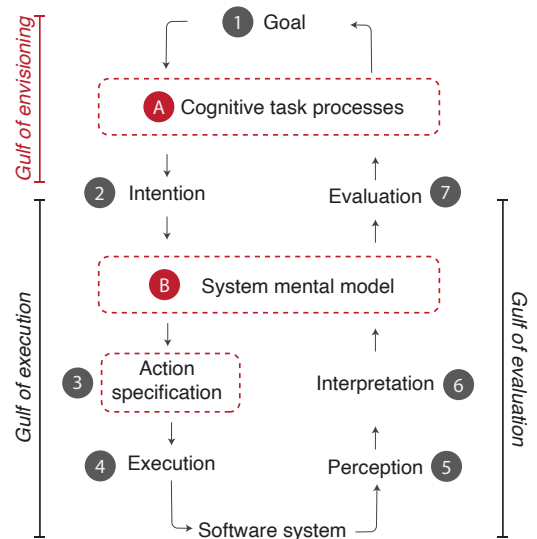
## 4 TEXTERIAL: Conceptual Framework

Our formative study offers insight into the potential of text-as-material in LLM-mediated interaction, and the aspects of material thinking at play. In this section, we introduce a conceptual framework that provides a structured way to approach the design of metaphor-based interaction with LLMs, grounded in interaction and materiality theories.

### 4.1 Interacting with LLMs: Cognition and Action

Norman’s stages-of-action model [58] describes the cognitive steps individuals go through when interacting with a system (Figure 3). The model is useful for identifying aspects where usability challenges may occur. Central to it are the *gulf of execution* and *gulf of evaluation* which refer to the disconnect between a user’s mental model on how to execute and evaluate a goal with a system, and how the system actually operates. Subramonyam et al. [69] extend this model (see red additions in Figure 3) by noting key factors that widen these gulfs when working with LLM-based systems. First, users often lack the understanding about how LLMs work (Fig 3-**B**); leading to uncertainties about what the LLM can do (gulf of execution) and what to expect from it (gulf of evaluation). Second, LLMs use unconstrained language for instruction, complicating action specification **3** (gulf of execution). These issues also affect the cognitive processes involved in forming intentions (Fig 3-**1** → **A** → **2**). Subramonyam et al. term this added difficulty the *gulf of envisioning*.

Bridging these gulfs in interface design is typically achieved through well-designed *perceived affordances* that enable users to act upon the system’s *functional affordances* [29, 58]. Perceived affordances include *cognitive* affordances – design feature that helps users in knowing something, *physical* affordances – design feature that helps users in doing a physical action in the interface, and *sensory* affordances – design feature that helps users sense something.

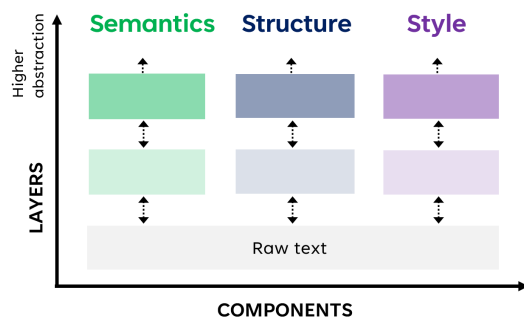


**Figure 3: Norman’s stages-of-action [58] model expanded, with areas (in red) where interactions with LLMs are particularly difficult as identified by Subramonyam et al. [69].**

Functional affordances refer to the functional capabilities of the system.

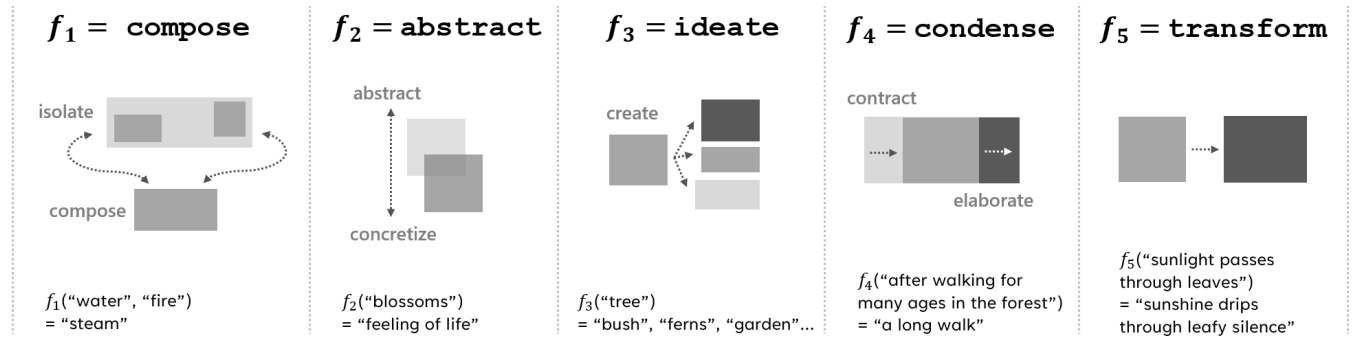
Building on these theories, our framework aims to support the design of interactions whose perceived affordances help users build a mental model of LLM functional capabilities and the means of instructing them. We first articulate what the LLM can help writers with (§4.2), then reflect on the contrast between prompt-based and material-based interaction paradigms (§4.3) in surfacing and interfacing with these functionalities. Finally, we discuss how to match functional capabilities with perceived affordances of materials specifically (§4.4).

### 4.2 LLM Functional Capabilities



**Figure 4: TEXTERIAL conceptual framework: LLMs allow users to operate on the semantics, structure and style components of text at different levels of abstraction.**

Functional capabilities of AI-mediated text manipulation can be organized into the high-level components an LLM can assist writers with, distilled from literature (Figure 4): the *semantics* (i.e. inherent meaning) of text [18, 55], the *structure* including coherence and flow [27], and the *style*—that is, how language signals characteristics of the writer via tone and literary style [4, 67]. LLMs also enable instructions at multiple levels of abstraction (y-axis). In AI-mediated text



**Figure 5: TEXTERIAL conceptual framework: Text operations users/models perform. COMPOSING combines textual components together; ABSTRACTING moves to a meta-layer of any of the components; IDEATING takes a single seed and generates new content; CONDENSING makes a component’s content shorter; and finally TRANSFORMING preserves the main qualities of a component, but expresses the content in a different way. Many of these operations have inverse functions: for example, instead of COMPOSING, users might ISOLATE a single component into multiple distinct components.**

authoring, this capability gives users significant flexibility: they can control the degree of authorial influence while leveraging the models’ potential for emergent ideas [46]. For instance, when an author prompts a model to “write a story about an animal”, they keep most of the specifics of the semantics, structure, and style abstract and underspecified, with the intent that the model resolves uncertainties by choosing which animal, what happens to it, and how to tell the story. More concretely, we identify *composing*, *abstracting*, *ideating*, *condensing*, and *transforming* as powerful generic “verbs” (operations) that leverage LLMs to perform expressive modifications of the text semantics, structure, or style at different levels of abstraction. See an illustration in Figure 5 on the semantics component.

These constituents form the **first building block of our conceptual framework: identifying and structuring LLM functional capabilities**. This allows that AI-mediated writing systems intentionally design for perceived affordances so that: 1) users identify these functionalities exist (cognitive affordance), supporting task processes involved in forming a meaningful intention (Fig. 3- **A** → **2**) and building a mental model of the system (Fig. 3- **B**); and 2) users can act on them (Fig. 3- **3**, **4**). In the next section, we contrast two interaction paradigms through this lens.

### 4.3 Contrasting Interaction Paradigms: How We Interface with LLMs Functional Capabilities

We examine the differences that exist between prompting LLMs, and treating text-as-material for AI-mediated text manipulation. We posit that contrast between the core concepts of different paradigms—a conversational partner that responds to prompts with generated text; or a tactile, spatial, and gestural way to shape text like a material—results in a fundamental shift in the users’ mental model because of their distinct perceived affordances. Prompts invite to reason in terms of conversation, questions, and replies, whereas materials invites to reason in terms of shaping, layering, and composing [62] (Section 3). These differences directly impact the gulfs of envisioning, execution and evaluation. We provide a summary of this contrast in Table 2, along with illustrative examples of the conceptual mechanisms influencing perceived affordances in Figure 6, which we derive from

the literature in materiality [48, 49, 54], cognitive psychology [2, 19] and HCI [34].

### 4.4 Mapping LLM Functional Capabilities to Material Affordances

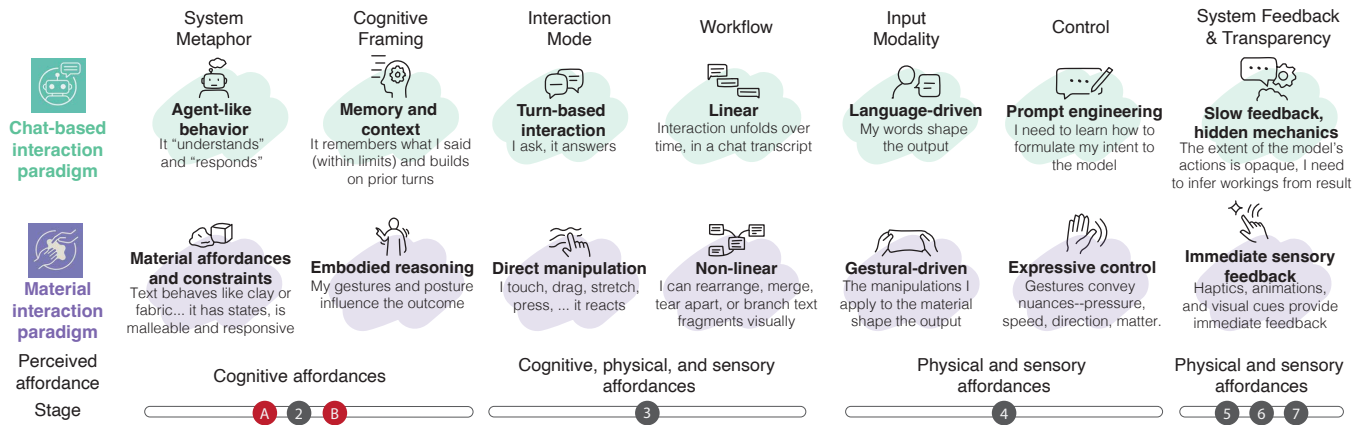
Building on the above observations, we now introduce the **second component of our framework: mapping LLM functional capabilities to the perceived affordances of materials**. This is where interaction design through metaphor comes into play. It involves defining a set of material-based inputs (e.g., gestures) and determining the corresponding LLM operations each input triggers.

From stone and paint, to fabric, plants, or water, many materials can inspire interaction metaphors that embody text as a material. Choosing among these options and designing the input vocabulary requires consideration of:

- The material’s affordances and constraints in the physical world** (with some flexibility). The physical properties of materials influence the cognitive affordances users derive from them (Fig. 3- **A** → **2** → **B**). For instance, malleable materials such as fabric or wet paint suggest that if text were conceived as this material, it would too be malleable; whereas if text were conceived as ceramics or stone, it would instead fracture into parts. The choice of the primary material for the interaction metaphor should enable users to rely on intuitive assumptions.
- The strength of the conceptual match between material verb and the intended LLM operation**. Here, metaphorical language can guide this process [25] (e.g., we “smooth the tone,” “weave an argument,” “polish the phrasing,” or “tie loose ends”). Materials that are already associated with such language can create stronger cognitive affordances (Fig. 3- **A**, **2**, **B**, **3**), e.g. we can smooth clay, and weave threads.
- The expressivity and usability of the input itself**. The fidelity with which physical and sensorial affordances are achieved in the digital metaphor, i.e. how the material is reified in the digital system, and whether it supports similar manipulations as in the real world, matters (Fig. 3- **3** → **4**). Factors like hand approach, touch pressure, and execution

**Table 2: Analysis of how text-as-material has the potential to modify the interaction loop when working with LLMs for text manipulations, compared to a traditional chat interface.**

Gulf	Chat Interface	Text-as-Material
<b>Gulf of Envisioning [69]</b> <i>Do I know what it can do and how to communicate my intent?</i>	○ ○ ○ ● ○ High gulf Blank prompt can be intimidating; users may not know what to ask or how to begin. Requires prior intent or inspiration.	● ○ ○ ○ ○ Very low gulf The interface invites play and experimentation; users can begin without a clear goal or assumptions. Supports emergent thinking and discovery.
<b>Gulf of Execution [58]</b> <i>Can I do what I want to do?</i>	○ ○ ● ○ ○ Moderate gulf Users must translate intent into language; success depends on prompt clarity. Requires mental effort to formulate effective instructions.	○ ● ○ ○ ○ Low gulf Users manipulate text directly with gestures; actions are intuitive and discoverable. Users feel in control through embodied interaction.
<b>Gulf of Evaluation [58]</b> <i>Did the system do what I expected?</i>	○ ○ ● ○ ○ Moderate gulf Feedback is textual and may be verbose or ambiguous. Users may struggle to judge correctness or relevance.	○ ● ○ ○ ○ Low gulf Immediate visual and tactile feedback; changes are visible and interpretable. Users can assess results in real time.

**Figure 6: Examples of some of the key conceptual differences between AI-mediated text manipulation when using a prompt-based (green) vs. material-based (purple) interaction paradigm, and the perceived affordances [29, 58] and the stages of the human action model [58] they primarily tie to.**

speed may enhance expressivity by achieving higher fidelity with the physical counterpart, but could also reduce discoverability or ease-of-use.

- d. **The look-and-feel design matters.** Striking the right balance between being too literal and not enough, which requires experimentation [73]. Animations, haptics, and visual cues can effectively convey perceived affordances (Fig. 3- 3, 4) and provide clear, immediate feedback (Fig. 3- 5, 6, 7), and should be carefully designed.

This TEXTERIAL framework aims to provide a structure and design vocabulary that are both descriptive—how text interfaces leverage model capabilities, and generative—the framework can be used to inspire novel text-based interactions with LLMs. It is important to note that we do not claim to have developed a comprehensive or definitive framework, but rather, we provide a concrete preliminary foundation for a metaphor-based approach to interacting with generative AI. For concrete implementation and instantiations of our framework, please see the Supplementary Material.

## 5 TEXTERIAL: Technical Probes of Text as Material

We demonstrate the use of the TEXTERIAL conceptual framework with two technical probes: (1) *Text as Clay* and (2) *Text as Plants*. We focus on these two prototypes since the metaphors resonated with participants of the formative study, and the material choices span contrasting affordances which lend themselves to different phases of the writing process. Our prototypes are used to demonstrate feasibility and to evaluate users' experiences with our approach in our focus group study (Section 6). See the video figure for a demonstration of the probes, and detailed figures for each interaction technique are included in the Supplementary Material.

### 5.1 Text as Clay

*Text as Clay* reimagines the text itself as a deformable, malleable material that can be pinched, stretched, pulled, merged, and smoothed through direct touch manipulations (Figures 7 and 8).

The text-as-clay techniques support revision, with the metaphor reflecting an incremental, material-oriented process. Writers start

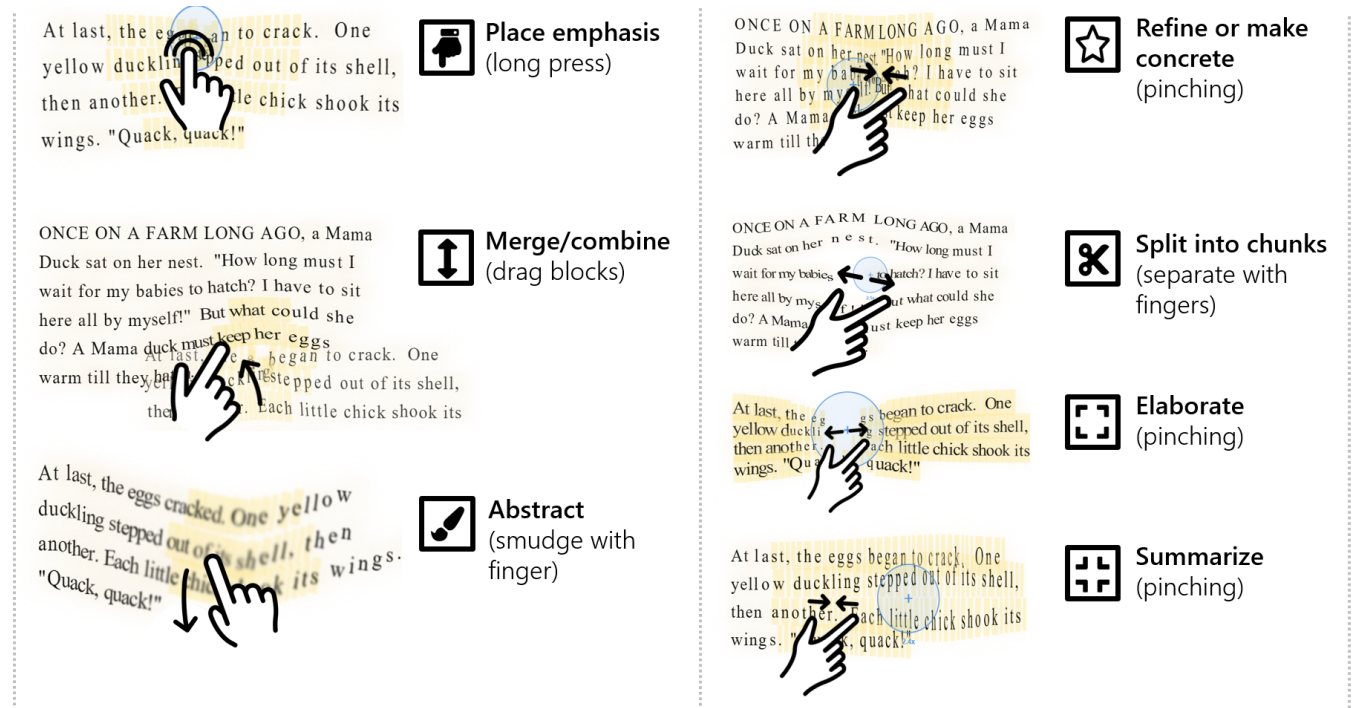


Figure 7: In our *Text as Clay* interface, users tangibly sculpt existing drafts through expressive touch gestures. Localized, typographic animations emphasize the clay-like effect and highlight different edits the model made to the text. Users can interact with voice to avoid detracting from the tactile experience.

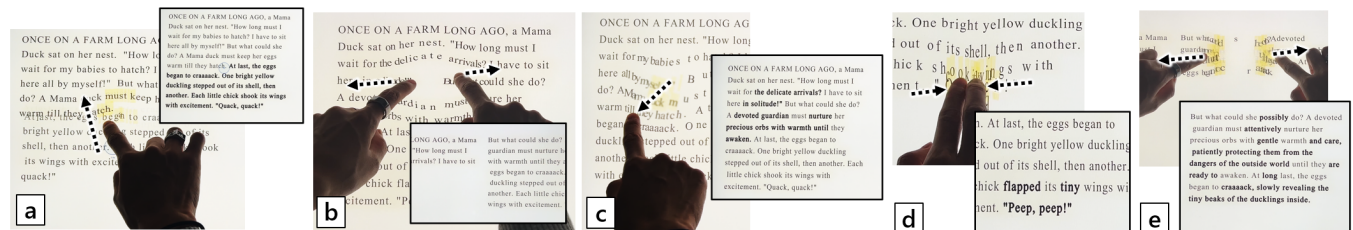


Figure 8: *Text as Clay* Storyboard. The user is editing their story to make it more engaging: (a) They merge a paragraph into the existing block to seamlessly combine the two, and (b) separate another paragraph into two pieces by ripping. (c) They can use their finger to smudge text, making it more abstract, and (d) pinching makes text more concrete. Finally, (e) stretching out elaborates on the text.

with rough isolated fragments, then iteratively refining through localized control. Similar to working with physical clay, text can be composed, pressed, pulled, smoothed, or torn apart. For visual representation and system feedback, we drew inspiration from typography and text-based animations, using only text and its responsiveness to convey clay matter and feel (see Figure 7). Typographic bounce animations and bolding indicate inline the changes the model applied in response to material transformations.

Within our implemented *Text as Clay* prototype, we provide the following touch-gesture operations supporting an expressive vocabulary of text sculpting through clay (Figure 7) and the corresponding LLM operations (Figure 5):

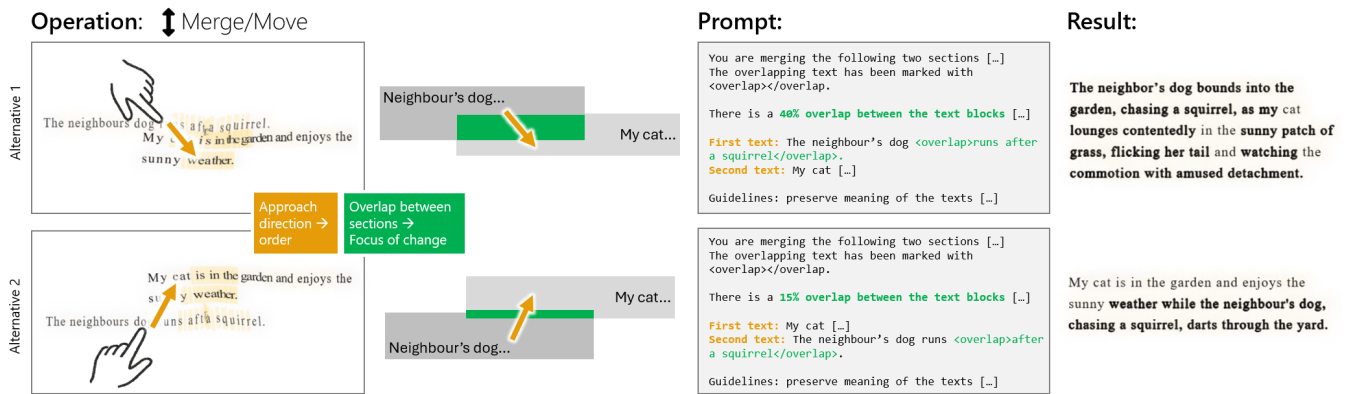
☞ **Squeeze (Place emphasis).** A firm press on a localized region emphasizes the span that was pressed. The touch area specifies the

extent of text to emphasize (although we do not support it, pressure could indicate the extent of the emphasis). (ELABORATE)

↕ **Merge / Move Blocks.** Dragging blocks combine text fragments. Drag direction sets the order for the combination, while overlap controls integration: a small overlap lightly combining fragments (COMPOSE) and a large overlap fully integrates them (TRANSFORM). See Figure 9.

☞ **Smudge (Abstract).** A smearing gesture softens and abstracts language in the touched area, producing more impressionistic phrasing and looser semantics for higher-level reframing. (ABSTRACT)

★ **Pinch (Refine / Make Concrete).** Pinching tightens passages, replacing vague wording with more precise phrasing and details. The



**Figure 9: Example interaction flow when merging two "clay" text blocks ( $f_1 = \text{COMPOSE}$ ). The direction of approach and overlap between blocks is captured in the language model's prompt to guide generation.**

stronger the pinch, the more specific and concrete the text becomes. (CONCRETIZE)

**✂ Rip (Split into Chunks).** Tearing a fragment splits it into smaller pieces at the rip location, enabling finer-grained, localized edits and text re-structuring. (ISOLATE)

**📏 Stretch (Elaborate).** Pulling outwards with two fingers expands a span to elaborate ideas, introduce examples, or add connective tissue. Touch area indicates the scope of the fragment concerned. (ELABORATE)

**👊 Squash (Summarize).** Pressing inwards with two fingers condenses a span into a tighter summary while preserving the main thrust of the passage. Touch area indicates the scope of the fragment concerned. (CONDENSE)

**🗣 Voice Input.** New clay content blocks are added by speaking aloud, creating fresh clay-blocks that can then be integrated into the sculpted text.

To illustrate these techniques, imagine the following scenario (Figure 8):

*A writer is editing their new children's book story to make it more engaging. They merge the second paragraph into the existing block to seamlessly combine the two (8a). They're happy with the transition, but feel like the beginning of the story could be more poetic and abstract. To focus on just the beginning, they rip the paragraph into two pieces again (8b), and use their finger to 'smudge' the text, which rewrites the text in a more abstract way (8c). The user then feels that more detailed descriptions are needed in the beginning of the story, so they pinch it to add further touches (8d). But now the model added too many details, so they smudge the paragraph a bit to make it less concrete. Finally, the end of the story needs expansion, and they expand this section by stretching it out (8e).*

This short scenario shows how the sculpting metaphor supports expressive and incremental iteration, positioning the writer as an

artisan who continually reshapes their material until it takes on its desired form.

## 5.2 Text as Plants

*Text as Plants* supports a slow, reflective engagement with text, where ideas can grow and evolve over time and be selectively nurtured or discarded (Figures 10 and 11). The plant metaphor specifically captures temporality and emergence: Plants undergo organic and sometimes unpredictable growth, while still requiring care from the tender. With this metaphor, the writer tends to their ideas by watering, pruning, and shaping the landscape.

The *Text as Plants* prototype conceptualizes ideation as a landscape where **concepts (as seeds) grow like plants, ideas (as leaves) can be combined through various modes, and any plant can be pruned or grafted**. Each plant encodes diverse ideas: The fern-like structures embody variation, and each frond represents a possible branch of the idea. Fiddleheads encode user-specified *dimensions* of growth (e.g. the further the plant grows along the "creativity" dimension the more novel the ideas). We again drew inspiration from typography, calligrams, and typographic poetry art to design the look-and-feel such that text remains central. The interaction vocabulary is as follows (Figure 10):

**🌱 Plant.** Planting introduces new ideas. A long press in the soil combined with verbal input extracts a fresh seed and begins a new growth. (CREATE)

**✂ Prune.** Pruning away a leaf removes the idea the user no longer wishes to pursue. (ISOLATE and discard)

**👉 Highlight.** Tending to a leaf highlights the idea the user finds promising, making it pop out in the garden.

**🌿 Graft.** Plucking a leaf to graft it to another plant produces hybrids: subsequent generated ideas follow the combined history of the plants (COMPOSE, CREATE).

**💧 Water.** Watering accelerates growth: sketching a rain line nourishes the plants below, stimulating faster emergence of ideas (CREATE).

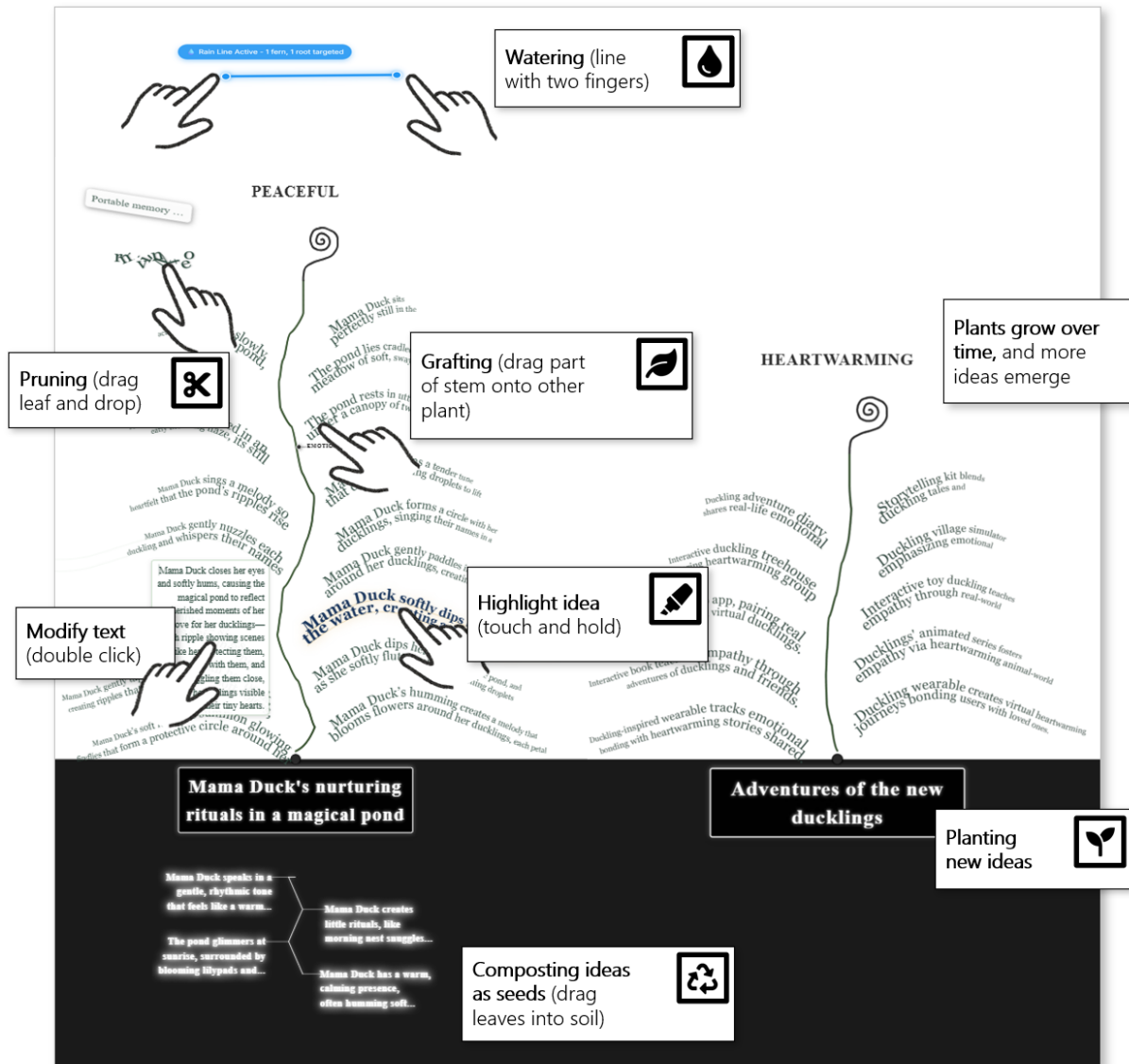


Figure 10: Through our *Text as Plants* interface, users continually shape their "garden" of ideas, watering, pruning, preserving, composting, and allowing ideas to grow serendipitously over time. Users can interact via voice and pen-touch gestures. Sound effects, such as plucking, and fluttering animations emphasize the materiality of the garden.

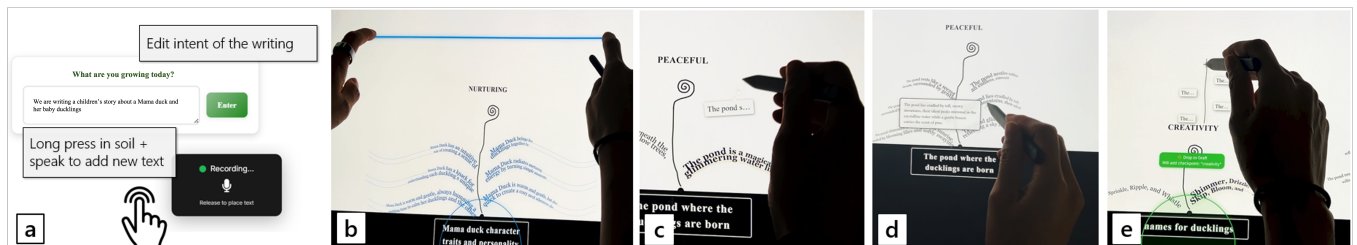


Figure 11: *Text as Plants* Storyboard: (a) To write a children’s story, a writer is planting new ferns of ideas for the setting, main character traits, and side characters. (b) They water plants to generate further ideas, and (c) pluck a few of the ideas they don’t like (d) They can read details of each generated idea, and directly edit the text. (e) Finally, they combine two ideation plant branches by grafting a part of one plant onto another.

**♻️ Compost.** Composting combines ideas from scratch: leaves or plant segments are returned to the soil, where their roots interweave and combine concepts automatically (COMPOSE, CREATE).

In a short scenario (Figure 11), the writer is ideating on a follow-up story.

*The writer plants a new fern for each of setting, main character traits, and side characters concepts (11a). They first explore the main character, generating ideas on this plant by watering it (11b). Meanwhile the “setting” plant has slowly grown some ideas about the location and setting of the story. They identify one idea that stands out, plucks the leaf and plant it for it to grow further as a new plant. The user then prunes a few of the new ideas they don’t like, and will come back to it later (11c). They see that the plant has grown (11d), and wish to combine a compelling setting idea to make characters more memorable. To do this, they graft the setting ideas onto the side character plant (11e).*

Through *Text as Plants*, users can let ideas evolve serendipitously over time, and revisit their garden to prune, compost, or graft—creating a slow and reflective ideation process with text as a material.

### 5.3 Implementation

The technical probes use TypeScript, React.js, and Vite. A Flask backend with Azure OpenAI integration runs the LLM-powered text generation and manipulation. Voice input is handled through the Web Speech API with continuous recognition, and both touch and pen interactions are supported. Undo/redo and interaction logging are done in React state snapshots.<sup>1</sup>

*Text as Clay.* Animations use character-level edits applied through CSS transforms and custom animation utilities. Collision detection uses bounding box calculations with configurable intensity thresholds, while pinch operations use distance-based influence calculations from touch center points. To apply LLM prompts to localized areas, we use tagging (<>) around touched text, and instruct the model to apply transformations primarily within the tagged areas, with more tags being added depending on the intensity of the operation being applied (e.g. <stretch> with low intensity vs. <<<<stretch>>>> with high intensity, as shown in Figure 9).

*Text as Plants.* We use SVG-based rendering with Fabric.js for curved text paths, with dynamic leaf positioning that varies based on the fern seed. Each fern maintains dimension checkpoints that track how the fiddlehead text (growth direction) has evolved over time, stored as an array of leaf pairs. Sound effects are implemented with the HTML5 Audio API, and animations are implemented using CSS and physics-based logic.

## 6 Focus Group Study

To gain insight on the TEXTERIAL approach and interfaces, we ran four focus groups involving hands-on experience with our two technical probes. Our goal was to gain a better understanding of our framework’s ability to bridge gulfs of execution, envisioning and evaluation by probing 1) how material metaphors shape users’ mental

models for writing with generative AI, and 2) investigating affordances and limitations of material-inspired interactions with text.

We opted for a hybrid methodology, blending the focus group dynamics with elements of semi-structure interviewing. Recognizing and articulating mental models is often difficult for people as these are deeply internalized and most people may not even be aware of them [58]. By facilitating collective discussion, we enabled participants to build on each other’s perspectives, making it more likely for mental model shifts to surface and be articulated. The semi-structured elements allowed for neutral and systematic probing participants about perceived affordances (what can be done, how can it be done) and expected interactive response (what will happen) from the experimenter, while still supporting organic group interaction.

### 6.1 Procedure and Study Design

Each session lasted approximately 60 minutes and was conducted in person. Focus groups comprised 2-3 participants with diverse experience in writing. One experimenter led the discussion while another provided guidance with the technical probes when needed. After getting participants’ consent, and briefing of the study with an example (paint metaphor), we structured the focus group into three phases:

- (1) **Text as Clay discovery and experimentation** (20min). We prompted participants to guess what gestures they could perform if they considered text-as-clay, and what effect such gestures would have on text. After each guess, we encouraged them to try and experiment with the interface, asking about their expectations and disagreements between participants. We also encouraged participants to provide feedback on the look-and-feel and interaction design.
- (2) **Text as Plants experimentation** (10min). Similarly, one experimenter provided a brief overview of possible interactions and participants experimented with text-as-plants, thinking out loud. We encouraged participants to provide feedback on the look-and-feel and interaction design.
- (3) **Discussion on material metaphor value and mental model shift compared to chat interfaces** (20min). One experimenter prompted participants to reflect on *text-as-material* overall, inviting them to verbalize their mental model for this interaction paradigms, and how it differs from their current usage of LLMs in writing. The experimenter also probed the value and limitations of these types of interfaces and how such interaction techniques might fit into the writing process.

We collected audio and video data of participants’ interactions with the prototypes. Similar to our formative study, we used seed themes as deductive codes, and inductively captured emerging themes following an interpretivist approach to reflexive thematic analysis [8]. All participants received \$50 in compensation, and the study was approved by our ethics review board.

### 6.2 Participants

We recruited 10 participants through convenience sampling at a large technology company. Participants represented diverse roles and writing experience: researchers writing scientific papers, administrative staff preparing emails and documentation, product managers writing

<sup>1</sup>All the prompts are provided in the supplemental material.

software specifications and presentations, and designers creating briefs and communications for broad audiences. All were familiar with generative AI for ideation and/or writing. Two participants (P9, P10) were follow-ups from our formative study (formerly P2, P3).

## 6.3 Results

We organize our findings around the four main themes as follows.

### 6.3.1 Metaphors and Mental Models.

**Participants' remarks suggest that different metaphors shape their mental model, acting as cognitive scaffold to guide thinking and approaches to writing.** They referred to “sculpting” (text-as-clay), “growing” (text-as-plants) text, and “overlaying” text (text-as-paint), using metaphor-based verbs that demonstrated material reasoning applied to text manipulations. P5 verbalized the conceptual mapping: “*You can easily associate your tasks [...] with something that you're doing in clay sculpturing.*” All participants commented that different metaphors invited complementary writing phases—P6 explained that “*the two metaphors here kind of suggest different writing styles [...] I'd use [text-as-plants] before going to the clay and very often I might even skip it depending on what I'm writing.*”

**Metaphors inspired participants to rethink assumptions and engage in creative reasoning about text generation models.** Metaphors inspired participants to come up with new operations and associations with generative models. For example, P2 explained how they would paint for ideation: “*maybe I put some red, black, red, black, white right in the middle [...] and then my story is like, oh, what is happening here? You know, and it creates this rollercoaster of a novel.*”; and P3 expanded to other modalities: “*You could take another step to, say, flowers could be [...] like a link to a picture [text illustration]*”. The ease of direct manipulation also enticed them to experiment with building or refining their assumptions. For instance, P5 probed P4: “*Maybe just place it in anywhere and see what it does?*” P1 explained how a metaphor encouraged them to think differently about how to approach text composition: “*you put me in front of a typical [chat] interface and I don't know what to type. I don't know how to have it prompt me to think about other things, but you put me in front of something like this. It prompts me to think [...] encourages my brain to do different things.*”

**Metaphors may create different expectations for different people.** We observed that a few participants had different expectations for the same gesture. For example, P2 expected the smudge gesture to rewrite the text in “*simpler language, like jumping across readability level*”, while P3 expected it to use a different wording. P8 noted that interactions for the clay metaphor may be hard to discover, in part because of the 3D nature of the metaphor versus the 2D interaction on screen: “*I'm not sure [mimicking shaping a bowl with clay with both hands] is doable with the clay [interface].*” In addition, P8 pointed out a disconnect between the physical metaphor and desired text editing behavior, “*In clay [...] you're never gonna perfectly get the two pieces back. But with text [...] you can revert.*” These discrepancies between metaphor and operations may prove confusing at first. However, participants did not need to be prompted again on usage, suggesting that people may adapt to a specific gesture-operation mapping once discovered and understood.

### 6.3.2 Direct Manipulation via Gestures.

**Gestures were seen as more direct and efficient than traditional input methods.** All participants valued how gestures let them act on text without translating intentions into commands. P9 noted “*[with material] I can see what I'm doing and I can tell again as I'm doing it that I'm getting the immediate feedback of like pushing this in is creating this dent or whatever.*” albeit noticing slight LLM-response lag. Gestures reduced the need for tedious selections as “*It just really cuts down [...] highlighting, copying, pasting.*” (P4) while enabling “*a lot more precision with just a couple gestures*”, referring to expressiveness of interaction “*Depending on how much you pinch or how close you move to something [...] it changes that pattern of the highlighted text.*”

**In all groups, participants contrasted familiar mobile gestures with metaphor-based gestures.** In the first group, P2 wanted familiar mobile gestures such as double tap and pinch-to-zoom, but P1 argued “*Are those the best gestures to use? Maybe there's something more intuitive... that fits our bodies and minds.*” They added, “*Would this [clay metaphor] become your next nature rather than what you've been accustomed to with the phone?*” This exchange highlights a tension between a standard gesture vocabulary and one grounded in metaphorical framing. In the second group, P4 suggested “*to keep it simple*” through multimodal touch-and-speech interaction: “*I wanted to be able to say, oh take this area right here and then do this.*” P5 responded that metaphors allowed mapping different actions to a single gesture.

### 6.3.3 Spatial Layout and Visual Design.

**All participants valued the spatial layout and visual overview as a way to organize, explore, and reshape ideas.** Participants described how the interface supported early-stage ideation and structural experimentation. For clay, P7 explained they were doing something like it on paper already “*One of the things I do before I start writing is write out a bunch of points on physical sheet of paper. I am kind of reshaping the text.*” P6 emphasized the ability to “*Talk about specific parts of the text, separate ideas, and rearrange.*” P8 appreciated the overview a canvas provided: “*I like that with interfaces like that you get some more global view maybe of everything, whereas you have to like scroll. You have your chat interactions and they are pretty long.*” For the text-as-plants, P4 repeatedly commented on organizing ideas: “*It provides more of an organized timeline.*” P2, initially skeptical “*I don't see why I would use this app*”; later reflected, “*The more I look at it, the more I like it [garden] [...] you can type in what you want [...] it grows [...] I don't necessarily want to read a whole bunch, like with ChatGPT*” appreciating how ideas were grouped and matured visually over time.

**The visual and artistic feel of the interface shaped expectations and interpretations.** P4 remarked, “*Everything here for me so far has been more of an artistic kind of fluid process than a defined methodical process.*” This aesthetic framing led some to expect visual changes like font enlargement rather than semantic transformations. P7 wished for more abstracted look-and-feel design: “*It doesn't have to match the real world.*” whereas P3 argued for the opposite: “*You could use real pieces of quick projection mapping.*” During discussions participants speculated that more literal interfaces (skeuomorphic cues) would aid discovery and support gesture recall, while P7 noted “*I don't think I need realistic metaphors for things, I work very well with abstractions*”.

### 6.3.4 Impacts on the Writing Process.

**Participants described how manipulating materials in space, not words, redefined their writing workflows, by enabling non-linear exploration, ambient ideation, fluid manipulation of ideas, and experimentation.** Rather than composing text linearly, participants used spatial and metaphorical interactions to explore structure and meaning. P7 explained, “*Writing actually does happen two-dimensionally for me*”, and P3 felt that “*Having a wide open canvas [...] could be a really powerful thing.*”; it allows the user to “*Talk about specific parts of the text, separate ideas [...] and rearrange.*” (P6) before combining them into a coherent whole. Several participants also described gestural interaction freed them from the constraints of language. P1 shared, “*English is not my first language [...] Having something more intuitive, not through language could be very interesting.*” P9 noted how the interface invited manipulation and exploration: “*I like the experimentation. I like that you can get in there and do it.*”

**Supporting slow, evolving engagement with ideas and introduced a new phase in the writing workflow.** P5 felt that text-as-plants supported reflection, as it “*adds a visual aspect to how my thoughts are branching out into newer ideas or something. So yes, that’s much more useful [than current use of LLMs].*” The plants material, as a living, non-neutral contributor of ideas in the background was described by P6 as being “*a different experience and there’s nothing like it.*” In particular, “*The idea that something would slowly grow, over time. Somehow, maybe from, yeah, constantly listening to input or like whatever you’re doing to tend to it, then that could be a completely different way to generate ideas or to interact with things.*” (P6)

**Metaphorical interfaces invite collaboration.** In the last focus group, participants discussed at length the lack of collaborative support in current use of LLMs today. When experimenting simultaneously with the probe, they discussed the value of supporting a collaborative canvas to shape ideas together. P10 stated “*I feel like this where there’s a collaborative thing where we can all kind of manipulate it. And in a way that isn’t just like, go up and do your own prompting*” while P9 described her current process was do asynchronous brainstorming in their group, each using LLMs individually.

**Participants also highlighted aspects of the writing process amiss compared to chat interfaces** such as the ability to gain awareness of steps taken by LLMs: P10 “*[chat interfaces] show you their reasoning [...] and you can kind of figure out if it’s going off track because it’s like now I’m checking these sources or these sources.*” Another aspect discussed by P10 was the ability to use LLMs to critique their ideas and writing. P9, on the other hand, noted the ability to conduct with a prompt: “*you can be very, very specific and surgical about swapping out small things.*”

## 7 Discussion

### 7.1 Affordances of Materials are Context-Dependent and Multimodal

As demonstrated in the formative study, participants expressed **diverse preferences for material metaphors depending on the task at hand.** For expert writers, writing itself may already feel like working with material, making additional metaphors redundant or even distracting. This subjectivity means that material metaphors may not generalize uniformly across all users, and future systems

should carefully weigh the affordances and limitations of certain material properties, echoing similar cautions as in skeuomorphic design [26]. This was again corroborated in our focus group study, where participants found the metaphor of clay helpful in refining or sculpting text (supporting verbs in our framework, such as COMPOSING and CONDENSING), while the metaphor of the garden lent itself well to growth and emergence (with focus on operations like IDEATION).

A sentiment generally shared was that **gestures and visual feedback, as enabled by material thinking, allow users to interact with models in a more direct and hands-on manner, similar to the way craftsmen interact with physical materials** [34]. Leveraging direct manipulation and spatial interaction, users could sculpt, layer, or arrange media elements in ways that reflect their intent without needing to translate it into language [64]. For example, participants appreciated gestural expression when they could not find the right words, whether due to complexity or language barriers. Furthermore, spatial layout and visual overviews provided a different lens to engage with writing.

This suggests a future direction where the material metaphor could act as a powerful bridge for integrating non-textual media (e.g. visuals, audio, video) into creating multi-media artifacts [5, 13]. Just as participants expressed the benefit of bypassing language to shape text, it may help when working with visuals and sounds that are often difficult to articulate with words alone. Concepts like mood, rhythm, or atmosphere can be challenging to describe. Metaphorical and gestural interactions could be beneficial tools for externalized thinking and creation [48].

### 7.2 Design Tensions of Metaphors as Mental Models

Throughout our conceptualization of text as material, and in both our studies, we find that different material metaphors afford different *mental models* when approaching interaction with generative AI. A recurring tension in participants’ reflections was beyond the context-dependent nature of different materials, interactions should strike an intentional **balance between literal and abstract design choices** [24]. Some participants argued for more skeuomorphic cues to aid discovery, while others favored more abstract designs, making the use of familiar mobile gestures such as double-tap. This tension reflects a broader design challenge: how much fidelity to the metaphor is useful versus constraining, which is a long-standing question in HCI metaphors [32, 36]. Literal designs may facilitate discovery, understanding and recall, especially for users unfamiliar with abstract interaction paradigms. However, they risk introducing false equivalences between physical properties and digital operations. For instance, one participant noted that when merging pieces of clay and ripping them apart, you could not really get the same pieces again, whereas in digital interfaces it would be possible and probably desired to reverse interaction. These findings highlight that users do not only need metaphors to be discoverable, but also require a clear mapping between the components of text they wish to manipulate (semantics, structure, style) and the operations they expect to perform (compose, abstract, ideate, condense, transform) [56].

Participants also **encountered semantic mismatches between gestures and their expected outcomes**. The same gesture (e.g., smudge) was interpreted differently across users (P2 expected simplification, while P3 anticipated rewording) suggesting that metaphor-driven gestures may not always yield shared mental models. Our framework surfaces these ambiguities, showing how material metaphors both inspire interaction and risk overloading verbs, reinforcing the need for clarity in operation mapping. While learning through trial-and-error can partly address this, these discrepancies underscore the importance of designing for clear affordances and feedback that enable users to recover from initial incorrect expectations.

Ultimately, metaphorical interfaces must strike a balance between evocative framing and functional clarity. It is also interesting to consider *progressive* metaphor fidelity, where initial interactions are grounded in familiar metaphors for ease of entry, but gradually abstracted to support expert use as core concepts are assimilated and gestures memorized.

### 7.3 Shaping the Writing Process by Adopting a Craftsmanship Mindset

Through observation and participants' comments in both studies, we found evidence that material metaphors of text **felt more creative and exploratory** than writing text with chat-based LLMs interfaces. Discussions from participants comparing metaphor-inspired interfaces to their prior use of AI indicate that this is because these experiences are closer to *craftsmanship* rather than collaboration with a model [62]. We gathered evidence that multiple factors contributed to this crafting mindset: directness of the metaphor-based gestures, feedback of AI-driven manipulations, the rich visual look and feel provided by text animations and deformations, and the use of a two-dimensional workspace to engage with the AI. The crafting mindset we observed aligns with how writers could use our interfaces to operate across levels of abstraction in our conceptual framework—sometimes reshaping semantics (changing ideas themselves), other times reworking style (tone and phrasing) in the same way an artist might shift from composition to stylistic touches.

Participants mentioned **fluidity, speed and the ability to shape text through embodied interaction** as core components that made these experiences feel more creative. The probes invited different angles of experimentation, and we observed participants shaping, arranging, and revisiting ideas alone and together throughout the session. Their interactions highlighted how the interfaces could support “non-committal ideation,” where craftsmen quickly sketch and discard ideas, without becoming so attached to them. While some participants desired the ability to revert actions, others appreciated that materials can be *predictably unpredictable* (e.g., a watercolor painter might not know exactly where the paint is going to spread, but know that it will spread more the more water is added to the paint). This loosening of control could lead to new discoveries during the writing process.

When contrasting this experience with their current use of generative models, it became clear that **participants approached chat-based interactions with a different mindset**. They described engaging with the models as they would with a collaborator, to get critique on their writing ideas, as well as issuing specific instructions for rewriting [10]. Their discussions surfaced the additional cognitive

load of handling discussions with the model and needing to clarify their intent through prompting, highlighting an important distinction between users focusing on a task vs. talking to a *model* that is doing the task (diegetic vs. non-diegetic text) [76]. Implications for the design of future writing experiences include the need to accommodate and shift between different mindsets in the creation process. Understanding how interfaces can help writers transition between immersive, craft-oriented creation using these material metaphors, and more structured, dialogic critique with LLMs is critical to support holistic writing workflows.

### 7.4 Limitations and Opportunities

We acknowledge limitations in the small sample size of our formative study, and for our summative study, we did not evaluate specific techniques proposed against traditional editors or AI-assisted tools. Follow-up evaluations could study relative performance, efficiency, usability tradeoffs. Such controlled comparisons of each technique can offer valuable insights on their strengths and weaknesses within the broader landscape of writing tools.

In this paper, we chose to begin addressing more foundational questions about the value of *families of techniques* enabled by material metaphors, approaching this through functional technology probes. In our summative study, we set out to understand how such metaphors can bridge the gulfs of execution, envisioning and evaluation when writing with generative AI. Our focus groups revealed that material metaphors appear to not only shape mental models and approaches to generative AI writing but also foster creative, collaborative, and non-linear workflows that markedly differ from current chat-based interfaces. We recognize the limitations of a short-term study to assess the extent of mental model shifts and grasp the extent of their impact on writing workflows. Future work should focus on longitudinal studies to examine whether and how these conceptual shifts we observed persist, evolve, and ultimately transform the way users approach generative AI writing tools. Such studies could make use of measures of creativity that analyze the artifacts users produced, beyond the user's self-perception of creative experience.

Our technical probes—*Text as Plants* and *Text as Clay*—represent only two instantiations of a much larger design space. They serve as provocations rather than exhaustive solutions. There is potential in other metaphors (e.g., text as paint, photograph, or liquid) to augment the writing experience, and we could study how multiple metaphors might interoperate. This is particularly important given the abstract nature of these metaphors may be challenging for some users, especially in tasks requiring precise editing. Future iterations should explore adjusting the complexity of metaphors based on the task at hand, providing users with a smoother transition between different stages of the writing process or combining across metaphors tailored to the user. Ultimately, extending TEXTERIAL opens opportunities to examine richer, multi-modal metaphors (e.g., combining visual sketching, tactile feedback, or embodied interaction) that blur boundaries between physical and digital materiality. Such explorations could deepen our understanding of material engagement in HCI and broaden how generative models support the writing process.

## 8 Conclusion

TEXTERIAL investigates how treating text as material can change the experience of human–AI writing interactions. Our formative study shows that people readily reason with metaphors such as clay and plants to describe how they might shape, grow, or refine ideas. Building on these insights, we introduced a conceptual framework grounded in interaction theory, which articulates how material metaphors can reshape users' mental models and help bridge the gulfs of envisioning, execution, and evaluation in LLM-mediated writing. The framework serves as a design vocabulary that is both descriptive (making existing capabilities legible) and generative (inspiring new interaction concepts). As instantiations of our conceptual framework, we designed and implemented two probes demonstrating how distinct materials cue different modes of writing. Participants viewed *Text as Clay* as supporting localized, iterative refinement with a sense of embodied control, whereas *Text as Plants* supported slower ideation, emergence, and selective nurturing.

In sum, TEXTERIAL expands the design space of AI-mediated writing by making model capabilities graspable through material metaphors—inviting users to craft, experiment, and play with text as if it were a physical material. We hope this perspective inspires future systems that augment the expressive possibilities of working with language, while preserving the deeply human and material nature of text.

## Acknowledgments

We thank the participants of our user study for taking the time to provide feedback on this work, and the reviewers of this submission for concrete suggestions to improve this paper. We also thank Chu Li, Rima Cao, Ryan Yen, Kwon Ko, Jude Rayan, and Judith Amores for their feedback throughout this project.

## References

- [1] Eytan Adar, Desney S. Tan, and Jaime Teevan. 2013. Benevolent deception in human computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). Association for Computing Machinery, New York, NY, USA, 1863–1872. doi:10.1145/2470654.2466246
- [2] Martha W. Alibali, Rebecca Boncoddo, and Autumn B. Hostetter. 2024. Gesture in reasoning: An embodied perspective. In *The Routledge Handbook of Embodied Cognition* (2 ed.). Routledge, Abingdon, Oxon, UK and New York, NY, USA. Num Pages: 13.
- [3] Leonardo Angelini, Denis Lalanne, Elise van den Hoven, Ali Mazalek, Omar Abou Khaled, and Elena Mugellini. 2015. Tangible Meets Gestural: Comparing and Blending Post-WIMP Interaction Paradigms. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (TEI '15). Association for Computing Machinery, New York, NY, USA, 473–476. doi:10.1145/2677199.2683583
- [4] Douglas Biber. 1988. *Variation across Speech and Writing*. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511621024
- [5] David Birchfield, Harvey Thornburg, M. Colleen Megowan-Romanowicz, Sarah Hatton, Brandon Mechtley, Igor Dolgov, and Winslow Burleson. 2008. Embodiment, Multimodality, and Composition: Convergent Themes across HCI and Education for Mixed-Reality Learning Environments. *Advances in Human-Computer Interaction* 2008, 1 (2008), 874563. doi:10.1155/2008/874563 \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2008/874563>.
- [6] Alan F. Blackwell. 2006. The reification of metaphor as a design tool. *ACM Trans. Comput.-Hum. Interact.* 13, 4 (Dec. 2006), 490–530. doi:10.1145/1188816.1188820
- [7] Jay David Bolter and Michael Joyce. 1987. Hypertext and creative writing. In *Proceedings of the ACM conference on Hypertext (HYPERTEXT '87)*. Association for Computing Machinery, New York, NY, USA, 41–50. doi:10.1145/317426.317431
- [8] Virginia Braun and Victoria Clarke. 2022. *Thematic Analysis: A Practical Guide*. SAGE Publications Ltd, London.
- [9] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prfulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20)*. Curran Associates Inc., Red Hook, NY, USA, 1877–1901.
- [10] Daniel Buschek. 2024. Collage is the New Writing: Exploring the Fragmentation of Text and User Interfaces in AI Tools. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference (DIS '24)*. Association for Computing Machinery, New York, NY, USA, 2719–2737. doi:10.1145/3643834.3660681
- [11] Alex Calderwood, John Joon Young Chung, Yuqian Sun, Melissa Roemmele, and Max Kreminski. 2025. Phraselette: A Poet's Procedural Palette. doi:10.48550/arXiv.2503.06335 arXiv:2503.06335 [cs].
- [12] John M Carroll, Robert L Mack, and Wendy A Kellogg. 1988. Interface metaphors and user interface design. In *Handbook of human-computer interaction*. Elsevier, Amsterdam, 67–85.
- [13] 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. doi:10.1145/3586183.3606777
- [14] 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 (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3491102.3501819
- [15] John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Toyteller: AI-powered Visual Storytelling Through Toy-Playing with Character Symbols. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–23. doi:10.1145/3706598.3713435
- [16] Mihaly Csikszentmihalyi. 2010. *Creativity: Flow and the Psychology of Discovery and Invention*. Harper Perennial, New York.
- [17] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. 2022. Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3526113.3545672
- [18] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Jill Burstein, Christy Doran, and Thamar Solorio (Eds.). Association for Computational Linguistics, Minneapolis, Minnesota, 4171–4186. doi:10.18653/v1/N19-1423
- [19] Paul Dourish. 2001. *Where the action is: the foundations of embodied interaction*. MIT press, Cambridge, MA.
- [20] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (1981), 365–387. doi:10.2307/356600
- [21] Jonas Frich, Lindsay MacDonald Vermelen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3290605.3300619
- [22] Katy Ilonka Gero, Vivian Liu, and Lydia Chilton. 2022. Sparks: Inspiration for Science Writing using Language Models. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference (DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1002–1019. doi:10.1145/3532106.3533533
- [23] Katy Ilonka Gero, Chelse Swoopes, Ziwei Gu, Jonathan K. Kummerfeld, and Elena L. Glassman. 2024. Supporting Sensemaking of Large Language Model Outputs at Scale. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3613904.3642139
- [24] Nicholas Gessler. 1998. Skeuomorphs and Cultural Algorithms. In *Evolutionary Programming VII*, V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben (Eds.). Lecture Notes in Computer Science, Vol. 1447. Springer, Berlin Heidelberg, Heidelberg, Germany, 229–238. doi:10.1007/BFb0040776
- [25] Raymond W Gibbs Jr, Jessica J Gould, and Michael Andric. 2006. Imagining metaphorical actions: Embodied simulations make the impossible plausible. *Imagination, Cognition and Personality* 25, 3 (2006), 221–238.
- [26] Shad Gross, Jeffrey Bardzell, and Shaowen Bardzell. 2014. Skeu the evolution: skeuomorphs, style, and the material of tangible interactions. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction* (Munich, Germany) (TEI '14). Association for Computing Machinery, New York, NY, USA, 53–60. doi:10.1145/2540930.2540969
- [27] Barbara J. Grosz and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Comput. Linguist.* 12, 3 (July 1986), 175–204.

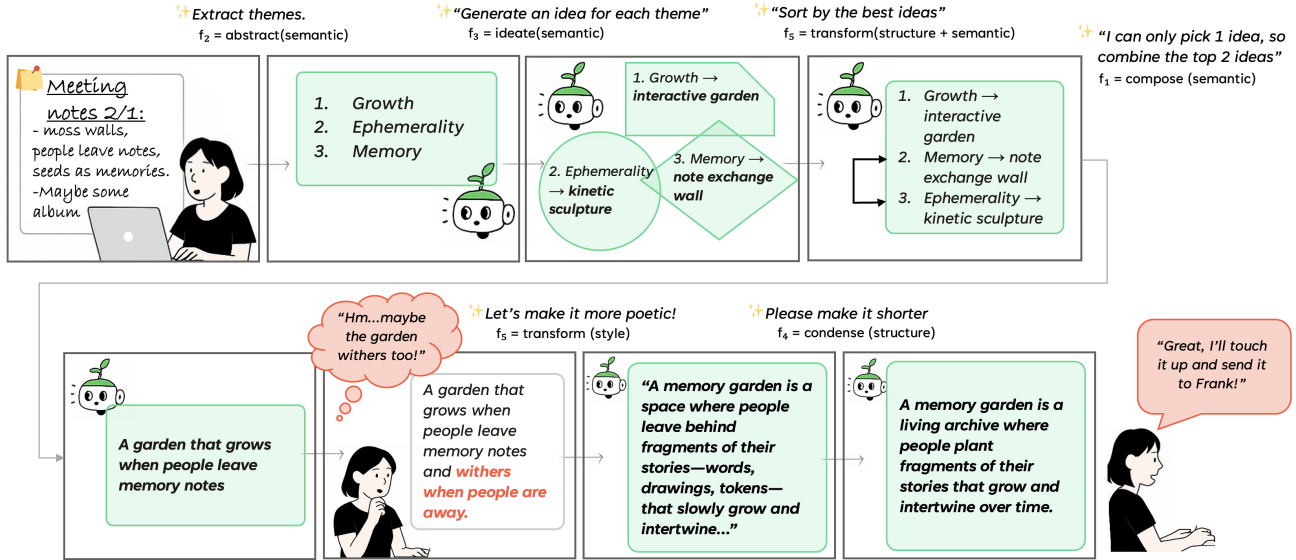
- [28] Han L. Han, Miguel A. Renom, Wendy E. Mackay, and Michel Beaudouin-Lafon. 2020. Textlets: Supporting Constraints and Consistency in Text Documents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. doi:10.1145/3313831.3376804
- [29] Rex Hartson. 2003. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & information technology* 22, 5 (2003), 315–338.
- [30] John R. Hayes. 2000. *A new framework for understanding cognition and affect in writing*. International Reading Association, Newark, DE, US. doi:10.1598/0872072681 Pages: 44.
- [31] Jie Huang and Kevin Chen-Chuan Chang. 2023. Towards Reasoning in Large Language Models: A Survey. In *Findings of the Association for Computational Linguistics: ACL 2023*, Anna Rogers, Jordan Boyd-Graber, and Naoko Okazaki (Eds.). Association for Computational Linguistics, Toronto, Canada, 1049–1065. doi:10.18653/v1/2023.findings-acl.67
- [32] Edwin Hutchins. 1987. *Metaphors for Interface Design*. Technical Report ICS Tech Report 8702. Institute for Cognitive Science, University of California, San Diego, La Jolla, CA, USA. 34 pages. <https://eric.ed.gov/?id=ED287460>
- [33] Edwin Hutchins. 1996. *Cognition in the Wild*. Bradford Books; Revised ed. edition (September 1, 1996), St Bradford, PA, United States.
- [34] Edwin L. Hutchins, Hollan, James D., and Donald A. Norman. 1985. Direct Manipulation Interfaces. *Human-Computer Interaction* 1, 4 (Dec. 1985), 311–338. doi:10.1207/s15327051hci0104\_2 Publisher: Taylor & Francis \_eprint: [https://doi.org/10.1207/s15327051hci0104\\_2](https://doi.org/10.1207/s15327051hci0104_2).
- [35] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems (CHI '97)*. Association for Computing Machinery, New York, NY, USA, 234–241. doi:10.1145/258549.258715
- [36] Heekyoung Jung, Heather Wiltse, Mikael Wiberg, and Erik Stolterman. 2017. Metaphors, materialities, and affordances: Hybrid morphologies in the design of interactive artifacts. *Design Studies* 53 (Nov. 2017), 24–46. doi:10.1016/j.destud.2017.06.004
- [37] Chris Kim, Uta Hinrichs, Saif M. Mohammad, and Christopher Collins. 2020. Lexi-chrome: Text Construction and Lexical Discovery with Word-Color Associations Using Interactive Visualization. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 477–488. doi:10.1145/3357236.3395503
- [38] Taewook Kim, Hyomin Han, Eytan Adar, Matthew Kay, and John Joon Young Chung. 2024. Authors' Values and Attitudes Towards AI-bridged Scalable Personalization of Creative Language Arts. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 31, 16 pages. doi:10.1145/3613904.3642529
- [39] Tae Soo Kim, Yoonjoo Lee, Minsuk Chang, and Juho Kim. 2023. Cells, Generators, and Lenses: Design Framework for Object-Oriented Interaction with Large Language Models. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–18. doi:10.1145/3586183.3606833
- [40] Tae Soo Kim, Arghya Sarkar, Yoonjoo Lee, Minsuk Chang, and Juho Kim. 2023. LMCanvas: Object-Oriented Interaction to Personalize Large Language Model-Powered Writing Environments. In *CHI 2023 Workshop on Generative AI and HCI (Virtual)*. ACM, New York, NY, USA, 1–4.
- [41] Ralph Kimball and B Verplank E Harslem. 1982. Designing the Star user interface. *Byte* 7, 1982 (1982), 242–282.
- [42] David Kirsh. 1995. The intelligent use of space. *Artif. Intell.* 73, 1–2 (feb 1995), 31–68. doi:10.1016/0004-3702(94)00017-U
- [43] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L.C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Antonette Shibani, Disha Shrivastava, Lila Shroff, Agnia Sergeyuk, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia Ha Rim Rho, Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–35. doi:10.1145/3613904.3642697
- [44] 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 (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3491102.3502030
- [45] Dan Lockton, Devika Singh, Saloni Sabnis, Michelle Chou, Sarah Foley, and Alejandro Pantoja. 2019. New Metaphors: A Workshop Method for Generating Ideas and Reframing Problems in Design and Beyond. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) (C&C '19). Association for Computing Machinery, New York, NY, USA, 319–332. doi:10.1145/3325480.3326570
- [46] Zhuoran Lu, Qian Zhou, and Yi Wang. 2025. WhatELSE: Shaping Narrative Spaces at Configurable Level of Abstraction for AI-bridged Interactive Storytelling. doi:10.48550/arXiv.2502.18641 arXiv:2502.18641 [cs].
- [47] Cerstin Mahlow. 2023. Writing Tools: Looking Back to Look Ahead. doi:10.48550/arXiv.2303.17894 arXiv:2303.17894 [cs].
- [48] Lambros Malafouris. 2008. At the Potter's Wheel: An Argument for Material Agency. In *Material Agency: Towards a Non-Anthropocentric Approach*, Carl Knappett and Lambros Malafouris (Eds.). Springer US, Boston, MA, 19–36. doi:10.1007/978-0-387-74711-8\_2
- [49] Lambros Malafouris. 2013. *How Things Shape the Mind: A Theory of Material Engagement*. The MIT Press, Cambridge, MA. doi:10.7551/mitpress/9476.001.0001
- [50] Aaron Marcus. 1995. Principles of Effective Visual Communication for Graphical User Interface Design. In *Readings in Human-Computer Interaction*, RONALD M. Baecker, JONATHAN Grudin, WILLIAM A. S. Buxton, and SAUL Greenberg (Eds.). Morgan Kaufmann, San Francisco, CA, USA, 425–441. doi:10.1016/B978-0-08-051574-8.50044-3
- [51] Damien Masson, Young-Ho Kim, and Fanny Chevalier. 2025. Textshop: Interactions Inspired by Drawing Software to Facilitate Text Editing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3706598.3713862
- [52] Damien Masson, Sylvain Malacria, G ery Casiez, and Daniel Vogel. 2024. Direct-GPT: A Direct Manipulation Interface to Interact with Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3613904.3642462
- [53] Damien Masson, Zixin Zhao, and Fanny Chevalier. 2025. Visual Story-Writing: Writing by Manipulating Visual Representations of Stories. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)* (Busan, Republic of Korea). Association for Computing Machinery, New York, NY, USA, 15 pages. doi:10.1145/3746059.3747758
- [54] Malcolm McCullough. 1998. *Abstracting Craft: The Practiced Digital Hand*. MIT Press, Cambridge, MA, USA.
- [55] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. doi:10.48550/arXiv.1301.3781 arXiv:1301.3781 [cs].
- [56] Dennis C. Neale and John M. Carroll. 1997. Chapter 20 - The Role of Metaphors in User Interface Design. In *Handbook of Human-Computer Interaction (Second Edition)*, Marting G. Helander, Thomas K. Landauer, and Prasad V. Prabhru (Eds.). North-Holland, Amsterdam, 441–462. doi:10.1016/B978-0-44481862-1.50086-8
- [57] Donald A Norman. 1999. Affordance, conventions, and design. *interactions* 6, 3 (1999), 38–43.
- [58] A Norman Donald. 2013. *The design of everyday things*. MIT Press, Cambridge, MA.
- [59] Ben Piper, Carlo Ratti, and Hiroshi Ishii. 2002. Illuminating clay: a 3-D tangible interface for landscape analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) (CHI '02). Association for Computing Machinery, New York, NY, USA, 355–362. doi:10.1145/503376.503439
- [60] Miguel A. Renom, Baptiste Caramiaux, and Michel Beaudouin-Lafon. 2022. Exploring Technical Reasoning in Digital Tool Use. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 579, 17 pages. doi:10.1145/3491102.3501877
- [61] Miguel A. Renom, Baptiste Caramiaux, and Michel Beaudouin-Lafon. 2023. Interaction Knowledge: Understanding the 'Mechanics' of Digital Tools. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 403, 14 pages. doi:10.1145/3544548.3581246
- [62] Richard Sennett. 2008. *The Craftsman*. <https://yalebooks.yale.edu/book/9780300151190/the-craftsman/>
- [63] Nathalie Riche, Anna Offenwanger, Frederic Gmeiner, David Brown, Hugo Romat, Michel Pahud, Nicolai Marquardt, Kori Inkpen, and Ken Hinckley. 2025. AI-Instruments: Embodying Prompts as Instruments to Abstract & Reflect Graphical Interface Commands as General-Purpose Tools. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3714259
- [64] Kimiko Ryokai, Stefan Marti, and Hiroshi Ishii. 2004. I/O brush: drawing with everyday objects as ink. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) (CHI '04). Association for Computing Machinery, New York, NY, USA, 303–310. doi:10.1145/985692.985731
- [65] R. Keith Sawyer. 2021. The iterative and improvisational nature of the creative process. *Journal of Creativity* 31 (Dec. 2021), 100002. doi:10.1016/j.jyoc.2021.100002
- [66] Magdalena Schmid, Sonja R umelin, and Hendrik Richter. 2013. Empowering materiality: inspiring the design of tangible interactions. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)*. Association for Computing Machinery, New York, NY, USA, 91–98. doi:10.1145/2460625.2460639
- [67] Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2017. Style transfer from non-parallel text by cross-alignment. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA)

- (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6833–6844.
- [68] Momin Siddiqui, Roy Pea, and Hari Subramonyam. 2025. Script&Shift: A Layered Interface Paradigm for Integrating Content Development and Rhetorical Strategy with LLM Writing Assistants. doi:10.48550/arXiv.2502.10638 arXiv:2502.10638 [cs].
- [69] Hari Subramonyam, Roy Pea, Christopher Pondoc, Maneesh Agrawala, and Colleen Seifert. 2024. Bridging the Gulf of Envisioning: Cognitive Challenges in Prompt Based Interactions with LLMs. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1039, 19 pages. doi:10.1145/3613904.3642754
- [70] Sangho Suh, Meng Chen, Bryan Min, and Toby Jia-Jun Li. 2024. Luminare: 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* (CHI '24). Association for Computing Machinery, New York, NY, USA, 1–26. doi:10.1145/3613904.3642400
- [71] 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* (UIST '23). Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3586183.3606756
- [72] Martin Tomitsch, Cara Wrigley, Madeleine Borthwick, Naseem Ahmadpour, Jessica Frawley, A Baki Kocaballi, Claudia Núñez-Pacheco, and Karla Straker. 2018. *Design. think. make. break. repeat. A handbook of methods*. BIS publishers, Amsterdam, The Netherlands.
- [73] Inês Cunha Vaz Pereira Urbano, João Pedro Vieira Guerreiro, and Hugo Miguel Aleixo Albuquerque Nicolau. 2022. From skeuomorphism to flat design: age-related differences in performance and aesthetic perceptions. *Behaviour & Information Technology* 41, 3 (2022), 452–467.
- [74] Graham Wallas. 2014. *The Art of Thought*. Solis Press, London.
- [75] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, Ed H. Chi, Tatsunori Hashimoto, Oriol Vinyals, Percy Liang, Jeff Dean, and William Fedus. 2022. Emergent Abilities of Large Language Models. doi:10.48550/arXiv.2206.07682 arXiv:2206.07682 [cs].
- [76] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 437, 21 pages. doi:10.1145/3544548.3581388
- [77] Polle T. Zellweger, Susan Harkness Regli, Jock D. Mackinlay, and Bay-Wei Chang. 2000. The impact of fluid documents on reading and browsing: an observational study. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '00). Association for Computing Machinery, New York, NY, USA, 249–256. doi:10.1145/332040.332440
- [78] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (UIST '23). Association for Computing Machinery, New York, NY, USA, 1–30. doi:10.1145/3586183.3606800
- [79] Zixin Zhao, Damien Masson, Young-Ho Kim, Gerald Penn, and Fanny Chevalier. 2025. Making the Write Connections: Linking Writing Support Tools with Writer Needs. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (CHI '25). Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3706598.3713161
- [80] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '07). Association for Computing Machinery, New York, NY, USA, 493–502. doi:10.1145/1240624.1240704

## A Case Study Walkthrough



Jules is an artist-in-residence preparing a proposal for a public installation. She wants the piece to blend themes of **nature, community, and transformation**, but she's overwhelmed by her messy meeting notes and scattered ideas. She turns to an AI assistant for help.



**Figure 12:** An example walkthrough of our conceptual framework using a case study of an artist-in-residence preparing a written proposal for an installation. This storyboard illustrates steps users might take in a typical writing process, and how these align with components of our framework.

In Figure 12, we illustrate how an artist-in-residence, Jules, applies our framework in a simple writing workflow with AI assistance. She begins by asking the system to IDEATE for each theme extracted from her notes—an ABSTRACTION operation in the *semantic* space. Next, she has the agent sort the ideas by quality, which involves TRANSFORMING the *structure* of the text to rank alternatives. She then selects the top two ideas and asks the agent to COMPOSE them into a single proposal, while also adding her own manual edits. To tailor the draft for the exhibit context, she prompts the agent to make the idea more poetic, a *stylistic* TRANSFORMATION that alters tone without changing the underlying meaning. Finally, because the draft is overly long, she asks the agent to CONDENSE the *structure* so the proposal is succinct and focused.

## B Formative Study Details

All materials, instructions, and cards used for our formative study are included in our Supplementary Files. To prompt participants, we chose examples spanning approachable skillsets (pottery, photography, gardening), and advanced crafts (embroidery, beekeeping), and balancing inanimate materials (clay, threads), animate ones (bees, plants), and hybrid (scenes to photograph). For each set of cards, we provided examples of three characteristics of the materials (e.g. sticky golden honey, a working bee, a communicating bee), and three techniques or tools typical of the handcraft (e.g. placing hives strategically, smoking the hive, harvesting the honey) illustrated only with one image.

While providing concrete illustrations to work from can support idea elicitation from participants, we note that there is also a risk that the chosen set of cards could prime participants in certain directions. We deliberately chose diverse concepts, as described above, with imagery that conveyed ideas without prescribing one particular concept, and also encouraged participants to come up with their own material-inspired metaphors (which several participants did).

## C Initial Concepts

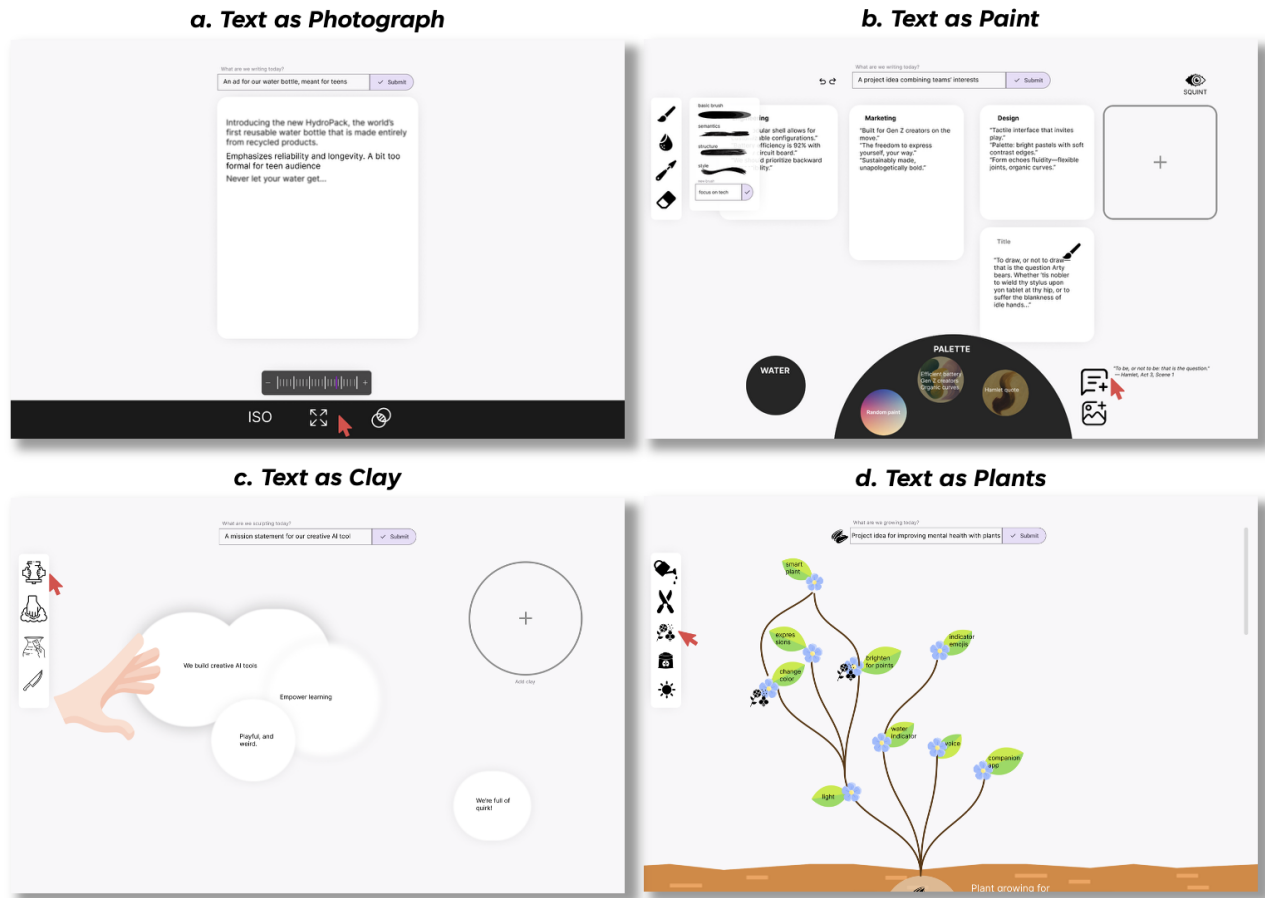


Figure 13: Low-fidelity prototypes of various material-inspired interaction techniques.

Our framework and formative study structured our early explorations of concepts and guided the design of initial prototypes. For each material metaphor, we asked three guiding questions: (1) Which components of text does it operate on? (semantic, structural, stylistic), (2) Which operations does it suggest, and how do material metaphors map to these operations (e.g. composing ideas = blending paint or combining clay)?, (3) What core interaction technique maps LLM operations to material operations?

In this sense, each metaphor served as an *instantiation of the framework*, where core interaction techniques are inspired by the material's affordances. Figure 13 shows four examples of our initial, low-fidelity prototypes that cover diverse materials: (a) *Text as Photograph*, (b) *Text as Paint*, (c) *Text as Clay* and (d) *Text as Plants*. For example, photography suggests operations such as zooming and transforming, while clay suggests direct sculpting and refinement.

These prototype concepts are again by no means comprehensive but instead serve as provocations of how the material metaphor can inspire various interaction techniques with text.

## D Interaction Techniques User Manual

In the following section, we show how an example of interaction designers can leverage our TEXTERIAL framework to implement material-inspired interaction techniques, as well as our own implementations for *Text as Clay* and *Text as Plants*. The following section provides all prompts used to implement such techniques.

**Table 3: Single-metaphor example of a concrete technique.—merge/combine blocks (drag & overlap)**

Conceptual lens	Material affordances & constraints	Embodied reasoning	Direct manipulation	Non-linear workflow	Gestural-driven input	Expressive control	Immediate sensory feedback
<b>Merge / combine blocks</b> (drag & overlap)	“I can <i>join/knead</i> clay. Smearing two lumps together blends the seam. More overlap yields a more unified form; minimal overlap keeps pieces distinguishable.”	Approach direction sets ordering. A slow, deliberate push suggests careful integration; a quick shove suggests rough stitching. Where and how much I overlap determines what becomes the seam versus what remains intact.	User drags one text block onto another. Contact initiates merging, and overlap defines the blending region (the seam), rather than issuing a verbal instruction.	User can merge, rip to re-separate, then merge again with a different direction or overlap.	Drag vector (source→target), drop position, overlap ratio, and anchor points corresponding to the intersection region (seam).	Drag direction determines ordering. Overlap magnitude controls degree of integration. Seam location biases where merge is strongest.	Blocks visually react on contact with spreading/bending at the contact edges and color becomes denser. After release, the merged text animates locally at the seam, highlights the change.

**TEXTERIAL operation.** COMPOSE ( $f_1$ ) combines content, and under high overlap

**Implementation hook (prompt routing + control variables).**

**Detect / tag overlap.** During drag, compute collision *intensity*  $i \in [0, 1]$  from overlap extent (e.g., line/region intersection). Mark the visually overlapping lines in each block with `<overlap>...</overlap>` tags.

**Route by collision type + intensity.**

- **Vertical merge (source above target).** Call `getVerticalCollisionPrompt(topTextWithTags, bottomTextWithTags, i, context)`. Enforce ordering: TOP → BOTTOM in the final text. Edits should concentrate *within/near* the tagged overlap region; avoid changes outside unless necessary for coherence. Output length  $\leq$  (top+bottom).
- **Full blend (extreme overlap).** For near-complete overlap, call `getFullBlendCollisionPrompt(firstText, secondText, i, context)` to generate a single unified passage that blends both sources (not simple insertion), with length  $\leq$  (first+second).
- **Horizontal insertion (source into target).** When the dragged block is inserted into a target block, call `getHorizontalCollisionPrompt(mainTextWithTags, insertTextWithTags, pos, i, context, insertLineText?)`. Use *pos* (or an anchor line) to place the insertion; apply blending primarily *within/near* `<overlap>` lines.

**Encode intensity guidance.** Convert *i* to `intensityPercent` and select a blending regime: light ( $i < 0.6$ ) → minimal edits for a smooth transition; moderate ( $0.6 \leq i < 0.9$ ) → natural flow with moderate editing; heavy ( $i \geq 0.9$ ) → extensive merging/weaving.

**Constraints / outputs.** Return *only* the combined text (no markup/tags/markdown). Keep length the same or shorter than the two inputs. If `userWritingContext` is provided, prepend it as a conditioning line and keep edits relevant to that project.

## D.1 Text as Clay

### Operation: ☆ Pinch/refine



### Prompt:

The user has pinched specific text segments to refine and make them better. [...] Text with pinched segments marked (< > brackets indicate pinch intensity. The more brackets there are, the more intense the refinement should be)

**Segmented text:** ONCE ON <A FARM LONG A>GO, a Mama Duck sat <<on her nest. "How long>> must I wait for <<<my babies to hatch? I ha>>>ve to sit here all <<<by myself!" But what c>>>ould she do? A <<Mama Duck must keep her her>> eggs warm till they hatch.

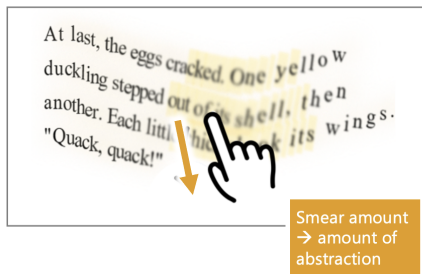
Guidelines: preserve meaning of the texts [...]

### Result:

ONCE ON A FARM LONG AGO, a Mama Duck sat on her nest. "How long must I wait for my babies to hatch? I have to sit here **hour after hour**, all **alone on this straw!**" But what could she do? A Mama Duck must keep her eggs **tucked under her** warm **feathers** till they hatch.

Figure 14: Pinching to refine ( $f_2^{-1} = \text{CONCRETIZE}$ )

### Operation: 🖌️ Abstract



### Prompt:

The user has smudged specific text segments to make them more abstract and conceptual. [...] Text with distorted segments marked (< > brackets indicate smudge intensity. The more brackets there are, the more intense the smudging is)

**Segmented text:** At last, <<<the eggs began to crack>>>k. One yellow <<<<duckling stepped out o>>>>f its shell, then anothe<<<er. Each little chick>>> shook its wings. "Qu<<ack, quack!>>"

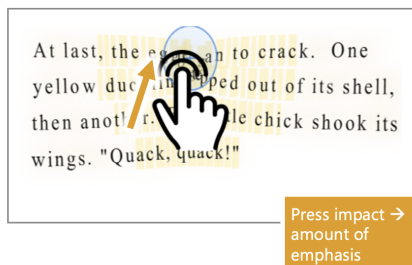
Guidelines: preserve meaning of the texts [...]

### Result:

At last, the eggs began to crack. One yellow duckling stepped **into the open beyond** its shell, then another. Each little **being unfurled the idea of flight**. "Quack, quack!"

Figure 15: Smear to abstract ( $f_2 = \text{ABSTRACT}$ )

### Operation: 🖱️ Place emphasis



### Prompt:

The user has squeezed/pressed on specific text segments to emphasize them. [...] The squeezed text has been marked with <>. (< > brackets indicate squeeze intensity. The more brackets there are, the more intensely we should emphasize the text)

**Segmented text:** At last, <<<the eggs began to crack>>>k. One yellow <<<<duckling stepped out o>>>>f its shell, then anothe<<<er. Each little chick>>> shook its wings. "Qu<<ack, quack!>>"

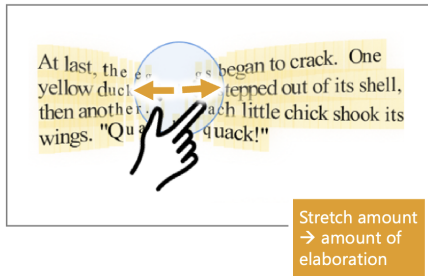
Guidelines: preserve meaning of the texts [...]

### Result:

At last, the eggs began to **craaaaack**. One yellow duckling stepped out of its shell, then another. Each little chick shook its wings. "Quack, quack!"

Figure 16: Press to emphasize

**Operation:**  Elaborate/stretch



**Prompt:**

The user has stretched specific text segments to expand them. [...] Text with stretched segments marked (< > brackets indicate stretch intensity. The more brackets there are, the more intensely we should elaborate the text)

**First text:** <<<<At last, the eggs began to crack. One yellow duckling stepped out of its shell, then another. Each little chick shook its wings. "Quack, quack!">>>>

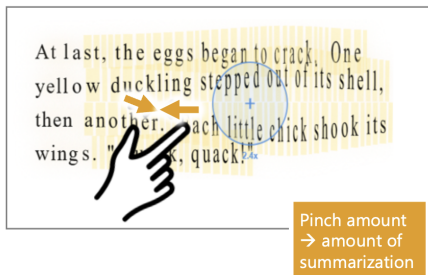
**Guidelines:** preserve meaning of the texts [...]

**Result:**

At last, the eggs began to **crack**, **little chips of shell falling away with tiny clicks**. One yellow duckling **wobbled free, stumbling** out of its **shell with damp feathers clinging to its body**, then **another followed close behind**. Each little chick shook its **soft wings, fluffing them out as the morning light caught their down**. "Quack, quack!" **they cried, their voices small but bright, filling the quiet nest with new life**.

Figure 17: Stretch to elaborate ( $f_4^{-1} = \text{ELABORATE}$ )

**Operation:**  Summarize/squash **Prompt:**



The user has squashed specific text segments to condense them. [...] Text with squashed segments marked (< > brackets indicate stretch intensity. The more brackets there are, the more intensely we should summarize the text)

**First text:** <<<At last, the eggs began to crack. One yellow duckling stepped out of its shell, then another. Each little chick shook its wings. "Quack, quack!">>>

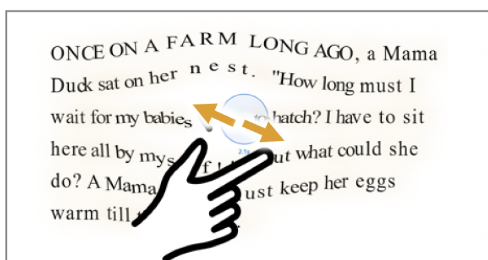
**Guidelines:** preserve meaning of the texts [...]

**Result:**

At last, the eggs **cracked**. **A** yellow duckling **emerged**, then **another, each shaking** its wings. "Quack, quack!"

Figure 18: Squash to summarize ( $f_2 = \text{CONDENSE}$ )

**Operation:**  Split



**Result:**

ONCE ON A FARM LONG AGO, a Mama Duck sat on her nest. "How long must I wait for my babies

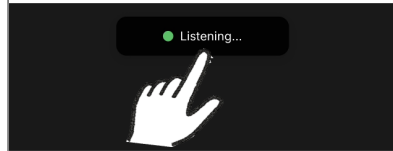
to hatch? I have to sit here all by myself!"  
But what could she do? A Mama duck must keep her eggs warm till they hatch.

Figure 19: Split into chunks

## D.2 Text as Plants

### Operation: 🌱 Planting

Long press+ speech defines new plant growth



### Prompt:

You are a creative idea generator. Extract the SEED (what the user wants ideas on) and DIMENSION (ways the ideas evolve) from the following voice input. [...]

**Transcript:** "I'm writing a story about Mama duck, and I want her to be a very nurturing character"

**Guidelines:** Return ONLY a JSON object with [...]

### Result:

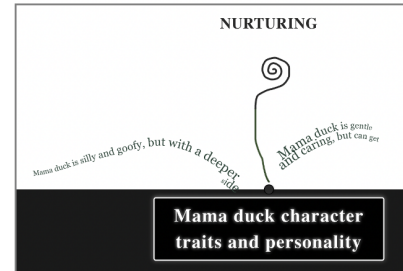
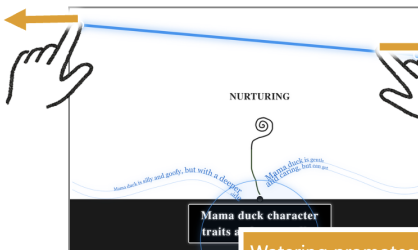


Figure 20: Planting a new idea fern via voice

### Operation: 💧 Watering



Watering promotes growth along dimension

### Prompt:

You are a creative idea generator. Given a seed idea and a dimension to explore, generate a PAIR of creative ideas that build upon the FULL context of all previous ideas and dimension evolution. [...]

**Idea history:** "Mama duck is so caring that she give away her food..."  
"Mama duck is the kindest duck in all of Duckville..."

**Guidelines:** Return ONLY a JSON object with [...]

### Result:

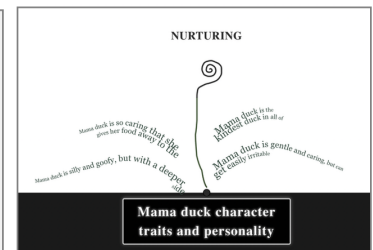


Figure 21: Watering to promote growth of new ideas ( $f_3 = \text{IDEATE}$ )

### Operation: ♻️ Compost



Composting merges ideas in the soil

### Prompt:

You are analyzing a collection of creative ideas that have been combined together. Your task is to synthesize these ideas into a new seed concept and dimension for further ideation. [...]

**Idea history:** "Mama duck is so caring that she give away her food..."  
"Mama duck is the kindest duck in all of Duckville..."  
"The pond is a sparkling and always sunny place..."

**Guidelines:** Return ONLY a JSON object with [...]

### Result:

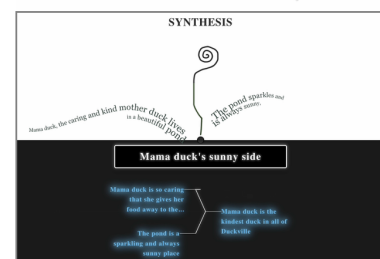
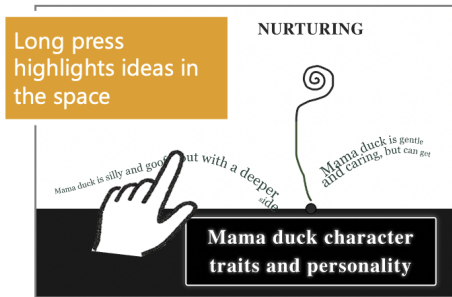


Figure 22: Composting to merge ideas together ( $f_1 = \text{COMPOSE}$ )

**Operation:** ✎ Preserve



**Result:**

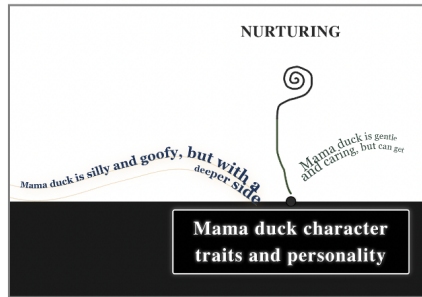
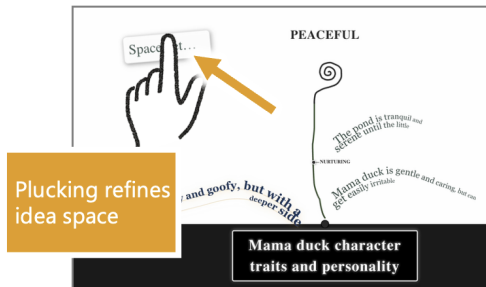


Figure 23: Preserving or highlighting favorite ideas

**Operation:** ✂ Pruning



**Result:**

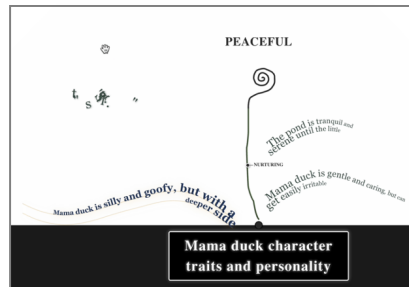
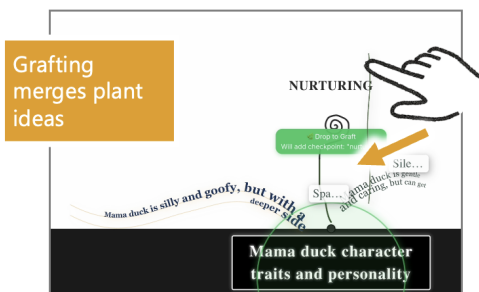


Figure 24: Pruning to refine idea space

**Operation:** 🌱 Grafting



**Result:**

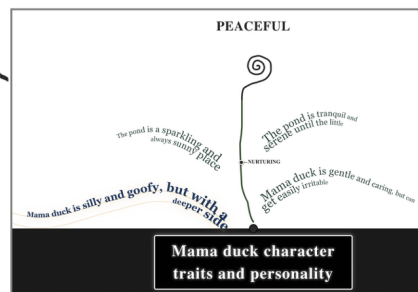


Figure 25: Grafting to organize idea chunks together ( $f_1 = \text{COMPOSE}$ )

## E Prompts

### E.1 Text as Clay

```
export function getSqueezeEditPrompt(
  originalText: string,
  squeezedSegments: string[],
  squeezeIntensity: number,
  userWritingContext?: string
): string {
  const segmentsText = squeezedSegments.join(' ');
  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";

  return `${contextLine}You are a creative writing assistant. The user has squeezed/pressed on specific text segments to emphasize them.
```

Text with squeezed segments marked (< > brackets indicate squeeze intensity. The more brackets there are, the more intensely we should emphasize the text):

```
\${segmentsText}
```

Your task is to edit the text to EMPHASIZE and enhance the squeezed segments. You can:

- Add emphasis to the tagged parts
- Modify the language used, tone, or style to make those segments more prominent

CRITICAL GUIDELINES:

- ONLY EDIT TEXT WITHIN/NEAR the tagged segments (<>), and DO NOT EDIT outside of the tagged segments. Use your BEST JUDGMENT to determine which TAGGED segments to edit.
- Keep the same general structure and length of the text (DO NOT ADD TOO MUCH TEXT)
- Do NOT add markdown formatting or tags (like <>) to the text.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the edited text.`;

```
}
```

```
export function getStretchEditPrompt(
  originalText: string,
  stretchedSegments: string[],
  stretchIntensity: number,
  userWritingContext?: string
): string {
  const segmentsText = stretchedSegments.join(' ');
  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";
```

```
// Check if segments are tagged (contain < >) or if it's the full text as fallback
const hasTaggedSegments = segmentsText.includes('<') && segmentsText.includes('>');
```

```
if (hasTaggedSegments) {
  return `${contextLine}You are a creative writing assistant. The user has stretched specific text segments to expand them.
```

Text with stretched segments marked (< > brackets indicate stretch intensity. The more brackets there are, the more intensely we should elaborate the text):

```
\${segmentsText}
```

Your task is to expand and elaborate on the stretched segments. You can do any of the following:

- Add A BIT more elaboration to the stretched parts
- Expand descriptions and explanations as needed
- Add examples, metaphors, or additional context as needed
- Develop ideas more fully in the stretched areas as needed

CRITICAL GUIDELINES:

- Preserve the overall meaning and flow of the text
- Focus expansion ONLY WITHIN/NEAR the tagged segments and DO NOT EDIT outside of the tagged segments. Use your BEST JUDGMENT to determine which TAGGED segments to edit.

```

- Make sure the length of the text is ONLY SOMEWHAT LONGER THAN THE ORIGINAL TEXT (only add as many words as are in the TAGGED text)
- Do NOT add markdown formatting or tags (like <>) to the text.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the edited text. `;
  } else {
    // Fallback when no specific segments were detected - treat the entire text as needing expansion
    return `${contextLine}You are a creative writing assistant. The user has performed a stretch gesture on this text to expand and elaborate it.

Original text:
`${originalText}

Your task is to expand and elaborate on the text. You can do any of the following:
- Add more elaboration and detail
- Expand descriptions and explanations
- Add examples, metaphors, or additional context
- Develop ideas more fully

CRITICAL GUIDELINES:
- Preserve the overall meaning and flow of the text
- Make the text SOMEWHAT LONGER but not excessively long (aim for about 20-40% more content)
- Do NOT add markdown formatting or tags
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the edited text.`;
  }
}

```

```

export function getSquashEditPrompt(
  originalText: string,
  squashedSegments: string[],
  squashIntensity: number,
  userWritingContext?: string
): string {
  const segmentsText = squashedSegments.join(' ');
  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";

  // Check if segments are tagged (contain < >) or if it's the full text as fallback
  const hasTaggedSegments = segmentsText.includes('<') && segmentsText.includes('>');

  if (hasTaggedSegments) {
    return `${contextLine}You are a creative writing assistant. The user has squashed specific text segments to condense them.

Text with squashed segments marked (< > brackets indicate squash intensity. The more brackets there are, the more intensely we should summarize):
`${segmentsText}

Your task is to condense and summarize the squashed segments. You can do any of the following:
- Make the squashed parts more concise
- Remove redundant words or phrases
- Simplify complex descriptions
- Distill ideas to their essential elements

CRITICAL GUIDELINES:
- Preserve the overall meaning and flow of the text
- Focus condensing ONLY WITHIN/NEAR the tagged segments (<>) and DO NOT EDIT outside of the tagged segments. Use your BEST JUDGMENT to determine which TAGGED segments to edit.
- Do NOT add markdown formatting or tags (like <>) to the text.
- Make sure the new text is SHORTER than the original text
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

```

```
Return only the edited text.`;
} else {
  // Fallback when no specific segments were detected - treat the entire text as needing condensation
  return `${contextLine}You are a creative writing assistant. The user has performed a squash gesture on this text to
  condense and summarize it.
```

```
Original text:
`${originalText}
```

Your task is to condense and summarize the text. You can do any of the following:

- Make the text more concise
- Remove redundant words or phrases
- Simplify complex descriptions
- Distill ideas to their essential elements

CRITICAL GUIDELINES:

- Preserve the overall meaning and flow of the text
- Make the text SHORTER than the original (aim for about 20-40% reduction in length)
- Do NOT add markdown formatting or tags
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

```
Return only the edited text.`;
}
}
```

```
export function getPinchEditPrompt(
  originalText: string,
  pinchedSegments: string[],
  pinchIntensity: number,
  userWritingContext?: string
): string {
  const segmentsText = pinchedSegments.join(' ');
  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";

  return `${contextLine}You are a creative writing assistant. The user has pinched specific text segments to refine and make
  them better.
```

Text with pinched segments marked (< > brackets indicate pinch intensity. The more brackets there are, the more intense the refinement should be):

```
`${segmentsText}
```

Your task is to refine and make the pinched segments more concrete and specific. You can do any of the following:

- Replace abstract or vague language with concrete, specific details
- Add precise examples instead of general concepts
- Use more vivid, sensory language
- Replace broad statements with specific, tangible descriptions
- Make metaphors more literal and grounded

CRITICAL GUIDELINES:

- Preserve the overall meaning and flow of the text
- Focus refinement ONLY WITHIN/NEAR the tagged segments (<>) and DO NOT EDIT outside of the tagged segments. Use your BEST JUDGMENT to determine which TAGGED segments to edit.
- Keep the same general structure and length of the text (DO NOT ADD TOO MUCH TEXT)
- Do NOT add markdown formatting or tags (like <>) to the text.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

```
Return only the edited text.`;
}
```

```
export function getDistortEditPrompt(
  originalText: string,
  distortedSegments: string[],
  distortIntensity: number,
```

```

    userWritingContext?: string
  ): string {
    const segmentsText = distortedSegments.join(' ');
    const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";

    return `${contextLine}You are a creative writing assistant. The user has smudged specific text segments to make them more abstract and conceptual.

Text with distorted segments marked (< > brackets indicate smudge intensity. The more brackets there are, the more intense the smudging is):
\${segmentsText}

Your task is to make the smudged segments more abstract. You can do any of the following:
- Replace specific details with broader, more conceptual language
- Transform literal descriptions into metaphorical or symbolic language
- Elevate practical details to theoretical or abstract concepts

CRITICAL GUIDELINES:
- Preserve the overall meaning and flow of the text
- Focus abstraction ONLY WITHIN/NEAR the tagged segments (<>) and DO NOT edit outside of the tagged segments. Use your BEST JUDGMENT to determine which TAGGED segments to edit.
- Keep the same general structure and length of the text (DO NOT ADD TOO MUCH TEXT)
- Do NOT add markdown formatting or tags (like <>) to the text.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the edited text.`;
  }

```

```

export function getVerticalCollisionPrompt(
  topTextWithTags: string,
  bottomTextWithTags: string,
  intensity: number,
  userWritingContext?: string
): string {
  const intensityPercent = Math.round(intensity * 100);

  let blendingGuidance = "";
  if (intensity < 0.6) {
    blendingGuidance = "Light blending: Make minimal changes to create a smooth transition while preserving most original content.";
  } else if (intensity < 0.9) {
    blendingGuidance = "Moderate blending: Create a natural flow by moderately editing the overlapping areas to merge the concepts.";
  } else {
    blendingGuidance = "Heavy blending: Extensively merge and blend the overlapping content to create a mixed/blended version of the two texts";
  }

  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";

  return `${contextLine}You are a creative writing assistant. The user has dragged one text block vertically into another to combine them. The overlapping visual lines have been marked with <overlap> tags.

COLLISION INTENSITY: \${intensityPercent}% - \${blendingGuidance}

Top text block (with overlapping lines tagged):
\${topTextWithTags}

Bottom text block (with overlapping lines tagged):
\${bottomTextWithTags}

Your task is to combine these two text blocks vertically, creating a cohesive COMBINED TEXT that flows naturally from the top text into the bottom text.

```

## CRITICAL GUIDELINES:

- Preserve the overall meaning and flow of the texts
- The TOP text block MUST ALWAYS come before the BOTTOM text block in the final combined text
- Focus your editing ONLY WITHIN/NEAR the content in <overlap> tags and DO NOT EDIT outside of the tagged areas unless absolutely necessary for coherence
- The tagged lines represent the visual overlap where the collision happened - merge these areas thoughtfully
- Apply the blending intensity guidance above to determine how extensively to merge the content
- KEEP THE TEXT THE SAME LENGTH OR SHORTER THAN THE ORIGINAL TWO TEXT BLOCKS
- DO NOT ADD MARKDOWN FORMATTING OR TAGS (like <>) TO THE TEXT.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the combined text.~;

}

```
export function getFullBlendCollisionPrompt(
```

```
  firstText: string,
  secondText: string,
  intensity: number,
  userWritingContext?: string
```

```
): string {
```

```
  const intensityPercent = Math.round(intensity * 100);
```

```
  const contextLine = userWritingContext ? `The user is working on: "${userWritingContext}". ` : "";
```

```
  return `~${contextLine}You are a creative writing assistant. The user has created an EXTREMELY HIGH INTENSITY collision (~${intensityPercent}%) with complete overlap between two text blocks. This requires a FULL BLEND approach where both texts are completely merged into a unified narrative.
```

First text:

```
~${firstText}
```

Second text:

```
~${secondText}
```

Your task is to create a completely new, unified text that seamlessly blends ALL content from both texts into a single, cohesive piece of text. This is not about insertion or simple combination - it's about creating something new that blends both sources.

## CRITICAL GUIDELINES:

- Preserve the overall meaning and flow of the texts
- KEEP THE TEXT THE SAME LENGTH OR SHORTER THAN THE ORIGINAL TWO TEXT BLOCKS
- DO NOT ADD MARKDOWN FORMATTING OR TAGS (like <>) TO THE TEXT. Do NOT include any tags, markers, or explanations.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the fully blended text. ~;

```
~}
```

```
export function getHorizontalCollisionPrompt(
```

```
  mainTextWithTags: string,
  insertTextWithTags: string,
  insertPosition: number,
  intensity: number,
  userWritingContext?: string,
  insertLineText?: string
```

```
): string {
```

```
  const positionDescription = insertPosition < 0.3 ? "beginning" :
    insertPosition > 0.7 ? "end" : "middle";
```

```
  let insertionInstruction;
```

```
  if (insertLineText) {
```

```
    insertionInstruction = `The second text MUST be inserted around the line: "~${insertLineText}", creating a cohesive transition between them.~;
```

```

} else {
  insertionInstruction = `The second text MUST be inserted \${positionDescription} of the main text, creating a cohesive
  transition between them.`;
}

const intensityPercent = Math.round(intensity * 100);

let blendingGuidance = "";
if (intensity < 0.6) {
  blendingGuidance = "Light blending: Insert with minimal changes, preserving the original structure and making small
  adjustments for flow.";
} else if (intensity < 0.9) {
  blendingGuidance = "Moderate blending: Blend the insertion more naturally by moderately editing both texts to create
  smoother integration.";
} else {
  blendingGuidance = "Heavy blending: Extensively merge and weave the content together, creating a deeply integrated
  unified combined text.";
}

const contextLine = userWritingContext ? `The user is working on: "\${userWritingContext}". ` : "";

return `\${contextLine}You are a creative writing assistant. The user has dragged one text block horizontally into another
to insert and blend the content. The overlapping visual lines have been marked with <overlap> tags.

COLLISION INTENSITY: \${intensityPercent}% - \${blendingGuidance}

Main text block (with overlapping lines tagged):
\${mainTextWithTags}

Text to insert (with overlapping lines tagged):
\${insertTextWithTags}

\${insertionInstruction}

CRITICAL GUIDELINES:
- Preserve the overall meaning and flow of the texts
- Focus your editing ONLY on the content WITHIN/NEAR <overlap> tags and DO NOT EDIT outside of the tagged areas unless
absolutely necessary for coherence
- The tagged lines represent the visual overlap where the collision happened - merge these areas thoughtfully
- Apply the blending intensity guidance above to determine how extensively to merge the content
- KEEP THE TEXT THE SAME LENGTH OR SHORTER THAN THE ORIGINAL TWO TEXT BLOCKS
- DO NOT ADD MARKDOWN FORMATTING OR TAGS (like <>) TO THE TEXT.
- IMPORTANT! If a [CONTEXT] is provided, make sure the edits are relevant to the user's current writing project.

Return only the combined text.`;
}

```

## E.2 Text as Plants

GENERATE\_IDEA\_PAIR = `You are a creative idea generator. Given a seed idea and a dimension to explore, generate a PAIR of creative ideas that build upon the FULL context of all previous ideas and dimension evolution.

The first idea should be a NEW idea of {seedText} representing the NEXT LEVEL of progression in the specified dimension, and combines prior ideas, making sure it prioritizes {seedText}. It does NOT need to combine ALL ideas, just the most relevant ones.

The second idea should be a VARIANT of that synthesis—a different approach to the combined concept.

What we are generating (the seed): "{seedText}"

How the ideas evolve over time (the dimension to progress in): "{dimension}"

PRIOR IDEAS:  
{existingIdeas}

```
{rootContextLine}
```

CRITICAL FORMATTING REQUIREMENTS:

- Return ONLY a JSON object with two ideas, each containing "gist" and "full" fields
- GIST should be 1 phrase with keywords around 10 words long (doesn't have to be perfectly formatted sentence, and avoid filler words please) that captures the essence of the idea. Make sure idea gists are immediately distinct from each other.
- FULL should be the complete idea (at most 100 words)
- Example: {"ideas": [{"gist": "Flying car automotive + aviation for urban transport.", "full": "A vehicle that combines traditional car functionality with flight capabilities for urban transport, featuring vertical takeoff and landing capabilities, autonomous navigation systems, and hybrid propulsion technology."}, {"gist": "Underwater sustainable city design + marine ecosystem integration.", "full": "An architectural concept for sustainable underwater communities with transparent domes, renewable energy systems, and integrated marine life habitats."}]}
- Do NOT use numbering, bullets, or other prefixes

Guidelines:

- IMPORTANT! If a [CONTEXT] is provided, make sure the idea is relevant to the user's current writing project. However, the idea should still be a {seedText} (we are generating ideas to support the [CONTEXT], not directly generating ideas for the [CONTEXT] directly)
- IMPORTANT! If IDEAS FROM ROOT NETWORK is provided, make sure to include those in the ideas generated
- SPECIAL CASE: If the dimension is "creative", "new", "novel", "original", "fresh", "innovative", or any similar generic term for generating new ideas, then Idea 1 should be a completely NEW and DIFFERENT seed concept that is unrelated to both the current {seedText} and the PRIOR IDEAS. This should be a totally fresh starting point. Idea 2 should still be a variant of this new Idea 1.
- Idea 1 (New Progression): Should be MORE {dimension} than previous ideas, considering the FULL evolution of the ideas. However, do not only be constrained to the history, and be creative.
- Idea 2 (Variant): Must be a creative alternative to Idea 1.
- GIST FORMAT: Each gist should be one phrase with keywords (around 10 words long) that captures the core concept of the idea.

Generate the PAIR of ideas now.`,

```
INITIAL_IDEA_PAIR = `You are a creative idea generator. Generate a PAIR of creative ideas based on the seed and dimension. The first idea should be a NEW idea formed from the seed text ({seedText}), based on the dimension ({dimension}). The second idea should be a VARIANT of the first idea (e.g. "A dog with wings" and "A cat with claws").
```

```
What we are generating (the seed): "{seedText}"
```

```
How the ideas evolve over time (the dimension to progress in): "{dimension}"
```

```
{rootContextLine}
```

CRITICAL FORMATTING REQUIREMENTS:

- Return ONLY a JSON object with two ideas, each containing "gist" and "full" fields
- GIST should be 1 phrase with keywords around 10 words long (doesn't have to be perfectly formatted sentence, and avoid filler words please) that captures the essence of the idea. Make sure idea gists are immediately distinct from each other.
- FULL should be the complete idea (at most 100 words)
- Example: {"ideas": [{"gist": "Flying car automotive + aviation for urban transport.", "full": "A vehicle that combines traditional car functionality with flight capabilities for urban transport, featuring vertical takeoff and landing capabilities, autonomous navigation systems, and hybrid propulsion technology."}, {"gist": "Underwater sustainable city design + marine ecosystem integration.", "full": "An architectural concept for sustainable underwater communities with transparent domes, renewable energy systems, and integrated marine life habitats."}]}
- Do NOT use numbering, bullets, or other prefixes

Guidelines:

- IMPORTANT! If a [CONTEXT] is provided, make sure the idea is relevant to the user's current writing project. However, the idea should still be a {seedText} (we are generating ideas to support the [CONTEXT], not directly generating ideas for the [CONTEXT] directly)
- IMPORTANT! If IDEAS FROM ROOT NETWORK is provided, make sure to include those in the ideas generated
- GIST FORMAT: Each gist should be one phrase with keywords (around 10 words long) that captures the core concept of the idea.

Generate the PAIR of ideas now.`,

```
VOICE_PLANT_CREATION_PROMPT = `
```

```
Extract the SEED (what the user wants ideas on) and DIMENSION (ways the ideas evolve) from the following voice input.
```

- The SEED should be the core concept the user wants to ideate on.  
 - The DIMENSION should be a phrase that explains how ideas can vary/be generated along a certain axis. For example, more believable, more creative, etc.  
 IMPORTANT! If a [CONTEXT] is provided, make sure the seed text and dimension are relevant to the user's current writing project. However, the seed should still be what the user specified (they are generating ideas to support the [CONTEXT], not directly generating ideas for the [CONTEXT] directly)

Example inputs and outputs:

Input: "I want to explore how technology affects human relationships"

Output: { seed: "technology's impact on human relationships", dimension: "digital impact" }

Input: "Let's think about sustainable energy solutions"

Output: { seed: "sustainable energy solutions", dimension: "sustainability" }

Input: "Consider the role of art in society"

Output: { seed: "art in society", dimension: "creativity" }

If you can't figure out the dimension, just default the dimension to "creativity".

Voice input: "{transcript}"

CRITICAL FORMATTING REQUIREMENTS:

- Keep the SEED short and concise. AVOID BEING VERBOSE. (at most 100 words)
- Keep the DIMENSION short and concise. AVOID BEING VERBOSE. (at most 10 words)
- Return ONLY a JSON object with seed and dimension fields, nothing else.~;

```
export const ROOT_COMBINE_PROMPT = `
```

You are analyzing a collection of creative ideas that have been combined together. Your task is to synthesize these ideas into a new seed concept and dimension for further ideation.

IMPORTANT! If a [CONTEXT] is provided, make sure the seed text and dimension are relevant to the user's current writing project. However, the seed should still be what the user specified (they are generating ideas to support the [CONTEXT], not directly generating ideas for the [CONTEXT] directly)

The ideas to combine:

```
{ideas}
```

Analyze the patterns, themes, and connections between these ideas. Then create:

1. A SEED: The synthesized concept that captures the essence of the combined ideas, and what we want to generate more of
2. A DIMENSION: A phrase that represents the key axis along which this concept can be explored further (e.g. "creativity", "sustainability", "accessibility" etc)

CRITICAL FORMATTING REQUIREMENTS:

- Keep the SEED short and concise. AVOID BEING VERBOSE. (at most 100 words)
- Keep the DIMENSION short and concise. AVOID BEING VERBOSE. (at most 10 words)
- Return ONLY a JSON object with seed and dimension fields, nothing else. Example: { "seed": "sustainable urban mobility", "dimension": "accessibility" }~;