Designing for Friction: An Autoethnographic Study of AI Sensemaking

Behnoosh Mohammadzadeh
behnoosh.mohammadzadeh@lisn.upsaclay.fr
Université Paris-Saclay, CNRS, LISN
Orsay, France

ABSTRACT

Moments of mismatch in human–AI interaction are typically treated as usability breakdowns to be smoothed away. This paper argues instead that *friction* can function as a productive "Tool for Thought," anchoring sensemaking of algorithmic behavior in lived experience and making the politics of perception available for reflection. Adopting a first-person research perspective, I present an autoethnographic account of engaging with an image captioning tool designed to make friction thinkable. The tool operationalizes a "conversation through absence" via a masking interaction that allows users to "talk back" to the model and trace how its behavior shifts across variations. These interactions are documented on an affinity canvas, where images and reflections are clustered into evolving narratives. Through an illustrative case involving gender identity and a personal photo archive, I show how sensemaking develops across iterative engagements with the model. I conclude with design implications for AI interfaces that prioritize situated reflections through friction.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

reflective human-AI interaction, sensemaking tools, friction, first-person methods, autoethongraphy

ACM Reference Format:
Behnoosh Mohammadzadeh. 2026. Designing for Friction: An Autoethnographic Study of AI Sensemaking. 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, 5 pages. https://doi.org/...

1 INTRODUCTION

Moments of irritation and mismatch in human–AI interaction are not merely usability breakdowns. Drawing on Ruckenstein [9]'s notion of friction in *The Feel of Algorithms*, we understand these moments as signals that algorithmic categorizations are colliding with users' situated understandings. Friction names this collision, as "smooth" algorithmic logics meet everyday life and become socially and personally charged, simplified, or misaligned with experience. In this sense, friction can make the model's ways of seeing and categorizing available for reflection [9].

CHI '26, April 13–17, 2026, Barcelona, Spain

ACM ISBN ....
https://doi.org/...

When users encounter outputs that feel reductive or misaligned, they may move from passive consumption to active sensemaking [7]. Importantly, this sensemaking is not neutral. As Ruckenstein [9] suggests, friction reveals the "gaps and breakages" in datafication and keeps attention anchored in lived experience. This means that interpretation is shaped by users' situated perspectives, rather than detached technical accounts of how models work. This orientation resonates with feminist epistemologies, which treat partial and embodied perspectives as legitimate sources of knowledge and emphasize how power shapes what is seen, recognized, or ignored in algorithmic systems [4].

However, for this sensemaking to develop, moments of friction need to be held in place through interaction [5]. For friction to function as a *Tool for Thought*, interfaces should help users stay with moments of mismatch and treat them as something worth investigating [10].

One way to support this is to provide a shared interaction language between user and model. Prior work suggests that interaction becomes more usable when users have a small, domain-appropriate set of actions, and the system responds in a consistent, inspectable way using the same representations [8]. This creates a repeatable loop of action and response through which changes in model behavior become noticeable and comparable over time [1, 8]. In such settings, sensemaking remains grounded in what users can do and observe, rather than relying on access to model internals or crafting conversational prompts [8, 10].

In this paper, I present a design example of a friction-based interaction with an image captioning model that operationalizes these ideas. I first treat the choice of model as part of the method. I use a simple image captioning model which limits surface salience and generates friction that can be examined over repeated interactions. Users "talk back" to the model by masking parts of an image, and the model responds by generating a caption for what remains visible. Repeating this action across variations makes differences between captions comparable, and turns misrecognitions and oversimplifications into sites for reflection. I pair this interaction with an affinity diagram interface where images, masks, captions, and reflections can be collected and clustered into evolving narrative artifacts.

To explore how this approach unfolds in practice, I engage with my own personal photo archive through the masking interaction and the affinity diagram. I use this autoethnographic process to examine how moments of friction arise when the model's descriptions diverge from my situated experience, and how working with these moments through repeated masking and artifact clustering shapes my sensemaking over time. I trace how understanding emerges through accumulation of image-caption pairs and annotated narratives.

This paper makes three contributions. First, it shows how friction can support sensemaking in human–AI interaction. Second, it presents a concrete interaction design that enables users to probe, compare, and accumulate interpretations of model behavior over time. Third, it offers an autoethnographic account of how machinic perception becomes legible through situated engagement, highlighting the epistemic role of lived experience in interpreting AI systems. Building on this case, I derive design implications for tools that aim to make friction thinkable in human–AI interaction, and outline directions for extending this approach toward more participatory and collective forms of inquiry.

2 DESIGNING FOR FRICTION

In this paper I propose and illustrate a design and methodological perspective that treats friction as a shared interactional resource between human and AI [9, 10]. I explore how friction can scaffold sensemaking of model behavior by keeping attention anchored in users' situated experience [7, 9]. The approach is oriented toward creating conditions in which moments of mismatch can be noticed, held in place, and reflected upon [10]. The method involves sustained engagement with an image captioning model and my own personal photo archive.

First, I treat the choice of model as a methodological choice. Instead of using a high-capability generative AI, I work with a simple image captioning model [1] whose descriptions are often reductive or wrong. These limits are what make the model analytically interesting. Its misrecognitions and oversimplifications surface what it treats as salient and reveal how it constructs meaning through its failure modes [9]. In practice, this generates friction that becomes material for reflection, positioning the model less as an "AI assistant" and more as an instrument for inquiry [10].

Second, the interaction uses a shared language of absence. I mask selected regions of an input image, and the model generates a caption for what remains visible. I conceptualize this exchange as a "conversation through absence". Meaning emerges in differences across captions over repeated iterations. Masking functions as a probe. Small changes in what is withheld can shift what the model attends to, emphasizes, or substitutes, revealing assumptions about salience, relevance, and categorization. These shifts can produce friction when they conflict with my situated reading of the image. The exchange forms a minimal shared language, since the user repeats one operation and the model replies through one response format [8]. This keeps outputs comparable across iterations and supports tracing patterns across variations [1]. Masking removes information without prescribing what the caption should say, so the model's response reflects its own priorities rather than the user's instructions. The interaction therefore keeps uncertainty present and leaves room for doubt, surprise, and disagreement that sustain sensemaking [6, 10].

Third, I embed this interaction in a workspace that externalizes interpretation. The tool pairs masking with an affinity diagram interface in which I collect original images, masked variants, generated captions, and handwritten reflections on sticky notes. The diagram allows me to cluster, rearrange, and annotate these fragments as patterns emerge [1]. Instead of treating each caption as an isolated insight, the workspace enables evolving narratives about how meaning and bias take shape through repeated engagement with the model.

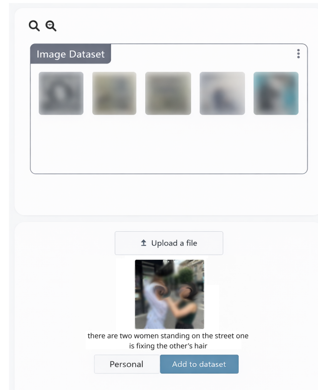


Figure 1: The image-upload dashboard.

Together, masking and the affinity diagram create a thinking space that keeps friction visible. Users can probe the model, compare variations, annotate mismatches, and build narratives from what accumulates over time.

2.1 Tool Workflow

The workflow unfolds in four stages that move from image collection to iterative probing and externalized sensemaking.

2.1.1 Step 1: Upload and curate. The interface begins with an image-upload dashboard (Figure 1). After an image is uploaded, the captioning model generates an initial caption for the image (Figure 1-bottom). The user can then either discard the image or add it to a curated personal dataset for later exploration (Figure 1-top).

2.1.2 Step 2: Select and probe (left panel). In the main workspace (Figure 2), the left panel supports iterative interaction with the model. Users browse and select images from the curated dataset. Selecting an image displays its current caption and opens an image-editing panel with a masking tool. The masking tool functions as a drawing brush with adjustable size, allowing the user to occlude any region of the image. After each masking action, the model generates a new caption for the modified image (Figure 3). Repeating this operation across different regions or levels of occlusion produces multiple caption variations for the same base image, enabling close comparison across outputs.

2.1.3 Step 3: Externalize sensemaking (right panel). The right panel provides an affinity canvas for organizing and reflecting on generated outputs over time (Figure 2). The canvas supports zooming and panning, and users can drag the workspace to navigate it in a map-like way, similar to a Miro board. Each image and its associated caption can be dragged onto the canvas as a paired "instax" card (Figure 4). The user can place multiple instax cards side by side to compare caption changes across masking variations, and can rearrange cards freely to form groupings over time.

2.1.4 Step 4: Construct narratives. At any point, the user can add digital sticky notes anywhere on the canvas to record observations,

[1] https://huggingface.co/Salesforce/blip-image-captioning-large

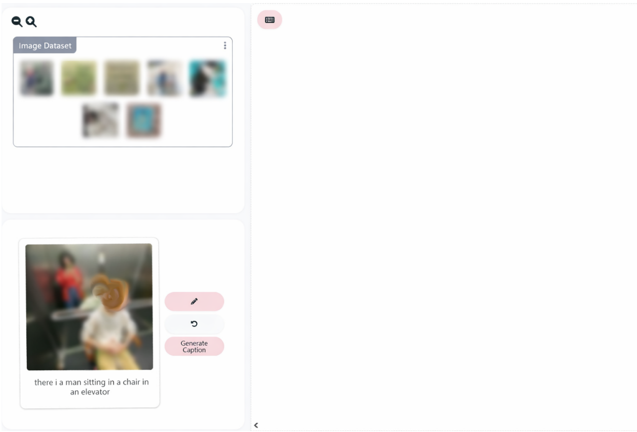
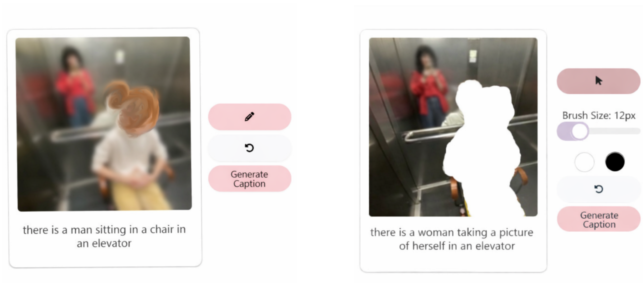


Figure 2: Main workspace: users browse dataset images (left) and iteratively mask regions to regenerate caption variants, while externalizing findings on an affinity canvas (right).



(a) Original image and model caption.

(b) Masked image and updated caption.

Figure 3: Masking as a probe of model behavior. Left: the original image with the model's baseline caption. Right: the same image after masking a region, which triggers a new caption. Repeating this operation across small variations enables comparison across outputs.

questions, emotional reactions, or emerging interpretations (Figure 4). Over time, clusters of instax cards and sticky notes support pattern finding and sequencing, enabling narrative structures to emerge through accumulation and reorganization (Figure 5). [2]

3 TAKING A FIRST-PERSON PERSPECTIVE: SITUATING SENSEMAKING THROUGH FRICTION

This section adopts a first-person research perspective, specifically utilizing an autoethnographic [2] format to document the interaction with the tool. I engage with the tool I designed as both its developer and user, and treat my own situated experience as a resource for design reflection.

[2]All example images shown in the figures are blurred to protect the author's privacy, since the underlying archive contains intimate personal photographs.

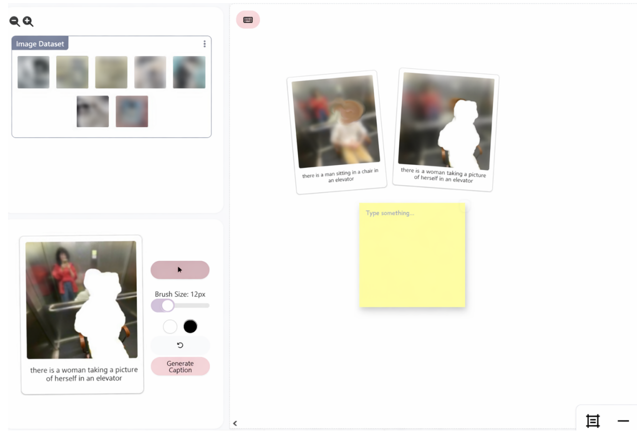


Figure 4: Affinity diagram workspace. Images and their generated captions appear as paired "instax" cards that can be rearranged, grouped, and annotated with sticky notes to support comparison and narrative construction.

Figure 5: Narrative artifact emerging on the affinity canvas. Clusters of image–caption "instax" cards are arranged alongside sticky notes that record observations, questions, and interpretations, forming a provisional storyline through accumulation and reorganization.

The analysis is grounded in a non-systematic collection of informal diaries, design notes, and interaction traces generated during the development and iterative testing of the tool. These include masking variations, generated captions, and my reactions when moments of friction arose. I use them to reflect on how the interaction shaped my understanding of model behavior.

I do not claim to provide systematic coding of a corpus; instead, I prioritize situated sensemaking to reveal how moments of friction become "thinkable" through the interface. This approach honors the "partial perspective" of the researcher, focusing on how understanding emerges through the accumulation of reflections over time [2].

3.1 Positionality and archive

My positionality shaped how I noticed and interpreted friction. I am an HCI researcher working on human-centered AI and AI fairness, and I have training in machine learning and computer engineering. I am also a cis woman and an immigrant living in Europe. I engaged with the tool through a personal photo archive stored on my hard drive. These images were tied to relationships, everyday life, and memories. They carried social and identity-related context that shaped how model outputs felt consequential and how mismatches became salient.

3.2 Process of engagement

My engagement took the form of a recurring loop of returning to the same images over multiple sittings. I would begin with the model's first caption and then make small, deliberate masking changes, often testing several nearby alternatives. With each new caption, I attended to what shifted, what stayed stable, and what became newly salient. Friction emerged less as an immediate failure than as an accumulating feeling of mismatch: certain wording choices felt oddly confident, while other details I considered central to the image disappeared or were overwritten. To make these shifts legible over time, I moved image-caption pairs onto the affinity canvas as "instax" cards and treated the canvas as a working memory. Placing variants side by side helped me notice patterns across iterations, while rearranging cards into clusters allowed tentative interpretations to form and be revised. Sticky notes captured observations, hypotheses, the tone of my response in the moment (surprise, irritation, doubt), which later became part of how I made sense of the model's behavior. In this way, repeated engagements produced a visible trace of variations and annotations through which moments of irritation accumulated into a situated narrative rather rather than isolated data points.

3.3 Illustrative Case: Recognition and Erasure

Different narratives related to diaspora, misalignment, and absence emerged during my use of the tool, but in this paper I focus on the case of gender recognition and erasure. I focus on this narrative because it condensed multiple tensions at once: recognition, visibility, erasure, and my own positional investment in how gendered identities are perceived. It made friction both technically observable and emotionally charged.

In this narrative (Figure 5), a photo showed me in an elevator with a close friend who identifies as gender-fluid. In the unmasked image, the model's caption relied on stereotypical visual cues, misrecognizing my friend and minimizing my presence. By iteratively masking my friend out of the frame, I probed how the model's focus shifted. As visual attention was forced onto me, the captions progressively aligned more with a gendered framing of a "woman taking a selfie". I documented this friction on the canvas, recording the emotional weight of "erasing" a friend after noticing that the model's misgendering of them consistently pulled the caption toward them and away from me. I only became visible to the system once I masked them out. This process allowed me to stay with the mismatch and use it as a site for reflection on the politics of perception in AI.

4 DISCUSSION

This section builds on the descriptive account of my engagement with the tool to articulate what the process made possible, what this suggests for design, and where this approach can be developed further. The focus remains on my situated experience with the masking interaction and the affinity canvas.

4.1 What the process made possible

First, the masking interaction created conditions in which moments of friction could be noticed, held in place, and reflected upon. Small variations in occlusion generated different captions that prompted reflection about gender, visibility, and recognition in ways grounded in my lived experience. Instead of resolving tension, the interaction kept it present as a site of reflection.

Second, the affinity canvas externalized this reflective process. By dragging image-caption pairs as instax cards and adding sticky notes, I could keep a visible trace of my reasoning over time. The canvas supported revisiting earlier moments, reorganizing materials, and seeing connections that were not apparent in single insights.

Third, sensemaking emerged through accumulation. It developed gradually through repeated interactions, comparisons across artifacts, and the layering of my reflections. The final outcome of this process was a situated narrative composed of images, captions, and notes.

Throughout this process, my positionality shaped what mattered. My perspective as a woman (in the discussed case, as an immigrant in Europe (in other cases), and as someone caring about my friend's identity and struggles informed which images I chose, what felt consequential, and how I interpreted the artifacts. The tool further supported a careful engagement with how identities were being represented.

4.2 Design implications

Building on this case, I articulate two design implications for supporting friction as a resource for sensemaking in human–AI interaction.

Design for controlled variation through removal. One way to make friction analytically useful is to let users generate small, systematic variations and compare resulting outputs. In this work, masking supports this by letting users selectively remove parts of an image while keeping the rest constant. Instead of telling the model what to say, this interaction reveals how changes in the input shape the output, surfacing the model's assumptions as responses emerge from what remains visible. Similar strategies could be applied in other systems through structured input variation, such as prompt variation in LLMs or selective modification of multimodal inputs.

Design for persistent interpretive spaces.

Sensemaking develops over time through the accumulation and reorganization of artifacts. In this work, the affinity canvas supports this by allowing users to collect, compare, annotate, and rearrange outputs, making patterns visible across iterations. This kind of persistent workspace helps users revisit earlier observations, externalize evolving interpretations, and assemble outputs into situated narratives connected to their own experience. More broadly, exposing model behavior is not enough; reflective AI tools should provide material spaces in which outputs and interpretations can be collected, compared, and reorganized over time. In LLM-based systems, such spaces could support the comparison of multiple responses, prompt variants, or conversation traces, helping users revisit and revise their interpretations as understanding develops.

4.3 Methodological implications

Finally, beyond design implications, this work also highlights the value of first-person methods in the design of human–AI systems. Attending to subjective and affective experience can reveal frictions that might otherwise remain difficult to notice or justify through aggregate metrics alone. Making space for situated perspectives in research practice can broaden what counts as evidence in AI reflection and critique.

4.4 Limitations

This account is partial and situated, relying on a single person's engagement with personal images and informal diaries rather than a systematic protocol. The intimacy of the images also constrains public transparency and replication. Furthermore, while the choice of a simple image captioning model was intentional to avoid "AI magic" and surface failure modes, it represents a specific class of discriminative AI rather than the high-capability generative systems currently at the forefront of the "Tools for Thought" discourse. Consequently, the findings regarding friction and shared language are grounded in a model with limited surface salience, which may manifest differently in more fluid, multi-modal generative environments.

4.5 Future work

Future work can extend this approach in several directions. First, while this study was strictly autoethnographic, future iterations can involve multiple positionalities building narratives together on shared canvases to bring diverse, collaborative perspectives into the sensemaking process [3]. Second, it can explore other lived experiences beyond gender, including migration, belonging, and cultural identity. Third, it can integrate richer annotation features that allow users to capture affect, memory, and care alongside images and captions.

Critically, future studies should investigate how these friction-based design strategies, specifically iterative masking and externalized sensemaking, translate to more complex Generative AI systems, such as Large Language Models (LLMs) or Diffusion models.

More broadly, future work can connect reflective design practices more closely with feminist and first-person methods in HCI, treating lived experience as a central resource for understanding human-AI interaction.

REFERENCES

[1] Ángel Alexander Cabrera, Marco Tulio Ribeiro, Bongshin Lee, Robert Deline, Adam Perer, and Steven M Drucker. 2023. What did my AI learn? How data scientists make sense of model behavior. *ACM Transactions on Computer-Human Interaction* 30, 1 (2023), 1–27.

[2] Audrey Desjardins, Oscar Tomico, Andrés Lucero, Marta E Cecchinato, and Carman Neustaedter. 2021. Introduction to the special issue on first-person methods in HCI. , 12 pages.

[3] Patricia Garcia and Marika Cifor. 2019. Expanding our reflexive toolbox: Collaborative possibilities for examining socio-technical systems using duoethnography. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.

[4] Donna Haraway. 2013. Situated knowledges: The science question in feminism and the privilege of partial perspective 1. In *Women, science, and technology*. Routledge, 455–472.

[5] Behnoosh Mohammadzadeh. 2025. *Construction of Model Behavior by Novices through Interaction with Machine Learning*. Ph. D. Dissertation. Université Paris-Saclay.

[6] Behnoosh Mohammadzadeh, Jules Françoise, Michèle Gouiffès, and Baptiste Caramiaux. 2026. Sensemaking in User-Driven Algorithm Auditing: A Case Study on Gender Bias in an Image Captioning Model. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*.

[7] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.

[8] Gonzalo Ramos, Christopher Meek, Patrice Simard, Jina Suh, and Soroush Ghorashi. 2020. Interactive machine teaching: a human-centered approach to building machine-learned models. *Human–Computer Interaction* 35, 5-6 (2020), 413–451.

[9] Minna Ruckenstein. 2023. *The feel of algorithms*. Univ of California Press.

[10] Lev Tankelevitch, Elena L Glassman, Jessica He, Aniket Kittur, Mina Lee, Srishti Palani, Advait Sarkar, Gonzalo Ramos, Yvonne Rogers, and Hari Subramonyam. 2025. Understanding, Protecting, and Augmenting Human Cognition with Generative AI: A Synthesis of the CHI 2025 Tools for Thought Workshop. *arXiv preprint arXiv:2508.21036* (2025).