

# Supporting Data-Frame Dynamics in AI-assisted Decision Making

CHENGBO ZHENG, University of Queensland, Australia

TIM MILLER, University of Queensland, Australia

ALINA BIALKOWSKI, University of Queensland, Australia

H PETER SOYER, University of Queensland, Australia

MONIKA JANDA, University of Queensland, Australia

High stakes decision making often requires a continuous interplay between evolving evidence and shifting hypotheses, a dynamic that is not well supported by current AI decision support systems. In this paper, we introduce a mixed-initiative framework for AI assisted decision making that is grounded in the data frame theory of sensemaking and the evaluative AI paradigm. Our approach enables both humans and AI to collaboratively construct, validate, and adapt hypotheses. We demonstrate our framework with an AI assisted skin cancer diagnosis prototype that leverages a concept bottleneck model to facilitate interpretable interactions and dynamic updates to diagnostic hypotheses.

Additional Key Words and Phrases: AI-assisted Decision Making, Explainable AI, Human-AI Collaboration, Sensemaking

## ACM Reference Format:

Chengbo Zheng, Tim Miller, Alina Bialkowski, H Peter Soyer, and Monika Janda. 2025. Supporting Data-Frame Dynamics in AI-assisted Decision Making. 1, 1 (June 2025), 6 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 Introduction

Decision-making in high-stakes scenarios is inherently dynamic. Consider the clinical context: a physician may initially identify a few key features in a patient’s profile that narrow the diagnostic options. Yet rather than committing immediately to these options, the doctor is likely to continuously integrate additional data, such as subtle clinical signs or evolving test results, to either reinforce the initial hypothesis or necessitate a shift to an alternative hypothesis. This iterative process is at the essence of the **data-frame theory** of sensemaking, which is developed through field observations of people decision making in complex and real-world environments [7]. In the data-frame theory, data refers to the signals extracted from the contexts (e.g., a dermatologist’s observation of a lesion’s dark color or irregular shape), while the frame represents the model or hypothesis that structure those observations (e.g., the lesion is melanoma). At its essence, the data-frame theory describes how people iteratively fit data into frames and fit frames into data, ultimately achieving their congruence.

This paper was presented at the 2025 ACM Workshop on Human-AI Interaction for Augmented Reasoning (AIREASONING-2025-01). This is the authors’ version for arXiv.

Authors’ Contact Information: Chengbo Zheng, University of Queensland, Brisbane, Queensland, Australia; Tim Miller, University of Queensland, Brisbane, Queensland, Australia; Alina Bialkowski, University of Queensland, Brisbane, Queensland, Australia; H Peter Soyer, University of Queensland, Brisbane, Queensland, Australia; Monika Janda, University of Queensland, Brisbane, Queensland, Australia.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

Manuscript submitted to ACM

1

In various domains [4, 10, 12], the dynamics of data and frames have been well documented, yet many AI-driven decision support systems do not adequately incorporate these interactions into their design. Most current systems focus on providing decision recommendations (i.e., frames) and possibly explanations linking the data to those frames [3, 9, 17]. By fixating on a single data-frame structure, such systems often overlook alternative possibilities and stray from the cognitive needs of human decision makers. When a user’s hypothesis diverges from AI’s, the human-AI relationship can shift from collaboration to competition (Figure 1 (A)), increasing the cognitive burden on the user or leading to overreliance if they abandon their own perspective.

Recently, Miller [11] introduced “evaluative AI,” a conceptual framework for AI-assisted decision making rooted in abductive reasoning and data-frame theory. The framework advocates for AI systems that, rather than merely offering recommendations, empower decision makers to formulate multiple hypotheses (i.e., frames) and evaluate them against data (Figure 1 (B)), thereby aligning the reasoning processes of both human and machine. While this approach lays a conceptual framework for collaborative decision support, it leaves several key implementation challenges unresolved. For example, what specific support should the AI system provide to facilitate the data-frame sensemaking process? If users find the system’s proposed data-frame questionable, how can they intervene in the process, and how should AI adapt to offer continuous, context-aware support?

In this paper, we provide a method that instantiates Miller [11]’s evaluative AI via a *mixed-initiative human-AI collaborative sensemaking* workflow that accounts for data-frame dynamics (Figure 1 (C)). By “mixed-initiative” [1], we refer to a process where both the human and the AI can contribute to the data-frame dynamics, thereby jointly shaping their understanding of the task. By formalizing the sensemaking process, we propose the decision support system should clearly communicate AI-driven processes of evidence extraction, hypothesis retrieval, and hypothesis scoring. On the other hand, human decision makers should be empowered to easily express their own reasoning across these three critical sensemaking functions. Human inputs should also directly influence AI reasoning, progressively aligning the human-AI team’s collaborative sensemaking. We see two key advantages in this method. First, users are not compelled to make a binary judgment on AI’s logic (i.e., accept or reject), potentially reducing overreliance and the cognitive load. Second, AI can continue offering decision support over time grounded on the users’ just-in-time reasoning, which might facilitate engagement, thus benefit skill learning over time.

In the following sections, we first present our vision for human-AI collaborative sensemaking based on the data-frame theory. Then, to illustrate this concept, we introduce our AI-assisted skin cancer diagnosis prototype.

## 2 Support Data-Frame Dynamics through Human-AI Collaboration

While there are other theoretical models characterizing the sensemaking process [13, 14, 16], Klein et al. [7]’s data-frame theory is developed with a focus on high-stakes decision making [6].

Here, we formalize Klein et al. [7]’s data-frame theory. In order to fit the context of decision making – the focus of this paper – we adopt the terminology of *evidence* for data and *hypotheses* for frames, following Miller [11].

Let  $I$  represent the set of raw information available,  $E_t$  denote the set of evidence derived from this raw information at time step  $t$ , and  $\mathcal{H}_t$  indicate the set of hypotheses under consideration at time  $t$ . The sensemaking process begins by transforming raw information into evidence via an extraction function  $f_t : I_t \rightarrow E_t$ , such that  $E_t = f_t(I_t)$ . Given the current evidence  $E_t$ , a retrieval function  $r_t : E_t \rightarrow 2^{\mathcal{H}}$  selects a subset of relevant hypotheses, denoted as  $\mathcal{H}_t = r_t(E_t)$ . Each hypothesis  $H \in \mathcal{H}_t$  is evaluated against the evidence using a scoring function  $s : E_t \times \mathcal{H}_t \rightarrow [0, 1]$ , resulting in a score  $s(E_t, H)$  that indicates how well each hypothesis explains the evidence.

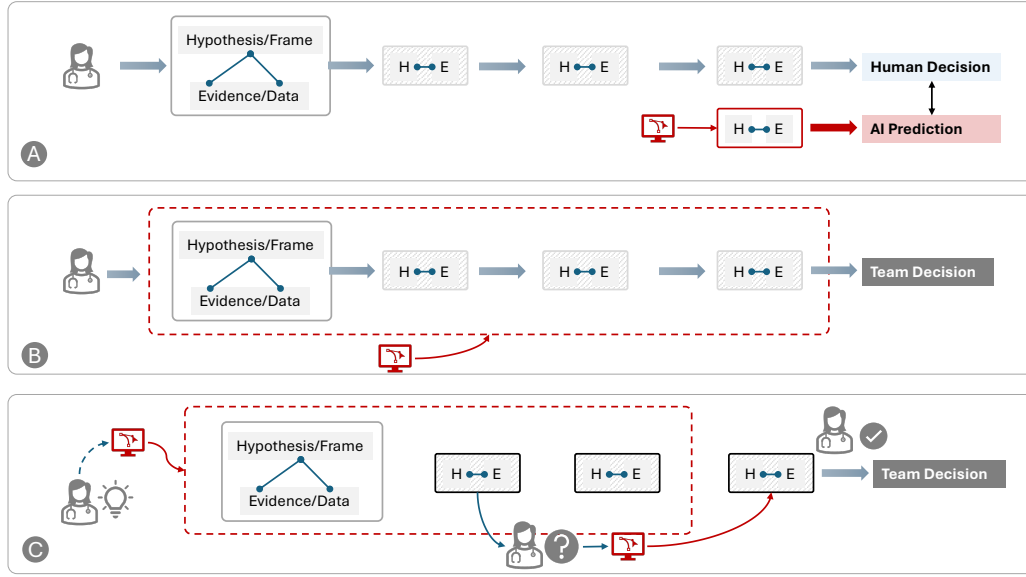


Fig. 1. (A) Typical AI decision support: The human’s decision and the AI’s prediction can be at odds, requiring the human to interpret the AI’s output and resolve conflicts independently; (B) The evaluative AI framework: The AI generates hypotheses and corresponding evidence to facilitate sensemaking. However, how to generate the support to well synchronize with human reasoning is unclear; (C) Instantiate evaluative AI with mixed-initiative sensemaking: Humans can propose reasoning, and the AI actively integrates these updates to offer support, enabling collaborative, team-based decision making.

If at time  $t$  the highest-scoring hypothesis does not meet or exceed an acceptance threshold  $\delta$ , i.e., if  $\max_{H \in \mathcal{H}_t} s(E_t, H) < \delta$ , one of, or both the two possible iterative refinement actions can occur at the next time step  $t + 1$ : 1) additional evidence is generated from raw information,  $E_{t+1} = E_t \cup f_{t+1}(I)$ ; 2) additional hypotheses are retrieved to expand the candidate set,  $\mathcal{H}_{t+1} = \mathcal{H}_t \cup r_{t+1}(E_t)$  or  $\mathcal{H}_{t+1} = \mathcal{H}_t \cup r_{t+1}(E_{t+1})$ . This iterative updating continues dynamically over multiple time steps until at least one hypothesis exceeds the threshold  $\delta$ , formally expressed as:

$$H_t^* = \arg \max_{H \in \mathcal{H}_t} s(E_t, H), \quad \text{subject to} \quad s(E_t, H_t^*) \geq \delta.$$

To support the dynamics of the sensemaking, we propose that the AI decision support should assist in evidence extraction  $f_t$ , hypotheses retrieval  $r_t$ , and scoring of hypotheses versus evidence  $s_t$ . Besides, the human decision makers might have different functions, thus user-initiated interactions should be supported in intervening the above three functions. Such an intervention should also introduce updates to AI’s reasoning by introducing  $f_{t+1}$ ,  $r_{t+1}$  and  $s_{t+1}$ .

More specifically, first, for **the evidence extraction function**  $f_t$ , AI could display the extracted evidence in an interpretable format and allow users to annotate, modify, or augment this evidence. Feedback from the human decision maker introduces  $f_{t+1}$ ,  $E_{t+1}$  as well as  $\mathcal{H}_{t+1}$  from  $r_t(E_{t+1})$ . The latter two updates should be clearly communicated with the users to lay down the common ground. Second, for **the hypotheses retrieval function**  $r_t$ , AI should enable users to review the candidate hypotheses, suggest additional ones, or exclude those that appear irrelevant. Third, for **the hypotheses scoring function**  $s(E_t, H)$ , AI should clear communicate its weightings of the evidences that influence the evaluation. Users should be empowered to discuss new weightings to shape the team scoring of the hypotheses.

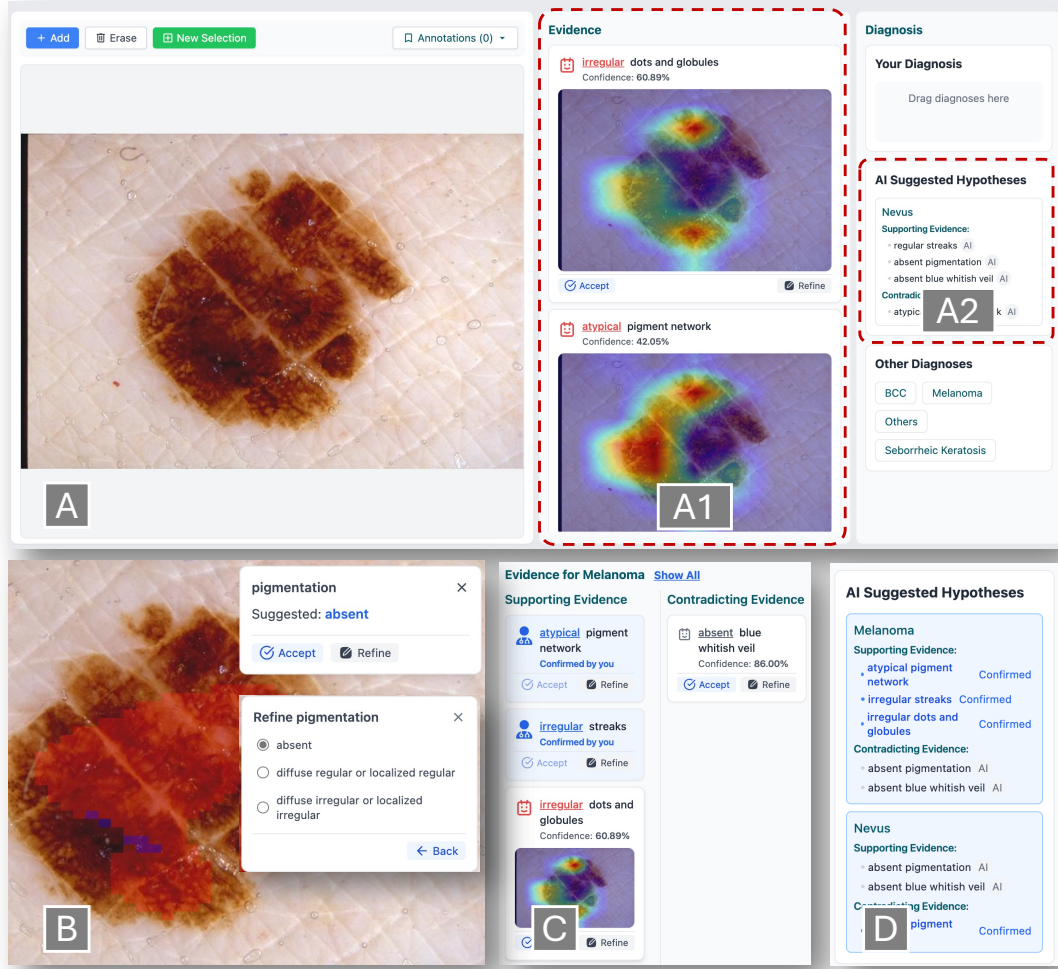


Fig. 2. An example system that implements the mixed-initiative sensemaking support for skin cancer diagnosis (the image is from the derm7pt dataset [5]).

### 3 Example: AI-assisted Skin Cancer Diagnosis

In this section, we present an AI-assisted skin cancer diagnosis prototype (Figure 2) to exemplify the implementation of the three key functions for supporting data-frame dynamics outlined in section 2. The interface comprises three primary views (from left to right in Figure 2 (A)):

- **Image view:** displays the dermoscopy image.
- **Evidence view:** lists evidences extracted by the AI or provided by the user.
- **Diagnosis view:** allows the user to manage hypotheses and finalize the diagnosis.

We trained a concept bottleneck model (CBM)[8] on the derm7pt dataset[5] to power the prototype. CBM is an interpretable deep learning approach that first maps image representations to task-relevant concepts, then makes final predictions based on these concepts. The seven dermoscopy concept annotations (e.g., pigment network, streaks) in the derm7pt dataset make it well-suited for training a CBM. These concepts can also be viewed as critical clinical evidences that corresponds to the data in the Klein et al. [7]’s sensemaking theory. Moreover, CBM enables users to edit the predicted concepts, thus influencing the final diagnostic outcome [8].

The prototype implements the three key functions mentioned in section 2 to facilitate the sensemaking process in dermatological diagnosis:

- **Evidence Extraction.** Initially, the system presents AI-extracted evidence in the *Evidence View*, utilizing a heatmap to visually highlight how the evidence was derived from the raw input (computed using GradCam [15] based on concept predictions within CBM). Upon first viewing the raw input, human decision-makers might not check the AI-extracted evidence but quickly anchor on specific pieces of evidence (key anchoring data, as discussed by Klein et al. [7]). Besides, they might find discrepancies between AI-generated evidence and their own understanding, or determine that evidence updates are necessary when evaluating hypotheses against existing evidence. To address these scenarios, the system supports user-initiated interventions for evidence extraction. If a user considers certain evidence inaccurate or incomplete, they can refine it using the “Refine” button (see Figure 2(B) and Figure 2 (C)). Such refinements update the corresponding concept predictions in the CBM, consequently adjusting AI-generated hypotheses. Additionally, if users identify suspicious regions, they can annotate these regions (purple patches in Figure 2(B)). Although users can manually label these annotations, the system proactively proposes potential evidence related to the user’s annotation (illustrated as red patches in Figure 2(B) and within pop-up windows). Users may accept or further modify these AI-proposed evidences
- **Hypotheses Retrieval.** The system automatically suggests hypotheses based on the current set of evidence (Figure 2(A2)). The hypotheses set is generated through conformal prediction [2]. This set is also dynamic. When newly confirmed evidence introduces additional plausible hypotheses, the system recommends these new options (e.g., “Melanoma” appearing in Figure 2(D) after user refinement of evidence). Additionally, users can introduce new hypotheses if the AI suggestions do not fully align with their reasoning or if updates required as no hypothesis satisfies the acceptance threshold after the scoring. This is done by dragging options from the “Other Diagnosis” section to the “Your Diagnosis” section within the Diagnosis View.
- **Hypotheses Scoring.** Each hypothesis is accompanied by clearly indicated supporting and contradicting evidence (Figure 2 (D)), generated based on weights from the final prediction layer in the CBM, using concept predictions as inputs. Confirmed evidence by users and AI-generated, user-unverified evidence are visually distinguished in the interface (e.g., blue points for confirmed evidence and grey points for unverified evidence in Figure 2 (D)). Users can also explore the scoring of individual hypotheses in greater depth by selecting a specific hypothesis within the Evidence View (Figure 2 (C)). They can remove evidences from the supporting and contradicting groups, or move evidence between groups (though such cases are expected to be rare in skin cancer diagnosis as general consensus on the weights of the clinical concepts exists [5]).

## 4 Conclusion

In this work, we instantiates evaluative AI [11] by introducing a mixed-initiative interaction framework that bridges human insight and machine reasoning for enhanced decision making. Grounded in Klein et al. [7]’s data-frame theory,

our approach suggests supporting three sensemaking functions – evidence extraction, hypotheses retrieval, and hypotheses scoring – that empower human decision makers to collaborate with AI. We introduce an AI-assisted skin cancer diagnosis prototype as a proof-of-concept of the interaction framework. Further study with dermatologists on the prototype will be conducted to evaluate the effectiveness and tradeoffs of the interaction framework.

## References

- [1] James E Allen, Curry I Guinn, and Eric Horvitz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14, 5 (1999), 14–23.
- [2] Anastasios Angelopoulos, Stephen Bates, Jitendra Malik, and Michael I Jordan. 2020. Uncertainty sets for image classifiers using conformal prediction. *arXiv preprint arXiv:2009.14193* (2020).
- [3] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. 2021. Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–16.
- [4] Darren Hudson and Gurmeet Singh. 2017. Expert medical decision-making: How the data-frame theory can explain physician sense-making. In *Building Capacity for Health Informatics in the Future*. IOS Press, 167–171.
- [5] Jeremy Kawahara, Sara Daneshvar, Giuseppe Argenziano, and Ghassan Hamarneh. 2018. Seven-point checklist and skin lesion classification using multitask multimodal neural nets. *IEEE journal of biomedical and health informatics* 23, 2 (2018), 538–546.
- [6] G Klein. 2005. Recognition-primed decisions. *Ergonomics: major writings* 271 (2005), 271–307.
- [7] Gary Klein, Jennifer K Phillips, Erica L Rall, and Deborah A Peluso. 2007. A data-frame theory of sensemaking. In *Expertise out of context*. Psychology Press, 118–160.
- [8] Pang Wei Koh, Thao Nguyen, Yew Siang Tang, Stephen Mussmann, Emma Pierson, Been Kim, and Percy Liang. 2020. Concept bottleneck models. In *International conference on machine learning*. PMLR, 5338–5348.
- [9] Shuai Ma, Ying Lei, Xinru Wang, Chengbo Zheng, Chuhan Shi, Ming Yin, and Xiaojuan Ma. 2023. Who should i trust: Ai or myself? leveraging human and ai correctness likelihood to promote appropriate trust in ai-assisted decision-making. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [10] Laura G Militello, Shilo Anders, Sarah M Downs, Julie DiIulio, Elizabeth C Danielson, Robert W Hurley, and Christopher A Harle. 2018. Understanding how primary care clinicians make sense of chronic pain. *Cognition, Technology & Work* 20, 4 (2018), 575–584.
- [11] Tim Miller. 2023. Explainable ai is dead, long live explainable ai! hypothesis-driven decision support using evaluative ai. In *Proceedings of the 2023 ACM conference on fairness, accountability, and transparency*. 333–342.
- [12] David T Moore and Robert R Hoffman. 2011. Data-frame theory of sensemaking as a best model for intelligence. *American Intelligence Journal* 29, 2 (2011), 145–158.
- [13] 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.
- [14] Daniel M Russell, Mark J Stefik, Peter Pirolli, and Stuart K Card. 1993. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on human factors in computing systems*. 269–276.
- [15] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*. 618–626.
- [16] Karl E Weick and Karl E Weick. 1995. *Sensemaking in organizations*. Vol. 3. Sage publications Thousand Oaks, CA.
- [17] Ming Yin, Jennifer Wortman Vaughan, and Hanna Wallach. 2019. Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–12.