Interactive AI: Designing for the "Two Black Boxes" Problem

John Wenskovitch
Virginia Tech

Chris North
Virginia Tech

***Abstract*—The nascent field of Explainable AI seeks to unmask the underlying details of black box learning algorithms, enabling these algorithms to explain their state and results to human analysts. However, to truly enable interactive AI, we argue that there exists a second black box representing the cognitive process of the user, containing information which must be communicated to the algorithm. Using this "Two Black Boxes" problem as motivation, we present a symmetric, collaborative human-AI model using Semantic Interaction as a design philosophy to connect human and machine. We discuss challenges associated with each phase of communication between the pair of cooperatively-learning entities and the benefits that emerge from combining the expertise of the human and the AI.**

■ **IN DATA** analytics, the "black box" problem denotes the challenge that artificial intelligence (AI) algorithms in general, and neural network models in particular, suffer from opaqueness. These algorithms can supply useful results, such as finding novel latent structure in otherwise difficult to comprehend data. However, they typically do not provide any justification or rationale for their output. Users of these algorithms are therefore faced with the decision of whether to accept the results at face value, without the ability to question or understand the underlying process. This problem has resulted in the "Explainable AI" (XAI) research agenda, which seeks to open the black box of these algorithms and explain their results to human analysts. Analysts can thereby inspect the algorithms and gain insight into how the analytical results were discovered by the algorithm, the process trail, analytical provenance, and supporting data. This is represented by the right-pointing arrow in Figure 1.

However, this is only half the problem in human-AI interaction for data analytics. We posit that there is another black box in the equation — the black box of human cognition. Analysts conduct cognitive sensemaking activities, and as a result of these thought processes, also want to be

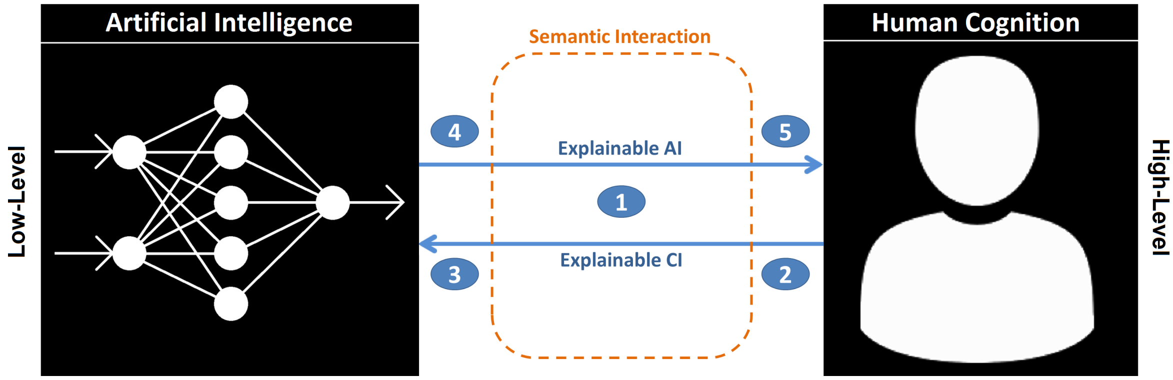

Figure 1. The two black boxes of interactive AI, human cognition and AI algorithm, are connected via two communication channels that can be mediated by Semantic Interaction (SI) systems (#1). Explainable Cognitive Intelligence (XCI) transmits information from the human to the AI (#2–3), while Explainable Artificial Intelligence (XAI) transmits information from the AI to the human (#4–5).

able to influence the algorithms to produce results of interest to their sensemaking. However, from the perspective of the algorithm, the human mind is a black box that is equally (or perhaps even more) difficult to interpret. How can the human "explain" herself to the algorithm, so that the algorithm can respond to her internal thought processes, goals, motivations, expert domain knowledge, and intents? How can the machine learn from user interaction? To parallel XAI, we refer to this communication channel as "Explainable Cognitive Intelligence" (XCI), represented by the left-pointing arrow in Figure 1. We refer to both communication challenges together as the "Two Black Boxes" problem.

A goal of research in this area is to optimize the strengths and mitigate the weaknesses of the human and the machine. The human is skilled at conducting cognitive sensemaking such as abductive reasoning, inference, making judgements, and bringing domain expertise to bear, while the machine can efficiently discover implicit knowledge or hidden patterns in large-scale data. Many intelligent systems are designed so that humans can focus on the high-level analysis, supplemented by the low-level details provided by the machine. A truly hybrid human-AI co-learning system will perform this task while also fostering increased trust and communication between the two entities, enabling the human to understand and apply the analyses of the machine (XAI) while the machine learns and adapts from the human cognition (XCI), thereby enabling bigger and better analysis.

Our contribution in this work is a discussion of five notable challenges that must be addressed to create such hybrid human-AI co-learning. These challenges are focused on the communication channels between the human and the AI. As shown in Figure 1, the black boxes of the human and the AI can be accessed by interactive analytical systems, but the actions of contextualizing, externalizing, and interpreting the knowledge of both entities are challenges that must be addressed. Throughout this article, we use Semantic Interaction [1] as a running example to discuss these challenges within the context of interactive sensemaking of textual information.

An Overview of Semantic Interaction

Semantic Interaction (SI) is a design philosophy featuring a means to directly interact with modeled representations of data in interactive systems. Often also referred to as *demonstrational interaction*, the high-level goal of SI is to shield the user from the complexity of tuning underlying computational models and maintaining their focus on the data while still benefiting from those models. To do so, the user performs actions within the visual representation to communicate a desired outcome, and the system attempts to learn an update to those underlying models that will create that outcome. As a result, the system can infer the intent of the user from a sequence of interactions.

For example, the StarSPIRE [2] and Cosmos [3] systems enable analysts to spatially organize collections of textual documents. These systems model the human sensemaking process as two sub-processes: a foraging process and a synthesis process (see Pirolli and Card [4] for a thorough discussion of this sensemaking process). As analysts organize important documents, the AI provides support by foraging relevant documents onto the screen and synthesizing those documents into the structure of the user. Indeed, using computational support in the background to support the analysis of the user has been shown to enhance the capabilities of the user in several performance measures [5].

The visual interface of StarSPIRE consists of documents (or glyphs representing documents) within an open workspace. By manipulating these documents through interactions such as highlighting words, annotating notes on a document, and changing the position of documents in the workspace, the system can infer both documents the user may be interested in and why those documents appear to be interesting. Using this inference, the system can search for additional relevant documents, restructuring the workspace to better match the needs of the user. In a similar manner, the AxiSketcher system [6] permits a user to sketch a path over a workspace; in response, the system forms the axes that best combine attributes of the data to simplify that path.

Challenge 1: What is the Loop?

Several phrases are used in visual analytics to describe the roles of humans and AI in the workflows of interactive systems. "Human in the Loop" describes interactive systems that are commonly designed so that analytic algorithms occasionally consult humans for expert feedback. In other words, the human is one step away from being a bystander in the computational process, only occasionally chiming in to provide course correction or to respond to system inquiries.

At the other extreme is "Machine in the Loop," in which the human is primarily in charge of the analysis process, occasionally consulting the machine for suggestions or assistance in problems that require computational prowess [7]. Such a model is useful for creative exercises driven by human agency, in which occasional tasks such as sorting or otherwise arranging a large collection of documents by some priority metric support a larger human activity.

Each of these extremes can be appropriate for certain classes of problems, but are not universally applicable. The amount of automation and the amount of human intervention required in an arbitrary workflow is dependent upon factors such as the data, tasks, and users. Work by Chen and Ebert [8] argues that building data intelligence workflows to consider these factors necessitates an optimization-based design strategy.

For the case of sensemaking, the "Human *is* the Loop" approach proposed by Endert et al. [9] emphasizes understanding the underlying cognitive sensemaking processes of users, and then fitting computational support into the existing workflow. Thus, the focus is on sensemaking loop [4], modeled and augmented with AI methods for the foraging and synthesis sub-loops.

Understanding the best method for integrating AI into existing human-centric processes is a challenge, but we argue that the challenge goes one step deeper. The AI should not just assist the user, but should also improve its understanding of the goals and processes of the user in order to provide better assistance over time. Such sensemaking is not just a human-centric process, but should also be a goal of the machine, forming a *cooperative learning loop*. The "Machine Learning from User Interactions" (MLUI) Workshop[1] is a venue focused in this area of computationally understanding the behaviors of a user.

Considering again the examples of StarSPIRE and Cosmos, we see that both systems are designed so that the human and the machine cooperate in exploring and making sense of document corpora. An iteration of the learning process begins with a human interaction in the synthesis phase, performing some action to better understand the on-screen information. The AI detects this action and incrementally learns the interests of the user based on a sequence of such actions. From this learning, the AI can then forage for other relevant documents to present to the user. This *synthesis-driven foraging* [5] then leads to an updated layout structured around the interests and goals of the human, in which the new information is placed in context. The human can then ingest

[1] https://learningfromusersworkshop.github.io/

this new information and situate it appropriately into their mental model before beginning the next step in their analysis. Overall, this interaction loop works in such a way that the AI learns the interests of the human, providing computational support to aid the human cognitive process.

For this cooperative learning loop to function optimally, both the human and the AI need to understand the other. If the human black box remains closed to the AI, the AI cannot provide the best computational assistance. Likewise, if the AI black box remains closed to the human, the human may not trust the suggestions that they receive. In order to open both black boxes fully, we need to address four additional communication challenges: extracting cognition from the human, presenting that cognition in a meaningful way to the AI, extracting information from the AI, and presenting that information in a meaning way to the human.

Challenge 2: Externalizing Cognition

In the XCI communication channel, extracting the cognitive state of the human is a major challenge that must be addressed. Without knowledge of the thoughts and conclusions made by the human, the AI has no information to draw upon in order to learn and improve. Distributed cognition [10], the process of externalizing and sharing cognitive resources in order to extend individual cognitive limits, serves as one means of extracting such information. Humans often make use of physical notes, diagrams, and lists to externalize knowledge, reveal relationships, and recall information. Converting such artifacts to digital form provides a mechanism for the machine to learn the thoughts of the human.

Embodied interactive systems provide means for humans to use virtual spaces to externalize their thoughts. For example, the "Space to Think" study [11] discovered how large, high-resolution displays (Figure 2) support human sensemaking processes by becoming a part of the distributed cognitive process. In addition to demonstrating how the large display served as an external memory, these studies also uncovered the semantic layers of the space. Analysts used order, proximity, and alignment to encode pairwise relationships, create clusters and lists, and generally express mental associations in a digital space. The space provides the interaction medium by which the human naturally and efficiently externalizes their cognition.

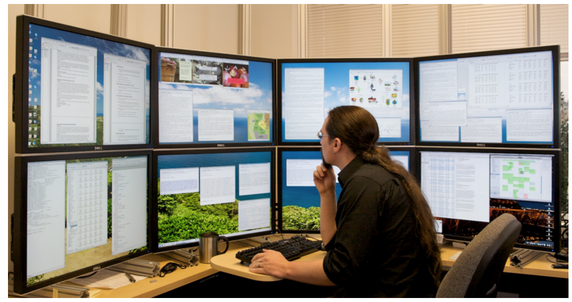

Figure 2. Large displays can be used to externalize the thoughts and conclusions drawn by a human.

The space not only helps human cognition, but also can provide AI a window into human cognition. The semantic relationships that an analyst externalizes can be utilized to model the cognition of the human via SI. This technique has been demonstrated in StarSPIRE [2], where the "space to think" was modeled as a metric learning and dimension reduction process. Distributed cognition spaces can serve as the common ground between the two black boxes.

Other structural metaphors exist to enable humans to externalize their cognition. Trees and related hierarchical structures such as file systems enable users to interactively categorize and form associations between groups of documents. Smaller-scale interactive spaces enable users to interact with both documents and high-dimensional quantitative data in virtual large spaces [3], [12], [13] rather than relying upon the physically large spaces seen in "Space to Think." Sketching interactions represent a more complex medium to externalize a thought process, such as demonstrating a desired axis in AxiSketcher [6]. Additionally, future research can extend these techniques into augmented reality (AR) and virtual reality (VR) spaces, providing even richer "space to think."

The key to enabling AI to learn from these metaphors is discovering how cognition reveals itself in particular interactions and scenarios, making use of human-computer interaction (HCI) techniques such as think-aloud studies and semi-structured interviews. Discover the natural way that humans want to perform an action rather than forcing them into an artificial externalization. Would a user rather drag two documents closer together to demonstrate a degree of similarity, or should they select both documents, open a dialog box, and drag a slider?

Challenge 3: Cognitive Input into AI

The XCI process is not complete without converting the externalized cognition of the human into a form understandable by a machine. In other words, the high-level interactions performed by the human must be transformed into a set of parameter and value updates that reflect the intentions underlying the interactions. The AI can then learn about the needs and interests of the user from these parameter updates.

An important usability challenge is that typical machine parameters are low-level and require premature formalization, both mismatches for human cognition. A common example is the k-means clustering algorithm, which requires prior specification of the parameter k, the number of clusters, well before the human analyst has cognitively formalized how many clusters she is interested in. Similarly, model parameters such as term weights are too low-level, as human sensemaking emphasizes higher-level concepts and events. Instead, a design goal is to *contextualize* the inputs within the cognitive space to think by recasting natural sensemaking interactions (e.g. moving, highlighting, annotating, searching) into underlying parameter manipulations via methods such as heuristic interpretation or machine learning.

Identifying the appropriate mapping between interaction and parameter is a problem-specific process, as each system will have a set of unique interactions and unique parameters to connect. To continue the StarSPIRE example from previous sections, the system maps a “relevance” weight to the words contained within each document. As a user performs interactions on those words such as highlighting, this weight is increased, demonstrating increased user interest in that term (Figure 3). Similar interactions also exist at the document level, including opening and closing documents and overlapping them. These interactions inform the importance of each term, which further drives the foraging process.

The goal of the SI paradigm is to shield users from the complexity of these low-level parameters, making use of existing interactions that users will perform in their natural sensemaking process. In doing so, a human can naturally interact with the system while supported by the AI in an unobtrusive manner. The mathematical underpinnings of this technique are summarized in Visual to Parametric Interaction (V2PI) [14]. As the human will be performing such actions during their sensemaking process, making use of them to drive AI support is effectively a “cost-free” extension of those interactions.

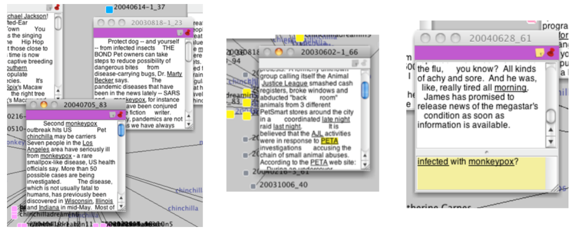

Figure 3. Clustering documents, highlighting, and annotating are actions that an analyst would ordinarily perform while exploring a document collection. SI uses these actions to learn the interests of the human, converting them into model updates.

A parallel can also be drawn to research in the area of adaptive search algorithms. Such algorithms change their behavior based on new information available as they continue to run, providing better results over time as they optimize based upon their knowledge. Indeed, StarSPIRE improves its understanding of the search interests of a user as they continue to interact with the system, providing additional relevant documents to the user as their interactions iteratively provide new information to the algorithm.

Challenge 4: Contextual Output from AI

Explainable AI (XAI) is an active area of research in both machine learning and visualization [15]. Among other contributions, the goal of XAI is to convey information about the state of the AI to the human as a means to build trust in the model output, demonstrate the level of confidence vs. uncertainty in results, and show computation in context. For our discussion, we divide this XAI process into two components: extracting information from the AI to display to the human (this section) and the processes undertaken by the human with that information

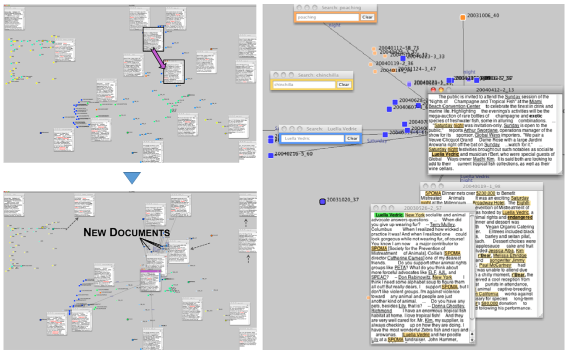

Figure 4. StarSPIRE uses the result of interactions such as overlapping a pair of documents **(left)** to forage for related relevant documents, placing them in context within the display **(right)**.

(next section).

The overarching goal of SI is to shield users from the complexity of underlying models so that their focus can remain on the analysis task at hand, rather than making the cognitive switch to tuning model parameters manually. Similarly, the XAI phase embedded within a system supporting SI should also be non-intrusive. The relevant information extracted from the AI should be visually integrated into the existing visual metaphor, either into or onto the artifacts on which the analyst is working. As a result, a system can augment the mental model of the human, providing new information that is relevant to their interests within the corresponding locations in their space to think.

In StarSPIRE, information from the AI is *contextualized* into the existing workspace in three ways. First, the visual workspace displays the documents via a force-directed layout. With this layout, users can pin selected documents into specified locations, and the remainder of the documents are automatically positioned within the display so that similar documents are positioned close to each other, while dissimilar documents are separated. When the system learns a new set of weights to apply to the terms, the layout is restructured with this new information in mind.

Second, the newly-foraged documents are inserted into the workspace using that same layout algorithm, so that each new document appears in context with the others (Figure 4). As a result, if a user highlights the term “chinchilla” in a pinned document, a group of new documents relevant to chinchillas will appear in that area, simultaneously contracting the cluster of these documents.

Finally, the AI reveals its internal weights by automatically highlighting terms within the documents of investigation, mapping the color intensity to the weight of the terms in the model. Such contextualized output directs the analyst to the useful detailed information within those documents, while also enabling the user to gain insight into what the AI has inferred is important.

Challenge 5: Understanding the Machine

Human sensemaking is incremental, but machine learning and analysis is often batch-oriented. A significant challenge is designing algorithms that embrace human *incremental formalism* [16]: algorithms that learn incrementally from iterations of human feedback and provide incremental results. This enables the human to absorb the AI into their cognitive process. If a magical AI provided the correct solution up front, the human likely would not recognize it as such. The incremental co-learning process is important.

A challenge inherent in human-AI collaboration, particularly when a human is receiving recommendations from an AI entity when exploring a large data set, is the human leaping to a conclusion based upon partial evidence, prematurely inferring relationships and conjecturing connections that may not be supported by other evidence. If the most relevant documents are provided before the supporting evidence, then the human is drawing conclusions without knowing the full context of the information presented. Similarly, the AI should not conclude the interests of the human too early in the analysis process, as those interests may shift as new information is presented or as new chains of evidence lead the human towards a different conclusion than their initial hypothesis.

This issue is addressed by StarSPIRE through the continuous, incremental updates to the underlying models, layout, and displayed document set (Figure 5). Rather than providing a number of interesting documents for the human to consider simultaneously, the human is required to open each document sequentially. As they do so, the

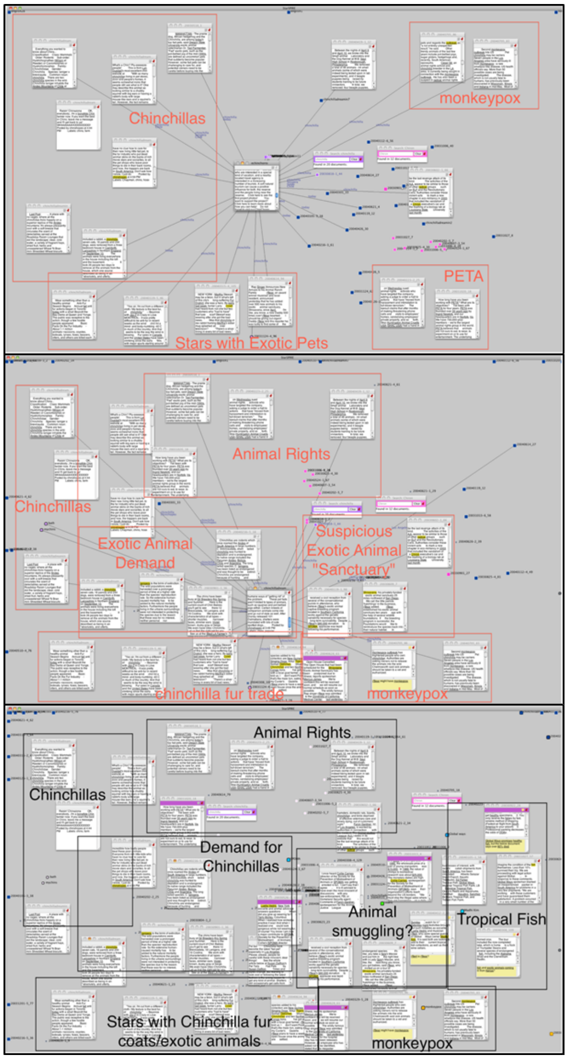

Figure 5. As a user interacts with StarSPIRE, they progressively build a more thorough understanding of the document collection they are exploring, with corresponding structures externalized in the workspace.

system begins to incrementally learn about the interests of the human, updating term weights slowly rather than drastically. Term weights also decay over time, so that past interactions are not judged to be as important as recent interactions. With incremental changes to the model, the layout updates are also incremental. Documents do not drift far from their last position after an interaction, maintaining the mental model of the human and preventing disorientation. This incremental formalism approach [16] considers the history of interactions by the human to gradually construct a user interest model.

This tight coupling between the AI incrementally learning about the human and the human incrementally learning about the data (with AI assistance) demonstrates the co-learning relationship between the human and the AI. Each provide their skills (the deduction of the human and the computational processing of the AI) to collaboratively solve a problem.

Discussion

The combination of the XCI and XAI communication channels serves to pull the algorithm into the human cognitive space, affording a collaborative learning environment that improves the performance of both participants. Algorithms such as those in SI learn from relatively simple interactions in the workspace rather than relying upon complex semantics, incrementally gaining clues about the user intent.

Because the human and AI are tightly integrated in the process, SI aids in relieving some explainability challenges studied in XAI research. This can produce more trust in results, as well as the development of clear high-level and low-level roles. Further, learning from the exploration process of the user permits system designers to quantify and overcome bias. We briefly discuss these benefits in this section.

Trust

Cooperation between a pair of humans is dependent upon trust, and the same is true for cooperation between a human and an AI. A core goal of XAI research is to build systems that humans can trust, with this goal achieved by designing ways for the AI to be understood and its decisions accurately interpreted. As a result, the human in the pair has additional insight on the behavior of the AI, and can subsequently better determine whether or not its output is accurate.

A similar relationship exists for XCI, though in a more extended manner. Rather than the AI needing to trust the human, the human needs to trust that the AI understands their goals and intentions. In other words, "Is the AI actually listening to me?" The incremental formalism process supports this need, as the human is able to see progressive changes occur within the visualization in response to their actions. Since the human was involved throughout the process, the need for extensive explanation about the arrived solution is reduced or obviated. This stands in contrast to more typical XAI approaches which better model the "Human in the Loop" paradigm, performing a lengthy computation and then presenting a set of results for a human to validate or correct.

High- and Low-Level Roles

In human-AI collaborative learning, each participant plays separate roles to support the other. The role of the AI is low-level processing, using its computational power to quickly search through documents, identify co-occurring terms, identify targets in images, and other computationally-intensive tasks. The role of the human is high-level thinking, identifying connections through knowledge that may not be contained in files, spotting patterns and trends that are computationally-difficult to locate, and driving the course of the analysis.

A benefit of collaborative learning is permitting the AI to free the cognitive resources of the human to focus on the bigger picture; the human no longer has to perform low-level search and sort tasks. The machine fills in the concepts with relevant detailed information. When the human repositions a single document, the AI handles updating the position of every other on-screen document in response to that interaction. The human can then spend their time contemplating the concept represented by the new cluster they created and the supporting evidence.

Quantifying and Overcoming Bias

By learning from the human's interactions, AI can quantify many aspects of human exploration, including potential bias. An important feature of the synthesis-driven foraging in StarSPIRE is its relevance-based foraging. Rather than searching for documents supporting the hypothesis of the user, it instead retrieves all relevant documents, including those which may refute the hypothesis. Further, because StarSPIRE builds a model of user interests during the sensemaking process, it is able to query a broader set of relevant documents than simple keyword search [5]. Indeed, keyword search overfocuses the results on the smaller set of known relevant terms. By quantifying this model, the system helps to overcome the bias of the search keywords entered by the user.

Related solutions include artificially broadening queries to introduce alternative analysis paths, as well as reorganizing the documents presented to the human in order to draw their attention to these alternate paths.Computing these alternatives requires analysis of the exploration of the human. By measuring the analysis trajectory of the human against multiple possibilities, the bias of current analysis can be quantified and corrected. Systems such as ModelSpace [17] permit the human to view their analysis paths and identify portions of the document space that they may not have inspected.

Semantic Interaction Alternatives and Extensions

While we use SI as a paradigm for mediating communication between human and AI, alternative techniques do exist. These techniques primarily differ in the XCI communication channel; using interactions for external cognition is certainly not the only means by which the cognition of a human can be extracted and conveyed to the AI. For example, a human can use natural language to communicate their interests and intentions to the AI [18]. A sequence of utterances can be used to continue to subsequent phases of analysis or to correct misunderstandings in the interpretation of a previous instruction. The intent of the human does not even necessarily need to be inferred; an AI can use active learning to directly query the user [19].

The ideas that underlie SI can also be extended further. For example, StarSPIRE requires a mapping between interaction and model update to be defined in advance. In contrast, the Metatation system does not require such a mapping. Instead, Metatation combines a linguistic data-model with the interaction sequence from the human to learn the meaning behind free-form annotations, thereby recommending next analysis steps [20]. Another technique for model explainability is seen in systems such as Andromeda, which displays the current model state via a set of interactive sliders beside the workspace rather than within the space itself [12].

Conclusion

Understanding the cognition of a human and conveying that information accurately to a machine is a promising direction for future research. By examining the broader problem of Interactive AI, not just Explainable AI, we hope that methodologies such as SI will lead toward a more human-centered approach to AI design.

REFERENCES

1. A. Endert, P. Fiaux, and C. North, "Semantic Interaction for Sensemaking: Inferring Analytical Reasoning for Model Steering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, pp. 2879–2888, Dec 2012.
2. L. Bradel, C. North, L. House, and S. Leman, "Multi-Model Semantic Interaction for Text Analytics," in *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 163–172, Oct 2014.
3. M. Dowling, N. Wycoff, B. Mayer, J. Wenskovitch, S. Leman, L. House, N. Polys, C. North, and P. Hauck, "Interactive Visual Analytics for Sensemaking with Big Text," *Big Data Research*, vol. 16, pp. 49 – 58, 2019.
4. P. Pirolli and S. Card, "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis," *Proceedings of International Conference on Intelligence Analysis*, vol. 5, pp. 2–4, 2005.
5. J. Wenskovitch, L. Bradel, M. Dowling, L. House, and C. North, "The Effect of Semantic Interaction on Foraging in Text Analysis," in *2018 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 13–24, Oct 2018.
6. B. C. Kwon, H. Kim, E. Wall, J. Choo, H. Park, and A. Endert, "AxiSketcher: Interactive Nonlinear Axis Mapping of Visualizations through User Drawings," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, pp. 221–230, Jan 2017.
7. C. Tan, "Human-Centered Machine Learning: A Machine-in-the-Loop Approach." https://medium.com/@ChenhaoTan/human-centered-machine-learning-a-machine-in-the-loop-approach-ed024db34fe7, 2018.
8. M. Chen and D. S. Ebert, "An ontological framework for supporting the design and evaluation of visual analytics systems," *Computer Graphics Forum*, vol. 38, no. 3, pp. 131–144, 2019.
9. A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews, "The Human is the Loop: New Directions for Visual Analytics," *Journal of Intelligent Information Systems*, vol. 43, pp. 411–435, Dec 2014.
10. E. Lehtinen, K. Hakkarainen, L. Lipponen, M. Rahikainen, and H. Muukkonen, "Computer Supported Collaborative Learning: A Review," *The JHGI Giesbers Reports on Education*, vol. 10, p. 1999, 1999.
11. C. Andrews, A. Endert, and C. North, "Space to Think: Large High-Resolution Displays for Sensemaking," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, (New York, NY, USA), pp. 55–64, ACM, 2010.
12. J. Z. Self, M. Dowling, J. Wenskovitch, I. Crandell, M. Wang, L. House, S. Leman, and C. North, "Observation-Level and Parametric Interaction for High-Dimensional Data Analysis," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 8, pp. 15:1–15:36, June 2018.
13. J. Wenskovitch and C. North, "Observation-Level Interaction with Clustering and Dimension Reduction Algorithms," in *Proceedings of the 2nd Workshop on Human-In-the-Loop Data Analytics*, HILDA'17, (New York, NY, USA), pp. 14:1–14:6, ACM, 2017.
14. S. C. Leman, L. House, D. Maiti, A. Endert, and C. North, "Visual to Parametric Interaction (V2PI)," *PLoS ONE*, vol. 8, pp. 1–12, 03 2013.
15. F. K. Došilović, M. Brčić, and N. Hlupić, "Explainable Artificial Intelligence: A Survey," in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 0210–0215, May 2018.
16. F. M. Shipman III and R. McCall, "Supporting Knowledge-Base Evolution with Incremental Formalization," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 285–291, 1994.
17. E. T. Brown, S. Yarlagadda, K. Cook, R. Chang, and A. Endert, "ModelSpace: Visualizing the Trails of Data Models in Visual Analytics Systems," in *Proceedings of the Machine Learning from User Interaction for Visualization and Analytics Workshop at IEEE VIS 2018*, Oct 2018.
18. M. Tory and V. Setlur, "Do What I Mean, Not What I Say! Design Considerations for Supporting Intent and Context in Analytical Conversation," *IEEE Transactions on Visualization and Computer Graphics*, 2019.
19. Y. S. Chan and H. T. Ng, "Domain Adaptation with Active Learning for Word Sense Disambiguation," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, (Prague, Czech Republic), pp. 49–56, Association for Computational Linguistics, June 2007.

20. H. Mehta, A. Bradley, M. Hancock, and C. Collins, "Metatation: Annotation As Implicit Interaction to Bridge Close and Distant Reading," *ACM Trans. Comput.-Hum. Interact.*, vol. 24, pp. 35:1–35:41, Nov. 2017.

John Wenskovitch is a Visiting Assistant Professor of Computer Science at Virginia Tech. He has previously worked as a research intern at FXPAL, and also taught as a Mathematics/Statistics adjunct at Chatham University and as a Computer Science Visiting Assistant Professor at Allegheny College. His current work at the Discovery Analytics Center seeks to enable human-AI interaction for big data analytics, with specific foci on high-dimensional data, sensemaking, and clustering. John is a member of the IEEE and the IEEE Computer Society. Contact him at jw87@cs.vt.edu.

Chris North is a Professor of Computer Science at Virginia Tech. He is Associate Director of the Discovery Analytics Center, and leads the Visual Analytics research group. He is a member of the Center for Human-Computer Interaction and the Hume Center for National Security. The major themes of his research are data visualization and analytics, usability in interactive human-in-the-loop machine learning, novel display and interaction techniques, large high-resolution display workspaces, visual text analytics for sensemaking, and insight-based evaluation. Contact him at north@vt.edu.