How Does Explainability Look in Hybrid User Interfaces?

Julián Méndez*
Interactive Media Lab Dresden
TU Dresden, Germany

Marc Satkowski*†
Interactive Media Lab Dresden
TU Dresden, Germany

Rufat Rzayev*
Interactive Media Lab Dresden
TU Dresden, Germany

1 Introduction

The recent growth of Artificial Intelligence (AI) based systems considerably widespread their use in many application areas and our daily lives [28]. For instance, AI models are nowadays imbued into web search engines, self-autonomous vehicles, recommendation systems, games, and healthcare [24]. Accordingly, the demand for eXplainable AI (XAI) has risen [20, 21] to help users cope with the growing complexity and opaqueness of the emerging generation of AI models [20, 21]. By allowing users to perceive and make sense of the behaviors and outcomes of such models, XAI enables users of diverse levels of expertise to trust, manage, design, inspect, and develop AI models [13].

Since visualizations are the most common form of explanations [24], one of the most prominent ways to achieve XAI is by leveraging Visual Analytics (VA) [22]. This is also presented by hundreds of examples discussed in recent surveys [1, 11, 28]. VA uses interactive data visualization techniques to support users in understanding, reasoning, and making decisions based on large and complex datasets [26]. However, as the amount and complexity of data on VA tasks increase, a single device can be insufficient to cover the user needs [51]. In response to this, hybrid user interfaces (HUIs) provide a meaningful combination of devices and modalities to increase the quality of interaction and the display possibilities [51]. HUIs allow users to benefit from a multi-device ecology for a given task in a specific context. Furthermore, complementing traditional displays with mixed reality (MR) devices provides advantages such as stereoscopic rendering, unlimited display space, and natural interaction. Considering that XAI deals with large, complex, multi-faceted datasets and given evidence that it can be explored in MR (e.g., [50]), using hybrid or complementary interfaces for XAI is a natural combination to consider. Yet, research on this is very rarely, if at all, present in the XAI surveys mentioned before.

In this position paper, we discuss the under-explored intersection between XAI and HUI incorporating MR devices based on our previous experiences in developing tools for explainability [3, 14, 34] and multi-device systems [43] for VA [29, 38, 42]. We discuss the challenges of combining these areas and sketch scenarios where clear research opportunities arise in hopes of kick-starting the discussion on combining and making the most of both domains.

2 Combining HUI and XAI

As AI has consistently proven capable of permeating every aspect of our lives, we assume its role in future HUI containing MR devices will be highly important. Therefore, we first clarify our positions regarding the future of HUI (Sec. 2.1) and the broad types of XAI that can be seen in the wild (Sec. 2.4). Afterward, we describe the challenges and opportunities related to the use of explanations for AI models in HUIs (Sec. 2.3). Lastly, we highlight the importance of targeting XAI as opposed to general VA for HUIs (Sec. 2.4).

*e-mail: [julian.mendez2, marc.satkowski, rufat.rzayev]@tu-dresden.de
†Also with Centre for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig, Germany

2.1 HUIs incorporating MR

Hybrid user interfaces (HUIs), more recently referred to as complementary interfaces, present a *"symbiosis of interfaces, where each component purposefully increases the quality of interaction and further supports users in their current activity"* [51]. Reviewing the development of the consumer market over the last decades, we can see that the introduction of laptops, smartphones, tablets, and smartwatches expanded the device ecology and catered to specific features and needs of the users. Another emerging technology is MR head-mounted displays (HMDs), which still need to enter the consumer market successfully. In general, MR HMDs possess several properties that are impossible to imitate with commonly used devices, making the introduction of MR HMDs into the existing device ecologies promising. Such properties include:

MR-1: Unlimited Display Space: MR HMDs enable the spatial placement of virtual information in any real-world or virtual environment, allowing users to interact with near-unlimited display space. To complement 2D displays, MR enables extending their available space by placing virtual content around, above, or between them [29] or even in front or behind given displays [38].

MR-2: Stereoscopic Rendering: MR HMDs render images for each eye of the user separately, enabling a stereoscopic perception of the virtual content. Therefore, complementing a traditional 2D display with an MR HMD can allow a user to benefit from the provided feeling of immersion and presence [33]. Moreover, concepts such as embedded or situated visualization can be used to view virtual content in proximity to the related referent [9].

MR-3: Altering the Real-World: As MR content can be placed everywhere in the environment, it is possible to (temporarily) alter the visual perception of other objects. With regard to desktop monitors, MR can add additional information in empty spaces, can hide objects on the monitor screen, or can change the appearance of already existing objects [15].

MR-4: Natural Interaction: HUIs incorporating MR devices can not only facilitate the display possibilities but also provide natural interaction capabilities. Through additional sensors within an HMD, it is possible to, e.g. track the hands and directly interact with the presented virtual content. Additionally, physical navigation [7] is beneficial in finding content spatially placed in the immediate environment.

MR-5: Personal View: As MR HMDs are worn on a user's head, they are inherently personal devices. Such devices allow users to view private or personalized content without being visible to bystanders. For example, while a HUI can provide a complex and detailed visualization for an expert user to interact in an MR environment, a non-expert user can simultaneously view a simplified version of the exact visualization on a 2D display.

Following those benefits, we believe that MR devices will become an integral part of our future device ecology, allowing users to overcome some limitations of traditional display and interaction modalities. However, a distinct device combination (i.e., HUI) can be used even when not all or any benefits from a specific device type are exhausted in the given context. One straightforward example is the transfer of desktop content into the MR space due to missing additional monitors or the limited display space while still resulting in an *"increased cognitive load [or] high transaction costs"* [51].

To summarize, we agree that HUIs combining traditional devices (i.e., desktops, smartphones) and MR HMDs will find broad applicability in the future. However, we believe that not every possible HUI will aspire to *"purposefully increase the quality of interaction"* [51]. HUI may also be used due to, for example, legacy bias, as users may not want to leave behind their desktop devices altogether but may be willing to try out alternatives simultaneously.

2.2 Explainability through Visual Analytics

The area of eXplainable Artificial Intelligence (XAI) pertains to tools that *"enable human users to understand, appropriately trust and effectively manage the emerging generation of AI systems"* [20]. While this definition is quite broad (i.e., it does not pertain to only machine learning but any AI system), we focus on approaches that leverage Visual Analytics (VA) since the most common form of explanation is visualization [24]. This focus allows developers to decouple from the specifics of each model and reduce the problem to an instance of high-dimensional, multi-faceted abstract data visualization, which can be related to arbitrary use cases.

Rosa et al. [28] abridge the topic of explanation tools using VA, based on a cross-analysis of general XAI surveys. We slightly rephrase some of their points to illustrate a general overview of the types of XAI we cover with our ideas, as opposed to their more specific version for only deep learning models. Firstly, they note two ways to differentiate explanations, based on *the moment when the explanations are generated* (ad hoc: while the model is building, or post hoc: after the model has created outputs) and on *the range that the explanation covers* (local: a single input to output relation, or global: a wide range of inputs and outputs). Then, the authors propose 5 categories of explanations, which we adapted to general XAI as follows:

XAI-1. Input Attribution: These explanations focus on (and extract insights on) which parts (or features) of the input are responsible for the model's outcome. For example, in image classification, certain areas of the input image may contain the most relevant parts needed for a prediction [10, 40]. On the other hand, for reasoning systems where multiple proofs can be computed for consequences, some of the proofs do not use the entirety of the input [4]. On similarity learning (e.g., for fraud detection), it is also possible to decompose the inputs to identify which data attributes had the higher weights leading to the predictions [37].

XAI-2. Model Attribution: In contrast to Input attribution, these explanations identify which components of the AI model were triggered or had an effect on producing an output. For example, looking at which neurons or layers activate the most may lead to insights into the design of a neural network and the role of its components [44]. Another example of OWL ontologies is the computation of atomic decompositions, which reveal a subset of the ontology responsible for the computation of a consequence [46]. Using these types of explanations, it is also possible to extract equivalent models that are less opaque, such as natural language-like syntaxes that describe neural network behaviors (e.g., Rule Extraction [6]).

XAI-3. Explanations by examples: Generally speaking, this explanation type illustrates instances (data samples or cases) that the AI model considers similar to the input from the user. For instance, a recommendation system may provide examples of previously recorded cases where the outcome is comparable to the current one. More explicitly, Amazon's "Users like you also bought...", or Netflix's "Because you liked X, you may also like Y" hint at this type of explanation and could also show specific instances of purchase or view histories to support the explanations. It is also possible to combine these explanations with others. For example, the system Melody [10] provides **examples** in which specific features are similar to those of the user input (**XAI-1**). The concept of explanations by example could also extend to the generation of examples that represent an outcome of an AI system in a more graspable manner. For instance, if the output of a reasoner system says that "All birds have wings, but not all fly", it may be easier for the user to digest the explanation when given an explicit example, like penguins.

XAI-4. Counterfactuals: These explanations indicate the minimal changes needed for the outcome of the models to change. As stated in [28], they answer the question: *"What do I have to change to obtain a different outcome?"*. Similar to Model Attribution, these type of explanations are very *actionable* when the users only have agency over the inputs that the AI receives. For instance, in AI planning, these explanations enable users to decide which goals to remove with the least amount of impact on their original plans. On generative models (e.g., ChatGPT, Dahl-e, GANs), one could provide options for slight alterations of the inputs, which would result in dramatic differences in the output of the models. An example of this explanation method is the Recast system [49], which highlights specific words within a textual input with respect to certain metrics influencing the model output.

XAI-5. Model behavior: These approaches support understanding the models as entities or agents that react to inputs. From human-in-the-loop techniques, users can freely test the model's behavior to understand its shortcomings and capabilities. Examples of these can be seen in visual representations of the model structure and/or behaviors using various display technologies. Some examples include showing a node-link diagram of a neural network in Virtual Reality [32] and probing an AI with real-world objects using mobile devices [47]. Colley et al. have also conceptualized how Tangible user interfaces may be used for making sense of the AI models [12]. It is also possible to create abstract visualizations that focus primarily on post-hoc metrics of the AI models under analysis, like it is done with the What-if-tool [48].

With the outlined categories of explanations, we can now envision how they may play a role together with HUIs that involve MR. It is worth repeating that, to the best of our knowledge, while there are examples in the wild for XAI on just MR (e.g., [50]), research on XAI on HUIs is missing. In the following, we will lay out several reasons we see for this research gap, which in turn become challenges to overcome for integrating the XAI and HUI research venues.

2.3 Challenges in Combining XAI and HUI

Research projects on HUIs, such as MARVIS [29], PARVIS [38], and STREAM [23], illustrate the rich exploration space that comes to VA on HUIs involving mobile devices, MR HMDs, wall displays, etc. However, this research usually exists in a theoretical space, unconcerned with real-life applications and focusing on datasets for which a single device may not suffice. As stated in the previous section, looking at the field of XAI from the lens of VA allows us to examine AI models as high-dimensional, multifaceted abstract data. This implies that the amount, complexity, and structure of the XAI data can be already challenging for a single device setup, similar to the datasets used in the research for VA on HUI. While that by itself could suffice as motivation to explore the intersection of XAI and HUI, many of the challenges that affect VA in HUI are inherited from XAI, as described in the following.

Technical limitations hinder practicality: XAI tools are evaluated based on how well they support users in practical tasks related to the AI models in question [28]. This means that the XAI systems themselves have to also be reliable and trustworthy to transmit the same sentiments about the underlying models. MR systems, however, are infamous for limitations purely on a technical level (e.g., inconsistent calibration, complicated setups involving external tracking systems, limited gesture sets, limited field of view), which are occasionally deliberately overlooked in MR research, considering an optimistic future where these limitations have been surpassed. Thus, complementing, for example, a traditional desktop computer with an MR HMD can introduce more technical challenges than the ones it alleviates.

Abstract data vs spatial data: AI models are by nature abstract, and thus the data they produce is also abstract. While one of the most vital benefits of MR is the ability to render stereoscopic imagery (**MR-2**), the absence of inherently spatial/3D data means that there is no intrinsic need to make use of stereoscopy to display it. There is, however, a notable trend indicating that stereoscopy has the potential for overcoming the known limitations of 3D visualizations of abstract data, as summarized in the survey by Kraus et al. [27]. Likewise, Data Visceralization [30] proposes that virtual reality can enable a deeper understanding of abstract data through direct and immersive representations of the abstract data. Another point to make here is that although the AI models are abstract, they may work over a domain that involves spatial data (e.g., robot-assisted surgery, trajectory analysis, architecture, and production systems). This presents an opportunity to embed (part of) the visual explanations on spatial simulations or artifacts of the application domain itself.

Legacy bias and skepticism: Although this is a general problem with unconventional and novel technologies, it is more critical when convincing stakeholders that the value obtained from such technologies is higher than their cost (both monetarily and in terms of learning curves, comfort, and efficiency in solving their tasks). For instance, general software development, despite its spread, importance, and influence, has yet to incorporate fully hybrid interfaces in the industry space. After introducing AI to this process (e.g., Github's Copilot[1] or other generative models), we believe that the addition of XAI does not merit by itself the adoption of complementary or different display technologies. We experienced this resistance to unfamiliar technologies while developing Evonne [34], an explainability system for AI reasoning. In this project, our design sessions with the domain experts led us to prioritize familiarity with the device setups supported. Evonne has some features that support complementary interfaces, i.e., its views can be off-loaded onto any device with a web browser, and the views remain connected for linked interactions through sockets. However, it is clear that the full potential of hybrid setups is not utilized here yet.

2.4 XAI as concrete research for VA on HUIs

The challenges described so far extend to general VA systems dealing with high-dimensional, multi-faceted abstract data. Therefore, one must ask what is gained in HUI by targeting XAI as opposed to general VA. XAI presents a *concrete* application scenario for VA, as opposed to the more *theoretical* cases usually discussed in VA HUI research. In some of our HUI projects (e.g., [29, 38]), we have encountered difficulties justifying theoretical data scenarios. With concrete application areas such as XAI, it is possible to directly compare the effectiveness of solutions designed for systems using HUI versus those that exist in traditional, single-device setups.

Furthermore, aspects such as *trust*, *interpretability*, and *visualization literacy* have recently received more attention in visualization research (e.g., [31, 41]). This is in part thanks to the wide use of AI systems, amplifying these requirements (e.g., [18]). However, such aspects have not yet been at the focus of HUI for VA, to the best of our knowledge. Motivated by XAI, we find better footing for research questions such as: *How can HUI systems be designed to effectively deliver explanations to users of varying visualization literacy and domain knowledge?*; *Is there a relation between the display technology (or a combination thereof) and acceptance of explanations?*, etc. Lastly, since the end goal of XAI is typically to *convince* or *justify*, the research on the psychological effects and ethical implications of HUIs also becomes more relevant. For instance, *how can we ensure that HUIs do not conceive, mislead, or trick users into e.g., fallacies about the models?*

Overall, we believe that the research venue of XAI is rather oriented on providing solutions to *current problems* that are directly affecting users in their abrupt incorporation of AI into their daily lives, while research for VA on HUI involving MR HMDs exists on a more *future-oriented*, exploratory space. While it is possible to see XAI as subsumed by VA, the convergence of these research directions strengthens the relevance of research on VA on HUI and promises exciting results that we cannot predict.

3 Envisioning HUIs for Explainability

So far, we have laid out five types of AI explanations, five major benefits that MR brings to HUI, and our view on why the research on HUIs for XAI has remained largely unexplored, despite the promises that such an intersection inherently holds. Our initial proposal to think about this field simply consists of combining the two previously described sets of types and benefits, resulting in a 5x5 design space. The following scenarios illustrate the benefits of MR (see Sec. 2.1) with respect to the explanation types presented in Sec. 2.4.

Data on Unlimited Display Space (MR-1) Perhaps the easiest-to-justify benefit of combining a traditional desktop with an MR HMD for XAI is the ability to place an arbitrary amount of information in the space beyond the desktop screen. While the field of view limitation affects current technologies, for the simplest version of this scenario, one could consider the usage of virtual monitors together with physical monitors, as described by Pavanatto et al. [36]. Technically then, every XAI tool covered in the current surveys could be put onto a HUI with minimal effort (and obtaining minimal benefits). At this point, one could investigate the following: What is the most efficient virtual vs. real monitor configuration for XAI through VA? Where to locate the monitors? or; What are the preferred dimensions? Although we speculate that these questions can be reduced to matters of personal preference. Instead of full virtual monitors, one could off-load single views onto the MR space, achieving a similar effect. However, the benefits of arbitrary information allocation in unlimited display space are more apparent when considering what-if analysis and comparison tasks. With this, it is now possible to superpose, juxtapose, or freely organize views (or the entirety) of the VA dashboards. This allows for more flexible and natural interactions (**MR-4**) with the abstract data representations, partially emulating the benefits of data physicalization [25]. Generally speaking, all XAI tasks (**XAI-1 to 5**) can benefit from such flexibility.

Leveraging Stereoscopy (MR-2) Suppose an AI model is working on an application domain with inherently 3D data. In that case, it is relatively easy to imagine that embedding the explanations onto a stereoscopic representation can be helpful for explanations by example (**XAI-3**). For instance, imagine an architect working on a hybrid setup consisting of an augmented reality (AR) model of the building under construction and interactive surfaces (this is the ARchitecture setup proposed by Reipschläger et al. [39]). To this, we incorporate an AI recommendation system that suggests adjustments to the architect's design, trying to minimize the chance of structural malfunctions in the building. An explanation by example would be much easier to convey (and for the architect to accept) if the system could illustrate, directly in AR, buildings that historically presented similar design flaws.

If the AI model works on abstract data, we can use the previously mentioned concept of data visceralization. Stereoscopy by itself may support understanding explanations through, e.g., the use of

[1] https://github.com/features/copilot

spatiality to assist the user in sense-making of the model behavior (**XAI-5**). Imagine a social network engineer working with a graph neural network (GNN) for social recommendations (e.g., whom to follow, what interests to subscribe to). This is the case illustrated by Fan et al. [17]. We can envision a HUI setup, where the engineer has the editor interface for the model on a traditional desktop application. At the same time, a stereoscopic representation of the social recommendation graphs complements the 2D monitor display. Following the argument that stereoscopic representations of graphs can facilitate analysis of their structures [27], it may be easier for the engineer to detect artifacts of the networks, such as over-squashing or over-smoothing [2].

Explaining by Altering the Real World (MR-3) Inspired by the research on augmented forensics [19], we imagine a forensic pathologist making a postmortem diagnosis on a patient. To do this, the doctor has a handheld device for their notes and an AR system that lets them look at the patient's body while overlaying 3D representations of the organs. An AI model assists the pathologist in diagnosing and suggests a cause of death that confuses the doctor. To make sense of this confusion, the AR system may highlight what characteristics of the body contributed to the predicted cause of death more heavily (**XAI-1**, input attribution). Likewise, this convenient scenario would benefit significantly from an XAI modality that enables the AR overlays to show with a counterfactual (**XAI-4**) what would need to be different for the cause of death to be different or no longer certain. However, in this case, the existing AR overlays would need to accommodate the visual explanation, so using the handheld device to show the explanation on demand may be more convenient.

Trust and safety on personal spaces (MR-5) As introduced and motivated earlier, XAI systems have among their goals to generate trust towards the AI models. Suppose we envision collaborative HUI systems where multiple users interact with a central interface while requesting and querying private information. In that case, it is easy to imagine an increase in the sense of privacy (and therefore trust) simply by using devices (e.g., head-mounted displays, mobile) that do not expose personal views to other users. Consider personal AI assistants that provide hints/suggestions to individuals in collaborative scenarios, and which the users can monitor and inspect using XAI views within their personal space or publicly. This scenario is motivated by recent research on the use of large language models like ChatGPT for personal education [45] and mental health [5], as well as the digital twin environments [35], exemplifying the relevance of research on psychological and ethical implications of HUIs for XAI.

Through these examples, we hope to have illustrated how a design space for XAI on HUI can be initialized by brainstorming over explanation types (XAI-1 to 5) with respect to the benefits of the complementary technology – in our case, MR (MR-1 to 5). The granularity of this design space can certainly be increased given a specific application scenario, and as hinted in our descriptions, many questions arise from this thought experiment, forming a quite broad research space for the future.

4 Discussion and Outlook

While we believe it is too early to claim that we understand the full impact that both of these venues will have, it is certainly possible to not just envision but also already design systems that enhance XAI thanks to the strengths of HUIs.

With that in mind, we believe the following points are worth investigating for researchers interested in XAI and HUI simultaneously. The **placement** of the explanations within the device setup needs to be considered. This includes the relation between the placement and the semantic relation of the distributed information [16]. Is there a clear benefit to having the explanations on one display technology over the others? Subsequently, the **representation** of the explanations should convey the intended message effectively. This is highly dependent on who is the user that will receive the explanation. Is the expertise with respect to the AI model sufficient for more abstract representations to suffice, or should there be an additional effort in simplifying, visceralizing, or re-interpreting the explanations? Should the explanations display adaptiveness? To what extent? How can transitional interfaces [8] support users with varying visualization and AI literacy in making sense of the models? Furthermore, the **perspective** of the explanation should maximize sense-making. Is redundancy helpful or obtrusive (e.g., showing the same explanation in multiple ways simultaneously)? Are formal and exhaustive details needed, or are approximations sufficient? Lastly, the **context** of the explanations (e.g., individual versus collaborative settings, synchronous versus asynchronous tasks) certainly also influences how the explanations should be delivered, and should therefore be carefully studied as well.

Besides our proposed design space (which we believe can be both extended and more granular), one could think of other methods to board HUI XAI. For instance, one could take existing XAI tools and redesign them to leverage HUIs, in what can only be described as completely hypothetical, exploratory research. Then, through user studies that emulate the evaluations of the original tools, one could assess if the resulting approaches outperform the originals with respect to task completion, sense-making, or even user satisfaction. Although only briefly hinted at so far, we make note that the novelty effect of HUIs may also influence the sentiment of users with respect to AI models and their explanations.

5 Conclusion

In this position paper, we sketched our approach to combine two highly contrasting fields: Hybrid User Interfaces (HUIs) and eXplainable Artificial Intelligence (XAI). This was motivated by their similarly widespread influence and applicability to interdisciplinary fields. To explore this combination, we discussed possible benefits through MR devices in HUIs, the types of XAI that use visual analytics, and challenges related to their combination. Lastly, we envisioned possible future scenarios where HUI XAI systems can find application.

We hold no doubts that both XAI and HUI will continue to rapidly integrate into our daily lives and work environments. Therefore, we want to motivate the following research on XAI to investigate how it can leverage HUI. Furthermore, we firmly believe that AI will even become an integral part of future HUI systems to, for example, automatically place distributed information in an immersive and shared environment, or configure the hybrid space. It follows that eventually, this also has to be explained – within the same HUI it supports, or from an external perspective.

Acknowledgments

We thank Ricardo Langner for his valuable input on the discussions that motivated this position paper. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) by grant 389792660 as part of TRR 248 – CPEC (see `https://cpec.science`), and by the German Federal Ministry of Education and Research (BMBF, SCADS22B) and the Saxon State Ministry for Science, Culture and Tourism (SMWK) by funding the competence center for Big Data and AI "ScaDS.AI Dresden/Leipzig".

References

[1] G. Alicioglu and B. Sun. A survey of visual analytics for explainable artificial intelligence methods. *Computers & Graphics*, 102:502–520, 2022. doi: 10.1016/j.cag.2021.09.002

[2] U. Alon and E. Yahav. On the bottleneck of graph neural networks and its practical implications. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021.

[3] C. Alrabbaa, F. Baader, S. Borgwardt, R. Dachselt, P. Koopmann, and J. Méndez. Evonne: Interactive proof visualization for description logics (system description). In *Automated Reasoning - 11th International Joint Conference, IJCAR 2022*, Lecture Notes in Computer Science. Springer, 8 2022. doi: 10.1007/978-3-031-10769-6_16

[4] C. Alrabbaa, F. Baader, S. Borgwardt, P. Koopmann, and A. Kovtunova. Finding good proofs for description logic entailments using recursive quality measures. In A. Platzer and G. Sutcliffe, eds., *Automated Deduction – CADE 28*, pp. 291–308. Springer International Publishing, Cham, 2021. doi: 10.1007/978-3-030-79876-5_17

[5] S. Aminah, N. Hidayah, and M. Ramli. Considering ChatGPT to be the first aid for young adults on mental health issues. *Journal of Public Health*, p. fdad065, 05 2023. doi: 10.1093/pubmed/fdad065

[6] M. G. Augasta and T. Kathirvalavakumar. Rule extraction from neural networks — a comparative study. In *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)*, pp. 404–408, 2012. doi: 10.1109/ICPRIME.2012.6208380

[7] R. Ball, C. North, and D. A. Bowman. Move to improve: Promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pp. 191–200. Association for Computing Machinery. doi: 10.1145/1240624.1240656

[8] M. Billinghurst, H. Kato, and I. Poupyrev. The MagicBook - moving seamlessly between reality and virtuality. 21(3):6–8. doi: 10.1109/38.920621

[9] N. Bressa, H. Korsgaard, A. Tabard, S. Houben, and J. Vermeulen. What's the situation with situated visualization? a survey and perspectives on situatedness. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):107–117, 2022. doi: 10.1109/TVCG.2021.3114835

[10] G. Y. Chan, E. Bertini, L. G. Nonato, B. Barr, and C. T. Silva. Melody: Generating and visualizing machine learning model summary to understand data and classifiers together. *CoRR*, abs/2007.10614, 2020.

[11] J. Choo and S. Liu. Visual analytics for explainable deep learning. *IEEE Computer Graphics and Applications*, 38(4):84–92, 2018. doi: 10.1109/MCG.2018.042731661

[12] A. Colley, K. Väänänen, and J. Häkkilä. Tangible explainable ai - an initial conceptual framework. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*, MUM '22, p. 22–27. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3568444.3568456

[13] R. Confalonieri, L. Coba, B. Wagner, and T. R. Besold. A historical perspective of explainable artificial intelligence. *WIREs Data Mining and Knowledge Discovery*, 11(1):e1391, 2021. doi: 10.1002/widm.1391

[14] R. Dachselt, S. Gaggl, M. Krötzsch, J. Méndez, D. Rusovac, and M. Yang. Nexas: A visual tool for navigating and exploring argumentation solution spaces. In *Proceedings of the 9th International Conference on Computational Models of Argument*, 9 2022. doi: 10.3233/FAIA220146

[15] R. Doerner, W. Broll, B. Jung, P. Grimm, M. Göbel, and R. Kruse. *Introduction to Virtual and Augmented Reality*, pp. 1–37. Springer International Publishing, Cham, 2022. doi: 10.1007/978-3-030-79062-2_1

[16] M. O. Ellenberg, M. Satkowski, W. Luo, and R. Dachselt. Spatiality and Semantics - Towards Understanding Content Placement in Mixed Reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA'23. ACM. doi: 10.1145/3544549.3585853

[17] W. Fan, Y. Ma, Q. Li, Y. He, E. Zhao, J. Tang, and D. Yin. Graph neural networks for social recommendation. In *The World Wide Web Conference*, WWW '19, p. 417–426. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3308558.3313488

[18] L. W. Ge, Y. Cui, and M. Kay. Calvi: Critical thinking assessment for literacy in visualizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581406

[19] A. P. Gee, P. J. Escamilla-Ambrosio, M. Webb, W. Mayol-Cuevas, and A. Calway. Augmented crime scenes: Virtual annotation of physical environments for forensic investigation. In *Proceedings of the 2nd ACM Workshop on Multimedia in Forensics, Security and Intelligence*, MiFor '10, p. 105–110. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1877972.1877999

[20] D. Gunning and D. W. Aha. Darpa's explainable artificial intelligence program. *AI Magazine*, 40(2):44–58, 2019. doi: 10.1609/aimag.v40i2.2850

[21] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang. Xai—explainable artificial intelligence. *Science Robotics*, 4(37):eaay7120, 2019. doi: 10.1126/scirobotics.aay7120

[22] F. Hohman, M. Kahng, R. Pienta, and D. H. Chau. Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2674–2693, 2019. doi: 10.1109/TVCG.2018.2843369

[23] S. Hubenschmid, J. Zagermann, S. Butscher, and H. Reiterer. Stream: Exploring the combination of spatially-aware tablets with augmented reality head-mounted displays for immersive analytics. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery. doi: 10.1145/3411764.3445298

[24] M. R. Islam, M. U. Ahmed, S. Barua, and S. Begum. A systematic review of explainable artificial intelligence in terms of different application domains and tasks. *Applied Sciences*, 12(3), 2022. doi: 10.3390/app12031353

[25] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk. Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 3227–3236. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2702123.2702180

[26] D. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann. *Mastering the information age solving problems with visual analytics*. Eurographics Association, 2010. doi: 10.2312/14803

[27] M. Kraus, J. Fuchs, B. Sommer, K. Klein, U. Engelke, D. Keim, and F. Schreiber. Immersive analytics with abstract 3d visualizations: A survey. *Computer Graphics Forum*, 41(1):201–229, 2022. doi: 10.1111/cgf.14430

[28] B. La Rosa, G. Blasilli, R. Bourqui, D. Auber, G. Santucci, R. Capobianco, E. Bertini, R. Giot, and M. Angelini. State of the art of visual analytics for explainable deep learning. *Computer Graphics Forum*, 42(1):319–355, 2023. doi: 10.1111/cgf.14733

[29] R. Langner, M. Satkowski, W. Büschel, R. Dachselt, W. Buschel, and R. Dachselt. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. In *Conference on Human Factors in Computing Systems - Proceedings*. ACM. doi: 10.1145/3411764.3445593

[30] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. Data visceralization: Enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1095–1105, 2021. doi: 10.1109/TVCG.2020.3030435

[31] M. Lisnic, C. Polychronis, A. Lex, and M. Kogan. Misleading beyond visual tricks: How people actually lie with charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580910

[32] Z. Lyu, J. Li, and B. Wang. Aiive: Interactive visualization and sonification of neural networks in virtual reality. In *2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 251–255, 2021. doi: 10.1109/AIVR52153.2021.00057

[33] N. Meissler, A. Wohlan, N. Hochgeschwender, and A. Schreiber. Using visualization of convolutional neural networks in virtual reality for machine learning newcomers. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 152–1526, 2019. doi: 10.1109/AIVR46125.2019.00031

[34] J. Méndez, C. Alrabbaa, P. Koopmann, R. Langner, F. Baader, and R. Dachselt. Evonne: A visual tool for explaining reasoning with owl ontologies and supporting interactive debugging. *Computer Graphics*

Forum, 3 2023. doi: 10.1111/cgf.14730

[35] M. Pascher, K. Kronhardt, and J. Gerken. Physicaltwin: Mixed reality interaction environment for ai-supported assistive robots. 2023.

[36] L. Pavanatto, C. North, D. A. Bowman, C. Badea, and R. Stoakley. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 759–767, 2021. doi: 10.1109/VR50410.2021.00103

[37] I. Psychoula, A. Gutmann, P. Mainali, S. H. Lee, P. Dunphy, and F. A. P. Petitcolas. Explainable machine learning for fraud detection. *CoRR*, abs/2105.06314, 2021.

[38] P. Reipschlager, T. Flemisch, and R. Dachselt. Personal Augmented Reality for Information Visualization on Large Interactive Displays. 27(2):1182–1192. doi: 10.1109/TVCG.2020.3030460

[39] P. Reipschläger, S. Engert, and R. Dachselt. Augmented displays: Seamlessly extending interactive surfaces with head-mounted augmented reality. In *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1–4. Association for Computing Machinery. doi: 10.1145/3334480.3383138

[40] M. T. Ribeiro, S. Singh, and C. Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, p. 1135–1144. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2939672.2939778

[41] M. Satkowski, F. Kessler, S. Narciss, and R. Dachselt. Who benefits from Visualization Adaptations? Towards a better Understanding of the Influence of Visualization Literacy. In *2022 IEEE Visualization and Visual Analytics (VIS)*, pp. 90–94. doi: 10.1109/VIS54862.2022.00027

[42] M. Satkowski, W. Luo, and R. Dachselt. Towards In-situ Authoring of AR Visualizations with Mobile Devices. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 324–325. IEEE. doi: 10.1109/ISMAR-Adjunct54149.2021.00073

[43] M. Satkowski, R. Rzayev, E. Goebel, and R. Dachselt. ABOVE & BELOW: Investigating Ceiling and Floor for Augmented Reality Content Placement. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 518–527. doi: 10.1109/ISMAR55827.2022.00068

[44] M. Sundararajan, A. Taly, and Q. Yan. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, p. 3319–3328. JMLR.org, 2017.

[45] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang. What if the devil is my guardian angel: Chatgpt as a case study of using chatbots in education. *Smart Learning Environments*, 10(1):15, Feb 2023. doi: 10.1186/s40561-023-00237-x

[46] C. D. Vescovo, M. Horridge, B. Parsia, U. Sattler, T. Schneider, and H. Zhao. Modular structures and atomic decomposition in ontologies. *J. Artif. Intell. Res.*, 69:963–1021, 2020. doi: 10.1613/jair.1.12151

[47] C. Wang and P. An. A mobile tool that helps nonexperts make sense of pretrained cnn by interacting with their daily surroundings. In *Adjunct Publication of the 23rd International Conference on Mobile Human-Computer Interaction*, MobileHCI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3447527.3474873

[48] J. Wexler, M. Pushkarna, T. Bolukbasi, M. Wattenberg, F. Viegas, and J. Wilson. The what-if tool: Interactive probing of machine learning models. *IEEE Transactions on Visualization & Computer Graphics*, 26(01):56–65, jan 2020. doi: 10.1109/TVCG.2019.2934619

[49] A. P. Wright, O. Shaikh, H. Park, W. Epperson, M. Ahmed, S. Pinel, D. H. P. Chau, and D. Yang. Recast: Enabling user recourse and interpretability of toxicity detection models with interactive visualization. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW1), apr 2021. doi: 10.1145/3449280

[50] X. Xu, A. Yu, T. R. Jonker, K. Todi, F. Lu, X. Qian, J. a. M. Evangelista Belo, T. Wang, M. Li, A. Mun, T.-Y. Wu, J. Shen, T. Zhang, N. Kokhlikyan, F. Wang, P. Sorenson, S. Kim, and H. Benko. Xair: A framework of explainable ai in augmented reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581500

[51] J. Zagermann, S. Hubenschmid, P. Balestrucci, T. Feuchtner, S. Mayer, M. O. Ernst, A. Schmidt, and H. Reiterer. Complementary interfaces for visual computing. 64(4-5):145–154. doi: 10.1515/itit-2022-0031