Explainable Artificial Intelligence Across Various Scales of Interaction and Experience, From Wearable to Ambient

Radu-Daniel Vatavu

MintViz Lab, Ştefan cel Mare University of Suceava, Suceava 720229, Romania

Abstract

AI embodiments take on a diversity of forms, effectively sustaining diverse contexts of use for mobile, wearable, and ambient device users. This diversity enables a variety of interactions with AI systems and devices as well as heterogeneous user experiences, from the scale of a tiny wearable device to that of a large smart building, which are notably different in their nature and complexity. This paper presents an overview of possible scales of interaction and user experience, abstracted through Kuniavsky's hierarchy, and discusses implications of how the explanator level, from Schwalbe and Finzel's taxonomy of XAI, should adapt in terms of output and interactivity to users engaging with AI at various physical scales.

Keywords

ambient intelligence, wearable intelligence, explainability, scales of experience

1. Introduction

AI embodies a wide range of form factors, from the personalized recommendations offered by a video streaming application on a TV [1] to the thematically relevant suggestions provided by a writing assistant on a desktop PC [2], customized navigation cues delivered through a mobile device [3], and the context-aware guidance offered by a personal assistant that is always available via the voice interface of a tiny device worn at the wrist [4]. Recent consumer gadgets, such as Bee,[1] Limitless,[2] and AI Pin,[3] provide always-available access to AI assistance through their inconspicuous form factors—for example, "Bee sits quietly in the background, learning your patterns, preferences and relationships over time, building a deeper understanding of your world without demanding your attention."[1] Beyond personal devices, AI is increasingly integrated into our everyday living environments, transforming them into smart spaces that incorporate a heterogeneity of interactive devices [5] and evoking a sense of ambient-distributed intelligence [6]. Moreover, beyond fixed environments, AI permeates larger spaces, addressing smart mobility, smart buildings, and smart cities. These varied forms of presence and form

Joint Proceedings of the ACM IUI Workshops 2025, March 24-27, 2025, Cagliari, Italy
radu.vatavu@usm.ro (R. Vatavu)
http://eed.usv.ro/~vatavu (R. Vatavu)
0000-0002-7631-6445 (R. Vatavu)

[1] https://www.bee.computer
[2] https://www.limitless.ai
[3] https://humane.com

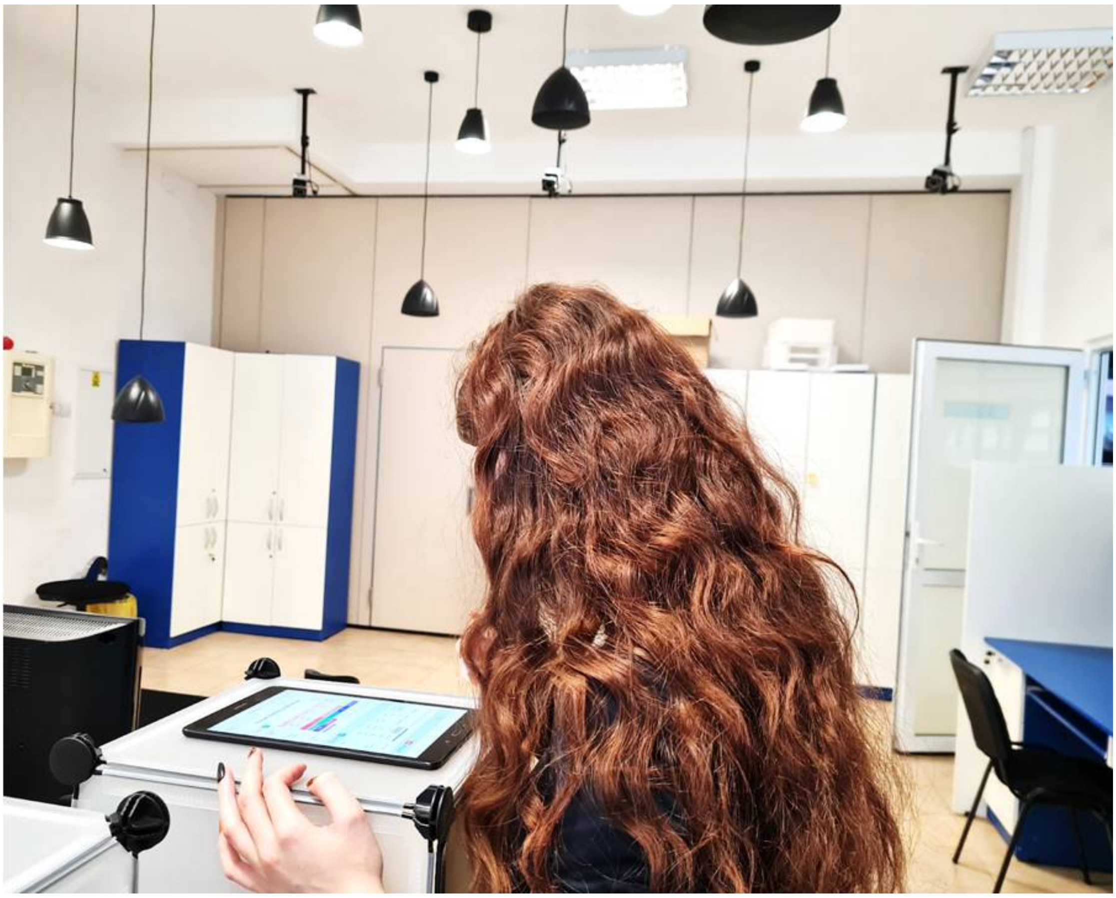

Figure 1: An ambient intelligence environment featuring smart lighting, sound, and airflow installations, which are directly controllable through a user interface available on a mobile or wearable device, but which can also serve as novel modalities for implementing explainability in a way that is characteristic to the specific scale of interaction and experience enabled by the environment.

factors make engaging with AI an incredibly versatile experience in our lives, from the tiniest wearable devices to the largest urban spaces.

In this context, we argue that the scale of the experience when engaging with AI in various forms necessitates adaptations to existing explainability techniques, making XAI inherently dependent on the scale at which interactions between users and AI may take place. For example, explainability delivered through a small wrist wearable, a public interactive display, or a smart building is likely to take different forms in terms of output modalities and the level of interactivity expected from the user. The smart environment illustrated in Figure 1 reveals a variety of opportunities for interaction, ranging from personal devices to natural modalities, and correspondingly, offers different ways for an explanator to engage with users through ambient media [7] and augmented reality media [8], respectively. The specific embodiment of AI and the interactive experience it evokes are thus crucial to effective explainability and, consequently, to end-user trust that the AI is making sound decisions. In this context, explainability implemented for smart chairs [9] could adopt the output modalities that smart chairs would naturally support, such as actuation of their constituent parts [10]; explainability implemented for a smart room could adopt subtle displays of light and sound [7]; whereas explainability implemented for an wearable delivering electrical muscle stimulation [11] would use body pose and movement.

This position paper represents a preliminary exploration into the relationship between the scale of user experience and interaction and the necessary qualities of an explanator that would align with that experience. It argues that specific AI embodiments necessitate specific modalities to implement explainability in ways that would not break, but rather amplify and augment the user experience of engaging with the AI.

2. Scales of User Experience and Interaction and Implications for AI Explainability

Following Schwalbe and Finzel's [12] taxonomy for XAI, the high-level overview for building an explanation system involves the following three levels:

- The *Problem Definition* level encompasses the traits of the task (task type, such as classification, clustering, regression, etc., and input data types, such as symbolic or non-symbolic data) and the level of interpretability of the explanandum (i.e., what is to be explained).
- At the *Explanator* level, functionally is divided into input (required inputs, degree of portability to other input types, explanation locality), output (object of explanation, output type, and presentation aspects), interactivity with the user, and any further formal constraints posed on the explanator (such as the number of iterations).
- The *Metrics* level considers the metrics that can be used to assess the quality of XAI methods according to subjective human evaluation, with categories represented by functionally-grounded metrics (which are independent of human judgment), human-grounded metrics (where subjective human judgment is required), and application-grounded metrics (representing full human-AI collaboration).

We also rely on Kuniavsky's [13] hierarchy of user experience scales relevant for ambient intelligence systems and interactions, which encompasses six distinct scales—*Covert, Mobile, Personal, Environmental, Architectural,* and *Urban*—as follows:

- At the *Covert* scale, characterized by interactions requiring a physical space on the order of centimeters, the experience is entirely delivered through wearable devices in close contact with the user's body. Examples include smartwatches, armbands, rings and devices for finger augmentation [14] that feature specific input modalities based on voice, gesture, motion, and proximity, as well as specific output modalities, such as haptics, due to their closeness to the human body. This scale also includes clip-on consumer electronics, such as lifelogging devices [15], designed to operate independently and create digital records of their users' life experiences.
- At the *Mobile* scale, characterized by physical interactions on the order of tens of centimeters, the experience is delivered through smartphones, which represent the prevalent form of mobile computing today.

- At the *Personal* scale, characterized by a range on the order of one meter, the experience is delivered by systems and devices of a person-sized magnitude, such as public terminals and interactive kiosks. This scale involves interactions involving different and more parts of the body [16, 17] and, possibly, assisted interaction [18].
- At the *Environmental* scale, the experience occurs at the order of tens of meters and takes place at the level of a room. The experience usually involves more users, and potentially many, as in the case of mass-computer interaction [19].
- At the *Architectural* scale, on the order of hundreds of meters, the experience involves an entire building, combining elements of human-computer interaction, architecture, and building information modeling, sensing, and actuation technologies [20].
- At the *Urban* scale, on the order of kilometers, the experience involves an entire city or urban area, as in location-based mobile exergames [21] or location-based film experiences in an augmented place [22].

The intersection between the *Explanator* level of XAI [12] and possible *UX Scales* in ambient intelligence environments [13] is interesting to explore in how output and interactivity can be tailored to match specific user interactions with various AI embodiments. From the smallest, *covert* scale to the largest, *urban* scale, the way explanations are provided in terms of output modalities must align with user expectations for those scales. For example, non-visual output modalities, designed to minimize interference with user tasks, are preferable at the *covert* scale, as is multimodal visual-audio output for smartphones on the *mobile* scale, aligning with mobile content consumption paradigms. The *personal* scale may potentially involve more extensive explanatory dialogues, using layered or step-by-step explanations and involving the location of where the interactive system is installed. At the *environmental* scale, explanations should be adapted to address multiple users, likely to be involved into the same interaction. At the *architectural* scale, the explanations should consider the complexity of the built environment and integrated devices. Lastly, at the *urban* scale, explanations should integrate the data that becomes available from different sources involving various aspects of city life, such as living, transportation, and cultural elements.

Matching explanations to the scale of interaction and experience is key to ensuring that explanations are effective, well-received by users, and delivered through the most suitable mechanisms. We are looking forward to both theoretical and practical explorations in this direction, where adaptive XAI interfaces actively consider the scale of the experience created by engaging with AI in various embodiments and form factors.

Acknowledgments

This work was conducted within the MintViz Research Laboratory. This work is also supported by the NetZeRoCities Competence Center, funded by European Union-NextGenerationEU and Romanian Government, under the National Recovery and Resilience Plan for Romania, contract no. 760007/30.12.2022 with the Romanian Ministry of Research, Innovation and Digitalization through the specific research project P3, Smart and Sustainable Buildings. The photograph in Figure 1 was taken by Alexandru-Tudor Andrei.

References

[1] H. Steck, L. Baltrunas, E. Elahi, D. Liang, Y. Raimond, J. Basilico, Deep learning for recommender systems: A Netflix case study, AI Magazine 42 (2021) 7–18. doi:10.1609/aimag.v42i3.18140.

[2] A. Guo, L. Wang, J. Heer, A. Zhang, Preserving writer values in AI writing assistance tools, in: Proceedings of the Third Workshop on Intelligent and Interactive Writing Assistants, In2Writing '24, ACM, New York, NY, USA, 2024, p. 58–61. doi:10.1145/3690712.3690727.

[3] H. Zhang, N. J. Falletta, J. Xie, R. Yu, S. Lee, S. M. Billah, J. M. Carroll, Enhancing the travel experience for people with visual impairments through multimodal interaction: NaviGPT, a real-time AI-driven mobile navigation system, in: Companion Proceedings of the 2025 ACM International Conference on Supporting Group Work, GROUP '25, ACM, New York, NY, USA, 2025, p. 29–35. doi:10.1145/3688828.3699636.

[4] R. Arakawa, J. F. Lehman, M. Goel, PrISM-Q&A: Step-aware voice assistant on a smartwatch enabled by multimodal procedure tracking and large language models, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 8 (2024). doi:10.1145/3699759.

[5] O.-A. Schipor, R.-D. Vatavu, J. Vanderdonckt, Euphoria: A scalable, event-driven architecture for designing interactions across heterogeneous devices in smart environments, Information and Software Technology 109 (2019) 43–59. doi:10.1016/j.infsof.2019.01.006.

[6] R. Dunne, T. Morris, S. Harper, A survey of ambient intelligence, ACM Comput. Surv. 54 (2021). doi:10.1145/3447242.

[7] H. Ishii, C. Wisneski, S. Brave, A. Dahley, M. Gorbet, B. Ullmer, P. Yarin, ambientROOM: integrating ambient media with architectural space, in: CHI'98 Conference Summary on Human Factors in Computing Systems, CHI '98, ACM, New York, NY, USA, 1998, p. 173–174. doi:10.1145/286498.286652.

[8] R. T. Azuma, The most important challenge facing augmented reality, Presence: Teleoper. Virtual Environ. 25 (2016) 234–238. doi:10.1162/PRES_a_00264.

[9] A.-T. Andrei, L.-B. Bilius, R.-D. Vatavu, Take a seat, make a gesture: Charting user preferences for on-chair and from-chair gesture input, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24, ACM, New York, NY, USA, 2024, pp. 1–17. doi:10.1145/3613904.3642028.

[10] D. Menheere, I. Damen, C. Lallemand, S. Vos, Ivy: A qualitative interface to reduce sedentary behavior in the office context, in: Companion Publication of the 2020 ACM Designing Interactive Systems Conference, DIS' 20 Companion, ACM, New York, NY, USA, 2020, p. 329–332. doi:10.1145/3393914.3395822.

[11] T. Duente, J. Schulte, M. Pfeiffer, M. Rohs, MuscleIO: Muscle-based input and output for casual notifications, Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2 (2018). doi:10.1145/3214267.

[12] G. Schwalbe, B. Finzel, A comprehensive taxonomy for explainable artificial intelligence: A systematic survey of surveys on methods and concepts, Data Mining and Knowledge Discovery 38 (2024) 3043–3101. doi:10.1007/s10618-022-00867-8.

[13] M. Kuniavsky, Smart Things: Ubiquitous Computing User Experience Design, Morgan

Kaufmann, USA, 2010. URL: https://www.elsevier.com/books/smart-things/kuniavsky/978-0-12-374899-7.

[14] R. Shilkrot, J. Huber, J. Steimle, S. Nanayakkara, P. Maes, Digital digits: A comprehensive survey of finger augmentation devices, ACM Comput. Surv. 48 (2015). doi:10.1145/2828993.

[15] C. Gurrin, A. F. Smeaton, A. R. Doherty, Lifelogging: Personal big data, Found. Trends Inf. Retr. 8 (2014) 1–125. doi:10.1561/1500000033.

[16] R. Jota, P. Lopes, D. Wigdor, J. Jorge, Let's kick it: How to stop wasting the bottom third of your large screen display, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, ACM, New York, NY, USA, 2014, p. 1411–1414. doi:10.1145/2556288.2557316.

[17] J. Müller, R. Walter, G. Bailly, M. Nischt, F. Alt, Looking glass: A field study on noticing interactivity of a shop window, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, ACM, New York, NY, USA, 2012, p. 297–306. doi:10.1145/2207676.2207718.

[18] R.-D. Vatavu, O.-C. Ungurean, L.-B. Bilius, Interactive public displays and wheelchair users: Between direct, personal and indirect, assisted interaction, in: Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology, UIST '22, ACM, New York, NY, USA, 2022, pp. 1–17. doi:10.1145/3526113.3545662.

[19] J.-Y. L. Lawson, J. Vanderdonckt, R.-D. Vatavu, Mass-computer interaction for thousands of users and beyond, in: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, CHI EA '18, ACM, New York, NY, USA, 2018, pp. 1–6. doi:10.1145/3170427.3188465.

[20] E. Economidou, A. Itzlinger, C. Frauenberger, Lived experience in human-building interaction (HBI): an initial framework, Frontiers in Computer Science 5 (2024). doi:10.3389/fcomp.2023.1233904.

[21] M. Saaty, D. Haqq, D. B. Toms, I. Eltahir, D. S. McCrickard, A study on Pokémon GO: exploring the potential of location-based mobile exergames in connecting players with nature, in: Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '21, ACM, New York, NY, USA, 2021, p. 128–132. doi:10.1145/3450337.3483481.

[22] H. Park, W. Woo, Metadata design for location-based film experience in augmented places, in: Proceedings of the 2015 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design, 2015, pp. 40–45. doi:10.1109/ISMAR-MASHD.2015.12.