

The Perceptual Gap: Why We Need Accessible XAI for Assistive Technologies

Shadab H. Choudhury

shadabc1@umbc.edu

University of Maryland, Baltimore County

Baltimore, Maryland, USA

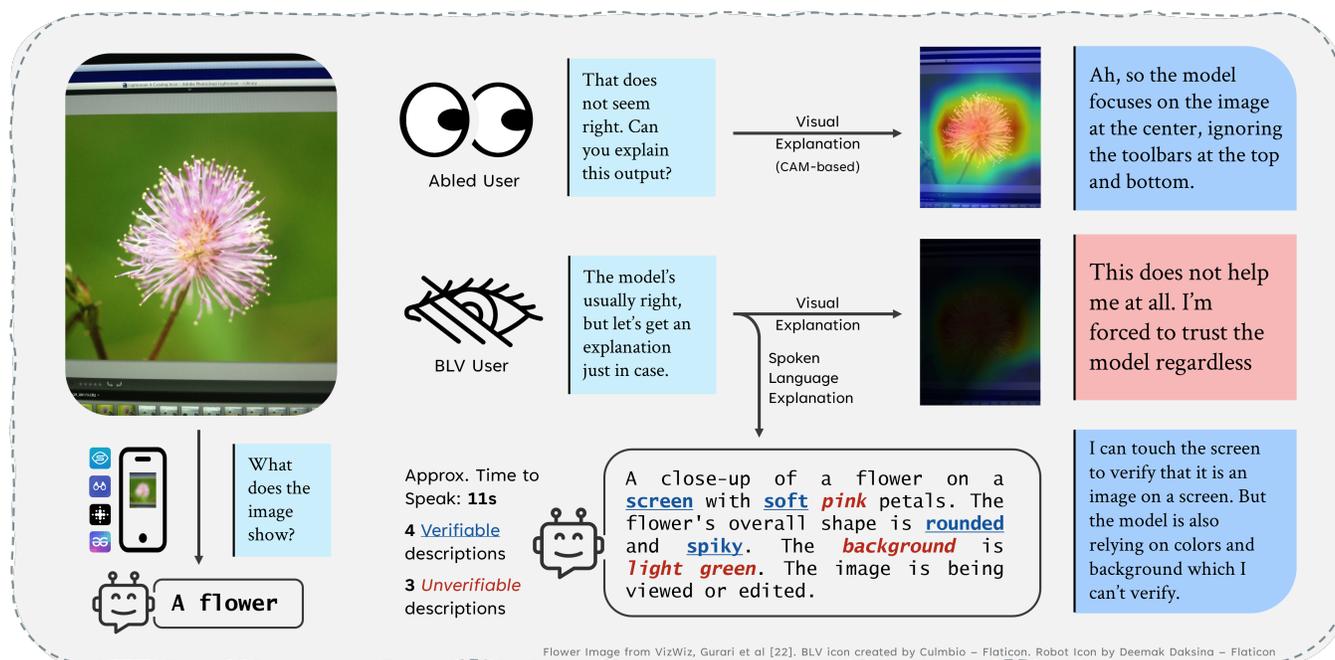


Figure 1: A depiction of the process of using an image recognition app, how visual explainability methods are flawed for the primary target userbase of such applications, and how verifiability remains a problem with language descriptions.

Abstract

Artificial intelligence systems are widely used by people with sensory disabilities, like loss of vision or hearing, to help perceive or navigate the world around them. This includes tasks like describing an image or object they cannot touch, reading documents, automatically captioning speech, and so on. Presently, models used for these tasks are based on deep neural networks and are thusly black boxes. Explainable AI (XAI) describes methods that can explain why a model gave the output it did. However, existing XAI methodologies are rarely accessible or designed with disabled users in mind. In this paper, we survey existing work in XAI with a focus on human-centered and accessibility-centered approaches or evaluations. We show that there is next-to-no XAI work that accounts for people

with sensory disabilities, that many typical explanations are difficult for them to comprehend, and propose possible avenues for future work in Accessible Human-Centered XAI.

CCS Concepts

• **Human-centered computing** → **Accessibility**; *Accessibility technologies*; • **Computing methodologies** → **Artificial intelligence**; *Speech recognition*; *Computer vision*.

Keywords

Accessibility, Explainable AI, XAI, Artificial Intelligence, Position, Review



This work is licensed under a Creative Commons Attribution 4.0 International License. CHI EA '26, Barcelona, Spain

© 2026 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/10.1145/3772363.3799105>

ACM Reference Format:

Shadab H. Choudhury. 2026. The Perceptual Gap: Why We Need Accessible XAI for Assistive Technologies. In *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3772363.3799105>

1 Introduction

Using Artificial Intelligence (AI), digital systems can now emulate human behavior under a strict, narrow scope. For example, an object recognition model, given an image of an object, can return the name or description of that object, just like a person looking at the picture could. Assistive devices and software designed today regularly take advantage of these capabilities. An individual who is blind or has low-vision (BLV) could use such a model to “see” nearby objects without having to touch and feel them. Apps such as *Seeing AI*¹, *Lookout*², *Envision*³, and *TapTapSee*⁴ all use AI. These are not niche applications either — Lookout and TapTapSee both have over 500k downloads on Android devices, while the other two have over 100k downloads.

Similarly, an individual who is deaf or hard-of-hearing (DHH) can use a speech captioning model to convert spoken language that they can’t hear to written language that they can read. *Otter*⁵ is the most widely used standalone app for this (with north of 5 million downloads on Android), but most modern devices and video calling software now also have built-in closed captioning.

How does one know if the output of a model is accurate or reliable? One way is to have an intimate knowledge of how the model was designed and trained. The other way is to study the way the model processes an input and arrives to a particular conclusion. ‘Interpretability’ and ‘Explainability’ both cover this from different perspectives — Interpretability focuses on understanding how the model works in a way that machine learning practitioners can use the insights [31]. Explainability attempts to describe or exhibit the model’s behaviour in such a way that a layperson can understand how the model works to some extent [14]. Explainability is generally *post-hoc*, which means there is no need to dive into a model’s internals or design processes to understand it.

Explainability methods like SHAP [35] or LIME [42] explain the outputs by listing how important the different components of the input (a.k.a. features) were for the final output. However, this is only feasible for inputs with few features, such as a table of values. CAM or GRAD-CAM [43, 55] are more useful for the object recognition models typically used in apps for BLV users; they generate a ‘heat map’ overlaid on top of the original input, showing which parts of the image were used to recognize the object.

At this stage, one can recognize the obvious issue: GRAD-CAM points to different parts of the image. But someone who is blind or low-vision may not be able to perceive the image in the first place. The explanation is useless to them! There’s a *perceptual gap* created by the fact the inputs and explanations are both in a particular modality, while the user’s perception or mode of sensing isn’t.

Users without any sensory disabilities are always able to examine the input to determine if a model is reliable or not [8, 17, 24]. But disabled users are typically unable to do that without external support [3], which is a major blow to their independence. To make

matters worse, the majority of the applications listed above are proprietary, so it is also not possible for external reviewers or auditors to analyze these models either.

Explainability is also crucial for accountability [4, 13]. If an AI system makes a mistake, who is at fault? The designers of the system for not making it sufficiently robust? The user for using it in scenarios it was unintended for? The annotators further up in the pipeline for inaccurately labelling the training data? The blame game is an easy one to play, but without any explanation for the model’s behaviour it’s also a fruitless endeavour. And once again, disabled users are unable to make effective use of this facility.

Hence, we ask two questions:

RQ1: To what extent does existing work in Explainable AI support or center the needs of users with disabilities?

RQ2: How can researchers in Accessibility and Explainable AI work together to improve the accessibility and usage of XAI in assistive technologies?

To that end, we survey existing literature in Accessibility and XAI to understand where researchers have highlighted such concerns previously. We find only a handful of works have considered this problem, and that there has been almost no work done on integrating Accessible XAI in assistive technology. Following the survey, we offer some suggestions on how can XAI methodology be made more accessible.

Broadly, we offer a provocation to Explainable AI: *How can people participate as shareholders by acknowledging explanations and contest errors in AI outputs when they are incapable of perceiving the ground truth?*

1.1 Explainability and Accessibility

The ways in which XAI fails people with disabilities may not be intuitive to abled people. Before focusing on the two research questions, an accounting of existing XAI methods, the scenarios where they are commonly applied, and how they fail people with disabilities, is needed.

1.1.1 Explaining Vision. Vision, both as a modality and as an accessibility problem, is more thoroughly represented than any other. A majority of accessibility work is done for the benefit of BLV people [36], and a majority of research into human senses in general also focuses on visual perception [26]. An extremely wide body of research exists on Visual XAI, arising mainly from the original CAM paper, such as GradCAM, SmoothCam, ScoreCAM, LayerCAM [27, 39, 43, 50]. An alternative are Prototypes [12, 18], which take parts of the original input directly rather than generating heat maps. Visual models also easily lend themselves to counterfactual explanations [10, 49].

BLV users are excluded from almost all of those aforementioned methods, since they can perceive neither the original image, nor the generated heat map or prototype. Existing apps for BLV users tend to rely on uncertainty percentages, which don’t ‘explain’ much [3]. That said, there is one method they aren’t excluded from: Language.

Visual models can also have language-based explanations [23], as shown in Figure 1. Language is versatile and can be expressed in multiple modalities: visually (text), aurally (speech), and tactilely (braille), making it uniquely suited to accessibility contexts. However, they still have drawbacks — speaking can be slow, while

¹<https://www.seeingai.com/>

²<https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal>

³<https://www.letsenvision.com/>

⁴<https://taptapseeapp.com/>

⁵<https://otter.ai/>

braille displays are specialized devices not available everywhere. Explanations can be unhelpful, focusing on the wrong details, or simply be incorrect [11, 21, 46].

1.1.2 Explaining Speech and Language. There is not as much work on the Explainability of speech and language models in general. A wide range of tasks fall under this category: captioning speech, describing or classifying sounds in general, or recognizing and generating sign language.

For speech captioning, feature-based explanations are often used [1], alongside breaking down individual phonemes [53]. Confidence scores can also help give users some idea of which parts of the captions are correct or incorrect [33], but there are challenges for comprehensibility in real-world applications [5].

For audio descriptions, features, spectrogram images (which are difficult for laypeople to parse), and language descriptions [2] are used. Sign language is typically visual, so there’s no perceptual gap for deaf users. However, deaf-blind users still face challenges, and research on tactile sign languages is very sparse.

The perceptual gap in Audio XAI isn’t as severe as in Visual XAI. Some existing explanation methods in audio, such as spectrograms, are already visual rather than auditory. Reading text is also much faster than listening to speech, so language descriptions work better. That said, simply explaining isn’t the entire goal here: good explanations can teach users how the models work, and very little work has been done in this aspect for audio.

2 Previous Work

2.1 Search Methodology

For **RQ1**, we focus our search on the ACM Digital Library, where the major venues for disseminating accessibility and human-computer interaction research publish their proceedings [36]. We select four primary venues for accessibility research: ASSETS, TACCESS, CHI, and TOCHI. Works that were published in other venues, were known to the authors prior to the search, or were found in preliminary searches, are listed under ‘Other’. We applied the following filters progressively, and Table 1. shows the resulting counts:

Access: accessib* OR disab* OR “impaired”

AI/ML: “machine learning” OR “ML” OR “computer vision” OR “object recognition” OR “image captioning” OR “automatic speech recognition” OR “ASR” OR “speech synthesis” OR “natural language processing” OR “NLP” OR “language model”

XAI: explainab* OR interpretab* OR “XAI”

Many of these papers simply use the term “explainable” or “interpretability” in a way unrelated to AI, or mention Explainable AI as a part of the introduction or background only. Hence, we manually verified each result from the final XAI filter. Other terms related to XAI, such as ‘uncertainty’, ‘calibration’ or ‘trust’ returned too many results to verify, and the majority were still unrelated to the problem at hand. So, while there may be some gaps in our review, it’s clear that both the Human-Centered AI and the Accessibility communities at large have yet to focus on this intersection.

Other venues and workshops that focus on XAI or Accessibility specifically [16, 48] also do not have any work relevant to the accessibility of XAI.

Venue	Access	AI/ML	XAI	Verified
ASSETS	1,097	426	34	3
CHI	1,083	393	54	1
TACCESS	261	117	12	1
TOCHI	422	156	47	0
Other	-	-	-	5
Total	2936	1,092	147	10

Table 1: Count of papers per venue based on the above filters.

2.2 Discussion

These results are unfortunate, but not unexpected. Previous work [41, 45, 47] has shown only a small fraction of XAI research actually validates their systems through user studies, and the intersection of sensory disabilities and XAI is even more niche.

The closest work to this paper is a similar survey on Accessible Explainable AI by Nwokoye et al. [38]. Compared to that survey, we focus more more narrowly on the HCI and Accessibility communities, while casting a broader net in terms of **RQ1** – we also discuss upstream problems that lead to the need for XAI, or parallel works that help users understand the models without explicitly using XAI methodology. As a position paper, we aim to also situate this problem in the broader canon of human-centered XAI.

At first glance, [15] seems unrelated to accessibility. However, it uses the example of BLV or DHH visitors experiencing cultural heritages as a place where accessible XAI may be needed. Wolf and Ringland [52] similarly notes that all explanations should be tailored to specific audiences and use appropriate interfaces. Peixoto et al. [41] both surveys XAI research and sets forth similar guidelines.

Both Alharbi et al. [3] and Fernando et al. [19] approach the problem of trust and verifiability for BLV users of accessibility software qualitatively. The interviewed participants highlight multiple problems with existing applications: that they only express uncertainty through a single ‘confidence’ score value, and it takes a long time or a lot of effort to interact with responses from the app. These issues boil down to the fact that XAI methods are not designed for users like them. Khan et al. [32] focuses specifically on the issues brought up in the reviews of apps used by BLV users. While most reviews are positive, some reviews also mention similar problems with accuracy and the consequences of an incorrect output. Hong and Kacorri [25] go into the details of how blind users handle errors, and note how unsuitable existing XAI techniques are for blind users.

ImageExplorer [34] uses scene graphs to break down images into components for easier interrogation. This lets users interact with the image and understand if an output from a model is accurate or not. While not strictly using an XAI method, this approach is similar to how abled users might interact with prototypical or heat map explanations.

Finally, Peixoto et al. [40] directly engages with the problem of XAI in Accessibility tools. They discuss how BLV users can interact with audio descriptions as XAI for automatic video captioning, and also survey existing work in sound-based XAI.

All of the above focus on visually-disabled users or visual models. What about audio and speech? As mentioned in Section 1.1.2, audio suffers from less of a perceptual gap. However, XAI is still both

underutilized and understudied in this area, and a few works such as [9, 30] have called attention to it.

3 Making XAI Accessible

In this section, we focus on **RQ2** and offer some suggestions to XAI, Accessibility and even Visualization researchers to help solve the problems described so far. This is not an exhaustive list, and is phrased as questions and suggestions for future research directions that would improve the trustworthiness of assistive technology that uses AI, allowing disabled users to maintain a greater degree of independence.

3.1 What is Verifiable and what is not?

As Figure 1 shows, a language description may have some components that can be verified using other senses (touch, hearing or smell) while other components may be hard or impossible to verify (background details or colour). The 'verifiability' of a component can also depend on the context. E.g., the background of an image inside one's own house is verifiable, but not of an outdoors image.

In the near term, future work could consider integrating the idea of verifiability into multimodal models (Sections 3.2 and 3.5). In the long run, we need to understand what people with sensory disabilities consider 'verifiable', and how to personalize such systems to account for varying contexts, use cases, and individual needs.

3.2 Where do Language Descriptions fall short, and how can they be improved?

Existing language descriptions rely on standard visual captioning models, which are annotated by abled people. When used in accessibility applications downstream, they often veer away from disabled users' needs [20]. Gurari et al. [22]'s VizWiz dataset takes a step in the right direction, involving disabled people in the data collection part. However, the labelling and captioning afterwards were still done by abled people.

In the near term, researchers should aim to simply create more accurate vision-language models that give verifiable (Section 3.1) descriptions, or emulate the perceptual experience of disabled users better. In the long term, people with disabilities should be included at every step of the process of training models, which means we must design new interfaces and procedures for them to interact with training data and model evaluations in detail.

3.3 Can XAI Visualizations be made more Accessible?

Feature-based explanations are typically visualized as graphs or charts [6], which can be effectively described [44]. Heat-map-based visual explanations pose a greater challenge, but could be combined with methods like [34] and [7] to create accessible explanations.

Many methods of visualization are now more accessible [25]. In future research, we encourage Visualization, XAI and Accessibility researchers to collaborate and implement these advances into existing assistive software or into the training process of machine learning models.

3.4 Can better Explanations lead to better Teaching?

Teachable AI describes systems that are designed from the ground up to be personalizable to the needs of a very small group or even a single individual [28, 29, 37, 51].

Since Teachable AI systems are not infallible, future work in this area should consider integrating XAI into them. That would both help disabled users better understand the cause behind those failures [31] and 'teach' the systems better, while also improving education and trust about such AI systems in general.

3.5 Can Explaining Multimodal *inputs* improve Trust and Verifiability?

People with sensory disabilities rely on multiple other senses. For instance, A DHH person might feel vibrations on the floor, and intuit someone's running around nearby even if they cannot see the other person. This principle can be extended to modern multimodal models as well. Even tactile models [54] could be used here, basing outputs off vibrations, impact, and the feel of different materials.

In future work, multimodal inputs can be a pathway to improving verifiability (Section 3.1) in assistive technologies.

4 Conclusion

Many modern assistive apps and tools now use machine learning models. Since these models are not always reliable, it is important that users are able to understand the processes or verify the outputs. Explainable AI is the primary tool for solving this problem. However, due to a perceptual gap, users with sensory disabilities are excluded from most XAI methods.

Hence, we propose some general research directions focused on two key points — that explanations and outputs should be designed in such a way as to be verifiable by disabled users, and that disabled users should be able to participate at all steps of the training process of the models they will use. We hope this paper encourages researchers to look deeper at the largely understudied intersection of Accessibility and Explainable AI, to improve trust and reliability, and allow people with disabilities to empower their use of assistive technologies and maintain independent lives.

References

- [1] Alican Akman and Björn W. Schuller. 2024. Audio Explainable Artificial Intelligence: A Review. *Intelligent Computing* 3 (Jan. 2024), 0074. doi:10.34133/icomputing.0074 Publisher: American Association for the Advancement of Science.
- [2] Alican Akman, Qiyang Sun, and Björn W. Schuller. 2025. Improving Audio Explanations Using Audio Language Models. *IEEE Signal Processing Letters* 32 (2025), 741–745. doi:10.1109/LSP.2025.3532218
- [3] Rahaf Alharbi, Pa Lor, Jaylin Herskovitz, Sarita Schoenebeck, and Robin N. Brewer. 2024. Misfitting With AI: How Blind People Verify and Contest AI Errors. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3663548.3675659
- [4] Victoria Bellotti and Keith Edwards. 2001. Intelligibility and Accountability: Human Considerations in Context-Aware Systems. *Human-Computer Interaction* (Dec. 2001). doi:10.1207/S15327051HCI16234_05 Publisher: Lawrence Erlbaum Associates, Inc..
- [5] Larwan Berke, Christopher Caulfield, and Matt Huenerfauth. 2017. Deaf and Hard-of-Hearing Perspectives on Imperfect Automatic Speech Recognition for Captioning One-on-One Meetings. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. Association

- for Computing Machinery, New York, NY, USA, 155–164. doi:10.1145/3132525.3132541
- [6] Deepshikha Bhati, MD Amiruzzaman, Ye Zhao, Angela Guercio, and Tram Le. 2025. A Survey of Post-Hoc XAI Methods From a Visualization Perspective: Challenges and Opportunities. *IEEE Access* 13 (2025), 120785–120806. doi:10.1109/ACCESS.2025.3581136
- [7] Brandon Biggs, Lena Yusim, and Peter Coppin. 2018. The audio game laboratory: Building maps from games. Houghton, USA. http://icad2018.icad.org/wp-content/uploads/2018/06/ICAD2018_paper_10.pdf
- [8] Daniel Bolya, Sean Foley, James Hays, and Judy Hoffman. 2020. TIDE: A General Toolbox for Identifying Object Detection Errors. In *Computer Vision – ECCV 2020*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer International Publishing, Cham, 558–573. doi:10.1007/978-3-030-58580-8_33
- [9] Danielle Bragg, Naomi Caselli, Julie A. Hochgesang, Matt Huenerfauth, Leah Katz-Hernandez, Oscar Koller, Raja Kushalnagar, Christian Vogler, and Richard E. Ladner. 2021. The FATE Landscape of Sign Language AI Datasets: An Interdisciplinary Perspective. *ACM Trans. Access. Comput.* 14, 2 (July 2021), 7:1–7:45. doi:10.1145/3436996
- [10] Ruth M. J. Byrne. 2019. Counterfactuals in Explainable Artificial Intelligence (XAI): Evidence from Human Reasoning. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence* (Aug. 2019), 6276–6282. doi:10.24963/ijcai.2019/876 Publisher: International Joint Conferences on Artificial Intelligence Organization.
- [11] Rwei-Che Chang, Yuxuan Liu, and Anhong Guo. 2024. WorldScribe: Towards Context-Aware Live Visual Descriptions. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3654777.3676375
- [12] Chaofan Chen, Oscar Li, Daniel Tao, Alina Barnett, Cynthia Rudin, and Jonathan K Su. 2019. This Looks Like That: Deep Learning for Interpretable Image Recognition. In *Advances in Neural Information Processing Systems*, Vol. 32. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2019/hash/ad77ee2dfc142b0e11888e72b43fb75-Abstract.html>
- [13] Jennifer Cobbe, Michelle Seng Ah Lee, and Jatinder Singh. 2021. Reviewable Automated Decision-Making: A Framework for Accountable Algorithmic Systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Virtual Event Canada, 598–609. doi:10.1145/3442188.3445921
- [14] Rudresh Dwivedi, Devam Dave, Het Naik, Smriti Singhal, Rana Omer, Pankesh Patel, Bin Qian, Zhenyu Wen, Tejal Shah, Graham Morgan, and Rajiv Ranjan. 2023. Explainable AI (XAI): Core Ideas, Techniques, and Solutions. *ACM Comput. Surv.* 55, 9 (Jan. 2023), 194:1–194:33. doi:10.1145/3561048
- [15] Natalia Diaz-Rodríguez and Galena Pisoni. 2020. Accessible Cultural Heritage through Explainable Artificial Intelligence. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 317–324. doi:10.1145/3386392.3399276
- [16] Upol Ehsan, Elizabeth A Watkins, Philipp Wintersberger, Carina Manger, Nina Hubig, Saiph Savage, Justin D. Weisz, and Andreas Rieker. 2025. New Frontiers of Human-centered Explainable AI (HCXAI): Participatory Civic AI, Benchmarking LLMs, XAI Hallucinations, and Responsible AI Audits. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3706599.3706713
- [17] Rahlhal Errattahi, Asmaa El Hannani, and Hassan Ouahmane. 2018. Automatic Speech Recognition Errors Detection and Correction: A Review. *Procedia Computer Science* 128 (Jan. 2018), 32–37. doi:10.1016/j.procs.2018.03.005
- [18] Andrea Fedele, Riccardo Guidotti, and Dino Pedreschi. 2025. This Sounds Like That: Explainable Audio Classification via Prototypical Parts. In *Discovery Science: 27th International Conference, DS 2024, Pisa, Italy, October 14–16, 2024, Proceedings, Part II*. Springer-Verlag, Berlin, Heidelberg, 348–363. doi:10.1007/978-3-031-78980-9_22
- [19] Sandra Fernando, Chiemela Ndukwe, Bal Vrdee, and Ramzi Djemai. 2025. Image Recognition Tools for Blind and Visually Impaired Users: An Emphasis on the Design Considerations. *ACM Trans. Access. Comput.* 18, 1 (Jan. 2025), 1:1–1:21. doi:10.1145/3702208
- [20] Kapil Garg, Xinru Tang, Jimin Heo, Dwayne R. Morgan, Darren Gergle, Erik B. Sudderth, and Anne Marie Piper. 2025. "It's trained by non-disabled people": Evaluating How Image Quality Affects Product Captioning with VLMs. doi:10.48550/arXiv.2511.08917 arXiv:2511.08917 [cs].
- [21] Ricardo E Gonzalez Penuela, Jazmin Collins, Cynthia Bennett, and Shiri Azenkot. 2024. Investigating Use Cases of AI-Powered Scene Description Applications for Blind and Low Vision People. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3613904.3642211
- [22] Danna Gurari, Qing Li, Abigale J. Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P. Bigham. 2018. VizWiz Grand Challenge: Answering Visual Questions From Blind People. 3608–3617. https://openaccess.thecvf.com/content_cvpr_2018/html/Gurari_VizWiz_Grand_Challenge_CVPR_2018_paper.html
- [23] Lisa Anne Hendricks, Zeynep Akata, Marcus Rohrbach, Jeff Donahue, Bernt Schiele, and Trevor Darrell. 2016. Generating Visual Explanations. In *Computer Vision – ECCV 2016*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 3–19. doi:10.1007/978-3-319-46493-0_1
- [24] Derek Hoiem, Yodsawalai Chodpathumwan, and Qieyun Dai. 2012. Diagnosing Error in Object Detectors. In *Computer Vision – ECCV 2012*, Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid (Eds.). Springer, Berlin, Heidelberg, 340–353. doi:10.1007/978-3-642-33712-3_25
- [25] Jonggi Hong and Hernisa Kacorri. 2024. Understanding How Blind Users Handle Object Recognition Errors: Strategies and Challenges. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, St. John's NL Canada, 1–15. doi:10.1145/3663548.3675635
- [26] Fabian Huttmacher. 2019. Why Is There So Much More Research on Vision Than on Any Other Sensory Modality? *Frontiers in Psychology* 10 (Oct. 2019), 2246. doi:10.3389/fpsyg.2019.02246
- [27] Peng-Tao Jiang, Chang-Bin Zhang, Qibin Hou, Ming-Ming Cheng, and Yun-chao Wei. 2021. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Transactions on Image Processing* 30 (2021), 5875–5888. doi:10.1109/TIP.2021.3089943
- [28] Hernisa Kacorri. 2017. Teachable machines for accessibility. *SIGACCESS Access. Comput.* 119 (Nov. 2017), 10–18. doi:10.1145/3167902.3167904
- [29] Hernisa Kacorri, Kris M. Kitani, Jeffrey P. Bigham, and Chieko Asakawa. 2017. People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5839–5849. doi:10.1145/3025453.3025899
- [30] Sushant Kafle, Abraham Glasser, Sedeeq Al-khazraji, Larwan Berke, Matthew Seita, and Matt Huenerfauth. 2020. Artificial intelligence fairness in the context of accessibility research on intelligent systems for people who are deaf or hard of hearing. *SIGACCESS Access. Comput.* 125 (March 2020), 4:1. doi:10.1145/3386296.3386300
- [31] Harmanpreet Kaur, Harsha Nori, Samuel Jenkins, Rich Caruana, Hanna Wallach, and Jennifer Wortman Vaughan. 2020. Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376219
- [32] Waqar Hassan Khan, Isaac Buo, and Ariane Middel. 2025. "Tell Me What That Is": Examining Blind and Low Vision Individuals' Reviews on Assistive Smartphone Applications. In *Proceedings of the 27th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Denver Colorado USA, 1–6. doi:10.1145/3663547.3759699
- [33] Korbinian Kuhn, Verena Kersken, and Gottfried Zimmermann. 2025. Evaluating ASR Confidence Scores for Automated Error Detection in User-Assisted Correction Interfaces. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3706599.3720038
- [34] Jaewook Lee, Jaylin Herskovitz, Yi-Hao Peng, and Anhong Guo. 2022. Image-Explorer: Multi-Layered Touch Exploration to Encourage Skepticism Towards Imperfect AI-Generated Image Captions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3491102.3501966
- [35] Scott M. Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 4768–4777.
- [36] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by "Accessibility Research"? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–18. doi:10.1145/3411764.3445412
- [37] Cecily Morrison, Martin Grayson, Rita Faia Marques, Daniela Massiceti, Camilla Longden, Linda Wen, and Edward Cutrell. 2023. Understanding Personalized Accessibility through Teachable AI: Designing and Evaluating Find My Things for People who are Blind or Low Vision. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3597638.3608395
- [38] Chukwunonso Henry Nwokoye, Maria J. P. Peixoto, Akriti Pandey, Lauren Pardy, Mahadeo Sukhai, and Peter R. Lewis. 2024. A Survey of Accessible Explainable Artificial Intelligence Research. doi:10.48550/arXiv.2407.17484 arXiv:2407.17484 version: 1.
- [39] Daniel Omeiza, Skyler Speakman, Celia Cintas, and Komminist Weldermariam. 2019. Smooth Grad-CAM++: An Enhanced Inference Level Visualization Technique for Deep Convolutional Neural Network Models. doi:10.48550/arXiv.1908.01224 arXiv:1908.01224 [cs].

- [40] Maria JP Peixoto, Chukwunonso Henry Nwokoye, Akriti Pandey, Ahsan Zaman, and Peter R. Lewis. 2025. Exploring Accessible Explainable AI: Promising Avenues. *The Journal on Technology and Persons with Disabilities* (2025), 350. <https://scholarworks.calstate.edu/downloads/2f75rj705#page=358>
- [41] Maria J. P. Peixoto, Akriti Pandey, Ahsan Zaman, and Peter R. Lewis. 2025. Who Benefits from AI Explanations? Towards Accessible and Interpretable Systems. doi:10.48550/arXiv.2508.10806 arXiv:2508.10806 [cs].
- [42] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "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)*. Association for Computing Machinery, New York, NY, USA, 1135–1144. doi:10.1145/2939672.2939778
- [43] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*. 618–626. doi:10.1109/ICCV.2017.74 ISSN: 2380-7504.
- [44] Alexa Siu, Gene S-H Kim, Sile O'Modhrain, and Sean Follmer. 2022. Supporting Accessible Data Visualization Through Audio Data Narratives. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3491102.3517678
- [45] Ho Chit Siu, Ashley Suh, Nora Smith, and Isabelle Hurley. 2025. "Explainable" AI Has Some Explaining to Do. *MIT Case Studies in Social and Ethical Responsibilities of Computing Winter 2025* (March 2025). doi:10.21428/2c646de5.e8e32375 Publisher: MIT Schwarzman College of Computing.
- [46] Abigale Stangl, Nitin Verma, Kenneth R. Fleischmann, Meredith Ringel Morris, and Danna Gurari. 2021. Going Beyond One-Size-Fits-All Image Descriptions to Satisfy the Information Wants of People Who are Blind or Have Low Vision. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3441852.3471233
- [47] Ashley Suh, Isabelle Hurley, Nora Smith, and Ho Chit Siu. 2025. Fewer Than 1% of Explainable AI Papers Validate Explainability with Humans. doi:10.48550/arXiv.2503.16507 arXiv:2503.16507 [cs].
- [48] Sukrut Rao, Indu Panigrahi, Sunnie S. Y. Kim, Vikram V. Ramaswamy, Rajat Sahay, Avinab Saha, Dahye Kim, Miguel-Ángel Fernández-Torres, Lenka Tětková, Teresa Dorszewski, Bartłomiej Sobieski, Marina Gavrilova, Yuhui Zhang, and Pushkar Shukla. 2025. The 4th Explainable AI for Computer Vision (XAI4CV) Workshop. https://xai4cv.github.io/workshop_cvpr25
- [49] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology (Harvard JOLT)* 31 (2017), 841. <https://heinonline.org/HOL/Page?handle=hein.journals/hjlt31&id=859&div=&collection=>
- [50] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. 2020. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks. 24–25. https://openaccess.thecvf.com/content_CVPRW_2020/html/w1/Wang_Score-CAM_Score-Weighted_Visual_Explanations_for_Convolutional_Neural_Networks_CVPRW_2020_paper.html
- [51] Linda Yilin Wen, Cecily Morrison, Martin Grayson, Rita Faia Marques, Daniela Massiceti, Camilla Longden, and Edward Cutrell. 2024. Find My Things: Personalized Accessibility through Teachable AI for People who are Blind or Low Vision. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3613905.3648641
- [52] Christine T. Wolf and Kathryn E. Ringland. 2020. Designing accessible, explainable AI (XAI) experiences. *SIGACCESS Access. Comput.* 125 (March 2020), 6:1. doi:10.1145/3386296.3386302
- [53] Xiaoliang Wu, Peter Bell, and Ajitha Rajan. 2024. Can We Trust Explainable AI Methods on ASR? An Evaluation on Phoneme Recognition. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 10296–10300. doi:10.1109/ICASSP48485.2024.10445989 ISSN: 2379-190X.
- [54] Fengyu Yang, Chao Feng, Ziyang Chen, Hyoungseob Park, Daniel Wang, Yiming Dou, Ziyao Zeng, Xien Chen, Rit Gangopadhyay, Andrew Owens, and Alex Wong. 2024. Binding Touch to Everything: Learning Unified Multimodal Tactile Representations. 26340–26353. https://openaccess.thecvf.com/content/CVPR2024/html/Yang_Binding_Touch_to_Everything_Learning_Unified_Multimodal_Tactile_Representations_CVPR_2024_paper.html
- [55] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. 2016. Learning Deep Features for Discriminative Localization. 2921–2929. https://openaccess.thecvf.com/content_cvpr_2016/html/Zhou_Learning_Deep_Features_CVPR_2016_paper.html