

“Zooming In” on Agentic Web Browsers as Assistive Technologies: A Case Study with a Low-Vision Technology Expert

LAURA COLAZZO, Politecnico di Milano, Italy

GIUSEPPE ANZILLOTTI, Politecnico di Milano, Italy

Agentic Web Browsers (AWBs), powered by Large Language Models (LLMs), are emerging as autonomous systems capable of navigating the Web on behalf of users. Beyond enhancing productivity, they offer significant promise as Assistive Technologies (ATs) for visually-impaired individuals, transforming web interaction into a fluid conversational exchange. In this paper, we present a case study with a low-vision technology expert, examining how AWBs can support visually-impaired users in web navigation. The findings show that, despite the current limitations, the navigation experience is notably fluid and satisfying compared to traditional assistive tools, underscoring the strong potential of AWBs to enhance accessibility and reduce barriers in web interaction.

CCS Concepts: • **Human-centered computing** → **Empirical studies in accessibility**; **Accessibility technologies**.

Additional Key Words and Phrases: Agentic Web Browsers, Web Accessibility, Assistive Technologies, Large Language Models, Visual Impairments, Voice Interaction

ACM Reference Format:

Laura Colazzo and Giuseppe Anzillotti. 2018. “Zooming In” on Agentic Web Browsers as Assistive Technologies: A Case Study with a Low-Vision Technology Expert. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 4 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Large Language Models (LLMs) are evolving from passive tools into active agents, capable of perceiving, reasoning, and acting in digital spaces ranging from individual applications to complete operating systems. Agentic Web Browsers (AWBs) exemplify this shift by leveraging LLMs to autonomously navigate and interact with web environments, converting natural language goals into interface actions. Beyond the advantages AWBs can introduce in terms of efficiency and productivity, they also hold promise as Assistive Technologies (ATs) for people with vision impairments, lowering accessibility barriers by enabling natural language interaction instead of relying on visual access and direct interface manipulation. In this respect, they may serve as a compelling alternative to traditional ATs for visually-impaired users, as they can effectively reduce web interaction to a natural and fluid conversational exchange.

In the following, we present a case study involving a low-vision technology expert exploring the use of AWBs as assistive technologies for web navigation. The study aimed to gather preliminary insights into the overall experience, as well as to elicit informed perspectives, grounded in the subject’s technological expertise, on the potential of these systems to reshape web accessibility for visually-impaired users.

Authors’ Contact Information: Laura Colazzo, Politecnico di Milano, Milan, Italy, laura.colazzo@polimi.it; Giuseppe Anzillotti, Politecnico di Milano, Milan, Italy, giuseppe.anzillotti@mail.polimi.it.

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

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

Manuscript submitted to ACM

2 Related Work

Agentic Web Browsers. Compared to Computer-Use Agents (CUAs), such as Anthropic’s Computer Use Tool, which are capable of simulating human interaction (e.g., moving a cursor, clicking buttons, scrolling) with operating systems and a broad range of applications[1, 2, 4], AWBs can be viewed as a specialized subset of these systems operating exclusively within the browser environment. Similarly to CUAs, AWBs such as ChatGPT Atlas¹[6] or Perplexity Comet² are powered by LLMs and are capable of interpreting webpage content leveraging the DOM structure, as well as screenshots taken in real-time during navigation. Taking advantage of this data, they can autonomously plan and execute actions on the Web in response to natural language user instructions. Their graphical interface typically consists of a central web content area and a side chat panel for issuing commands to the system. When control is handed over to the agent, a colored animation appears around the frame of the central content area, while interactions with UI elements, such as clicks or scrolls, are visually displayed in real time. Some AWBs interfaces like Comet also include a dedicated voice mode or a text-to-speech button for vocal interaction.

LLMs and Web accessibility. Some prior studies on LLMs and web accessibility have preserved the existing assistive technology paradigm, leveraging generative AI capabilities to reduce the effort associated with screen-reader-based navigation rather than replacing it [3, 5]. Some recent studies move away from this approach and instead explore CUAs as an alternative to traditional assistive technologies for supporting visually-impaired users. For example, Shin et al. [7] explored the online shopping experience of a group of visually-impaired users using a voice-first LLM-powered CUA. At the same time, Gubbi Mohanbabu et al. [4] highlight the accessibility gap of CUAs, showing that interaction with such systems remains challenging for blind and low-vision users, as they are primarily designed with sighted people in mind, making oversight and control difficult for users with visual impairments.

Building on this body of work, this paper narrows its scope to AWBs, providing a preliminary exploration of their role in web accessibility for visually-impaired users, based on the experience of a low-vision user.

3 Methodology

Participant. The study involved a low-vision male individual aged 31 with congenital low vision. His condition is characterized by complete blindness in the right eye, and residual vision in the left eye (<1/20) accompanied by a visual field reduction below 3%. The subject has used assistive technologies since early childhood, starting with his first use of a personal computer, specifically screen zoom and color inversion, while reporting no use of speech synthesis. Additionally, the participant is a Computer Engineer and currently works as a professional in the technology sector. He reported advanced knowledge of Generative AI and general familiarity with AWBs.

Procedure. The session took place at the end of May 2026 and was conducted by two researchers: one led the interaction with the participant, while the other observed and took notes. The study employed the Voice User Interface (VUI) of Comet, an AWB developed by Perplexity, which the participant had never interacted with before. Inside the Comet interface, the VUI can be accessed through a dedicated button in the top-right corner of the window or via a configurable keyboard shortcut. Upon activation, an auditory cue signals that the system is ready to receive voice input, after which the user can start interacting with the assistant through voice and interrupt its responses at any time. After collecting demographic data, the participant was given time to familiarize himself with the VUI and was then observed performing tasks in two different scenarios. In the first scenario, he was invited to complete an exploratory task on a commercial website consisting of locating and configuring a product, while the second one involved a form-filling task on a public

¹<https://chatgpt.com/atlas/>

²<https://www.perplexity.ai/comet>

administration portal. Following tasks completion, an in-depth semi-structured interview was conducted to gather reflections on the overall experience and elicit perspectives on the potential of this technology as an AT for navigating the Web for visually-impaired individuals.

4 Results

In this section, we present the strengths and limitations that emerged from the case study, focusing on the user experience and the implications these have for the future success of this technology as an assistive tool.

High conversational fluidity and interaction flexibility. The participant expressed strong appreciation for the conversational fluidity and overall flexibility of the interaction with Comet, identifying these as key strengths of the system and as clear evidence of the potential of this interaction paradigm for the future of web accessibility. In particular, he appreciated the system’s ability to adjust the level of detail in its responses according to user contextual needs.

Lack of non-visual feedback. A key issue that emerged during the interaction with Comet is the lack of non-visual feedback while the system processes requests and executes actions autonomously. This became particularly critical when the participant experienced a longer delay while waiting for a system response. Although in this case he understood that the system was still working, thanks to the colored animation perceived on the screen and his knowledge of the technology, this lack of non-visual feedback creates in general transparency issues, leaving users uncertain about what is happening behind the scenes.

Limited control and transparency. Another observed issue is the lack of user control during the system’s autonomous execution. In particular, in the filling-form scenario, Comet autonomously inserted fabricated data into the fields without asking for confirmation. Likewise, in the other scenario, the agent failed to inform the user of an available configuration option and chose one on his behalf. These behaviors reduced the user’s sense of control, compromised system transparency and ultimately undermined trust.

Overall, the participant’s optimism about the crucial role AWBs will play in the future of web accessibility emerged consistently throughout the session. In his view, the limitations encountered can be overcome by conducting further user studies and refining the underlying technology based on user needs, steering the web agent to implement interaction mechanisms that foster user control, enhance transparency and trust, and still provide a fluid interaction experience.

5 Conclusions and Future Work

In this paper, we presented a case study on the use of Comet as an assistive tool for web navigation by a low-vision technology expert. Although findings are based on a single-case study and cannot be generalized to the broader population of visually-impaired users, they offer an initial positive perspective on the potential of AWBs as ATs. This perspective is shaped by the participant’s technical expertise and his ability to recognize the system’s potential beyond its current limitations. Future work should therefore continue this line of research, extending the investigation to a larger and more diverse group of participants, including blind individuals, with particular attention to issues of transparency, user control, and trust. Comparative evaluations with other existing ATs, such as screen readers, could also be conducted to better contextualize the role of AWBs within the accessibility landscape.

References

- [1] Anthropic. 2024. Introducing computer use, a new Claude 3.5 Sonnet, and Claude 3.5 Haiku. <https://www.anthropic.com/news/3-5-models-and-computer-use>
- [2] Ruijia Cheng, Jenny T Liang, Eldon Schoop, and Jeffrey Nichols. 2026. Mapping the Design Space of User Experience for Computer Use Agents. In *Proceedings of the 31st International Conference on Intelligent User Interfaces (IUI '26)*. Association for Computing Machinery, New York, NY, USA,

- 646–662. doi:10.1145/3742413.3789132
- [3] Anujay Ghosh, Monalika Padma Reddy, Satwik Ram Kodandaram, Utku Uçkun, Vikas Ashok, Xiaojun Bi, and I. V. Ramakrishnan. 2024. Enabling Uniform Computer Interaction Experience for Blind Users through Large Language Models. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24)*. Association for Computing Machinery. doi:10.1145/3663548.3675605
- [4] Ananya Gubbi Mohanbabu, Rosiana Natalie, Brandon Kim, Anhong Guo, and Amy Pavel. 2026. A11y-CUA Dataset: Characterizing the Accessibility Gap in Computer Use Agents. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for Computing Machinery. doi:10.1145/3772318.3791896
- [5] Ananya Gubbi Mohanbabu, Yotam Sechayk, and Amy Pavel. 2025. Task Mode: Dynamic Filtering for Task-Specific Web Navigation using LLMs. In *Proceedings of the 27th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '25)*. Association for Computing Machinery. doi:10.1145/3663547.3746401
- [6] OpenAI. 2025. Introducing ChatGPT Atlas. <https://chatgpt.com/atlas/>
- [7] Subin Shin, Jeesun Oh, Suhyun Kim, Seoyeon Eom, and Sangwon Lee. 2026. Toward Independent Online Shopping of the Visually Impaired Through Voice-based Computer-Using Agent. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3772318.3791681

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009