International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Driving Digital Inclusion: Voice-First AI Systems & Accessibility in Mobile Infrastructure

Rohit Jarubula

Submitted: 01/07/2023 **Revised:** 10/08/2023 **Accepted:** 20/08/2023

Abstract: The paper examined how voice-first systems based on artificial intelligence (AI) can be used to promote digital inclusion and access in mobile settings. It explores how AI technologies, 5G, and edge computing could be used to develop inclusive experiences by the people with disabilities and older adults. The paper utilizes qualitative analysis of the recent studies to identify primary themes in the design of accessibility, adoption by the user, and support of the infrastructure and ethical governance. These findings reveal that accessible voice systems do not only enhance the usability of the system but also induce social equity and responsible innovation, and assist in evolving mobile networks to more inclusive and human-centric digital ecosystems.

***Keywords*:** *Mobile Infrastructure, Digital Inclusion, AI, Accessibility*

I. INTRODUCTION

Even though the industry in mobile technology has been very rapid in its development, there are still a number of people who are still barriers to the utilization of digital services. The aged and those who have disability does not easily deal with complex screens, tiny text and manual inputs. The voice-first AI represents a form of new interface to technology through natural speech and dialogue. These systems together with 5G and edge computing can provide reliable, quicker, and all-inclusive mobile experiences. Sixty01, a voice-based AI research, will be analyzed within the framework of voice-based AI capabilities to lead digital inclusion and boost the accessibility of diverse digital services and equitable involvement in the digital economy, through accountable and ethically attuned technology development.

II. RELATED WORKS

Role of AI in Digital Accessibility

Digital accessibility is a reaction to the development of digital systems to make them accessible to all through the accessibility of people with disabilities into the system. According to recent studies, artificial intelligence (AI) has emerged as one of the potent forces in enhancing the digital accessibility and inclusion.

An analysed literature of scholarly papers published in the past three years showed that AI applications have been used to support individuals with visual impairments, although there is a lack of emphasis on individuals with speech, hearing, or motor disabilities [1]. Such an imbalance creates a void in the inclusive research that takes into account the needs of such different disabilities. Another finding of the same study is the occurrence of several present-day AI systems not meeting the established standards of accessibility, and it should be reinforced in terms of compliance and design of the assistive technology should be treated more holistically.

The human capability is a major facilitator of equitable involvement in digital ecosystems driven by AI since it interprets and responds to human signals. The machine learning and natural language processing (NLP) techniques are currently being integrated into mobile and web applications and improving accessibility to users with disabilities [4]. The AI-based accessibility services are being used in the realms of computer vision and audio processing, and NLP in mobile platforms like Android.

The capabilities offered by these systems include real time translation, speech to text transcription and customized user interface that facilitates multilingual

Agile Scrum Master/Project Manager

and contextual interaction. Nevertheless, there are still difficulties with keeping the models correct and ethical use of data. Research has demanded the use of universal design architectures and global accessibility requirements in helping the enforcement of AI based solutions [4].

The interaction process between individuals with disabilities and educational, cultural, and social spheres is also changing with the help of AI. As an example, studies performed in science museums reveal that such AI technologies as audio guides (intelligent), image recognition, and tactile input devices enhance inclusivity towards individuals with visual, auditory, and mobility deficiencies [3].

The interactivity of such systems makes their user experiences more personal, more interactive, and fair. The associated risks associated with these developments are the issue of data privacy and inequities in the deployment and thus should be followed by responsible innovation on accessibility technologies [3].

Human-Centered Interaction

Voice assistants and smart speakers are voice-first AI systems that transformed the way human beings engage with technology. Such systems are based on speech recognition, NLP, and machine learning which interpret and answer commands that are given by a human voice which is understood to be hands-free and eyes-free. They have enormous potentials in the improvement of access by people with blind eyes or handicap problems and older adults with difficulty using traditional interfaces [6].

Voice assistants provided as AI, such as Siri, Alexa, and Google Assistant, have been the primary devices of inclusive digital interaction and are available with personalized, intuitive user experiences. Things are still being raised however about voice recognition bias, unequal support of languages and the ethical application of biometric information [6].

Studies to determine whether voice assistants are acceptable among the elderly have had differing findings. Although such devices can support the quality of life, the perceptions of usefulness, enjoyment, and technological self-efficacy tend to shape its level of acceptance [7]. The elderly who are more confident about their ability to use technology, and who are receptive to change, have high chances of benefiting by these systems.

Research has found that the elderly users might not be able to formulate structured voice expressions or interpret device responses and thus have difficulties with usability [2][7]. In this regard, it is possible to make multimodal designs that incorporate both voice and visual interfaces such as the ones in the devices of Echo Show flexible and comfortable to users. Interestingly, despite having pictorial choices, majority of the older participants favored the speech-based interaction to touch input which implies that speech is the most instinctual interaction mode to this group of people [2].

Besides helping to serve seniors, voice-first AI systems can enable users with visual or physical impairments because it reduces mental and physical efforts required to interact with digital space. The speech recognition systems with the usage of AI allow a smooth interaction with the education platforms, as well as any other services, where people with visual impairments receive real-time help [10].

Even with the few drawbacks in recognition accuracy relative to commercial systems, the latest systems have shown high potential of providing affordable and readily customizable solutions of accessibility [10]. Since voice-first systems are evolving with mobile infrastructure, one of the core objectives of providing digital inclusion is by making sure that speech recognition models are accommodating of different speech patterns, accents, and disabilities [9].

Infrastructure Enablers

Integrity of voice-first AI - Voice first applications rely on strong mobile infrastructure to provide real time, dependable and inclusive user experience. Accessibility features are being deployed on a large scale with the integration of the 5G and edge computing technologies, changing how these features will be implemented.

Edge computing enables the computation and data processing to be nearer to the user, which lowers the latency, and enhances the responsiveness, one of the key conditions of such tasks as live captioning, voice translation, and conversational AI [8]. Using edge computing in 5G-enable networks, data transmission will be faster and a localized process will be done this helps in the accessibility applications where a continuous voice or video input is required.

With the growing aspirations of digital inclusion, speed, cost, and accessibility are the three main factors that infrastructure providers are grappling with. With the integration of the cloud-based AI services and the

distributed edge resources, it will be possible to use assistive systems in a variety of settings, including the rural communities and high-dense urban centers with the ability to scale [8].

This potential is likely to be extended by the shift to 6G networks to make the network intelligence embedded to enable the delivery of services in an adaptive and context-aware manner. This development would allow more efficient and fairer accessibility options, which would cut digital divide along socioeconomic and geographic lines.

The privacy preserving models are also supported by edge computing, which is a major worry to voice-based systems where sensitive biometrical data are required. Edge AI has the capability to minimize the potential risk of exposing significant amounts of user data to the outside world as it processes user data at the edge sources without the need to reduce the degree of personalization and awareness about the surrounding context.

There is a need to have strong governance systems that will guarantee openness and confidence, especially in handling of biometric identifiers and speech data. Research in both the AI field and the telecommunications sector has reinforced the necessity to focus the accessibility initiative on accountable information administration and lawful actions [3][4][8].

Dimensions of Inclusive AI

As much as accessibility requires technological advancements, the same should be backed with ethical and policy platforms that ensure the safety of the consumers and also encourage fair play. Voice and AI-powered accessibility solutions amass a vast quantity of personal data, voice samples, behavioral patterns, and contextual cues, which tends to put its users at risk of privacy or algorithmic bias.

Researchers have indicated that most accessibility systems do not adhere completely to the current standards and ethical principles and hence, the disability groups are not fairly benefited [1][5]. This means that the element of inclusivity should be incorporated in the technical design as well as in the aspect of institutional governance and social responsibility.

There are potentials as well as drawbacks in AI technologies applied in educational and social inclusion initiatives. Advanced autonomy and engagement in students with disabilities has proven as a positive impact because of AI-powered screen readers and NLP learning tools [5]. Nevertheless, the lack of teacher preparedness, inadequate infrastructure and algorithm bias keep it at bay.

To fill these gaps, researchers propose the use of a mix of specific policy changes, teacher education, and open AI governance systems [5]. These strategies can make the accessibility technologies to be the agents of social equalization and innovation.

It is also significant that we create fair and available speech AI. The existing systems do not tend to identify individuals with speech diversities and so, they fail to prevent the exclusion of these individuals in the automated customer service, recruitment, and communication tools [9]. To overcome this, interdisciplinary studies in the fields of AI, human-computer interaction, speech pathology, and disability studies are needed.

The development of technical and policy principles to define inclusive speech AI is also being implemented, and involves fair training of models and participatory design involving communities who are impacted [9]. These attempts will play an important role in making the future of voice-first AI not efficient but also just and representative.

Being a digital citizen should be acknowledged as a social and a technological goal. With accessibility being incorporated at the design and implementation phase of AI and mobile infrastructure, organizational performance, user satisfaction and social equity may be realized and measured. Accessibility then becomes a performance difference maker instead of a compliance.

III. METHODOLOGY

This paper adheres to a qualitative research methodology to examine the ways that voice-first AI applications with the help of 5G and edge computing may enhance digital inclusion and accessibility. The reason why the qualitative approach was adopted is that it could be approached more profoundly in terms of user experiences, design practices, and associated social effects of accessibility. This work is not founded on the numerical data but on the patterns, meanings, and insights basing on the real-life examples, the academic literature, and the views of the experts.

Research Design

The study will be an exploratory qualitative study that will employ the thematic synthesis and secondary data

analysis. As the issue involves a combination of the technology, accessibility, and policy aspect, one research approach, say, an experiment, would not suffice.

As such, the findings of various scholarly studies, policy reports, or case analyses, published in 2018-24 are incorporated in this paper. The research papers that are covered in the review are those in the realms of artificial intelligence, mobile computing, human-computer interaction, accessibility design, and digital inclusion.

The purpose is primarily to find out the role of voice-first AI and associated infrastructure technologies in digital accessibility and what are the challenges and governance concerns surrounding its implementation. The study also seeks to emphasize some of the guidelines in design that would enhance inclusivity of voice-enabled systems with the elderly and people with disabilities.

Data Sources

The paper relies on ten peer-reviewed documents that present comprehensive evidence on the AI ranges of application in the areas of accessibility, voice assistant user experience, and infrastructure support, including 5G and edge computing. These sources have been chosen on the grounds of their relevance, credibility, and variety of opinion. The chosen articles are systematic reviews, empirical experiments, case studies, and conceptual articles on AI-based accessibility solutions.

The inclusion criteria were based on papers that:

- Attend to the issue of digital accessibility or inclusion with the help of AI technologies.
- Investigate the use of voice first/conversational interfaces.
- Jurisdiction or infrastructure, e.g. privacy or equity.
- have been published on official scientific journals or conference proceedings in 2018-24.

The analyzed references discussed a range of users, such as older adults and people with visual or hearing disabilities as well as speech/motor disabilities. These sources were such that they presented a rich interpretation in terms of the thematic aspects.

Data Analysis

Thematic analysis approach was used to plan and analyze the qualitative data. All the papers were thoroughly read and frequent concepts, ideas, and conclusions were coded as under leading themes. Categories that were obtained using the initial coding process included AI-based accessibility innovations, voice-first usability, ethical and privacy issues, and infrastructure enablers. These codes were then summarized into general themes that exhibited the objectives of the study.

In a bid to be sure, the results of various studies were juxtaposed to find the trends, similarities and discrepancies. In the case, there was research that has shown the advantage of multimodal voice interface to older adults and some studies have come out to point out the difficulty in the rate of accurate speech recognition and safety of the data. The study developed a holistic appreciation of opportunities as well as the hindrances of digital inclusion through the process.

Ethical Considerations

Being a qualitative synthesis, this paper will be based on existing academic and technical publications instead of the interaction with the user. As such, it does not require human subjects and it does not need to be approved by some ethical foundation. Nevertheless, the ethical awareness was preserved in the whole process by ensuring the crediting of everything that was of original nature and addressing the matters that included privacy of data, bias in the algorithms and fair access.

The paper acknowledges weaknesses such as the possibility of bias due to publication by other researchers and lack of primary data gathering. In future research they may involve interviews or participatory workshops with the users and developers to confirm these findings and broaden them.

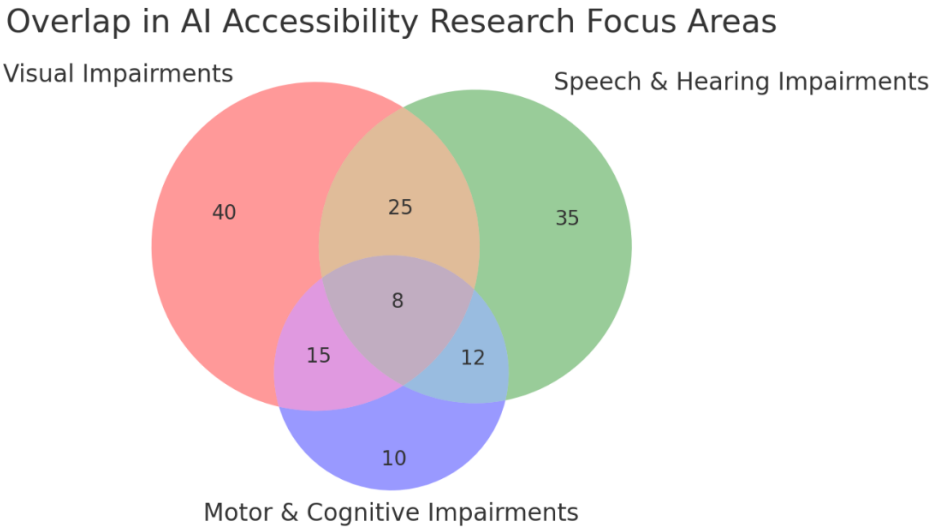
IV. RESULTS

AI-Driven Accessibility

The discussion has indicated that the technologies of artificial intelligence (AI) have become the priority of enhancing the accessibility of mobile and digital systems. In the studies reviewed, the scholars all highlighted how machine learning, natural language processing (NLP) and computer vision allow users with disabilities to interact with digital service in a more autonomous way [1][4][6].

The systems that have been developed due to these technologies have the capability of identifying speech, understanding gestures, and providing real-time feedback. Indicatively, AI-enabled screen readers and transcription have become an immensely popular

solution in education, healthcare and in subsistence [5][10].



Individuals with visual issues are also assisted by the AI systems, speaking a language and translating it into the text or the opposite, or with image identification and description of the surroundings [3][5]. Most of the literature observed that most AI accessibility projects still prioritize visual impairment although there is less concentration on speech, hearing, and motor disability [1]. Such imbalance indicates that the field of accessibility is yet to be made more accommodative to all categories of users.

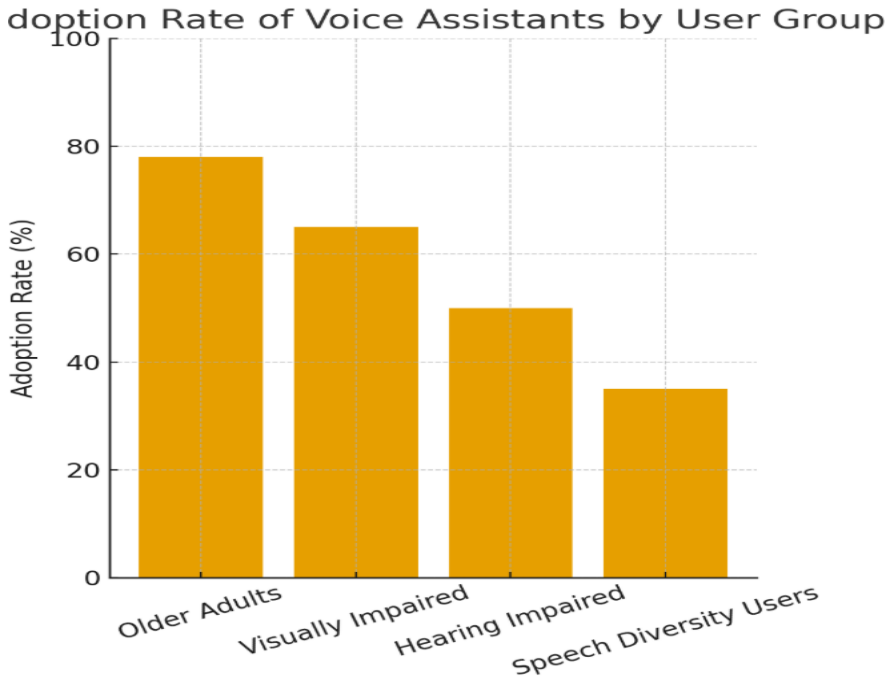
Table 1. AI-driven accessibility findings

Focus Area	**Description**	**Example Findings**
Visual and Hearing Accessibility	AI technologies are based on speech recognition and image analysis to enhance content accessibility.	AI is applied in museums and educational platforms to give visitors with disabilities audio tour guides and text descriptions [3][5].
Underrepresented Disabilities	Current systems have little attention to the speech, hearing, and motor impairments.	Research presents a research gap by excluding various needs of users [1].
Real-time Assistance	Adaptive and responsive tools facilitate communication, which are facilitated by AI.	The AI translation and transcription applications are becoming an extension of mobile devices [4][6].

It has been discovered that the AI systems only work well in situations where accessibility features are incorporated very early in the design process. Late accessibility retrofitting will tend to be expensive and tend to have poor usability results. This trend allows sustaining the argument that inclusive design is a societal need and an asset of performance of technology firms.

Multimodal Systems

Voice-first types of AI, e.g. Alexa, Siri and Google assistant, are now used as a routine communication and access to information tools. These systems allow hands free and eyes free communication and this is specifically helpful to those individuals with mobility or visual difficulties.



In the literature, voice assistants are typified as being helpful to older adults, especially regarding the use of voice assistants to remind them about simple things like weather, or to communicate with them [2][7]. Usability issues have not been eliminated because of inappropriate feedback, speech recognition, and inadequate contextual cognition.

A study of multimodal voice assistants, i.e. those that have both voice and touchscreen interfaces, established that, although graphical feedback can help you comprehend what the assistant is doing, most users reply verbally instead of using a screen touch [2]. The inclination toward the speech interaction indicates that voice-first-based systems are to be more about clarity, feedback of responses, and adjusting language instead of elaborated graphical features.

Table 2. User experience themes

User Group	**Experience Summary**	**Design Implications**
Older Adults	The person likes hands-free communication but requires less complicated instructions and feedback.	Slow speech, repetition and confirmation prompts should be supported by the systems [2][7].
Visually Impaired Users	Take advantage of text to speech feedback and audio feedback.	Presentation interfaces have to employ standard tone and are not to be screen-based [5][10].
Users with Speech Diversities	Best on frequent occurrence of face recognition error and system interruptions.	Voice AI should also be equipped with a variety of datasets along with wide tolerance of input [9].

Results also indicate that the voice assistant acceptance by users is reliant on perceptions of usefulness, enjoyment, and attitude to technology [7]. The elders who are happy to experiment with devices are more satisfied whereas those with low technological self-efficacy would not bother. Thus, major adoption can be achieved by community training, easy-to-follow tutorials, simplified onboarding processes.

The other significant message is that voice AI is able to provide the link between accessibility and digital literacy. As an illustration, a conversational system may take people through the steps of digital tasks such as appointing or paying to eliminate fear or confusion. These results support the conclusion that voice-first interfaces are not assistive systems only but also enabling digital inclusion.

Infrastructure Support

One of the outstanding findings of the literature review is the highly positive relationship between access and the availability of mobile infrastructure. Voice-first systems are based on high-speed, efficient and expediency network to operate successfully.

Some of the technologies like 5G and edge computing are stated to be among the vital enablers through which real-time processing of speech, live transcription, and translation can be done without delays [8]. Edge computing enables the user to process data much nearer which enhances efficiency and diminishes chances of privacy breach.

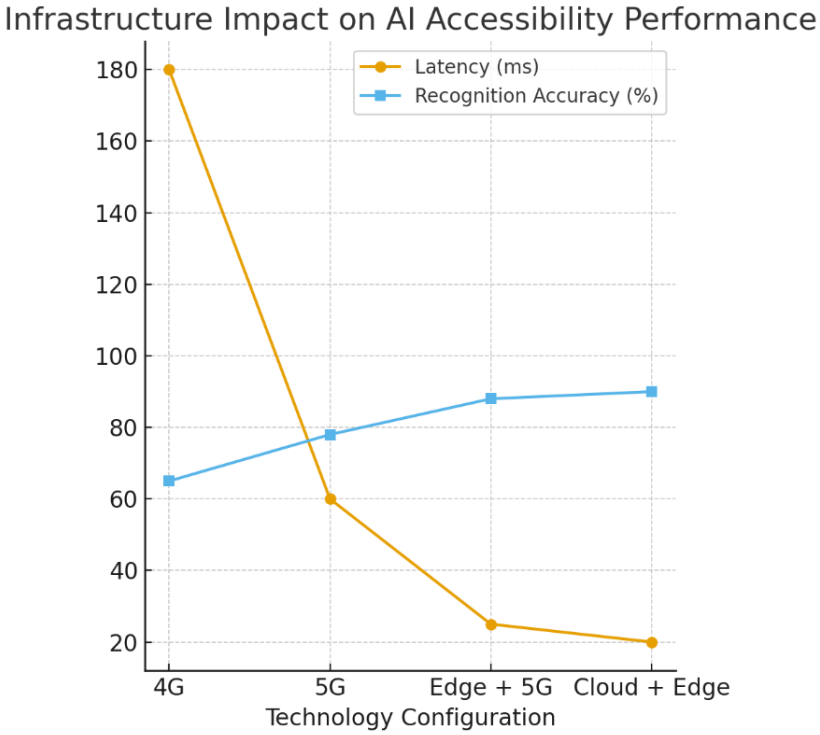


Table 3. Infrastructure and accessibility integration

Technology	Role in Accessibility	Observed Benefits
5G Networks	Supports low-latency, high-speed communication of the AI voice systems.	The translation, captioning, and interaction are made easier in real-time [8].
Edge Computing	Calculates calculation close to the user to save time and safeguard information.	Enhances the privacy of users, offline, and less reliance on central servers [8].
Cloud Services	Archive and transport AI models to become Global accessible.	Allows updating in large scales, but poses an ethical and data governance risk [3][4].

These results indicate that inclusion concerns more than by means of the user interface design but also network structure and resource allocations. With the development of mobile infrastructure, speed, cost and accessibility will set the scale on which the under tuned areas will achieve equitable application of such technologies. Besides, 6G innovations will further embed additional intelligence into the networks, and the accessibility functions will become self-operational and versatile to adjusting to changes [8].

Another threat of digital inequality in infrastructure is also found in the research. Accessibility tools that rely on cloud resources or edge resources can be affected negatively in areas that lack coverage of 5G. Such a gap underscores the necessity of collaborations between the government, telecom providers, and accessibility advocates in such a way that voice-first systems are deployed in an even manner.

Policy Implications

The last theme is centered on ethical, privacy and policy themes. In all of the reviewed articles, it was noted that researchers in the business sector prioritize proper AI governance to promote fairness, transparency, and user trust [1][3][5][9]. Voice-first systems are typically known to gather highly sensitive biometric information about speech patterns, accent gauge and tone of voice. Such data may be abused or abused incorrectly, even causing a bias in recognition algorithms without proper controls.

The users with disabilities and speech differences particularly are put in a very serious danger of ethical risks because they might be already excluded in digital spaces. Fair and accessible speech AI studies note that existing systems are developed primarily on a sample of so-called typical speech data, which is why people with stutters, speech-delays or accents are misrecognised [9]. These problems not only make usability lower but may also lead to emotional pain as well as strengthen social inequality.

Table 4. Ethical and policy concerns

Issue	Description	Proposed Actions
Data Privacy	Voice assistants gather big quantities of personal and biometric information.	Emerge overt regularity of explanation, local data processing through edge computing [3][8].
Algorithmic Bias	The models represent not users with speech or language disabilities.	Provide a variety of data and engage with affected users through experimental work [9].
Policy and Standards	Inadequate accessibility and privacy policies in most areas.	Bring AI policies and international standards of access on board [1][5].

The results indicate the need to implement ethical design across the AI development life cycle - data gathering to model training and testing with users. The policy frameworks should help companies to share the responsibility of accessibility and not consider it as an additional feature. There is also the fact that the direct participation of the disabled in co-design processes can result in equitable and more efficient systems.

A larger perception of digital inclusion as a social performance indicator is also indicated in the reviewed literature. By focusing on accessibility, organizations not only such as the organizations meet the legal requirements, but also enhance their public popularity and reach the markets. The factor of competitive advantage and an image of responsible innovation, consequently, becomes accessible.

Summary of Findings

The paper revealed that voice-first AI applications have the potential to redefine mobile accessibility provided that they are created using an inclusive design, comprehensive infrastructure, and established governance. The potential of AI technologies to help increase the barrier-free environment in people with disabilities and older adults, particularly conversational interfaces and multimodal ones, is high. Nevertheless, the threats of ethical concerns, technical challenges, as well as uneven savings of infrastructure are also major problems.

Through the integration of AI innovation with 5G and edge computing, organizations are able to create mobile ecosystems that are quicker, more equitable and human-oriented. The issues and findings of the research should also be further investigated concerning the practical application in various culture and language contexts so that the advantages of digital inclusion can be extended to all people.

V. CONCLUSION

The paper points to the conclusion that voice-only AI can substantially enhance the level of digital accessibility in case the system is backed by a powerful mobile infrastructure and ethically designed. Combined with inclusive interfaces, high-speed 5G connectivity, as well as edge computing, more equal digital experiences are built.

Fairness, privacy, and data security should also become the core of the further development. The accessibility should not be considered as the feature but the principle on which the innovations and performance can be based. With accessibility as its main feature, organizations may build a higher level of user trust,

enhance social inclusion, and contribute to a digital ecosystem that may equally benefit all individuals.

REFERENCES

[1] de Freitas, M. P., Piai, V. A., Farias, R. H., Fernandes, A. M. R., de Moraes Rossetto, A. G., & Leithardt, V. R. Q. (2022). Artificial Intelligence of Things Applied to Assistive Technology: A Systematic Literature Review. *Sensors*, *22*(21), 8531. https://doi.org/10.3390/s22218531

[2] Chen, C., Lifset, E. T., Han, Y., Roy, A., Hogarth, M., Moore, A. A., Farcas, E., & Weibel, N. (2023). Screen or No Screen? Lessons Learnt from a Real-World Deployment Study of Using Voice Assistants With and Without Touchscreen for Older Adults. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2307.07723

[3] Eardley, A. F., Thompson, H., Fineman, A., Hutchinson, R., Bywood, L., & Cock, M. (2022). Devisualizing the Museum: From access to inclusion. *Journal of Museum Education*, *47*(2), 150–165. https://doi.org/10.1080/10598650.2022.2077067

[4] AlOmar, E. A., Aljedaani, W., Tamjeed, M., Mkaouer, M. W., & Elglaly, Y. N. (2022). Finding the needle in a haystack: on the Automatic identification of Accessibility user reviews. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2210.09947

[5] Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2022). Assistive technology for the inclusion of students with disabilities: a systematic review. *Educational Technology Research and Development*, *70*(5), 1911–1930. https://doi.org/10.1007/s11423-022-10127-7

[6] Kim, S. (2020). Exploring how older adults use a Smart Speaker–Based Voice Assistant in their first interactions: Qualitative study. *JMIR Mhealth and Uhealth*, *9*(1), e20427. https://doi.org/10.2196/20427

[7] Chung, J., Bleich, M., Wheeler, D. C., Winship, J. M., McDowell, B., Baker, D., & Parsons, P. (2021). Attitudes and Perceptions toward Voice-Operated Smart Speakers among Low-Income Senior Housing Residents: Comparison of Pre- and Post-Installation Surveys. *Gerontology and Geriatric Medicine*, *7*, 23337214211005869. https://doi.org/10.1177/23337214211005869

[8] Filali, A., Abouaomar, A., Cherkaoui, S., Kobbane, A., & Guizani, M. (2020). Multi-Access Edge Computing: a survey. *IEEE Access*, *8*, 197017–197046. https://doi.org/10.1109/access.2020.3034136

[9] Geng, M., Xie, X., Ye, Z., Wang, T., Li, G., Hu, S., Liu, X., & Meng, H. (2022). Speaker adaptation using Spectro-Temporal deep features for dysarthric and elderly speech recognition. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2202.10290

[10] Abhishek, S., Sathish, H., K, A. K., & T, A. (2022). Aiding the Visually Impaired using Artificial Intelligence and Speech Recognition Technology. *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, 1356–1362. https://doi.org/10.1109/icirca54612.2022.9985659