CONTEXT-AWARE AI STENOGRAPHY - ENHANCING MEDIA ACCESSIBILITY FOR THE DEAF AND HARD OF HEARING

BREAKING THE BLACK BOX - INCLUSIVITY AND MINORITY CONTRIBUTIONS IN FREE AND OPEN SOURCE SOFTWARE

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Tony Chang

November 8, 2024

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Kent Wayland, Department of Engineering and Society

Mark Floryan, Department of Computer Science

Introduction

How should media software be designed for all, not most, to make for an equitable, more inclusive society?

The overarching question this paper tackles is how technology, particularly media-presenting and software development tools, can be designed to be more equitable and inclusive, addressing the diverse needs of all social groups rather than catering predominantly to the majority. Inclusivity in technology design is essential not only for empowering underrepresented communities but also for ensuring that technological advancements benefit society as a whole. One area of concern is the persistent underrepresentation of certain groups in software development. For example, in 2022, only 21% of self-identified software developers in the U.S. were female, a proportion that has been shrinking in recent years. While racial diversity among software developers shows a gradual increase—with 44% of developers identifying as BIPOC—there is still significant room for improvement, particularly in fostering equitable participation and addressing the barriers faced by these groups (Diversity in Software Development - BairesDev, 2022). This paper explores two dimensions of inclusivity in technology. The technical portion proposes an AI-driven model to enhance media accessibility for individuals with hearing impairments. This model generates context-aware captions tailored to users' specific needs, aiming to make digital media more inclusive. The STS analysis examines the participation of underrepresented groups—such as women, people with disabilities, and individuals from diverse cultural backgrounds—in the Free and Open Source Software (FOSS) community. It delves into how these groups contribute to and shape the development of software and how their unique needs influence design priorities. By combining these technical and social perspectives, the project advances a broader vision of inclusivity in technology. It demonstrates how integrating diverse viewpoints can create digital tools and environments that serve a wider range of users, fostering equity and innovation in technology design.

Technical Project: Context-Aware AI Stenography - Enhancing Media Accessibility for the Deaf and Hard of Hearing

This project proposes an advanced AI model to assist deaf and hard-of-hearing individuals in consuming audio-based media, including internet videos, recorded lectures, and live meetings, by creating a readable and visual text stream for the viewer, offering a more refined version of closed captions. The need for closed captions has a long history associated with television, as the social and technological system for producing and distributing captioned television to viewers in America had its roots in the history of subtitled film while also departing from that history in significant and necessary ways (G.J. Downey, 2008), where there were fears of when making television accessible to the deaf that it would make it unpalatable to the hearing. Many people within the deaf community desired media access, from educators to people

watching television in their living rooms. Current solutions like closed captioning and human stenography have notable limitations: closed captions, produced by simple speech-to-text models, often misinterpret audio, resulting in inaccuracies and incoherent sentences due to a lack of contextual and grammatical awareness, while human stenographers, though more accurate, are limited by typing speed, prone to error, and potentially introduce human biases. Furthermore, while stenographers type, on average more than 200 words per minute (What Is a Stenographer?, n.d.), their reported words lag behind when they are streamed to the viewer.

This AI model leverages natural language processing (NLP) and machine learning (ML) to provide an accurate, coherent, and contextually aware captioning solution that could surpass these limitations. Unlike traditional closed captioning, which outputs words based on direct transcription, the AI will analyze and construct sentences that align with natural language grammar rules and maintain the intended meaning of the speaker by referencing contextual cues from surrounding dialogue. This would result in more readable and accurate captions that enhance the user's understanding of complex audio content.

The model will be trained on diverse datasets that include variations in accent, pronunciation, and speech patterns, enabling it to adapt to different speakers' unique ways of speaking, whether they have an accent, lisp, or other distinctive pronunciations. I will gather speech samples of these different speakers through recordings of the professors at UVA. In multi-speaker environments, such as meetings or panel discussions, the model will also be capable of differentiating between speakers, allowing for speaker-specific captioning. For this, I will gather a different set of speech samples where multiple people are speaking simultaneously. This ensures that each speaker’s unique manner of speaking is accurately represented and that the captions maintain clarity and relevance, particularly in environments with overlapping dialogue. Furthermore, the AI will operate with minimal latency in real-time, ensuring that captions appear instantaneously, making it suitable for live settings like meetings or lectures. With these requirements and design features in mind, the software will deliver an inclusive and high-quality media experience, significantly improving accessibility to audio-based content.

STS Project: Breaking the Black Box - Inclusivity and Minority Contributions in Free and Open Source Software

The STS portion of this project addresses Free and Open Source Software (FOSS) and the role that less-represented minority groups, such as individuals from diverse cultural backgrounds and people with disabilities, play within it. FOSS, unlike proprietary software, is built on principles of transparency and collaboration, allowing anyone to access, modify, and distribute its source code. This model is increasingly popular, with developers often prioritizing open-source frameworks in their projects.

Background

Traditionally, software is distributed as executable files rather than source code, concealing its inner workings in a "black box" approach. By contrast, Free and Open Source Software (FOSS) seeks to democratize access, breaking open this black box to empower users to understand and modify the software to suit their needs. Actants in this ecosystem include developers, users, platforms, and marginalized groups who rely on or contribute to open-source tools. The values central to FOSS—transparency, collaboration, and accessibility—underscore its potential to serve as a platform for inclusivity and empowerment. However, these ideals are not always realized in practice, as structural and social barriers limit participation from minority groups. Key barriers include accessibility challenges for individuals with disabilities, underrepresentation of diverse cultural perspectives, and the difficulties of navigating development communities that remain predominantly homogeneous. Historical and situational factors, such as the dominance of English as the primary language for documentation and communication, exacerbate these challenges, making it harder for non-native speakers to engage fully. Similarly, the voluntary nature of FOSS contributions often excludes those who lack the financial or time resources to participate, perpetuating inequities. These dynamics raise critical questions: To what extent do the FOSS community's goals of transparency and inclusivity align with its actual practices? What software innovations have emerged to make open-source tools more accessible to minority groups? And how have these groups contributed to shaping FOSS projects to better meet their needs? Historically, open-source software has aimed to bridge gaps in digital equity, yet unintended consequences—such as designs that exclude marginalized users—highlight the importance of revisiting these goals (Damian, n.d.).

This project will explore the interactions between developers, FOSS platforms, and marginalized communities as key elements of this sociotechnical system. It will investigate the types of FOSS that prioritize accessibility for underrepresented groups and examine the contributions of minority developers and users. Through analyzing initiatives aimed at improving inclusivity, studying the norms and beliefs guiding FOSS development, and understanding the sociotechnical dynamics at play, this research will provide insights into how inclusive FOSS truly is. By shedding light on the interplay between technology, inclusivity, and empowerment, the study aims to illuminate paths toward a more equitable open-source ecosystem.

Literature Exploration

The groups of interest include software developers, software consumers, framework developers, and the minority communities within each of these groups. By contributing to FOSS, developers and designers from underrepresented groups bring unique perspectives that enhance inclusivity and accessibility. For example, the GNOME platform has benefited significantly from contributions by disabled developers. GNOME 2.4, noted for its accessibility features and polished interface, represents a collaborative effort among commercial vendors like Sun, Red Hat, and Ximian, as well as the GNOME community (Castro, 2003). Such contributions not only

foster inclusivity but also embody the FOSS community's commitment to diversity, equity, and empowering individuals from all backgrounds to shape technology for the collective good.

The scale and impact of FOSS projects is further illustrated by the open-source repositories maintained by the U.S. Department of Energy's national laboratories. These labs operate independently but contribute to a shared pool of over 8,000 projects, reflecting the broad scope of research software in the FOSS ecosystem. Community engagement is evident in the popularity of 2,005 repositories, of which more than half were updated within the last 12 weeks. However, a significant portion of these repositories, particularly the 260 identified as requiring sustainability support, underscores the ongoing challenges of maintaining and developing open-source software.

FOSS also serves as a platform for free expression, embodying the principles of free speech. Software is not merely a functional tool but a creative medium for sharing, modifying, and innovating ideas (E Gabriella Coleman, 2013). This perspective reinforces the need to protect FOSS as a space where diverse voices, including those from minority groups, can contribute without the limitations imposed by proprietary software models. Furthermore, by adhering to guidelines like the FAIR-USE4OS, which build upon the FAIR principles (Findable, Accessible, Interoperable, and Reusable), open-source projects can prioritize user-centered design, sustainability, and equitable access, increasing their long-term utility and impact (Schwartz et al., 2024).

However, ensuring that FOSS remains inclusive and impactful requires navigating complex decision-making processes within projects. Models such as the Stochastic Actor-Oriented Model and the Garbage Can Model highlight the multifaceted nature of these systems but fall short of fully capturing their intricacies (Lomi, Conaldi, & Tonellato, 2012). Decision-making often hinges on factors like contributor communication, bug severity, and contributor experience, with productivity among seasoned developers playing a critical role in problem-solving effectiveness. Efforts like the FLOSSMetrics project, funded by the European Commission, address these challenges by collecting and analyzing data from over 3,000 FOSS projects. This database supports diverse analyses, such as the study of cooperation and openness in VoIP client projects, illustrating how robust data infrastructures can enhance our understanding of FOSS dynamics.

Methods & Evidence

To investigate the inclusivity of the FOSS community and the contributions made by minority groups, I will gather qualitative and quantitative data from multiple sources. First, I will delve into case studies of inclusive FOSS projects, analyzing documentation, code contributions, and community discussions to see how accessibility needs are addressed. Contribution data from platforms like GitHub will reveal participation trends, helping identify gaps in minority involvement. Additionally, I will survey FOSS community members to quantify attitudes toward inclusivity, comparing FOSS ideals with practical experiences. Finally, I will review policies and guidelines from major FOSS projects, such as Linux and Mozilla, to assess their formal stance

on diversity and inclusion. Together, these sources will provide a comprehensive view of the role of minority contributors, the accessibility of FOSS projects, and how well the community's principles align with its practices on inclusivity.

Conclusion

This paper seeks to explore how technology, particularly media software, can be designed to be more inclusive and accessible, addressing the needs of all social groups, not just the majority. Through the technical project, I aim to develop an AI-driven model that improves media accessibility for the deaf and hard of hearing, offering more accurate and context-aware captions for diverse users. The STS component will examine the role of underrepresented groups in the FOSS community, exploring how they contribute to the development of more inclusive technologies and influence the design of accessible software. Together, these projects will highlight the intersection of technology and inclusivity, shedding light on how diverse perspectives can enhance the development and accessibility of digital tools. By focusing on both the technical and social aspects, this research aims to promote a more equitable digital society, where technology is tailored to meet the needs of all users. Future work could further explore the systemic barriers to minority participation in tech and the broader implications of inclusive design on societal equity.

Works Cited

Castro, J. O. (2003, September 10). *Inside the GNOME 2.4 Desktop and Developer Platform*. Ars Technica. https://arstechnica.com/information-technology/2003/09/gnome2-4/

Coleman, E.G.. (2013). *Coding freedom: the ethics and aesthetics of hacking*. Princeton University Press.

Damian, D. (n.d.). *Equity, Diversity, and Inclusion in Software Engineering*. Springer Nature.

Diversity in Software Development - BairesDev. (2022, October 24). *BairesDev Blog: Insights on Software Development & Tech Talent.* https://www.bairesdev.com/blog/diversity-in-software-development/

Erenkrantz, J. (2010). *Proceedings of the 3rd International Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Development*.

G.J. Downey. (2008). *Closed captioning: subtitling, stenography, and the digital convergence of text with television*. Johns Hopkins University Press.

Lomi, A., Conaldi, G., & Tonellato, M. (2012). *Organized Anarchies and the Network Dynamics of Decision Opportunities in an Open Source Software Project*. Research in the Sociology of Organizations, 36, 363–397. https://doi.org/10.1108/S0733-558X(2012)0000036017

Schwartz, S. D., Fickas, S. F., Norris, B., & Dubey, A. (2024). *A Survey of Open Source Software Repositories in the US Department of Energy's National Laboratories. Computing in Science and Engineering*, 1–8 https://doi-org.proxy1.library.virginia.edu/10.1109/MCSE.2024.3414951

Sonabend, R., Gruson, H., Wolansky, L., Kiragga, A., & Katz, D. S. (2024). *FAIR-USE4OS: Guidelines for creating impactful open-source software*. PLoS Computational Biology, 20(5 May). https://doi-org.proxy1.library.virginia.edu/10.1371/journal.pcbi.1012045

What is a stenographer? (n.d.). Planet Depos. https://planetdepos.com/trending/content/what-stenographer