A Set of Eyes (ASoE)

David Calvo, Jada Campbell, Jonathan Edwards, Hannah Lynn, Kavya Tumkur

Johns Hopkins University | Whiting School of Engineering | Baltimore, MD
Design Day 2021

Introduction

Despite the ubiquity of the web, the web browsing experience can differ between users as much as their daily lives do, particularly for people with visual disabilities. Blind and visually impaired users rely on screen readers to convey what users with typical eyesight can see on a display.

A screen reader works by looking at the code of a webpage and relays information about the webpage to the user, either through speech or Braille output. The user can then use their keyboard to navigate the web page. However, the screen reader's ability to understand the contents of a webpage depends wholly on how the webpage's underlying code is constructed, placing the responsibility of accessibility squarely in the developer's hands, who either don't know about the problem or don't care enough to mitigate it.

Problem

There are 285 million visually impaired people in the world and 7.7 million who reside in the United States,.

We want to make a better way for blind and low-sighted people to use a computer. Currently, they're dependent on technology which is in turn, dependent on people who don't have their best interests in mind.

We want to make a solution that makes them independent, empowering them to use the full capacity and range of technology just like everyone else.

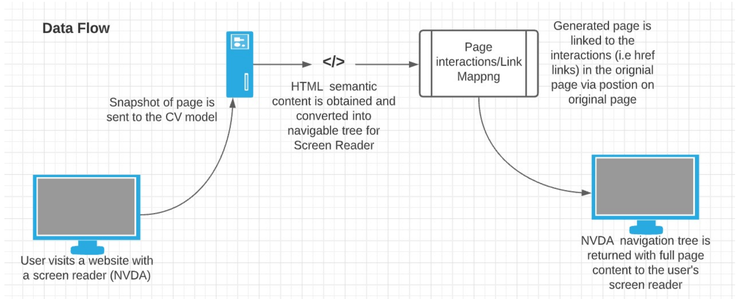
Solution

A Set of Eyes

A Set of Eyes is an AI-powered screen reader that ensures any website a visually impaired user visits is accessible regardless of developer implementation. What makes our screen reader different is that instead of looking at code, it uses artificial intelligence to look at the screen and intuitively describe what's on it. If a website is navigable by a user with typical eyesight, then our screen reader should be able to extract the same information that a sighted individual would and inform the user.

Our solution also bridges the gap between the underlying code of the webpage that a traditional screen reader navigates and the visual reality. A benefit to this approach is that our screen reader derives relationships from the visual elements of a webpage in much the same way a sighted user would, thereby improving the web browsing user experience in a way that other screen readers cannot.

Data Flow



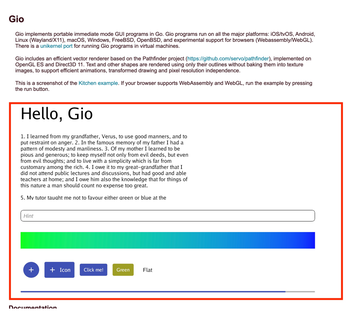
A Set of Eyes consists of a computer vision model backend and a flexible front end. When someone wants is using A Set of Eyes, a digital snapshot of the webpage is sent to the CV model, which recognizes the different parts of the webpage as semantic content, which in turn is generated through HTML. This new HTML representation of the webpage is then injected into the current webpage's HTML, filling in the semantic and visual gaps that the code doesn't translate on the screen.
This webpage is then passed to an NVDA endpoint that returns the navigation tree (i.e., how the code is organized via a DOM), back to the screen reader. The user then can navigate the rendered webpage through a series of interfaces: either traditionally, through the keyboard, or through a voice-activated UI.

Results

Our computer vision model went through a preliminary training of several hundred generated web pages to recognize five different types of web elements. What we found was that our model was not only able to accurately recognize which parts of the webpage corresponded to a section of code, but was also able to generalize to other well-known websites.

Displayed below is a web page documenting the JavaScript library Gio. On the left is the original webpage - and the box in red is a canvas element, which under a normal screen reader would be impossible to navigate. On the right is a visualization of our rendered webpage, which recognized the different types of elements inside the canvas element.

While our work so far details a model that recognizes 5 web elements, we can scale this technology to train on the many different types of items one can run into on the web.



Future Plans

Our market entry strategy is focused on the education sector, specifically targeting schools for the blind and visually impaired. The specific customers we have in mind are blind and visually impaired students who will be dependent on a screen reader for their entire lives. This makes visually impaired students exceptional lead users who can guide the development of ASoE towards a novel, intuitive user experience that will set us apart from other products as we enter follow-up markets. As the visually impaired community is small but tightly knit, we plan to reach them through word of mouth and direct marketing. We believe that word of mouth and positive testimonials within the community will be the most impactful way to get ASoE to capture the potential target market.