**1. Could you tell me a bit about you and your field?**

I’m an AI engineer and I specialize in the development of accessibility features for digital platforms. My area of focus is predominantly text-to-speech (TTS) technology, which enables digital content to be more accessible to individuals who are visually impaired, have reading disabilities or prefer to consume content audibly.

I hold a degree in computer science, with a specialization in artificial intelligence. Over the years, I've acquired valuable experience working on various projects involving machine learning, natural language processing, and speech synthesis technologies.

In my current role, I work on the same online bookstore app as Ding Shi, but I specifically focus on the computer website page version of the app. I develop and improve the TTS feature to ensure that it provides a smooth and pleasant listening experience for our users.

**2. In your experience, what are the challenges visually impaired users encounter when using Apps?**

The primary issue is that much of the content on apps may not be designed to be accessible. Textual content may not support text-to-speech technology, and images or videos may not have descriptive alt text or captions, leaving visually impaired users unable to fully understand the content.

Another challenge is that Apps not designed with accessibility in mind may have complex or unclear navigation, making it difficult for visually impaired users to move around and use the app effectively. This includes improperly labeled buttons, lack of clear hierarchy, or over-reliance on visual cues for navigation.

Many apps are designed primarily for touch or mouse input and do not fully support keyboard navigation. This can be a major barrier for visually impaired users who use a keyboard or assistive devices to navigate.

**3. Can you give me an example of how you apply different tools, applications, and methods to help users overcome their challenges?**

Our TTS technology leverages several key AI techniques to produce natural and clear spoken output from written text. Within our TTS function, NLP plays a critical role in understanding sentence structure, grammatical nuances, and contextual cues to accurately interpret and read the text. Once the text is processed, the next step is converting it into speech. We use advanced AI techniques to generate human-like speech. One technique we use is concatenative TTS, which pieces together segments of recorded speech from a voice actor. This can create a very natural-sounding voice, but can sometimes sound unnatural when piecing together complex sentences.

To overcome this, we also use parametric TTS, which uses rule-based systems to generate speech. It can be tuned and adjusted, which makes it highly flexible but can sometimes sound robotic.

Most recently, we've started exploring neural TTS systems, which use deep learning techniques to generate speech. This is very promising and can create highly natural-sounding speech even in complex sentences.

Also, we make sure our web app is fully navigable using a keyboard. This includes logical tab order, visible focus indicators, and support for keyboard shortcuts. This benefits users who rely on keyboard navigation or assistive devices that emulate keyboard input.

**4. What aspects of creating accessible content is most challenging, why are they challenging? and how have you overcome these challenges? Can you tell me about your process when you make these changes?**

A. Technical Limitations: Text-to-Speech (TTS) technology has evolved significantly, but it still has its challenges. One of these challenges is ensuring a natural speech flow that correctly interprets punctuation, intonations, and emotions in the text. This involves complex language processing and speech synthesis tasks that push the boundaries of current AI capabilities.

Another technical challenge is maintaining compatibility with various assistive technologies. Visually impaired users employ a range of assistive tools, including screen readers, Braille displays, and voice control systems. Ensuring our TTS feature and other accessibility enhancements function well with these diverse tools requires careful engineering and extensive testing.

B. Design Constraints: In striving for an accessible web app, certain design elements can indeed pose challenges. For instance, trendy design features like parallax scrolling or complex animations can be visually impressive, but they may not be compatible with screen readers and could be distracting or even disorientating for some users.

Similarly, design elements that rely on color (such as indicating active states or categories) can be problematic for users with color blindness. Therefore, we have to find ways to convey the same information that doesn't rely on color alone, like using icons or text labels.

C. Copyright Challenges with TTS: copyright in AI-generated content is a complex and evolving area of law. There have been debates about whether a synthesized voice can infrally copyright if it too closely mimics a famous person's voice. As of now, we take great care to ensure our synthesized voices are unique and do not mimic any specific individual to avoid any potential legal complications.

**5. Can you share any experiences where you modified content? What changes were made?**

A. When we initially rolled out our TTS feature, we noticed that certain complex sentences or phrases weren't being pronounced as naturally as we would have liked. The speech sounded choppy and disjointed, which was due to limitations in our early TTS technology.

To address this, we implemented a more advanced NLP algorithm and updated our speech synthesis techniques. We also began exploring Neural TTS systems, which use deep learning to produce more natural and fluent speech.

After these changes, the TTS feature could better handle complex sentence structures and the spoken output was much more natural and human-like. However, as this remains an area of active development and research, we continue to refine and improve upon these algorithms.

B. In the early days of our web app, we used a design aesthetic that was minimalist and heavily reliant on color to differentiate between different sections and actions. However, we received feedback from users with color blindness that this was causing confusion and usability issues.

In response, we re-evaluated our design approach and decided to modify our design language. We incorporated more icons and text labels and used texture in addition to color to differentiate between elements. We also implemented a high-contrast mode for users with low vision. While these changes meant moving away from our original minimalist aesthetic, they greatly improved the usability of our app for all users.

C. While working on the TTS feature, we initially used a concatenative speech synthesis approach that involved piecing together segments of recorded speech from voice actors. However, we wanted to be able to offer a wider range of voice options to our users, and recording new speech segments for each option would have been time-consuming and potentially problematic from a copyright perspective.

To address this, we started using parametric and neural speech synthesis techniques that generate speech algorithmically, which gave us greater flexibility to offer different voice options. We made sure all the voice data used to train our TTS models was explicitly recorded for this purpose and that the usage rights were clearly agreed upon to avoid any copyright issues.

**6. What aspects of creating accessible content is most challenging, why are they challenging? and how have you overcome these challenges? Can you tell me about your process when you make these changes?**

One of the biggest challenges is designing content that is equally accessible to all users, regardless of their physical abilities or disabilities. It's about trying to anticipate the needs of a diverse user base, and engineering a solution that can meet those needs without compromising functionality or aesthetics. Technology advances rapidly. While this can enable more powerful solutions, it also means that our team needs to stay current with the latest advances and understand how they can be leveraged to improve accessibility.

Our process begins with the user. We use techniques like interviews, surveys, and user testing to understand their needs and how they interact with our app. The insights gained are invaluable in guiding our design and development process. Based on our understanding of the user needs, we develop prototypes and implement changes. This could be tweaking the UI for better readability, modifying our TTS engine for more natural speech, or adjusting navigational features for easier access. We then test our changes, both internally and with users. This helps us validate our solutions and uncover any issues that we may have missed. We employ a mix of automated and manual testing, and where possible, we use real-world user testing to ensure our solutions are effective. We work closely with our legal team to ensure that all our content adheres to copyright laws and regulations. In the case of our TTS feature, we ensure we have the necessary permissions for any voice data we use. Technology, standards, and user expectations are always changing. We strive to keep abreast of these changes and regularly update our app to ensure it remains accessible and user-friendly.