

TorTalk eases reading challenges for dyslexics with Nuance OmniPage SDK

Challenge

- Develop an easy to use text-to-speech application to read books to people with dyslexia or to the sight impaired population
- Process text rapidly for audio
- Recognize and speak text from images and PDFs where text quality is sometimes less than perfect

Solution

- Nuance® OmniPage® SDK
- Nuance® Vocalizer SDK

Results

- Fast, reliable, accurate OCR
- Reads both fixed and editable text quickly
- Allowed TorTalk to capture 75% of the Swedish university market

Summary

Approximately 6 percent of Swedes struggle with dyslexia, including Swedish software developer Tor Ghai. For most of his life, he struggled with reading comprehension as he battled his way through required textbooks. As an adult, he searched for a way to make reading easier – and even enjoyable – so that he could stay on top of the latest in software development and maybe even read for fun. After developing a program that would read text to sight-impaired people and becoming familiar with text-to-speech engines, Tor began to develop a version for dyslexics like himself.

“The value of Nuance OmniPage SDK is in the TorTalk product and the difference it makes for the user. For me, when I started using text to speech, I could increase my ability to take in information five-fold. All of a sudden, I could learn so much so quickly. I went from not feeling so smart to feeling as smart as everyone else – it became a big boost for my self confidence and my productivity at work.”

Tor Ghai, software developer

As Tor continued to develop his solution, known as TorTalk, he discovered that a lot of the text-to-speech products on the market couldn't process images, PDFs, or any document that did not have editable text. He started using a desktop product to take a screenshot, turn it into editable text, and have it read out loud by a text-to-speech program but found that cumbersome and too slow for consuming longer works. He knew there was a better way to do this, and he began his search for a product that could work with fixed and editable text.

Accuracy is paramount for reading

However, finding an OCR engine to work with images of text was harder than it sounded, especially when it came to pairing the OCR engine with a voice engine that would speak the text to the user. While plenty of solutions exist for both, most of the vocalizer solutions would mangle approximately 20 percent of the words on a page – not terrible if you're only reading one page, but cumbersome if you're trying to read along with the audio in a multi-page document or a book. Misspeaking the word could derail the user's train of thought. And, sometimes, that missing 20 percent of information can mean the difference between understanding what you just read and being utterly confused. Additionally, most OCR products required the images to be high-quality, which isn't always the case in practice. Building an OCR engine was out of the question, as it would take a lot of time.

Of all the products tested, Nuance OmniPage SDK and Nuance Vocalizer had the highest rates of accuracy for recognizing, processing, and speaking text. Combining the two would make it possible for TorTalk users to get the most accurate readings. While the product will sometimes miss a word, the flow is smoother for those reading along with the text.

Additionally, Nuance OmniPage SDK provides real-time OCR capabilities. In one to two seconds, TorTalk activates Nuance OmniPage SDK to process the text, then read it to the user.

The user simply clicks a button, and the program starts reading the text, whether it's a single page or an e-book.

Nuance support eases development

Tor wanted to ensure TorTalk would work exactly as planned. He had a lot of questions for Nuance support, probably more questions than the average developer, he admitted. However, Nuance, and its London representative in particular, was very responsive to questions and provided him with the information he needed. Tor recently attended the OmniPage CSDK Developer's Forum in Palma de Mallorca and was able to learn more about Nuance OmniPage SDK and how he can develop with it, as well as network with other developers and share ideas.

Capturing the university market

With Nuance OmniPage SDK as its OCR backbone, and Nuance Vocalizer as the text-to-speech engine, Swedish universities quickly became interested in using TorTalk for its students. Tor had originally developed TorTalk for Windows operating systems, but with universities clamoring at his door for a Mac version, he found that, with the Nuance OmniPage SDK, he could just move most of the original code to a Mac solution.

Today, 75 percent of Swedish universities, including Goteborgs Universitet and Uppsala Universitet, use TorTalk for their students. In the most common case, a student will load a PDF or ebook onto a computer screen, then place the TorTalk window around the portion to be read. If it's electronic text, the student will highlight the text and press play. The text to speech engine then reads the text to the student. The biggest value is in the accuracy and speed, which helps students regain confidence in their ability to learn – something that Tor himself felt as he used TorTalk.

“The value of Nuance OmniPage SDK is in the TorTalk product and the difference it makes for the user,” Tor said. “I went from not feeling so smart to feeling as smart as everyone else – it became a big boost for my self-confidence and my productivity at work.”

About Nuance Communications, Inc.

Nuance Communications, Inc. is a leading provider of voice and language solutions for businesses and consumers around the world. Its technologies, applications and services make the user experience more compelling by transforming the way people interact with devices and systems. Every day, millions of users and thousands of businesses experience Nuance's proven applications. For more information, please visit: www.nuance.com.