

מנגנון שאילת שאלות ומתן התשובה לקצה לקצה לשיפור הבנת הנקרא עבור ילדים עם קשיי קריאה

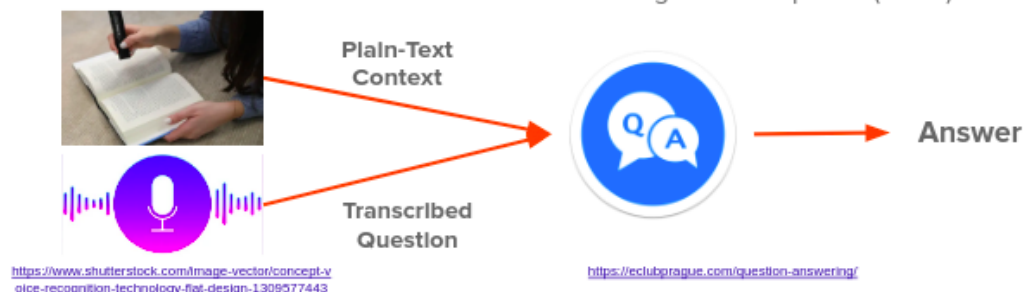
טל רוזנוין

תקציר: דיסלקציה משפיעה על 15-20% מאוכלוסיית העולם. אנשים דיסלקטים חווים קושי בקישורים שפתיים כגון איות, הגייה והבנת הנקרא.

בהרצאה זו, אציג יכולת של שאילת שאלות ומתן התשובה המשפרת את יכולת הבנת הנקרא עבור ילדים עם קשיי קריאה. היכולת פותחה עבור ה OrCamLearn, ועובדת כדלקמן: ראשית, תמונה של מסמך פיזי נרכשת (לדוגמא ספר) ומנוע OCR מחלץ את הטקסט מהתמונה. לאחר מכן המשתמש שואל שאלה על ידי דיבור למכשיר, והמכשיר מתמלל את התשובה באמצעות מנוע תמלול (ASR). התוצרים של מנועי ה OCR וה- ASR מוזנים מודל NLP המתמחה במתן תשובה לשאלה, והמנוע פולט את התשובה (אם קיימת) טקסטואלית. התשובה הטקסטואלית מועברת למשתמש באמצעות מנוע שמתרגם טקסט לדיבור (TTS) דרך רמקול שממוקם על גבי המכשיר.

A question answering feature that works as follows:

- The OrCam Learn captures an image of a physical document that we want to query.
- An OCR engine extracts the text.
- The user asks a question (by speaking to the OrCam Learn).
- The OrCam Learn transcribes the user's question using an ASR model.
- OCR and ASR outputs are passed to NLP Q&A model that returns the answer (if exists).
- The OrCam Learn communicates with the user using a text to speech (audio) interface.



Title: End-to-End Question Answering on an Edge Device For The Benefit of People With Reading Difficulties

Abstract: Dyslexia affects 15-20% of the world's population; it is a language-based learning disability that results in difficulties with specific language skills, particularly reading. Dyslexic people usually experience difficulties with other language skills such as spelling, pronouncing words, and reading comprehension.

In this talk, I'll present a new feature that we are developing for the OrCamLearn, a device that helps to improve the comprehension capabilities of people with reading difficulties. Specifically, I'll present a new question-answering feature that works as follows: an image of a physical document (i.e., book, document, etc.) is captured. Then, an OCR engine extracts the text. Next, the user asks a question by speaking, and the device transcribes

the user's question using an ASR model. Both OCR and ASR outputs feed a NLP Q&A model that returns the answer (if it exists) via a text to speech audio interface.