# iSinkwe: Automatic Synchronisation of Text Ebooks with Audio

**Rynhardt Kruger**

**21 March 2024**

# Introduction

- iSinkwe synchronises text ebooks with audio
- Audio either computer generated or human-narrated
- Designed for EPUB3 documents
- iSinkwe means bush baby in isiZulu, a nocturnal African monkey with big eyes and ears
- iSinkwe enables multisensory reading

# CSIR (Council for Scientific and Industrial Research)

- National scientific and research organisation

- Researches, develops, localises and diffuses technologies to accelerate socioeconomic prosperity in South Africa

- Human language technologies (ASR, TTS, NLP) in all 11 official spoken languages

# Motivation and Context

- Developed in South Africa

- Multilingual, eleven official spoken languages

- Reading barriers: print disabilities, low literacy

- Sporadic internet connectivity

# ISinkwe Overview

## Components

iSinkwe consist of three components:

- iSinkwe Convert: converts various formats to EPUB3
- iSinkwe Synchronise: synchronises EPUB3 with audio
- iSinkwe Read: EPUB3 reader supporting media overlays

# ISinkwe Convert

- Accessible via web interface
- User uploads document in various formats
- Document converted to EPUB3
- EPUB3 available for download

# Convert DOCX or PDF

Select DOCX or PDF file*

| Choose File | No file chosen |

Select EPUB 3 layout*

Reflowable layout ⌄

Enter EPUB 3 title

Enter EPUB 3 author

🗗 Convert

# ISinkwe Synchronise

- Accessible via web interface

- User uploads EPUB3

- User chooses human or computer audio

- Human-narrated audio fills in chapters with computer audio

- User notified when synchronisation completed

- Augmented EPUB3 available for download

# iSinkwe

EMPOWERING EYES AND EARS

Convert ▾  Synchronise ▾  Read  Help  Log out test

# Synchronise EPUB 3

Select EPUB 3 file*

Choose File | No file chosen

Select EPUB 3 language*

---------

Select audio type*

Computer-generated audio

Select voice characteristics*

---------

⇄ Synchronise

# ISinkwe Read

- EPUB3 Reader supporting media overlays
- Allows navigation in audio by
  word, sentence, or paragraph
- Synchronised text highlighted on screen
- Playback continuous or on demand
- Search function

In another moment down went Alice after it, never once considering how in the world she was to get out again.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went

# Technical Description

## Computer Generated Audio

- Text synthesised via TTS
- Qfrency TTS for South Africa's official spoken languages
- TTS audio added to EPUB3

# Human-narrated Audio

- TTS compared with human audio for time alignments
- Comparison via dynamic time warping algorithm
- DTW requires language-specific TTS
- South African languages facilitated by Qfrency TTS
- Human audio added to EPUB3

# Previous Usability Study

## Candidates

- Ten candidates
- Five blind and low-vision users
- Two users from publishing industry
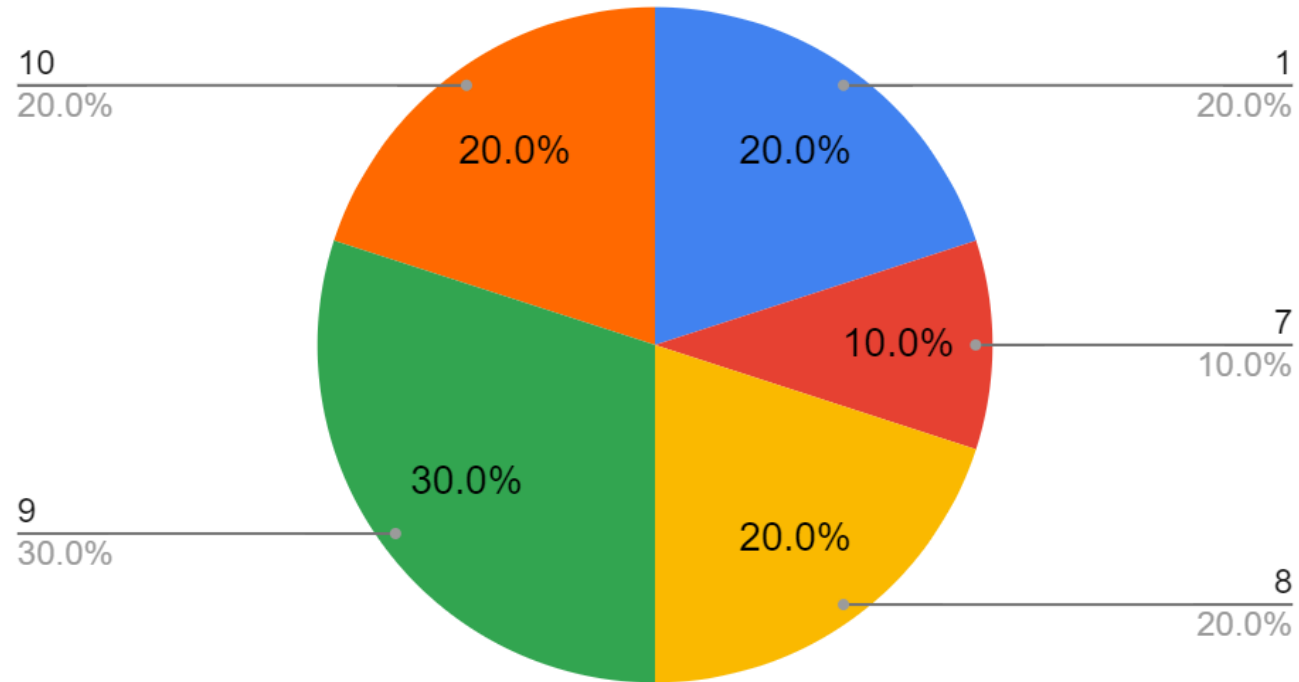- Three educators of print-disabled learners

# Evaluation Description

- Users evaluated iSinkwe Convert, Synchronise, and Read

- Asked to perform tasks with own and provided documents

- Qualitative and quantitative questions, including:
  - What users liked the most

  - What users liked the least

  - Whether they experienced any issues

  - What they would change

  - Rating of the component from 1 to 10
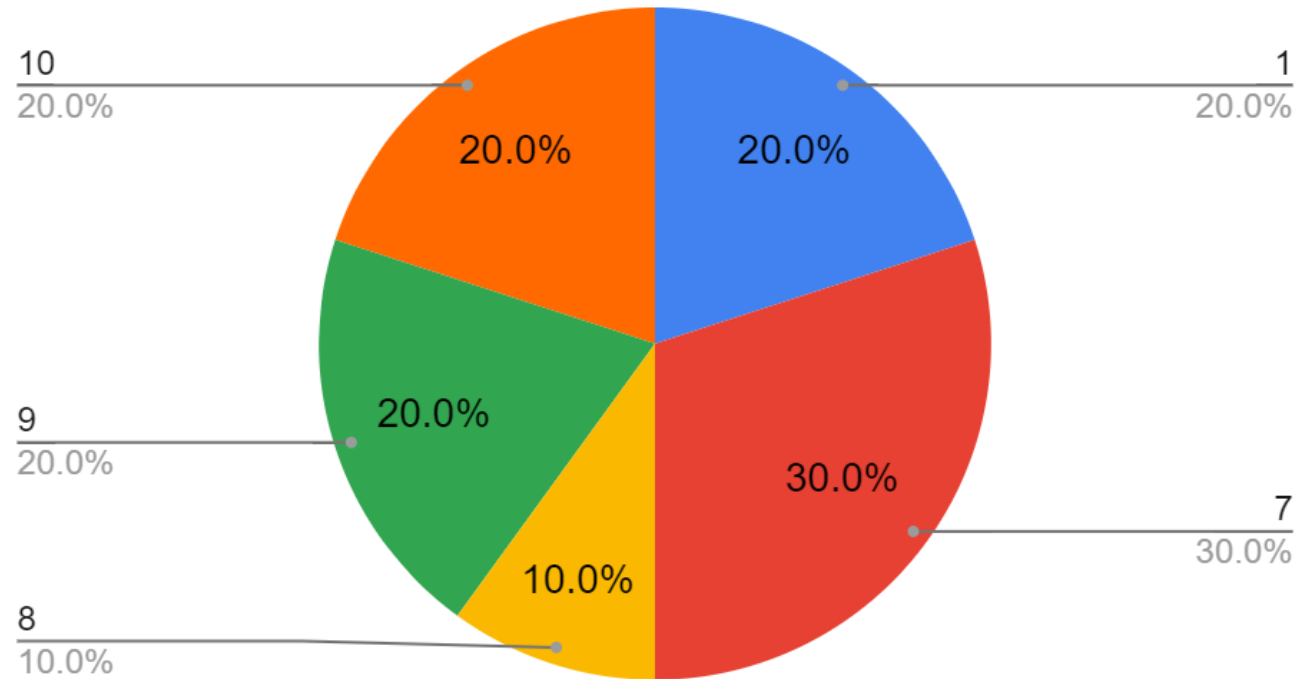
# Component Ratings

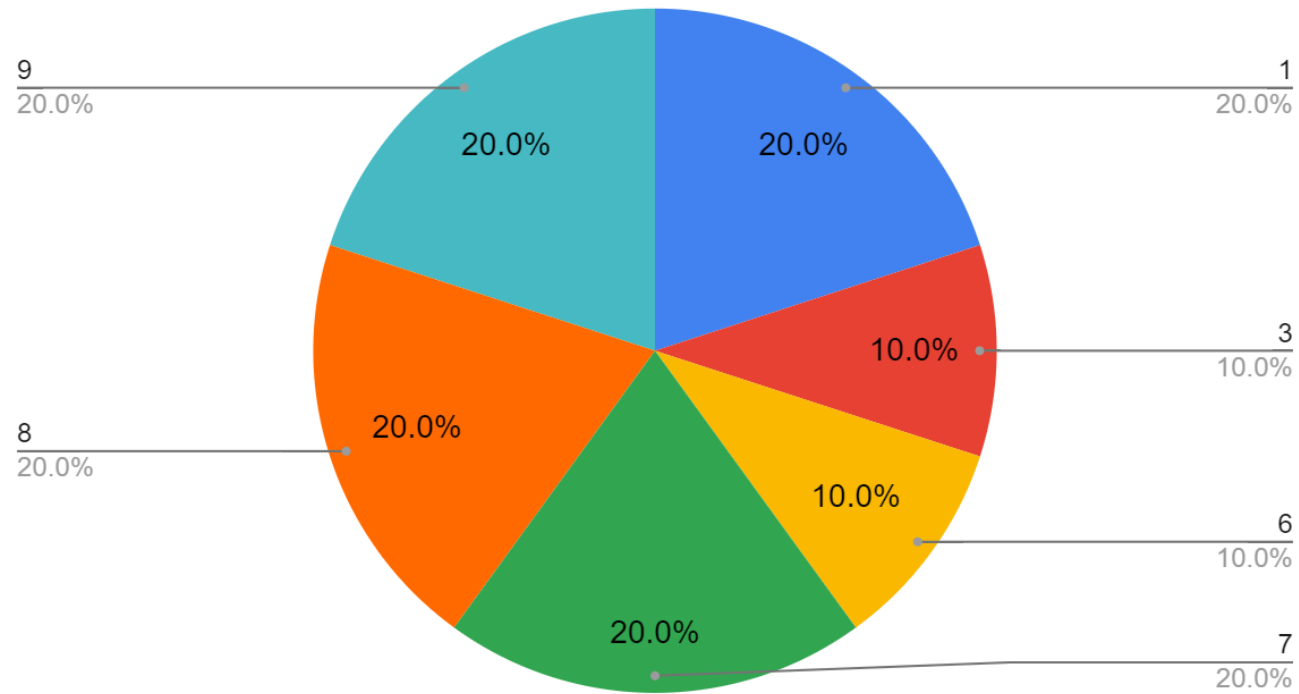## ISinkwe Convert



Possibility of recommending the Converter

- 10 — 20.0%
- 1 — 20.0%
- 7 — 10.0%
- 8 — 20.0%
- 9 — 30.0%

# ISinkwe Synchronise



Possibility of recommending the Augmenter

| Label | Percentage |
|---|---|
| 10 | 20.0% |
| 1 | 20.0% |
| 7 | 30.0% |
| 8 | 10.0% |
| 9 | 20.0% |

# ISinkwe Read

Possibility of recommending the Reader

# Discussion

- BLV users positive on iSinkwe Convert and Synchronise, suggested WCAG improvements
- BLV users mixed reviews on iSinkwe Read, we focus on existing readers as well
- Publishing industry mainly interested in iSinkwe Convert and Synchronise
- Educators least positive experience with iSinkwe
- Educators use existing AT allowing reading and writing in same document

# Future Work

- In development: recognition of mathematics

- Exploration of diagrams

# Contact Information

- isinkwe.com

- info@isinkwe.com

- rkruger@csir.co.za

# Demo

- Uploading a book to iSinkwe Synchronise
- Reading a book using Thorium
- Reading a book using iSinkwe Read