# 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

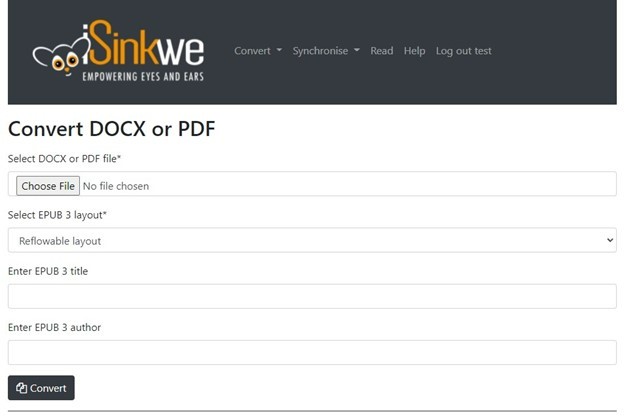
## Components

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

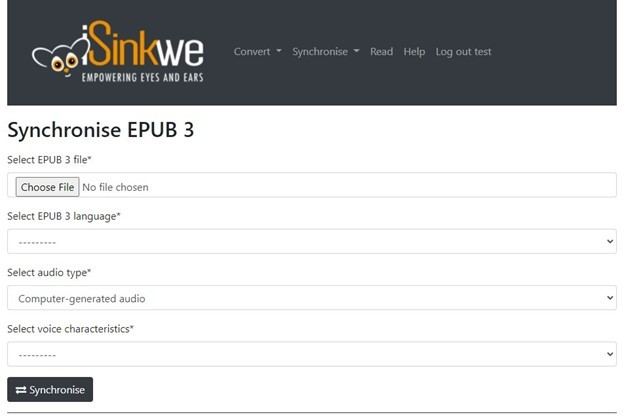
* Accessible via web interface
* User uploads document in various formats
* Document converted to EPUB3
* EPUB3 available for download



Screenshot of the web interface for iSinkwe Convert width=18cm

## 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



Screenshot of the web interface for iSinkwe Synchronise width=18cm

## 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



Screenshot of the web interface for iSinkwe Read height=15cm

# 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

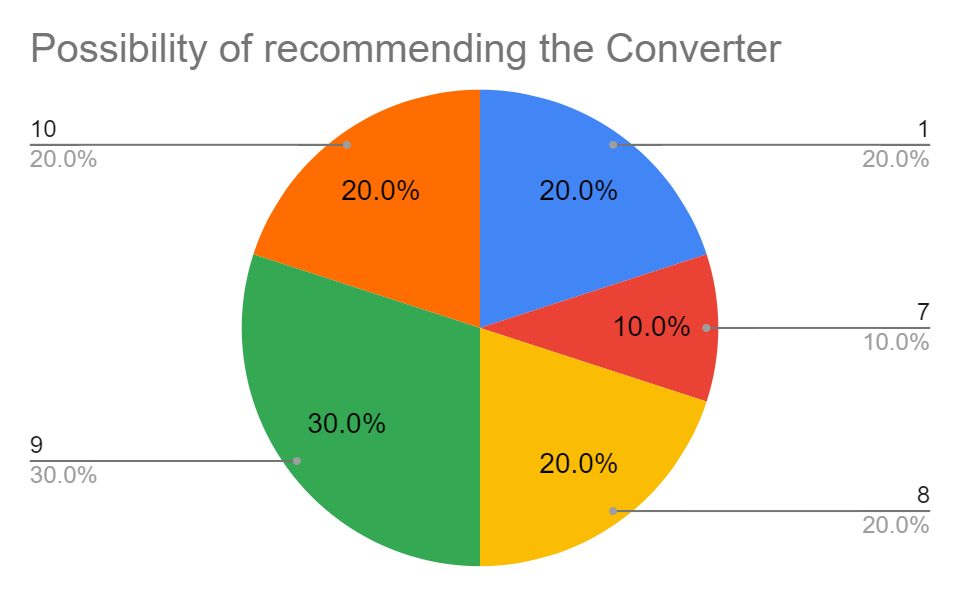
* Ten candidates
* Five blind and low-vision users
* Two users from publishing industry
* Three educators of print-disabled learners

## 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

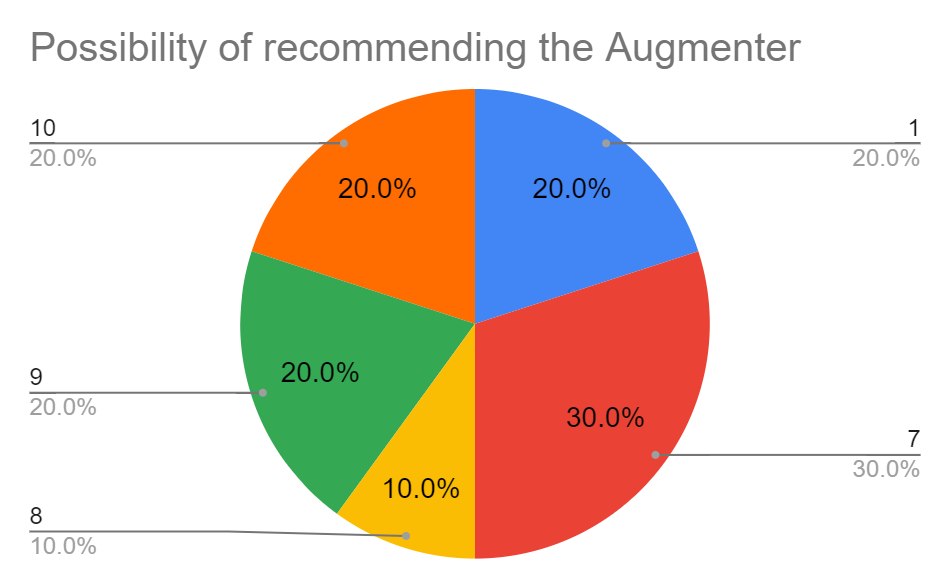
## Component Ratings

### ISinkwe Convert



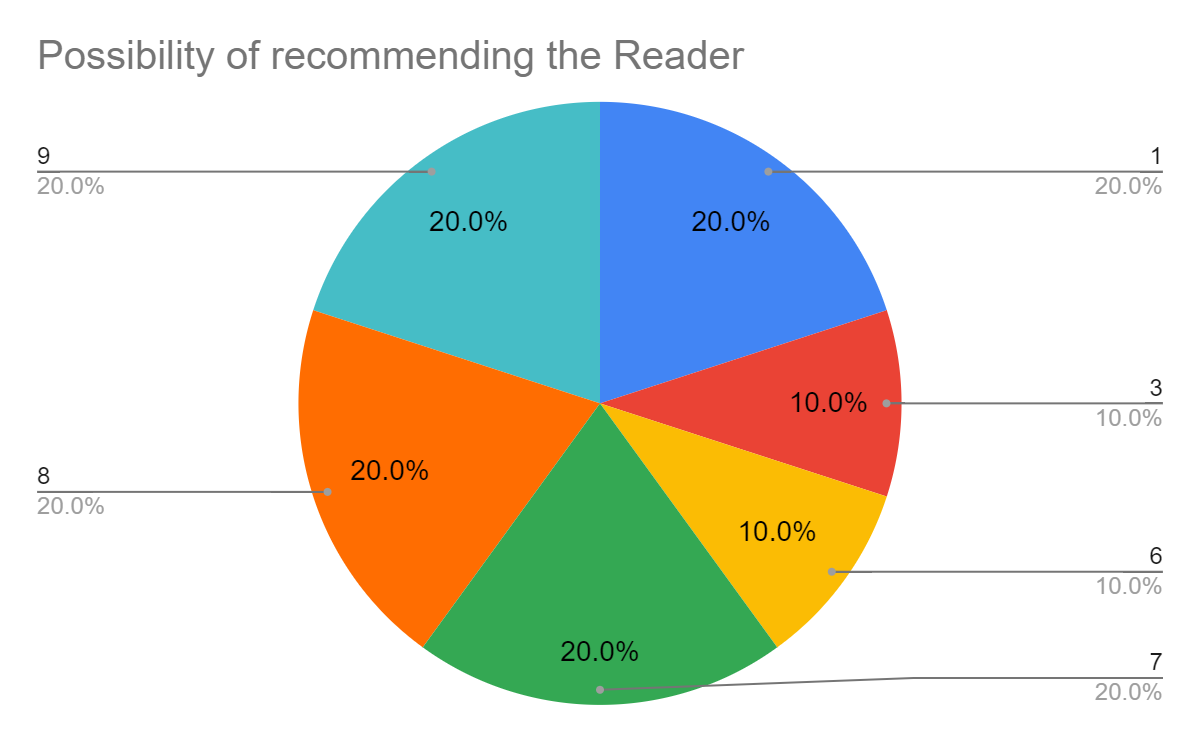
A pi chart depicting the possibility of users recommending iSinkwe Convert to someone they know. 20 percent of users gave a score of one, 10 percent gave a score of 7, 20 percent a score of eight, 30 percent a score of nine, and 20 percent a score of ten. width:18cm

### ISinkwe Synchronise



A pi chart depicting the possibility of users recommending iSinkwe Synchronise to someone they know. 20 percent of users gave a score of one, 30 percent a score of seven, 10 percent a score of eight, 20 percent a score of nine, and 20 percent a score of ten. width:18cm

### ISinkwe Read



A pi chart depicting the possibility of users recommending iSinkwe Read to someone they know. 20 percent of users gave a score of one, 10 percent a score of three, 10 percent a score of six, 20 percent a score of seven, 20 percent a score of eight, and 20 percent a score of nine. width:18cm

## 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