Software Requirements Specification

for

<Book Reader of Blind People>

Version 1.0 approved

Prepared by <Shahir Abdullah Bin Shahnoor (201714054) , Shariar Riday (201714037), Nazifa Hamid (201714044), Takik Hasan (201714038), Zubair Hasan (201714041)>

<Group 4 Section B>

<24/08/2019>

Table of Contents

1. Introduction

1.1 Purpose

The product is an audio book reader for the blind students or any blind people who wants to know the content of a text document. They can get this information by taking the photo of the document page and then the product or system will read the information of the document out loud with the help of text to speech technology. This SRS document contains the whole system architecture and model of the system along with the scope of the system and necessary functionalities of the system.

1.2 Intended Audience and Reading Suggestions

This document is intended for the developers of the product or doing relevant works on this subject of engineering. This document contains all the necessary components and architecture model of the product and system. Anyone who is interested in such kind of field or work can get necessary information about it reading this document.

1.3 Product Scope

This product is for blind students in Bangladesh who face several obstacles everyday life during their study period. Although there are several systems for their ease of education there aren't much which provides a good education system based on Bangla language. And since this product has also an android version so the students or users can also read and know about any contents of a text document remotely with just taking the photo of the page of that document. The scope of the product is totally based on the students.

1.4 References

1. *https://ieeexplore.ieee.org/document/7732406*
2. *https://paperswithcode.com/task/text-to-speech-synthesis/codeless*
3. *http://www.academia.edu/Documents/in/Text_to_Speech*
4. *https://www.researchgate.net/publication/319406874_Text_-_To_-_Speech_Synthesis_TTS*
5. *https://www.researchgate.net/post/What_are_the_must-read_papers_on_text-to-speech_synthesis*
6. *https://waset.org/publications/8303/the-main-principles-of-text-to-speech-synthesis-system*
7. *https://www.naturalreaders.com*
8. *https://www.aclweb.org/anthology/L18-1354*

2. Overall Description

2.1 Project Objective

To build a device that can read text out loud from any printed book for the blind or visually impaired by

i. Extracting text from an image.
ii. Converting extracted texts to speech.

2.2 Product Perspective

This product is for helping the blind people read any books without the help of any third person. Our product aims to help the blind people study and enjoy Bengali literature. This type of product already exists but there is no option for Bengali language.

2.3 Product Functions

- Scan Bengali books
- Read Bengali books to the user
- Book placement correction
- Reading speed settings

2.4 Operating Environment

There will be a stand for the user to keep the book on.There will be an app for the android mobile that will be kept on the stand and will be used to read the book aloud to the user.

2.5 Design and Implementation Constraints

The main constraint in our project is the scarcity of data. There are not many available datasets and research in the field of Bengali OCR. So, we have to create some datasets from the scratch which can reduce the accuracy of the machine.

3. Requirement Elicitation Process

3.1 Requirement Discovery:

i. We have found out several journals and IEEE papers about text to speech conversion and also real time text to speech conversion. And how Phonetic speech analysis is used for text conversion into speech.
ii. We have also looked for resources and projects on the internet about this text to speech technology. We found Google text to speech API and Amazon Polly which uses this technology. We also found a software named Jaws which is a window plugin for windows operating system which helps the blind people to use the computer by converting each text into audio.
iii. We also visited 2 blind university students in Jagannath University and asked few questions about their study methods and how they tackle their physical disability everyday while studying. The Q/A were :

1. How do you study and manage to read books which doesn't have brail system ?
2. What are the obstacles you face everyday during study?
3. Is their any system or softwares you use for helping you study?
4. If there are any softwares what are those?
5. Do you think the product will help you study in a better and efficient way?

6. Do you think there should be a mobile(android / ios) version of the product?

Through these activities and questions we gathered our necessary requirements to illustrate the elicitation process.

3.2 Requirements Classification and organization:

i. Robust architecture of the book reader and audio conversion mechanism for converting any kind of Bangla text to speech.
ii. Mobile application for portability and convenience for the user to convert text in anywhere in the world.
iii. Machine learning for getting better at converting and recognizing any kinds of fonts and texts.
iv. Navigating the user by audio commands for reaching for specific contents for user friendliness and convenience.
v. Controlling the system and using the system by audio commands.
vi. Algorithms for finding a specific contents and specific points for a paragraph quickly.

3.3 Requirement prioritization and negotiation:

We found out through our requirement discoveries that there are already several systems and products that the blind students are already using. One of them is a software named Jaws which helps the blind students in using operating system and computer by converting every texts into speech. But that software doesn't have an android version so the user can't convert any kinds of text the want to know about unless they take a photo of that content and then scan it to make a pdf file and put it in their computer. So it is a very lengthy process and sometimes tough too because they need a friend to help them out to convert those contents and putting those in their computers. Also there are no Bangla text to speech software either for android version. So the students have to relay on their friends to read the contents of Bangla books or papers. So our product will give all the supports that the products are giving which the students are currently using as well as make their readability much easier and independent. So our top priority will be making an android or ios version of the product so that users can use it anywhere also put machine learning voice command technology so that they don't have to relay on others to scan the contents of the paper or book they want to read or know about.

4. User Requirements

- The user should be able to place any types of books.
- There should be feedback for the end of page.
- The voice from the speaker should be audible.
- User should be able to set reading speed.
- User feedback to start reading.
- There should be a voice command system to use the product.

5. System Requirements

5.1 Book Placing
5.2 Cursor detection
5.3 Voice audibility
5.4 Setting Reading Speed
5.5 Reading initiation feedback

5.6 Voice command

5.1 Book Placing

The image of the page will be taken through pi-camera. Holder can be adjusted for wide angle of camera to capture different sizes of books.

5.2 Cursor Detection

When the end of the page will be reached there should be a command from the system to change the pages to the next page or another page the user want to listen to. Through image recognition the system will know the end of page. A vibrator can also be attached to let the user know that the end of the page has been reached along with the voice command.

5.3 Voice Audibility

A smart speaker will be used with raspberry pi.

5.4 Setting Reading Speed

There will be a feature which will help the user to set the reading speed of the system. A time difference will be the variable in a loop. Speed can be adjusted with a knob or voice command by the user.

5.5 Reading Initiation Feedback

A voice command will start the reading of the book or page the user wants to read listen to.

5.6 Voice Command

The voice command will be used by the user so that he/she doesn't have to relay on others to user the system. We will use speech recognition technology for this voice command feature.

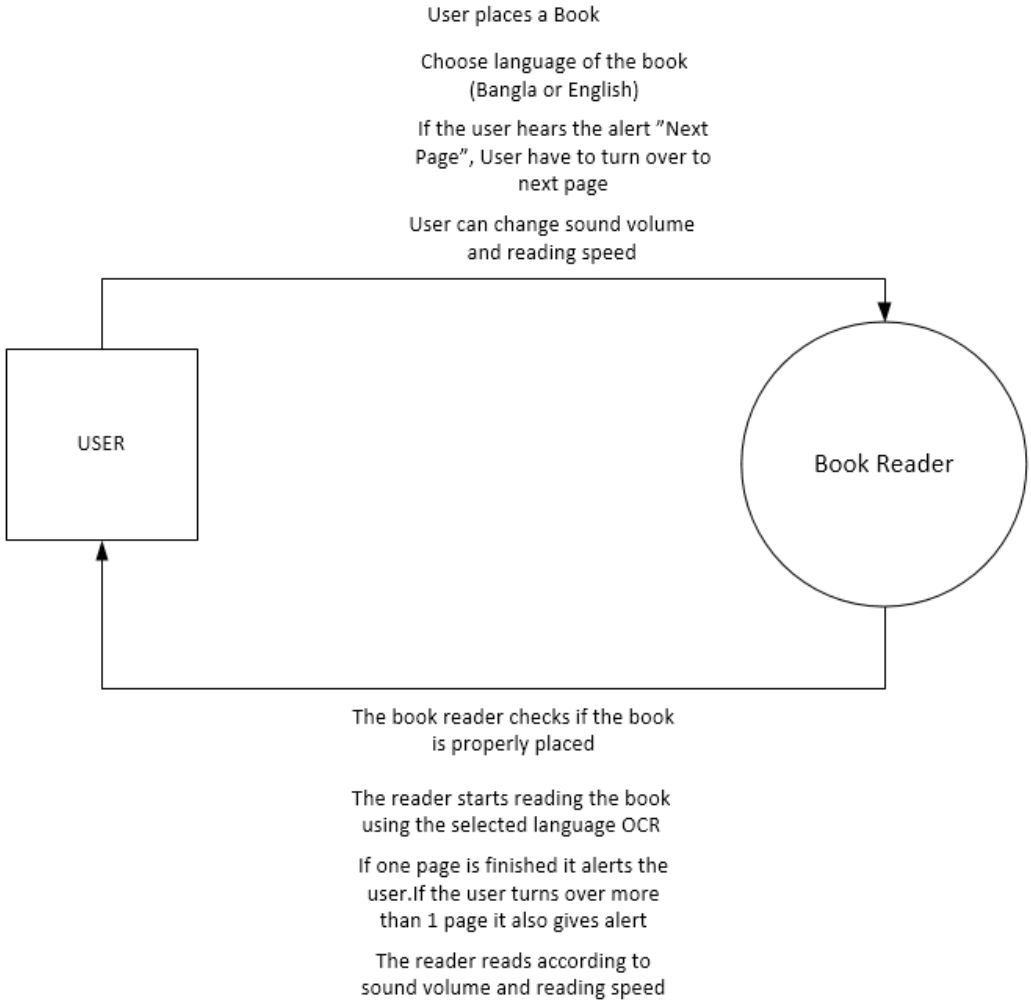
6. Requirements Classification

Serial	User Requirements	Types of Requirements (Functional)	Types of Requirements (Non Functional)
01.	Placing any types of books	✓	
02.	Cursor detection	✓	✓
03.	Voice audibility	✓	✓

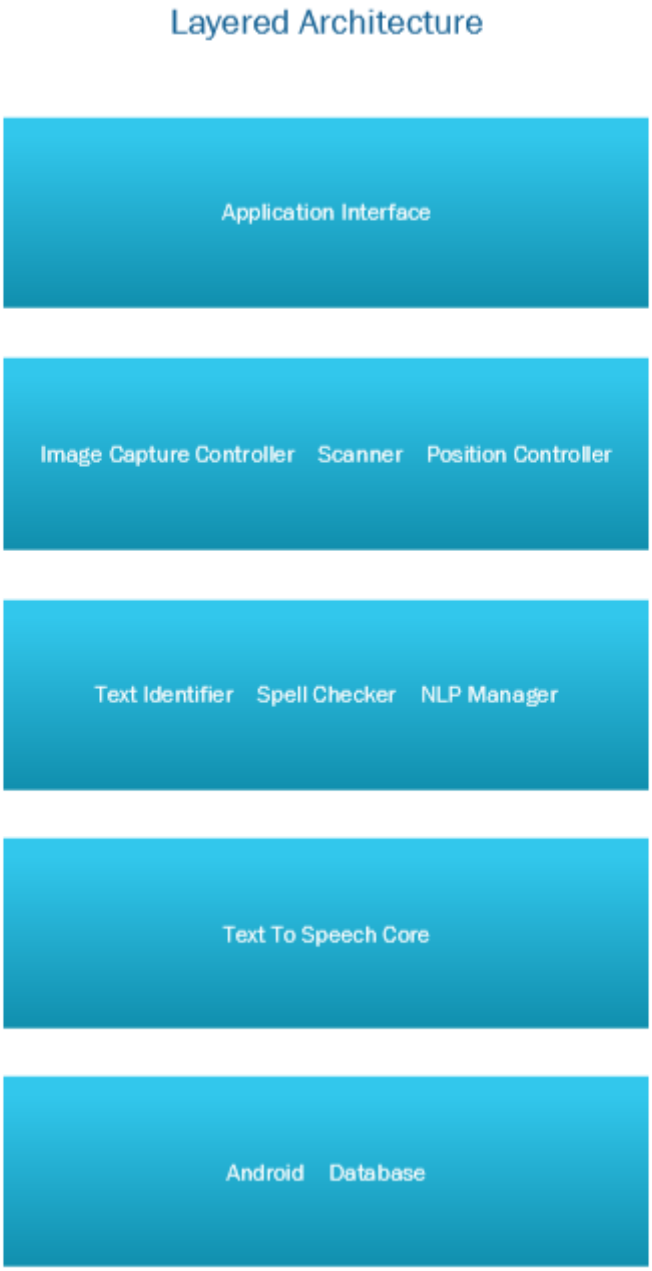
04.	Setting reading speed	✓	
05.	Reading initiation feedback	✓	✓
06.	Voice command	✓	

7. Graphical Representation

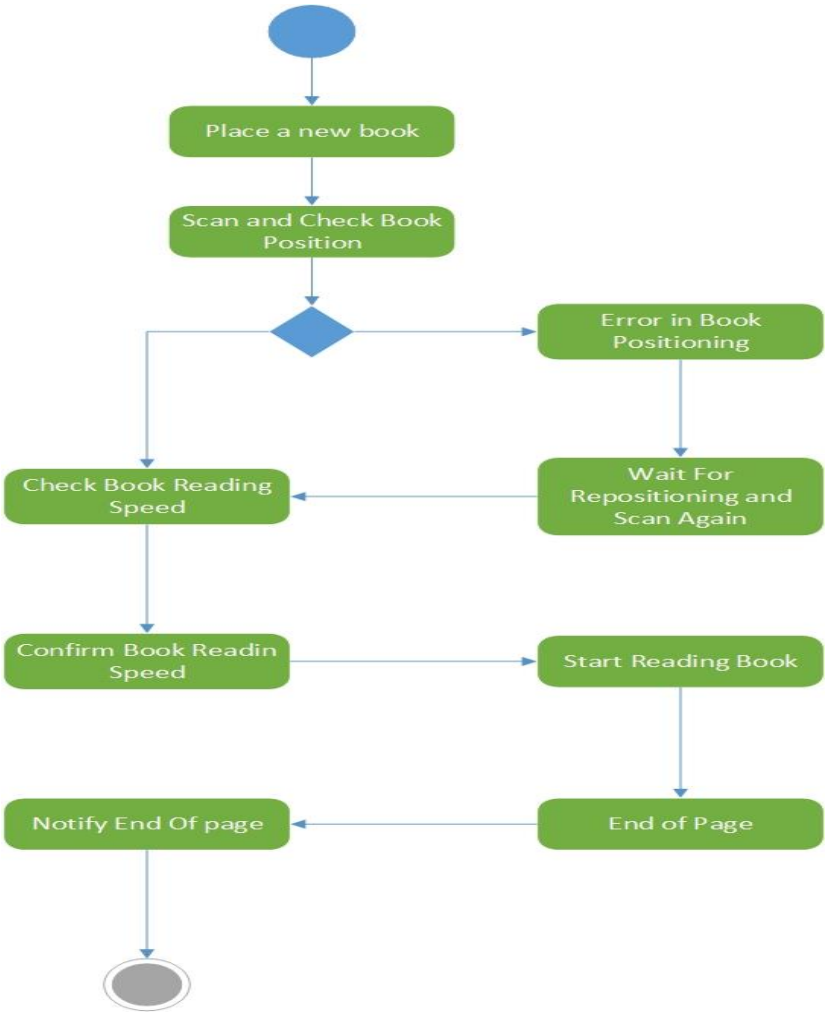
7.1 Context Diagram



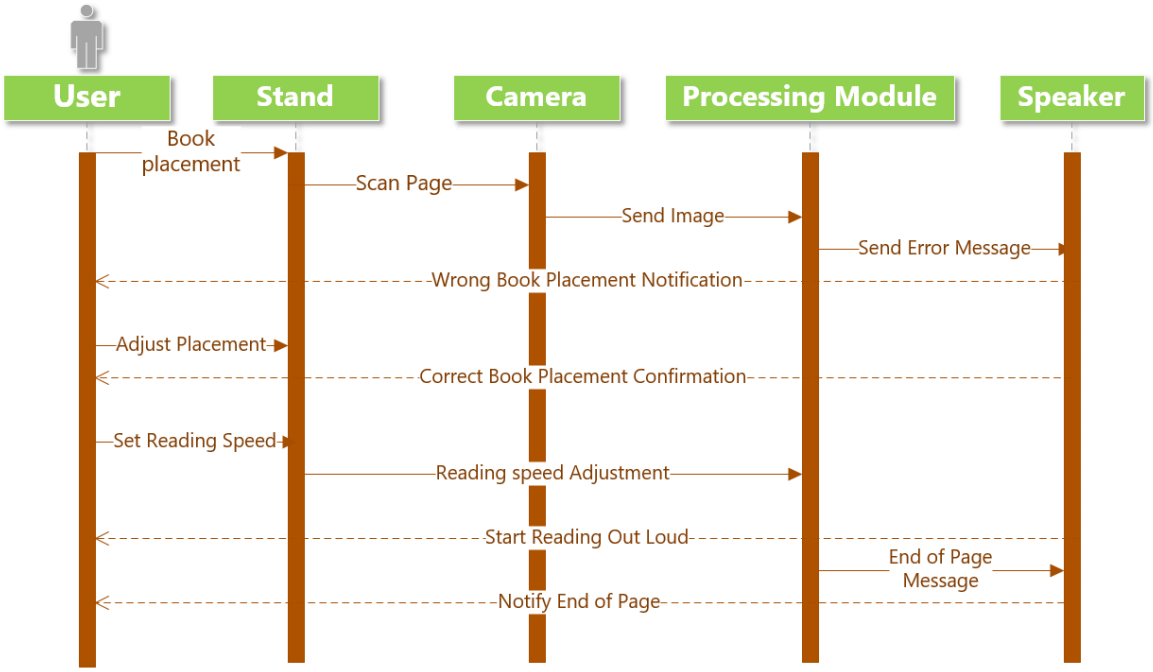
7.2 System Architecture



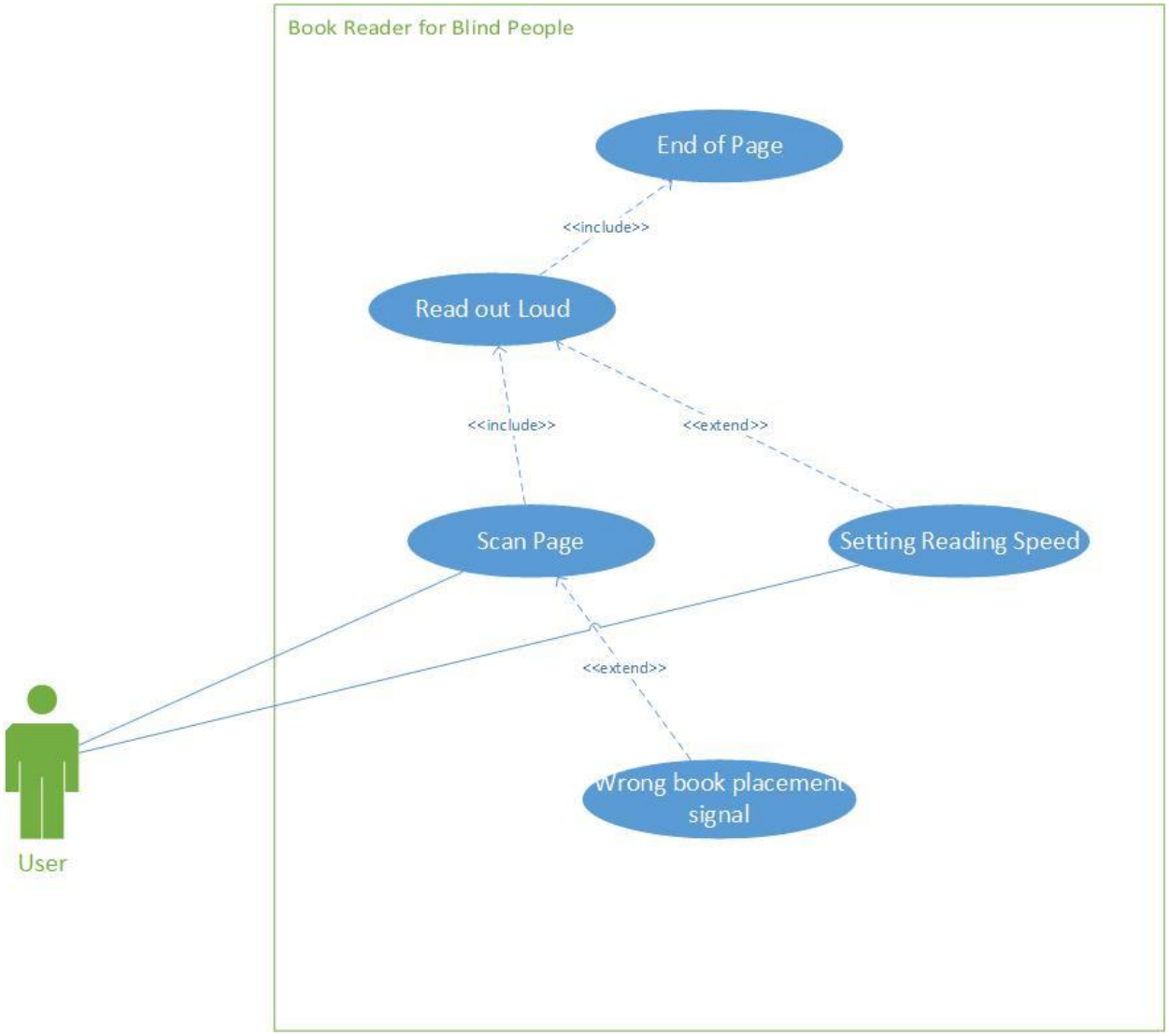
7.3 Activity Diagram



7.4 Sequence Diagram



7.5 Use Case Diagram



Use Case : Scan Page.

Primary actor : User

Secondary actors : -

Pre-condition: The system has a book for scanning.

Post-condition: -

Main Flow :

1. The use case is activated when the user requests it.
2. The user puts a book or a page to listen to.
3. The book is placed correctly.
4. The system initiates scanning.
5. If the scanning is finished, and [include "Read out loud"].

Alternative Flow :

3a. The book is not placed correctly.
3a1. [Extends use case "Wrong Book Placement Signal"].

Use Case : Wrong book placement signal.

Primary actor : Scan page use case.

Secondary actor : - User.

Pre-condition : The book is placed wrongly.

Post condition : The book is placed correctly.

Main flow :

1. The wrong book placement signal is activated by the scan page use case extention**.**
2. The user is notified that the book has been placed wrongly.
3. If the user places the book correctly, the use case is concluded.

Alternative flow :

3a. The book has not been placed correctly.
3a1. Return to the "Wrong book placement signal" use case.

Use Case : Read out loud.

Primary actor : User.

Secondary actor : -

Pre-condition : The book page is scanned.

Post condition : The scanning has reached the end of the page.

Main flow :

1. The use case is activated by the request of the user**.**
2. The use case has converted the scanned text file into audio file.
3. The audio file is being read by the system.
4. If the user is happy with the reading speed, reading is being continued at the default speed.
5. If the end of page is reached, the use case is concluded and [include use case "End of Page"].

Alternative flow :

4a. The user wants to change the reading speed of the system.
4a1. The use case extends the use case "Setting Reading Speed".
4a2. Return to the main flow.

Use Case : End of Page.

Primary actor : User.

Secondary actor : - System.

Pre-condition : The book has been read till the end.

Post condition :-

Main flow :

1. The use case is activated by the end of the use case "Read out loud"**.**
2. The use case concludes the reading of the current page in the system.

Use Case : Setting reading speed.

Primary actor : User.

Secondary actor : - Read out loud use case.

Pre-condition : The user requests setting reading speed use case.

Post condition : The reading speed of the system has been changed according to the user's need.

Main flow :

1. The use case is activated by the request of the user**.**
2. The use case has changed the reading speed of the system.
3. If the reading speed is set to the user's need the use case concludes the shifting of speed and returns to the "Read out loud" use case.

Alternative flow :

3a. The user wants to change the reading speed of the system.
3a1. The use case returns to the main flow.

Use Case Scenario :

1. Listening a page :

Goal in context: The user will listen to the system and know the contents of the page.
Scope: System.
Pre condition: User will place the book correctly on the system.
Success end condition: User will know the content of the page.
Failed end condition: User will not know the content of the page.
Primary actor: User.

Main Success Scenario:

1. System scans the page correctly.
2. System reads the page out loud.
3. Reading speed is right for the user.
4. When the end of page is being reached, reading ends.

Extensions:

3a. Reading speed is not right for the user.
3a1. Returns to "setting reading speed" case.
3a2. Return to main scenario.

2. Wrong book placement:

Goal in context: The user will place the book in wrong direction.
Scope: System.
Pre condition: User will not place the book correctly on the system.
Success end condition: User will know the content of the page.
Failed end condition: User will not know the content of the page.
Primary actor: User.

Main Success Scenario:

1. System doesn't scan the page correctly.

2. System calls "Wrong book placement" use case.

8. References

9. *https://ieeexplore.ieee.org/document/7732406*
10. *https://paperswithcode.com/task/text-to-speech-synthesis/codeless*
11. *http://www.academia.edu/Documents/in/Text_to_Speech*
12. *https://www.researchgate.net/publication/319406874_Text_-_To_-_Speech_Synthesis_TTS*
13. *https://www.researchgate.net/post/What_are_the_must-read_papers_on_text-to-speech_synthesis*
14. *https://waset.org/publications/8303/the-main-principles-of-text-to-speech-synthesis-system*
15. *https://www.naturalreaders.com*
16. *https://www.aclweb.org/anthology/L18-1354*