



The University of Jordan

**School of Engineering
Department of Computer Engineering**

Detecting and reading Arabic texts for vision impaired people

Supervisor:
Eng. Saadeh Sweadan

Author(s):

Yara Hasan Ahmed Al Jarzi	0186584
Tuqa Mohammad AbuRaddahi	0187524

May 25th, 2021

Submitted in partial fulfillment of the requirements of B.Sc. Degree in Computer Engineering

This page is intentionally left blank

ETHICAL STATEMENT

We, the undersigned students, certify and confirm that the work submitted in this project report is entirely our own and has not been copied from any other source. Any material that has been used from other sources has been properly cited and acknowledged in the report.

We are fully aware that any copying or improper citation of references / sources used in this report will be considered plagiarism, which is a clear violation of the Code of Ethics of the University of Jordan.

We further certify and confirm that we had no external help without the approval of our supervisor and proper acknowledgment when it is due. We certify and affirm that we never at any point commissioned a 3rd. party to do the work or any part of it on our behalf regardless of the amount charged or lack thereof. We also acknowledge that if suspected and thereafter proven that we commissioned a 3rd. party to do any part of this project that we risk failing the entire project.

We certify and confirm that all results presented in this project are true with no manipulation of data or fraud, that any statistics done, or surveys collected are conducted with the highest degree of scientific fidelity and integrity, and that if proven otherwise, we risk failing the entire project. We acknowledge that for any data collected, we have taken all the steps necessary in applying for proper authorizations if deemed necessary, and that all user data collected is subject to the utmost degrees of privacy and anonymity.

Replace student 1 name here: Yara Al-Jarzi

Replace date here:

Signature:

Replace student 2 name here: Tuqa

AbuRaddahi

Replace date here

Signature:

This page is intentionally left blank

SUPERVISOR CERTIFICATION

I hereby certify that the students in this project have **successfully finished** their senior year project and by submitting this report they have fulfilled in partial the requirements of B.Sc. Degree in Computer Engineering.

☐

I hereby certify that the students in this project have not completed their senior year graduation project and **I do not approve** that they proceed to the discussion.

☐

I suspect that the students have **violated** one or more of the clauses in the **ethical statement** and I suggest that an investigation committee look into the matter.

☐

Prof. / Dr. / Eng. Saadeh Sweadan

Signature:

This page is intentionally left blank

DEDICATION

We would like to express our sincere gratitude to our supervisor, Eng. Saadeh Sweadan, for his exceptional guidance and unwavering support throughout our graduation project. His expertise and dedication have been invaluable in shaping the success of our work. Additionally, we extend our deepest appreciation to our families for their constant love, encouragement, and belief in our abilities. Their unwavering support has been a driving force behind our accomplishments.

We would also like to acknowledge and thank our department professors and doctors, whose expertise and guidance have played a vital role in our personal and professional growth over the past five years. Their mentorship and care have been instrumental in nurturing our development. Finally, we are grateful to all those who have supported us on this journey, including friends, mentors, and colleagues. Your unwavering belief in our capabilities and your support has been truly uplifting.

SYMBOLS, ABBREVIATIONS, AND ACRONYMS

AI	Artificial Intelligence
API	Application Programming Interface
DL	Deep Learning
ERD	Entity Relationship Diagram
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
ML	Machine Learning
NLP	Natural Language Processing
OCR	Optical Character Recognition
PHP	Hypertext Preprocessor
SQL	Structured Query Language
SSML	Speech Synthesis Markup Language
TTS	Text-to-Speech
UI	User Interface
USB	Universal Serial Bus

ABSTRACT

This report presents the development and evaluation of Notoq, a user-friendly mobile application designed to aid visually impaired individuals by converting text to speech. The objective of this project was to create an accessible and intuitive tool that enhances the accessibility and independence of visually impaired users in accessing textual content.

The report begins by identifying the problem of limited accessibility faced by visually impaired individuals and the importance of a text-to-speech application for Arabic text. The development of Notoq involved integrating a responsive and accurate text-to-speech engine, accompanied by a carefully designed user interface for seamless navigation and operation. To evaluate the effectiveness of Notoq, testing was conducted on various Android devices with different versions, ensuring compatibility and functionality. The results showed that Notoq successfully worked on 100% of the tested devices, allowing users to create accounts and utilize all the application features.

A survey among university students who used the Notoq application provided valuable insights into its quality and effectiveness. The survey encompassed 12 questions, assessing user interaction, performance, usability, and usefulness. The statistical analysis of the survey results indicated positive feedback, with most participants finding the application useful for reading Arabic text and appreciating the clarity and accuracy of the audio messages. The user-friendly Graphical User Interface (GUI) received high ratings, and the feature of customizing voice settings was evaluated positively. Additionally, the ability to save speech output as audio files for later listening was widely liked.

The report concludes with a discussion of the accomplishments and contributions of this project. It highlights the successful integration of the Text-to-Speech (TTS) engine, the user-friendly design, and the positive impact of Notoq on accessibility and independence for visually impaired individuals. For future work, several recommendations are provided to enhance the capabilities and reach of Notoq. These include improving the text-to-speech engine's pronunciation accuracy, expanding language support, integrating file-to-speech functionality, implementing voice command capabilities, and exploring integration with other applications for comprehensive accessibility.

In summary, the Notoq project has successfully developed a user-friendly mobile application that addresses the accessibility needs of visually impaired individuals. The positive feedback from users and the potential for further improvements and enhancements highlight the significance of Notoq in improving the quality of life for visually impaired individuals and promoting inclusive technology solutions.

TABLE OF CONTENTS

ETHICAL STATEMENT	III
SUPERVISOR CERTIFICATION.....	V
DEDICATION.....	VII
SYMBOLS, ABBREVIATIONS, AND ACRONYMS	VIII
ABSTRACT.....	X
TABLE OF CONTENTS	XI
LIST OF FIGURES.....	XIII
LIST OF TABLES.....	XIV
INTRODUCTION.....	1
1.1. Problem Definition	1
1.2. Proposed Solution	2
1.3. Project Deliverables	2
1.4. Project Impact	3
1.4.1 Broader Societal Impact	3
1.4.2 Potential Challenges and Limitations	4
RELATED WORK.....	5
2.1 Natural Reader	5
2.2 Narrator's Voice	6
2.3 Speechify	7
SOLUTION DESCRIPTION AND IMPLEMENTATION	9
3.1 Application Theme and UI	9
3.2 Functional requirements.....	10
3.2.1 Notoq services.....	10
3.3 Non-Functional Requirements	12
3.4 Cloud-based AI Tools used in Notoq	13
3.5 Notoq Mobile Application Development	14
3.6 Graphical User Interface (GUI)	16
3.7 Database Implementation.....	17
3.7.1 MySQL Workbench 8.0 CE.....	17
3.7.2 Implementation	18
3.7.3 Entity Relationship Diagram (ERD).....	21
3.8 OCR and TTS Performance	22
3.8.1 OCR with different types of scripts	22

3.8.2 OCR with different text colors	25
3.8.3 Text to speech API performance.....	26
RESULTS AND DISCUSSION.....	27
4.1 Scenario.....	27
4.2 Testing.....	39
4.3 Survey.....	40
CONCLUSIONS AND FUTURE WORK	43
5.1 Conclusions.....	43
5.2 Future Work.....	43
APPENDICES	46
APPENDIX A: Application Layout and Features Overview	46
APPENDIX B: Sample Sound Test Records.....	53

LIST OF FIGURES

FIGURE 1 - NUMBER OF ARABS AROUND THE WORLD [1]	1
FIGURE 2 - HIGH AND LOW CONTRAST COLORS [3]	3
FIGURE 3 – NATURALREADER HOME PAGE [4]	5
FIGURE 4 – NARRATOR’S VOICE HOME PAGE [5]	6
FIGURE 5 – SPEECHIFY HOME PAGE [6]	7
FIGURE 6 – NOTOQ LOGO	9
FIGURE 7 – MAIN PAGE OF NOTOQ	13
FIGURE 8 – SOME OF NOTOQ INTERFACES	17
FIGURE 9 – USER ACCOUNT TABLE	19
FIGURE 10 – FOLDERS TABLE	19
FIGURE 11 – FILES TABLE	21
FIGURE 12 – ERD OF NOTOQ	21
FIGURE 13 – THE KUFIC SCRIPT	22
FIGURE 14 – NASKH SCRIPT	23
FIGURE 15 – RUQ’AH SCRIPT	23
FIGURE 16 – HANDWRITTEN SCRIPT 1	23
FIGURE 17 – HANDWRITTEN SCRIPT 2	24
FIGURE 18 – THULUTH SCRIPT	24
FIGURE 19 – DIWANI SCRIPT	24
FIGURE 20 – LOW CONTRAST TEXT COLOR	25
FIGURE 21 – HIGH CONTRAST TEXT COLOR	25
FIGURE 22 - SIGN UP OPERATION	27
FIGURE 23 – THE WELCOME PAGE	28
FIGURE 24 - HOME PAGE	29
FIGURE 25 - MY FOLDERS	30
FIGURE 26 - INSIDE PAGES FOLDER	30
FIGURE 27 - PASTE TEXT PAGE	31
FIGURE 28 - SCAN PAGES PREVIEW	32
FIGURE 29 - SCAN PAGES FEATURE	33
FIGURE 30 - UPLOAD PHOTO FEATURE	34
FIGURE 31 - SPEECH TO TEXT FEATURE	35
FIGURE 32 - PASTE TEXT FEATURE WITH ENGLISH TEXT	36
FIGURE 33 - SETTINGS TAB	37
FIGURE 34 - PERSONAL INFO	38
FIGURE 35 - INTERFACE LANGUAGE	39
FIGURE 36 - MAIN AND SIGN-UP PAGES	46
FIGURE 37 - LOGIN OPERATION	47
FIGURE 38 - IMAGE UPLOAD	48
FIGURE 39 - SPEECH-TO-TEXT FEATURE	49
FIGURE 40 – SETTINGS	49
FIGURE 41 - PERSONAL INFORMATION	50
FIGURE 42 - CHANGE PASSWORD AND BIO	51
FIGURE 43 - LANGUAGE TAB	52

LIST OF TABLES

TABLE 1– RELATED WORK SUMMARY	8
TABLE 2- USER ACCOUNT TABLE.....	18
TABLE 3- FOLDERS TABLE.....	19
TABLE 4- FILES TABLE.....	20
TABLE 5- TESTING NOTOQ ON DIFFERENT DEVICES.....	39
TABLE 6 - SURVEY AND STATISTICAL RESULTS	40

This page is intentionally left blank

INTRODUCTION

The ability to access and read digital content is an essential part of modern life. However, vision-impaired individuals face significant challenges when using mobile devices to read text. While many Text-to-Speech applications exist, few of them support Arabic text, leaving a significant gap in the market. We aim to address this gap to help them overcome these challenges and feel more comfortable using an Arabic TTS application (Notoq) that will read Arabic texts aloud for them without even having to ask someone for help. Not only the visually impaired but also anyone can use it, for example, if they have received a long message and would like to listen to it than reading it with just a few clicks.

1.1. Problem Definition

Vision-impaired people face a lot of difficulties while keeping pace with technology development, such as exchanging messages, reading emails, books, reading from images, and investigating the internet. As the number of these people increases it becomes a must to facilitate their life and their smartphone accessibility. Statistics have shown that the number of Arabs is increasing around the world. Figure 1 below shows the number of Arabs around the world from 2011 to 2021 [1]. This means indeed that Arabs with vision impairment are increasing too.

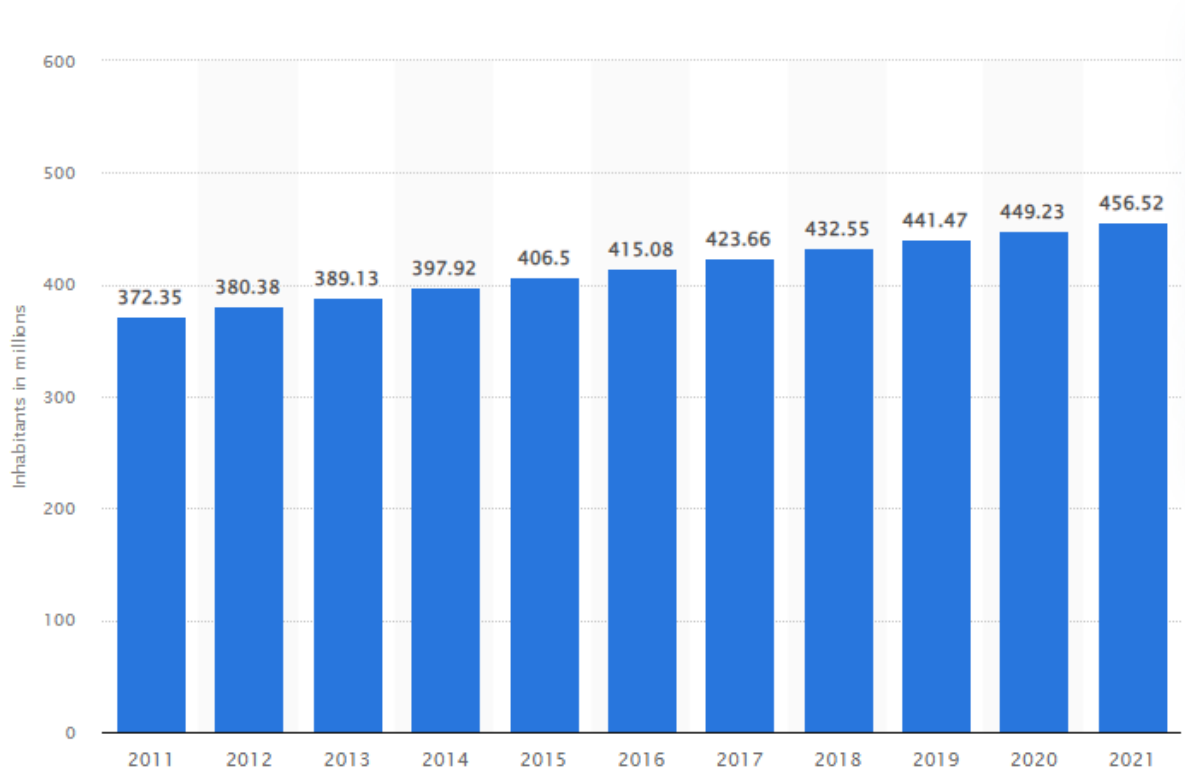


Figure 1 - Number of Arabs Around the World [1]

Not only for Arabs, but around the world there are over 2.2 billion people who are blind or visually impaired and the number is increasing [2], especially in poor countries. As long as the technology improves, it's a must to benefit of modern accessibility features like Text-to-Speech to allow these people to be more independent. One of the specific challenges that vision-impaired individuals face when using mobile devices to read text is the inability to distinguish between different colors and fonts. For example, if an app or website uses low contrast colors or small font sizes, it may be difficult for a visually impaired person to read the text. This can lead to frustration and a feeling of exclusion from the digital world.

Vision-impaired individuals may also struggle with accessing certain apps or websites. For example, if an app or website does not have proper accessibility features such as screen reader support or larger font sizes, it may be impossible for a visually impaired person to use it. This can limit their access to important information and services through the internet.

The impact of these challenges on the daily lives of vision-impaired individuals can be significant. For example, they may have difficulty reading important emails from their employer or completing work assignments that require reading and writing. They may also have difficulty staying connected with friends and family if they are unable to read text messages or social media updates. These challenges can lead to a feeling of isolation and dependence on others for assistance.

1.2. Proposed Solution

The proposed solution is a mobile application that uses cloud-based Artificial Intelligence (AI), Deep Learning (DL), and Natural Language Processing (NLP) tools, the first one for text detection, extraction, and processing and the second tool for converting the text to speech audio. A cloud-based storage bucket was used to store generated audios. Also, Speech to Text AI tool was used to extract the transcript of the given audio. The application can read Arabic text from a normal text pasted by the user, also it can detect and extract the text from an image and do the processing steps on it, even on images scanned using the camera. The application will provide a user-friendly and simple interface, makes it easy for any user to use the application with a smaller number of clicks and pages to navigate through, also it allows the user to control the reading speed and even download the audio generated on their phone if they wish to listen to it later offline or send it to their friends.

1.3. Project Deliverables

This project is a response to the challenges that vision-impaired individuals face when using mobile devices to read text. Our goal is to develop a mobile application that can convert Arabic text into read-aloud speech, making it easier for users to access and comprehend digital content. Our project aims to achieve several objectives, including developing a fully functional mobile application that can detect and read Arabic text. We also prioritize creating an accessible and friendly user interface (UI) that can be easily navigated by all users.

Our project aims to help vision-impaired people use their phones independently, giving them greater autonomy in their daily lives. Additionally, we aim to save time and effort for all users by allowing them to listen to any long message or scan any page and listen to it while driving their cars, exercising, or performing other tasks. This feature is not limited to vision-impaired individuals but can also benefit anyone who prefers to listen to text instead of reading it.

Furthermore, our application includes a unique feature that can extract text from images, making it easier to access information from sources that may not be compatible with traditional text-to-speech applications. For example, some images may have high light resistant colors or low contrast colors that make it difficult to read the text inside them. Our application can extract the text from these images and convert it into speech, allowing users to access this information with ease. Figure 2 illustrates the difference between high contrast and low contrast colors.



Figure 2 - High and Low Contrast Colors [3]

Finally, our application can also enhance reading skills for non-Arabian people who are interested in learning Arabic. By using our application with its English UI, users can listen to Arabic texts and improve their reading comprehension skills. In summary, our project aims to provide a useful, accessible, and user-friendly mobile application that can benefit a wide range of users.

1.4. Project Impact

1.4.1 Broader Societal Impact

The impact of the Notoq application goes beyond its ability to assist vision-impaired individuals. By promoting inclusivity and accessibility in the tech industry, the application

contributes to a more equitable and just society. Accessible technology allows individuals with disabilities to participate more fully in society, breaking down the barriers that have historically excluded them from important activities such as education, work, and social interactions.

Moreover, Notoq has the potential to help bridge language barriers between different cultures and communities. With the ability to read Arabic text aloud, non-Arabic speakers can improve their language skills and expand their understanding of Arabic culture. This can lead to increased cross-cultural communication and understanding, which can ultimately help to reduce prejudice and promote tolerance.

1.4.2 Potential Challenges and Limitations

Despite the benefits of the Notoq application, there are potential challenges and limitations to consider. One challenge is the accuracy and speed of the Text-to-Speech conversion process. While the application uses advanced AI and NLP technologies, there may be cases where the conversion is not 100% accurate or takes longer than expected. This could be especially challenging for users with time-sensitive needs, such as reading important emails or documents.

Another potential limitation is privacy and security concerns with the storage and processing of user data. The application requires access to the user's camera and storage to read images and scan pages, which raises questions about data privacy and security. To address these concerns, the application must implement robust security measures to protect user data and ensure user privacy.

Additionally, the application may require internet connectivity for some features, such as cloud-based text processing and storage. This could be challenging for users with limited access to the internet or for users who are traveling to areas with limited internet connectivity. In these cases, the application may need to provide offline functionalities or alternative methods of accessing Text-to-Speech conversion tools.

RELATED WORK

Assistive technology is an important field that aims to help individuals with disabilities to perform everyday activities with ease. One area of assistive technology that has been gaining increasing attention is the development of apps that assist visually impaired users such as Text to Speech applications.

2.1 Natural Reader

To begin using Natural Reader, users are required to create an account, enabling them to access the application's functionalities. The application offers various methods for adding text to the user's library, which can be read aloud. These methods include uploading documents, pasting text, scanning documents using the device's camera, uploading photos, and even processing webpages. The user interface of Natural Reader is shown in Figure 3.

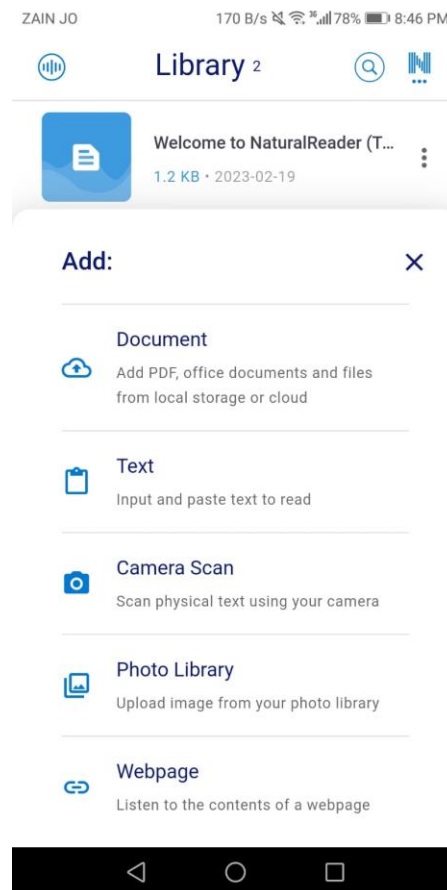


Figure 3 – NaturalReader Home Page [4]

The application boasts several noteworthy features, including the ability to read text in multiple languages and accents with clarity. However, it should be noted that Arabic language support is not available in Natural Reader, as the application is primarily designed for

languages other than Arabic. This limitation highlights the significance of Notoq, which is specifically built for Arabic text processing.

Another distinguishing aspect of Natural Reader is its capacity to extract handwritten texts. However, in contrast to Notoq, where users can upload and scan handwritten texts for audio playback, Natural Reader lacks this functionality. Furthermore, Natural Reader allows users to upload various file types. In contrast, Notoq focuses exclusively on image uploads.

In terms of controlling the speed of reading, Natural Reader offers a range of speeds within the $[-4, 9]$ range. On the other hand, Notoq provides users with precise control over the playback speed of audio, offering options such as 0.5x, 0.75x, 1.0x, 1.25x, 1.5x, and 2.0x.

2.2 Narrator's Voice

Another notable application in the realm of assistive technology is Narrator's Voice, which specializes in generating audio recordings that closely resemble a professional narrator reading text aloud. The home screen of Narrator's Voice is shown in Figure 4.

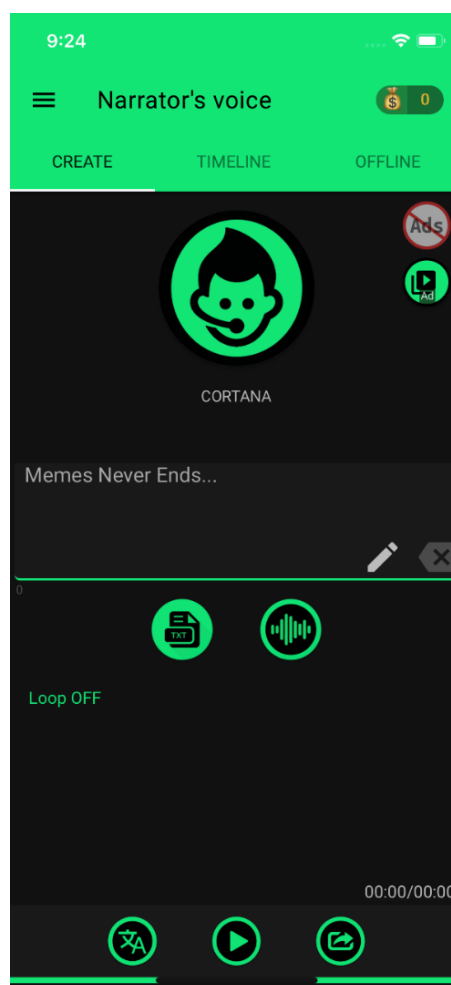


Figure 4 – Narrator's Voice Home Page [5]

The application offers many features such as:

1. **Language Support:** Narrator's Voice supports multiple languages, including Arabic, and provides various accents for each language. However, it should be noted that the application may struggle with accurately differentiating between Arabic letters. In contrast, Notoq excels in processing Arabic text with precision and offers the option to choose between male and female voices for the text reading.
2. **Voice Effects:** Narrator's Voice offers an array of voice effects to enhance the audio experience. These effects add a touch of creativity and variation to the narrated text.

2.3 Speechify

Speechify is an application specifically designed to enhance the listening experience of digital content for users. With features such as content importation, adjustable playback speed, and voice options, speechify aims to optimize the efficiency of consuming digital texts. The home screen and key features of the application are shown in Figure 5 below:

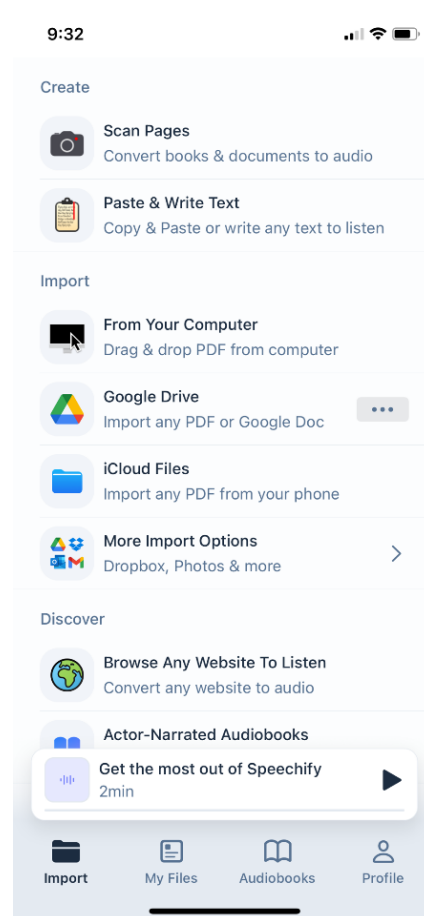


Figure 5 – Speechify Home Page [6]

The features of the application are:

1. Different languages with different accents and voices including Arabic.
2. Control the speed of reading. Notoq Also have this feature as mentioned above in section 2.1.
3. Can browse any website to listen.

In conclusion, a comprehensive analysis of our application and related work reveals several distinguishing factors, as summarized in Table 1:

Table 1– Related Work Summary

Features	Notoq (Our application)	Natural Reader	Narrator's Voice	Speechify
Arabic Language	Yes	No	Yes	Yes
Playback Speed Control	Yes	Yes	Yes	Yes
Handwritten text detection	Yes	Yes	No	No
Present different voices	Yes	Yes	Yes	Yes
Accepted file types	different types of images	All types of files	All types of files	All types of files

SOLUTION DESCRIPTION AND IMPLEMENTATION

This chapter will cover the application theme, UI, functional and non-functional requirements that are available in our Arabic TTS. Also, it will cover all the details about the features, design, and implementation of our application.

3.1 Application Theme and UI

Our project, "Notoq", is designed to provide users with an efficient and user-friendly platform for accessing information. To ensure that our project meets the needs of our users, we conducted extensive research to identify common questions and concerns that individuals may have when seeking information. Based on this research, we developed a set of services that our application would provide. We believe that this approach will allow our users to quickly and easily find the information that they need.

To build our Arabic TTS application in a way that meets our aimed people goals, we did some search on the combinations of colors and sizes of fonts to use in our application theme and UI to make it easier and accessible for visually impaired people who are going to use our application, so we need a high contrast combination of colors because High contrast color schemes allow users to be able to see information clearly and to read without putting additional strain on their eyes [7]. As a result, we ended up using a combination of dark blue (0xFF090088) and light yellow (0xFFFF1F7B5).

As the idea of our application is about reading text aloud. We have chosen an appropriate and representative name which is Notoq "نُطْق" written in Arabic bold, modern letters, and we choose the shaded eye symbol to sign to vision-impaired people (low vision in particular) We ended up with a logo that is simple, yet memorable, and reflects the values of our project, accompanied. As shown below in Figure 6.



Figure 6 – Notoq Logo

3.2 Functional requirements

Our application is specifically designed to cater to the needs of the visually impaired, with a particular focus on individuals with low vision. To ensure optimal efficiency and functionality, the application incorporates a range of carefully crafted tasks, prioritizing precision, and user-friendliness. These tasks are made possible through the implementation of our Arabic Text-to-Speech (TTS) technology:

1. **Account Registration:** As a fundamental step, users are required to create an account by providing essential details such as their email address, username, and password. This process is crucial for personalizing their experience and ensuring secure access to the application's features and services.
2. **User Authentication:** Once the account has been successfully set up, users can easily log in using the credentials established during the registration process. User authentication is a vital aspect of our application, guaranteeing that only authorized individuals can access their personalized content and settings.
3. **User Session Termination:** To provide users with complete control over their sessions, our application incorporates a straightforward 'Log Out' button. This feature allows users to securely terminate their session whenever they choose, ensuring the privacy and security of their personal information.
4. **Language Switch:** Recognizing the diversity of our user base, our application offers a bilingual interface. Users can effortlessly switch between Arabic and English languages by clicking on the language icon located within the settings tab. This language switch feature enhances the accessibility and user-friendliness of our application, accommodating both Arabic speakers and non-Arabic speakers who wish to utilize our services.

3.2.1 Notoq services

Notoq offers a variety of services designed to enhance the accessibility and usability of text content for visually impaired individuals. These services are aimed at empowering users to conveniently access and interact with textual information. The key services provided by Notoq include:

- **Paste Text:** With the "Paste Text" feature, users can copy any Arabic text they are unable to read and directly paste it into the application. Once pasted, users have the option to choose the gender of the audio reader, with the default setting being male. The text is then converted into audio, allowing users to conveniently listen to the content and gain access to the information they need. For this service, users can paste English text too.

- **Upload Photo:** Notoq enables users to upload images that contain text, regardless of readability issues caused by factors such as image colors or handwritten text. When an image is uploaded to the application, Notoq takes charge of extracting the text from the image. The extracted text is then converted into speech, delivering an audio to the user in less than a minute. This feature ensures efficient access to textual information contained within images, enhancing the overall user experience.
- **Scan pages:** With the "Scan Pages" feature, users can utilize their device's camera to capture images of pages containing text. By simply pressing the upload button, the scanned image undergoes the same process as uploading an image. Notoq extracts the text from the scanned image, converting it into audio format for easy comprehension and access.
- **Create Folder:** Notoq understands the importance of organization and accessibility. Therefore, the application allows users to create multiple folders where they can upload images. These folders serve as a repository for organizing and categorizing content, enabling users to easily locate and access the desired information without having to search through their entire gallery.
- **Speech to Text:** In addition to its core functionality, Notoq offers a unique "Speech to Text" feature. This feature caters to users who wish to transcribe audio content into written text. Users can upload audio files in .wav format, and Notoq utilizes specialized transcription tools to convert the audio into text. The transcribed text is then displayed on the application's screen, enabling users to conveniently access and utilize the transcribed content.
- **Settings:** The settings tab in Notoq provides users with a range of customization options and controls. Users can effectively manage their folders, update personal details including profile photo, password, and bio. Additionally, the language of the application interface can be easily switched between Arabic and English using the Language button. For added convenience, a "Log Out" button is conveniently located within the settings tab, allowing users to securely end their session and maintain privacy.

These services offered by Notoq are designed to cater to the specific needs of visually impaired individuals, providing them with a comprehensive and user-friendly platform to access and engage with textual content more effectively.

3.3 Non-Functional Requirements

Every system has some important non-functional requirements that are needed to achieve the system goals. Here are the most important non-functional requirements in Notoq:

- **Accessibility :**

The accessibility of an application refers to its ability to be easily used and interacted with by individuals with disabilities. In the case of Notoq, our focus is on providing accessibility to visually impaired individuals, particularly those with low vision. To ensure accessibility, our application is designed to be compatible with Android smartphones, leveraging the capabilities of mobile devices.

By choosing to develop our application for Android operating system, we enable visually impaired individuals to access and utilize Notoq on their smartphones, allowing users to use the application in various environments and situations, and assisting them in completing their daily tasks more efficiently.

The use of mobile devices, combined with the functionalities of Notoq, enhances the communication capabilities and independence of visually impaired individuals. It provides them with a tool that can assist them in accessing and interacting with textual information, facilitating their participation in different activities and tasks.

- **Ease of use:**

As mentioned many times before, Notoq is designed for vision impaired people, so we take the responsibility of choosing the best design that fits all, the fonts sizes, colors, buttons, and pages used in Notoq were chosen and designed carefully after a long search to make it more simple and easy to use, as they don't have many and complex pages to navigate through and some buttons are enough to achieve the purpose. Figure 7 below shows the main screen of our page that contains all the services mentioned in section 3.2 , it shows how simple our application is and that achieves our goal of ease of use.

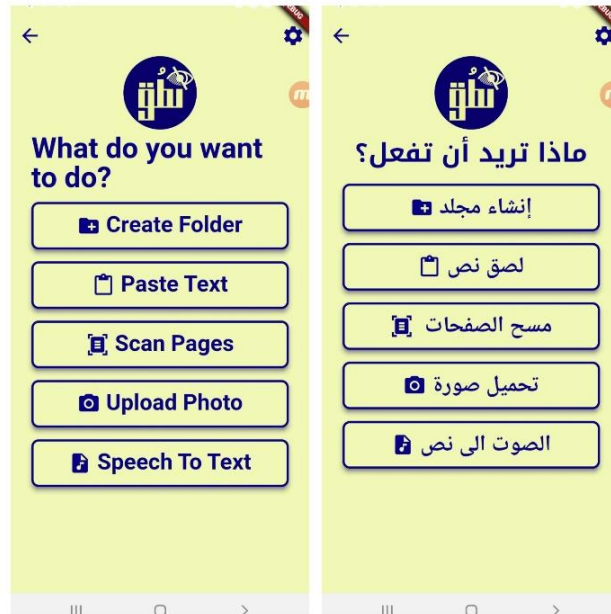


Figure 7 – Main Page of Notoq

3.4 Cloud-based AI Tools used in Notoq

In the development of Notoq, we utilized several cloud-based AI tools provided by Google Cloud to construct our application's core functionality. The tools we used are as follows:

- **Text-to-Speech:** This cloud-based Application Programming Interface (API) service enables the conversion of text into lifelike, natural-sounding speech. Utilizing advanced natural language processing and machine learning techniques, it synthesizes speech in various voices and languages, creating an auditory experience that closely mimics human speech. The service is designed to handle a variety of input formats such as plain text, Speech Synthesis Markup Language (SSML), and Text normalization markup. It's important to note that Google Text-to-Speech is a paid API, with cost structures based on the total number of characters processed.
- **Google Vision AI :** This is a versatile, machine learning-powered image analysis service offered by Google Cloud Platform. It utilizes advanced machine learning algorithms to extract text from images and perform optical character recognition. Its numerous applications include document processing, image classification, face detection, logo detection, and explicit content detection. For our purposes in Notoq, we primarily utilized the document processing feature, focusing specifically on Optical Character Recognition (OCR) and text detection in images and scanned pages. This was facilitated by Vision AI's pre-trained models. An impressive aspect of Vision AI is its capability to identify text in various formats like detecting text in images. It's also a paid tool with a different price for each feature per unit (unit like image). The images

that the user uploads will be sent to this tool for text extraction then the text extracted will be sent to the Text to speech tool.

- **Google storage bucket:** the bucket is like a container that holds the user's data, each bucket will have a unique name and a geographic location, the price differs depending on the location and the class of the bucket storage. In Notoq we used two buckets, one for storing the audios that were generated by the Text to speech tool before sending the audio to the user, and the second one for storing the images the user uploads before sending the image to the OCR tool also the second bucket is used to store the images that the users upload in the folders they create.
- **Speech to Text API:** this tool can transcribe the audio content, it supports wide range of audio extensions but .wav is the most supported directly so in Notoq only .wav audios are supported to upload and once the user uploads the audio, it will be stored in the bucket which was created for audios then a Uniform Resource Locator (URL) of the audio in the bucket will be generated and sent to the Speech to Text tool to extract transcription. its price differs based on the amount of audio processed by the API tool each month.

These cloud-based AI tools not only contributed significantly to the development of Notoq but also enhanced its functionality, providing a seamless and intuitive user experience.

3.5 Notoq Mobile Application Development

- **Flutter:**

Flutter, a creation of Google, is an open-source UI software development kit designed to facilitate the development of applications across various platforms, including Android, iOS, Linux, macOS, Windows, and the web, all from a single codebase. Its core programming language, Dart, is a high-level language that's relatively straightforward to learn and use. We've chosen Flutter as the backbone for our app development due to several reasons. Primarily, our target is to build an Android application, a task that Flutter seamlessly handles. Additionally, the simplicity of learning and using Flutter, compatibility with Android, ease of learning, coupled with its hot reloading feature and numerous other advantages, making it an optimal choice and the easiest for Notoq Frontend development.

- **Android Studio:**

Android Studio is the official IDE designed specifically for Android application development and has several features that enable developers to seamlessly build their applications. It was announced at the 2013 Google I/O conference and serves as a replacement

for the Eclipse Android Development Tools (E-ADT). This IDE can be run on Windows, macOS and Linux-based operating systems [8]. We relied on Android platform to represent Notoq in the most accurate way of making this project goal not only achievable but also reachable for our target users.

Android studio has many features that served us in building Notoq, such as:

- 1- **Live edit:** the developer doesn't have to restart the application to visualize each change, it will be made instantaneously.
- 2- **Android device emulator:** this is like a virtual device, allows the developer to test the application on different android devices such as phones and tablets, but for this to work properly it need high requirements in the developing device, like core i7 and 16G RAM and more, otherwise it will be deadly working as happened with us, we couldn't test our application on the emulator.
- 3- **Instant run:** the instant run button, if you make a small update on your code then a test is surely needed, due to this button the code will be updated on your device. Leading to save time. Hence, saving time is a major important point for all programmers when building any type of project.
- 4- **Capability of testing the application on a real device:** as a developer of course, you want to test your application and see how it looks like on a real mobile phone. Using android studio, you can connect your device and test the application on it. Two ways to do that, the first one is Universal Serial Bus (USB) debugging and the second one is by wireless debugging. We relied a lot on USB debugging while working on our application as it was better to test the application on a real device than emulator or web.
- 5- **Robust and flexible system build:** Android Studio makes an automation in dependency management, you just need to write the name of the library you want to use, and it will do the linking process on its own.

- **XAMPP and Hypertext Preprocessor (PHP):**

XAMPP is a completely free, easy to install Apache distribution containing MariaDB, PHP, and Perl. The XAMPP open-source package has been set up to be incredibly easy to install and to use [9]. The name "XAMPP" stands for Cross-Platform (X), Apache (A), MySQL (M), PHP (P), and Perl (P). It is available for Windows, macOS, Linux, and Solaris operating systems.

Here are the key components of XAMPP:

1. **Apache Hypertext Transfer Protocol (HTTP) Server:** XAMPP includes the Apache web server, which is one of the most widely used web servers globally. Apache

provides the infrastructure for hosting websites and handling HTTP requests. And we used it to test our application by starting the Apache server and start to send HTTP requests from our flutter frontend code. Which is about the features like when sending the text from the flutter application user interface to the server into the PHP backend code and then to Google cloud TTS.

2. **MySQL:** XAMPP bundles the MySQL database management system. MySQL is a popular relational database management system that allows you to store and manage data for your web applications. As we already work on MySQL workbench, we just started this option. This choice allowed us to efficiently handle database-related tasks and effectively manage the data within our project.
3. **PHP:** XAMPP includes PHP, a popular server-side scripting language used for web development. PHP enables dynamic content generation and interacts with databases, making it essential for building dynamic web applications. During the development of Notoq, we primarily utilized PHP as our backend code. This choice was driven by the language's user-friendly nature and its compatibility with Google Cloud features, making it effortless to integrate and contribute to the project. PHP's simplicity and wide adoption within the web development community made it an ideal choice for implementing the backend functionality of Notoq and leveraging the advanced capabilities provided by Google Cloud. And since Google Cloud offers a range of services and features that can be seamlessly integrated with PHP backend code. PHP was the best choice for Notoq backend development. By loading the service account key into the code, we can have wide range access to almost all google cloud services.
4. **Perl:** XAMPP also includes Perl, a versatile programming language used for various purposes, including web development, system administration, and network programming.

3.6 Graphical User Interface (GUI)

Graphical User Interface refers to the visual components of a software application that allows users to interact with it using graphical elements such as buttons, menus, icons, and windows, instead of using text-based commands. The pages (layout) of Notoq were designed to be as simple as possible to achieve the goal of our application for the target users. As mentioned before, the choice of the colors, button sizes, font sizes were fully analyzed and considered to make the application suitable for low vision people. Notoq interfaces were simulated via real smart phone with android operating system by connecting the phone to android studio through USB debugging. Figure 8 below shows some of the interfaces of Notoq, as they will be explained in detail in chapter 4.



Figure 8 – Some of Notoq Interfaces

3.7 Database Implementation

A database is an organized collection of structured information, or data, typically stored electronically in a computer system [10]. In our application Notoq, we used MySQL workbench to build our database.

3.7.1 MySQL Workbench 8.0 CE

It's a visual management tool for databases, it provides building and modeling databases, Structured Query Language (SQL) development and server configuration with administration tools, and a lot more. As our database is relational (which means that data are related to each other) we used SQL to build it, SQL is a programming language used by nearly all relational databases to query, manipulate, and define data, and to provide access control[11].

3.7.2 Implementation

Three tables were implemented to achieve the requirement of our application for storing the data of (users, folders, and files). Each table has many attributes. Our database is relational as we mentioned previously because there are relations between the tables like, each user creates many folders and each folder can contain many files.

We will have the tables and their attributes described below:

- 1- User account table: Each user can create an account using username, email, and password. Table 2 below shows the user account table with its attributes.

Table 2- User Account Table

Attribute	Type
user_id	int
user_name	varchar (255)
user_email	varchar (255)
user_password	varchar (255)
user_photo	varchar (255)
user_bio	text
member_since	datetime

- user_id: this attribute is the primary key for this table (each user has a unique id that can't be repeated)
- user_name: this attribute stores the name of the user.
- user_email: this attribute stores the email of the user.
- user_password: this attribute stores the password of the user but as hashed password for security.
- user_photo: this attribute stores the name of the users uploaded photo.
- user_bio: this attribute stores the bio if the user would like to have one for their account.
- member_since: this attribute to store the date and time when the user created the account and became a member, it uses current time stamp by default and stores it in the database as datetime.

Figure 9 shows the user account table in the database with its attributes and how the data is stored in each attribute.

user_id	user_name	user_email	user_password	user_photo	user_bio	member_since
1	Yara AlJarzi	yara123@gmail.com	\$2y\$10\$xmV8stvWGQw.cEiY7IAOuPtgmpZ60rF...	NULL	computer engineer	2023-05-20 15:34:2
3	Yara AlJarzi	yarahasan@gmail.com	\$2y\$10\$Xl/8vDtKhPTfIAVpiThrt.gI6xnJf5TFgde...	NULL	NULL	2023-05-25 21:29:2
4	Yara Hasan	yaraaljarzi@gmail.com	\$2y\$10\$BaFhFG/05BWbtoNZmULR7eAxSosY5r...	646faac7c498e.png	Computer Engineer	2023-05-25 21:32:1
5	Yara	yarahasan1@gmail.com	\$2y\$10\$EQEpOy7AED2tv1eZLe/WY0twyPffN2...	NULL	NULL	2023-05-25 21:41:4
6	Yara	yarahasan2@gmail.com	\$2y\$10\$DOL2wrB/1x.oTZ9OzbCa0e.zKt6e3d77...	646fade755904.png	Computer engineer	2023-05-25 21:45:1
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Figure 9 – User Account Table

2- Folders Table:

Each user can have many folders (one-to-many) and inside each folder they can upload many images. Table 3 shows Folders table with its attributes.

Table 3- Folders Table

Attribute	Type
folder_id	int
user_id	int
folder_name	varchar (255)
date_created	datetime

- folder_id: this attribute is the primary key for this table (each folder has a unique id that can't be repeated).
- user_id: this attribute will link the folder table with the user account table (foreign key).
- folder_name: this attribute will store the folder name and must not be null.
- date_created: this attribute will store the date on which the user created the folder and must not be null.

Figure 10 below shows the folders table in the database with its attribute and how the data is stored in each attribute.

folder_id	user_id	folder_name	date_created
1	1	First	2023-05-20 14:35:17
2	3	First Folder	2023-05-25 20:30:30
3	4	First Folder	2023-05-25 20:33:12
4	5	My images	2023-05-25 20:42:44
5	6	Folder1	2023-05-25 20:46:09
NULL	NULL	NULL	NULL

Figure 10 – Folders Table

3- Files Table:

Each user can upload many image files into the folders they create to keep the images and preview them at any time they want in an easy way rather than looking for them in gallery. Table 4 shows Files table with its attributes.

Table 4- Files Table

Attribute	Type
file_id	int
folder_id	int
file_name	varchar (255)
extension	varchar (255)
file_size	bigint
upload_date	Datetime
file_url	varchar(1000)

- file_id: this attribute is the primary key for this table (each folder has a unique id that can't be repeated).
- folder_id: this attribute will link the files table with the folder account table (foreign key) and by default, it will be linked with the user account table.
- file_name: this attribute will store the file name.
- extension: this attribute will store the extension of the file which the user uploads.
- file_size: this attribute will store the size of the uploaded file.
- upload_date: this attribute will store the date in which the user uploaded that file.
- file_url: this attribute will store the URL of the image that the user uploads stored in google cloud bucket.

Figure 11 shows the files table and its attributes and how the data is stored in each attribute.

	file_id	folder_id	file_name	extension	file_size	upload_date	file_url
1	1	1	IMG_20230506_122959.jpg	jpg	18947	2023-05-20 15:35:29	https://storage.googleapis.com/images_ocr_tts..
2	2	2	received_515355817337171.jpg	jpg	73196	2023-05-25 21:30:47	https://storage.googleapis.com/images_ocr_tts..
3	2	2	IMG_20230506_122959.jpg	jpg	18947	2023-05-25 21:30:55	https://storage.googleapis.com/images_ocr_tts..
4	3	3	received_515355817337171.jpg	jpg	73196	2023-05-25 21:33:19	https://storage.googleapis.com/images_ocr_tts..
5	3	3	IMG_20230506_122959.jpg	jpg	18947	2023-05-25 21:33:24	https://storage.googleapis.com/images_ocr_tts..
6	4	4	received_515355817337171.jpg	jpg	73196	2023-05-25 21:42:52	https://storage.googleapis.com/images_ocr_tts..
7	4	4	IMG_20230506_122959.jpg	jpg	18947	2023-05-25 21:42:57	https://storage.googleapis.com/images_ocr_tts..
8	5	5	received_515355817337171.jpg	jpg	73196	2023-05-25 21:46:16	https://storage.googleapis.com/images_ocr_tts..
9	5	5	IMG_20230506_122959.jpg	jpg	18947	2023-05-25 21:46:21	https://storage.googleapis.com/images_ocr_tts..

Figure 11 – Files Table

3.7.3 Entity Relationship Diagram (ERD)

The ERD is a diagram that describes how entities inside the database are linked and it shows the logical structure of the database. Figure 12 Below shows the ERD of Notoq.

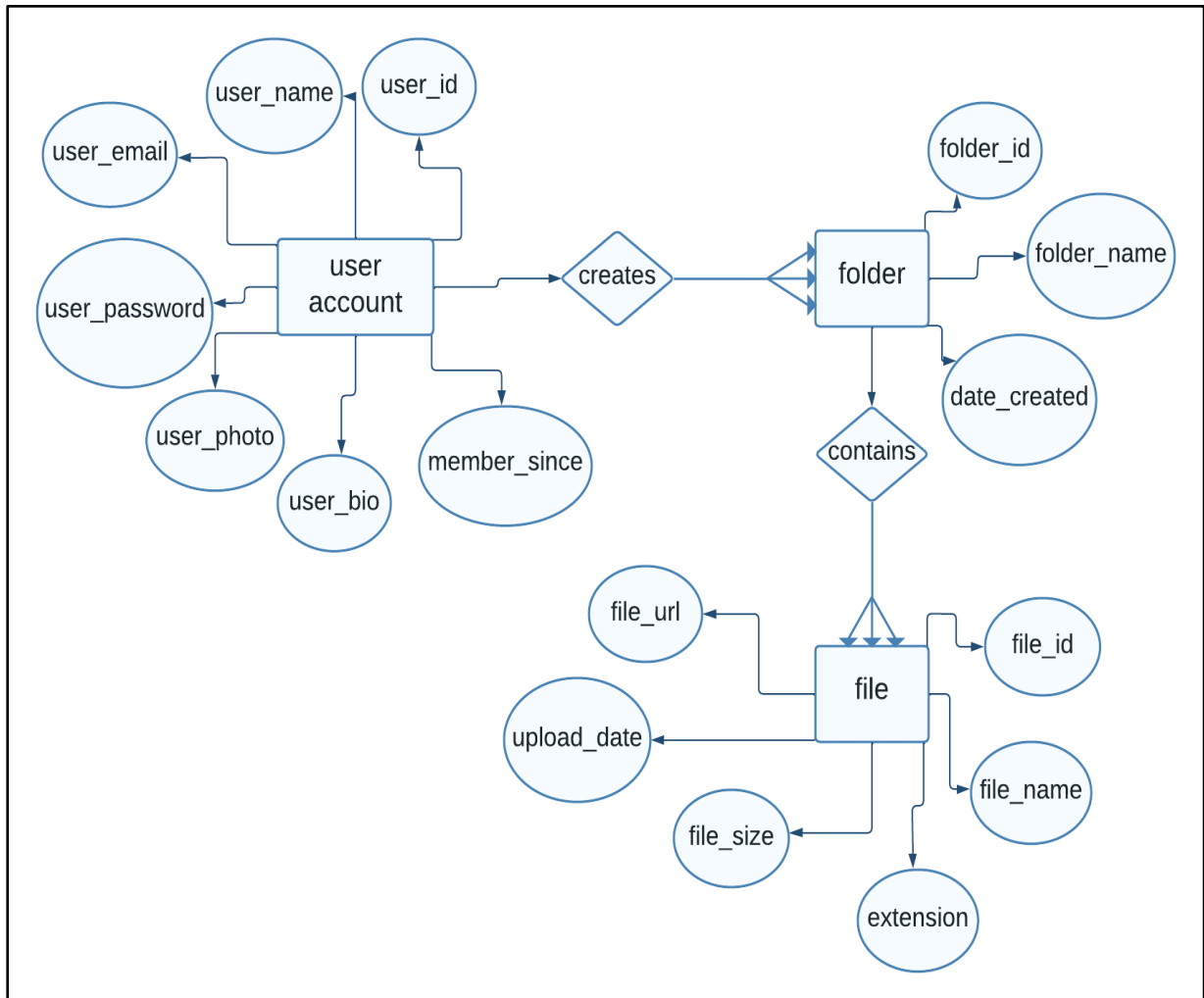


Figure 12 – ERD of Notoq

3.8 OCR and TTS Performance

In this study, we evaluated the performance of Google Cloud Vision - Optical Character Recognition (OCR) and Text-to-Speech (TTS) on Arabic fonts, handwriting, and different color contrasts.

3.8.1 OCR with different types of scripts

During the testing phase, the OCR functionality was subjected to some evaluation, encompassing a diverse range of Arabic fonts and handwritten images. The evaluation process aimed to assess the accuracy and effectiveness of the OCR tool in various scenarios, ensuring optimal performance and reliable text extraction, it was tested on different types of Arabic fonts and handwritten images:

- The Kufic script (الخط الكوفي):

Figure 13 shows sample of Kufic script, an instance where the OCR system encountered challenges in accurately detecting the text. In this case, the detected text "بِإِذْنِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ" exhibited some inaccuracies, resulting in an incorrect extraction. The intricacies of the Kufic script, with its angular and complex letterforms, proved to be a demanding challenge for the OCR algorithm.

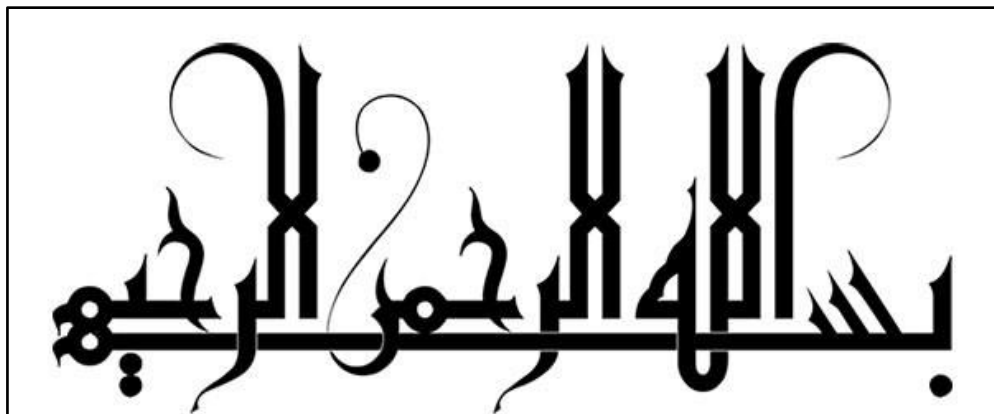


Figure 13 – The Kufic Script

- The Naskh script (خط النسخ):

Figure 14 showcases a sample written within Naskh script, which is known for its rounded and flowing letterforms. In this case, the OCR system successfully detected most of the text as "الله الرحمن الرحيم". However, it should be noted that the complete text was not correctly identified, indicating some limitations in accurately recognizing all elements of the Naskh script.

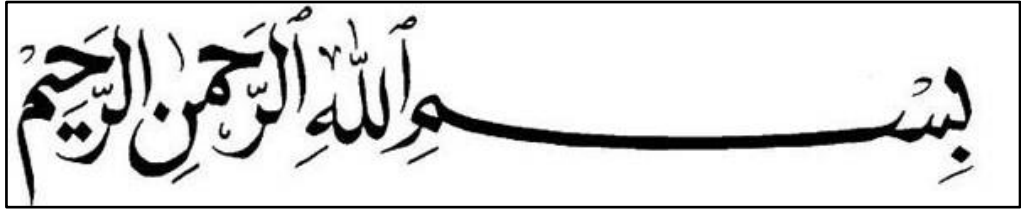


Figure 14 – Naskh Script

- The Ruq'ah script (خط الرقعة):

Figure 15 represents the Ruq'ah script, characterized by its bold and compact letterforms. The OCR tool performed admirably in this scenario, accurately detecting, and extracting the text as "بسم الله الرحمن الرحيم," showcasing its proficiency in handling this script type.



Figure 15 – Ruq'ah script

- Handwritten scripts:

The OCR tool was also put to the test with handwritten scripts, which often present unique challenges due to variations in handwriting styles. Figure 16 and Figure 17 display examples of handwritten text. In both cases, the OCR tool showcased its capabilities by accurately detecting the texts "بسم الله الرحمن الرحيم" and "مرحبا بكم في مشروع " and "تخرجنا (نطق)." This successful extraction of handwritten text demonstrates the tool's adaptability in deciphering individual writing variations.

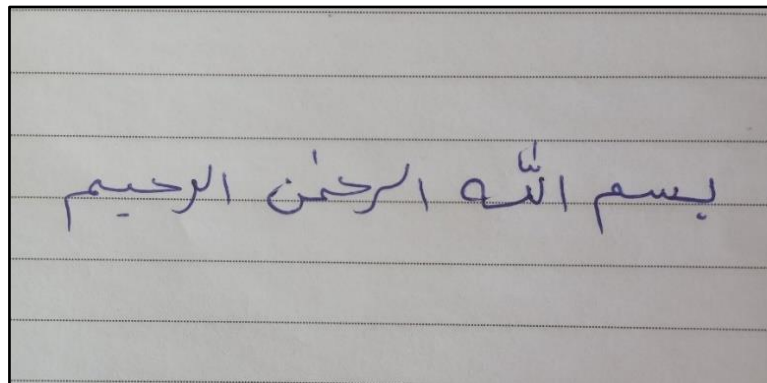


Figure 16 – Handwritten Script 1

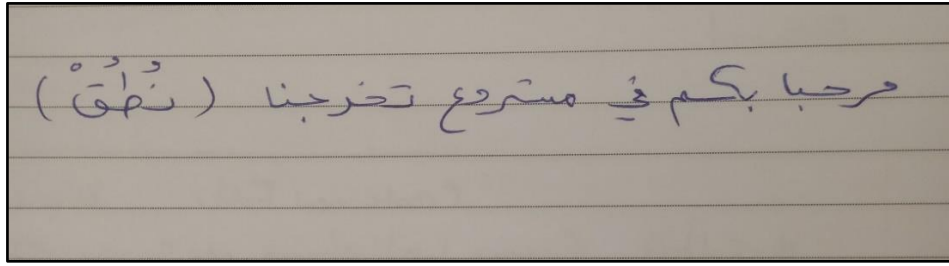


Figure 17 – Handwritten Script 2

- These findings highlight the impact of different script types on the OCR tool's detection accuracy. Notably, certain fonts, including the Diwani script (الخط الديواني), Thuluth script (خط الثلث), and specific variations of the Kufic script, presented challenges for the OCR tool. The intricate and ornamental letterforms inherent in these script types made it difficult for the tool to accurately recognize and extract the text, often resulting in the generation of unrelated symbols. These complexities posed significant hurdles for the OCR algorithm to interpret and process effectively.

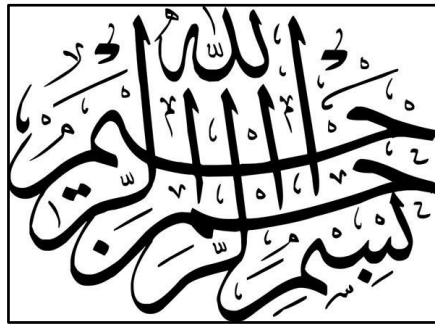


Figure 18 – Thuluth Script



Figure 19 – Diwani Script

However, it is worth noting that the OCR tool exhibited a remarkable 100% accuracy rate for the handwritten text that we tested. This success highlights the tool's adaptability and proficiency in deciphering individual writing variations. These findings emphasize the need for continuous efforts to improve the OCR tool's accuracy and broaden its compatibility with a wider range of script types. Ongoing work aims to enhance its performance in recognizing and

extracting text from complex script types, ensuring more reliable and accurate results across various fonts and handwriting styles for Arabic texts.

3.8.2 OCR with different text colors

To evaluate the OCR tool performance regarding text colors, we used images containing Arabic text with varying color contrast levels. The purpose was to determine whether the tool's detection and extraction capabilities were affected by these color variations. The OCR was tested on images that contain Arabic text with different color contrast levels.

Figure 20 below illustrates an example of low contrast text color, where the Arabic text appears in a light shade against a relatively light background. In this scenario, the OCR tool successfully detected and extracted the text with 100% accuracy, demonstrating its resilience to low contrast color combinations.



Figure 20 – Low Contrast Text Color

On the other hand, Figure 21 below represents an instance of high contrast text color, where the Arabic text stands out in a dark shade against a contrasting light background. Once again, the OCR tool proved its proficiency by accurately detecting and extracting the text with a 100% accuracy rate, reaffirming its ability to handle high contrast color combinations effectively.



Figure 21 – High Contrast Text Color

These results indicate that the OCR tool's performance remains unaffected by color contrast when it comes to detecting and extracting Arabic text. Whether the text is presented in low contrast or high contrast color combinations, the OCR tool consistently delivers accurate results, ensuring reliable text extraction regardless of color variations. This robust performance showcases the tool's capability to overcome color-related challenges and underscores its reliability in processing Arabic text across a wide range of color contrast scenarios.

3.8.3 Text to speech API performance

The performance of the Text-to-Speech (TTS) functionality was thoroughly evaluated to assess its effectiveness and accuracy in different text formats. The aim was to understand how the TTS tool performed under various conditions and identify any potential limitations for improvement.

- Without vowel marks (حركات):

The given text, which lacked vowel marks, was utilized to test the TTS tool's performance. It consisted of the following passage:

"علماء الأعصاب يقولون أنه في كل مرة تقاوم الانفعال والتصرف أثناء غضبك فأنت تقوم بإعادة برمجة دماغك ليصبح أكثر ذكاءً وهدوءاً في المستقبل أي بمعنى إذا نجحت في مقاومة غضبك باستمرار من الممكن أن تتغير شخصيتك كلياً لتصبح أكثر اتزاناً وحكمة."

During the evaluation, it was observed that the TTS tool encountered difficulties in accurately pronouncing some words in the absence of vowel marks. As a result, minor inaccuracies were detected in the generated audio output. As a reference, see Appendix B.1 for the audio output generated by the TTS tool.

- With vowel marks

To further explore the impact of vowel marks on the TTS performance, the same text was tested again, but this time with proper vowel marks added. The revised text for evaluation was as follows:

"علماء الأعصاب يقولون أنه في كل مرة تقاوم الإنفعال والتصرف أثناء غضبك فأنت تقوم بإعادة برمجة دماغك ليصبح أكثر ذكاءً وهدوءاً في المستقبل أي بمعنى إذا نجحت في مقاومة غضبك باستمرار من الممكن أن تتغير شخصيتك كلياً لتصبح أكثر إتراناً وحكمة."

In this case, only one word was mispronounced by the TTS tool, indicating an overall improvement in pronunciation accuracy when the text included vowel marks. As a reference, see Appendix B.2 for the audio output generated by the TTS tool.

Based on the findings, it can be inferred that the presence of vowel marks enhances the clarity and accuracy of the TTS reading. However, it is important to note that the OCR and TTS technologies provided by the Google Cloud Vision AI performed satisfactorily in most scenarios. Nonetheless, there is room for improvement in terms of enhancing the OCR tool's detection accuracy for all Arabic fonts and refining the TTS tool's pronunciation accuracy. Further developments and enhancements.

RESULTS AND DISCUSSION

In this chapter, we will mention the techniques we used to test all the project features, activity, and functionalities.

4.1 Scenario

To obtain a comprehensive understanding of the Notoq system's interaction from the user's perspective, a specific user was chosen, and their actions were adequately documented to construct a complete scenario. The primary objective of this scenario is to demonstrate the practical utilization of the system's services, prioritizing user experience over technical specifications.

Initially, the user proceeds to sign up for the application by providing their username, email, and password. Subsequently, the application verifies the validity of the email address and checks if it has already been registered. If the entered email address is already taken, the account creation process fails, and the user receives the following error message: "Error: Email is already taken." In such a case, the user remains on the signup page until a valid email address is provided.

When the user enters their password, the application assesses its strength and provides feedback indicating whether it is categorized as weak, medium, or strong. Figure 22 shows the signing up operation. Refer to Appendix A.1 for a comprehensive view of the Sign-Up GUI.



The screenshot displays a mobile application interface for signing up. It features three input fields: 'اسم المستخدم' (Username) with the value 'Ahmad Hussien', 'البريد الالكتروني' (Email) with the value 'ahmad_hussien@gmail.com', and 'كلمة السر' (Password) with masked characters. Below the password field, a strength indicator shows 'متوسطة' (Medium). At the bottom is a large blue button labeled 'سجل' (Sign Up). The interface has a light yellow background and a dark blue header bar with a back arrow.

Figure 22 - Sign Up Operation

Once the user has successfully registered, the application will display a welcome page that provides a general overview of the application and its objectives. This page aims to familiarize the user with the key features and benefits of the application. Additionally, to enhance user accessibility and accommodate different preferences, a button is included on the welcome page that allows users to listen to the text instead of reading it. This feature aligns with the application's concept and promotes inclusivity by catering to diverse user needs. As shown in Figure 23.

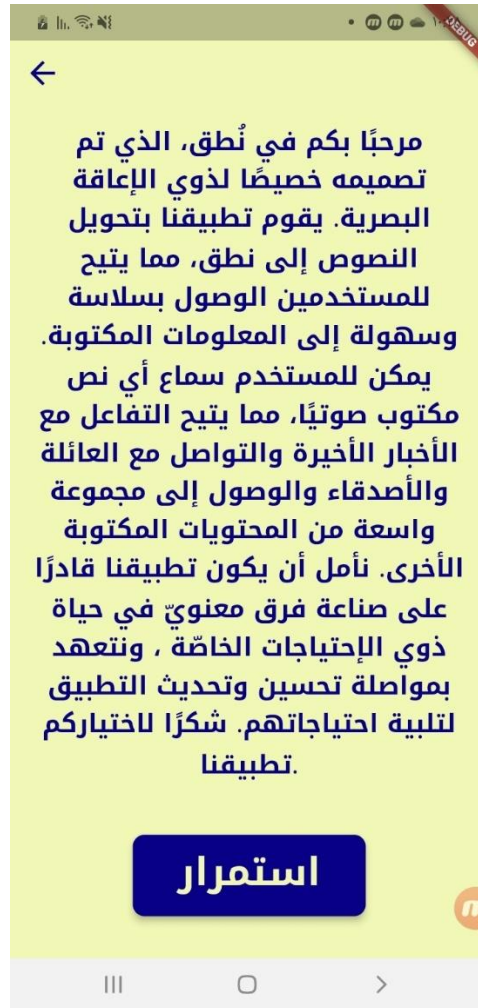


Figure 23 – The Welcome Page

Once the user has completed the registration process, they gain the ability to proceed with logging into their account by providing their registered email address and password. The application proceeds to validate the accuracy of the provided email and subsequently verifies the correctness of the entered password. If either the email or password is found to be incorrect, the login process will not succeed, and the user will be presented with an error message conveying the following information: "Login failed: The email or password you entered is incorrect.". Refer to Appendix A.2 for a comprehensive overview of the Login GUI, including detailed information and visual representations.

The first user, Ahmad, who has low vision, had successfully logged in. Ahmad is directed to the home page of the application. The home page presents Ahmad with five options to choose from: "Create Folder," "Paste Text," "Scan Pages," "Upload Photo," and "Speech to Text." Which are thoughtfully designed to facilitate Ahmad's seamless interaction with the application. This user interface, as shown in Figure 25, is designed to accommodate Ahmad's visual needs and provide accessible options for his interaction with the application.



Figure 24 - Home Page

Ahmad decides to select the first option, "Create Folder," which redirects him to the "My Folders" page. Here, he creates a new folder named "pages." Ahmad intends to use this folder to store pages from his favorite book, allowing him to easily access it whenever needed. Figure 25 illustrates the appearance of the "My Folders" page.

Taking advantage of the functionality provided by the application, Ahmad proceeds to upload a quote that resonates with him to the designated "pages" folder. The uploaded quote is saved within the folder with the exact same name as the image file stored on Ahmad's device. To gain a visual representation of this process, refer to Figure 26, which accurately portrays the interface and displays the relevant image file along with its corresponding name.



Figure 25 - My Folders

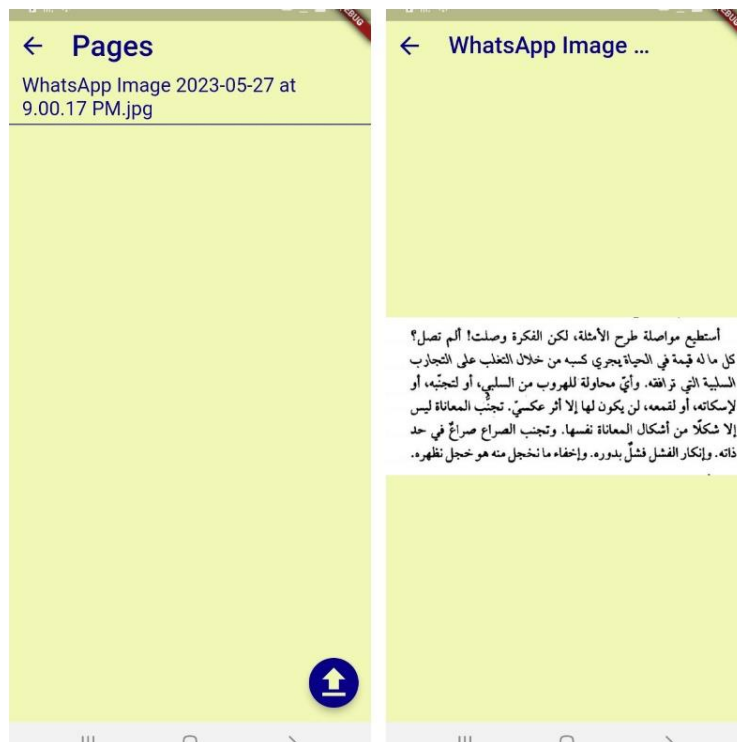


Figure 26 - Inside Pages Folder

Hussam, who is Ahmad's friend, sent him a quote through WhatsApp knowing Ahmad's interest in such quotes. However, Ahmad faced difficulty in reading the quote due to its long length and small font size. To overcome this challenge, Ahmad utilized Notoq's "Paste Text"

feature, which allows him to have the text read aloud. He copied the text from the WhatsApp message and pasted it into Notoq. Furthermore, Ahmad selected the male gender option for the text reader, as shown in Figure 27. This feature assists Ahmad in accessing and comprehending the content more effectively.

After Ahmad pasted the text into Notoq and clicked on "Convert to Audio," the application took a few seconds to process the conversion. Once completed, two buttons, namely "Download Audio" and "Play Audio," became visible on the screen. These buttons provide Ahmad with the options to either download the audio file or directly play the generated audio rendition of the text. As a reference, see Appendix B.3 for the audio output generated by the TTS tool.



Figure 27 - Paste Text Page

On a Sunday morning, Ahmad woke up feeling unwell, his body engulfed in a fever. Concerned for their son's health, his parents promptly took him to the hospital for medical attention. As a result of his illness, Ahmad was unable to attend his university classes, including an upcoming midterm examination scheduled for Tuesday, which covered the material discussed in today's lecture. Recognizing the significance of the lecture notes for his exam preparation, Ahmad reached out to one of his classmates and requested a copy. Understanding the urgency of the situation, his thoughtful colleague promptly printed the notes and handed Ahmad a copy to assist him in his studies.

However, when Ahmad arrived home and attempted to review the lecture notes, he encountered an unexpected obstacle. The font used in the printout was unfamiliar to him,

rendering the letters unrecognizable and hindering his ability to comprehend the content. Frustrated but determined, Ahmad sought out a solution to overcome this challenge.

Luckily, Ahmad had recently come across an innovative application called Notoq. It boasted various features designed to aid individuals in accessing and understanding text. Ahmad decided to utilize the third feature of Notoq, known as "Scan Pages," in the hopes that it would be the solution he desperately needed. With his mobile phone in hand, Ahmad positioned the camera over the printed lecture notes and pressed the "Take photo" button within the Notoq application then it shows him the photo with two options "Take again" and "Read Text" the "Take again" button will allow him to re-take the photo and scan the page. While Ahmad captured the photo correctly and fully, he pressed on "Read Text". After a brief processing period, Notoq displayed a preview page showcasing the photo Ahmad had just captured. The preview page offered two options: "Upload Image" and "Select Gender of Reader." As shown in figure 28.

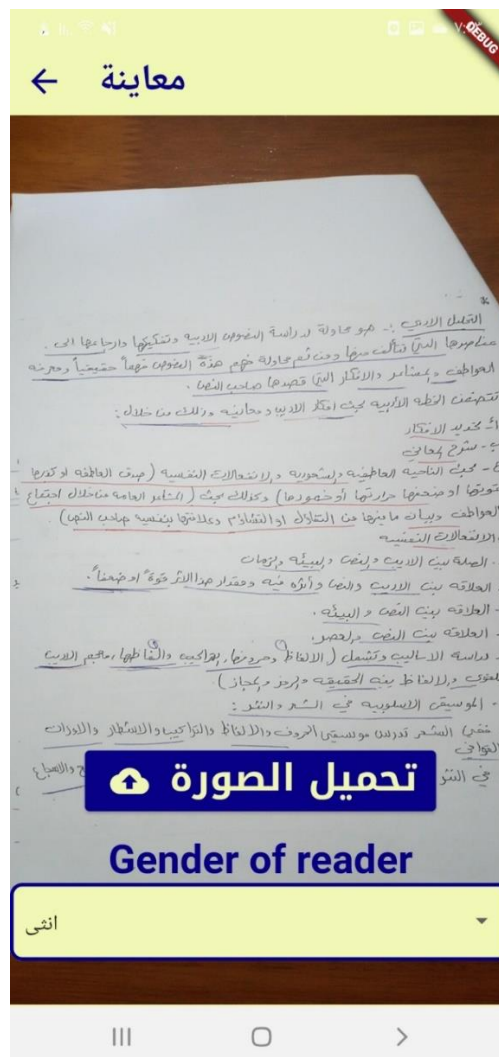


Figure 28 - Scan Pages Preview

Ahmad selected the appropriate gender (Female) and proceeded to press the "Upload" button. Almost instantaneously, Notoq employed Optical Character Recognition (OCR) technology and Text-to-Speech (TTS) functionality to convert the scanned text into an audio format. As a result, two buttons appeared on the screen: "Download Audio" and "Play Audio." As shown in Figure 29.

Understanding the value of having an audio version of the lecture notes, Ahmad chose to download the audio file onto his device. This way, he could listen to it at his convenience and even share it with his friends who may have also missed the lecture. After downloading the audio, Ahmad eagerly clicked the "Play Audio" button. As the audio played, Ahmad effortlessly assimilated the once intimidating material, allowing him to comprehend the concepts and study efficiently for his upcoming midterm examination. As a reference, see Appendix B.4 for the audio output generated by the TTS tool.

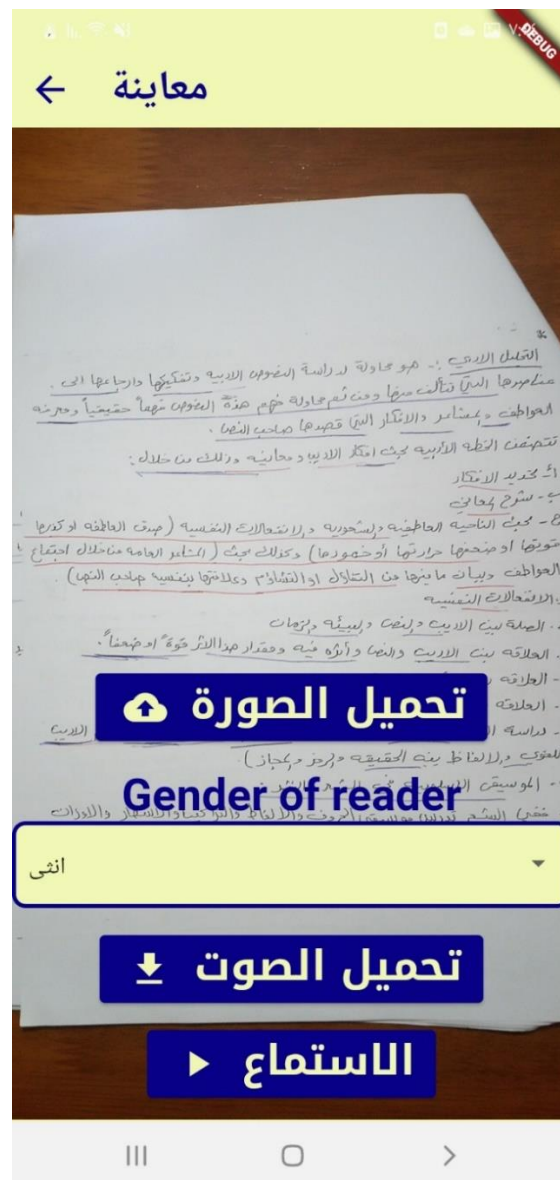


Figure 29 - Scan Pages Feature

One day, as Ahmad was casually scrolling through his social media feed, he stumbled upon a post where his mother had mentioned him. To his delight, the post contained a captivating photo adorned with a beautiful quote which is a type of content Ahmad truly appreciated. Recognizing his affinity for such quotes, his mother had thoughtfully shared it with him.

However, Ahmad faced a familiar predicament. The quote, embedded within the image, was not accessible to him due to his difficulty in perceiving visual content. Undeterred, Ahmad immediately turned to his trusted companion, Notoq, for assistance once again. Without delay, Ahmad saved the image to his device and opened the Notoq application. He navigated to the fourth feature, aptly named "Upload Photo," which had previously proven its worth in assisting him with printed text. Ahmad uploaded the image containing the beloved quote, hoping that Notoq would provide a solution. Much like the "Scan Pages" feature, Notoq embarked on a similar process with the uploaded photo. Utilizing its powerful OCR and TTS technologies, Notoq analyzed the visual content and converted the text within the image into an audio format. Ahmad eagerly awaited the results, anticipating the opportunity to engage with the quote that had caught his attention.

Within moments, Notoq fulfilled Ahmad's expectations by generating an audio file. Now, Ahmad had the option to either download the audio or play it directly from the application. The seamless integration of technology and accessibility offered by Notoq allowed Ahmad to fully immerse himself in the profound words of the quote his mother had shared. As shown in Figure 30. As a reference, see Appendix B.5 for the audio output generated by the TTS tool and for detailed information and visual representations of the "Upload Image" functionality, refer to Appendix A.3.



Figure 30 - Upload Photo Feature

In another unexpected turn of events, Professor Sarah, Ahmad's teacher at the university, sent him a voice note containing crucial information about the upcoming midterm examination, including the exact date, time, and location. Eager to have a physical copy of the details to reference and share with his friends, Ahmad once again turned to the versatile capabilities of Notoq as he was curious to test the last feature which is idea totally the opposite of the application idea by converting the speech into text.

Ahmad opened the Notoq application and swiftly navigated to its fifth feature, appropriately named "Speech to Text." This feature was specifically designed to transcribe spoken words into written text - an ideal solution for Ahmad's current predicament. Excited to convert the voice note into a readable format, Ahmad tapped on the feature button within Notoq.

Within the feature, Ahmad located the voice note sent by Professor Sarah and proceeded to upload it. Notoq's advanced technology, which is Speech to Text API provided by Google cloud effortlessly processed the audio content, accurately transcribing the spoken words into written text. Ahmad now had the text representation of Professor Sarah's voice note at his disposal, as shown in Figure 31. As a reference, see Appendix B.6 for the audio used to transcribe. For a comprehensive overview and visual demonstration of the "Speech to Text" feature, consult Appendix A.4.



Figure 31 - Speech to Text Feature

After some time, Ahmad had the opportunity to meet a new friend, Omar, at the gym. Interestingly, Omar, like Ahmad, also experienced vision impairment. Although Omar was born and had spent most of his life in London, his fluency in Arabic was limited. Consequently, the two friends communicated primarily in English.

During one of their conversations, Omar discovered Ahmad's passion for poems and quotes. Eager to share a profound quote with his newfound friend, Omar reached for his mobile phone and displayed a cherished line from Shakespeare: "All the world's a stage, and all the men and women merely players. They have their exits and their entrances; and one man in his time plays many parts" [12]. However, both Ahmad and Omar encountered difficulties in reading the quote due to their visual impairments. Yet, Ahmad, being well-acquainted with the accessibility tool Notoq, recognized an opportunity to enhance their experience and introduce Omar to this innovative application.

Ahmad swiftly resorted to Notoq's "Paste Text" feature, allowing him to conveniently copy and transfer the written content. By utilizing this feature, Ahmad aimed to assist both himself and Omar in fully appreciating the beauty of the poem. This action also served the purpose of introducing Omar to the capabilities of Notoq and guiding him towards exploring this application further as this application also could help Omar in his Arabic learning Journey. As shown in Figure 32, Ahmad seamlessly employed Notoq's "Paste Text" feature to transfer the Shakespearean quote onto the application to gain the audio format of it within a few seconds. By doing so, he facilitated an opportunity for both him and Omar to engage in a more accessible manner, transcending the limitations imposed by their visual impairments. As a reference, see Appendix B.7 for the audio output generated by the TTS tool.



Figure 32 - Paste Text Feature with English Text

After becoming aware of Notoq's capabilities, Omar decided to download and install the application on his personal device. With enthusiasm, he proceeded to sign up for an account, granting him access to the full range of features and personalized settings.

Excited to customize his Notoq profile, Omar proceeded to access the section dedicated to his personal information within the settings tab. This tab encompasses a total of four options for users to choose from. The first option, labeled as "Personal Info," enables users to view and edit their profile information. The second option, named "My Folders," allows users to access a page displaying the folders they have created and provides the functionality to create additional folders if desired. The third option, titled "Language," permits users to effortlessly switch between the Arabic and English interfaces of Notoq. Lastly, the fourth option, "Log out," simply logs the user out of their account, redirecting them to the initial page where they can choose to either log in or sign up anew. The visual representation of the settings tab can be observed in Figure 33.



Figure 33 - Settings Tab

To make his profile more engaging and relatable, in the personal info section, Omar chose to add a profile picture. He selected a photo that resonated with him, allowing others to connect a face to his name within the Notoq community. Additionally, Omar took the opportunity to provide a brief bio about himself. He crafted a concise yet informative description that captured his academic pursuits and personal interests: "computer engineer interested in poems" This bio not only reflected Omar's passion for his field of study but also highlighted his love for literature.

This page contains user information such as the username, email address, and password details. Additionally, it grants users the ability to modify their password if necessary. Moreover, the page includes the "Member Since" data, showcasing the date when the user initially joined the Notoq community and became a member. Furthermore, there is a dedicated bio section where users can provide a concise description about themselves. As shown in figure 34.

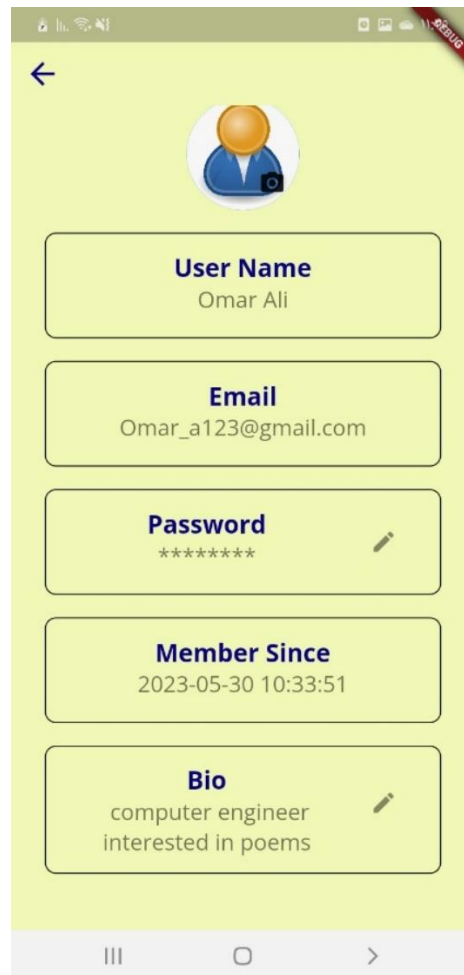


Figure 34 - Personal Info

Recognizing his limited proficiency in Arabic, Omar made the decision to set the language preference within Notoq to English. By selecting English as his primary language, Omar ensured that the application's interface and features would be readily accessible and comprehensible to him. As shown in Figure 35.

Conscious of the challenges his low vision posed and his interest in learning Arabic, Omar made a conscious commitment to leverage Notoq whenever he encountered difficulties due to his visual impairment or whenever he wants to read something written in Arabic. By utilizing this assistive tool, Omar embraced the opportunity to enhance his reading experiences, access a wealth of knowledge, and actively engage with various forms of textual content. For a detailed exploration of the settings tab and its various buttons, including their functions and configurations, please refer to Appendix A.5.



Figure 35 - Interface Language

4.2 Testing

Test cases are used to verify a feature or functionality of any software application. To verify Notoq functionalities, we must test Android availability on the devices and other functionalities that are implemented in the system.

As a result, our application is expected to work on 100% of android devices. We tested Notoq on different devices with different android versions to see if the user can create an account and if the features of our application work properly. The results are shown in Table 5.

Table 5- Testing Notoq on Different Devices

Device name	Android version	can create an account?	Notoq features work properly?
SAMSUNG A10	10	Yes	Yes
SAMSUNG A31	10	Yes	Yes
Blackview Tab 5	12	Yes	Yes

4.3 Survey

A survey was conducted among university students who used the Notoq mobile application to assess its quality and effectiveness. The survey, consisting of 12 questions, aimed to evaluate Notoq's achievement of goals and user interaction. Feedback on performance, usability, and usefulness was gathered to provide valuable insights for improvement. Analysis of the target user group's experiences identified successful aspects and potential enhancements. Detailed survey questions and statistical results are presented in Table 6.

Table 6 - Survey and statistical results

Q1: Have you ever used a Text-to-Speech application before?			
Yes		No	
48.7%		51.3%	
Q2: Have you ever used a Text-to-Speech application for Arabic text?			
Yes		No	
37.2%		62.8%	
Q3: How important is the availability of a Text-to-Speech application for Arabic text to you?			
Very important	Important	Not very important	Not important at all
48.7%	38.5%	9%	3.8%
Q4: How useful did you find that Text-to-Speech applications can be?			
Very useful	Quite useful	Not useful	Never tried before
60.3%	33.3%	3.8%	2.6%
Q5: After watching the demo, did you find it useful for reading Arabic text?			
Very useful	Quite useful	Not useful	Never tried before
71.8%	24.4%	2.6%	1.3%
Q6: How clear and accurate did you find the audio messages for the tested text samples?			
Very clear	Quite clear	Not very clear	Not clear at all
61.5%	33.3%	3.8%	1.3%
Q7: In general, how do you rate the GUI (Graphical User Interface) in Notoq?			
Excellent	Good	Fair	Poor
51.3%	37.2%	10.3%	1.3%
Q8: After watching the demo, do you think that vision-impaired users will have any struggles while using Notoq?			
Yes		No	
44.6%		56.4%	
Q9: How easy was it to navigate the Notoq app?			
Very easy and smooth	Easy	Difficult	Very difficult
53.8%	35.9%	10.3%	0%
Q10: How do you evaluate the feature of customizing the settings of the used voice(gender, speed, etc. ?			
Very useful	Quite useful	Not useful	Never tried before
64.1%	33.3%	1.3%	1.3%

Q11: Do you like the feature that allows you to save the speech output as an audio file for later listening?	
Yes	No
92.3%	7.7%
Q12: Please provide any additional comments or feedback that you think would help us improve the Notoq app.	
If you change the background color to white color or make it lighter, I think it will be better for some people.	
The rest is too good 😊.	
Notoq app seems to be a great idea 👍👏	
Well done!	
Improve the user interface and add more graphics.	
It is very useful for anyone who wants to learn the pronunciation of the Arabic language or even use the audio that comes out to design and produce beautiful videos.	

The survey was conducted among 87 students, providing valuable insights into the evaluation of the Notoq project. Analyzing the survey results provides valuable insights into the user's perspective. In response to the first question, it was found that 48.7% of the participants had previous experience using a Text-to-Speech (TTS) application, while 51.3% did not. This indicates that although a significant portion of the participants had prior exposure to TTS applications, there is still a considerable proportion that is relatively new to this technology.

When considering the use of TTS applications specifically for Arabic text, the survey revealed that only 37.2% of the participants had previous experience, while the majority (62.8%) had not used an Arabic TTS application before. This highlights the relatively lower adoption of Arabic TTS applications compared to their English counterparts, suggesting a potential area for improvement and growth in the market.

Examining the importance of having a Text-to-Speech application for Arabic text, it was found that 48.7% of the participants considered it very important, while 38.5% found it important. This indicates a significant demand and recognition of the value of Arabic TTS applications among the surveyed user group. In terms of usefulness, most participants (60.3%) found Text-to-Speech applications to be very useful, with 33.3% considering them quite useful. This reinforces the positive perception of TTS technology and its potential benefits for users.

After watching the demo, the survey revealed that 71.8% of the respondents found the Notoq application useful for reading Arabic text, while 24.4% found it quite useful. This positive feedback further supports the effectiveness of Notoq in meeting the needs of users. When evaluating the clarity and accuracy of the audio messages for the tested text samples, most participants (61.5%) found them very clear, with 33.3% considering them quite clear. This demonstrates the successful implementation of Notoq's text-to-speech engine, ensuring high-quality audio output for the users.

Regarding the Graphical User Interface (GUI) of Notoq, 51.3% of the participants rated it as excellent, while 37.2% considered it good. However, there were suggestions to improve

the user interface and add more graphics, indicating potential areas for enhancement. It is important to note that the GUI was carefully designed with consideration for individuals with low vision, using high contrast colors like light yellow and dark blue. The limited use of images was intentional to prioritize usability for individuals with visual impairments. Future updates can consider these suggestions while maintaining accessibility and usability for the target audience.

In terms of accessibility, 44.6% of the participants believed that vision-impaired users may face struggles while using Notoq, while a majority (56.4%) did not anticipate any difficulties. This insight can guide further improvements to ensure a seamless and inclusive user experience for individuals with visual impairments. When considering the ease of navigation within the Notoq app, 53.8% of participants found it very easy and smooth, with 35.9% finding it easy. This reflects the intuitive design and user-friendly interface of the application.

The feature of customizing voice settings, such as gender and speed, was evaluated positively by 64.1% of the participants, considering it very useful, while 33.3% found it quite useful. This feature was well-received, indicating the importance of customization options for enhancing user experience. The survey also highlighted the popularity of the feature that allows users to save the speech output as an audio file for later listening, with 92.3% expressing their liking for this feature. This demonstrates the practicality and usefulness of such functionality.

Overall, the survey feedback and comments suggest a positive perception of Notoq, with users recognizing its usefulness for language learning, audio production, and pronunciation improvement. The suggestions for improvement, such as changing the background color and enhancing the user interface, provide valuable insights for future updates and refinements of the Notoq app.

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In the scope of this project, we have successfully achieved the development of a user-friendly mobile application designed to aid visually impaired individuals by converting text to speech. The application, Notoq, was conceived with a focus on accessibility and usability, resulting in a tool that is not only functional but also intuitive to use. Throughout the project, we have made significant accomplishments:

- Integration of a responsive and accurate text-to-speech engine that seamlessly converts written text into audible speech, enhancing the accessibility of information for visually impaired users.
- Careful design of the user interface to ensure easy navigation and operation of the application.
- The successful realization of these features underscores the potential of technology to significantly improve the quality of life for visually impaired individuals, empowering them to access and engage with textual content more effectively.

Moreover, during the evaluation phase, we observed positive feedback from users who found Notoq to be a valuable tool in their daily lives. The application's user-friendly design and reliable text-to-speech capabilities have garnered appreciation for its impact on accessibility and independence.

5.2 Future Work

While this project has achieved its initial goals, there is ample room for future development and enhancement to further expand the capabilities and reach of Notoq. The following points outline potential areas for future work:

- Improve the text-to-speech engine's ability to understand and accurately pronounce technical terminology and informal language, expanding its usability in specialized or localized contexts.
- Integrate an AI-powered file-to-speech feature, enabling the application to upload text from various file formats, such as books, articles, and documents, broadening the scope of content that can be accessed and read aloud.
- Expand language support by collaborating with linguistics experts and native speakers to include more languages and dialects in the text-to-speech engine, ensuring accurate and natural-sounding speech output.
- Implement voice command functionality, allowing users to operate the application entirely hands-free, which would be particularly beneficial for individuals with additional mobility impairments.

- Explore integration with other popular applications such as email clients, e-book readers, and web browsers to extend the text-to-speech function to a wider range of digital activities, providing a seamless and comprehensive accessibility solution.
- Enhance the application's compatibility with different operating systems and devices, ensuring seamless functionality across various platforms such as iOS, Android, and web browsers. This would enable a wider user base to access and benefit from the application.
- Implement real-time translation capabilities, allowing users to convert text from one language to another and have it read out in the desired language. This feature would greatly assist individuals who are multilingual or those who need language support while traveling or communicating with people from different linguistic backgrounds.

By addressing these areas of future work, Notoq can continue to evolve and make a lasting impact on the lives of visually impaired individuals. The application stands as a powerful tool for accessibility, and with continued development and enhancement, it has the potential to bring about positive change and empower users to access information with greater ease and independence.

REFERENCES

- [1] Aaron O’Neil , “Arab world - total population 2011-2021,” *Statista*. [Online]. Available: <https://www.statista.com/statistics/806106/total-population-arab-league> [Accessed May 2023].
- [2] W. H. Organization, “Over 2.2 billion people are blind or visually impaired, and the numbers are rising. Here’s why.,” *Medium*, Oct. 19, 2020. <https://medium.com/who/over-2-2-billion-people-are-blind-or-visually-impaired-and-the-numbers-are-rising-heres-why-f0da6df0a191>
- [3] “Text Accessibility,” *Faculty Central*, Apr. 11, 2014. <https://facultycentral.wordpress.com/accessibility/text-accessibility/> [Accessed April 2023].
- [4] Natural reader, “text to speech online,” *Naturalreaders.com*, 2019. <https://www.naturalreaders.com/>
- [5] “Narrator’s Voice - TTS - Apps on Google Play,” *play.google.com*.
- [6] “Speechify - Listen to text with Speechify,” *speechify.com*. <https://speechify.com/>
- [7] Veronica Lewis , “Choosing High Contrast Color Schemes For Low Vision | Veroniiiica,” *veroniiiica.com*, Oct. 25, 2019. <https://veroniiiica.com/high-contrast-color-schemes-low-vision/>
- [8] Lubna Rasheed , “Core features of Android Studio,” *Educative: Interactive Courses for Software Developers*. <https://www.educative.io/answers/core-features-of-android-studio> [Accessed May 2023].
- [9] Apache Friends, “XAMPP Installers and Downloads for Apache Friends,” *www.apachefriends.org*, 2022. <https://www.apachefriends.org/>
- [10] Oracle, “What Is a Database? ” , *Oracle.com*, 2021. [Online]. Available: <https://www.oracle.com/database/what-is-database> [Accessed April 2023]
- [11] Oracle, “What Is a Database? ” , *Oracle.com*, 2021. [Online]. Available: <https://www.oracle.com/database/what-is-database> [Accessed April 2023]
- [12] Shakespeare, William. "As You Like It." Act II, Scene VII.

APPENDICES

APPENDIX A: Application Layout and Features Overview

In Appendix A, we present a visual overview of the application layout in both English and Arabic, highlighting key features and functionalities. The images provided illustrate the main pages, login process, image upload process, speech-to-text feature, and the settings tab. This section aims to provide a comprehensive understanding of the application's layout and functionality for a better user experience.

A.1 : Main page in English and Arabic, along with signup page that handles errors such as email duplication.



Figure 36 - Main and Sign-Up Pages

A.2 : The login page, showcasing notification errors for incorrect login credentials and other login-related issues.

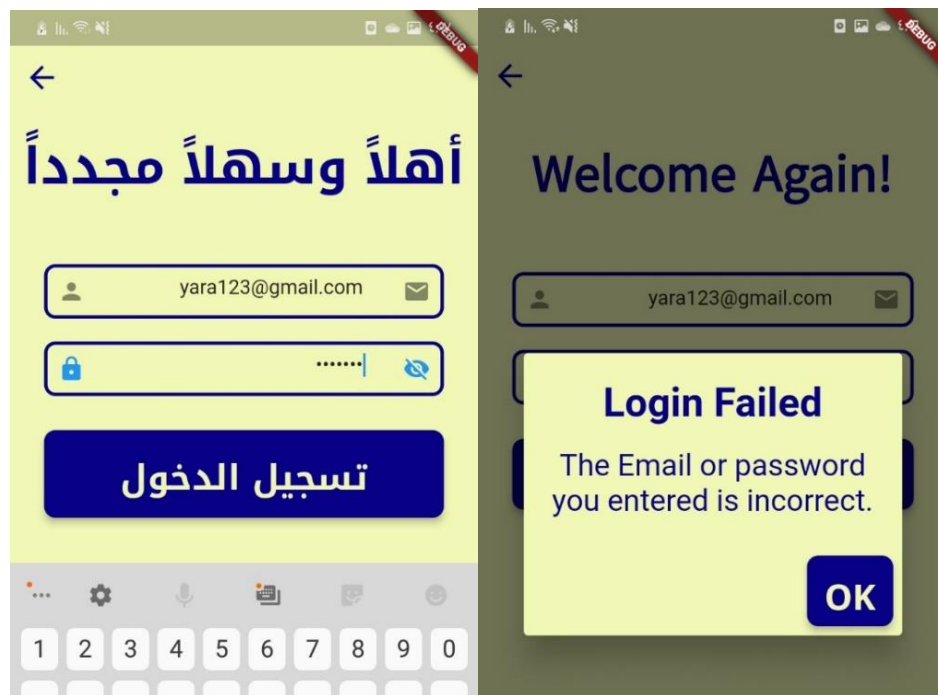


Figure 37 - Login operation

A.3 : Image upload process in both Arabic and English, including the steps of selecting, processing, and converting text to speech.

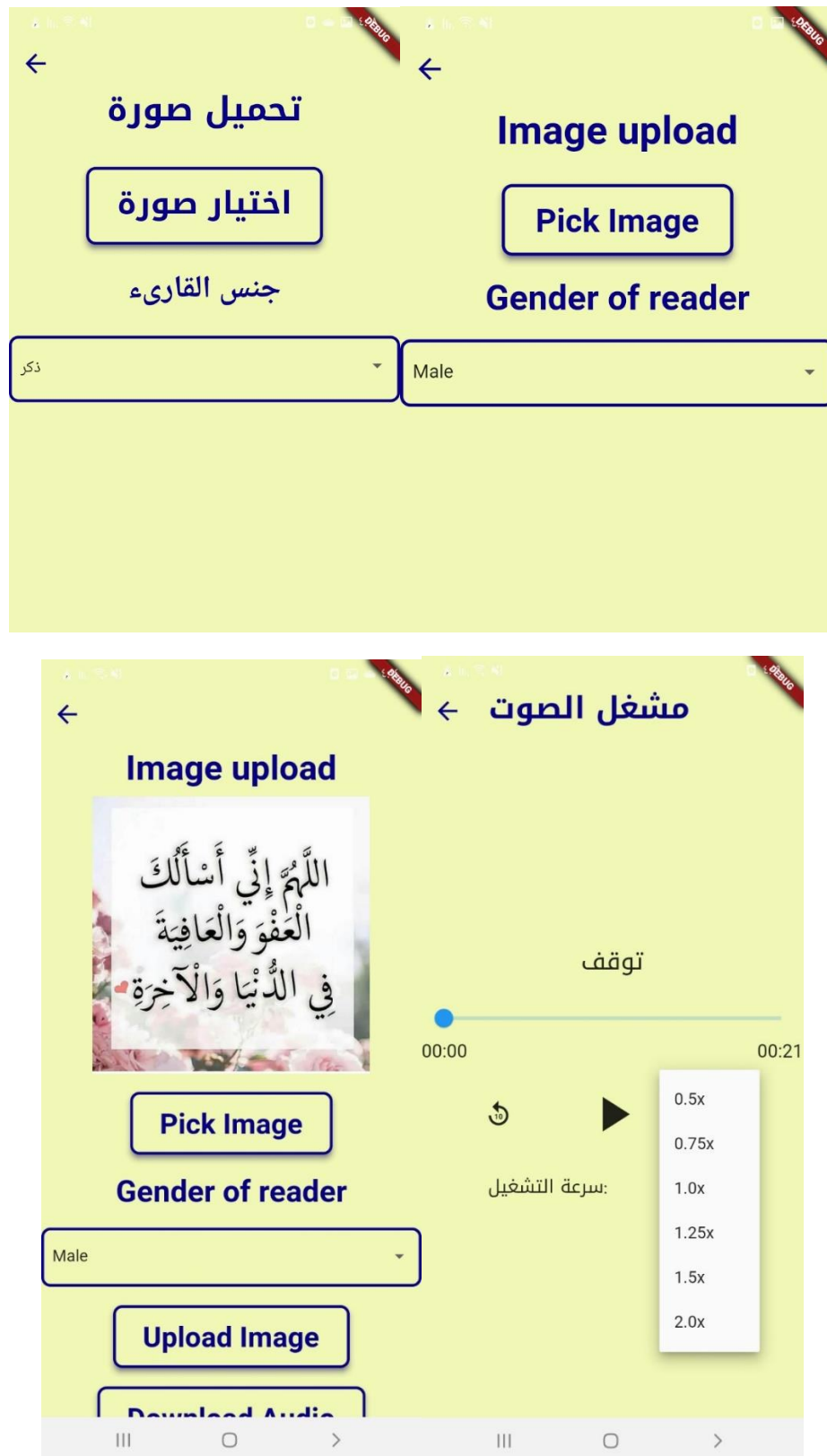


Figure 38 - Image Upload

A.4 : Speech-to-text feature supports Arabic and English languages, converting spoken words to written text.



Figure 39 - Speech-To-Text Feature

A.5 : Settings tab which provides users with control over various features, such as folder management, personal information updates, language selection, and a Log Out option.

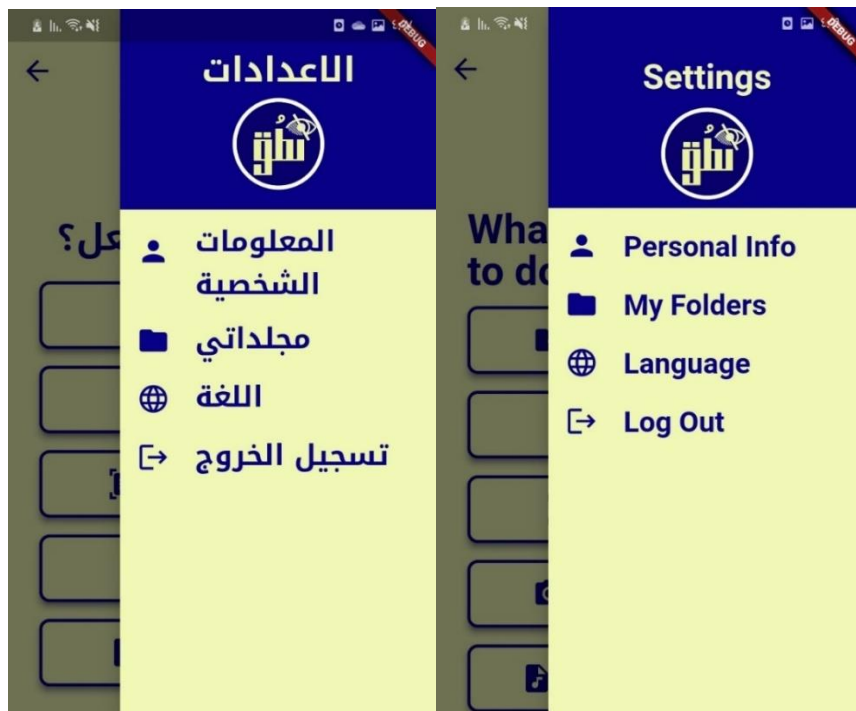


Figure 40 – Settings

A.6 : Personal Information where the user can check for the account information, change profile picture, bio and update password.



Figure 41 - Personal Information

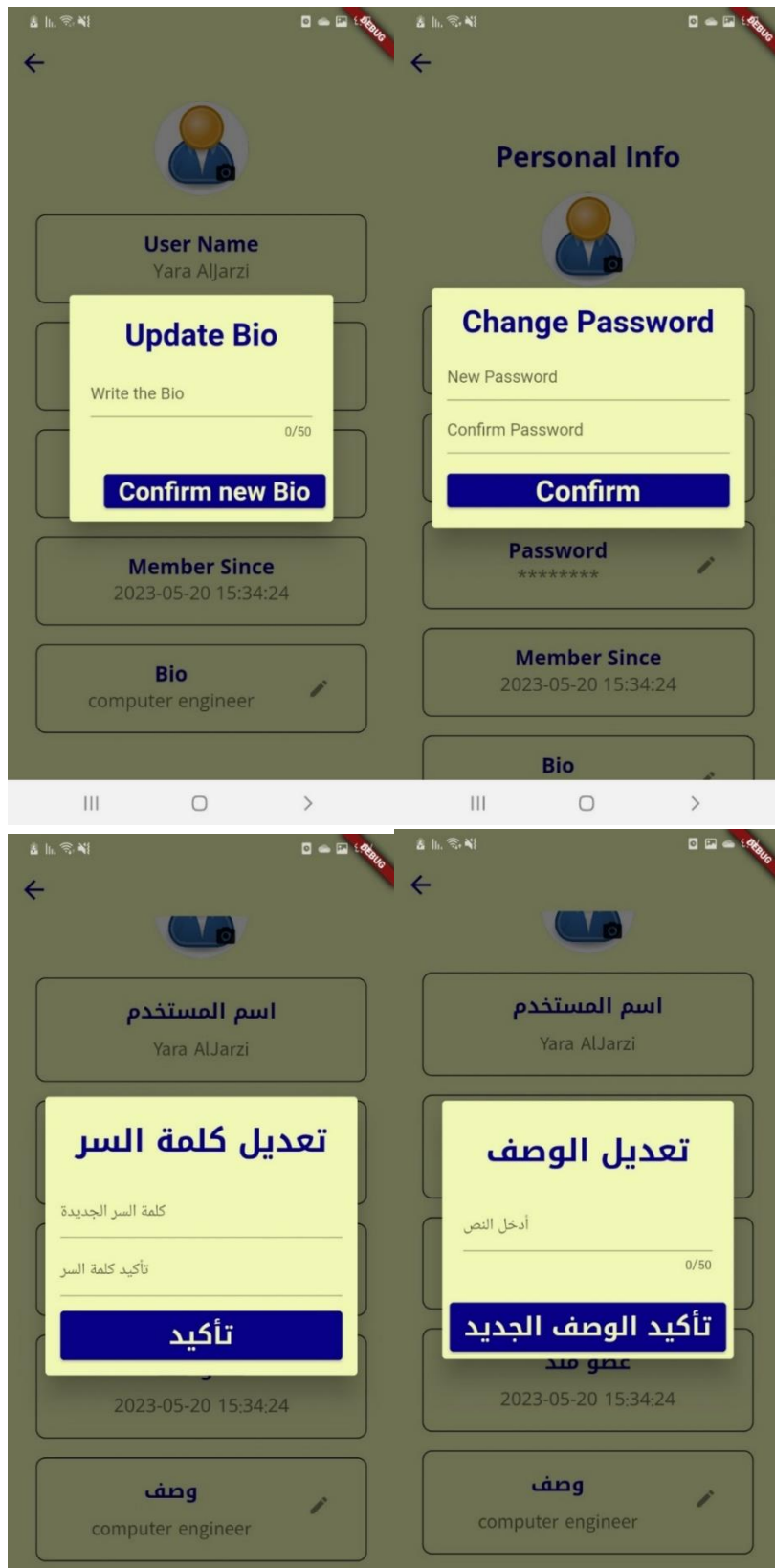


Figure 42 - Change Password and Bio



Figure 43 - Language Tab

APPENDIX B: Sample Sound Test Records

The following records provide examples of the audio output generated by the Text-to-Speech (TTS) tool in various scenarios, showcasing the capabilities and functionality of Notoq:

B.1 : Audio output generated by the TTS tool for the text without vowel marks:
https://storage.googleapis.com/audiosavephp/audio_6475e76d51bf41.46480692.mp3

B.2 : Audio output generated by the TTS tool for the text with vowel marks:
https://storage.googleapis.com/audiosavephp/audio_646fac48b1dd02.65527759.mp3

B.3 : Audio output generated by the Text-to-Speech (TTS) tool for the quote received through a messaging platform, demonstrating how Notoq converts text into an audio format for easier comprehension:

https://storage.googleapis.com/audiosavephp/audio_6472541389e3b2.64678750.mp3

B.4 : Audio output generated by the TTS tool for the lecture notes scanned using Notoq's "Scan Pages" feature, showcasing the conversion of printed text into an audio format to aid individuals with visual impairments in accessing and understanding the content:
https://storage.googleapis.com/images_ocr_tts/audio_647596544a7435.11975248.mp3

B.5 : Audio output generated by the TTS tool for the image containing a quote uploaded using Notoq's "Upload Photo" feature, illustrating the conversion of visual content into an audio format, enabling individuals with visual impairments to engage with the textual information:
https://storage.googleapis.com/images_ocr_tts/audio_64759cd4a3c8e5.25925426.mp3

B.6 : Audio for a voice note transcribed into text using Notoq's "Speech to Text" feature, demonstrating the conversion of spoken words into written text for improved accessibility and comprehension: https://storage.googleapis.com/audiosavephp/audio_6475a01ebb46d.wav

B.7 : Audio output generated by the TTS tool for a pasted quote in Notoq's "Paste Text" feature, exemplifying how text can be transformed into an audio format, allowing individuals with visual impairments to fully appreciate and engage with written content:
https://storage.googleapis.com/audiosavephp/audio_6475a450158e75.00864701.mp3