Evaluation and Redesign of a Document Scanning Application

Stavropoulos Konstantinos

Charalampopoulos Dimitrios

Table of Contents

1) Project Summary

In this paper, we analyze, evaluate, and redesign a document scanning application using human-centered design principles. Initially, we examine several available applications and from these, we choose one that seemed relatively unwieldy to us. Subsequently, we conduct a user analysis, meaning we analyze typical users, the scenarios in which they might use a scanning application, and define personas. In parallel, we perform a task analysis, meaning we study the context in which this application is used, for what purpose, and in what way. This allows us to define the functional and non-functional requirements of the application. Following this, we evaluate the application using both expert-based and user-based evaluation methods to derive observations. Finally, we redesign the application to address the observations we and the users had. Lastly, we compare the redesigned application with the original one and report conclusions.

2) Review of Scanning Applications

To choose a scanning application, we initially examine the available applications and select the one we believe requires the most improvement.

i) Tiny Scanner

This is an application with a free trial period of a few days that allows all scanning, digitization, processing, and document sharing functions. Furthermore, it is easy to understand how to use the application.

ii) Microsoft Lens

This specific application is free; however, we found it relatively unwieldy because during document scanning it displays a huge red square that confuses us. It also has some usability issues, such as when you take the first photo, it automatically takes you to the preview, regardless of whether you have more photos to scan. Therefore, the application can be greatly improved, and in the rest of this paper, we will focus on this application.

iii) PDF Scanner – EZTech Apps

This specific application is free, but we found it particularly problematic due to several designer choices. During scanning, the page size selection is made by displaying a long list of all available options, where the user has to search extensively to find the desired size (in fact, the most popular option, which is A4, was not quickly and easily accessible). At the same time, there is no option to directly replace pages of a file from one photo to another.

iv) Easy Scanner

Easy Scanner is a free application where we identified some problems during use. While the application works normally for a single page, when scanning multiple pages is required, it is not possible to edit the pages directly after scanning. Also, the page reordering option is difficult to locate, as it is placed in the extra options.

v) TurboScan

This specific application is free and allows all scanning, digitization, processing, and document sharing functions. The only problem we identified was that the design was overly simple, using icons without words, which could confuse some users.

3)User Analysis

To evaluate the application, we first conducted a user analysis. This involves defining the application's stakeholders, exploring various use case scenarios relevant to typical users, and developing personas – representative individuals who would utilize the application. To gain further insight into typical user behavior, we interviewed people we know regarding their experiences with scanning applications, focusing on their usage patterns, functionalities employed, and frequency of use.

Stakeholders

Stakeholders are individuals who interact with or are impacted by the application. We categorize them into three types: primary stakeholders, who are regular users; secondary stakeholders, encompassing occasional or indirect users; and tertiary stakeholders, who are impacted by the application without directly using it. For a document scanning application, primary stakeholders include all users who utilize the application for scanning documents. Secondary stakeholders are the system's development and maintenance team. Tertiary stakeholders encompass the application owner, developers of competing applications, and furthermore, the marketing and management teams of the developing company.

Scenarios

To further understand typical users, we initially inquired with people we know about their motivations for using scanning applications. Based on the responses and our experience, we outline below some typical use case scenarios for scanning applications.

Sending a Doctor's Prescription

This scenario arose from a neighbor who mentioned to us that when she was sick with flu, she needed to send to her daughter a doctor’s prescription, and she had asked us if we knew a good scanning application.

Kaiti is retired and visited her doctor to write her a prescription for the otitis that she has and is bothering her the last week. He told her what medicines to take and gave her a prescription, but because there is no pharmacy in the neighborhood and Kaiti does not drive, she wants to send the prescription to her daughter to get the medicines. For this reason, Kaiti used a scanning application to digitize and send the prescription to her daughter. However, it needed many attempts for her to succeed because her hands tremble and the photos were blurry.

Concluding an Agreement

This particular scenario arose from an internet search for the reasons why an employee uses a scanning application.

Charalampos closed a large contract with his client for the supply of computer equipment and wants to send the signed documents to his sales department manager for a copy. In fact, it is important for him to have a specialized application on his mobile phone that has a nice and user-friendly design to scan the contracts and send them. Specifically, he desires the ability to scan multiple pages and organize the files so that he can distinguish the different contracts he has and organize them according to his clients.

Submitting Homework

This scenario has arisen from our own participation in various distance learning exams.

Panos is a student studying philosophy remotely and wants to submit his answers for the exam in the course "Ethics and Science" within the next 30 minutes. He had used a different application in the past during online exams, but he has uninstalled it. Now he has installed a new application and wants to immediately scan 8 pages of A4 size and combine them into a pdf file, which he can rename and send quickly.

Changing Electricity Provider

This scenario comes from everyday life, as our parents are changing electricity providers.

Marina wants to change electricity providers because the provider she is currently with increased prices for the next month, and she needs to send a copy of her ID card. Because she does not have a scanner at home, Marina installs a scanning application on her mobile phone. She wants the application to produce good results even for smaller documents, such as an ID card, and at the same time, for all her personal details to be clearly visible.

Personas

Based on the analyzed scenarios, we developed personas representing typical users.

As previously mentioned, Kaiti requires the application to produce legible text and compensate for hand tremors in photos.

Figure 1: Kaiti, a retiree persona.

Charalampos prioritizes document clarity and high quality above all else.

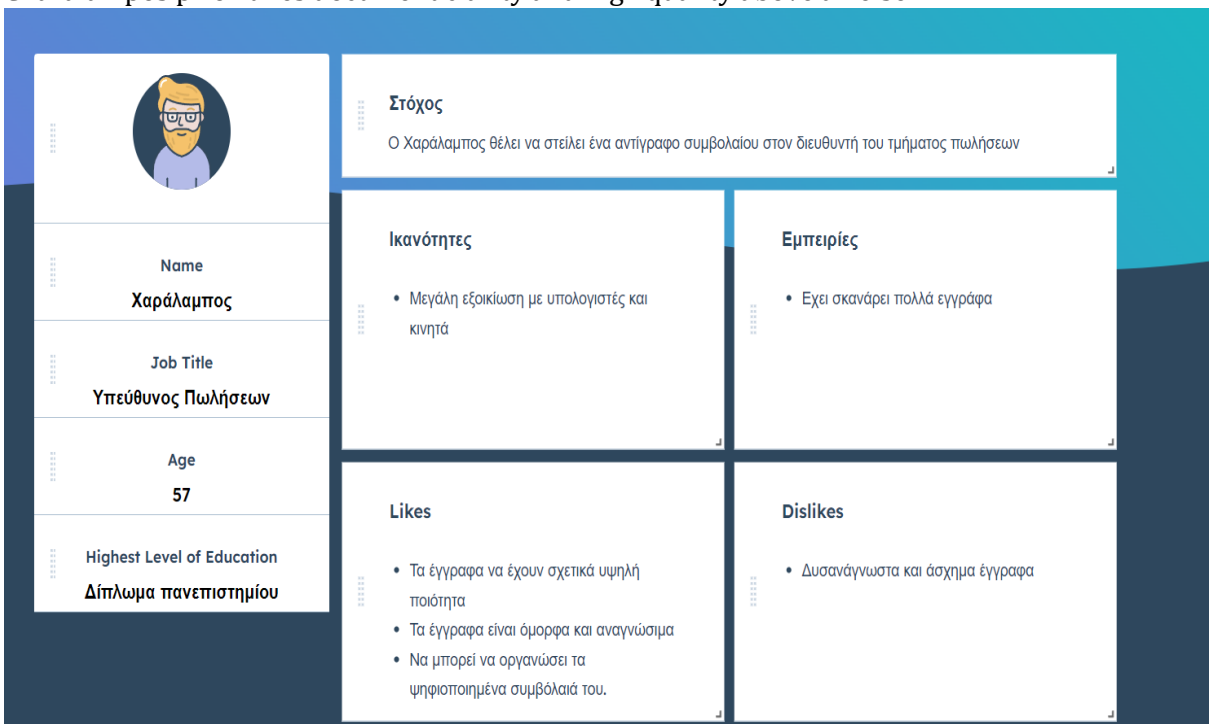


Figure 2: Charalampos, a professional persona.

Panos, pressed for time to scan his exam work, needs to quickly understand the application's functionality.

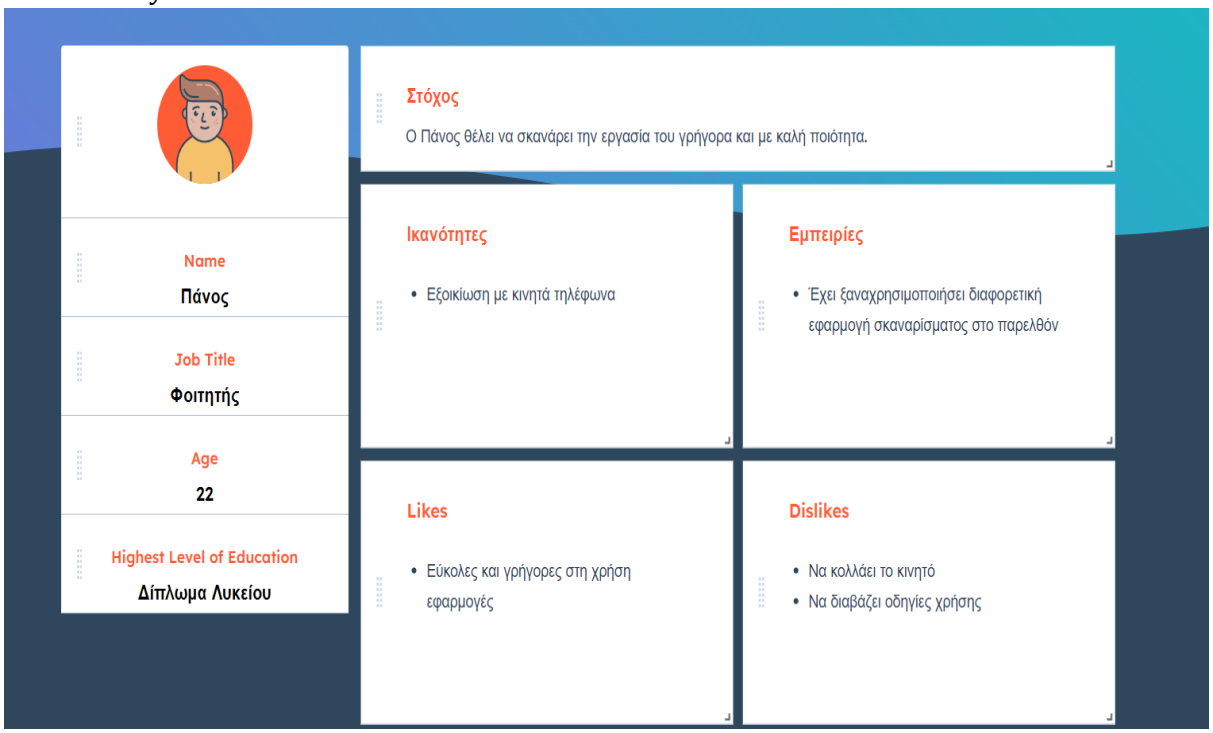


Figure 3: Panos, a student persona.

Marina values both the aesthetic appeal and legibility of her digitized ID.

Figure 4: Marina, an adult persona.

4) Task analysis

Context of Use

Within the task analysis framework, we initially explore how typical users interact with the application. As observed from the use case scenarios, individuals usually work in a space with a flat surface and sufficient room, such as a desk or table. Furthermore, users typically employ the application to send a digitized document, which they have in front of them, to a third party. This document often holds significant importance for either the user or the recipient; examples include a medical prescription, a contract, an identification card, or a student's exam answers. Notably, while some users may not be under time pressure, others might face strict deadlines for sending the document.

PACT Framework

The PACT framework (People, Activities, Context, Technologies) considers the individuals who will use the system, their actions within it, the context of use, and the technologies employed (Nis, 2023). In the user analysis section, we defined the typical users of the system and their actions. Additionally, in the context of use section above, we outlined the conditions under which the system is utilized. Regarding the technology underpinning the system, it is clear that user interaction with the application occurs through a mobile phone or tablet. This is necessary because the device must be easily portable, enabling users to photograph the document they wish to scan.

Indeed, examining the influence of technology on user experience is crucial. The application is exclusively used via mobile phones or tablets for scanning documents using the device's camera. Consequently, the quality of scanned documents heavily depends on the user's camera quality. This means the same application might create a positive impression for some users and a negative one for others, depending on their camera capabilities. Therefore, the technology utilized by the scanning application should ideally not assume users possess high-end cameras, ensuring the application operates smoothly across a wide range of devices.

Hierarchical Task Analysis

To understand the specific actions a user needs to perform when using the application to scan a handwritten document, we employ hierarchical task analysis. This involves breaking down the task into elemental sub-actions, as illustrated in Figure 5 below.

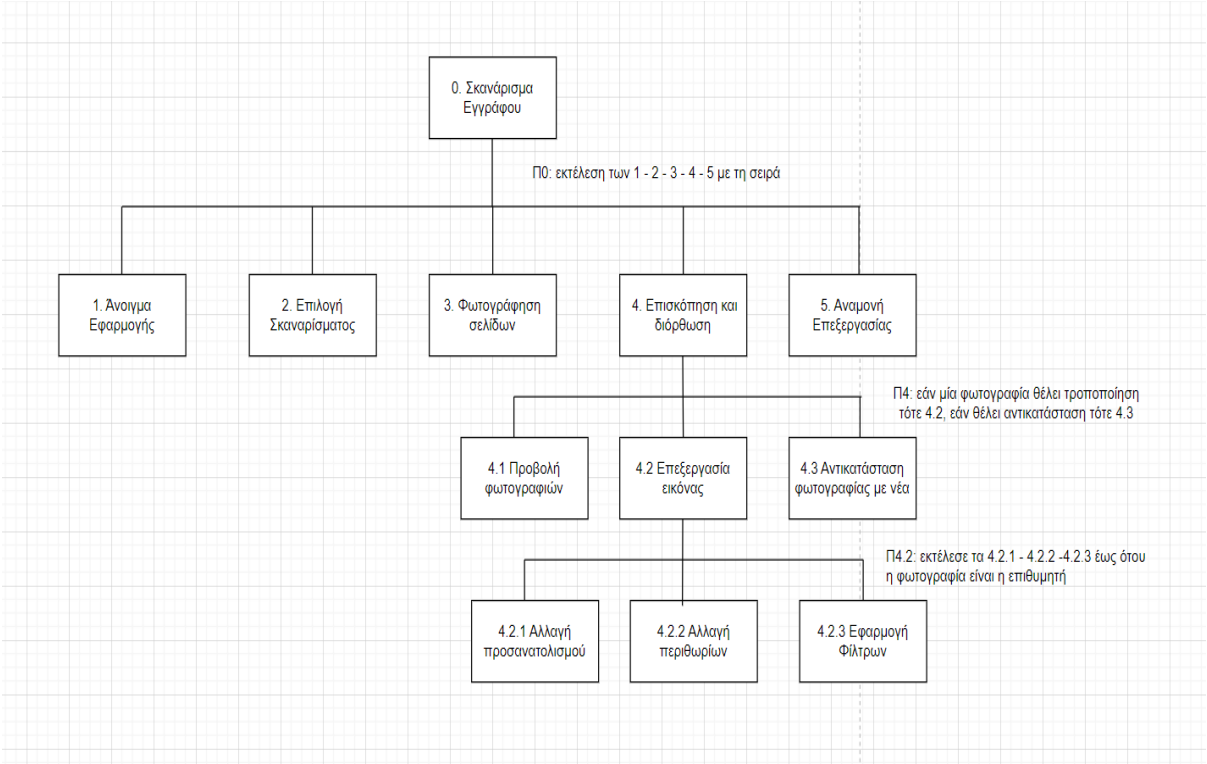


Figure 5: Hierarchical Task Analysis

We observe that an ideal application, to be effective and user-friendly, should offer users the option to select document scanning upon opening, or have it as the default option. Subsequently, the user should be able to photograph the document they wish to digitize and make necessary corrections. These corrections can involve either editing a photograph or replacing it with a new one. Finally, the user expects to wait until the document digitization process is complete.

Requirements

Functional Requirements

Functional requirements are defined as the functionalities we want the system to possess – essentially, what we want the system to do. Based on the user analysis, scenarios, and personas we developed, we specifically conclude that the functional requirements include the production of high-quality digitized documents, file organization, file compression, document editing, document sharing, and document protection with a password to prevent unauthorized access from anyone who might gain access to the mobile device. Furthermore, important functional requirements are the ability to replace a digitized page with a new one and the removal of hand tremor or hand shadow effects from the scanned photograph. Ideally, character recognition

(OCR) capability should be included for producing text documents, followed by the option to translate them into other languages. Additionally, the application could offer cloud storage functionality to ensure documents are accessible from other devices. We categorize these requirements based on their importance in the categories below:

Must-have:

- Ability to scan and produce high-quality digitized documents.
- Ability to replace a digitized page with a new one.
- Ability to share documents.
- Ability to organize documents.

Should-have:

- Removal of hand tremor effects.
- Removal of hand shadow effects.
- Document password protection.

Could-have:

- Compression capability for efficient document storage.
- Cloud storage capability.

Desired but not for initial release:

- Character Recognition (OCR) capability.
- Document translation capability to other languages.

Non-Functional Requirements

Non-functional requirements define how the application should be designed. Specifically, based on the user analysis, we conclude that the non-functional requirements include: the application being responsive across devices with varying screen sizes to ensure usability on mobile phones, and the application being easy to use, allowing first-time users to immediately understand how to operate it. Additionally, it is important for the language used within the application to be simple and accessible, ensuring comprehension even for older users. Finally, the application should be aesthetically pleasing, featuring an attractive color scheme and a well-designed interface. Lastly, some users would appreciate the ability to personalize the application. We subsequently categorize these requirements based on their importance in the categories below:

Must-have:

- Application responsiveness across devices with different screen sizes.
- Immediate understandability for new users regarding application usage.
- Simple and understandable language.

Should-have:

- Attractive color scheme.
- Well-designed interface.

Desired but not for initial release:

- User personalization options for the application.

5) Evaluation

Next, we evaluate the Microsoft Lens scanning application. First and foremost, we define the objectives of the evaluation. This evaluation is primarily formative, meaning the aim is to derive specific suggestions for improving the application, which we will subsequently use for its redesign. Naturally, we also wish to ascertain whether our redesigned version is superior, and therefore, we will also conduct a comparison of the applications using summative evaluation.

Moreover, the context of application use and the technology to be employed have been defined during the task analysis. Additionally, the typical users participating in the application evaluation will correspond to the typical users we defined in the user analysis.

Furthermore, the evaluation will be conducted both by us, making it expert-based, and by typical users, thus also being user-based. For the formative evaluation, we will utilize both heuristic evaluation, an expert-based method, and the think-aloud protocol, to gather user feedback on their interaction with the application. To assess the design quality of Microsoft Lens and perform summative evaluation from a user perspective, we will employ questionnaires, specifically the UEQ questionnaire. Finally, to compare the application with our redesign, we will use the KLM method, a summative expert-based evaluation technique.

Expert-Based Evaluation

Initially, an expert-based evaluation is conducted to pinpoint the specific issues of the application.

1. Heuristic Evaluation

Heuristic evaluation is a formative evaluation method relying on experts (Nielsen, 1993). There are ten heuristic principles, and if one is violated, a severity rating is assigned, ranging from 0, if it is not considered a problem, to 4, if it is deemed catastrophic for the user experience. A rating of 1 indicates a purely aesthetic problem, 2 a minor usability issue, and 3 a major usability problem. Specifically, these heuristic principles have emerged from the experience of numerous designers and are as follows:

1. Visibility of system status: The system should consistently inform users about its status, keeping them aware of what is happening.
2. Match between system and the real world: The application should use language familiar to its users.
3. User control and freedom: If a user makes a mistake, they should be able to easily exit the erroneous state.
4. Consistency and standards: The application should maintain consistent design characteristics across different screens.

5. Error prevention: The system should facilitate the prevention of user errors.
6. Recognition rather than recall: The system should enable users to recognize what they need to do at each step, minimizing the need to remember information.
7. Flexibility and efficiency of use: The system should be adaptable to user needs and efficient, allowing experienced users to quickly achieve their goals, while also ensuring ease of use for novice users.
8. Aesthetic and minimalist design: The system should be aesthetically pleasing and free from unnecessary elements.
9. Help users recognize, diagnose, and recover from errors: The system should recognize user errors and facilitate their correction.
10. Help and documentation: The application should provide adequate user instructions, if deemed necessary.

Both members of our team independently conducted heuristic evaluations. Subsequently, we discussed our findings to consolidate our results. It is worth noting that the problems we identified were largely common, with only slight differences in the assigned severity levels, for which we calculated an average.

The violations of the heuristic principles are as follows:

A)

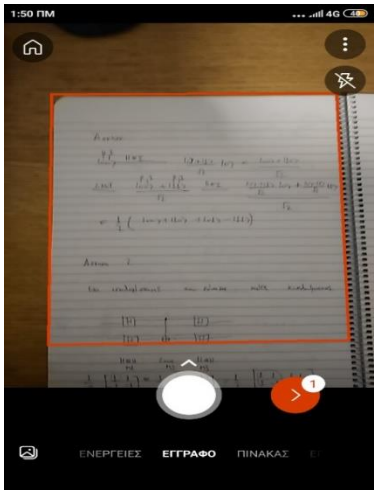


Figure 6: Red frame searching for page margins and changing rapidly

When photographing a page, the application displays a red rectangle indicating the area of the page that will be within the margins. When the mobile phone's camera is not of high quality, the application frequently misidentifies the page boundaries. Moreover, even if the user moves, the frame may change but still be incorrect. At times, the frame changes on its own without user movement. If these issues occur repeatedly, the user will have a poor experience and may choose to use another application. Furthermore, when the automatic margin recognition algorithm does not function correctly, the user is forced to take the photo while seeing that the margins are incorrect. This violates heuristic principle 5, which states that the system should help prevent

users from making errors. The severity of this problem is significant and is considered a major usability issue, especially for users who do not have devices with good cameras. Therefore, it is crucial to rectify this. To correct this error, the margins should not change so rapidly, and in terms of design, users should be given the option to disable the automatic margin recognition frame.

B)

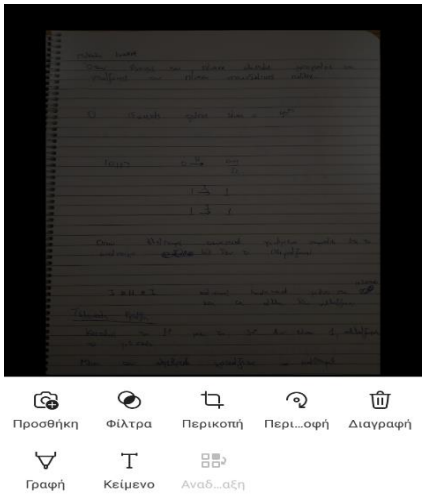


Figure 7: Image editing during preview

After the user finishes scanning pages and is in the preview phase, if they have made a mistake and decide they want to change a photo and replace it with another, they cannot easily do so, as evident from the editing options presented in Figure 7. For instance, suppose a user has scanned a 20-page document and realizes that page 3 is blurry. Ideally, there would be a "retake" button. When the user is on page 3 and presses it, they would be automatically taken to the scanning screen, and the next photo the user scans would automatically be placed in position 3, replacing the previous blurry photo. However, this specific option is not available in the application's menu. Consequently, the user must press the back button to return to the scanning screen and take a new photo, which will be placed at the end of the document. Subsequently, they must delete the original blurry photo and rearrange the images to position the new photo in position 3. This violates heuristic principle 7, regarding flexibility and efficiency of use, specifically shortcuts for frequent user actions. Indeed, this has a significant impact on those who make mistakes and is therefore considered a serious usability problem that needs to be addressed. The correction can be easily made by adding the option to replace a photo with a new one during the preview phase.

C)

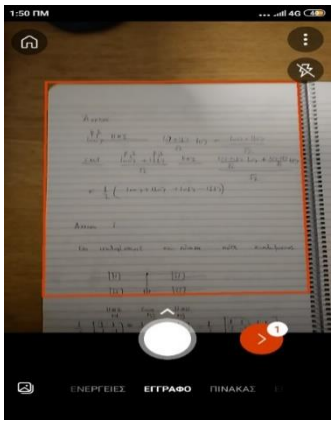


Figure 8: Intense red rectangle

The rectangle shown in Figure 8, indicating the selected margins of the photo during scanning, has an intense color and is particularly noticeable. Therefore, rule 8 regarding minimalism and avoiding unnecessary elements is violated. However, this is considered merely a minor aesthetic problem. To correct this, designers could give the user the option to hide the red rectangle. During preview, users can check if the margins have been selected correctly and adjust them themselves. Of course, if the margin selection algorithm functions correctly, there will be no need for frequent margin modification.

D)

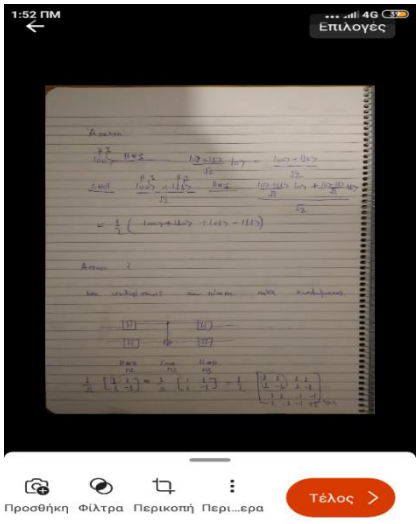


Figure 9: Menu to which the user is automatically taken after taking one photo

When the user takes the first photo, they are automatically taken to the preview menu, shown above. Furthermore, if the user wants to scan a document with multiple pages, they must press "Add" and continue scanning pages. However, they are not given the option to specify that they want to scan multiple pages before this menu opens. Consequently, rule 7 regarding flexibility is violated due to the absence of shortcuts for users who want to choose to scan multiple pages together. This is a significant usability problem as it takes more time to scan a document, resulting in an unpleasant user experience. To correct this problem, an option could be added during

scanning to specify that the user wishes to take multiple photos, preventing automatic redirection to the preview menu.

E)

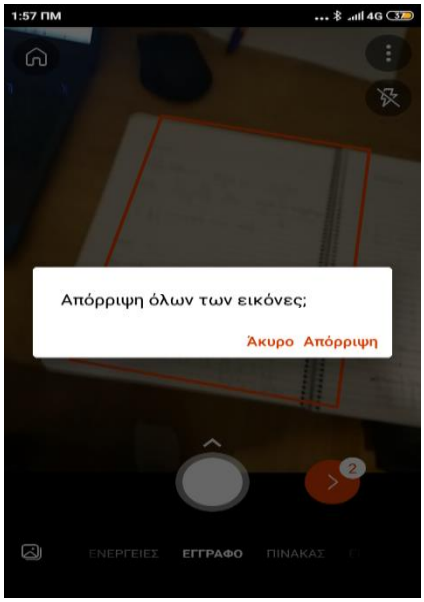


Figure 10: Termination of scanning

If the user is scanning a multi-page document and presses the "back" button on their mobile phone, the above message appears. Logically, based on the message displayed, the user expects to discard the photos and return to the scanning screen. However, the back button does not take them to the scanning screen but closes the application. Therefore, rule 3 is violated, as a clear and understandable exit is not provided to the user; they expect to return to the application's scanning screen, but instead, it closes completely. This is considered a minor usability problem. To correct this, the back button should either return them to the scanning screen, or the exit message displayed to the user should be changed to clearly indicate that the application will close.

F)

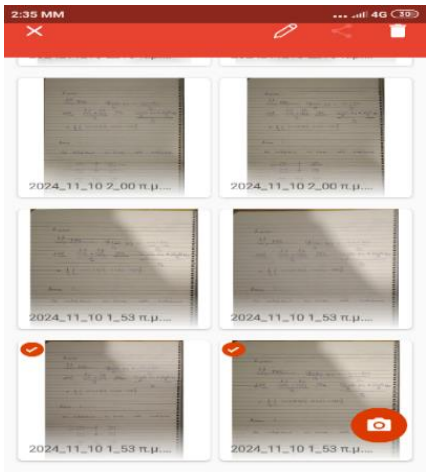


Figure 11: Sharing of two or more pages is not possible

Suppose a user has scanned a five-page document and saved it as photos (choosing the "save to gallery" option). The user may wish to send some of the scanned pages to someone. However, the application only allows sending one page at a time. As shown in Figure 11, if someone selects two photos, the share symbol is disabled. This violates rule 7 regarding adaptability to the various needs of users. However, this is considered a minor usability problem as the user can save the document in PDF format and send it all at once.

G)

As seen in Figures 9 and 11, some pages in the application have a dark background, while others have a light white background. This violates rule 4 regarding maintaining consistency across pages. Indeed, this is considered simply an aesthetic problem. To correct this, a slightly darker color, such as gray, could be chosen for the pages with a white background.

H)

A user who wishes to rearrange pages must first press the "More" button shown in Figure 9, and then the desired option will appear. Thus, the rearrangement option is located in a relatively inconspicuous place, and the user must remember where to find it. This violates rule 6 regarding minimizing the user's memory load. Indeed, it is judged to be a minor usability problem. Because the page rearrangement option is something many users utilize, the option could be placed in the main menu shown in Figure 9, without requiring users to press "More."

2. KLM/GLM Evaluation

The KLM (Keystroke Level Model) evaluation method originates from the Human Processor model (Avouris et al., 2018). We will use this method to calculate the time it takes for a user to scan a three-page document and save it as a PDF named "myfile" using the Microsoft Lens application. Subsequently, after designing our own scanning application, we will compare these times. Given that this is a mobile application, we will utilize the Gesture Level Model (GLM) variant (Nyström, 2018). The operators, their descriptions, and corresponding model times are shown below.

Operator	Description	Time (sec)
P	Finger movement	0.34
T	Tap on screen	0.73*
D	Drag element	0.44
M	Mental thought	1.35
H	Focus on different screen area	0.4
X	Distraction	Minor: +6% of Total Time Major: +21% of Total Time

Table 1: GLM Model Response Times (Typing time was calculated from experimental tests by team members)*

For the T operator, we measured our own typing times using a specialized application (*Free Mobile Typing Speed Test for Your Smartphone - Find out Your Wpm on IPhone and Android*, 2024). The average typing times we measured for the two team members were 1.2 seconds and 0.25 seconds, highlighting the significant variation in typing speeds between individuals. Therefore, for our analysis, we take the average of these two measurements and select T= $\frac{1.2+0.25}{2} = 0.73sec$.

The process we will examine involves a user scanning 3 photos with default dimensions and options, and finally saving the file as a PDF named "myfile."

Initially, the user opens the application and scans the 1st page, as shown in **Figure 12**. These steps correspond to the following elementary actions in the GLM model:

M: Open initial menu

H: Focus on the photo button

P: Finger movement

T: Tap photo button

Figure 12: Application initial menu

After the first photo, the application automatically opens the photo preview menu, shown in Figure 13.

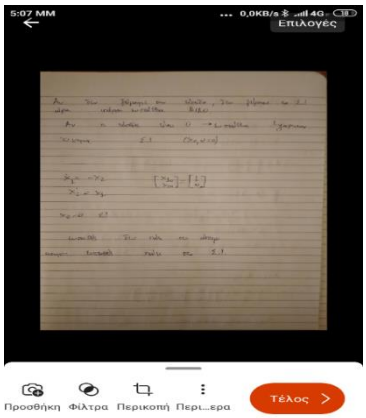


Figure 13: Preview menu opens automatically after the first photo

Next, the user chooses to add another photo to return to the photo-taking screen in Figure 12 and continue scanning. Therefore, the following actions occur:

M: Preview menu opens

H: Focus on the add photo button

P: Finger movement

T: Tap button

M: Photo-taking screen of Figure 12 opens

H: Focus on the photo button

P: Finger movement to photo button

T: Tap button to take the 2nd photo

After the first photo, the preview menu does not open automatically, as it did initially. Therefore, the user continues to photograph the third page and complete the scanning process, as shown in Figure 14.

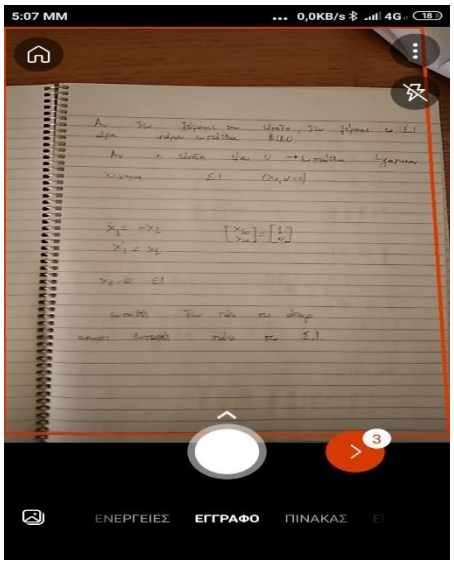


Figure 14: User has scanned the desired 3 pages

Consequently, the user performs the following actions:

H: Focus on the photo area to take the 3rd photo

H: Focus on again the photo button

T: Tap photo button for the 3rd photo

M: Think about how to proceed

H: Focus on the circle with the photos (red circle with the number 3)

P: Place fingers over the circle

T: Tap the circle with the photos

The user is then taken back to the preview menu of Figure 13. This time, they wish to proceed with completing the scan, so they perform the following actions:

M: Preview menu appears

H: Focus on the "Done" button

P: Finger movement

T: Tap "Done" button

Subsequently, they are taken to the save menu shown below and wish to change the default name and rename the file to "myfile."

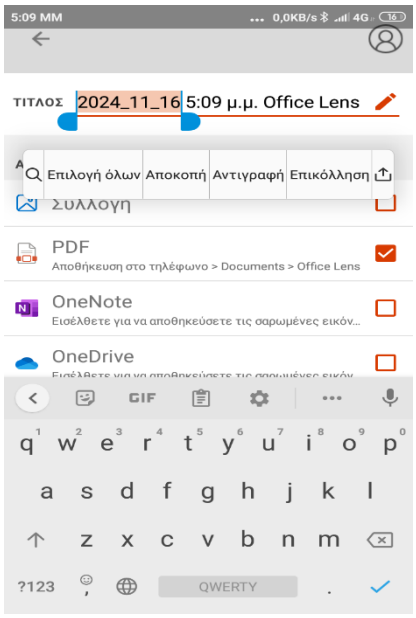


Figure 15: Save scanned file menu

The default name contains spaces (whitespaces), so when the user selects it to delete it, only one word is selected, and they must then drag the selectors to select the entire name. Therefore, the following actions occur:

M: Save menu opens

H: Focus on the name

P: Finger movement to the filename

2T: Long tap on the name. Because the default name has spaces (whitespaces), only one word is selected

D: Drag the right selector to select the entire name

H: Focus on the delete selected name button on the keyboard

P: Finger movement

T: Tap delete selected name

M: Think of a new name ("myfile")

6T: Type "myfile"

T: Tap "ok" for the name

Finally, the user chooses to save it in PDF format. Therefore, the actions are as follows:

M: Think about the desired save option

H: Focus on the save options

P: Finger movement

T: Deselect "Save to gallery"

T: Select "Save as PDF"

M: Think about how to proceed

T: Tap "Save"

The actions required for the user to scan a three-page document are summarized in Table 2.

Operator	Times Applied	Required Time (sec)
M	9	9*1.35=12.15
H	10	10*0.4=4
P	8	8*0.34=2.72
T	19	19*0.73=13.87
D	1	0.44

Table 2: Required Times for Scanning with Microsoft Lens Application

Therefore, the user needs a total of 12.15 + 4 + 2.72 + 13.87 + 0.44 = 33.18 seconds to scan a three-page document.

Assuming they are focused when scanning and their attention is not easily distracted, the distraction factor is 6%. Therefore, the required time is:

$$T = 1.06 * 33.18 = 35.17\ sec$$

In conclusion, the time we theoretically calculated using the GLM model is 35.17 seconds. Indeed, we conducted experiments to scan three pages with the application and observed that approximately this much time is required, which confirms the accuracy of our calculations. Finally, the process of calculating the scanning time will be repeated for the application we design to compare the applications.

User-Based Evaluation

For the user-based evaluation, we will utilize the think-aloud protocol, which is a formative evaluation method, and the questionnaire method, which is summative. Specifically, we will use the User Experience Questionnaire (UEQ) to evaluate whether typical users consider the Microsoft Lens application to be well-designed. Notably, the questionnaire is freely available online, along with tools for analyzing the results (Hinderks et al., 2018).

1. User Experience Questionnaire

In our study, N=10 typical users participated. Specifically, we asked them to scan three pages using the Microsoft Lens application and then answer the questionnaire shown in Figure 16. The questionnaire was created and distributed online using the Google Forms tool.

	1	2	3	4	5	6	7		
ενοχλητικό	○	○	○	○	○	○	○	απολαυστικό	1
δυσνόητο	○	○	○	○	○	○	○	κατανοητό	2
δημιουργικό	○	○	○	○	○	○	○	αναποτελεσματικό	3
εύκολο στη μάθηση	○	○	○	○	○	○	○	δύσκολο στη μάθηση	4
πολύτιμο	○	○	○	○	○	○	○	υποδεέστερο	5
βαρετό	○	○	○	○	○	○	○	συναρπαστικό	6
αδιάφορο	○	○	○	○	○	○	○	ενδιαφέρον	7
απρόβλεπτο	○	○	○	○	○	○	○	προβλέψιμο	8
γρήγορο	○	○	○	○	○	○	○	αργό	9
εφευρετικό	○	○	○	○	○	○	○	συμβατικό	10
παρελκυστικό	○	○	○	○	○	○	○	υποστηρικτικό	11
καλό	○	○	○	○	○	○	○	κακό	12
περίπλοκο	○	○	○	○	○	○	○	εύκολο	13
αντιπαθητικό	○	○	○	○	○	○	○	συμπαθητικό	14
συνηθισμένο	○	○	○	○	○	○	○	πρωτοπόρο	15
δυσάρεστο	○	○	○	○	○	○	○	ευχάριστο	16
ασφαλές	○	○	○	○	○	○	○	ανασφαλές	17
ενθαρρυντικό	○	○	○	○	○	○	○	αποθαρρυντικό	18
ανταποκρίνεται στις προσδοκίες	○	○	○	○	○	○	○	δεν ανταποκρίνεται στις προσδοκίες	19
ανεπαρκές	○	○	○	○	○	○	○	επαρκές	20
σαφές	○	○	○	○	○	○	○	μπερδεμένο	21
μη πρακτικό	○	○	○	○	○	○	○	πρακτικό	22
οργανωμένο	○	○	○	○	○	○	○	ανοργάνωτο	23
ελκυστικό	○	○	○	○	○	○	○	απωθητικό	24
φιλικό	○	○	○	○	○	○	○	εχθρικό	25
συντηρητικό	○	○	○	○	○	○	○	καινοτόμο	26

Figure 16: User Experience Questionnaire

This specific questionnaire consists of 26 pairs of characteristics, one with a positive connotation and the other with a negative one. Furthermore, the position of the positive answer, whether on the right or left, varies from question to question to avoid responses that are all skewed towards one end of the scale, e.g., all positive. The questions in the questionnaire are divided into six scales that measure: Attractiveness, Clarity-Understandability, Efficiency, Reliability, Stimulation-Excitement, and Novelty. Additionally, the questions are on a 7-point Likert scale corresponding to scores from -3 for the most negative answer to +3 for the most positive. For each scale, we calculate the mean user response, the standard deviation, as well as the 95% confidence intervals. A score in a scale greater than 0.8 corresponds to a positive evaluation, a score from -0.8 to +0.8 to a neutral evaluation, and a score less than -0.8 to a negative evaluation (Hinderks et al., 2018).

The results of the statistical processing of user responses are shown in the table below.

Category	Mean Value	Standard Deviation	95% Confidence Interval	
Attractiveness	-0,400	0,48	-0,828	0,028
Clarity	0,450	2,11	-0,450	1,350
Efficiency	0,150	0,41	-0,245	0,545
Reliability	-0,100	1,20	-0,778	0,578
Stimulation	-0,375	1,78	-1,202	0,452
Novelty	-0,125	0,89	-0,710	0,460

Table 3: User Evaluation Results

We observe that all scales have a mean value within the range [-0.8, +0.8], which corresponds to a neutral evaluation. This implies that, on average, users did not consider the application to be either well-designed or poorly designed. Furthermore, the standard deviation informs us about the likelihood of a response differing from the mean value (a small standard deviation means that all responses are very close to the mean). The largest standard deviation is observed in Clarity, indicating that some users found the application easy and understandable, while others did not. However, users generally agree on a neutral evaluation for the Attractiveness and Efficiency scales, as the standard deviation is small, and the mean value is also low, just -0.4 and 0.15 respectively. Finally, the 95% confidence interval shows the range of values within which 95% of the results are likely to fall concerning the mean value of each scale. These statistics were calculated using the tools provided on the website of the questionnaire creators (Hinderks et al., 2018).

These results are visualized in the graph below.

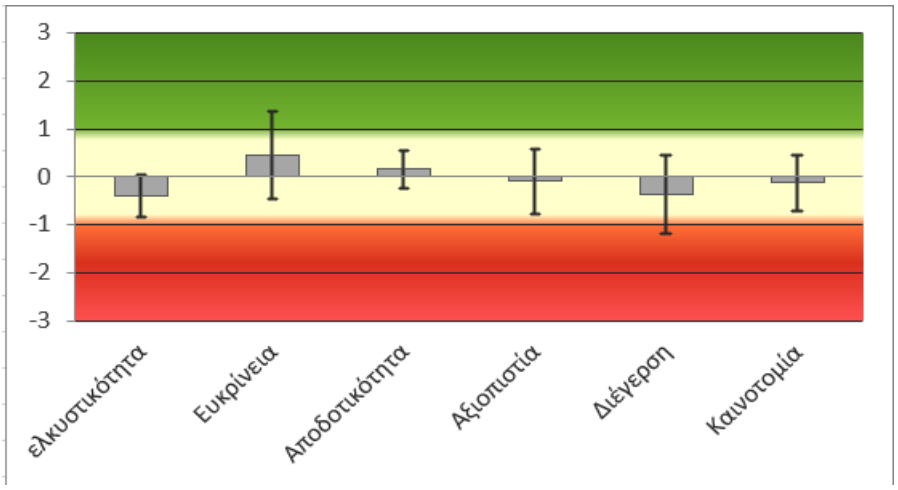


Figure 17: Mean value and variance for each scale

Additionally, the creators of the UEQ questionnaire allow for the comparison of our evaluation results with a database containing approximately 470 evaluations of various technology

applications, involving more than 21,000 participants. The comparison of our application in relation to the evaluations of other applications is shown in Figure 18.

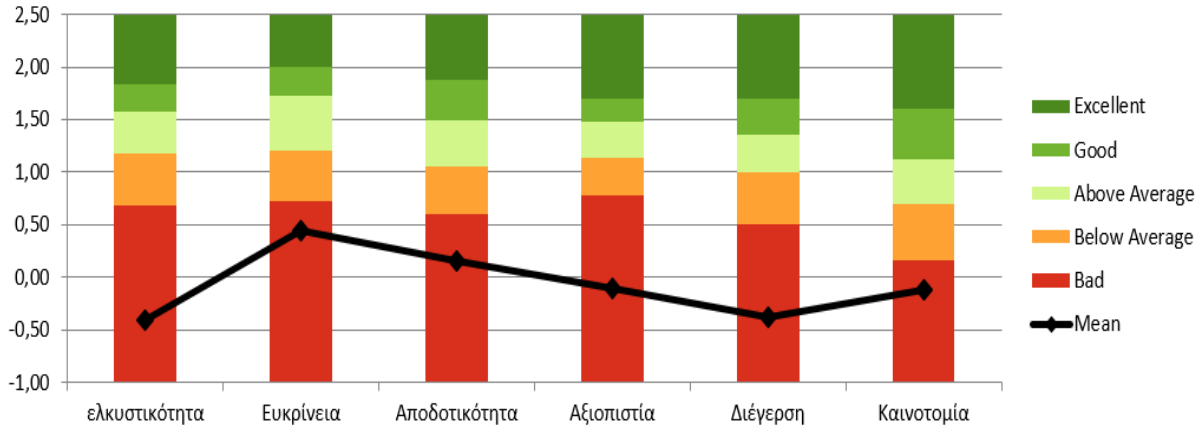


Figure 18: Microsoft Lens application score compared to other application evaluations.

Indeed, we observe that in all scales, the application is considered poorly designed, falling into the 25% of the worst evaluations. However, this may not fully reflect reality, as the above comparison is made across a broad range of applications and not solely scanning applications. The results presented in Figure 17, which reveal that users, on average, considered the application to be neutrally designed, are likely more representative of the actual user perception.

Therefore, the evaluation method using the User Experience Questionnaire confirmed that users believe the application could be improved. To understand specifically what difficulties they encountered with the application, we will subsequently use the think-aloud protocol.

2. Think-Aloud Protocol

In the think-aloud protocol, users spontaneously verbalize their thoughts while using the system (Nielsen, 2012). This method allows us to understand where they encountered difficulties and when unexpected events occurred.

Specifically, for our application, we asked users to scan 3 pages and save them as a PDF. While they were using the application, they verbalized their thoughts to us, and we recorded them. The feedback we received is shown in Table 4.

User	Comments	When they said it
1	This frame is annoying.	When they saw the red frame during scanning (Figure 6)
	Now what do I do?	When the application, after the first photo, automatically took them to the preview menu, but they wanted to scan more photos (problem D in the heuristic evaluation).
2	The frame is interesting.	When they saw the red frame during scanning (Figure 6)
3	The frame changes very quickly	When they saw the red frame during scanning (Figure 6).
	Where do I tap to finish scanning?	When they wanted to stop scanning (red circle Figure 6).
4	The filters are understandable.	Όταν προσπάθησαν να προσθέσουν φίλτρα, τα εικονίδια που εμφάνιζαν το αποτέλεσμα βοηθούσαν στην επιλογή κατάλληλου φίλτρου.
	How does page reordering work?	When they tried to change the order of pages in the program, the lack of instructions confused the user.
	The red frame moves too much.	When they saw the red frame during scanning (Figure 6).
5	The options are very understandable.	When they saw the different photo-taking options and the various image editing tools.
	The filters help with readability.	When they added a filter to a slightly blurry image to make it more readable.
	The page does not adjust well.	When they saw the result of the red frame after photographing the page.
6	The process is very simple.	After completing the scanning process, they commented on the simplicity and directness of the process.
	Inability to send more than one file.	When they tried to send 2 files simultaneously, the application did not allow it.

7	Very minimalist and simple application	Regarding the overall appearance of the application.
	Strange location for the delete button.	When they saw the delete button in image editing.
	The red frame moves intensely.	When they saw the red frame during scanning (Figure 6).

Table 4: User Comments from the Think-Aloud Protocol.

We observe that users agree with the problem highlighted by the heuristic evaluation, namely that the application automatically takes them to the preview menu after the first photo, even though they may want to scan more documents. Also, some users disliked the red frame, while others found it interesting. Additionally, one user had trouble understanding what to do to stop scanning. Therefore, these problems should be taken into account during the redesign.

6) Design

Subsequently, based on the results of the analysis and the errors we identified during the expert and user evaluations, we will design a new scanning application. In fact, the application we analyzed in Phase A, Microsoft Lens, is by Microsoft, and therefore we are assuming that we are designing a new scanning application *for* Microsoft. Initially, both members of our team created paper sketches to determine the design that we believe best addresses the requirements of our application. Following this, the design will be implemented using the Figma tool (www.figma.com), which allows both team members to simultaneously work on the design.

We initially created three sketches for the scanning screen, each with its own advantages and disadvantages. One team member created the first and third designs, and the other created the second.

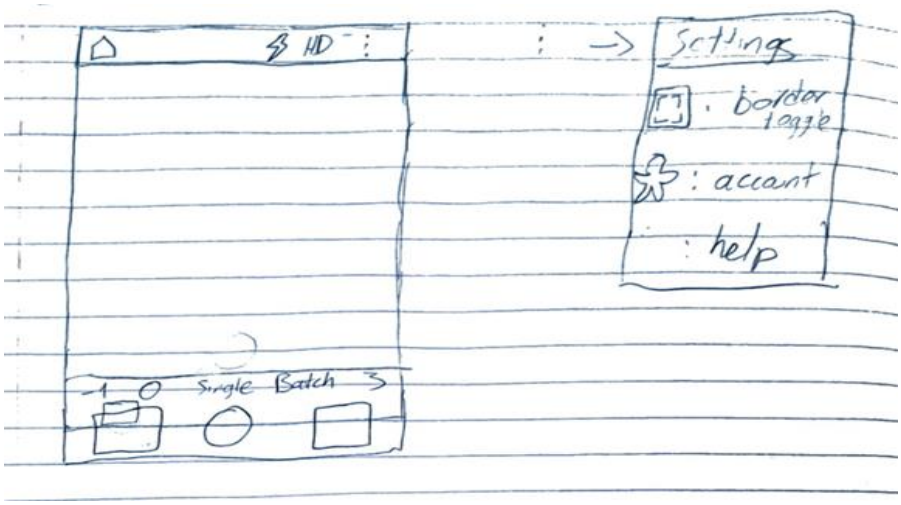


Figure 19: Sketch 1

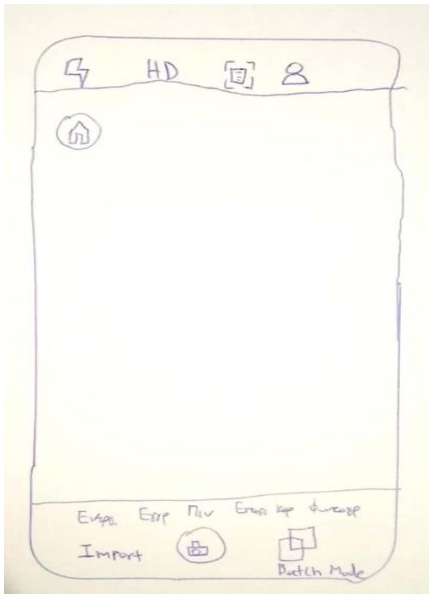


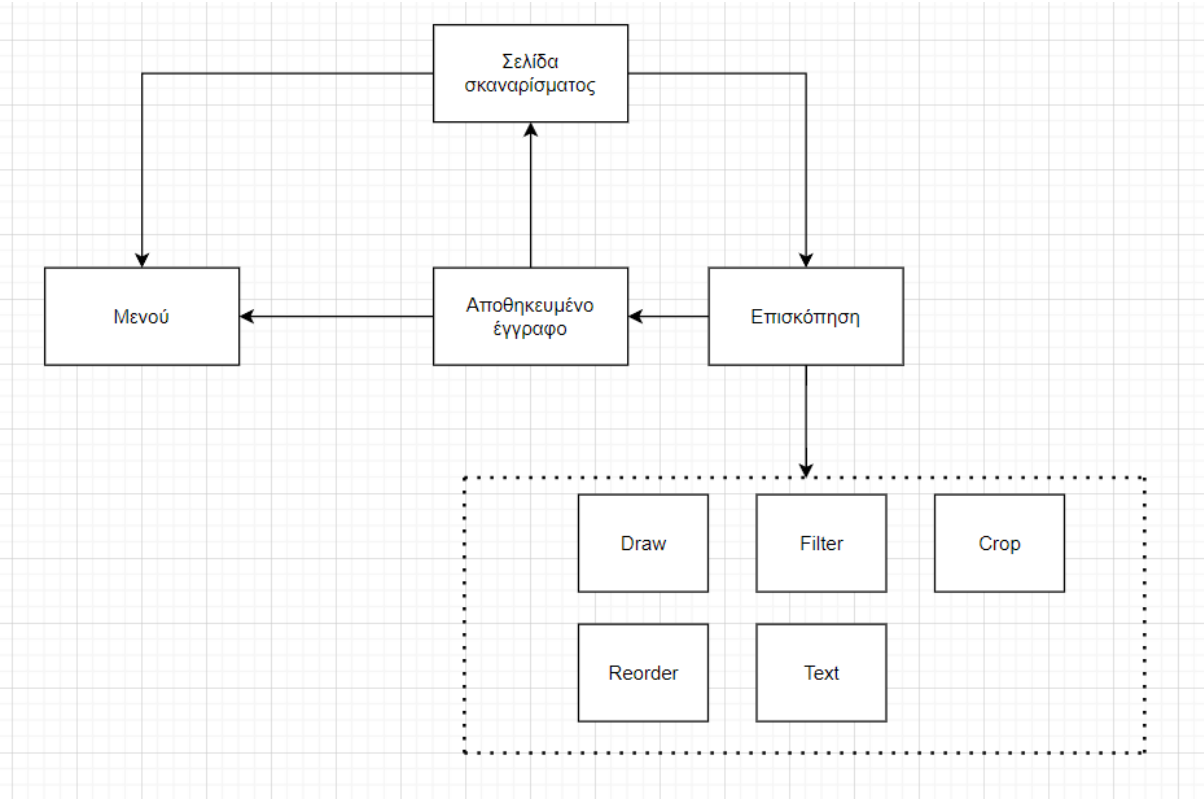
Figure 20: Sketch 2

Figure 21: Sketch 3

Sketch 2 features a circle with a camera icon for the photo button, while the other two sketches have only a white circle. In Sketch 3, the photo button is located outside the frame containing the options, whereas in Sketch 1, it is inside. Furthermore, Sketch 2 includes a separate "Batch Mode" button, indicating that a user wants to scan a multi-page document, while in Sketches 1 and 3, this functionality is grouped with the options. The options are the same as in the Microsoft Lens application, namely: options for whiteboard, business card, document, or photo.

Ultimately, we decided to have a separate button for "Batch Mode" and not group it with the other options. This is because regardless of *what* someone wants to scan (document, business card, etc.), they may need to take multiple photos, and therefore they would select the "Batch Mode" option. If "Batch Mode" is not selected, the user intends to scan only a single page and is thus automatically taken to the preview after scanning the page. However, if they *have* selected "Batch Mode," they continue scanning. This resolves problem D) that we identified during the heuristic evaluation and that several users mentioned in the think-aloud protocol. Additionally, for aesthetic reasons, we decided that the photo button should be within the frame containing the options, and all icons should be within the upper icon frame.

Moreover, the navigation map of our application is shown below.



Schema 1: Application Navigation Map

When the user opens the application, they are on the scanning screen, where they can scan a document. From this screen, they can directly access the application menu where all saved documents are located. Also, once the user completes scanning, they are taken to the preview screen, where they can check if the photos are satisfactory and edit them. If they want to add or change a photo, they can, of course, return to the scanning screen (the transitions are bidirectional; the arrows describe the typical sequence in which a user would visit the pages). On the preview page, the user should have the ability to edit the image, specifically to apply filters, adjust margins, rearrange images, add text, or even draw. After completing the preview, the user will save the document and be taken to another page where they will see the saved document. Finally, after saving the document, the user can return to the scanning screen to add another photo, or to the menu to view all the documents they have saved.

Below, we present mockups and some specific functionalities for each page. The prototypes have all the functions and transitions detailed and are available at the following link: https://www.figma.com/proto/DEWXOAiI9v0SlVEOiWMIV3/Scanner-App?node-id=0-1&t=IKaZQpYARTGxII1s-1

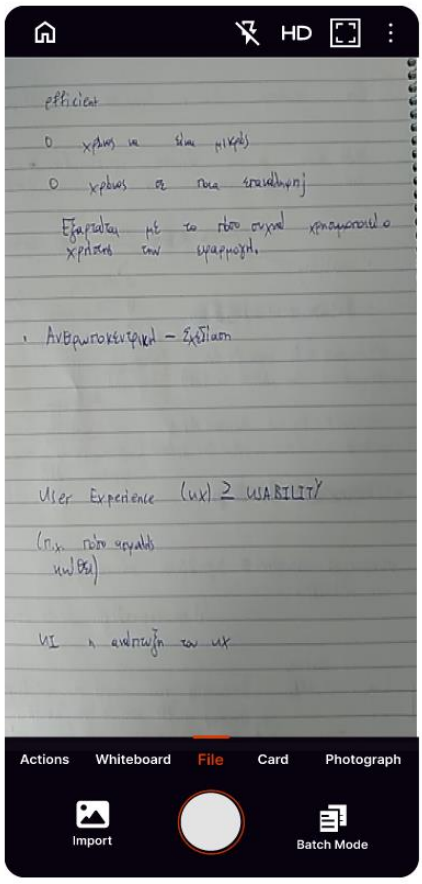


Figure 22: Initial Scanning Screen

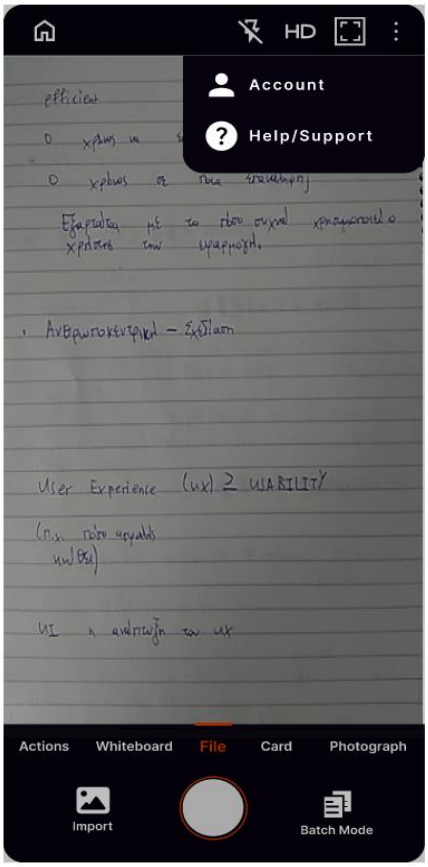


Figure 23 Settings

The user can select whether they want to scan a document, a card, a whiteboard, or take a photo. They can also choose another action via "Actions," such as QR scan or OCR. "Actions" is implemented in the prototype but is not included here to avoid filling the report with repetitive pages that are not crucial.

At the top of the screen, the user has options to go to "Home" (where all files are located), change the flash settings, image resolution, and choose whether to display the camera focus frame. Additionally, through the ellipsis (three dots), they can access help/support and their account. All of these are fully implemented in the prototype. Figure 23 shows the dropdown menu with settings.

The user selects "Batch Mode," as explained previously, to scan documents with more than one page. If "Batch Mode" is selected, then after taking a photo, the screen changes as shown in Figure 24.

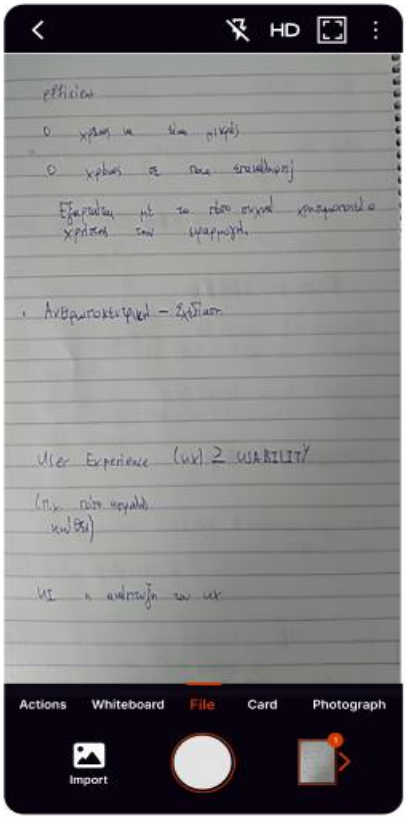


Figure 24: Scanning in Batch Mode

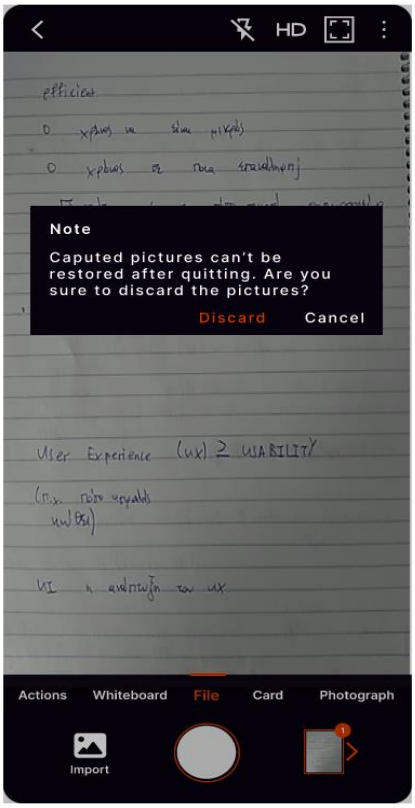


Figure 25 Discard scan

In the top left corner of the screen, there is an arrow that the user can tap to cancel scanning, as shown in Figure 25.

In fact, in Figure 23, which shows the settings, the account option is present because we are assuming we are creating an application for Microsoft. Therefore, the user will be able to log in to access Word, PowerPoint, PDF, etc. By tapping the account icon, the user is taken to the page in Figure 26.

Subsequently, by tapping "Sign up" or "Login," they are taken to the standard Microsoft login page, which is the same for all applications, and therefore we have not designed it. Similarly, the help/support feature in Figure 23 is the standard Microsoft help/support system..

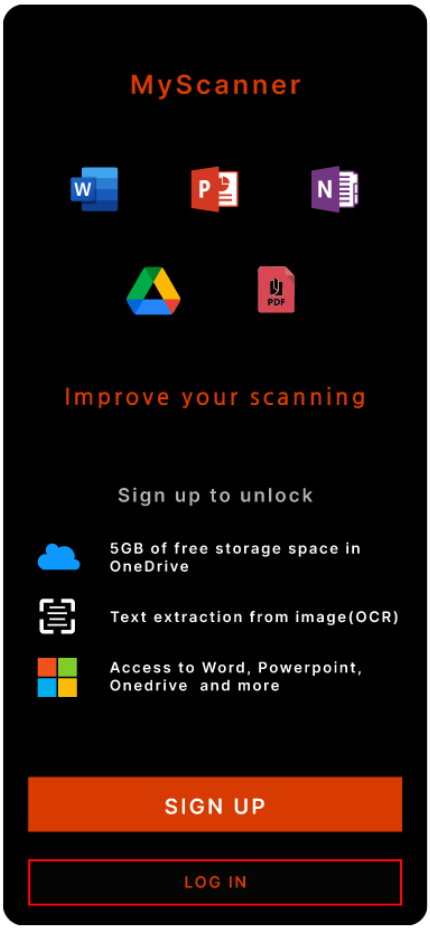


Figure 26: Login and Sign up

If the user has not selected "Batch Mode," when they scan the first photo, they are automatically taken to the preview screen in Figure 27, If they have selected "Batch Mode," then to access the preview, they tap the photo count indicator located at the bottom of Figure 24.

In the preview screen, the user can rename the file, view all photos, and modify them. In the prototype we have created, all transitions function exactly as they would on a real mobile phone. Furthermore, by tapping the "More" button, the user gains access to all the modification options shown in Figure 28.

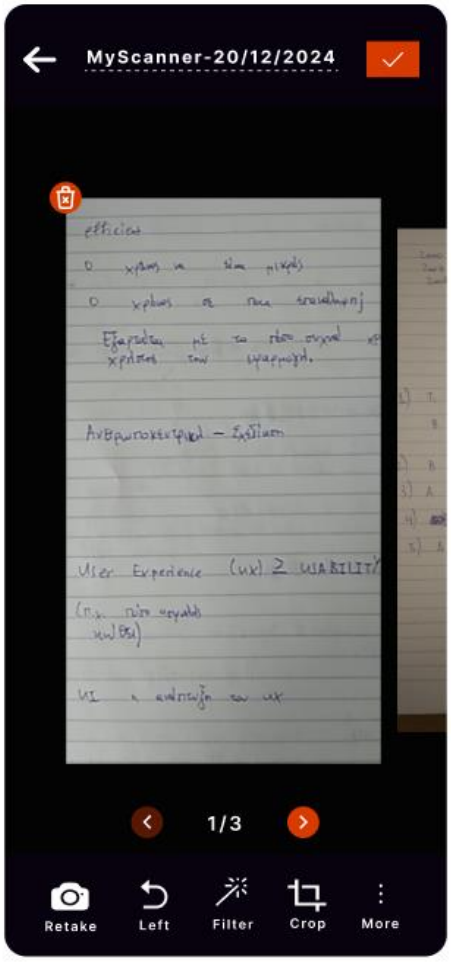


Figure 27: Preview Screen

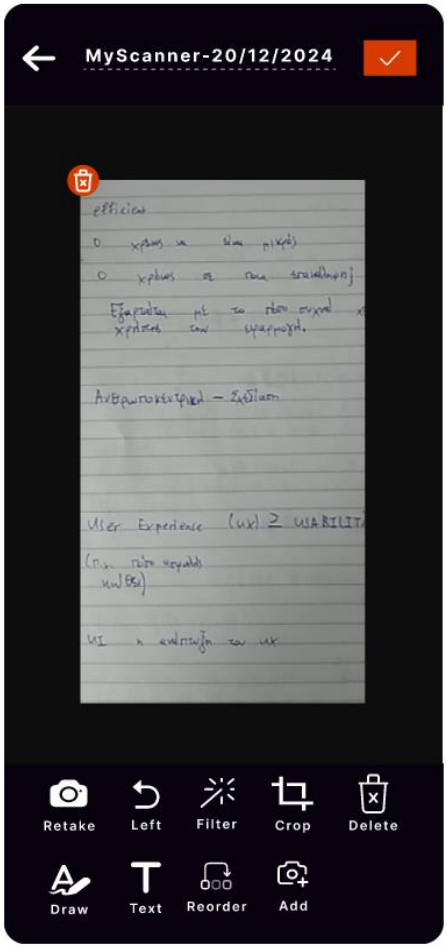


Figure 28: Modification Options

"Retake" addresses problem B) that we described during the heuristic evaluation, for cases where someone wants to quickly replace a photo, perhaps one that is blurry, with a new one. "Left" rotates the photo, "Filter" allows for filter selection, "Crop" adjusts the margins, "Delete" removes the photo, "Draw" is for drawing on the photo, "Text" is for inserting text, and "Reorder" is for rearranging. Finally, "Add" allows for adding another photo. These functionalities are shown in the following figures, Figures 29-35. Notably, when the user has selected a function to modify an image, they are taken to a slightly different page, where the document title is replaced with the function name, and the other functions are not visible. Also, in the top left corner, there is an "X" button for the user to go back without saving the change, and a red button with a checkmark to save the change and take the user back to the preview screen.

We note that the images in the report may have slightly different sizes due to pasting and varying dimensions, but in the prototype, they all have the same size.

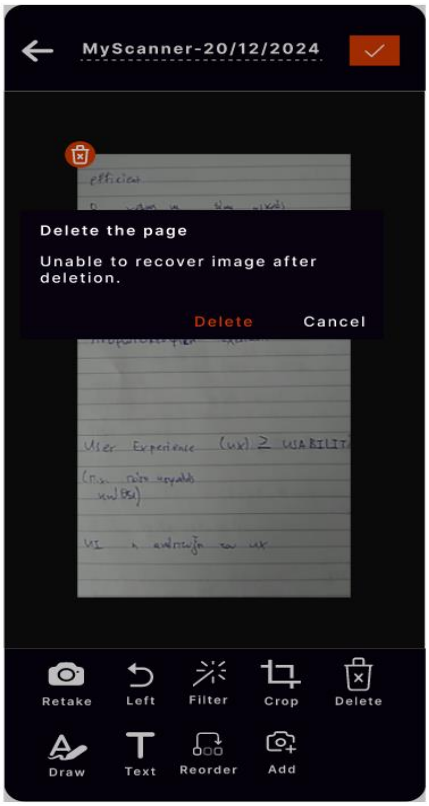


Figure 29 Delete

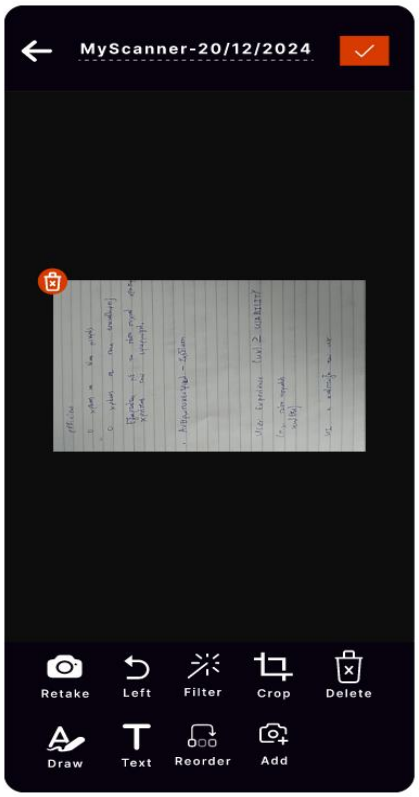


Figure 30 Left

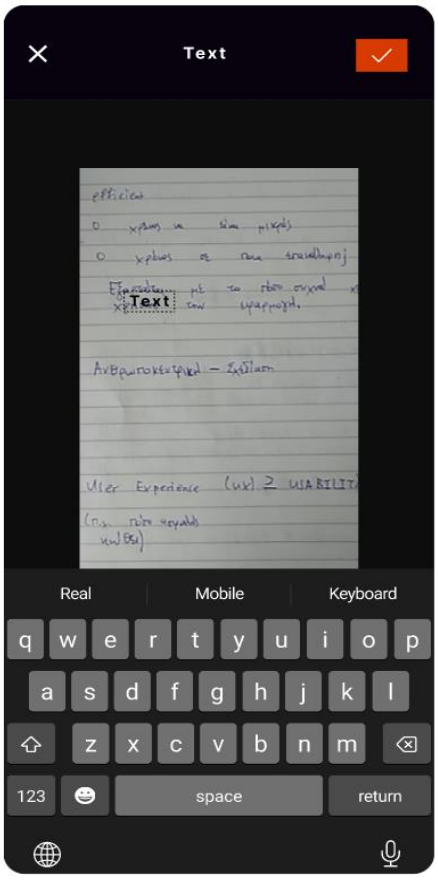


Figure 31 Text

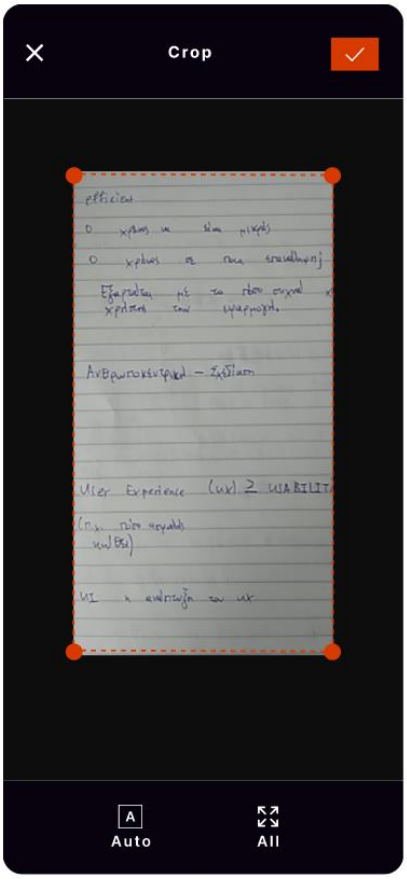


Figure 32 Crop

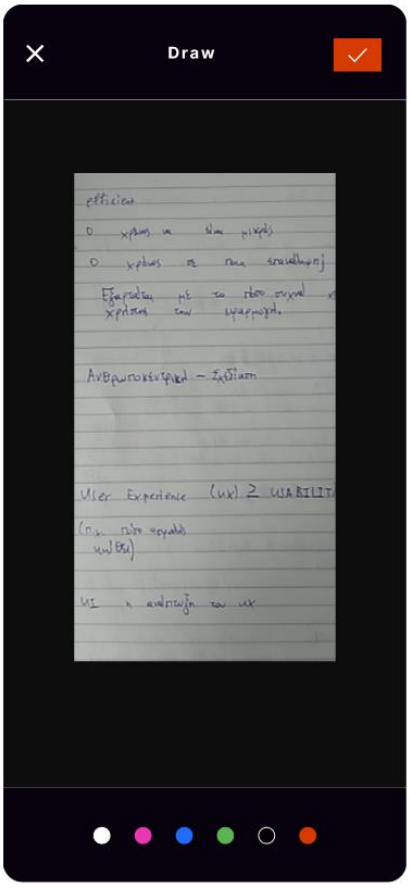


Figure 33 Draw

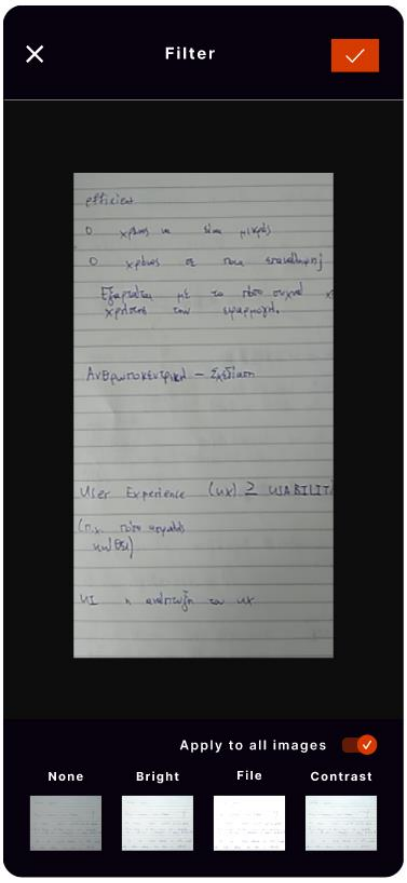


Figure 34 Filter

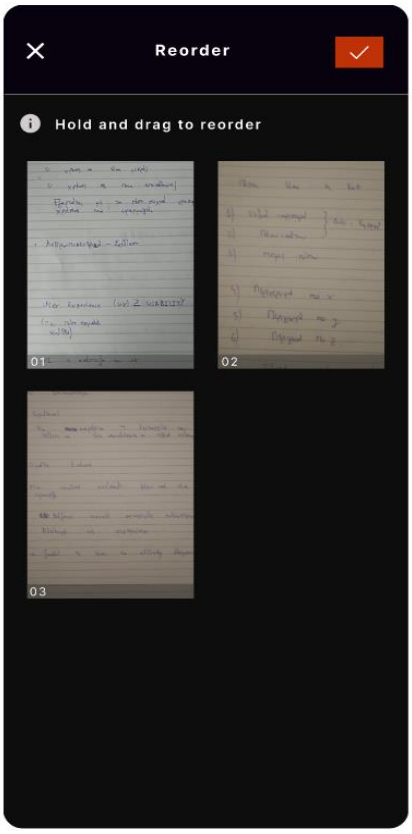


Figure 35 Reorder

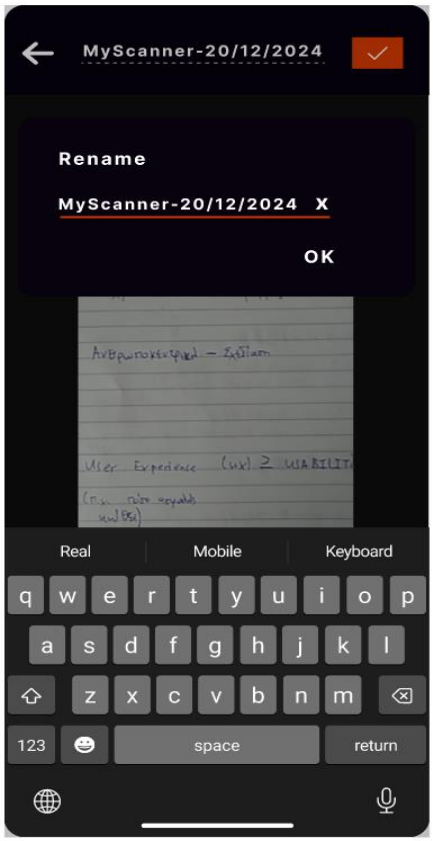


Figure 36 Rename

When the user has finished the preview and wishes to save the photos, they tap the red button next to the title. This displays the dropdown menu below, allowing them to choose the desired file format for saving their document. For example, they can select OneDrive, and their file will be saved directly to the cloud without being stored on their phone.

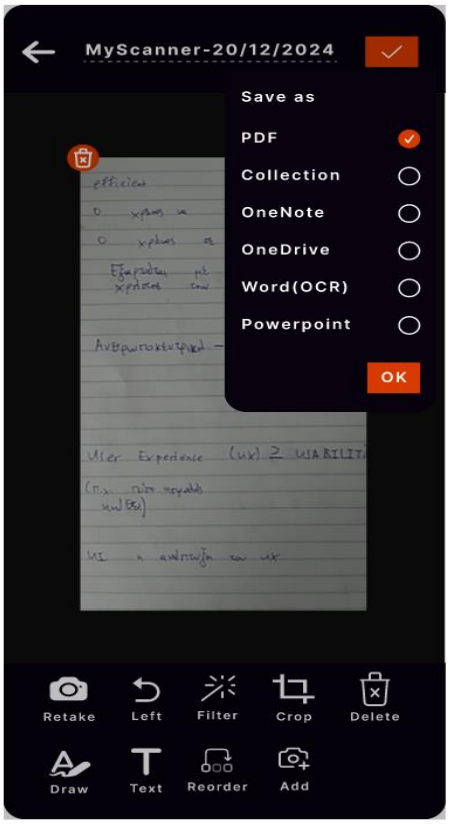


Figure 37: Save Format

Indeed, when the user taps "OK," they are taken to the screen shown in Figure 38.

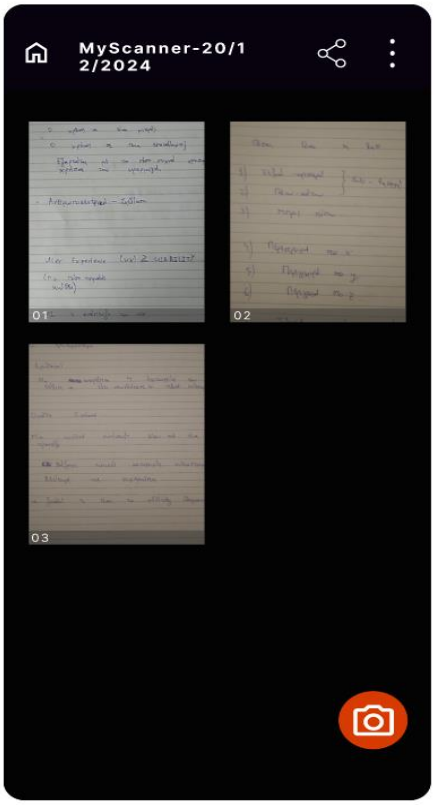


Figure 38: Preview of Saved File.

On this page, the user can share the file, rename it, edit it, and password-protect the file to secure it.

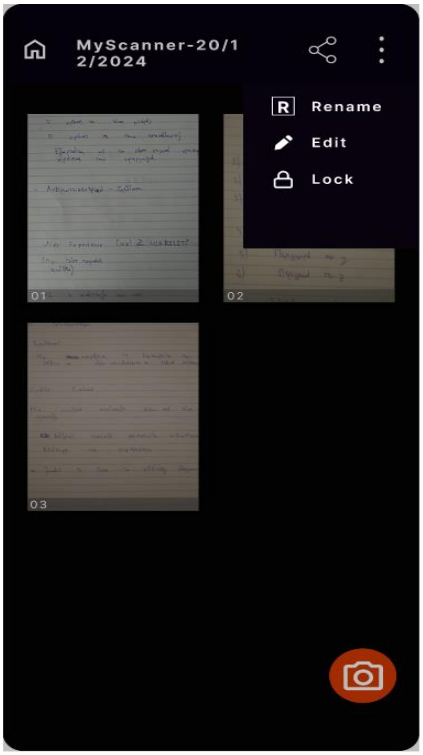


Figure 39: Settings

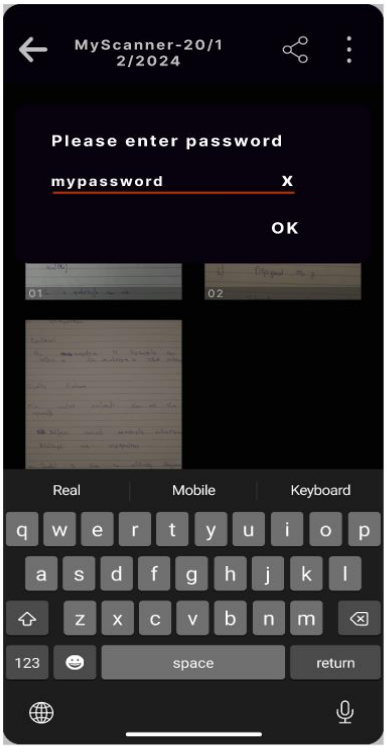


Figure 40: Lock

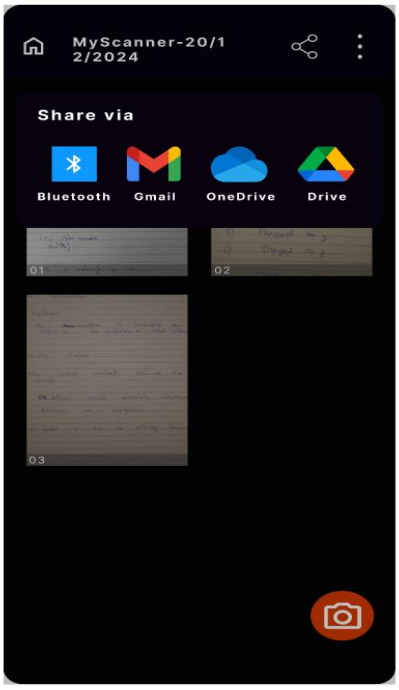


Figure 41: Sharing

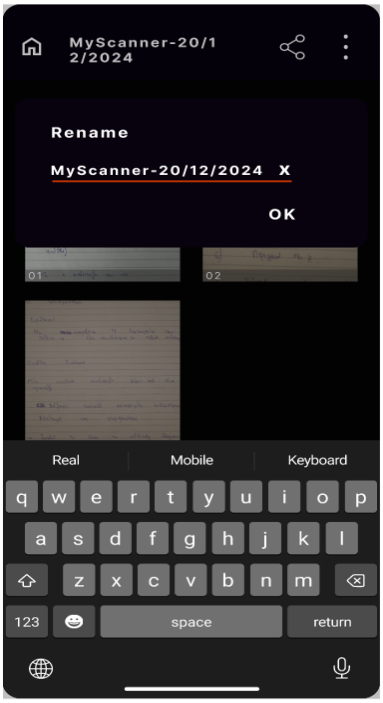


Figure 42 Rename

Furthermore, it is possible to select individual pages of the document. Each page can be shared, saved to the phone's gallery, or moved to a different document scanned by the application.

Figure 43: Selecting Individual Page.

Figure 44: Transferring selected pages to different document.

Finally, if the user taps the "Home" icon in Figure 38, they are taken to the main menu of the application, where they see all saved documents. For each document, they see basic information such as the title, the date it was saved, and the number of pages.

Figure 45:Main Menu

They can also search for a document or sort them according to various criteria

Figure 46: Document Sorting

Finally, the user can select a document and modify it.

Figure 47: Document Modification

7) Feedback and Modifications

The next step after creating the initial prototype of the application was to evaluate the prototype with users. This was carried out using cooperative evaluation, through which users work as co-evaluators with experts (Monk et al., 1993). This specific method was chosen because the prototype was not a complete, fully functional application, so the intervention of our team members was necessary to prevent and resolve any dead ends.

During this process, each user expressed positive and negative observations about the prototype. A team member asked questions such as, "How do you find this page/functionality?", "Does the system react as you would expect?", "Is any functionality or option missing that you would consider necessary?", "Is there anything you don't understand?". Finally, the team member requested a general opinion on the prototype. Specifically, this evaluation was conducted with 4 different individuals, and the results are shown in the following table.

User	Positive Observations	Negative Observations
1	The photo window and the menu with the files were very understandable and easy to use.	I would like icons on the file save page.
	Changing the page order seems very simple and understandable.	I would like to be able to read the file I created, as it would appear from a text reader application.
2	I like the file search option.	What does "collection" mean?
	I like the tools and the options within them.	I expected to see the final result in text format and not pages.
		I would like the icon to change color when the flash is activated.
3	I like that there are no icons on the image while I'm taking photos.	I would like a back button on the login/signup.
	Interesting page transfer to another file.	Why is edit hidden in the file preview?
	I like the simplicity of the windows.	I would like a preview of the file before I send it.
4	I like the options in the tools	I would like icons in the file type selection
	The windows are understandable.	They did not understand that they were on the file preview window.
	I like that I can choose the file type.	

Table 5: User Feedback on the Prototype from Cooperative Evaluation.

Most of the changes in the design and appearance of the application were positively received during the evaluation. However, it is evident that some minor user concerns still needed to be addressed. Specifically, all users had minor issues or desires related to the final file preview screen. Therefore, the following changes were implemented. First, all options that had already been implemented are now directly visible at the bottom of the screen. Simultaneously, switching between reading mode and thumbnail view was implemented by adding a corresponding button in the top right corner of the screen. The initial design is shown in Figure 38 and Figure 39, while the modified version is shown in Figure 48 and Figure 49.

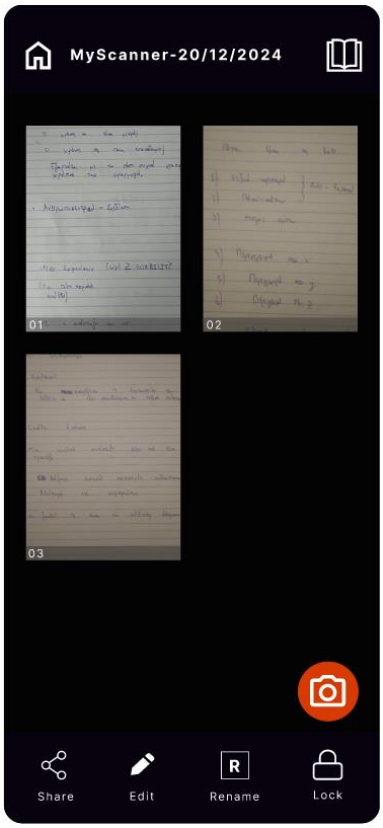


Figure 48: Final Preview Window

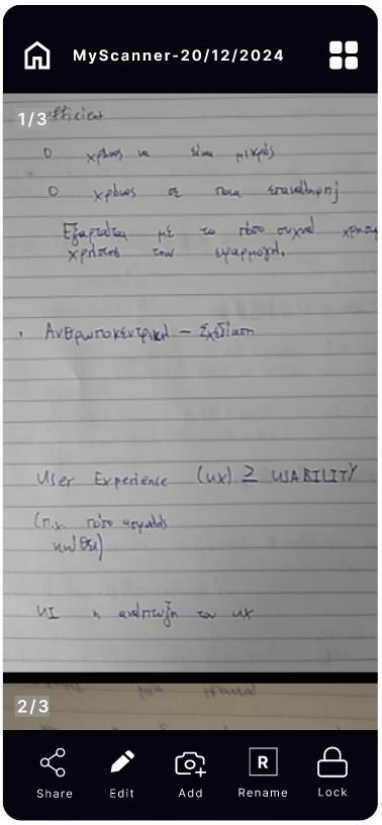


Figure 49: Read mode

The remaining functionalities of the preview page operate in the same way, with the only difference being the display location of the sharing window, which is shown in Figure 50.

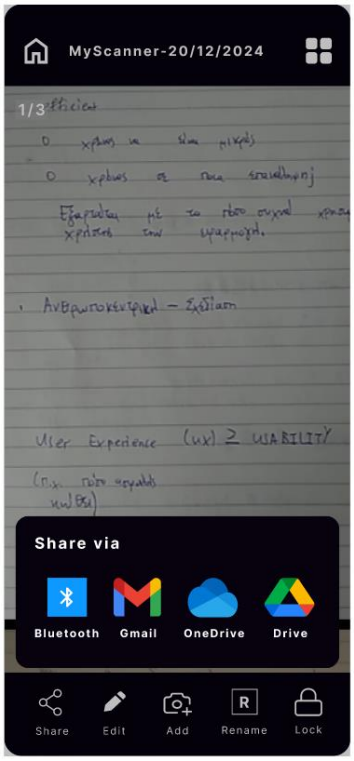


Figure 50: Final Sharing Popup

The next modification implemented was the addition of icons to the file format selection page and changing the name "Collection" to "Gallery." The final result is shown in Figure 51.

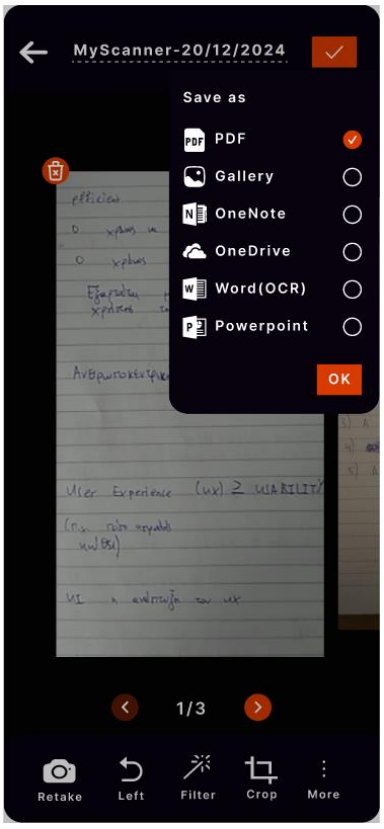


Figure 51: Final Save Format Selection

Finally, minor modifications were made to the account page and the photo-taking screen. On the account page, a back button to return to the previous page was added, as shown in Figure 52. On the photo-taking screen, the only change was the addition of an orange color to the flash icon when it is active, and this change is shown in Figure 53.

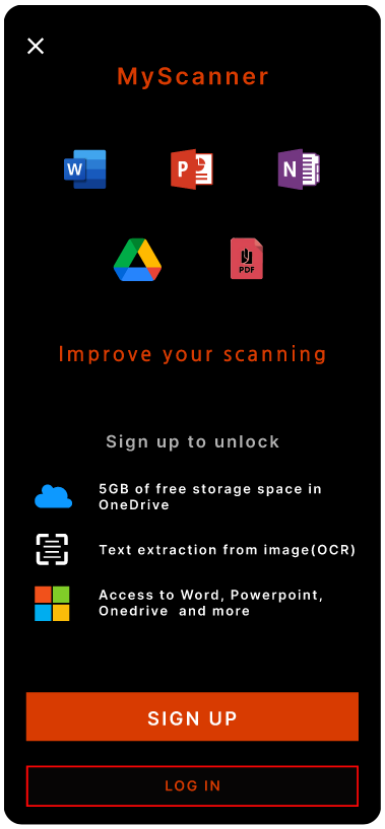


Figure 52: Final Account Window

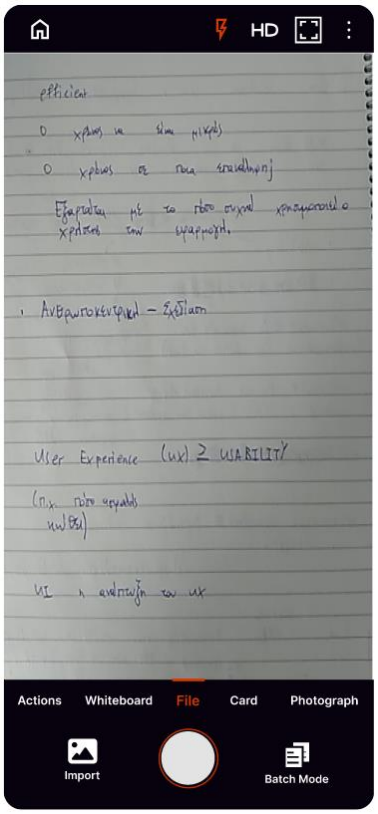


Figure 53: Final Photo-taking Screen (Flash Icon Highlighted)

8) Comparison

To compare our redesigned application with the original Microsoft Lens application, we will apply the GLM method to calculate the time it takes a user to scan a document. Subsequently, we will compare the times for the original application with those of our redesigned application.

As before, the process we will examine involves a user scanning 3 photos with default dimensions and options and finally saving the file as a PDF named "myfile."

Initially, the user opens the application and scans 3 pages (Figure 26). An experienced user knows to first tap "Batch Mode" and then take photos. Therefore, they do not need to engage in mental thought (M) for each transition between actions. These steps correspond to the following elementary actions in the GLM model:

M: Open initial scanning screen (Figure 26)

H: Focus on the "Batch Mode" button

P: Finger movement

T: Tap "Batch Mode" button

H: Focus on the photo button

P: Finger movement to the photo button

T: Tap photo button to scan the first page

H: Focus on the photo area to take the 2nd photo (looking at the scene to be photographed)

H: Focus again on the photo button

T: Tap photo button for the 2nd photo

H: Focus on the photo area to take the 3rd photo

H: Focus again on the photo button

T: Tap photo button for the 3rd photo

M: Think about how to access the preview

H: Focus on the photo thumbnail indicator (bottom right of Figure 28)

P: Finger movement

T: Tap the indicator

The user is then taken to the preview menu (Figure 27) and wishes to change the default name, renaming the file to "myfile," and save the document. Therefore, they perform the following actions:

M: Preview menu appears

H: Focus on the title

P: Finger movement

T: Tap the title

M: Menu of Figure 42 opens

H: Focus on the "x" (delete) button

P: Finger movement

T: Tap the "x" and the title is deleted

M: Think of a new name ("myfile")

6T: Type "myfile"

M: Think about how to proceed

H: Focus on the "OK" button

P: Finger movement

T: Tap "OK" for the name

Next, they tap the red checkmark button, which opens the menu of Figure 40, and select "OK." These actions correspond to:

M: Preview menu appears

H: Focus on the red checkmark button

P: Finger movement

T: Tap the red checkmark button

M: Dropdown menu for format selection opens; user verifies PDF is the default option

H: Focus on the "OK" button

P: Finger movement

T: Tap "OK"

The actions required for the user to scan a three-page document with our redesigned application are summarized in Table 6.

Τελεστής	**Φορές που εφαρμόζεται**	**Απαιτούμενος Χρόνος (sec)**
M	**8**	**8*1.35=10.8**
H	**12**	**12*0.4=4.8**
P	**8**	**8*0.34=2.72**
T	**16**	**16*0.73=11.68**

Table 6: Required Times for Scanning with the "myScanner" Application

Therefore, the user needs a total of 10.8 + 4.8 + 2.72 + 11.68 = 30 seconds to scan a three-page document using our redesigned application.

Assuming the user is focused during scanning and their attention is not easily distracted (6% distraction factor), the required time is:

$$T = 1.06 * 30 = 31.8\ sec$$

In conclusion, the calculated time for a user to scan three pages with our redesigned application is 31.8 seconds, compared to 35.17 seconds for the original Microsoft Lens application. Therefore, with each scan, each user saves approximately 35.17 - 31.8 ≈ 3.4 seconds. This is highly significant for an application like Microsoft Lens, which has been downloaded 50,000,000 times. If we assume each user uses the application once a month, this translates to a monthly time saving of 50,000,000 * 3 = 150,000,000 seconds.

Beyond time efficiency, our proposed design is effective because it addresses all the problems identified during the formative evaluation by both experts and users. Furthermore, we demonstrated our application to 3 typical users who had participated in the think-aloud evaluation, and they were impressed by the improvements and the overall design, stating that they would not change anything.

9) Conclusions

In this paper, we employed human-centered design methods and conducted user analysis to understand and define typical users. Subsequently, we performed task analysis to comprehend the specific tasks users perform with a scanning application. We also conducted requirements analysis and defined the functional and non-functional requirements of the application.

Following this, we evaluated the Microsoft Lens application using both expert-based and user-based evaluation methods. For formative user evaluation, we used the think-aloud protocol, and for expert formative evaluation, we used heuristic evaluation. For summative application evaluation, we employed questionnaires, specifically the User Experience Questionnaire (UEQ), for users, and the KLM method for expert-based summative evaluation. Both formative evaluations by users and experts highlighted significant usability issues. Furthermore, the questionnaire method indicated that typical users who participated in the application evaluation considered it to be neutrally usable, neither very well-designed nor very poorly designed.

Subsequently, we took into account all the feedback we had gathered and, combined with the requirements identified during the requirements analysis, redesigned the application in Figma to effectively address all the observed problems. Indeed, we created detailed prototypes showcasing all individual functionalities of the application and user interaction patterns. These prototypes were further refined through re-evaluation by users.

Finally, we compared the time required for a typical task – scanning 3 pages – using the original application and our redesigned application. We found that the time required with our redesigned application is approximately 3 seconds less. Additionally, we presented the redesigned prototypes to typical users to gather any further suggestions, but they indicated they would not change anything and considered the design highly user-friendly.

Bibliography

Free mobile typing speed test for your smartphone - Find out your wpm on iPhone and Android. (2024). Typing-Test-Mobile.com. https://www.typing-test-mobile.com

Hinderks, A., Schrepp, M., & Thomaschewski, J. (2018). *User experience questionnaire*. User Experience Questionnaire (UEQ). https://www.ueq-online.org/

Nielsen, J. (1993). *Usability engineering*. Boston : Academic Press.

Nielsen, J. (2012, January 15). *Thinking aloud: The #1 usability tool*. Nielsen Norman Group. https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

Nis (2023, Απρίλιος 24). *The PACT Analysis: A Human-Centered Design Framework*. Bornoe.org Blog. https://bornoe.org/blog/2023/04/the-pact-analysis-a-human-centered-design-framework/

Nyström, A. (2018). Gesture-level model : A modified Keystroke-level model for tasks on mobile touchscreen devices.

Avouris, Nikolaos, Katsanos, Christos, Tselios, Nikolaos, and Moustakas, Konstantinos. 2018. Introduction to Human-Computer Interaction. University of Patras Publications.