A Case Study on Accessibility of Document Scanning Applications

Kosmas Archontis
arxodiskosmas@gmail.com
University of Patras
Patras, Greece

Maria Aspioti
mariaaspioti64@gmail.com
University of Patras
Patras, Greece

Christos Sintoris
sintoris@upatras.gr
University of Patras
Patras, Greece

Nikolaos Avouris
avouris@upatras.gr
University of Patras
Patras, Greece

Abstract

In this paper, we report a case study to investigate usability and accessibility barriers in current mobile document scanning apps and propose design recommendations and prototypes. Guided by human-centered design principles, we included in the study blind and low-vision user interviews and expert evaluation based on heuristic evaluation. Our findings reveal significant challenges related to recognizing the system state, effectively using the camera, and accurately extracting text from scans. To address these issues, we designed a "Guided Mode" to assist with camera alignment and specified the key functionalities it should provide. Some important aspects of the app interface that were found challenging for users were redesigned. We used a modified Gesture Level Model for screen readers to evaluate the proposed design, which showcases improved task execution times for important functions. This work offers insights into the barriers blind and low-vision users face when interacting with a mobile document scanning app, identifies critical accessibility issues, and presents strategies to enhance accessibility.

CCS Concepts

• **Human-centered computing → Accessibility design and evaluation methods**; *User studies*; *Empirical studies in accessibility*; *Interaction design.*

Keywords

Case study, Accessibility, Blind and low-vision, User testing

ACM Reference Format:
Kosmas Archontis, Maria Aspioti, Christos Sintoris, and Nikolaos Avouris. 2026. A Case Study on Accessibility of Document Scanning Applications. In *Proceedings of 8th Summit on Gender Equality in Computing (GEC '26).* ACM, New York, NY, USA, 2 pages. https://doi.org/XXXXXXX.XXXXXXX

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
GEC '26, Patras, Greece
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/2026/05
https://doi.org/XXXXXXX.XXXXXXX

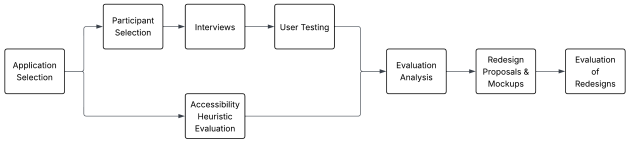

Figure 1: The flowchart showcasing the general design of the study.

1 Introduction

Mobile document scanning apps are increasingly used to convert physical documents into digital formats, enabling greater access to printed information. Blind and low-vision (BLV) people, however, who'd benefit the most from these apps, are often blocked from effectively using them due to accessibility issues, especially those related to screen reader compatibility and other assistive technologies, like screen magnification. Prior research on the matter focused on camera interaction primarily, rather than providing a comprehensive analysis of the overall BLV user experience [2].

As such, we conducted a broader usability and accessibility evaluation of a representative document scanning app, and proposed design recommendations with mock-ups illustrating their implementation, after identifying key problems.

2 Method

To investigate how BLV users interact with document scanning applications, we selected Microsoft Lens for its popularity, robust features, and Microsoft's accessibility focus. Two functionally blind male participants, ages 22 and 23, both familiar with assistive technologies, were recruited.

Each participant completed a remote user testing session via video call, verbalizing their thoughts throughout (Think-Aloud protocol) while performing five core tasks: **(1)** scan a three-page document, apply a filter, save it as PDF, and open it; **(2)** share the PDF; **(3)** edit and save it with a new filter; **(4)** delete saved PDFs; and **(5)** extract and copy text.

Audio and screen output were recorded for post-session analysis.

To complement user insights, we conducted an expert-based Accessibility Heuristic Evaluation by Deque Systems [3], which adapted Nielsen's 10 heuristics to align with Web Content Accessibility Guidelines (WCAG) standards.

A schematic overview of the study process is provided in Fig. 1.

Table 1: Comparison of estimated task execution times between *Microsoft Lens* and our redesign

	Task 1	Task 2	Task 3	Task 4	Task 5
Microsoft Lens	86.45	9.93	48.74	23.9	129.76
Proposed Redesign	83.93	10.35	39.81	24.32	97.2
Difference	-2.52	+0.42	-8.92	+0.42	-32.56

3 Evaluation Findings

Synthesizing the user testing and heuristic evaluation revealed four major accessibility barriers:

Lack of audiovisual alternatives. The app provides no contextual guidance (e.g., lighting, focus, framing), creating severe uncertainty. Both participants doubted their photo quality, either relying on prior sighted assistance or using physical tactile cues (longer side, texture). The heuristic evaluation confirmed a lack of non-visual feedback during auto-detection, mislabeled icons, and incorrect screen reader selection order.

Scanning session handling. Multi-page scanning navigation is unpredictable. After capturing the first page, the app fails to navigate to the editing screen automatically. Furthermore, previous session photos are retained without clear indication; One user unknowingly continued a session with incorrect images, while the other struggled when pages were silently appended to an imported PDF without audio feedback.

Convoluted text extraction. Contrary to expectations, extraction is limited to single pages and easily confused with the read-only "Immersive Reader." One user completely failed the text extraction task (Task 5) due to this UI ambiguity. Additionally, the system reads aloud nonsensical OCR results from poor-quality images without providing partial-failure error warnings.

Limited content sharing. The heuristic evaluation identified that users cannot share multiple files simultaneously or combine extracted text from multi-page documents, violating *Predictability and Consistency* principles and compounding the text extraction barriers.

These consolidated challenges directly informed our redesign proposals.

4 Redesign

Based on our evaluation phase, we propose the following key accessibility improvements: **(1)** A Guided Mode integrating screen readers to automate capture, notify users of lighting issues, and detect unsteady hands via concise prompts; **(2)** Streamlined Session Handling for linear, screen reader-friendly navigation; and **(3)** Restructured interface design to prioritize discoverability of multi-page scanning and text extraction for BLV users.

We created mock-ups for the redesigned application based on our suggestions, and applied Blind FLM [1] with substituted time values from GLM [4] to estimate task completion times, enabling comparison with the original app. Table 1 shows our design improved efficiency, especially for multi-page scanning and text extraction, with only minimal time increases in two tasks, in exchange for enhanced functionality (addition of buttons in the UI).

5 Conclusion

This study examined the usability and accessibility of a representative document scanning app for BLV users, revealing persistent barriers despite prior research on non-visual interaction. While accessibility issues like limited audiovisual feedback were anticipated, general usability problems were often amplified for BLV participants, leading to a compounded negative impact on their experience, and prompting them to seek sighted assistance. Through user testing and a heuristic evaluation, we identified challenges in session handling, text extraction, and lack of audiovisual alternatives. We proposed targeted redesigns, implemented as semi-interactive prototypes, and evaluated them using a hybrid Blind FLM–GLM analysis, which showed improved task efficiency.

References

[1] Shiroq Al-Megren, Wejdan Altamimi, and Hend S. Al-Khalifa. 2017. Blind FLM: An Enhanced Keystroke-Level Model for Visually Impaired Smartphone Interaction. In *Lecture Notes in Computer Science (Human-Computer Interaction - INTERACT 2017, Vol. LNCS-10513)*, Regina Bernhaupt, Girish Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer International Publishing, Bombay, India, 155–172. doi:10.1007/978-3-319-67744-6_10 Part 3: Assistive Technology for Blind Users.

[2] Michael Cutter and Roberto Manduchi. 2017. Improving the Accessibility of Mobile OCR Apps Via Interactive Modalities. *ACM Trans. Access. Comput.* 10, 4, Article 11 (Aug. 2017), 27 pages. doi:10.1145/3075300

[3] Deque Systems. 2019. *Supporting the Design Phase with Accessibility Heuristics Evaluations.* https://www.deque.com/blog/supporting-the-design-phase-with-accessibility-heuristics-evaluations/ Accessed: 2025-04-04.

[4] Anton Nyström. 2018. *Gesture-level model : A modified Keystroke-level model for tasks on mobile touchscreen devices.* Master's thesis. Uppsala University. https://api.semanticscholar.org/CorpusID:52036710