USER PERSONAS & POVS FOR ANONDOCS

PERSONA 1: ANA

1. Name: Ana

2. Age: 21

3. Occupation: Computer Science Student (English track)

4. Background: Ana is a junior in her Computer Science degree at university, specializing in the information engineering. She is passionate about both technology and communication, and often works on projects that combine coding with linguistic analysis. Ana is active in tech clubs at school and frequently volunteers to help organize student events. Outside of studies, she enjoys coding challenges, reading sci-fi novels, and learning new languages.

5. Goals and Needs:

- Quickly scan and redact sensitive details from documents related to her academic work (research papers, assignments, etc.).
- Ensure that the format of the documents remains intact when sharing with professors or classmates.
- Safeguard any personal or sensitive information when submitting documents online.

6. Pain Points:

- Redacting documents manually is a tedious task, especially when dealing with large amounts of text.
- Fears making mistakes when manually redacting sensitive information, leading to potential privacy issues.
- Needs to ensure documents are submitted in their original format for academic consistency.

7. Technology Proficiency: Very tech-savvy, comfortable with coding, using tech tools, and automating tasks.

8. Preferred Platforms: Primarily uses a laptop for schoolwork and a smartphone for quick tasks on the go.

9. Quote: "I need a tool that lets me secure my documents quickly and efficiently, without ruining the format or wasting too much time."

PERSONA 2: STEFAN

1. Name: Stefan

2. Age: 21

3. Occupation: Call Center Representative

4. Background: Stefan works at a call center where he handles customer information and data on a daily basis. His role involves verifying and correcting customer details, such as addresses, phone numbers, and other personal information. Outside of his job, he doesn't typically handle sensitive documents. Stefan is familiar with technology but prefers using trusted tools for work-related tasks. In his free time, Stefan enjoys gaming and keeping up with tech news.

5. Goals and Needs:

- Efficiently scan and redact sensitive customer data from documents when required.
- Use AI to help with data verification and ensure no mistakes are made during document handling.
- Simplify the document correction process by using a reliable and trusted tool.

6. Pain Points:

- Needs to verify and double-check information, such as addresses, which can be tedious.
- Concerned about using untrusted AI tools, like Deepseek, due to privacy and data security issues.
- Would appreciate more efficiency in the process of handling customer data and document redaction.

7. Technology Proficiency: Comfortable with using integrated applications at work but does not delve into complex tech outside of his job.

8. Preferred Platforms: Primarily uses his smartphone for scanning and handling documents.

9. Quote: "A trusted AI tool would make my work much easier, especially if it helps me quickly verify and redact customer data."

PERSONA 3: ABDELRAHMAN

1. Name: Abdelrahman

2. Age: 20

3. Occupation: Computer Science Student & Part-Time Business Analyst at a Retail Company

4. Background: Abdelrahman is a full-time computer science student who works part-time as a business analyst at a retail company. His role involves scanning and processing product information, including taking photos of products, filling in details like prices, descriptions, and stock quantities into Excel sheets. Outside of work, he enjoys coding, gaming, and participating in tech workshops. He is tech-savvy and always looking for ways to improve his workflow.

5. Goals and Needs:

- Scan product information quickly and efficiently, automatically populating fields in an Excel sheet.
- Reduce the time it takes to fill in product details by using AI-powered tools.
- Keep the process of data entry streamlined, accurate, and faster than using the Excel app on his phone.

6. Pain Points:

- The current Excel app on his phone is slow and tedious when filling in product details manually.
- Taking multiple photos of products and manually entering details for each can be very time-consuming.
- Needs a more efficient solution for handling product data and filling fields without errors or delays.

7. Technology Proficiency: Highly proficient in technology, particularly coding and working with software tools to automate tasks.

8. Preferred Platforms: Primarily uses his smartphone (Android) for scanning and taking photos of products and his laptop (Windows) for more complex data analysis and Excel work.

9. Quote: "I need a faster way to scan documents and fill out product details so I can focus on more important tasks instead of wasting time on manual data entry."

SHORT USER STORIES

User Story 1 (Ana):
"As Ana, I want to quickly scan and automatically redact sensitive details from my academic documents so that I can share them with professors and classmates without compromising personal or institutional data."

User Story 2 (Stefan):
"As Stefan, I want a reliable AI-based tool to efficiently verify and redact customer information so that I can handle sensitive data with confidence and speed at my call center job."

User Story 3 (Abdelrahman):
"As Abdelrahman, I want to scan product information directly into an Excel sheet so that I can minimize manual data entry and focus on higher-level analysis for my retail company."

POINTS OF VIEW (POVS)

POV #1 (Ana)

Ana is a **Computer Science student** who needs to **securely and accurately redact sensitive information** from academic documents while **preserving the original formatting**, so that she can **submit assignments and research safely** without risking privacy.

POV #2 (Stefan)

Stefan is a **Call Center Representative** who needs to **verify and redact sensitive customer information** quickly, so that he can **handle large volumes of customer data** without compromising **privacy or accuracy**.

POV #3 (Abdelrahman)

Abdelrahman is a **Computer Science student and part-time Business Analyst** who needs a **faster way to scan, capture, and populate product data** into Excel, so that he can **reduce manual entry** and focus on **higher-level tasks** in his retail job.

HOW MIGHT WE (HMW) STATEMENTS

From POV #1 (Ana)

1. **HMW (1A)**: How might we ensure no sensitive content is overlooked during the redaction process?

2. **HMW (1B)**: How might we preserve formatting and layout while automatically removing sensitive text?

From POV #2 (Stefan)

3. **HMW (2A)**: How might we give Stefan confidence that AI redaction tools are secure and trustworthy?

4. **HMW (2B)**: How might we automate verifying and redacting customer information to reduce manual effort and errors?

From POV #3 (Abdelrahman)

5. **HMW (3A)**: How might we speed up the process of capturing product info from images and translating it into structured Excel data?

6. **HMW (3B)**: How might we reduce the need for repetitive manual data entry, allowing Abdelrahman to focus on analysis and decision-making?

SELECTING THE BEST 3 HMW STATEMENTS

HMW (1B): How might we preserve formatting and layout while automatically removing sensitive text?

HMW (2A): How might we give Stefan confidence that AI redaction tools are secure and trustworthy?

HMW (3A): How might we speed up the process of capturing product info from images and translating it into structured Excel data?

POTENTIAL SOLUTIONS FOR EACH OF THE SELECTED HMWS

(1) HMW (1B)

How might we preserve formatting and layout while automatically removing sensitive text?

1. **Solution A**:

   o Develop an **AI-powered text-layer detection** system that separates text from underlying document structure (such as PDF layers).

   o The system **redacts only the text** while **retaining the original layer design**, so fonts, headers, footers, and styling remain unchanged.

2. **Solution B**:

   o Use **template-matching** or **document segmentation** that locks down each section's style/formatting.

   o Once the text is identified for redaction, replace it with an equivalent length of black bars or placeholders to **maintain spacing**, ensuring that the rest of the document's structure doesn't shift.

---

(2) HMW (2A)

How might we give Stefan confidence that AI redaction tools are secure and trustworthy?

1. **Solution A**:

   o Build an **auditable log** within the AI tool that tracks each step of the redaction process (timestamped records of what was scanned, how it was redacted, and by which algorithm).

   o Include **data encryption** and **compliance certifications** (e.g., GDPR, ISO) to convey transparency and reliability.

2. **Solution B**:

   o Create a **secure sandbox environment** or local/offline version of the redaction tool that **doesn't upload data** to external servers.

   o Provide a built-in **verification report** that highlights the redacted sections to reassure Stefan that no sensitive information remains.

---

(3) HMW (3A)

How might we speed up the process of capturing product info from images and translating it into structured Excel data?

1. **Solution A**:

   o Implement an **image-to-Excel AI pipeline**:

   ▪ Use **OCR (Optical Character Recognition)** to extract text from product images.

   ▪ Apply **natural language processing** to categorize product info (e.g., name, price, stock).

   ▪ **Auto-populate** an Excel spreadsheet in real-time, highlighting any anomalies.

2. **Solution B**:

   o Provide a **mobile companion app** where users snap product photos and quickly verify AI-suggested fields (price, SKU, etc.).

   o Once verified, the data is **synced** to a central Excel or database in the cloud, removing the need to do all data entry manually on a smartphone or laptop.

SELECTING THE TOP 3 SOLUTIONS

From the solutions proposed, here are three with the **biggest potential impact** and **ease of adoption**:

1. **Offline/Local Redaction Tool with Compliance Reporting** (Solution B from HMW 2A)

   o Ensures **data security** and addresses privacy concerns directly by avoiding external servers.

   o Builds **trust** through easy-to-read, automatic verification reports.

2. **Image-to-Excel AI Pipeline** (Solution A from HMW 3A)

   o Highly **scalable** for large product catalogs and significantly **reduces manual data entry**.

   o Improves **accuracy** via OCR and NLP, automating data structuring in Excel.

3. **AI-Powered Text-Layer Detection System** (Solution A from HMW 1B)

   o Preserves the **original design** and layout of academic documents while redacting.

   o Minimizes the risk of **formatting disruptions**, making it ideal for official or academic use.