# Scope

**Problem Statement** : To facilitate the transition from paper based operations to digital operations as seamlessly as possible for consumers and businesses lacking a dedicated infrastructure for digitization and to provide a document creation and management platform to users for a variety of purposes.

The objective is to offer a **comprehensive document management service** comprising -

1. Scan, edit, convert, sharing of documents
2. Ability to create, share, receive digital documents (eg. receipts, invoices, payslips, etc.)
3. Document storage services

- Current services offer **"slice" solutions** - pureplay scanning and editing solutions.
- We have to provide multiple pre and post scan services - **full stack solutions** to build a long term sustainable business (document creation, editing, storage, collaboration and editing, etc.)

**A standalone scanning solution will lose out to businesses that provide a scanning service as part of a larger ecosystem of services (eg. Adobe, Microsoft)**

With the **current push to a digital economy** by the central government, it is high time that all businesses and individuals follow suit and are able to **minimize the utilization of paper based documents, receipts, invoices** etc and move to more secure, permanent and convenient digital documents. With more than **442 million smartphone users** and data at **Rs 6.7 ($0.09) per gigabyte (GB)**, this is the most appropriate time to push for digitization. Also with the advent of COVID-19, **there will be a considerable change in how people look at contact with others and there might be a considerable decrease in exchange of physical items between people**. In such a situation a mobile based document management service can not only provide for the desired push towards a more digitized society, but will also reduce the exchange of paper based documents. Another major reason that this idea could succeed is the banning of CamScanner which created **a void for 100 million+ users** in the mobile scanner application segment, which is also our entry point into the market.

# Target Audience

| **Target Group** | **Use Case** |
|---|---|
| Students, Teachers | Notes, assignments |
| Businessmen | Tax invoices, delivery receipts |
| Professionals | Documents, expense receipts, visiting cards |
| Field Staff | Documents, expense receipts, visiting cards |
| Lawyers | Documents (eg. contracts) |
| Doctors | Prescriptions, patient records |

| | |
|---|---|
| Chartered Accountants | Client documents |

# Services

- Leverage user base by building a scan and edit functionality in app
- Work to build secure document storage and collaboration in app
- Develop vertical stack SaaS solutions for documentation management needs of priority users segments - **SMEs, lawyers, accountants, teachers**

## Sample Vertical Stacks

*To be initiated post reaching 50 mn users (target by Mar '21)*

### SME Pack

Allows business owners to:

- Tap into a library of templates: to generate quotations, delivery receipts, invoices
- Signature and delivery confirmation digitally for documents exchanged with customers
- Build folders to auto allocate documents to storage (eg. invoices, receipts, etc.)
- Share documents internally, with CAs, etc.
- Value Prop: Significant time and cost savings for small businesses. Reduction of physical contact with others.

### **Document Delivery Pack**:

- Partner with banks, financial services institutions to automatically deliver digital copies of bills, financial statements instead of physical copies direct to user folders
- More secure, more green, cheaper and faster delivery of documents; solves a consumer pain points of managing documents

### Other packs being considered

- Academia pack (for teachers and students),
- Expenses pack (to auto-read receipts, export to excel or expense software),
- Lawyer pack,
- CA pack, etc.

# Business Model

## Pricing

- Scanning as a service will be offered free of charge.
- Ads as a revenue model is limited given competitors offer ad-free services
- Monetisation models include:
    - charging subscription fees from businesses / professionals (for specific packs priced at **INR 25-100 per month**)
    - Charging users for AI-powered auto sorting of documents and storage (after having generated enough partners like banks, financial institutions, etc.)
    - Data harvesting and earning revenues from DMP (data management platform services)
- Path to Monetisation
    - Build user base with scanning
    - Expand engagement with cloud storage and collaboration
    - Set up vertical stacks facilitating document creation and sharing and sell as a subscription

## Timelines

| **Month** | **Milestone** |
|---|---|
| Sept ‘20 | Beta version app launched on Android and iOS |
| Oct ‘20 | Commencing cloud based storage services |
| Dec ‘20 | Beta testing of vertical stacks for SMEs |
| Mar ‘21 | Market launch of vertical stack #1 |

## User Growth

- The core challenge in this business will be the inability to spend on advertising, given limited revenue potential in scan.
- The business has to address this by ensuring word of mouth and building awareness.
- Awareness will drive downloads, but positive WOM must drive engagement.

Currently, we have identified **5 levers to drive awareness.**

| Levers | Description |
|---|---|
| Personal Networks | • Linkedin, Facebook, Whatsapp |
| Media | • Publishers focussed on startup news (Inc42, YourStory)<br>• Publishers / authors writing articles on the scan market<br>• Publishers writing about tech alternatives in the scan market |
| Influencers | • Citizens vocal about India's software ecosystem potential |
| User Communities | • Groups / netizens actively talking about the scan market<br>• Teacher, CA, doctor communities on FB, WA |
| Organic Growth | • Easy discoverability on Play / App Store<br>• WOM<br>• In-app user referrals |

# Features

## Product Roadmap

Transition from a beta product that has robust scanning to a full stack document creation and collaboration tool by Dec '20

| Features | Beta | V1.0 | V2.0 | V3.0 |
|---|---|---|---|---|
| By | Aug '20 | Sep '20 | Nov '20 | Dec '20 |
| **Category: Scanning** | | | | |
| Recognise borders for cropping | Yes | Yes | Yes | Yes |
| Autocapture | Yes | Yes | Yes | Yes |
| Angle Indicator for flat scans | No | No | Yes | Yes |
| **Category: Editing** | | | | |
| Crop | Yes | Yes | Yes | Yes |
| Annotate | No | Yes | Yes | Yes |
| Highlight | No | Yes | Yes | Yes |
| Add signature | No | No | Yes | Yes |
| **Category: Convert, Sharing** | | | | |

| | | | | |
|---|---|---|---|---|
| Export to formats | PDF, PNG, JPG, JPEG | | | |
| Sharing options | Yes, email (eg. Gmail), chat (eg. Whatsapp), file storage (eg. Drive) | | | + Own collaboration folders |
| **Category: OCR** | | | | |
| Convert text to OCR | No | No | Yes | Yes |
| Edit OCR and export | No | No | Yes | Yes |
| Search words from text | No | No | No | Yes |
| **Category: Storage** | | | | |
| Cloud Storage | No | No | Yes | Yes |
| Group Collaboration / Editing | No | No | Yes | Yes |
| Wireless Printing | No | No | No | Yes |

Green color - already completed in source code

# Feature Benchmarking

**v.s present market leaders**

Current Market Leaders of the entry segment (ie scanning) provide only “slice solutions”. Although solutions for a document management service are provided by these market leaders, the solutions are different products, designed separately not keeping in mind the use case of a document management service.
By Dec ‘20, our features will far exceed Adobe and Microsoft current features

| Features | Our Solution (V3.0, Dec ‘20) | Adobe Scan | Microsoft Lens |
|---|---|---|---|
| **Category: Scanning** | | | |
| Recognise borders for cropping | Yes | Yes | Yes |
| Autocapture | Yes | Yes | No |
| Angle Indicator for flat scans | No | No | No |
| **Category: Editing** | | | |

| | | | |
|---|---|---|---|
| Crop | Yes | Yes | Yes |
| Annotate | No | Yes | Yes |
| Highlight | No | Yes | Yes |
| Add signature | Yes | For premium users | No |
| **Category: Convert, Sharing** | | | |
| Export to formats | PDF, PNG, JPG, JPEG | PDF | PDF, PPTX, DOCX |
| Sharing options | Email, social media, cloud | Email, social media, cloud | One Drive |
| **Category: OCR** | | | |
| Convert text to OCR | Yes | Yes | No |
| Edit OCR and export | Yes | For premium users | No |
| Search words from text | Yes | For premium users | No |
| **Category: Storage** | | | |
| Local only storage (sensitive docs) | Yes | No | No |
| Cloud Storage | Yes | Yes | Yes |
| Group Collaboration / Editing | Yes | For premium users | No |
| Wireless Printing | Yes | No | No |

# Development Principles

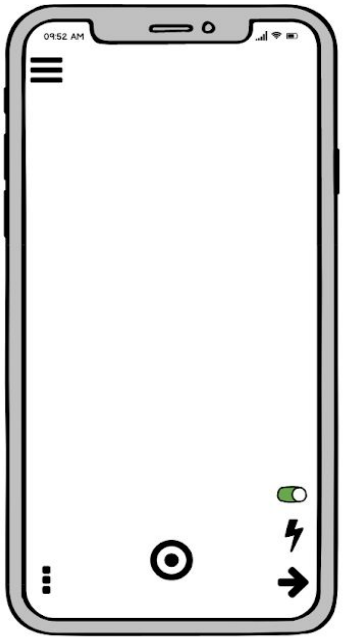
- We will offer a iOS and Android application for download from respective application stores
- A web-based service will go live in line with our collaboration stack rollout (in Nov)
- Some key guardrails:
    - A plain / clean UI to penetrate digital hinterlands and draw a large base of users
    - Light app (under 30MB) to ensure that max users can adopt this application

# Technology Stack

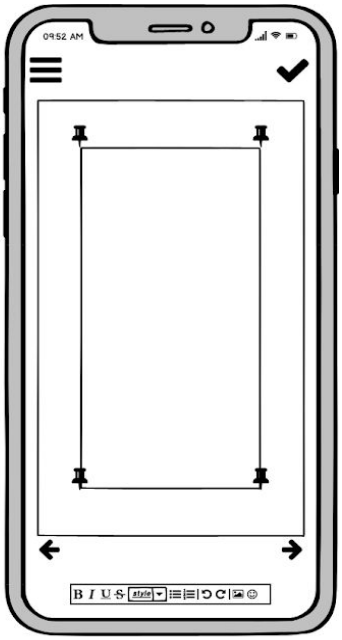
The services offered will be divided into two platforms-

- **A Mobile based client application**
    - The Mobile client, being offered on both iOS and Android will leverage the **JavaScript** framework called **React-Native** to present the same application to both the platforms, minimizing the efforts to ensure cross-platform compatibility of the code being written. React-Native allows for quick prototyping of cross-platform applications. Although not written in the native languages for each platform, React-Native apps perform as well as

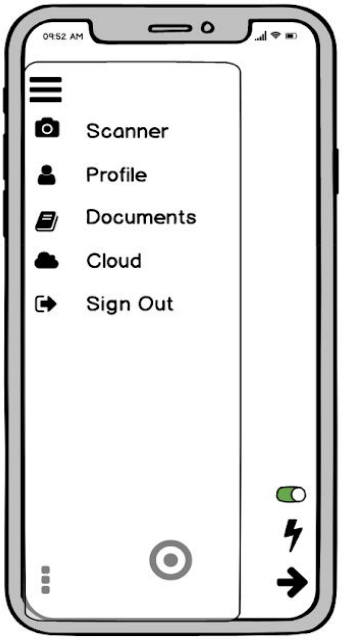
native applications owing to the javascript bridge that connects the javascript modules with their corresponding native modules.
    - Javascript, being the versatile language as it is, can also be used to deploy the Artificial Intelligence algorithms being developed as a part of the Image Understanding and OCR capabilities of the solution. **Tensorflow lite** and **Tensorflow.js** will be preferred technologies for deploying the developed models
    - **OpenCV** will be leveraged to develop the image editing capabilities for the application. To Extend the image Editing abilities to both Android and iOS, **Java** and **Objective C** modules using OpenCV will be developed. These modules will mirror each other in abilities and will be connected to the same JavaScript application.
- **A Web based server application**
    - The Web based server application will be used to provide any and all cloud based features as mentioned in the feature benchmarking system. The web application will also be used to deploy the large sized and memory intensive Artificial Intelligence algorithms to maintain a small application size and quick access. The **Flask microframework** will be used to develop a **RESTful API** that will serve the aforementioned Artificial Intelligence Models. The RESTful API will be served using a **Gunicorn** application server with an **Nginx** reverse proxy on a cloud service provider like **Amazon Web Services,** till the time the venture can afford to create and set-up our own cloud architecture. All user services, like login and personalized services will also be provided through this web application
    - The Web Application mentioned in the previous application is developed in **Python** and as such development and deployment of the Artificial Intelligence algorithms will also be done using python, for the models which can not be deployed directly onto the mobile frontend client. To develop and deploy these models, the technologies of **TensorFlow** and **TensorFlow Extended** will be leveraged
    - Database services will be implemented using a **Relational Database** to efficiently store and retrieve user information. The use and design of relational databases will allow for storage and retrieval of information according to predefined relationships. This will allow the databases to serve as the data input pipeline for the Artificial Intelligence Algorithms being deployed as a part of the solution for both Image Understanding and User Experience enhancement. **Postgresql** along with **SQLAlchemy** will be used to develop and deploy the database models.
    - Artificial Intelligence algorithms to be used in the initial solutions will be developed on top of open source and current state of the art OCR algorithms using Deep Learning. The following github repositories will serve as a base for development of the OCR capabilities-
        - https://github.com/githubharald/SimpleHTR
        - https://github.com/Breta01/handwriting-ocr
    - Both the aforementioned Deep Learning models make use of recurrent and convolutional neural networks to identify and learn words to convert an image with text into a computer editable document. Word Search models will also be implemented using the concepts used in these models.



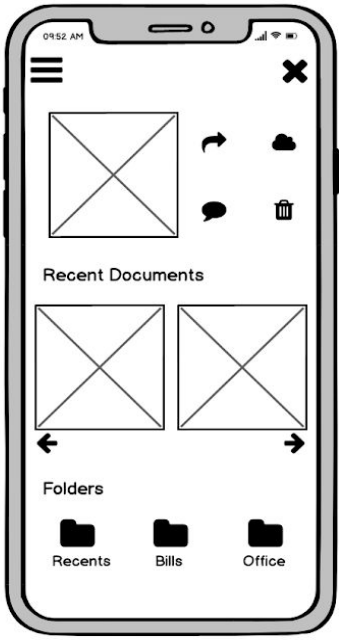

Capture Screen




Edit Screen




Drawer Screen




Save Screen

Wire Frame for Beta Application

# Key Evaluation Parameters

Data Storage

- For the first 90 days (till Nov ‘20), we will only offer local storage on user phone
- We plan to build secure internal cloud storage within India
    - Analysis on in house versus outsourced storage to be examined
    - Given steady requirement over time, internal storage likely to be more effective
- This capability will drive us to provide cloud storage as well as partner with businesses for the **digital document delivery pack** mentioned above

Scalability

- The stack is being build to **manage 100mn+ users** by next year
- Unlike a transactional platform (e-commerce, ride-hailing, payments, etc.) this platform has very steady server requirements, meaning scalability in architecture is not a major concern
- Being a mobile based platform the only major expenditure will be that of money for hosting services and time for the development of the perfect application.