**** 

**Lebanese Canadian University – LCU**

Faculty of Arts and Science – FSA

Bachelor in Computer Science

Final project

**PAPPER-HEX**

Presented by

**Aghnatios Steven**

**&**

**Khachkhachian Andrew**

Supervised by

**Dr Khalil Georgio**

June 20, 2021

Table of Contents

[1 Introduction 4](#_Toc16586292)

[1.1 Overview 4](#_Toc16586293)

[1.2 Scope 4](#_Toc16586294)

[1.3 Goals and Objectives 5](#_Toc16586295)

[1.4 Business Statements 5](#_Toc16586296)

[1.5 Acronyms and Abbreviations 5](#_Toc16586297)

[1.6 Points of Contact 6](#_Toc16586298)

[2 General Design Constraints 7](#_Toc16586299)

[2.1 Mandated Constraints 7](#_Toc16586300)

[2.2 Potential System Evolution 7](#_Toc16586301)

[3 Nonfunctional Requirements 8](#_Toc16586302)

[3.1 Performance Requirements 8](#_Toc16586303)

[3.2 Security Requirements 8](#_Toc16586304)

[3.3 Safety Requirements 8](#_Toc16586305)

[3.4 Legal Requirements 8](#_Toc16586306)

[3.5 Documentation and Training 8](#_Toc16586307)

[4 System Architecture 9](#_Toc16586308)

[4.1 Software Interface 9](#_Toc16586309)

[4.2 Process Diagram 9](#_Toc16586310)

[5 Technical Specifications 10](#_Toc16586311)

[5.1 Admin Module 10](#_Toc16586312)

[5.1.1 Flow Chart 10](#_Toc16586313)

[5.1.2 Features 10](#_Toc16586314)

[5.1.3 ERD Diagram 10](#_Toc16586315)

[5.2 User Module 10](#_Toc16586316)

[5.2.1 Flow Chart 10](#_Toc16586317)

[5.2.2 Features 10](#_Toc16586318)

[5.2.3 ERD Diagram 10](#_Toc16586319)

[5.3 Client Module 10](#_Toc16586320)

[5.3.1 Flow Chart 11](#_Toc16586321)

[5.3.2 Features 11](#_Toc16586322)

[5.3.3 ERD Diagram 11](#_Toc16586323)

[6 Contract 12](#_Toc16586324)

[6.1 Introduction 12](#_Toc16586325)

[6.2 The authors of the software license 12](#_Toc16586326)

[6.3 Preamble 12](#_Toc16586327)

[6.4 Objects delivered to the client 12](#_Toc16586328)

[6.5 Objects not delivered to the client 12](#_Toc16586329)

[6.6 Contract effective date 12](#_Toc16586330)

[6.7 Associated services 12](#_Toc16586331)

[6.8 Additional services 12](#_Toc16586332)

[6.9 Warranty and maintenance 12](#_Toc16586333)

[6.10 Languages 12](#_Toc16586334)

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Name** | **Description** |
|  |  |  |  |
|  |  |  |  |

**PAPER-HEX**



# Introduction

## Overview

People have nothing against paper. Paper is a brilliant invention of humanity. But in the 21st century people find it more appropriate for paper-based documents to be digitized (scanned). Once scanned, appropriate software can be used to find any document in a fraction of a second, just by typing a few keywords.

Paper-Hex is a document management system designed to work with scanned documents. As well as OCR with full text search, it provides the look and feel of major modern file browsers, **with tags** and (hierarchical structure) for files and folders, so that you can organize your documents in a similar way to Drop-box (via web) or Google Drive (but with tags and other cool features).

## Scope

Database Management System or DMS is type of system which stores a data and administrator can retrieve it any time with use of specific queries. Database Management System is used widely in almost all organization to keep record of various types of data. There is no better system present currently to manage and organize large number of data. It has features like tags, hierarchical folders and automations so that you can efficiently organize your documents.

What it does:

1. It [OCRs](https://papermerge.readthedocs.io/en/latest/glossary.html#ocrs) your documents.
2. It extracts text from your scans (which might be PDF, TIFF, JPEG or PNG) and indexes it.
3. Helps you fix scanned documents issues.
4. Reduced storage space.
5. Provides you nice user interface to easily browse your documents.
6. It allows you to save files online and access them anywhere from any smart-phone, tablet, or computer.
7. Helps you instantly find your documents:

🡪Based on Extracted-Text

🡪Based on Tags and Folders

🡪Based on [Metadata](https://papermerge.readthedocs.io/en/latest/metadata.html)

## Goals and Objectives

Database management systems have made a revolution in modern society. Business houses are making huge profits just because of database management system. DBMS provides a lot benefits in all spheres of human activity. The objectives are manifold. The main objectives of database management system are data availability, data integrity and data independence.

The three main goals of Paper-Hex are:

Data availability is a term used by computer storage manufacturers and storage service providers (SSPs) to describe products and services that ensure that data continues to be available at a required level of performance in situations ranging from normal through disastrous.

Data availability refers to the fact that the data are made available:

* To a wide variety of users
* In a meaningful format
* At reasonable cost
* With ease of access
* When and where required

Integrity is a critical aspect to the design, implementation and usage of any system which stores, processes, or retrieves data. Data integrity refers to the correctness of the data in the database. In other words, how reliable is the data available in the database. Integrity also means your data is authentic, accurate and consistent.

Database security professionals employ a number of practices to assure data integrity, including:

* Data encryption, which locks data by encryption
* Input validation, to prevent incorrect data entry
* Data validation, to certify uncorrupted transmission
* Data backup, which stores a copy of data in an alternate location
* Access controls, including assignment of read/write privileges

One of the main objectives of DBMS is to facilitate sharing of a database by current and future applications. The DBMS should not be tailored to a specific platform. One should be able to run DMS on any platform. DMS must ensure data independence for application programs.

Data independence allows:

* Change of database without affecting application programs
* Change of hardware or system software without affecting application programs
* Sharing of data by different applications by providing views appropriate for the application
* Control of Redundancy avoids duplication.

## Business Statements

## Acronyms and Abbreviations

**LCU** – an abbreviation for Lebanese Canadian University.

**HTML** – an abbreviation for Hyper Text Markup Language.

**SQL** – an abbreviation for Structured Query Language.

**Software –** is a set of instructions used to operate computers and execute specific tasks. Software is a generic term used to refer to applications, scripts and programs that run on a device.

**PDF** – Portable Document Format

**TIFF** – A convenient file container. Tagged Image File Format.

**JPEG** – Joint Photographic Experts Group

**PNG** – Portable Graphics Format

**VS** – An abbreviation for Visual Studio code.

**SSL**–An abbreviation for secure sockets layer.

**DB** – An abbreviation for Database.

**OCR**– An abbreviation for Optical Character Recognition.

**API**–API is the acronym for Application Programming Interface.

**JS**– an abbreviation for JavaScript.

**CRUD**–CRUD is an acronym that comes from the world of computer programming and refers to the four functions that are considered necessary to implement a persistent storage application: create, read, update and delete.

**XAMPP**–XAMPP is an abbreviation for cross-platform, Apache, MySQL, PHP and Perl, and it allows you to build Word-Press site offline, on a local web server on your computer.

**DBMS**– an abbreviation for Database Management Systems.

**Use case** – describes a goal-oriented interaction between the system and an actor. A use case may define several variants called scenarios that result in different paths through the use case and usually different outcomes.

**Scenario** – one path through a use case

**Actor** – user or other software system that receives value from a use case.

**Role** – category of users that share similar characteristics.

**Product** – what is being described here; the software system specified in this document.

**Project** – activities that will lead to the production of the product described here. Project issues are described in a separate project plan.

**Shall** – adverb used to indicate importance; indicates the requirement is mandatory. “Must” and “will” are synonyms for “shall”.

**Should** – adverb used to indicate importance; indicates the requirement is desired but not mandatory.

**May** – adverb used to indicate an option. For example, “The system may be taken offline for up to one hour every evening for maintenance.” Not used to express a requirement, but rather to specifically allow an option.

**Controls** – the individual elements of a user interface such as buttons and check boxes.

## Points of Contact

Below is a list of Point of Contacts relevant to this project:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Contact Name | Contact Type | Department | Telephone Number | Email | Oversight Function |
| Aghnatios  Steven | LCU  Lead Agency | Operation department | 76318864 | Steven\_aghnatios  @hotmail.com | Implementation  Problem Resolution  Decisions  Main developer  Financial |
| Khachkhachian  Andrew | LCU  Lead Agency | Operation Department | 71696621 | Andrew.k11198  @gmail.com | Implementation  Main developer  Database management |

# General Design Constraints

## Mandated Constraints

Paper-Hex is an open source document management system (DMS) primarily designed for archiving and retrieving your digital documents. Instead of having piles of paper documents all over your desk, office or drawers - you can quickly scan them and configure your scanner to directly upload to Paper-Hex DMS. Paper-Hex DMS on its turn will OCR the document and index it. You will be able to quickly find any scanned document using full text search capabilities. Paper-Hex is actively developed and we release quite frequently. This is web-based software. This means there is no executable file (aka no .exe files), and it must be run on a web server and accessed through a web browser. This software will use a local database software for the information storage. This database is php-my-admin SQL Server and the languages to be used in the creation of the DMS (Paper-hex) software is Visual Studio, JavaScript along with SQL.

## Potential System Evolution

This section suggests ways the system is likely to be extended or modified in the future such as customization, supports multiple languages, OCR in 27 languages, supports SSL, inbuilt calendar, email, sms, create documents using templates.

# Nonfunctional Requirements

Checkups will be performed every 1-3 months on all workstations for software or hardware failures in order to ensure that everything runs smoothly.

## Performance Requirements

System startup time should be less than 4 seconds. With 20 concurrent users no operation should take more than 3-6 seconds and 90% of the operations should take less than 2 seconds.

## Security Requirements

🡪Access to data and features may be limited to specific administrator(s)/user(s).

🡪Strong passwords should be enforced.

🡪Accounts will be locked after three or four login attempts.

🡪Password hashes will be stored encrypted and salted.

🡪A procedure should be put in place to ensure that accounts are deactivated when staff leave or move to different roles.

🡪The software must be accessed through an authentication screen before allowing any user to have access to the program.

🡪Access to data on the database is limited to search queries predefined in the DB and users must not have access to change it.

🡪 The input modules must be shielded from SQL injection attack.

## Safety Requirements

🡪 Working at a computer can cause back, neck and shoulder pains, headache, and eyestrain and overuse injuries of the arms and hands.

🡪Users must take a 10 minute break every 3 hours to reduce eye strain.

🡪Users must position the monitor at or slightly below eye level.

🡪Users must use their entire arm, not just their wrist, to move the mouse.

🡪Users must adjust the brightness on their screen to a comfortable level.

## Legal Requirements

You are entitled, upon request, to disclosure regarding your personal data that we are storing or are otherwise processing. You are also entitled to have any incorrect personal data corrected and rights to blocking or deletion of your personal data. We do not give, sell or transmit, neither your documents nor derivative information from your documents to 3rd parties. Your documents stay on our servers, and it is the core of our business to keep your documents safe and secure. Your data is yours only and our business is to guard your data securely. If you have any complaints regarding our processing of your personal data, we encourage you to contact us. Please address any requests in such matters to contact us @ Gmail: papper-hex@gmail.com.

## Documentation and Training

Paper-Hex is designed to be very intuitive. Because its user interface resembles a typical desktop file browser, you will be able to use Paper-Hex for basic document management tasks without any training at all. However, don’t be misled by familiar user interface, underneath it, there is a powerful tool which will bring your paperwork to the new level. Paper-Hex is a full-fledged document management system, not just a file browser, and to use it efficiently you need to understand few high-level concepts.

# System Architecture

## Software Interface

Paper-Hex is a web based application, addressable API’s using technologies such as web services. As with any web based application it can be accessed and used from any modern web browser, like Mozilla Firefox, Chrome, Edge, or Safari. Like a typical web application it runs on server-side Linux or Unix-like computer. Thus, if you want to deploy and run Paper-Hex on your own, you need a Linux/Unix compatible operating system.

Software Requirements:

🡪(Languages)

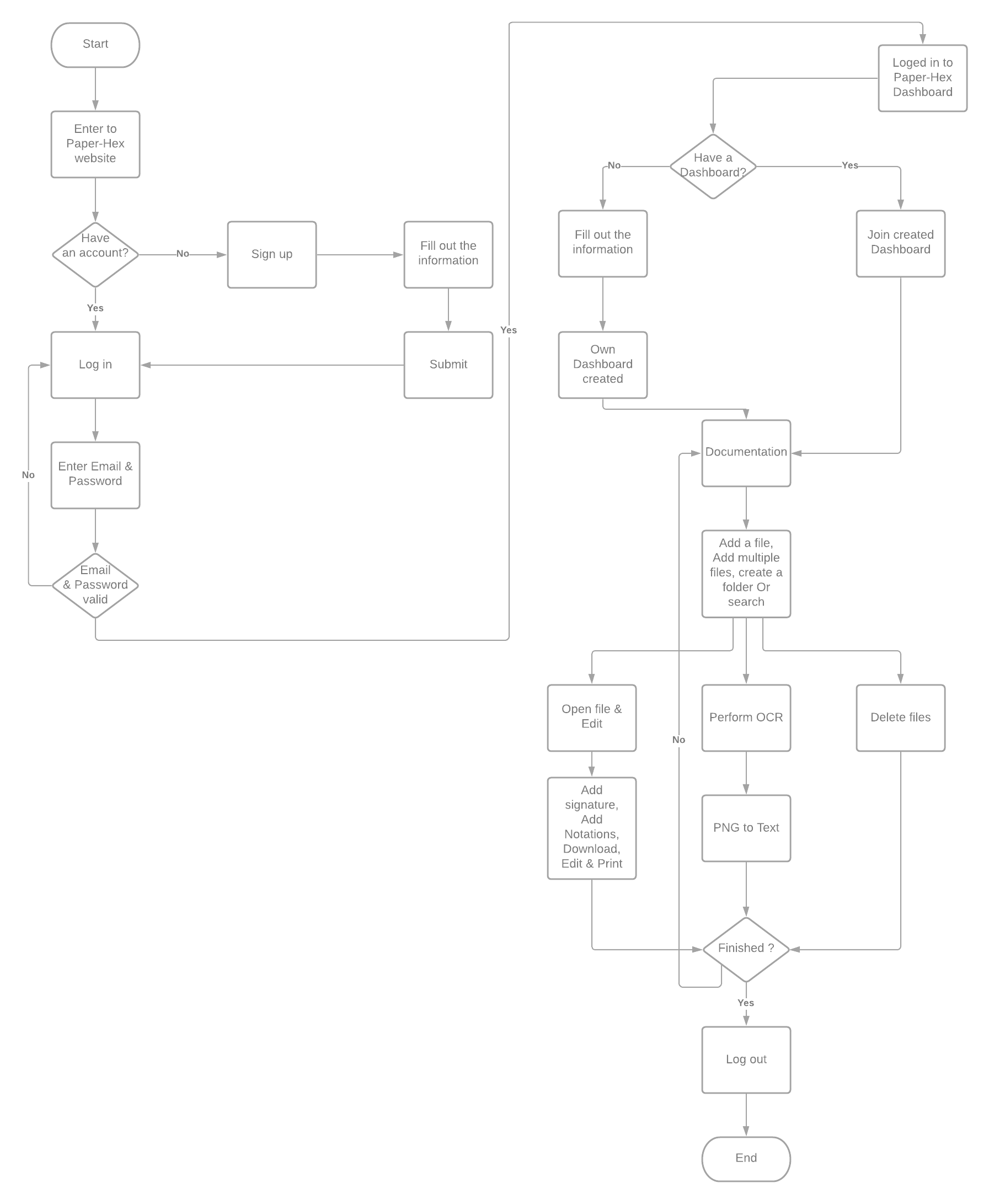
🡪(Languages)

🡪(Languages)

🡪(Languages)

🡪(Languages)

## Process Diagram



# Technical Specifications

The core requirements of the system are listed in this section. This template recommends organizing requirements by features rather than use cases. Features are system behaviors from the user’s point-of-view. The requirements of a feature are described by one or more use cases plus any additional narration that is necessary

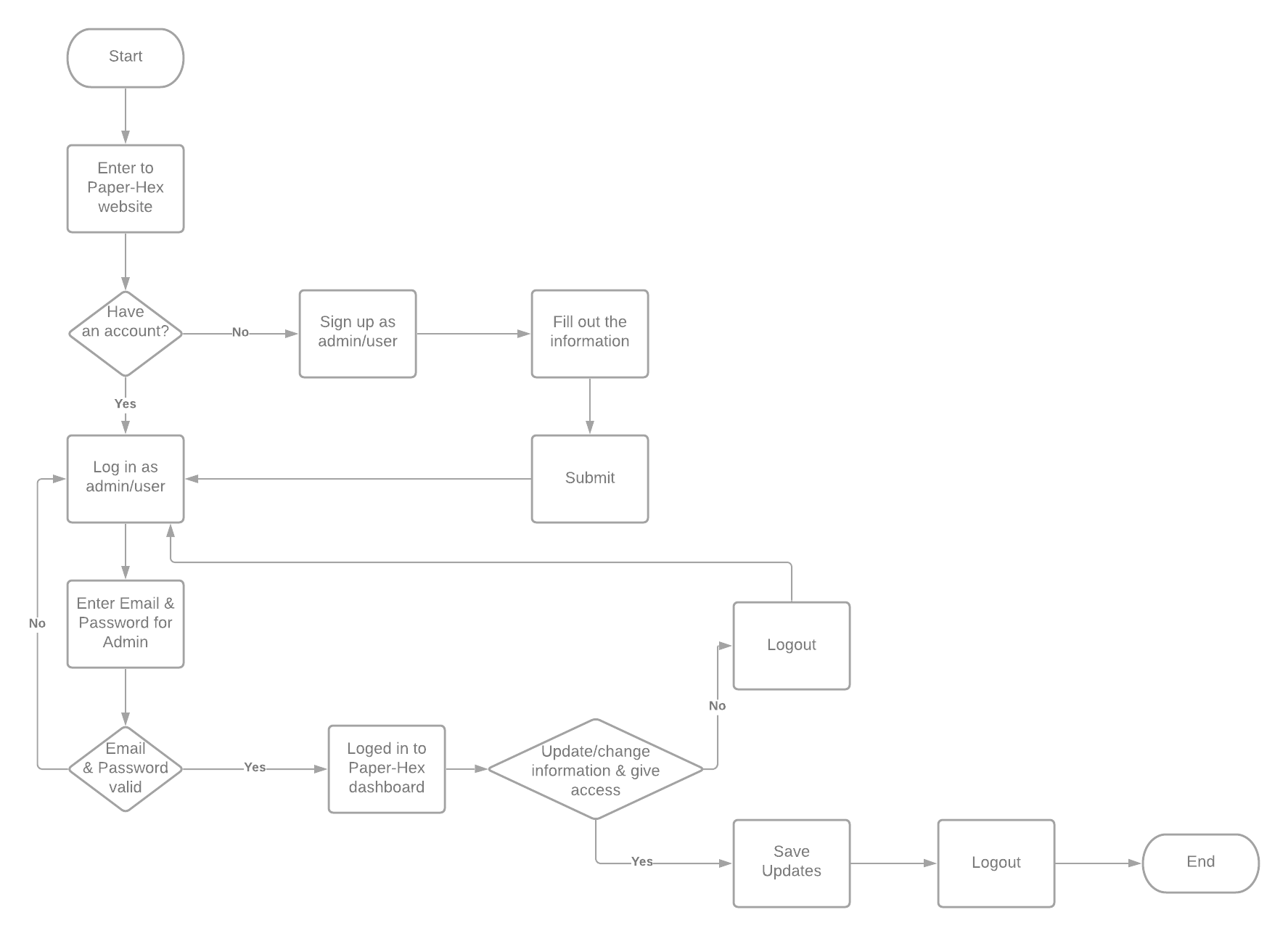
Features should be ranked and listed in priority order. Priority is determined by cost, risk and value. To prevent arguments over the exact values of these measures this template recommends using the values: high, medium and low. There should be a written understanding how the priorities listed here are used to determine what order features are delivered and what determines essential features, desirable features and optional features.

## Admin Module

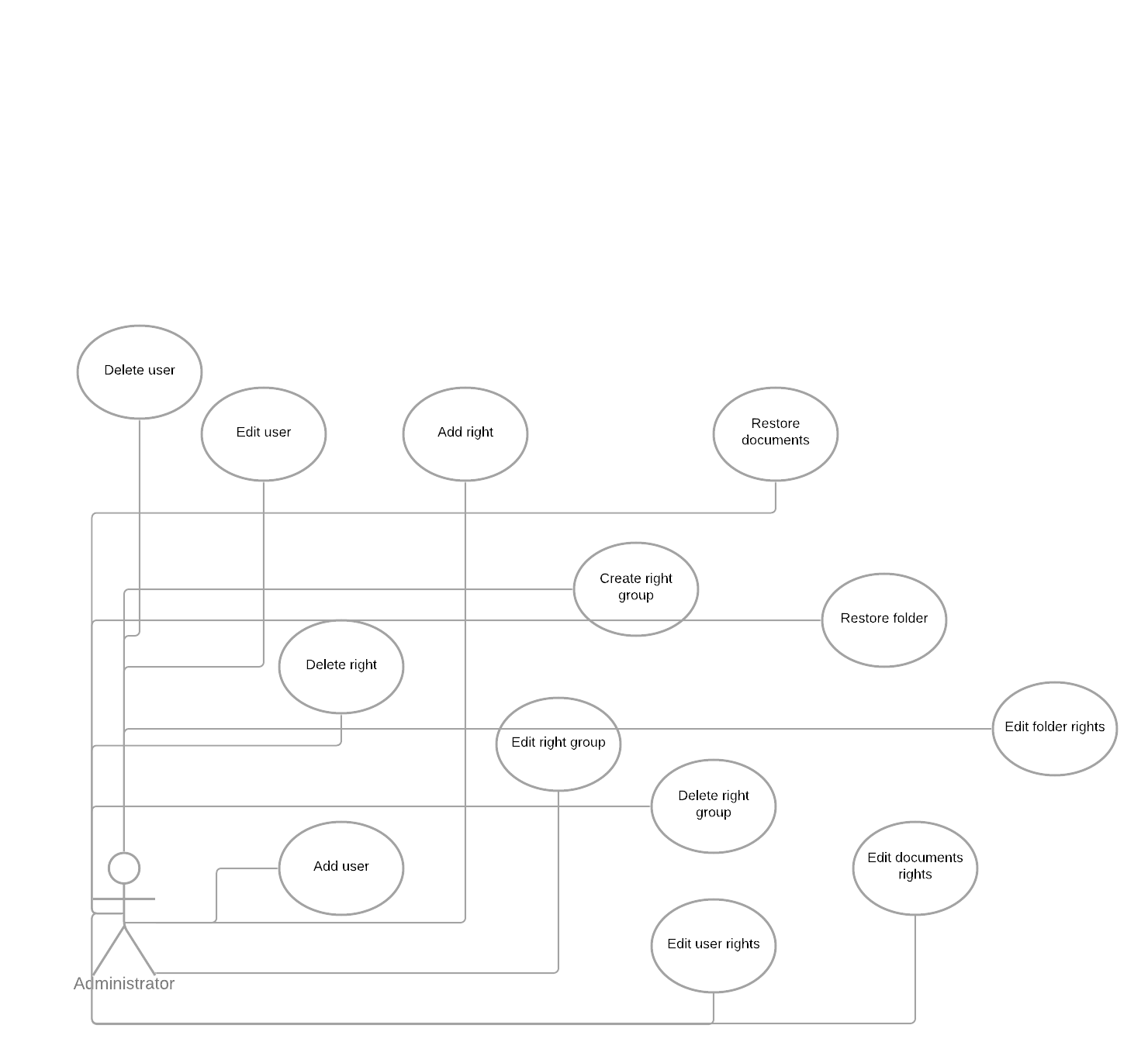
This module is used by Administrators of the staff to perform tasks that require administrative approval; therefore before any feature is applied the Administrator must authenticate them by inserting their usernames and passwords.

The features are present in the user module but can only be accessed after Administrative authentication.

### Flow Chart



### Features



Logout from current Admin account

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Intermediate |
| **Source** | **Availability** | **Effort** | |
| Organizational | 26-02-2021 | 1 personnel | 1.5 weeks |
| **Type** | | | |
| Input. | | | |
|  | | | |

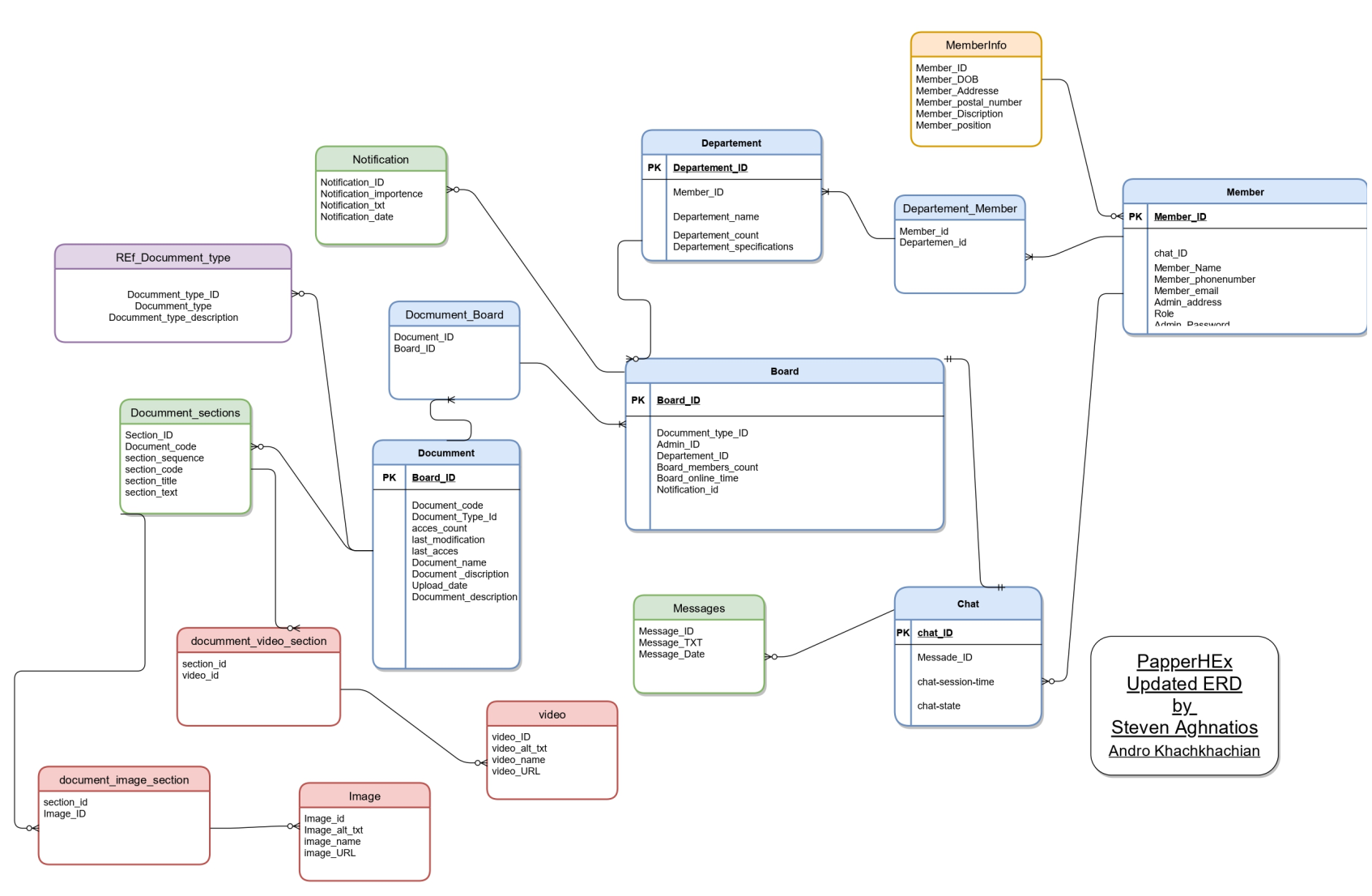
Giving access to users

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Difficult |
| **Source** | **Availability** | **Effort** | |
| Organizational | 04-04-2021 | 2 personnel | 1.5 months |
| **Type** | | | |
| Operation. | | | |

Add a new User.

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Intermediate |
| **Source** | **Availability** | **Effort** | |
| Organizational | 15-04-2021 | 1 personnel | 1 weeks |
| **Type** | | | |
| Input. | | | |

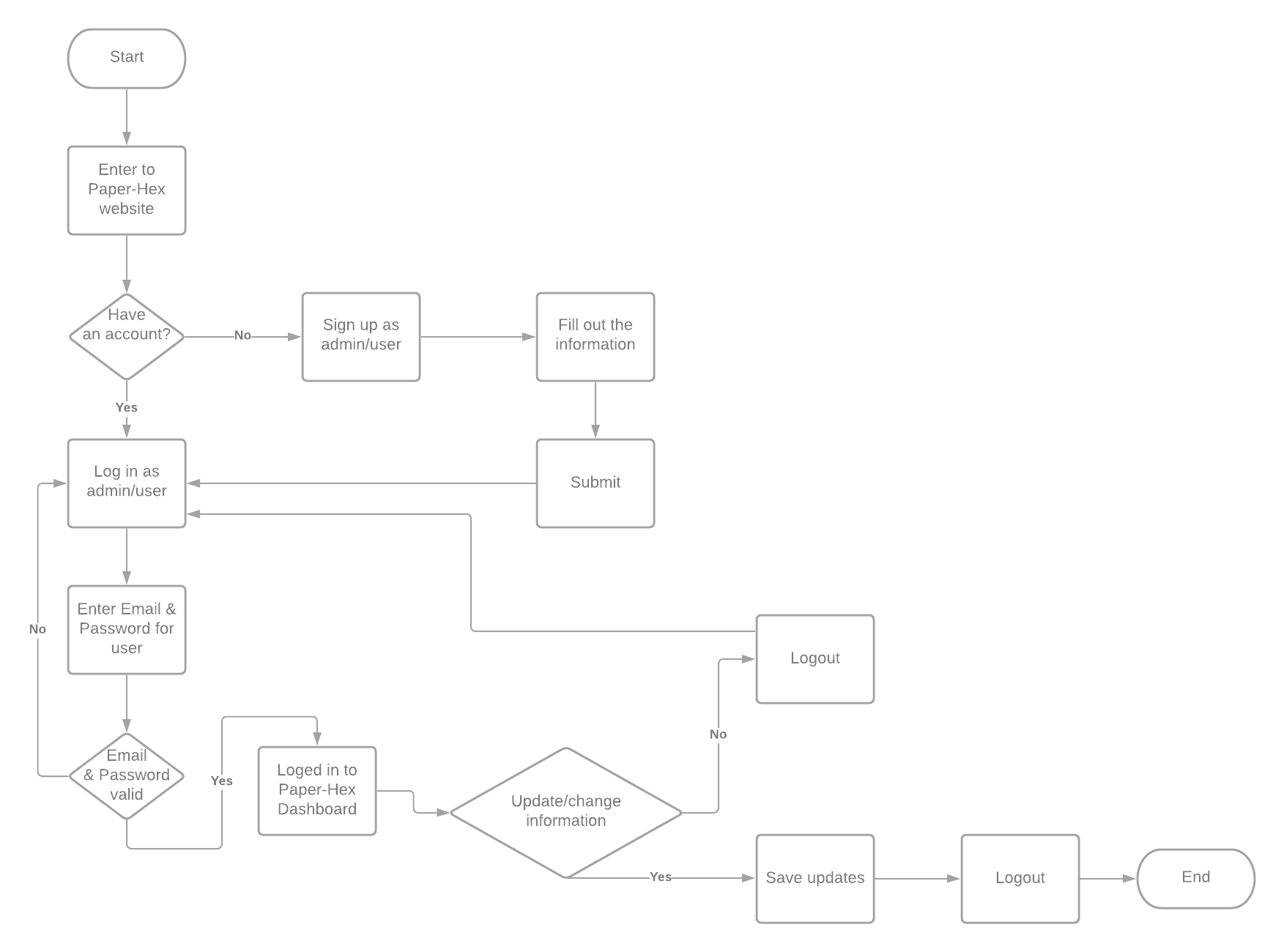
### ERD Diagram



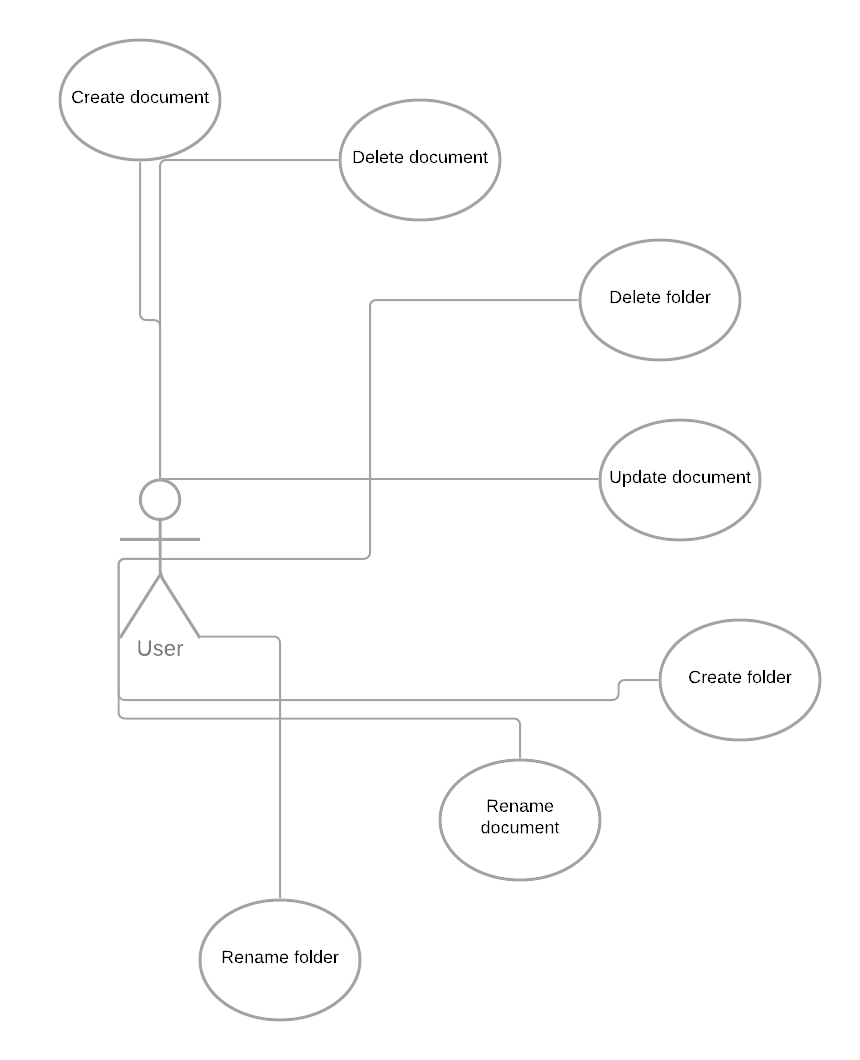
## User Module

This module is used by the user.

### Flow Chart



### Features



OCR

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Difficult |
| **Source** | **Availability** | **Effort** | |
| Organizational | 22-03-2021 | 2 personnel | 1 months |
| **Type** | | | |
| Input. | | | |
|  | | | |

Create Folder

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Intermediate |
| **Source** | **Availability** | **Effort** | |
| Organizational | 14-02-2021 | 1 personnel | 1 weeks |
| **Type** | | | |
| Input. | | | |

Delete files

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Easy |
| **Source** | **Availability** | **Effort** | |
| Organizational | 20-02-2021 | 1 personnel | 1 week |
| **Type** | | | |
| Input. | | | |

Search for information, folder, file & data

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Intermediate |
| **Source** | **Availability** | **Effort** | |
| Organizational | 7-2-2021 | 1 personnel | 1 month |
| **Type** | | | |
| Output. | | | |

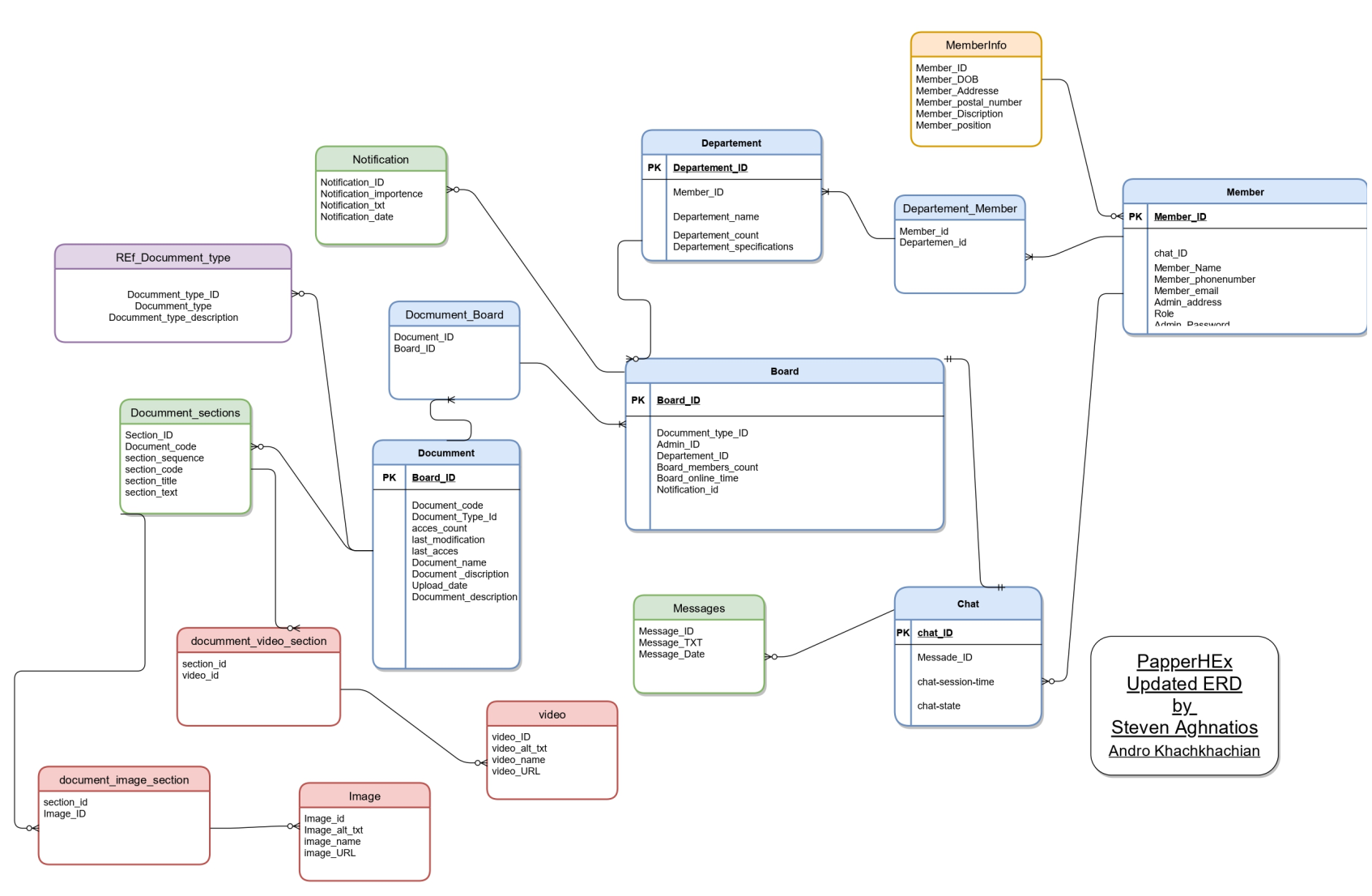
Update documents/folders

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Difficult |
| **Source** | **Availability** | **Effort** | |
| Organizational | 26-04-2019 | 1 personnel | 2 weeks |
| **Type** | | | |
| Input. | | | |

Logout from current user account

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Intermediate |
| **Source** | **Availability** | **Effort** | |
| Organizational | 10-02-2021 | 1 personnel | 2 weeks |
| **Type** | | | |
| Input. | | | |

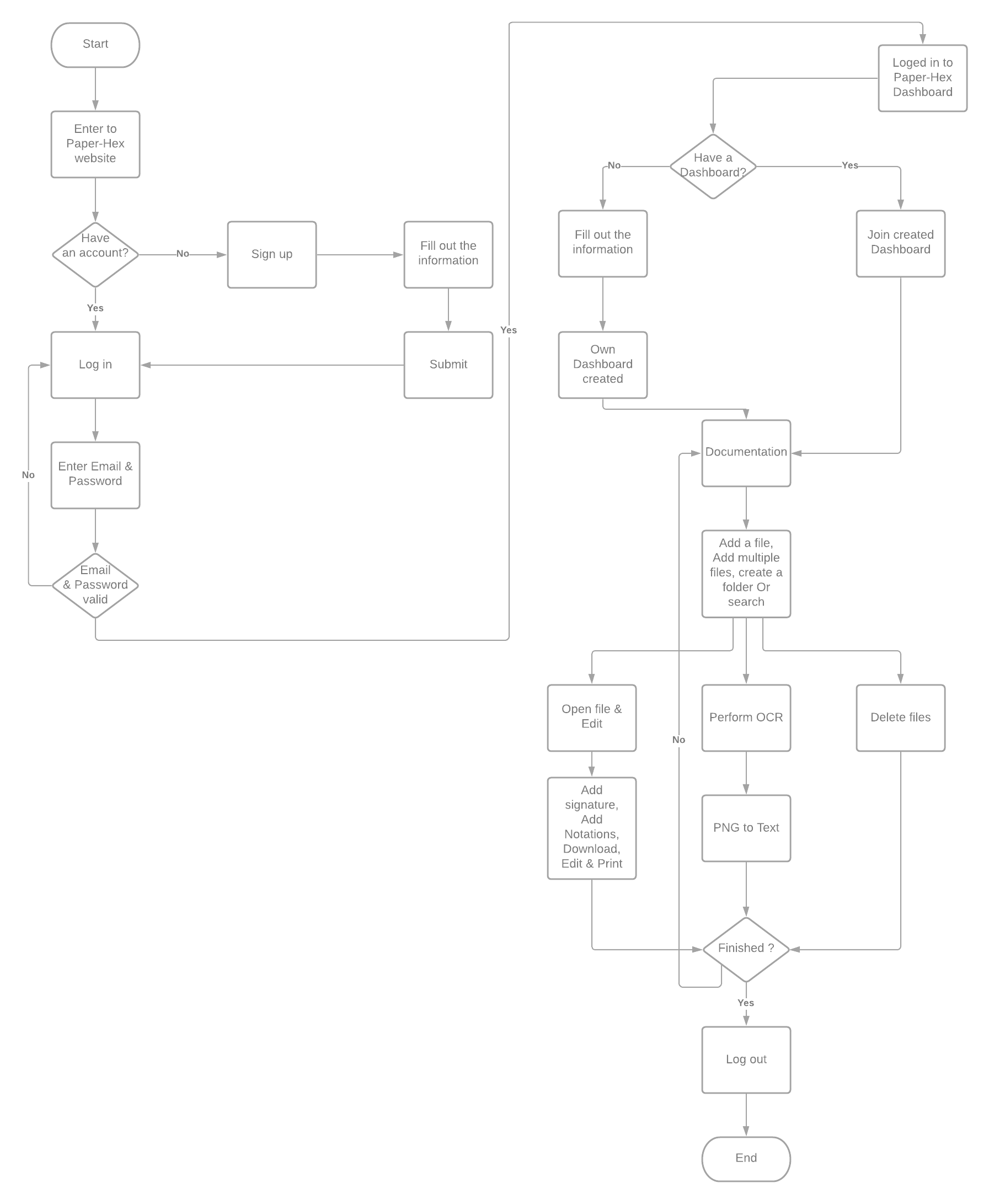
### ERD Diagram



## Client Module

This module is used by the client.

### Flow Chart

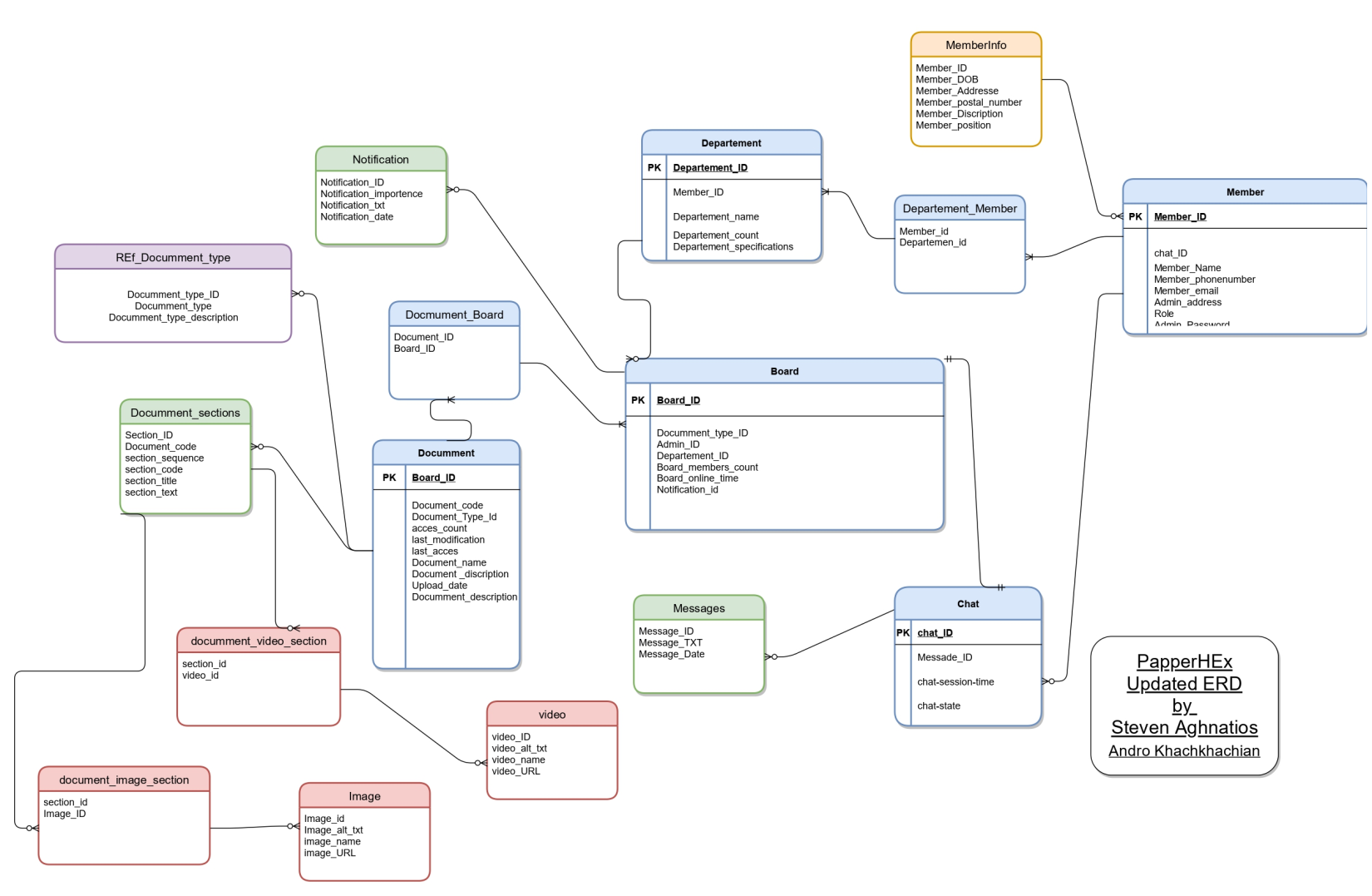


### Features

Save updated Information.

|  |  |  |  |
| --- | --- | --- | --- |
| **Priority** | **Criteria** | **Status** | **Feasibility** |
| High | Mandatory | Approved | Difficult |
| **Source** | **Availability** | **Effort** | |
| Organizational | 25-03-2021 | 1 personnel | 2 weeks |
| **Type** | | | |
| Input. | | | |

### ERD Diagram



# Contract

## Introduction

This will serve as a contract of production for Paper-Hex Windows Software. It details the requirements requested for this software and the details of each requirement proposed.

## The authors of the software license

Aghnatios Steven

&

Khachkhachian Andrew

## Preamble

The contract will detail all the objects to be delivered and the deliverable within them. It will serve as a guideline for the development team and a contract to ensure that the client’s requests are met.

## Objects delivered to the client

* It [OCRs](https://papermerge.readthedocs.io/en/latest/glossary.html#ocrs) your documents.
* It extracts text from your scans (which might be PDF, TIFF, JPEG or PNG) and indexes it.
* Helps you fix scanned documents issues.
* Reduced storage space.
* Provides you nice user interface to easily browse your documents.
* It allows you to save files online and access them anywhere from any smart-phone, tablet, or computer.
* Helps you instantly find your documents:
  + 🡪Based on Extracted-Text
  + 🡪Based on Tags and Folders
  + 🡪Based on [Metadata](https://papermerge.readthedocs.io/en/latest/metadata.html)

## Objects not delivered to the client

🡪Training to the staff will not be provided by the authors.

🡪Maintenance fees are paid by the client.

🡪It does not stay in your way when you make decisions about your data.

🡪It does not overwrite your original documents.

## Contract effective date

Effective immediately when software is installed.

## Associated services

Installation of the project.

## Additional services

Further follow-up in case of confusion.

## Warranty and maintenance

No warranty or free maintenance will be provided.

## Languages

English