

Florida International University
School of Computing and Information Sciences

Software Engineering Focus

Final Deliverable

Project Title:

Patentify - Data Labeling by Active Learning

Team Members:

Capstone I: Jordy Araujo, Katalina Keyser, Josercarlos Lusbel, Samuel Muvdi, Freddy Somarriba, Luis Gil

Capstone II: Simarjeet Chhabra, Carlos Cordero, Manuel Garcia, Eric Tu

Product Owner(s): Dr. Mark Finlayson, Tisa Islam Erana

Instructor: Masoud Sadjadi

The MIT License (MIT)

Copyright (c) 2016 *Florida International University*

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Abstract

This document presents the information necessary to gain a good understanding of the Patentify project. The project aims to create a solution for the United States Patent and Trademark Office (USPTO) to label and classify patents. Patentify is a web application that was developed using the Express.js, MongoDB, Node.js, and React.js

Table of Contents

Introduction

Current System.....	5
Purpose of New System.....	5

User Stories

Implemented User Stories.....	7
Pending User Stories.....	12

Project Plan

Hardware and Software Resources.....	14
Sprints Plan.....	15
Sprint 1.....	15
Sprint 2.....	16
Sprint 3.....	17
Sprint 4.....	18
Sprint 5.....	19
Sprint 6.....	20
Sprint 7.....	21

System Design

Architectural Patterns.....	24
System and Subsystem Decomposition.....	24
Deployment Diagram.....	25
Design Patterns.....	26

System Validation	27
--------------------------------	----

Glossary	28
-----------------------	----

Appendix	30
-----------------------	----

Appendix A - UML Diagrams.....	30
Eric Tu 0.....	3
Simarjeet Chhabra.....	33
Manuel Garcia.....	36
Carlos Cordero.....	39
Appendix B - User Interface Design.....	42
Appendix C - Sprint Review Reports.....	45
Appendix D - User Manuals, Installation/Maintenance Document, Shortcomings/Wishlist Document and other documents.....	57
Installation/Maintenance Document.....	57
Prerequisite Technologies.....	57
Shortcomings/Wishlist.....	68
References.....	70

INTRODUCTION

The United States Patent & Trademark Office, also known as USPTO, works with millions of patents a day. Every patent has its own label and people who annotated those patents. It is difficult to keep track of all patents that are being annotated and passed on. As a result, USPTO decided to work with FIU to an application that could allow users to annotate patents easily and have the database keep track of those annotations. The idea of the application is to have a list of patents, each categorized by different labels that users annotate individually. The database will then update the patents in the database automatically without any user interaction. This is significantly better than having to keep track of everything by hand.

Patentify is the web application where patents are uploaded onto the database and presented to users for annotation purposes. By having an automated way of dealing with the work, it provides a less time-consuming and simpler way of annotating patents without the need to track everything manually.

Current System

The current system being used for labeling patents is not standardized and it takes a great deal of time. There is also not a visual interface for annotators to label patents nor is there an established system to ensure consistency in the annotations. There is also not a simple way to check the quality of annotations. Since the annotations are currently stored in spreadsheets, the patents are separated from them and have to be manually opened when annotations are being done.

Purpose of New System

Through communication with the product owners, the team was able to gather what improvements they considered to be important this semester. Meetings were conducted with the product owner in every sprint to display progress on the requested features and gather feedback along with discussing possible next steps.

The main objective of this project is to implement machine learning onto the web application. This would enable active learning algorithms to label patents. Other goals for this semester consisted of fully deploying the project, implementing a patent search feature, and introducing workload distribution to the application. Fully deploying the project was the first goal of this semester as that needed to be completed before additional work could be done. Implementing a patent search feature was important as well since users now have the ability to search for any patent that they need. Lastly, introducing workload distribution was significant as well since users now have different patents assigned to them in a queue so that they can work on annotating another patent after they finish their current one.

Improvements were made to the front-end of the website as requested by the product owners. These improvements consisted of creating a larger patent view for the user and implementing hints to provide the user with any additional information that they need.

USER STORIES

Implemented User Stories

- User Story 1 - Setup older project onto current systems
 - Description - As a developer, I want to be able to download and setup the older project the Product Owner gave us so that we can work on the project.
- User Story 2 - Review MongoDB/Express
 - Description - As a developer, I want to learn Express so that I can access the back-end of the project.
- User Story 3 - Review Node/React
 - Description - As a developer, I want to review Node and React so that I can work with the project.
- User Story 4 - Review HTML/CSS/JS/Bootstrap
 - Description - As a developer, I want to be able to learn HTML/CSS/JS/Bootstrap so that I can work on the front-end of the project.
- User Story 5 - Fix any bugs that may occur during setup
 - Description - As a developer, I want to be able to solve issues regarding setup so that I can work on the project without any errors.
- User Story 6 - Review machine learning foundation
 - Description - As a developer, I want to learn more about machine learning so that we can train the AI in this project
- User Story 7 - Learn about Github
 - Description - As a Capstone 1 member I would like to learn how to utilize Github to help create and use branches with the team.
- User Story 8 -
 - Description - As a developer, I want to learn about Python to complete this project
- User Story 9 -
 - Description - As a developer, I want to understand the documentation, code, and get an idea of how the previous project was made so that I can work with the project efficiently

- User Story 10 -
 - Description - As a developer, I want to discuss deployment of the project so that the team has an idea of how to do so next Sprint
- User Story 11 -
 - Description - As a developer, I want to be able to have the database be functional and not have any bugs.
- User Story 12 -
 - Description - As a user, I want access to the project online so that I can work with the application
- User Story 13 -
 - Description - As a user, I want to be able to sign into the website so that I use my account
- User Story 14 -
 - Description - As a user, I want to be able to sign up to the website so that I can have an account with all my searches and patents saved
- User Story 15 -
 - Description - As a developer, I want to be able to test the sign-in and login features so that we can see if it works
- User Story 16 -
 - Description - As a developer, I want to make a template for a page to search patents so that users can see a display for this feature
- User Story 17 -
 - Description - As a developer, I want to create a GitHub Repo for the project so that we can save changes to the project as we make them
- User Story 18 -
 - Description - As a developer, I want to know more about the database and the AI so that I can work with it as we are applying new features
- User Story 19 -
 - Description - As a user, I want to be able to search a patent by number so that I can search for the patent easier
- User Story 20 -
 - Description - As a developer, I want to implement backend functionality for search patents by number so that users can access the backend for patents
- User Story 21 -

- Description - As a user, I want to be able to find the patents I want so that I can annotate on to them
- User Story 22
 - Description - As a user, I want to be able to load patents and see them so that I can annotate on those patents
- User Story 23
 - Description - As a user, I want to be able to look for a specific patent so that I can annotate on that specific patent I want
- User Story 24
 - Description - As a user, I want to have a page where I can use the patent search page so that I can look for a specific patent in the database
- User Story 25
 - Description - As a developer, I want to be able to test the patent search so that I know the patents are found in the database and displayed to user
- User Story 26
 - Description - As a developer, I want to continue looking at the machine learning model so that I can understand how that works and how to apply features to it
- User Story 27 -
 - Description - As a user, I want to access the patents quickly so that I do not have to wait too long for the application to load.
- User Story 28 -
 - Description - As a user, I want to have access to annotated patents so that I can see what changes have been made
- User Story 29 -
 - Description - As a developer, I want to discuss how to enhance the machine learning model so that it can find patents efficiently and effectively
- User Story 30 -
 - Description - As a user, I want to have keywords for each question so that it is easier to understand what each question entails
- User Story 31 -
 - Description - As a user, I want to be able to have a none of the above option so that I can mark that patent as none of the above options
- User Story 32 -

- Description - As a user, I want to have the buttons closer together so that I can access those buttons easier
- User Story 33 -
 - Description - As a user, I want to be able to read the whole patent so that I can get a better grasp of what the patent is before answering the questions
- User Story 34 -
 - Description - As a user, I want notes next to each question so that I can understand what entails for that specific patent category
- User Story 35 -
 - Description - As a user, I want to be able to use the app on Firefox so that the application can be used on all browsers
- User Story 37 -
 - Description - As a user, I want to work on other patents that are in other people's queues so that I can help others with queued patents that need to be worked on immediately
- User Story 38 -
 - Description - As a user, I want to be able to see the queue so that I can see what patents take priority and what do not
- User Story 39 -
 - Description - As a user, I want to be able to see the patent clearly so that I can determine the labels for it
- User Story 40 -
 - Description - As a developer, I need to fix bugs to patent search so that it does not effect the other parts of the application
- User Story 41
 - Description - As a developer, I want to be able to setup queues into the backend so that I can test the frontend users access to them
- User Story 42
 - Description - As a user, I want a page to see queues to that I can access the queues easily
- User Story 43
 - Description - As a user, I want to be able to hover on top of an icon so that I can see the information for each label
- User Story 44

- Description - As a user, I want to see how many patents are in queue so that I can gauge how much work I have
- User Story 45
 - Description - As a user, I want to have easier access to viewing the whole patents so that I do not have to go through each page one by one
- User Story 46
 - Description - As a user, I want a list of queues so that I can know how much work I have left and what patents they are
- User Story 47
 - Description - As a user, if I search for a patent, it will have its previous annotations so that I can fix them or see if they are correct.
- User Story 48
 - Description - As a user, I do not want to get the same patents repeatedly so that I can work on other patents once they have been skipped/submitted
- User Story 49
 - Description - As a developer, I want the database to keep track of all the patents a user has in their queue so that it is not randomly selected
- User Story 50
 - Description - As a user, I want to be able to keep the same patent after logging out so that I can work on it later when I am available
- User Story 51
 - Description - As a user, I do not want to see a patent that has already been annotated so that I can focus on other patents that haven't been worked on
- User Story 52 -
 - Description - As a user, I want the skip button under the submit button so that there is more space for the patent view
- User Story 53 -
 - Description - As a developer, I want to test scenarios to see if the queue is working such as when two people access and annotate a patent at the same time so that these situations may not cause problems
- User Story 54 -
 - Description - As a developer, I want to give a list of labeled patents to the PO so that the PO may have a list of everything that has been edited and changed

- User Story 57 -
 - Description - As a user admin, I want a page to see the active user queues so that I can access those lists easier
- User Story 58 -
 - Description - As a user, I want to only see my own annotations so that I do not have any biases when annotating on my own
- User Story 59 -
 - Description - As a user admin, I want to be able to see who annotated what patents and what they annotated so that I can get information on each user's annotations
- User Story 60 -
 - Description - As a user, I want an even larger patent view screen so that I can view the patent easier
- User Story 61 -
 - Description - As a developer, I want to edit the hint information so that users can get a clearer idea of what each label means
- User Story 62 -
 - Description - As a developer, I want to include information and comments about my code so that the future groups will be able to understand the code
- User Story 67 -
 - Description - As a user, I want an easier to use page for viewing active user queues so that I can see clearly who is working
- User Story 69 -
 - Description - As a developer, I want to make a backup of the program so that future groups can work on it without changing anything important

Pending User Stories

- User Story 36 -
 - Description - As a user, I want the option to save patents for later so that I can work on that specific patent for later
- User Story 63 -
 - Description - As a user, I want to know how many pages that the document has so that I can go through the document via next/previous page
- User Story 64 -
 - Description - As a user admin, I want to see new entries for each user annotation on a patent so that I can see what the users annotated for that patent
- User Story 65 -
 - Description - As a developer, I want to fix the bug where User 2's annotations on Patent 1 erases User 1's annotations on Patent 1 so that annotations on a patent can be added altogether to Patent 1
- User Story 66 -
 - Description - As a developer, I want to test to see if sending queues as an admin works so that I can fix any bugs that occur while working on the functionality
- User Story 68 -
 - Description - As a developer, I want to send patents in a logged out user's queue back to the backend list of non-annotated patents so that it does not get stuck in just one user's queue

PROJECT PLAN

This section describes the planning that went into the realization of this project. This project incorporated the agile development techniques and as such required the sprints to be planned. These sprint plannings are detailed in the section. This section also describes the components, both software and hardware, chosen for this project.

Hardware and Software Resources

Hardware

FIU computer 1 - Server Host

Software

MongoDB - Database

Node - back-end

React - front-end

Visual Studio Code - Development environment

GitHub - Code Repository

Python - AI implementation

Sprints Plan

Attendees:

Product Owners:

- Tisa Islam Erana

Developers:

- Capstone 2
 - Manuel Garcia (MG)
 - Simarjeet Chhabra (SC)
 - Carlos Cordero (CC)
 - Eric Tu (ET)
- Capstone 1
 - Jordy Araujo (JA)
 - Katalina Keyser (KK)
 - Josecarlos Lusbel (JL)
 - Samuel Muvdi (SM)
 - Freddy Somarriba (FS)
 - Luis Gil (LG)

Sprint 1

Date: 8/30/2021

Start Time: 8:00 PM

End Time: 8:35 PM

- User Story 1 - Setup older project onto current systems
 - Description - As a developer, I want to be able to download and setup the older project the Product Owner gave us so that we can work on the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 2 - Review MongoDB/Express
 - Description - As a developer, I want to learn Express so that I can access the back-end of the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

- User Story 3 - Review Node/React
 - Description - As a developer, I want to review Node and React so that I can work with the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 4 - Review HTML/CSS/JS/Bootstrap
 - Description - As a developer, I want to be able to learn HTML/CSS/JS/Bootstrap so that I can work on the front-end of the project.
 - Assigned to - JA,KK,JL,SM,FS, LG
- User Story 5 - Fix any bugs that may occur during setup
 - Description - As a developer, I want to be able to solve issues regarding setup so that I can work on the project without any errors.
 - Assigned to - MG, SC, CC, ET
- User Story 6 - Review machine learning foundation
 - Description - As a developer, I want to learn more about machine learning so that we can train the AI in this project
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 7 - Learn about Github
 - Description - As a Capstone 1 member I would like to learn how to utilize Github to help create and use branches with the team.
 - Assigned to - JA,KK,JL,SM,FS, LG

Sprint 2

Date: 9/13/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 11 -
 - Description - As a developer, I want to be able to have the database be functional and not have any bugs.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 12 -
 - Description - As a user, I want access to the project online so that I can work with the application
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 13 -

- Description - As a user, I want to be able to sign into the website so that I use my account
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 14 -
 - Description - As a user, I want to be able to sign up to the website so that I can have an account with all my searches and patents saved
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 15 -
 - Description - As a developer, I want to be able to test the sign-in and login features so that we can see if it works
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 16 -
 - Description - As a developer, I want to make a template for a page to search patents so that users can see a display for this feature
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

Sprint 3

Date: 9/27/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 21 -
 - Description - As a user, I want to be able to find the patents I want so that I can annotate on to them
 - Assigned to - CC, ET, JA, SC, MG, SM
- User Story 22
 - Description - As a user, I want to be able to load patents and see them so that I can annotate on those patents
 - Assigned to - CC, ET, JA, SC, MG, SM
- User Story 23
 - Description - As a user, I want to be able to look for a specific patent so that I can annotate on that specific patent I want
 - Assigned to - MG, SC, CC, ET, KK, JL, FS, LG
- User Story 24
 - Description - As a user, I want to have a page where I can use the patent search page so that I can look for a specific patent in the database

- Assigned to - MG, SC, CC, ET, KK, JL, FS, LG
- User Story 25
 - Description - As a developer, I want to be able to test the patent search so that I know the patents are found in the database and displayed to user
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 26
 - Description - As a developer, I want to continue looking at the machine learning model so that I can understand how that works and how to apply features to it
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

Sprint 4

Date: 10/11/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 30 -
 - Description - As a user, I want to have keywords for each question so that it is easier to understand what each question entails
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 31 -
 - Description - As a user, I want to be able to have a none of the above option so that I can mark that patent as none of the above options
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 32 -
 - Description - As a user, I want to have the buttons closer together so that I can access those buttons easier
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 33 -
 - Description - As a user, I want to be able to read the whole patent so that I can get a better grasp of what the patent is before answering the questions
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 34 -
 - Description - As a user, I want notes next to each question so that I can understand what entails for that specific patent category

- Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 35 -
 - Description - As a user, I want to be able to use the app on Firefox so that the application can be used on all browsers
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 36 -
 - Description - As a user, I want the option to save patents for later so that I can work on that specific patent for later
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 37 -
 - Description - As a user, I want to work on other patents that are in other people's queues so that I can help others with queued patents that need to be worked on immediately
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG

Sprint 5

Date: 10/25/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 41
 - Description - As a developer, I want to be able to setup queues into the backend so that I can test the frontend users access to them
 - Assigned to - MG, SC, CC, ET, JA
- User Story 42
 - Description - As a user, I want a page to see queues to that I can access the queues easily
 - Assigned to - KK, JL, SM, FS, LG
- User Story 43
 - Description - As a user, I want to be able to hover on top of an icon so that I can see the information for each label
 - Assigned to - SC, CC, ET
- User Story 44
 - Description - As a user, I want to see how many patents are in queue so that I can gauge how much work I have
 - Assigned to - MG, SC, CC, ET, JA, SM

- User Story 45
 - Description - As a user, I want to have easier access to viewing the whole patents so that I do not have to go through each page one by one
 - Assigned to - KK, JL, SM, FS, LG
- User Story 46
 - Description - As a user, I want a list of queues so that I can know how much work I have left and what patents they are
 - Assigned to - MG, SC, CC, ET, JA

Sprint 6

Date: 11/8/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 52 -
 - Description - As a user, I want the skip button under the submit button so that there is more space for the patent view
 - Assigned to - CC, KK, JL, SM, FS, LG
- User Story 53 -
 - Description - As a developer, I want to test scenarios to see if the queue is working such as when two people access and annotate a patent at the same time so that these situations may not cause problems
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 54 -
 - Description - As a developer, I want to give a list of labeled patents to the PO so that the PO may have a list of everything that has been edited and changed
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 55 -
 - Description - As a user admin, I want to be able to send one patent list via queue to users of my choice so that they can annotate the same patents and have multiple annotations on those patents
 - Assigned to - MG, SC, CC, ET, JA
- User Story 56 -
 - Description - As a user admin, I want to see all active user queues so that I can get an idea of what patents are being worked on

- Assigned to - MG, SC, CC, ET, JA
- User Story 57 -
 - Description - As a user admin, I want a page to see the active user queues so that I can access those lists easier
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 58 -
 - Description - As a user, I want to only see my own annotations so that I do not have any biases when annotating on my own
 - Assigned to - MG, SC, CC, ET, JA
- User Story 59 -
 - Description - As a user admin, I want to be able to see who annotated what patents and what they annotated so that I can get information on each user's annotations
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 60 -
 - Description - As a user, I want an even larger patent view screen so that I can view the patent easier
 - Assigned to - ET, KK, JL, SM, FS, LG
- User Story 61 -
 - Description - As a developer, I want to edit the hint information so that users can get a clearer idea of what each label means
 - Assigned to - CC, SC, ET, KK, JL, FS, LG

Sprint 7

Date: 11/22/2021

Start Time: 8:00 PM

End Time: 8:30 PM

- User Story 62 -
 - Description - As a developer, I want to include information and comments about my code so that the future groups will be able to understand the code
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 63 -

- Description - As a user, I want to know how many pages that the document has so that I can go through the document via next/previous page
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 64 -
 - Description - As a user admin, I want to see new entries for each user annotation on a patent so that I can see what the users annotated for that patent
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 65 -
 - Description - As a developer, I want to fix the bug where User 2's annotations on Patent 1 erases User 1's annotations on Patent 1 so that annotations on a patent can be added altogether to Patent 1
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 66 -
 - Description - As a developer, I want to test to see if sending queues as an admin works so that I can fix any bugs that occur while working on the functionality
 - Assigned to - MG, SC, CC, ET, JA

SYSTEM DESIGN

1. Frontend

The frontend is created using React.js along with bootstrap libraries for HTML and CSS. An iframe is utilized to show the sign-up and login forms and to show the patent on the website.

2. Backend

The backend is implemented using Express.js which utilizes the Node.js framework. Routes are implemented for loading different pages. They also only allow authorized users to access specific pages. The backend communicates with the database to retrieve and send information to and from the frontend part of the application.

3. Database

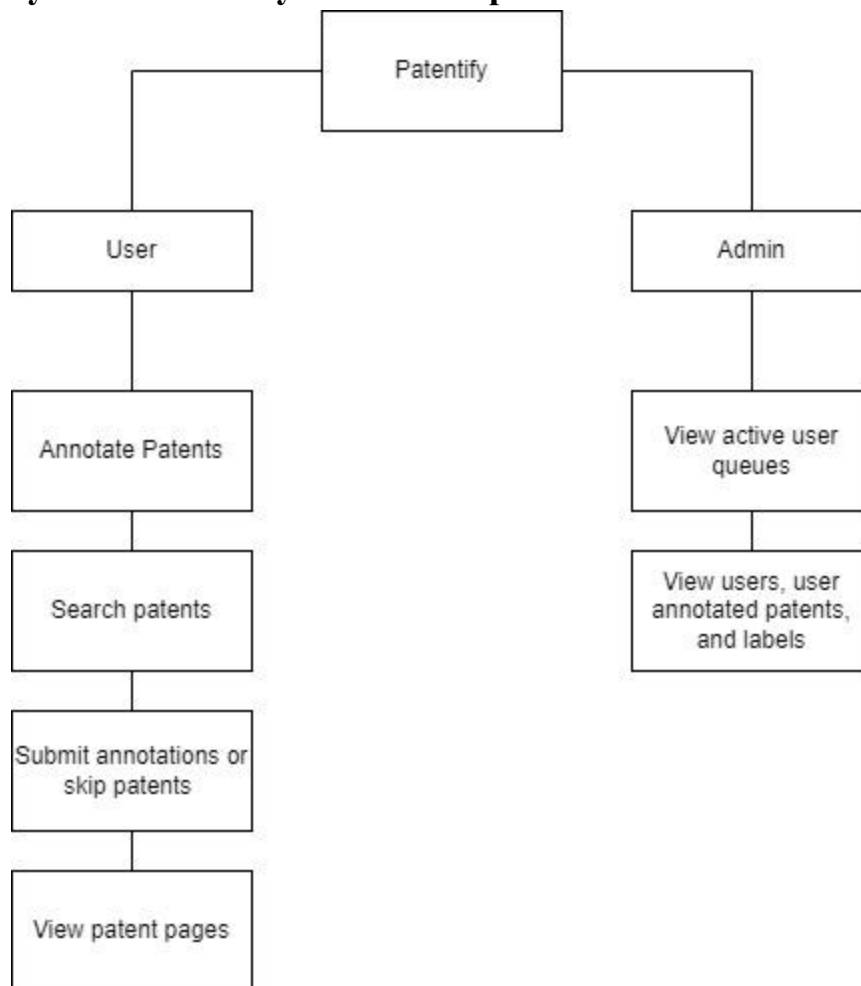
The data for this project is stored in MongoDB.

4. Active Learning

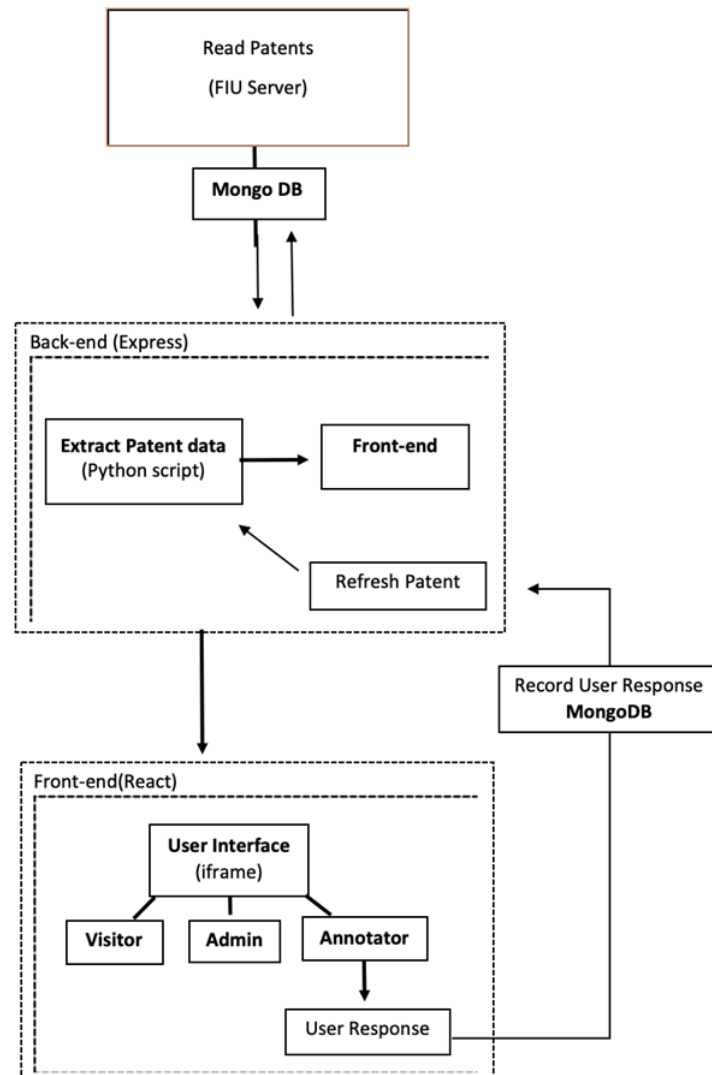
The active learning pipeline is implemented in Python. It allows for communication directly with the backend to read patents and apply machine learning algorithms to improve patent category prediction.

Architectural Patterns

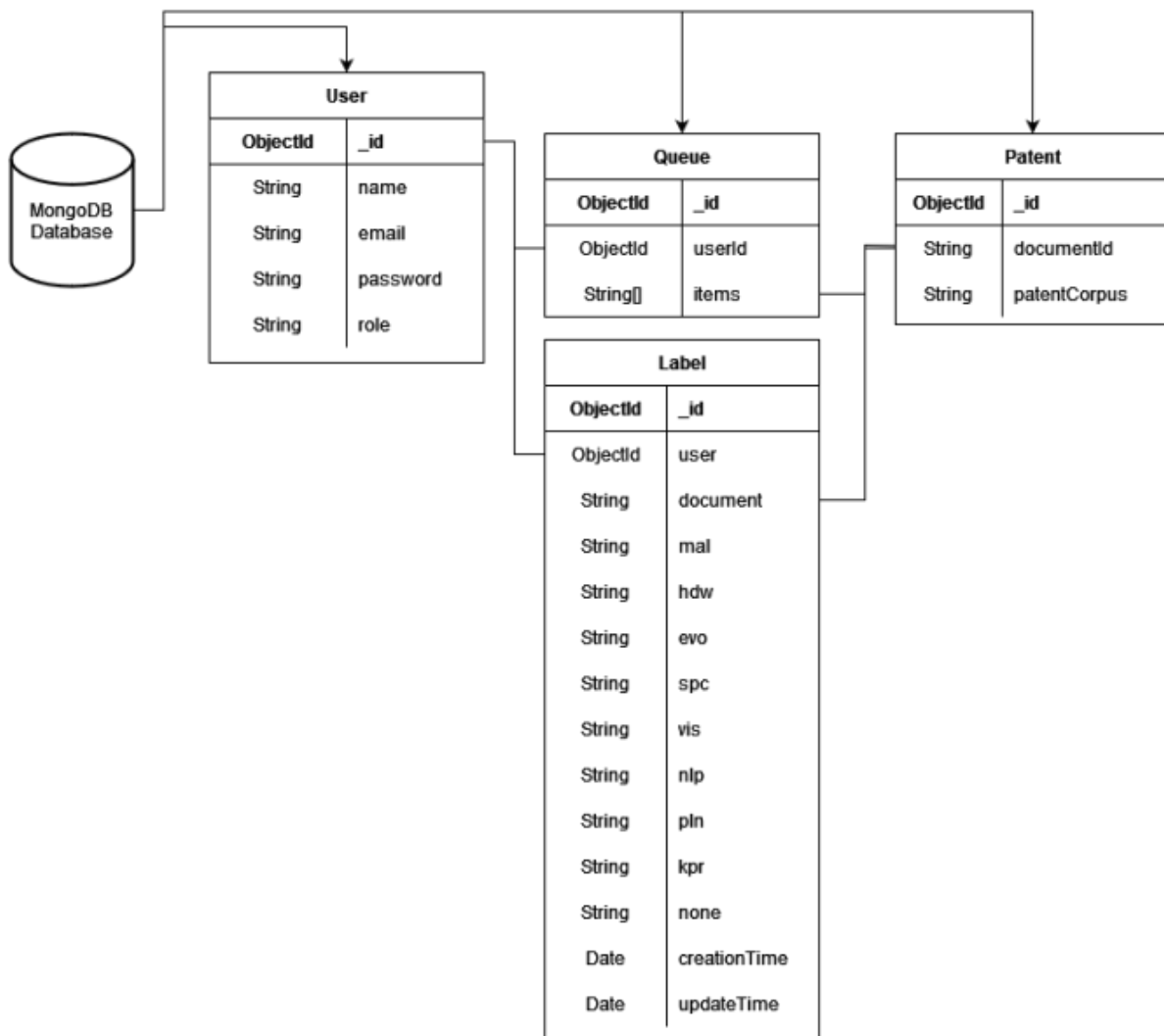
System and Subsystem Decomposition



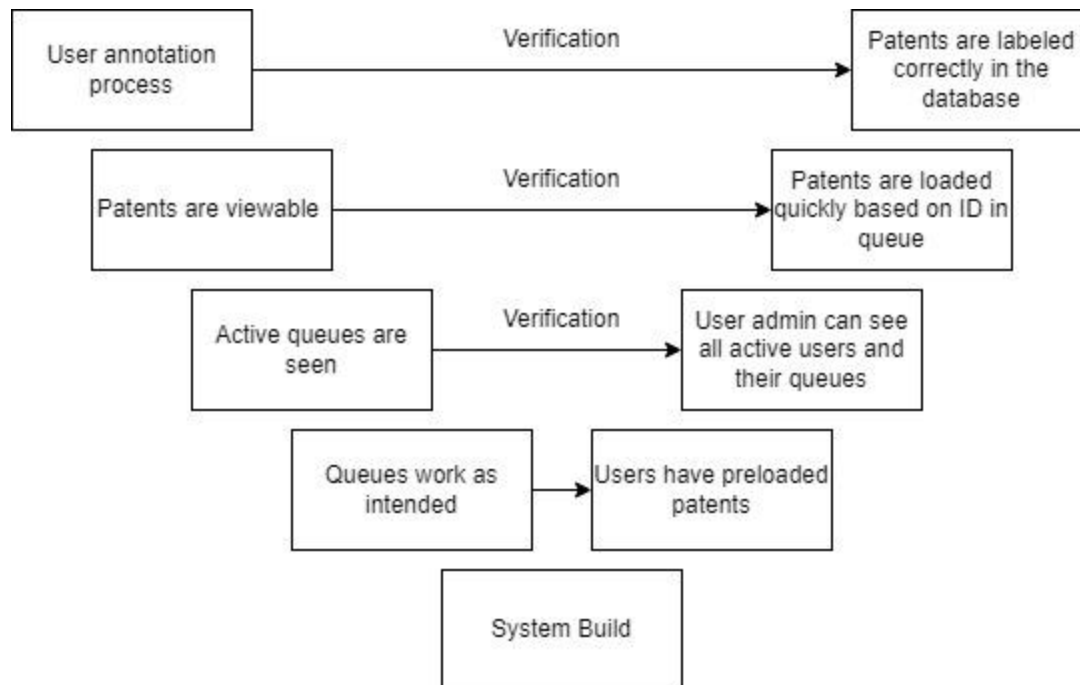
Deployment Diagram



Design Patterns



SYSTEM VALIDATION



GLOSSARY

Inter Annotator Agreement Measure: Cohen's Kappa is given by the following formula

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

AI Category Definitions:

Vision

Definition: This field is categorized as patents that deal with extracting information from images or videos such as edge detection that identifies edges in an image or color processing. (Russell et al. 2009)

Speech

Definition: Speech recognition is the ability to recognize words from a certain signal such as extracting information from a noisy signal. (Russell et al. 2009)

Natural Language Processing

Definition: This categorizes patents dealing with grabbing information and data from human language. (Russell et al. 2009)

Knowledge Processing

Definition: Knowledge processing deals with patents that are related to the world. In this case, it is anything with deriving new information from an existing base foundation of knowledge or even a collection of facts and relationships using many techniques. (Russell et al. 2009)

Evolutionary Computation

Definition: Evolutionary computation uses computation methods that deal with nature and evolution. Those algorithms can be that of genetic algorithms that select optimal mutations for a specific gene from a group of random mutations. (Russell et al. 2009)

Machine Learning

Definition: The field of Machine Learning deals with the computation learning models such as supervised learning classification algorithms that allow the AI to group certain inputs and data. (Russell et al. 2009)

Planning/Control

Definition: The field of planning and control are about AI methods that will execute plans to achieve a certain goal such as actions and states of the world as well as reasoning towards the consequences of such actions. (Russell et al. 2009)

AI Hardware

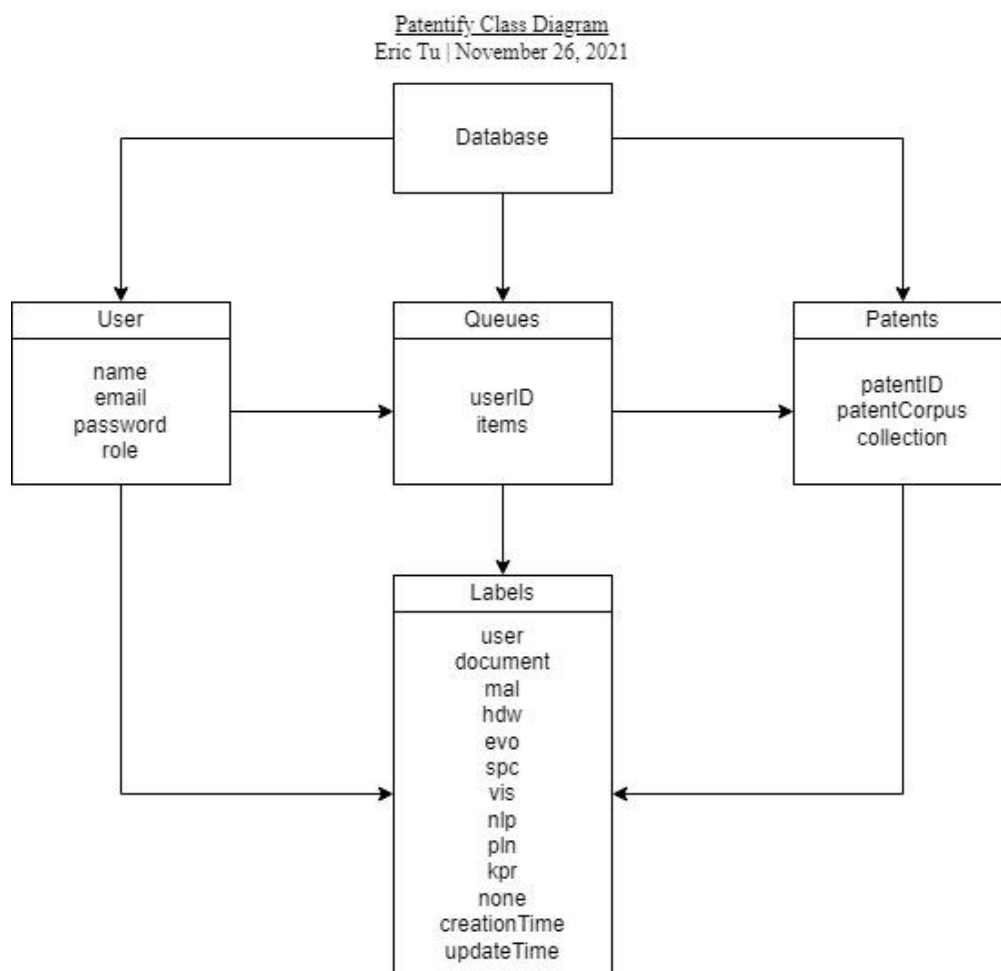
Definition: This field deals with all hardware that is used to store the AI software such as networks, computer components like memory, solid state technologies, and more. (Russell et al. 2009)

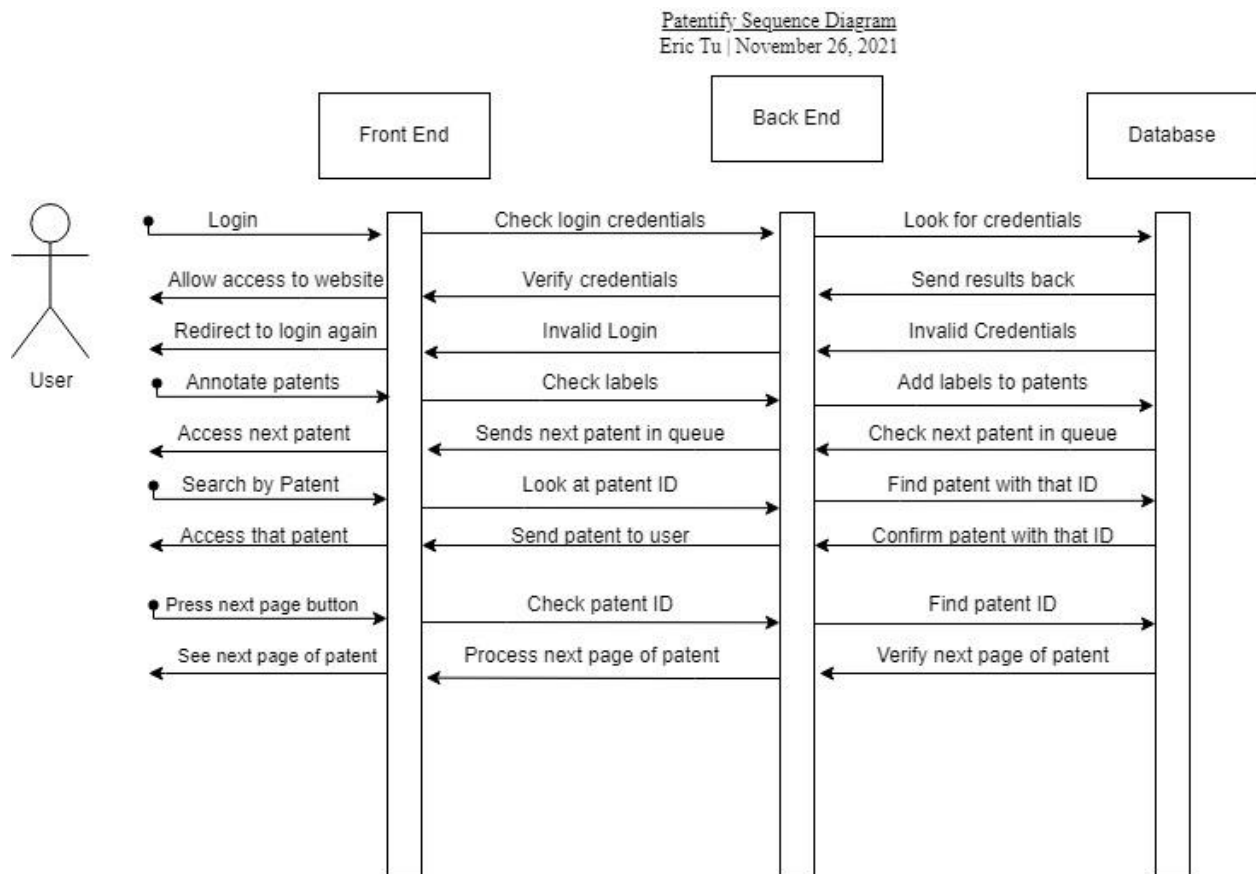
APPENDIX

Appendix A - UML Diagrams

Eric Tu

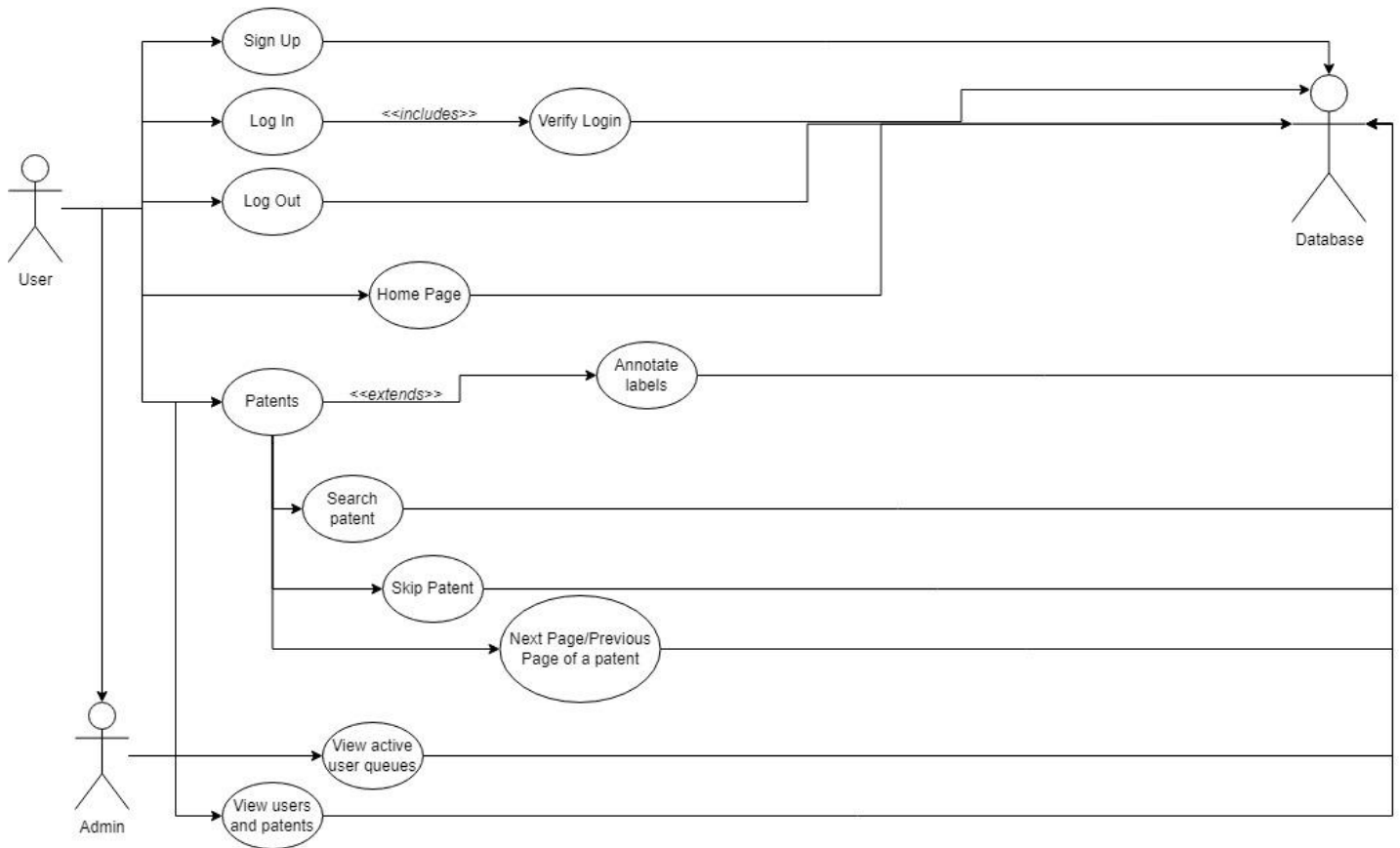
Eric Tu: Class Diagram



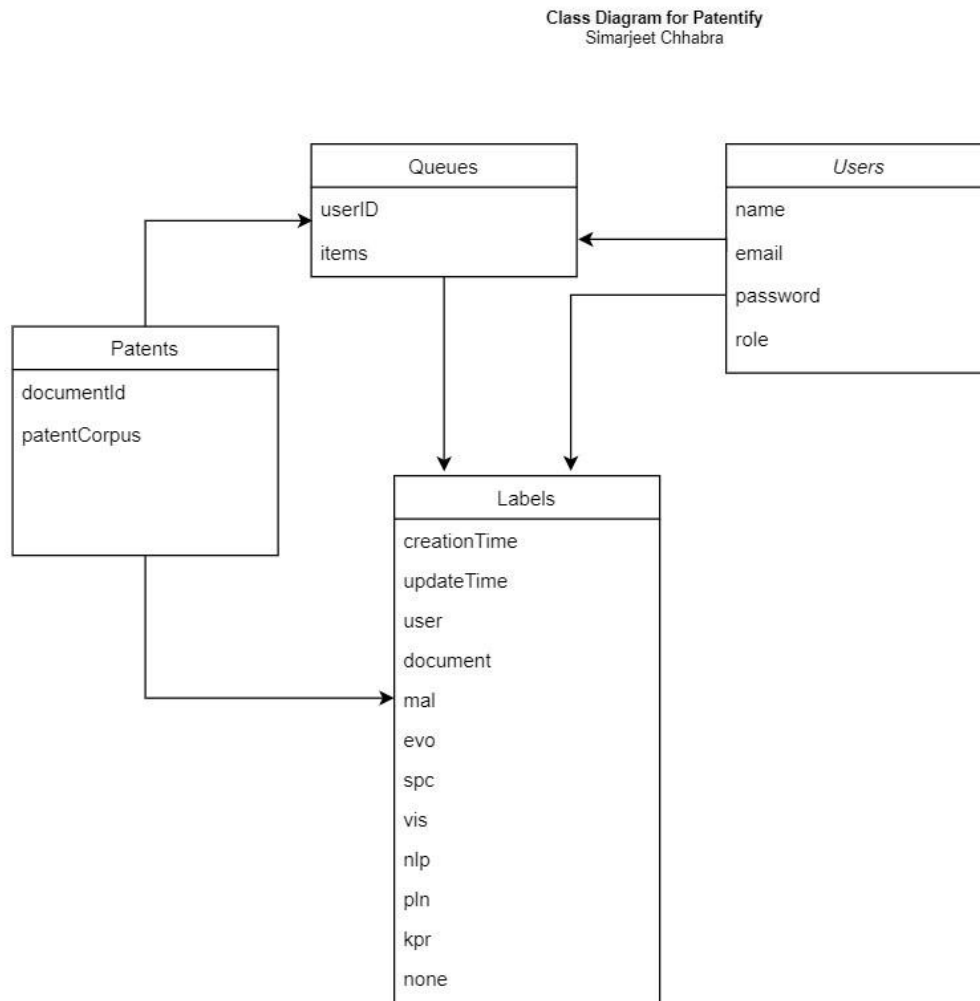
Eric Tu: Sequence Diagram

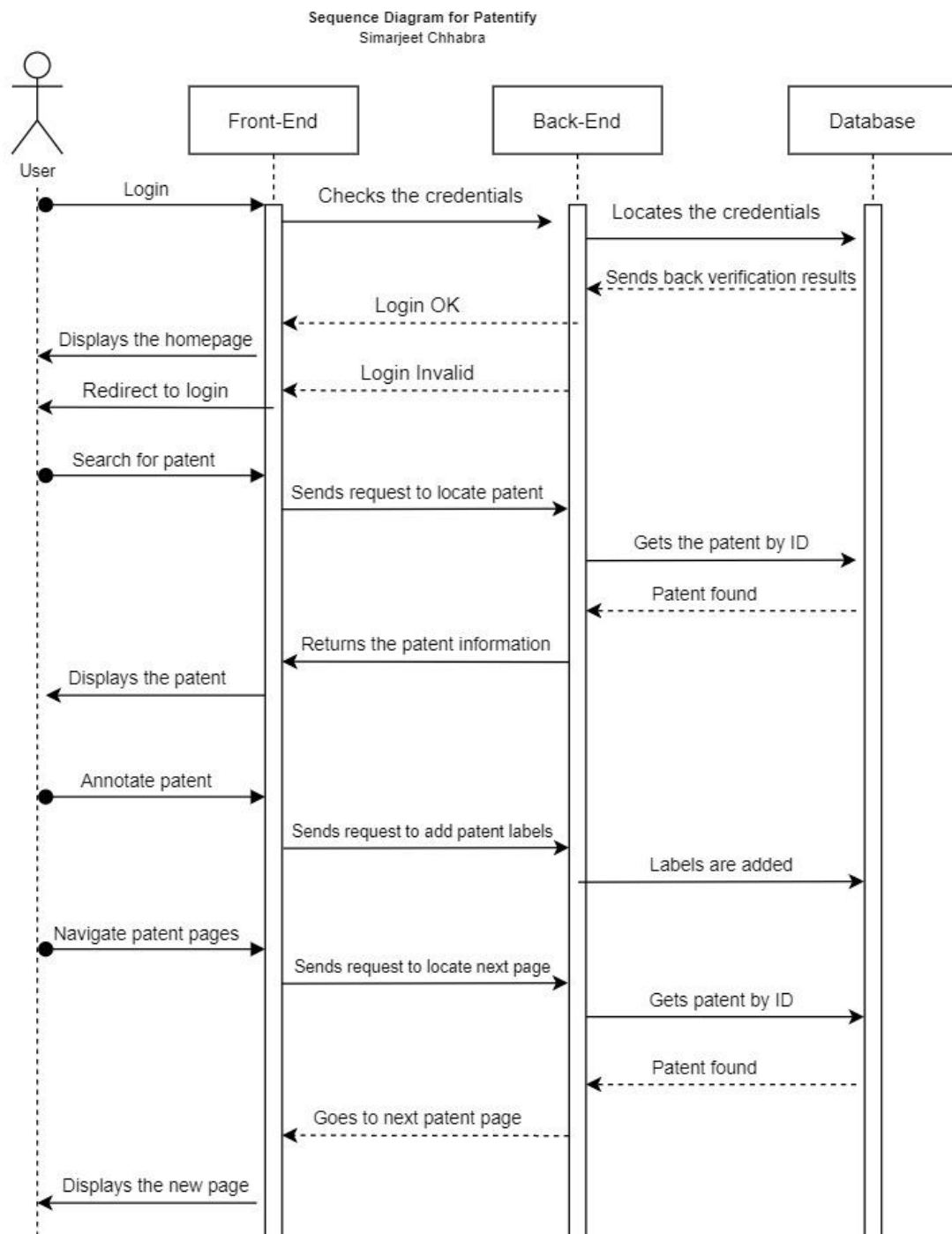
Eric Tu: Use Case Diagram

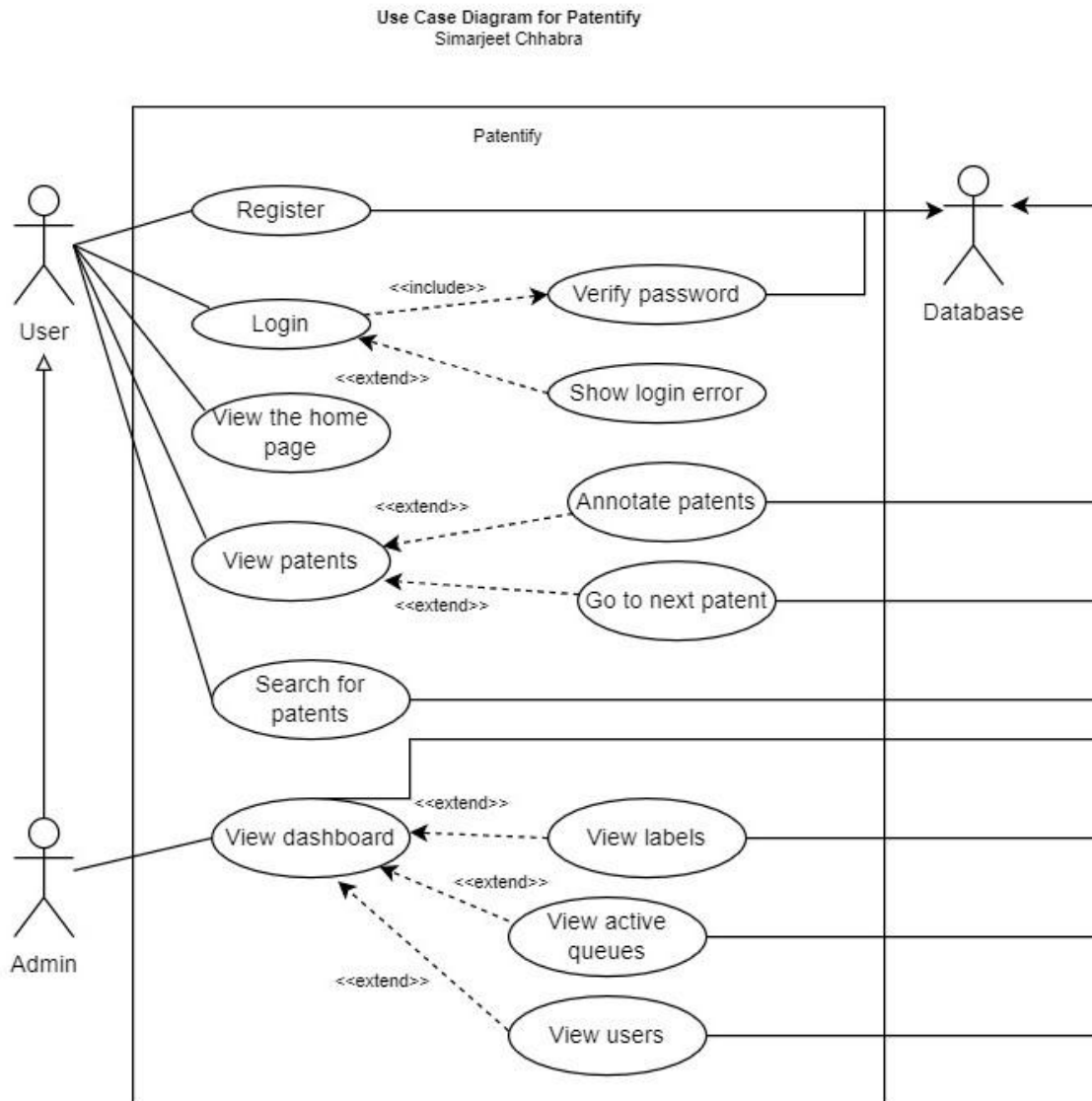
Patentify Use Case Diagram
Eric Tu | November 26, 2021



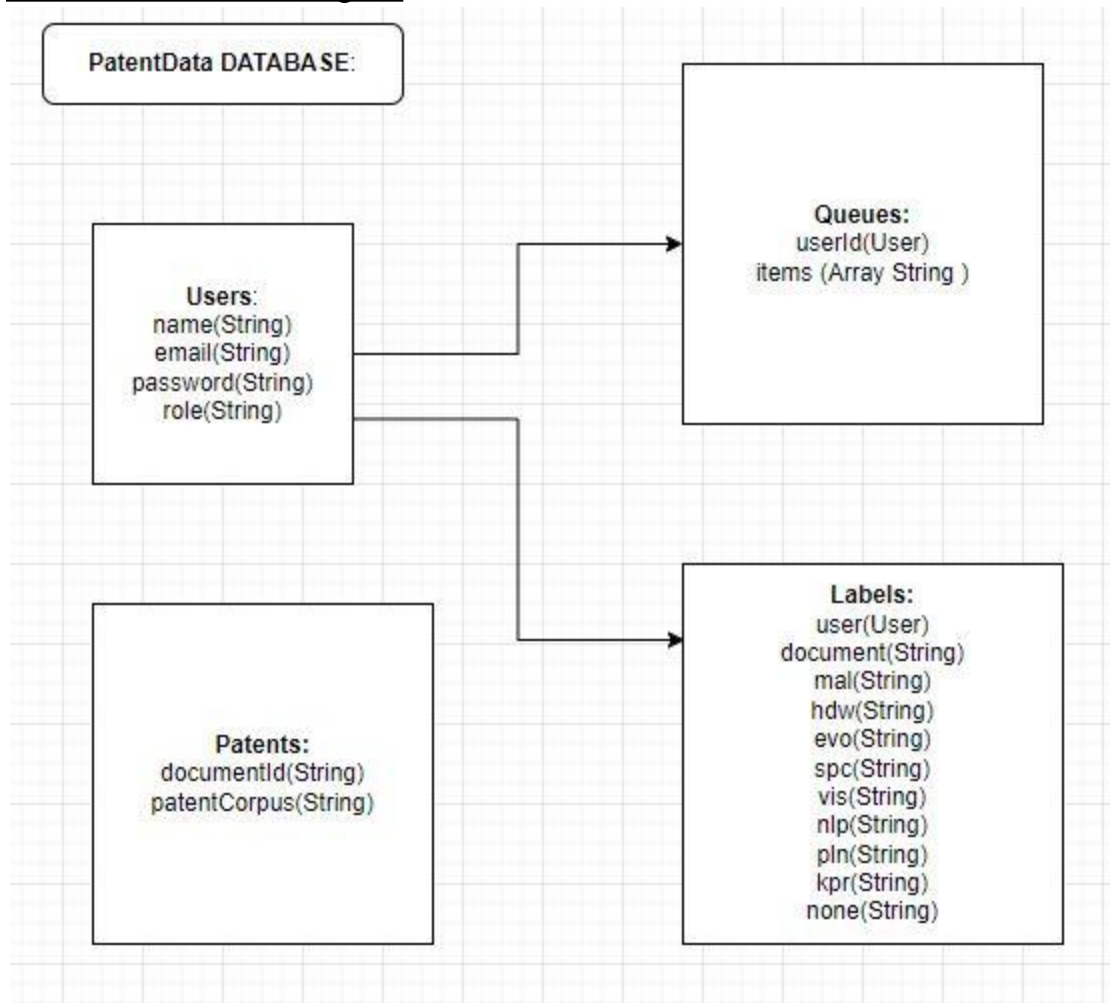
Simarjeet Chhabra

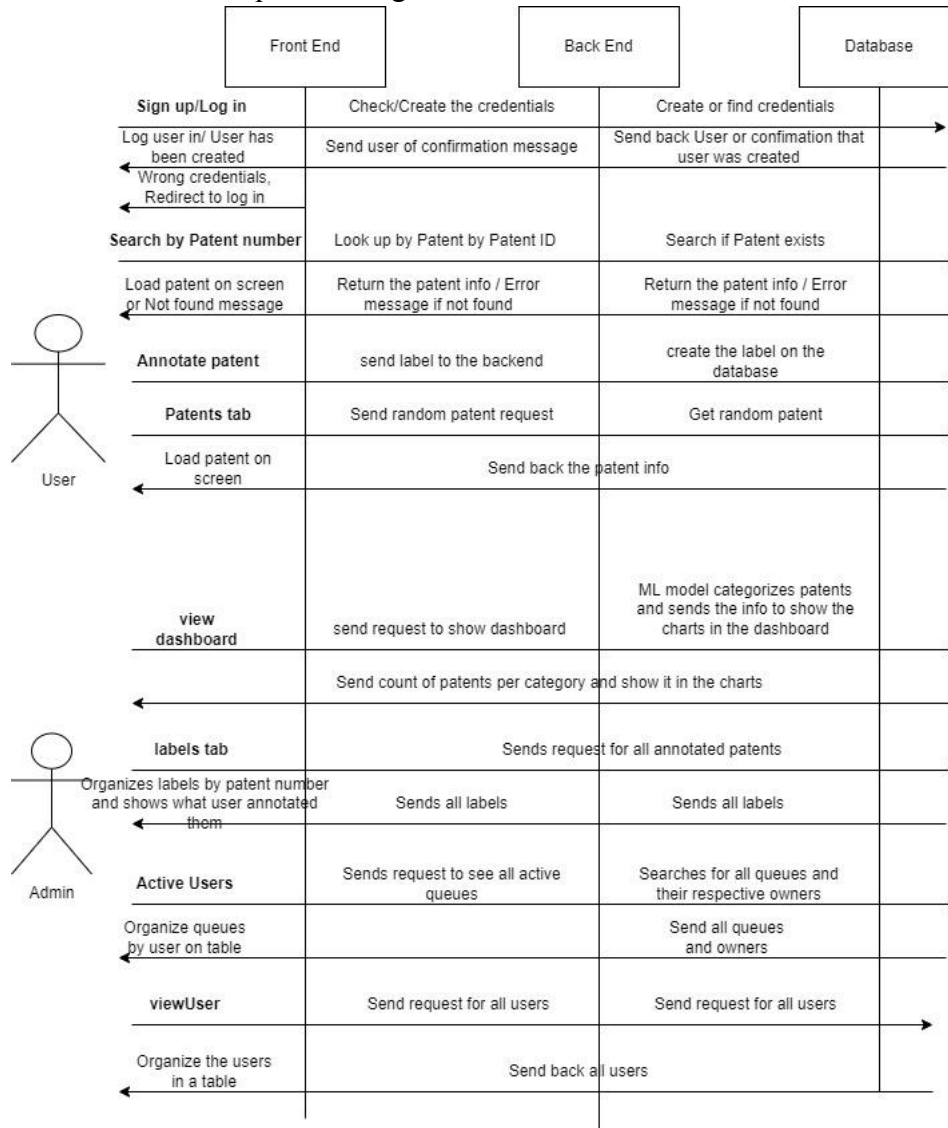
Simarjeet Chhabra: Class DiagramSimarjeet Chhabra: Sequence Diagram

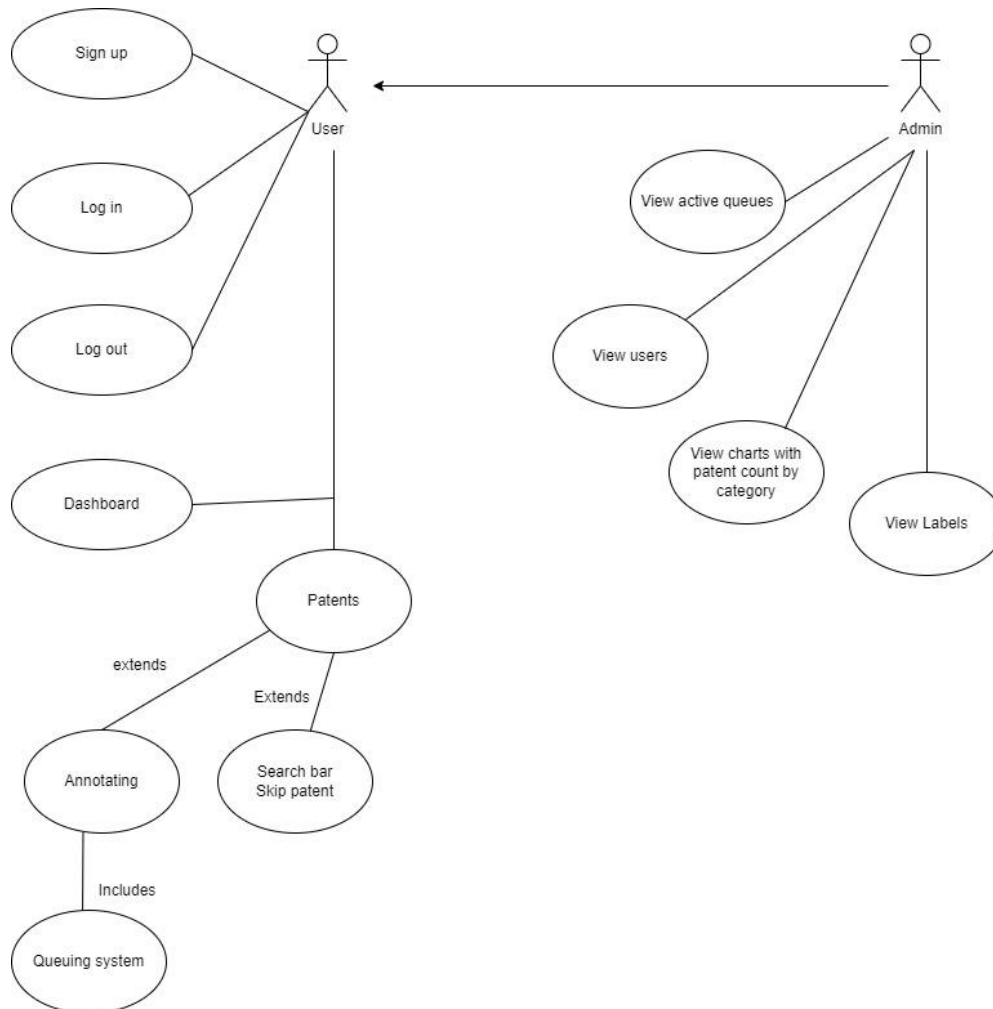


Simarjeet Chhabra: Use Case Diagram

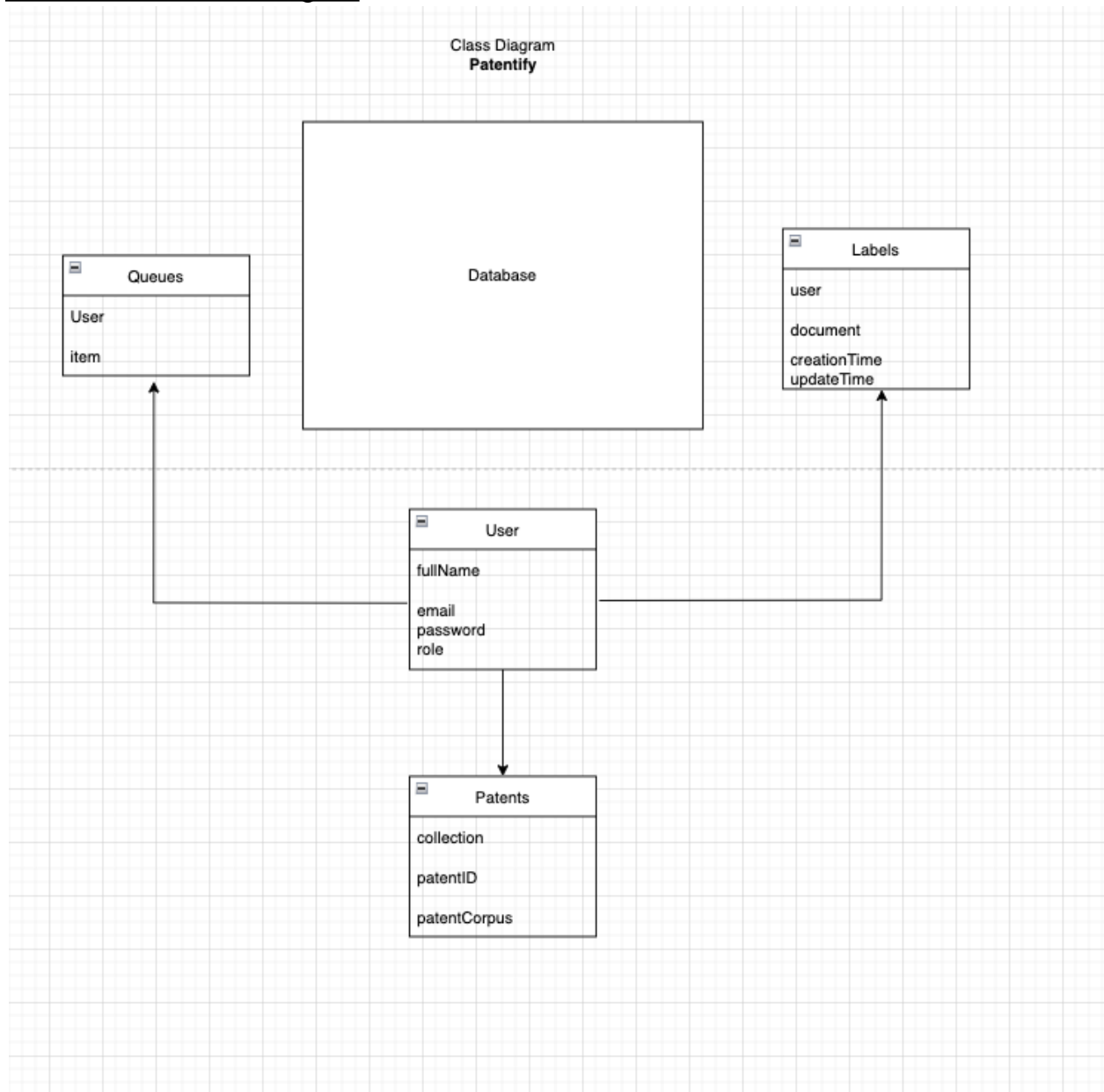
Manuel Garcia

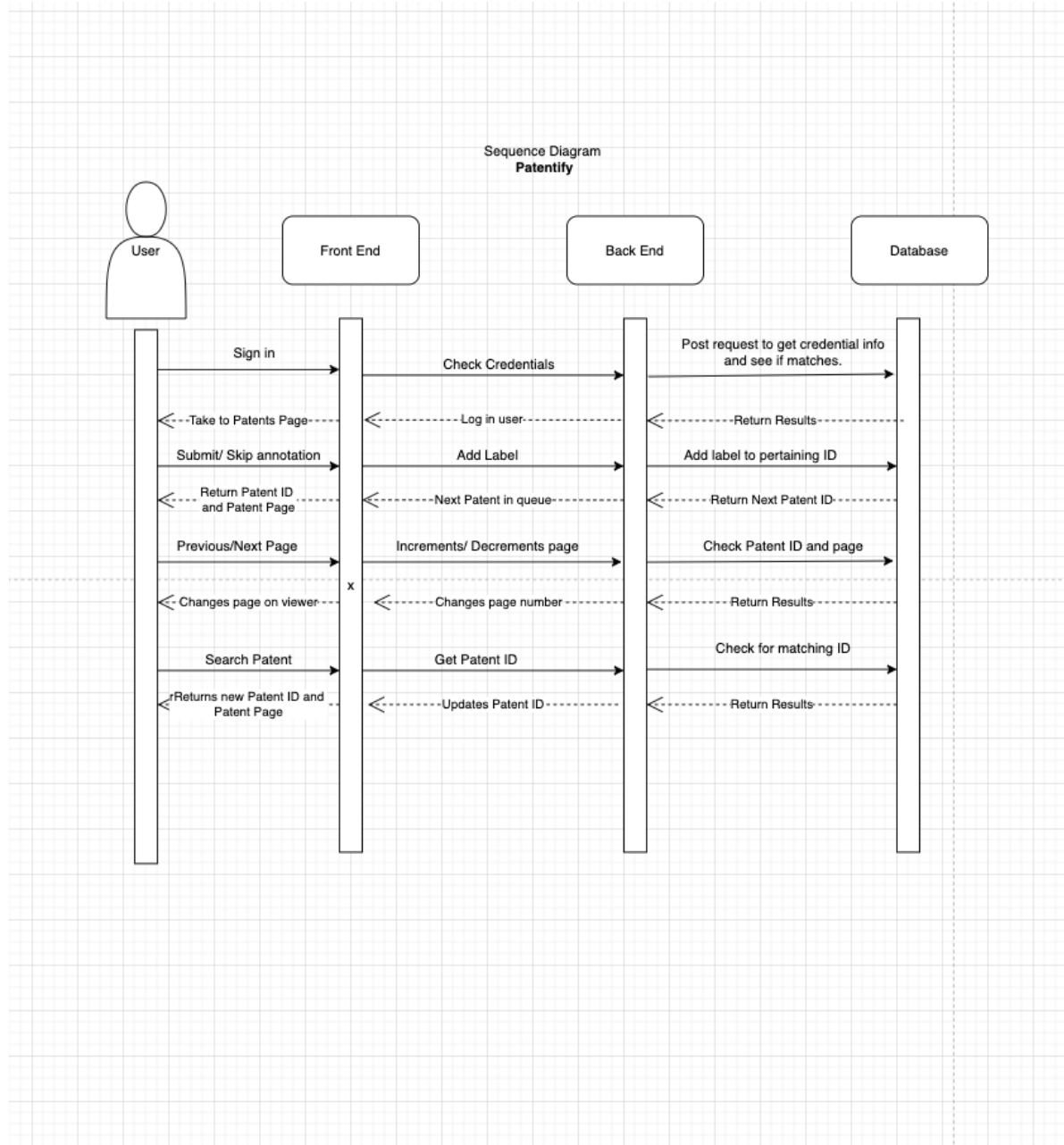
Manuel Garcia: Class Diagram

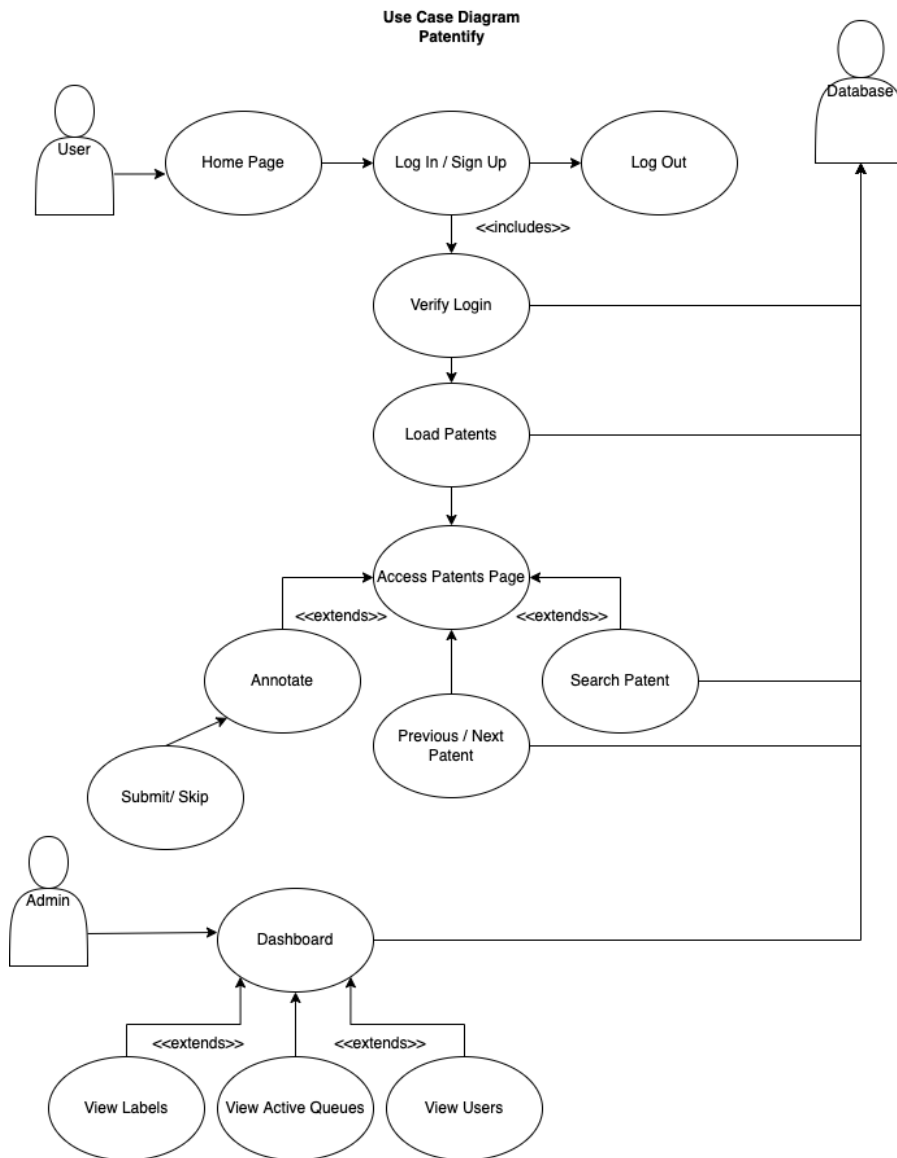
Manuel Garcia: Sequence Diagram

Manuel Garcia: Use Case Diagram

Carlos Cordero

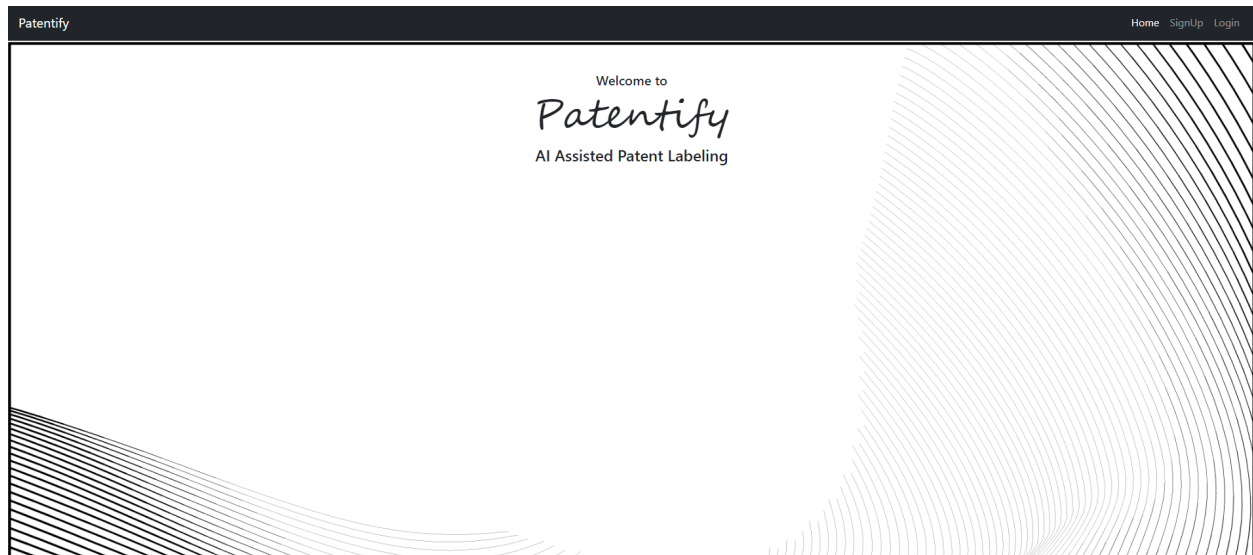
Carlos Cordero: Class Diagram

Carlos Cordero: Sequence Diagram

Carlos Cordero: Use Case Diagram

Appendix B - User Interface Design

This is the landing page and home page users will see first with a navigation bar on top.



This is the sign up page.

The image shows a sign-up page for 'Patentify'. It has the same dark navigation bar as the landing page. Below the navigation bar, the word 'SignUp' is centered. Underneath, there is a white box with a light gray border containing the following elements: the text 'Sign Up to start your session', three input fields labeled 'Enter Name' (with a person icon), 'Enter Email' (with an envelope icon), and 'Create Password' (with a lock icon), a dark 'Sign Up' button, and a link that says 'Have An Account? Login'.

This is the login page.

Patentify

HomeSignUpLogin

Login

Sign in to start your session

Enter Email

Enter Password

Sign In

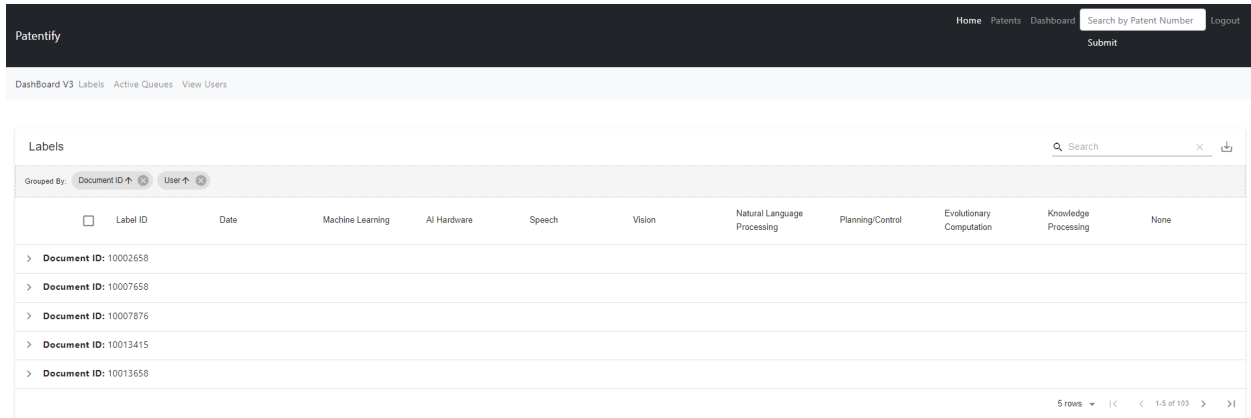
☐ Remember Me

Don't have An Account? [SignUp](#)

This is the patent page where users can annotate labels to patents or skip that current patent. They can search for a patent on the top left by putting a patent ID in.

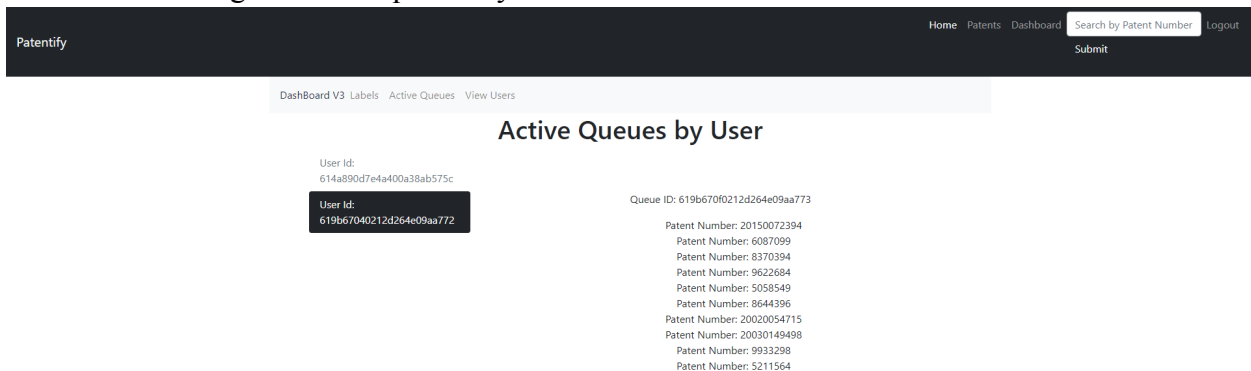
[illegible]

This is the dashboard for administrators. They can see users and each patent.



Label ID	Date	Machine Learning	AI Hardware	Speech	Vision	Natural Language Processing	Planning/Control	Evolutionary Computation	Knowledge Processing	None
> Document ID: 10002658										
> Document ID: 10007658										
> Document ID: 10007876										
> Document ID: 10013415										
> Document ID: 10013658										

This is the viewing for active queues by the administrator for each user.



User ID	Queue ID	Patent Number
614a890d7e4a400a38ab575c		
619b67040212d264e09aa772	619b670f0212d264e09aa773	Patent Number: 20150072394 Patent Number: 6087099 Patent Number: 8370394 Patent Number: 9622684 Patent Number: 5058549 Patent Number: 8644396 Patent Number: 20020054715 Patent Number: 20030149498 Patent Number: 9933298 Patent Number: 5211564

Appendix C - Sprint Review Reports

Attendees:

Product Owners:

- Tisa Islam Erana

Developers:

- Capstone 2
 - Manuel Garcia (MG)
 - Simarjeet Chhabra (SC)
 - Carlos Cordero (CC)
 - Eric Tu (ET)
- Capstone 1
 - Jordy Araujo (JA)
 - Katalina Keyser (KK)
 - Josecarlos Lusbel (JL)
 - Samuel Muvdi (SM)
 - Freddy Somarriba (FS)
 - Luis Gil (LG)

Sprint 1

Date: 9/10/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 1 - Setup older project onto current systems
 - Description - As a developer, I want to be able to download and setup the older project the Product Owner gave us so that we can work on the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

- User Story 2 - Review MongoDB/Express
 - Description - As a developer, I want to learn Express so that I can access the back-end of the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 3 - Review Node/React
 - Description - As a developer, I want to review Node and React so that I can work with the project.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 4 - Review HTML/CSS/JS/Bootstrap
 - Description - As a developer, I want to be able to learn HTML/CSS/JS/Bootstrap so that I can work on the front-end of the project.
 - Assigned to - JA,KK,JL,SM,FS, LG
- User Story 5 - Fix any bugs that may occur during setup
 - Description - As a developer, I want to be able to solve issues regarding setup so that I can work on the project without any errors.
 - Assigned to - MG, SC, CC, ET
- User Story 6 - Review machine learning foundation
 - Description - As a developer, I want to learn more about machine learning so that we can train the AI in this project
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 7 - Learn about Github
 - Description - As a Capstone 1 member I would like to learn how to utilize Github to help create and use branches with the team.
 - Assigned to - JA,KK,JL,SM,FS, LG
- User Story 8 -
 - Description - As a developer, I want to learn about Python to complete this project
 - Assigned to - JA, KK, JL, SM, FS, LG
- User Story 9 -
 - Description - As a developer, I want to understand the documentation, code, and get an idea of how the previous project was made so that I can work with the project efficiently
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 10 -

- Description - As a developer, I want to discuss deployment of the project so that the team has an idea of how to do so next Sprint
- Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

None

Sprint 2

Date: 9/24/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 11 -
 - Description - As a developer, I want to be able to have the database be functional and not have any bugs.
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 12 -
 - Description - As a user, I want access to the project online so that I can work with the application
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 13 -
 - Description - As a user, I want to be able to sign into the website so that I use my account
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 14 -
 - Description - As a user, I want to be able to sign up to the website so that I can have an account with all my searches and patents saved
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 15 -

- Description - As a developer, I want to be able to test the sign-in and login features so that we can see if it works
- Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 16 -
 - Description - As a developer, I want to make a template for a page to search patents so that users can see a display for this feature
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 17 -
 - Description - As a developer, I want to create a GitHub Repo for the project so that we can save changes to the project as we make them
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 18 -
 - Description - As a developer, I want to know more about the database and the AI so that I can work with it as we are applying new features
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 19 -
 - Description - As a user, I want to be able to search a patent by number so that I can search for the patent easier
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 20 -
 - Description - As a developer, I want to implement backend functionality for search patents by number so that users can access the backend for patents
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

None

Sprint 3

Date: 10/8/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 21 -
 - Description - As a user, I want to be able to find the patents I want so that I can annotate on to them
 - Assigned to - CC, ET, JA, SC, MG, SM
- User Story 22
 - Description - As a user, I want to be able to load patents and see them so that I can annotate on those patents
 - Assigned to - CC, ET, JA, SC, MG, SM
- User Story 23
 - Description - As a user, I want to be able to look for a specific patent so that I can annotate on that specific patent I want
 - Assigned to - MG, SC, CC, ET, KK, JL, FS, LG
- User Story 24
 - Description - As a user, I want to have a page where I can use the patent search page so that I can look for a specific patent in the database
 - Assigned to - MG, SC, CC, ET, KK, JL, FS, LG
- User Story 25
 - Description - As a developer, I want to be able to test the patent search so that I know the patents are found in the database and displayed to user
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 26
 - Description - As a developer, I want to continue looking at the machine learning model so that I can understand how that works and how to apply features to it
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 27 -
 - Description - As a user, I want to access the patents quickly so that I do not have to wait too long for the application to load.
 - Assigned to - MG, CC, ET, JA, SC, MG, SM
- User Story 28 -
 - Description - As a user, I want to have access to annotated patents so that I can see what changes have been made

- Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG
- User Story 29 -
 - Description - As a developer, I want to discuss how to enhance the machine learning model so that it can find patents efficiently and effectively
 - Assigned to - MG,SC,CC,ET,JA,KK,JL,SM,FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

None

Sprint 4

Date: 10/22/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 30 -
 - Description - As a user, I want to have keywords for each question so that it is easier to understand what each question entails
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 31 -
 - Description - As a user, I want to be able to have a none of the above option so that I can mark that patent as none of the above options
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 32 -
 - Description - As a user, I want to have the buttons closer together so that I can access those buttons easier
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 33 -
 - Description - As a user, I want to be able to read the whole patent so that I can get a better grasp of what the patent is before answering the questions

- Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 34 -
 - Description - As a user, I want notes next to each question so that I can understand what entails for that specific patent category
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 35 -
 - Description - As a user, I want to be able to use the app on Firefox so that the application can be used on all browsers
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 39 -
 - Description - As a user, I want to be able to see the patent clearly so that I can determine the labels for it
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 40 -
 - Description - As a developer, I need to fix bugs to patent search so that it does not effect the other parts of the application
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

- User Story 36 -
 - Description - As a user, I want the option to save patents for later so that I can work on that specific patent for later
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 37 -
 - Description - As a user, I want to work on other patents that are in other people's queues so that I can help others with queued patents that need to be worked on immediately
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 38 -
 - Description - As a user, I want to be able to see the queue so that I can see what patents take priority and what do not
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG

Sprint 5

Date: 11/5/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 41
 - Description - As a developer, I want to be able to setup queues into the backend so that I can test the frontend users access to them
 - Assigned to - MG, SC, CC, ET, JA
- User Story 42
 - Description - As a user, I want a page to see queues so that I can access the queues easily
 - Assigned to - KK, JL, SM, FS, LG
- User Story 43
 - Description - As a user, I want to be able to hover on top of an icon so that I can see the information for each label
 - Assigned to - SC, CC, ET
- User Story 44
 - Description - As a user, I want to see how many patents are in queue so that I can gauge how much work I have
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 45
 - Description - As a user, I want to have easier access to viewing the whole patents so that I do not have to go through each page one by one
 - Assigned to - KK, JL, SM, FS, LG
- User Story 46
 - Description - As a user, I want a list of queues so that I can know how much work I have left and what patents they are
 - Assigned to - MG, SC, CC, ET, JA
- User Story 47

- Description - As a user, if I search for a patent, it will have its previous annotations so that I can fix them or see if they are correct.
- Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 48
 - Description - As a user, I do not want to get the same patents repeatedly so that I can work on other patents once they have been skipped/submitted
 - Assigned to - MG, SC, CC, ET, JA, KK, JL
- User Story 49
 - Description - As a developer, I want the database to keep track of all the patents a user has in their queue so that it is not randomly selected
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 50
 - Description - As a user, I want to be able to keep the same patent after logging out so that I can work on it later when I am available
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 51
 - Description - As a user, I do not want to see a patent that has already been annotated so that I can focus on other patents that haven't been worked on
 - Assigned to - MG, SC, CC, ET, JA, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

None

Sprint 6

Date: 11/19/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 52 -
 - Description - As a user, I want the skip button under the submit button so that there is more space for the patent view
 - Assigned to - CC, KK, JL, SM, FS, LG
- User Story 53 -
 - Description - As a developer, I want to test scenarios to see if the queue is working such as when two people access and annotate a patent at the same time so that these situations may not cause problems
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 54 -
 - Description - As a developer, I want to give a list of labeled patents to the PO so that the PO may have a list of everything that has been edited and changed
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 57 -
 - Description - As a user admin, I want a page to see the active user queues so that I can access those lists easier
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 58 -
 - Description - As a user, I want to only see my own annotations so that I do not have any biases when annotating on my own
 - Assigned to - MG, SC, CC, ET, JA
- User Story 59 -
 - Description - As a user admin, I want to be able to see who annotated what patents and what they annotated so that I can get information on each user's annotations
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 60 -
 - Description - As a user, I want an even larger patent view screen so that I can view the patent easier
 - Assigned to - ET, KK, JL, SM, FS, LG
- User Story 61 -
 - Description - As a developer, I want to edit the hint information so that users can get a clearer idea of what each label means
 - Assigned to - CC, SC, ET, KK, JL, FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

- User Story 55 -
 - Description - As a user admin, I want to be able to send one patent list via queue to users of my choice so that they can annotate the same patents and have multiple annotations on those patents
 - Assigned to - MG, SC, CC, ET, JA
- User Story 56 -
 - Description - As a user admin, I want to see all active user queues so that I can get an idea of what patents are being worked on
 - Assigned to - MG, SC, CC, ET, JA

Sprint 7

Date: 11/28/2021

Start Time: 8:00 PM

End Time: 8:30 PM

After a show and tell presentation, the implementation of the following user stories were accepted by the product owners:

- User Story 62 -
 - Description - As a developer, I want to include information and comments about my code so that the future groups will be able to understand the code
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 67 -
 - Description - As a user, I want an easier to use page for viewing active user queues so that I can see clearly who is working
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 69 -

- Description - As a developer, I want to make a backup of the program so that future groups can work on it without changing anything important
- Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG

The following ones were rejected and moved back to the product backlog to be assigned to a future sprint at a future Sprint Planning meeting.

- User Story 63 -
 - Description - As a user, I want to know how many pages that the document has so that I can go through the document via next/previous page
 - Assigned to - MG, SC, CC, ET, JA, KK, JL, SM, FS, LG
- User Story 64 -
 - Description - As a user admin, I want to see new entries for each user annotation on a patent so that I can see what the users annotated for that patent
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 66 -
 - Description - As a developer, I want to test to see if sending queues as an admin works so that I can fix any bugs that occur while working on the functionality
 - Assigned to - MG, SC, CC, ET, JA
- User Story 65 -
 - Description - As a developer, I want to fix the bug where User 2's annotations on Patent 1 erases User 1's annotations on Patent 1 so that annotations on a patent can be added altogether to Patent 1
 - Assigned to - MG, SC, CC, ET, JA, SM
- User Story 68 -
 - Description - As a developer, I want to send patents in a logged out user's queue back to the backend list of non-annotated patents so that it does not get stuck in just one user's queue
 - Assigned to - MG, SC, CC, ET, JA, SM

Appendix D - User Manuals, Installation/Maintenance Document, Shortcomings/Wishlist Document and other documents

For the annotation of patents the USPTO office provides a set of guidelines that should be used to annotate patents.

Specific examples for each AI Category are provided below is a link to a document containing the guidelines for annotators to use when classifying patents.

https://drive.google.com/file/d/1wWxiVDheXrt_e5Y-8ErBuHrsyjXClUHJ/view?usp=sharing

Installation/Maintenance Document

Prerequisite Technologies

- Latest versions of Node, Python, MongoDB and Git
- Make a clone of the following repository
 - <https://github.com/64bit-awesome/Patentify>
- Using your terminal CD into the directory where you have cloned the repository.
 - CD into the frontend and run the following command **git pull**, this will install any missing dependencies, then run the following command **npm start**, localhost should open in your browser.
 - Now open a new terminal window and CD in the backend folder and run the following command **git pull**, this will install any missing dependencies, then run the following command **nodemon** to start the backend.
- After running those commands the application will be running.

- In order to start the active learning program CD into the backend folder then CD into the SVM folder and run the following command **python main.py**
- Please read our shortcomings because the part of the application should be integrated into the backend start script.

Exporting Annotations into JSON Format

1. Connect to Patentify VM via TigerVNC
2. Open a Terminal
3. Type: cd Desktop
4. type: mongoexport --db=PatentData --collection=labels --out=annotated-patents.json

if you only want to see the file STOP HERE:

there should be a file on the Desktop called 'annotated-patents.json'.

double-click it and it will open in Sublime Text.

done.

if you want to copy the file onto your computer:

5. type: mv annotated-patents.json /var/www/patentify.com/html/

the file should now be accessible here:

<https://patentify.cognac.cs.fiu.edu/annotated-patents.json>

open the link on a web browser:

via Google Chrome once the page loads, you have to:

1. right click

2. click 'Save as...'

3. give the file a name (make sure it ends in .json) and save it.

via Firefox once the page loads,

there will be an error because it will try to analyze the file,

something like: "SyntaxError: JSON.parse: unexpected character at line 1 column 1 of the JSON data"

do not panic, this is supposed to happen because MongoDB adds an extra line at the end:

1. right click

2. click 'Save Page As...'

3. give the file a name (make sure it ends in .json) and save it.

6. IMPORTANT: go back to the Patentify VM on TigerVNC

in the Terminal, type: `rm /var/www/patentify.com/html/annotated-patents.json`

otherwise, the annotations will be publically accessible for download to anyone.

test that the file is gone by visiting: <https://patentify.cognac.cs.fiu.edu/annotated-patents.json> it should redirect you to the Home page.

Full Deployment/Configuration Process for Patentify:

[warning: this assumes some knowledge of linux/shell usage, always think before you type]

Frontend:

0. Pre-Requisites:

```
/var/www/patentify.com/html/.htaccess
```

if it does not exist or is incorrectly configured, this is what it should look like (minus the dash lines ---):

```
-----  
-----  
Options -MultiViews
```

```
RewriteEngine On
```

```
RewriteCond %{REQUEST_FILENAME} !-f
```

```
RewriteRule ^ index.html [QSA,L]
```

```
-----  
-----  
this correctly sets up React Routing with Apache.
```

```
/etc/apache2/sites-available/patentify.com.conf
```

if it does not exist or is incorrectly configured, this is what it should look like (minus the dash lines ---):

```
-----  
-----  
<VirtualHost *:80>
```

```
ServerAdmin admin@patentify.com
```

```
ServerName patentify.cognac.cs.fiu.edu
```

```
ServerAlias patentify.cognac.cis.fiu.edu
```

```
ServerAlias patentify.com
```

```
DocumentRoot /var/www/patentify.com/html
```

```
ErrorLog ${APACHE_LOG_DIR}/error.log
```

```
CustomLog ${APACHE_LOG_DIR}/access.log combined
```

ProxyRequests on

ProxyPass /users/ http://localhost:4000/users/

ProxyPass /patents-api/ http://localhost:4000/patents-api/

ProxyPass /patents-api/status http://localhost:4000/patents-api/status

ProxyPass /patents-api/logout http://localhost:4000/patents-api/logout

ProxyPass /patents-api/search http://localhost:4000/patents-api/search

ProxyPass /patents-api/getAllQueues http://localhost:4000/patents-api/getAllQueues

ProxyPass /patents-api/queue/remove http://localhost:4000/patents-api/queue/remove

ProxyPassReverse /users/ http://localhost:4000/users/

ProxyPassReverse /patents-api/ http://localhost:4000/patents-api/

ProxyPassReverse /patents-api/status http://localhost:4000/patents-api/status

ProxyPassReverse /patents-api/logout http://localhost:4000/patents-api/logout

ProxyPassReverse /patents-api/search http://localhost:4000/patents-api/search

ProxyPassReverse /patents-api/getAllQueues http://localhost:4000/patents-api/getAllQueues

ProxyPassReverse /patents-api/queue/add http://localhost:4000/patents-api/queue/remove

</VirtualHost>

<VirtualHost *:443>

ServerAdmin admin@patentify.com

ServerName patentify.cognac.cs.fiu.edu

ServerAlias patentify.cognac.cis.fiu.edu

ServerAlias patentify.com

```
DocumentRoot /var/www/patentify.com/html
ErrorLog ${APACHE_LOG_DIR}/error.log
CustomLog ${APACHE_LOG_DIR}/access.log combined
```

ProxyRequests on

```
ProxyPass /users/ http://localhost:4000/users/
ProxyPass /patents-api http://localhost:4000/patents-api/
ProxyPass /patents-api/status http://localhost:4000/patents-api/status
ProxyPass /patents-api/logout http://localhost:4000/patents-api/logout
ProxyPass /patents-api/search http://localhost:4000/patents-api/search
ProxyPass /patents-api/queue/remove http://localhost:4000/patents-api/queue/remove
ProxyPass /patents-api/getAllQueues http://localhost:4000/patents-api/getAllQueues
```

```
ProxyPassReverse /users/ http://localhost:4000/users/
ProxyPassReverse /patents-api/ http://localhost:4000/patents-api/
ProxyPassReverse /patents-api/status http://localhost:4000/patents-api/status
ProxyPassReverse /patents-api/logout http://localhost:4000/patents-api/logout
ProxyPassReverse /patents-api/search http://localhost:4000/patents-api/search
ProxyPassReverse /patents-api/queue/remove http://localhost:4000/patents-api/queue/remove
ProxyPassReverse /patents-api/getAllQueues http://localhost:4000/patents-api/getAllQueues
</VirtualHost>
```


this correctly sets up the proxy between the frontend and the backend.

- for every new backend API route,

you must update this file to ProxyPass and ProxyPassReverse for both the http and https version of the site.

the first URL is that of the frontend, the second is that of the backend.

notice the backend is running on port 4000, you would need to update this, if it is ran on another port.

this is the best solution I could come up with for this, otherwise we would need to hardcode the links into the code.

1. Copy files over to VM:

If you have a GitHub repository set-up,

you can use "git clone" the first time and then "git pull" to get the latest changes.

If you have just "git cloned" the repository from GitHub,

it may be necessary to cd into frontend and "npm install".

this will install all the frontend package dependencies.

2. Build the app:

cd into frontend

build the app with "npm run build"

3. Push files to live Apache directory:

sudo systemctl stop apache2 [take the site offline]

cp -rv build/* /var/www/patentify.com/html/

sudo systemctl start apache2 [put the website back online]

if not already enabled,

```
sudo systemctl enable apache2
```

to have the service automatically start at reboot.

Backend:

0. Pre-Requisites:

```
/lib/systemd/system/backend.service
```

if it does not exist, create it and add this (minus the dash lines ---):

```
[Unit]
```

```
Description=runs USPTO backend
```

```
[Service]
```

```
Environment='PORT=4000'
```

```
Environment='MONGO_URL=mongodb://localhost:27017/PatentData'
```

```
WorkingDirectory=/home/cognac/Documents/Patentify/USPTOApp/backend
```

```
ExecStart= npm start
```

```
Restart=on-failure
```

```
[Install]
```

```
WantedBy=multi-user.target
```

note the WorkingDirectory, this is where the backend code needs to be copied to.

note the PORT number, this is the port the backend lives on in localhost.

note this is where we set up the MONGO_URL for the backend to connect to the right database.

security warning:

the backend is also accesible publically via:
<http://patentify.cognac.cs.fiu.edu:4000/>

1. Copy files over to VM:

If you have a GitHub repository set-up,

you can use "git clone" the first time and then "git pull" to get the latest changes.

If you have just "git cloned" the repository from GitHub,

it may be necessary to cd into backend and "npm install".

2. Copy to WorkingDirectory:

```
sudo systemctl stop backend [stop the backend service]
```

```
cd .. into the main GitHub project directory.
```

```
cp -rv backend ~/Documents/Patentify/USPTOApp/
```

```
sudo systemctl start backend [start the backend service]
```

if not already enabled,

```
sudo systemctl enable backend
```

to have the service automatically start at reboot.

MongoDB:

0. Pre-Requisites:

```
/etc/mongod.conf
```

if it does not exist or is incorrectly configured, this is what it should look like (minus the dash lines ---):

```
-----
```

```
# mongod.conf
```

```
# for documentation of all options, see:
# http://docs.mongodb.org/manual/reference/configuration-options/
# Where and how to store data.
storage:
  dbPath: /var/lib/mongodb
  journal:
    enabled: true
# engine:
# mmapv1:
# wiredTiger:
# where to write logging data.
systemLog:
  destination: file
  logAppend: true
  path: /var/log/mongodb/mongod.log
# network interfaces
net:
  port: 27017
  bindIp: 127.0.0.1
# how the process runs
processManagement:
  timeZoneInfo: /usr/share/zoneinfo
#security:
```

```
#authorization: 'enabled'
```

```
#operationProfiling:
```

```
replication:
```

```
  replSetName: rs1
```

```
#sharding:
```

```
## Enterprise-Only Options:
```

```
#auditLog:
```

```
#snmp:
```

```
-----
```

```
  security warning:
```

```
    it looks like there is already a firewall setup to prevent accessing the database
from the public internet.
```

```
    the MongoDB database should only be set up to be accessible from localhost.
```

```
    it should NOT be set to bind to patentify's public ip.
```

1. Mongod Service:

```
  sudo systemctl start mongod
```

```
  if it is not already enabled,
```

```
    sudo systemctl enable mongod
```

```
    so that it will start automatically at boot.
```

Python Active Learning (SVM):

0. Implementation Limitations:

The current system has not been integrated with the backend yet.

ideally, the python script should be called from the backend service.

1. Copy files over to VM:

If you have a GitHub repository set-up,

you can use "git clone" the first time and then "git pull" to get the latest changes.

2. Training the model:

cd into backend/svm

run it with "python main.py"

Shortcomings/Wishlist

1. The administrator can send a list of random patents as a queue to certain users. This way, if there are patents that need to be prioritized or be worked on, then the admin can provide that to the users. At the moment, there is no such functionality besides the fact that the user can view those active queues of each user. With that in mind, the admin can make changes to queues on that same page.

2. At the moment, users will annotate a patent and replace each other's annotations on the site, but not in the database. In a scenario where User 1 annotated Patent 1 as AI Hardware. If User 2 were to annotate Patent 1 by searching for it and mark it as Machine Learning, then it would replace User 1's annotation on the dashboard for the administrator. However, both User 1's and User 2's annotations will have marked and labeled the patent with those annotations. So, Patent 1 would be marked as both Machine Learning and AI Hardware. As a result, fixing this along with the admin view of seeing each user annotation for each patent is rather important.

3. Patents get stuck in queues for a long time. So, when a user logs out, the patents in the queue for that user will stay there until they are worked on. It would be better to fix this issue by taking the patents in that user queue and place it back in the list of non-annotated patents. This way, those patents can make their way to another user without being static in a queue. A solution

could be a time-based algorithm that will take patents from a queue based on a certain amount of time.

4. Although there is a next and previous page, setting up the pages for each patent to let the user know how many pages a patent has would be good. At the moment, there is a next and previous page that allows the user to go through the patent. However, the interface only shows one page which can lead a user to be confused and think that the patent only has one page. Even if the next and previous page icons are colored to let the user know that there is a way to go through the patent, it might not be enough. So, making a way for the user to see that there is a total amount of pages for that patent would be good. Another way would be to let the user have an option to view the whole patent as well.

REFERENCES

Pustejovsky, J. and Stubbs, A., 2012. *Natural Language Annotation for Machine Learning: A guide to corpus-building for applications*. O'Reilly Media, Inc.

Russell, S.J. and Norvig, P., 2016. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited

“Getting Started – React.” *React – A JavaScript Library for Building User Interfaces*, <https://reactjs.org/docs/getting-started.html>. Accessed 26 Apr. 2021.

Mark Otto, Jacob Thornton, and Bootstrap contributors. “Introduction · Bootstrap v5.0.” *Bootstrap · The Most Popular HTML, CSS, and JS Library in the World.*, <https://getbootstrap.com/docs/5.0/getting-started/introduction/>. Accessed 26 Apr. 2021.

“Welcome to the MongoDB Documentation — MongoDB Documentation.” *Welcome to the MongoDB Documentation — MongoDB Documentation*, <https://docs.mongodb.com/>. Accessed 26 Apr. 2021.

“Express Routing.” *Express - Node.js Web Application Framework*, <https://expressjs.com/en/guide/routing.html>. Accessed 26 Apr. 2021.

“Pandas Documentation — Pandas 1.2.4 Documentation.” *Pandas - Python Data Analysis Library*, <https://pandas.pydata.org/docs/>. Accessed 26 Apr. 2021.

modAL-python. “GitHub - ModAL-Python/ModAL: A Modular Active Learning Framework for Python.” *GitHub*, <https://github.com/modAL-python/modAL>. Accessed 26 Apr. 2021.

“BigQuery Documentation | Google Cloud.” *Google Cloud*, <https://cloud.google.com/bigquery/docs>. Accessed 26 Apr. 2021.