

IT Project - Syntax Sorcerers	2
Progress Checklist 1	4
System Structure	5
Design Class Diagram	6
Tech Stack	7
Architecture Diagram	8
Webpage Pathways	9
Database Diagram	10
API Documentation	11
Tech Stack Links	13
Project Requirements	14
Use Cases and User Stories	19
Personas	21
Research Project to Product	25
Use Case Diagram	26



IT Project - Syntax Sorcerers

About

How can we ensure that the written word maintains its integrity in an era of rapid digital transformation? Driven by curiosity and the desire to make a meaningful impact, we embarked on this quest.

We are a dynamic team of third-year university students on a mission to revolutionize the way authenticity is upheld in the realm of literature. As computer science students, we recognized the significance of preserving the voice and style of authors in an increasingly AI-focused world.

Our passion for technology and literature fuels the creation of our authorship verification web app. With a blend of cutting-edge algorithms and intuitive design, we empower users to distinguish genuine authorship, unveiling the essence of each writer's unique expression within the realm of education.

Mission and vision

As Syntax Sorcerers, we are dedicated to crafting top-tier software solutions that surpass client expectations and adapt to an ever-changing digital landscape. With a foundation of collaboration, creativity, and continuous improvement, we strive to deliver excellence and champion creativity.

Our mission is to reshape the realm of authorship verification, by merging cutting-edge algorithms with our tight bond (after less than a single semester of knowing each other 😎)

Meet the team

Bunny Phetmunee

Product Owner

Dialogue/Communication with the Client

Front-end Development

Yash Patil

Scrum Master

Tracking and Delegating Tasks Through Agile

Back-end Development

Talym Myler

Developer

Front-End Development and Team Support

Antonio Felicissimo

Developer

Back-End Development and Team Support

Muhammad Usman

Developer

Full Stack Development and Team Support

Important Pages

- [Project Requirements](#)
- [Authorship Front-End Design](#) (Figma Designs)
- [Authorship Front-End Design](#) (Interactive Prototype)
- [Personas](#)
- [Standup notes](#)
- [IT-Project-Syntax-Sorcerers](#)
- [GitHub - The-Syntax-Sorcerers/ava](#)

Progress Checklist 1

Process

- ☒ Adherence to Agile Ceremonies
- ☒ Team Structure
- ☒ Team Internal Communication
- ☒ Team Decision Making
- ☒ Communication with Client
- ☒ Tools Used

Artefacts

- ☒ Requirements
- ☒ Front-end Design: [Figma Designs](#) | [Interactive Prototype](#)
- ☒ Architectural Design
- ☒ Coding: [Flask App](#) | [ML Service API](#)
- ☐ Testing
- ☒ Deploying

System Structure

Git Repository Structure

Repo: [GitHub - The-Syntax-Sorcerers/ava](#)

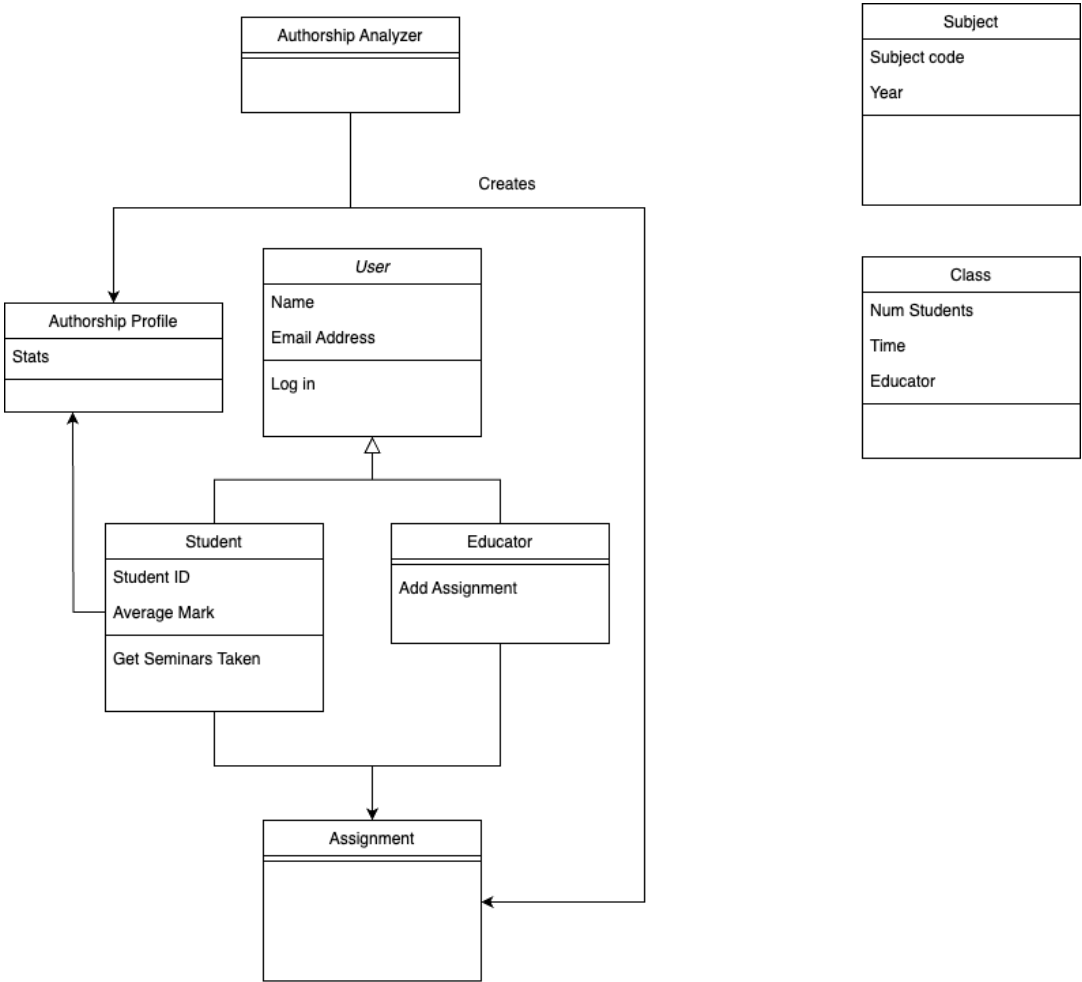
We'll try to be as consistent as possible. Please familiarise yourself with the repository structure below & try to follow this as closely as possible.

Note: `venv` directory is git ignored, so that folder will only exist in your local repo & it should be automatically configured by PyCharm IDE.

```
1  |─ flaskr/
2  |   |─ __init__.py
3  |   |─ db.py
4  |   |─ schema.sql
5  |   |─ auth.py
6  |   |─ blog.py
7  |   |─ templates/
8  |   |   |─ base.html
9  |   |   |─ auth/
10 |   |   |   |─ login.html
11 |   |   |   |─ register.html
12 |   |   |─ blog/
13 |   |   |   |─ create.html
14 |   |   |   |─ index.html
15 |   |   |   |─ update.html
16 |   |─ static/
17 |   |   |─ style.css
18 |   |   |─ images/
19 |─ tests/
20 |   |─ conftest.py
21 |   |─ data.sql
22 |   |─ test_factory.py
23 |   |─ test_db.py
24 |   |─ test_auth.py
25 |   |─ test_blog.py
26 |─ venv/
27 |─ setup.py
28 |─ MANIFEST.in
```

Design Class Diagram

Our Design Class illustrates the conceptual entities and relationships we want to represent in our web app.

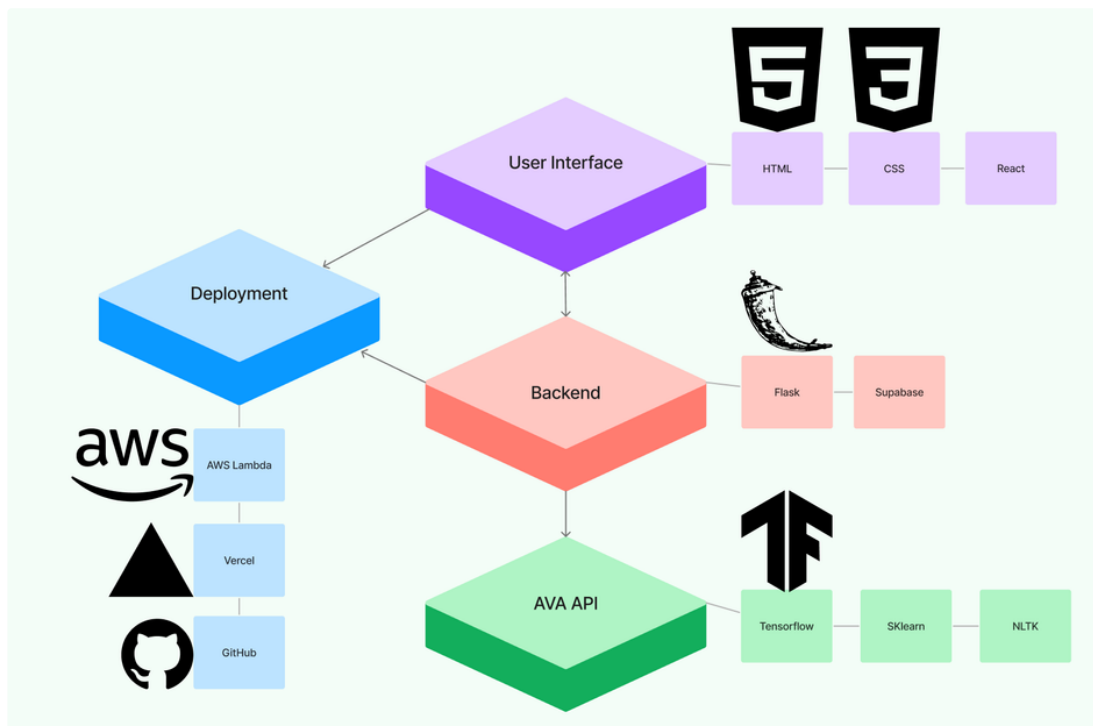


Tech Stack

AVA Tech Stack

Our current tech stack is split into four categories across 3 different layers.

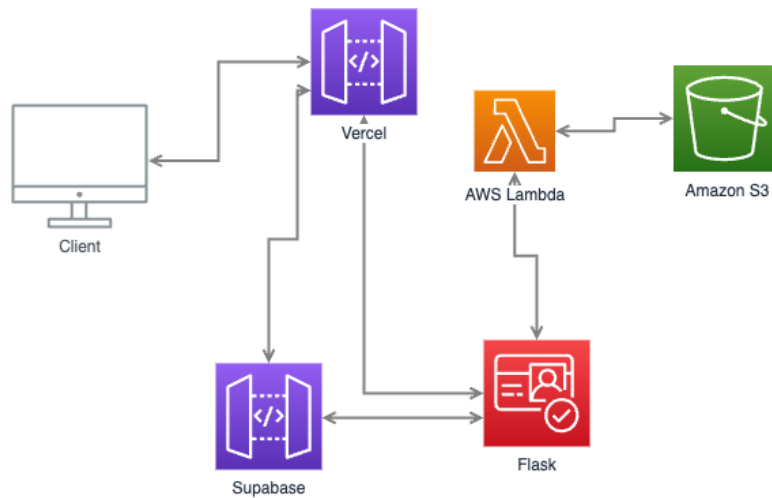
1. The top layer/client-side/user interface/front end will be written primarily using HTML, CSS, and React JS
2. Our second layer is our backend, which is divided into:
 - a. our server-side functionality provided by Flask and Supabase.
 - b. Hosting and deployment provided by Vercel & AWS
3. A separate API component which is the core focus of our product, the authorship verification algorithm, is already provided to us and uses Tensorflow, SKlearn, NLTK, and other Python libraries.



Architecture Diagram

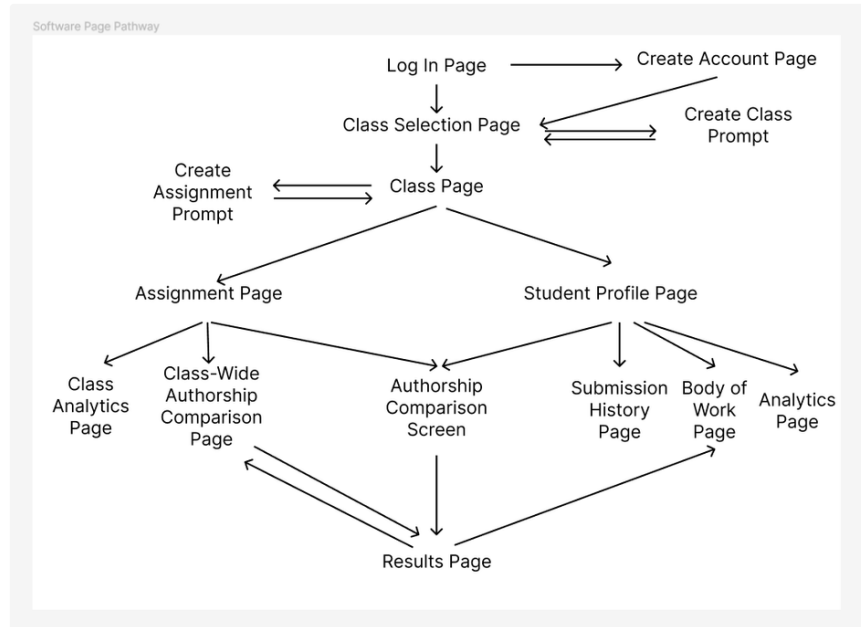
This diagram shows how data flows between the different endpoints in our software architecture.

1. The client will first log onto our app and the data will be sent through Vercel to our Flask App.
2. User data will be stored in Flask and there will be a 3-communication cycle between these three components
3. When the user submits a document, the data will be stored in Supabase and then sent to an AWS Lambda function, which will run the ML Service. The model weights for the ML model will be stored and loaded from S3.

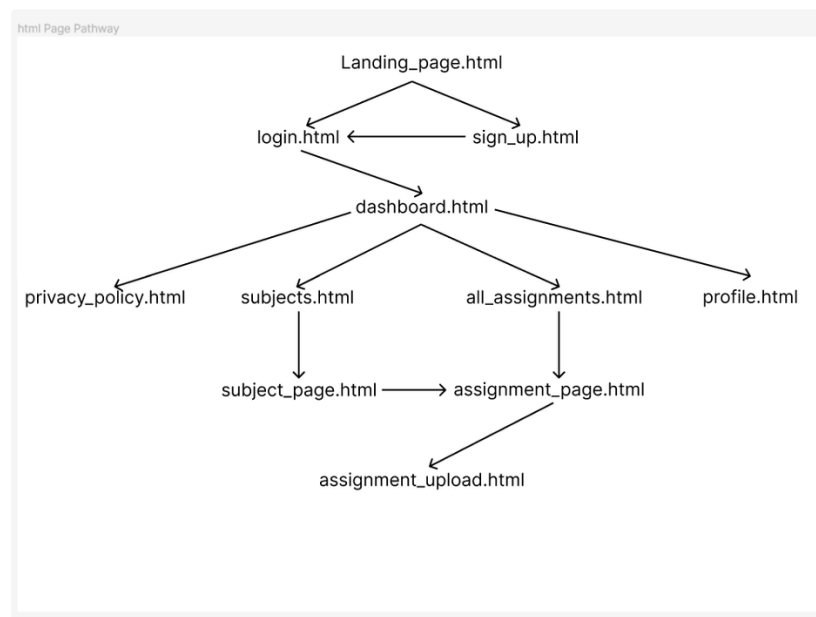


Webpage Pathways

When First designing our application, we had many pages and functionalities we initially wanted to include in our app. This diagram shows a map of how a user would navigate through our web app:



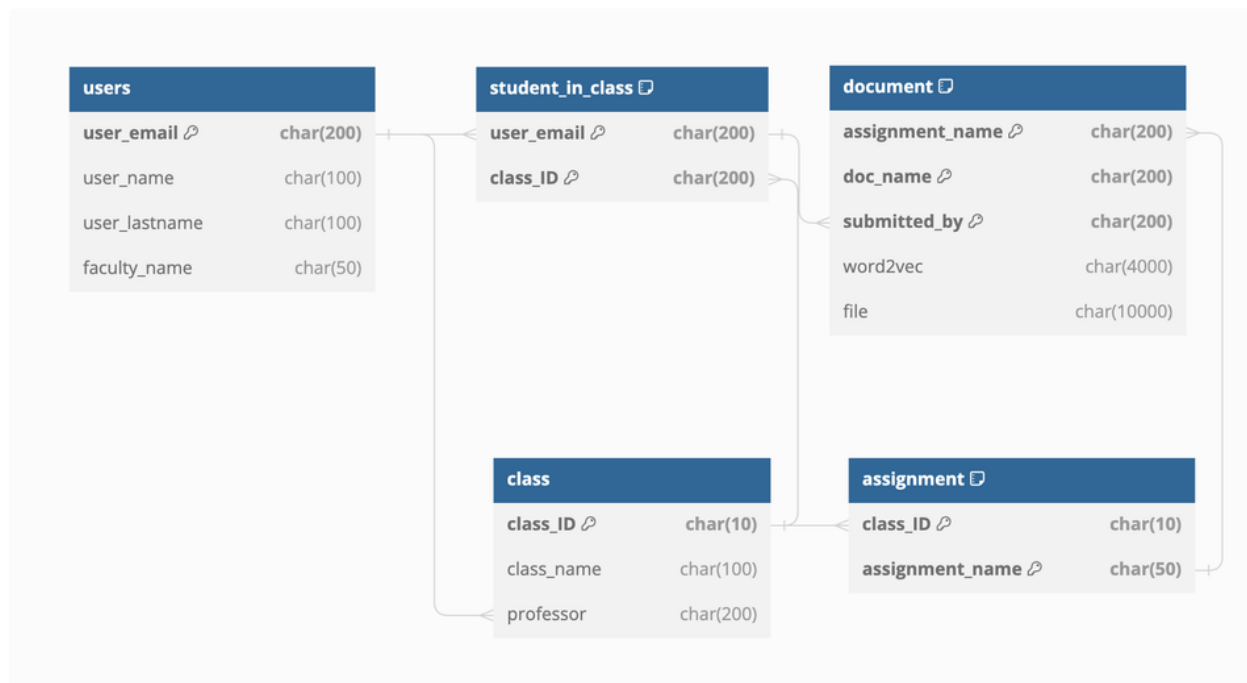
After reflecting on our first few sprints and refining the scope of our project (we were too ambitious in the beginning), these are the pages that we want to focus on to deliver a fully functioning product:



Database Diagram

This database diagram shows how will store data in our database management system (Supabase).

1. The user table will store both teachers and students, the unique identifier for each user will be their email address.
2. The student-in-class table will map each student to a class.
3. Each class will be assigned to one professor/educator to manage
4. The assignment table will contain information and details for each assignment (e.g. due date, description, possible marks awarded)
5. Each document submitted will be linked to a class, assignment, and a student, and we will have a field that stores a file location for the document and its pre-processed vector model



[A Free Database Designer for Developers and Analysts](#)

API Documentation

Overview

This API documentation covers two main modules of a web application: `bp.py` and `auth.py`. The `bp.py` module handles routes related to a user dashboard and assignment uploads, while the `auth.py` module manages authentication routes for login, signup, confirmation, and logout functionalities.

`bp.py`

1. Dashboard

- **Endpoint:** `/dashboard`
- **Method:** `GET`, `POST`
- **Authentication:** Required
- **Description:** Provides access to the user's dashboard.
- **Parameters:** None
- **Returns:** Renders the `dashboard.html` template.

2. Upload

- **Endpoint:** `/upload`
 - **Method:** `GET`, `POST`
 - **Authentication:** Not required
 - **Description:** Allows the user to upload an assignment.
 - **Parameters:** None
 - **Returns:**
 - `GET`: Renders the `assignment_upload.html` template.
 - `POST`: Redirects to the homepage.
-

`auth.py`

1. Login

- **Endpoint:** `/`
- **Method:** `GET`, `POST`
- **Authentication:** Not required
- **Description:** Provides login functionality.
- **Parameters:** None
- **Returns:**
 - `GET`: Renders the `auth/login.html` template.
 - `POST`:
 - Success: Redirects to the user's dashboard.

- Failure: Prints error message.

2. Signup

- **Endpoint:** `/signup`
- **Method:** `GET`, `POST`
- **Authentication:** Not required
- **Description:** Provides signup functionality for new users.
- **Parameters:** None
- **Returns:**
 - `GET`: Renders the `auth/signup.html` template.
 - `POST`:
 - Success: Redirects to the login page with a successful account creation message.
 - Failure: Prints error message.

3. Confirmation

- **Endpoint:** `/complete`
- **Method:** `GET`
- **Authentication:** Not required
- **Description:** Shows an email confirmation page.
- **Parameters:** None
- **Returns:** Renders the `auth/email_confirmation.html` template.

4. Logout

- **Endpoint:** `/logout`
- **Method:** `GET`
- **Authentication:** Required
- **Description:** Provides logout functionality.
- **Parameters:** None
- **Returns:** Redirects to the login page with a successful logout message.

Tech Stack Links

Jira: [IT-Project-Syntax-Sorcerers](#)

Confluence: [IT Project - Syntax Sorcerers](#)

Github: [GitHub - The-Syntax-Sorcerers/ava](#)

Supabase: [Supabase](#)

Vercel: [Dashboard – Vercel](#)

AWS: <https://ap-southeast-2.console.aws.amazon.com/console/home?region=ap-southeast-2#>

Resources:

Github: [GitHub - supabase-community/supabase-py: Python Client for Supabase](#)

PAN datasets: [PAN Data](#)

React: [React](#)

feel free to add more :)

Project Requirements

Objective

Create a visual interface for an Authorship Verification Algorithm.

Success metrics

Goal	Metrics
Students and Educators can create an account and log in	<ul style="list-style-type: none">• Data is stored securely in a DBMS• Passwords encrypted
Educators can enroll students into classes	<ul style="list-style-type: none">• Educators can view all the students currently enrolled into a class• Educators can view all current students and past students enrolled in a subject.
Educators can add assignments to a class and students can upload documents for the assignment	<ul style="list-style-type: none">• Documents stored in a cloud securely• Documents can be accessed and analyzed using the analyzer

Milestones

401Client must be authenticated to access this resource.

Functional Requirements

Requirement	User Story	Importance	Jira Issue
Log in/Sign up	<ul style="list-style-type: none"> As a student, I want to have access control over my account to ensure secure access to my work As a student, I want to be able to reset my password to maintain control of my account if I forget it As a teacher, I want to have access control over my account to ensure secure access to my structures and students' work As a teacher, I want to be able to reset my password to maintain control of my account if I forget it 	HIGH	IPSS-30 IPSS-43 IPSS-36 IPSS-37 IPSS-53
View and Create Classes	<ul style="list-style-type: none"> As a teacher, I want to be able to organise my students into classes for easy access and 'file'-keeping As a teacher, I want to be able to create classes for easy organisation of students and their work As a teacher, I want to be able to assign/dismiss students to classes for easy organisation of students As a student, I want to be able to have assigned classes to make assignment access easy 	HIGH	IPSS-31
View Student Profile	<ul style="list-style-type: none"> As a student, I want to be able to access my own profile to view my submission history and stylometry analytics As a student, I want to be able to access my own submission history from my profile to easily check I submitted the correct work and the given results As a teacher, I want to be able to access individual student profiles to troubleshoot issues they may be having with allocations and submissions As a teacher, I want to be able to review a student's submission history from their profile for classes and assignments to be able to validate results personally As a teacher, I want to be able to view a student's body of work from their profile to review the matching sections in the text 	HIGH	IPSS-32 IPSS-38

View and Create Assignments	<ul style="list-style-type: none"> • As a teacher, I want to be able to create assignment pages to organise student work • As a teacher I want to be able to create assignment pages to allow students to submit their work • As a teacher, I want to be able to edit assignment names/due-dates/descriptions/active statuses to control information for students • As a teacher, I want to be able to view class analytics by assignment to see if the quality of my assignment specs are up to scratch • As a student, I want to have assignment pages with descriptions to easily check requirements • As a student, I want to have assignment pages with upload options to submit my work • As a student, I want to have past assignment pages to review previous submission information and results 	HIGH	IPSS-34
Upload documents to assignment	<ul style="list-style-type: none"> • As a student, I want to have assignment pages with upload options to submit my work • As a teacher I want to be able to create assignment pages to allow students to submit their work 	HIGH	IPSS-35
Run Authorship Verification securely and efficiently	<ul style="list-style-type: none"> • As a student, I want to be able to upload and submit files for assignments to validate my authorship • As a teacher, I want to be able to compare a student's work to ensure authorship • As a teacher, I want to be able to compare an 'in doubt' piece of work across the whole class to check if actual authorship is local (Potential Ethics Issue?) • As a teacher, I want to be able to configure the authorship comparison parameters to have finer control over the results (e.g. adjust the n-gram size, word vector size, and model type) (Potential Ethics Issue?) (Not suited for target audience?) 	HIGH	
View results and analysis	<ul style="list-style-type: none"> • As a student, I want to be able to check my authorship comparison results after 	MED	


	<p>submitting an assignment to know the outcome of my result</p> <ul style="list-style-type: none"> • As a student, I want to be able to visually view highlighted sections of my work to know what helped or hindered my results • As a teacher, I want to be able to view the comparison results to ensure authorship • As a teacher, I want to be able to visually view highlighted sections of student work to know what helped or hindered their results so I can make my own judgments in an 'in doubt' case 		
Body of work storage	<ul style="list-style-type: none"> • As a student, I want my work that has passed inspection to be added to my reference body of work to most accurately reflect my writing style • As a teacher, I want my student's work that has passed inspection to be added to their body of work to most accurately reflect their writing style • As a student, I want my previous reference body of work stored so I don't have to manually add previous documents every time I make a comparison • As a student, I want my previous body of work to be secure and only accessed by parties deemed absolutely necessary to the authentication process • As a teacher, I want student's previous reference body of work stored so I don't have to manually add previous documents every time I make a comparison 	MED	

Non-Functional Requirements

Requirement	User Story	Approach
Security and Privacy	As a user, I want to make sure that none of the information I provide to this service will be misused or shared without consent.	<ul style="list-style-type: none"> • Write a privacy policy and terms of service on the website to let users know exactly how their data will be collected, stored, and used. • When creating an account, ask users to agree to a terms of service before proceeding. • Make sure authentication and encryption services are reliable. • Make sure database storage is reliable

Efficiency	As a teacher, I want to quickly perform authorship verification on a large cohort of students.	<ul style="list-style-type: none"> • Use an online service that can easily scale to serve multiple requests. • Leverage GPU to speed up inference if possible
Adaptability	<p>As a student/teacher, I should be able to rely on the output of the model as my writing style changes/as more data comes in.</p> <p>As a student/teacher, I should be able to trust that the model works for different types of papers/documents.</p>	<ul style="list-style-type: none"> • The model should periodically fine tuned on new data so that it can adapt to new information and patterns. • Make sure that the model has adequate regularization to prevent overfitting • Extensively test and evaluate whether the model is reliable using a variety of datasets.
User Feedback and Monitoring	As a user, I should be able to override or report incorrect predictions made by the model.	<ul style="list-style-type: none"> • Consistently ask for user feedback and allow users to report any clear mistakes made by the model. • Use feedback to further refine the model through Reinforcement Learning from Human Feedback (RLHF)

User interaction and design

 Use Cases and User Stories

 Use Case Diagram

 Personas

Assumptions

- Assumes that teachers are not also students (i.e. tutors who are also enrolled in other subjects cannot be educators?)
- Each assignment is written only by one student

Out of Scope

- Accessibility-supported UI for blind people
- Live Writing Style Analysis as you type
- Directly showing results to students through the website
- Verifying group projects

Use Cases and User Stories

User Stories

Site Access

- As a student, I want to be able to have access control over my account to ensure secure access to my work
- As a student, I want to be able to reset my password to maintain control of my account if I forget it
- As a teacher, I want to be able to have access control over my account to ensure secure access to my structures and students' work
- As a teacher, I want to be able to reset my password to maintain control of my account if I forget it

Organising students by Class

- As a teacher, I want to be able to organise my students into classes for easy access and 'file'-keeping
- As a teacher, I want to be able to create classes for easy organisation of students and their work
- As a teacher, I want to be able to assign/dismiss students to classes for easy organisation of students
- As a student, I want to be able to have assigned classes to make assignment access easy

Organising class work by assignments

- As a teacher, I want to be able to create assignment pages to organise student work
- As a teacher I want to be able to create assignment pages to allow students to submit their work
- As a teacher, I want to be able to edit assignment names/due-dates/descriptions/active statuses to control information for students
- As a teacher, I want to be able to view class analytics by assignment to see if the quality of my assignment specs are up to scratch
- As a student, I want to have assignment pages with descriptions to easily check requirements
- As a student, I want to have assignment pages with upload options to submit my work
- As a student, I want to have past assignment pages to review previous submission information and results

Student profiles

- As a student, I want to be able to access my own profile to view my submission history and stylometry analytics
- As a student, I want to be able to access my own submission history from my profile to easily check I submitted the correct work and the given results
- As a student, I want to be able to view my body of work from my profile to see my own records
- As a student, I want to be able to view my own analytics screen from my profile to gain a better understanding of my own writing style and see my trends over time
- As a teacher, I want to be able to access individual student profiles to troubleshoot issues they may be having with allocations and submissions
- As a teacher, I want to be able to review a student's submission history from their profile for classes and assignments to be able to validate results personally
- As a teacher, I want to be able to view a student's body of work from their profile to review the matching sections in the text
- As a teacher, I want to be able to view a student's analytics screen to check for inconsistencies in the most recently submitted work and the student's past submissions

Submission History

Body of Work

- As a student, I want my work that has passed inspection to be added to my reference body of work to most accurately reflect my writing style
- As a teacher, I want my student's work that has passed inspection to be added to their body of work to most accurately reflect their writing style
- As a student, I want my previous reference body of work stored so I don't have to manually add previous documents every time I make a comparison
- As a student, I want my previous body of work to be secure and only accessed by parties deemed absolutely necessary to the authentication process
- As a teacher, I want student's previous reference body of work stored so I don't have to manually add previous documents every time I make a comparison

Analytics

- As a student, I want to be able to view my stylometry analytics to better understand my writing style
- As a student, I want to be able to see the changes in my stylometry over time to better understand how my writing style is changing

Making authorship comparisons

- As a student, I want to be able to upload and submit files for assignments to validate my authorship
- As a teacher, I want to be able to compare a student's work to ensure authorship
- As a teacher, I want to be able to compare an 'in doubt' piece of work across the whole class to check if actual authorship is local (Potential Ethics Issue?)
- As a teacher, I want to be able to configure the authorship comparison parameters to have finer control over the results (e.g. adjust the n-gram size, word vector size, and model type) (Potential Ethics Issue?) (Not suited for target audience?)

View Results and Analytics

- As a student, I want to be able to check my authorship comparison results after submitting an assignment to know the outcome of my result
- As a student, I want to be able to visually view highlighted sections of my work to know what helped or hindered my results
- As a teacher, I want to be able to view the comparison results to ensure authorship
- As a teacher, I want to be able to visually view highlighted sections of student work to know what helped or hindered their results so I can make my own judgments in an 'in doubt' case

Personas

Educator 1



Persona name	Patricia Welling
User Type	Educator
Role description	University Professor
User ID	2153374

Institution

University name	University of Melbourne
University size	Large university with thousands of students
Subjects taught	History and Philosophy

Demographic information

Age	30
Gender	Female
Education level	Master's

Personal quote

“Authenticity and originality are necessary precursors to gaining true wisdom in the arts.”

Biography

Goals	Motivators
-------	------------

<ul style="list-style-type: none"> • Verify the authenticity of student submissions. • Prevent plagiarism within her classes. • Provide accurate and quick feedback to students. 	<ul style="list-style-type: none"> • Wants to challenge students to think for themselves • Wants to maintain the integrity of the university, so that they aren't allowing unqualified students to get a degree • Wants to build trust between the academic staff and the students
Challenges	Needs
<ul style="list-style-type: none"> • Checking for plagiarism is time consuming • It is difficult to detect small changes in writing style • There is no existing tool that can seamlessly integrate into her current workflow 	<ul style="list-style-type: none"> • Accuracy and precision in plagiarism detection • A secure and private platform to maintain anonymity and confidentiality to reduce subconscious biases in grading. • An efficient way of analyzing papers because she has to review hundreds of students' work.

Student 1



Persona name	Kevin Kwon
User Type	Student
Role description	3rd Year Bachelor of Art's Student at Unimelb
User ID	2917163

Institution

University name	University of Melbourne
University size	Large university with thousands of students
Subjects enrolled	Sociology and Philosophy

Demographic information

Age	21
------------	----

Gender	Male
Education level	High School

Personal quote

"P's get degrees"

Biography

Goals	Motivators
<ul style="list-style-type: none"> Graduate with minimal effort Relax as much as he can in undergrad 	<ul style="list-style-type: none"> His parents want him to go to college, so he's only doing it to make them happy.
Challenges	Needs
<ul style="list-style-type: none"> University work keeps getting harder to keep up with He stopped paying attention in lectures and has a lot of assignments piled up 	<ul style="list-style-type: none"> A system that can detect plagiarism so that students like him can learn the consequences of not putting in effort.

Student 2



Persona name	Linda Chen
User Type	Student
Role description	1st Year Masters Student at Unimelb
User ID	1239572

Institution

University name	University of Melbourne
University size	Large university with thousands of students

Subjects enrolled	Law and Criminal Psychology
--------------------------	-----------------------------

Demographic information

Age	24
Gender	Female
Education level	Bachelor's Degree

Personal quote

"If you can't even manage to write an assignment yourself for University, what makes you think you have what it takes to navigate the real world?"

Biography

Goals	Motivators
<ul style="list-style-type: none"> • Ensure her assignments are plagiarism-free. • Improve her writing skills based on feedback. 	<ul style="list-style-type: none"> • Wants to achieve personal growth • Maintain academic integrity so that she feels more confident working professionally • Fairness in assessment between herself and peers
Challenges	Needs
<ul style="list-style-type: none"> • Concerns about unintentional plagiarism. • Limited access to expensive plagiarism detection tools. 	<ul style="list-style-type: none"> • a tool that can identify inconsistencies in long writing projects. • Get insights into her writing strengths and weaknesses.

Research Project to Product



Steps involved in the research project code:

- Extract text from files - need to consider how we deal with different types of input files (pdf, txt, docx)
- Preprocessing - cleaning text
- Training a W2Vec model
- Text Vectorization - extracting stylistic features
- Load model weights - there are different models to choose from (SVM, SiameseNet)
 - Or we could train our own model from scratch using the data they provide, possibly modify some of the parameters
- Perform inference using the model using test/validation data



Issues:

- confusing to navigate
- too many function calls
- no user interface
- input data only seems to contain txt files – what if we want to upload a pdf?
- there are several different models available



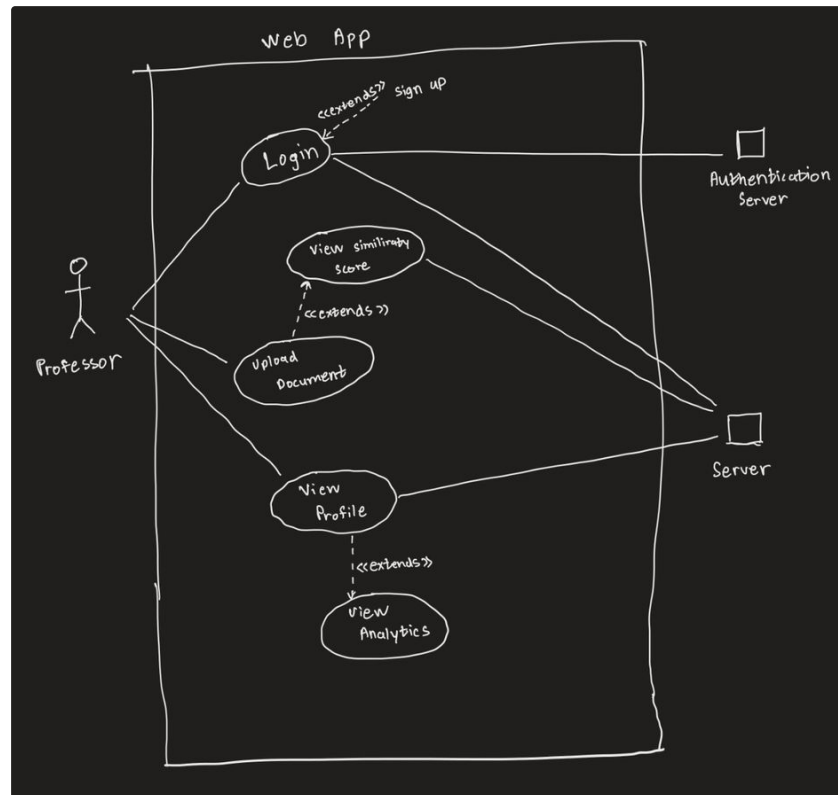
Solutions:

- Create a more streamlined program that can perform all of the steps with one function call (i.e. move functions out of jupyter notebook into a callable Python program)

Use Case Diagram

This is a preliminary design of a use case diagram which we used to inspire our project requirements and user story planning.

Prettier Version:



Original Version:

