

DocuMatcher: User Manual

Authors

Elijah Mullens Jack Blackwood

Fin O'Loughlin Aidan Power

Louis Lai Yiu Xuan Lok

Table of Contents

Installation Guide.....	1
Creating a Secret Key.....	1
Running Locally.....	1
Using The Pre-Built Image.....	1
Building From Source.....	2
Running On AWS Lightsail Containers (Recommended).....	2
Interface Guide.....	4
Upload Page (Home Page).....	4
What Files Can Be Uploaded?.....	5
Managing Duplicates.....	5
Actions Available in the Upload Interface.....	5
Summary Page (Match Page).....	6
Overview.....	6
List View.....	7
Grid View.....	7
Sorting Options.....	7
Search Functionality.....	8
Percentage Filtering.....	8
Export Options.....	9
Reuploading Files.....	9
RSID Comparison Page (Visualise Page).....	10
Docx Summary.....	10
Hover RSID Text.....	10
Border Colour Green/Red.....	11
Context Menu.....	11
Switch Button.....	11
Height.....	12
Change RSID Colour.....	13
Hide Single Colour.....	14
Colour Input Display.....	14
Exporting PDF.....	15
Performance.....	16
Test Case 1.....	16
Test Case 2.....	17

Installation Guide

There are two methods available to run DocuMatcher; locally on your own machine or, hosted on a server. Both methods will be described here, however the cloud option will only be described in the context of AWS Lightsail Containers (though the big cloud providers are all very similar).

Creating a Secret Key

A secret key is a required environment variable for the app to run securely. Python is required to create a secret key; you can install it from their [official page](#) for Windows. Once Python is installed run the command

```
python -c "import secrets; print(secrets.token_hex())"
```

And copy its output.

Running Locally

DocuMatcher is built entirely within a single Docker image, allowing for flawless compatibility and consistency across all platforms. Getting and starting the docker image can vary slightly depending on whether you're on Linux, Mac, or Windows.

The minimum prerequisite for running DocuMatcher locally is Docker Desktop. This can be installed from [Docker's official page](#).

Using The Pre-Built Image

For ease of use, a pre-built image is available on DockerHub, so you only need to download that image. It should be noted that all provided commands can be run from PowerShell as well as Docker Desktop's embedded terminal. All commands can be run from anywhere within PowerShell.

1. Once Docker Desktop is installed, you can open PowerShell and run the following command (this is one command)

```
docker run --name documatcher -d -p 8080:8080 --env SECRET_KEY=<secret-key>  
tulitque/documatcher:latest
```

DocuMatcher will now be accessible at <http://127.0.0.1:8080/>. To stop the container you can use the command

```
docker stop documatcher
```

If you want to restart the container use the command

```
docker start documatcher -d
```

Building From Source

Building the DocuMatcher from source requires both Docker Desktop and [GitHub Desktop](#) which is why it is recommended that you use the image.

1. Using GitHub Desktop clone the DocuMatcher's repository from <https://github.com/normit581/cits3200-proj> into your desired destination
2. Navigate to the 'deployment' folder of the cits3200-proj repository
3. Create a file called 'app.env' and put your secret key inside it. The app.env file should now look like

```
SECRET_KEY=<your-secret-key>
```

4. Go to the root of the app repository (the parent folder of 'deployment' and it will contain a 'docker-compose.yml' file)
5. Copy the path to the root folder. It should look something like

```
C:\path\to\repo\cits3200-proj
```

6. Open PowerShell and navigate to the repo's root directory using the command

```
cd C:\path\to\repo\cits3200-proj
```

7. Run the command

```
docker compose up --build -d
```

DocuMatcher will now be running and accessible from <http://127.0.0.1:8080/>. To stop the container you can run the following command from the repository's root directory.

```
docker compose down
```

For any subsequent runs where no changes have been made, you can use the following command, again from the repository's root directory.

```
docker compose up -d
```

Running On AWS Lightsail Containers (Recommended)

AWS' Lightsail Containers are the simplest of all the methods to run, especially with DocuMatcher having a pre-built image on DockerHub. You are still required to generate a secret key, however, besides Python, there is nothing you are required to install.

1. Login to your AWS Lightsail account in the browser and go to the 'Containers' section using the aside.
2. Click the Create Container service.

3. The first two options are ‘Service Location’ and ‘Service Capacity’; fill them in with your desired locations and capacity (I recommend ‘Sydney’ zone location, ‘Micro’ power, and ‘1’ scale).
4. Below the ‘Capacity’ section click ‘Set up deployment’, then select ‘Specify a custom deployment’.
5. Configuring Deployment (do not copy the “ “ “ when filling out the configuration)
 - a. Container Name: ‘documatcher’, or your desired name
 - b. Image: ‘tulitque/documatcher:latest’
 - c. Launch Command: leave blank
6. Select ‘Add environment variables’
 - a. Key: ‘SECRET_KEY’
 - b. Value: the secret key that you created
7. Select ‘Add open port’
 - a. Port: ‘8080’
 - b. Protocol: ensure it is set to ‘HTTP’
8. For the public endpoint choose the name of the container you just created.
9. Fill out ‘Identify Service’ with the name you want for your service i.e. ‘DocuMatcher-Service’
10. After selecting ‘Create Container Service’ you will be redirected back to the ‘Containers’ Page where you can wait for your newly created service to deploy (it will take a while).
11. Selecting the container will take you to the page where you can see more details regarding that container.
12. The link beside ‘Public Domain’ in the container details page will take you to the DocuMatcher site that you just deployed.

Container name
Container names must contain only alphanumeric characters and hyphens. A hyphen (-) can separate words but cannot be at the start or end of the name.

Image
Enter the image reference from a public registry, such as DockerHub.

Configuration
Optionally specify a command, the environment variables, and the ports to open on your container.
Launch command:

Environment variables
Key Value (optional)

SECRET_KEY	<your-secret-key>
------------	-------------------

+ Add variable

Open ports
Your application code for this container must listen to a port specified here.
Port Protocol

8080	HTTP
------	------

+ Add port

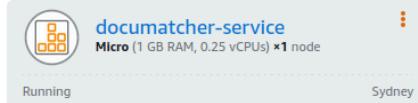
PUBLIC ENDPOINT
Choose a container in your deployment that you want to make available to the internet as a public endpoint.
Make sure to open an HTTP or HTTPS port on the selected container configuration, and then choose it as the port of your public endpoint.
The container you choose as your public endpoint must respond to traffic on the specified port.
Container:
Port: (?)
Health check path:
Show all settings

Good afternoon!

Filter by name, location, tag, or type

Sort by **Creation date** ▾

Create container service



documatcher-service

Container service

Micro (1 GB RAM, 0.25 vCPUs) ×1 node

Sydney

Disable

Status: **Running**

Public domain: documatcher-service.ytv19cvqksd3p.ap-southeast-2.cs.amazonaws.com ⓘ

How do I use my domain with my container service? ⓘ

Private domain: documatcher-service.service.local

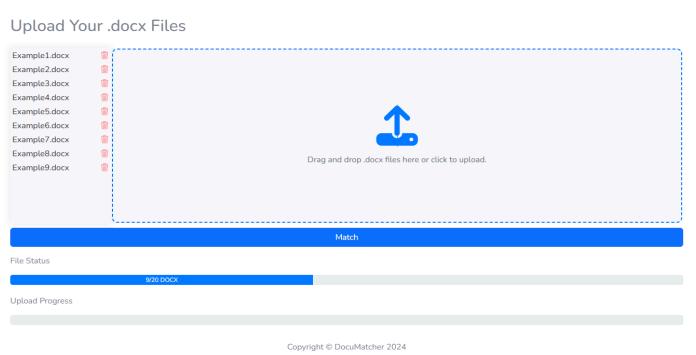
Interface Guide

Upload Page (Home Page)

This section explains how users can interact with the upload interface on DocuMatcher, including the process of uploading files, what types of files are accepted, how errors and duplication are handled, file size and quantity limits, and key actions available in the upload process.

1. Using the Drop Zone:

- Users will see a **large drop zone** with an upload icon and the message:
"Drag and drop .docx files here or click to upload."
- Users can either:
 - **Drag and drop .docx** files directly into the drop zone.
 - **Click on the drop zone** to open the **file explorer** and manually select files.



2. Real-Time Upload Tracking:

- As files are added, **progress bars** display upload progress:
 - **File Status Bar:** Shows how many files have been successfully uploaded.
 - **Upload Progress Bar:** Tracks the percentage of upload completion.

3. Reuploading Files:

- Users can click the **Reupload button** to modify their file selection without refreshing the page.

What Files Can Be Uploaded?

- **Accepted File Type:**
 - The system only supports **.docx** files for upload.
 - Attempting to upload other file types will generate an error:
"Only .docx files are allowed."
- **Maximum File Size:**
 - Each file must not exceed **220 MB**.
 - If the uploaded file exceeds this limit, users will see:
"File [filename] exceeds the maximum size of 220 MB."
- **Number of Files Allowed:**
 - **Minimum of 2 files** must be uploaded to perform a comparison.
 - In production, a pre-set 'limit' for the file upload cap can be adjusted during installation to suit server requirements and prevent overload.

Duplicate Files:

- The system automatically detects duplicate files. Users will see a warning icon () next to the file in the list if the same file is uploaded twice.
- A tooltip will provide additional information:
"Duplicate file with [filename]."

Note that all duplicated file name with warning will be removed.
Continue?

OK

Cancel

Managing Duplicates

- **Duplicate Detection:**
 - The system checks for identical files and prompts the user to either continue or modify the selection.
 - If a duplicate is detected, a warning icon () will appear in the file list, and users will receive a confirmation prompt:
"Duplicate detected. Do you wish to continue?"
- **Replacing Duplicate Files:**
 - Users can modify their selection by **removing the duplicate** using the **trash icon** next to the file name or by clicking **Reupload** to replace it with another file.

Example1.docx	
Example2.docx	
Example3.docx	
Example4.docx	
Example5.docx	
Example6.docx	
Example7.docx	
Example8.docx	
Example9.docx	
⚠ Example1.docx	

Actions Available in the Upload Interface

1. **Delete Files:**
 - Use the **trash icon** () next to a file to remove it from the upload list. This action updates the progress bar accordingly.
2. **Reupload Files:**
 - Users can click the Reupload button to restart the file selection process if they need to make changes.
3. **Submit Files for Comparison:**
 - Once the required files are uploaded, click the **Match button** to start the comparison process and progress to the **Summary Page**.

Summary Page (Match Page)

After successfully uploading files, users are taken to the **Post-Upload Summary Page**. This section outlines how users can explore their uploaded documents and compare results through different views, sorting options, search, percentage filtering, and export functionalities.

Overview

Once files are uploaded, DocuMatcher processes the documents and displays the comparison results through an intuitive file view. The **sidebar** offers an organised structure for navigating between different document pairs, while the **content area** displays the results of each comparison.

- **Sidebar:**
 - The **sidebar** offers an organised structure to navigate between different document pairs, while the **content area** displays the results of each comparison.
- **Main Content Area:**
 - Displays the comparison data in two formats:
 1. **List View** (detailed breakdown)
 2. **Grid View** (compact card layout)
- **Details Shown:**
 - **File Name:** Displays the name of the compared document.
 - **Match Percentage:** Shows the similarity percentage (e.g., 45.2%) with a **coloured badge** (green, yellow, or red) based on the degree of similarity with the selected document.
 - **Match Count:** Shows the total number of matched RSID codes.
 - **Interactive Elements:** Clicking any row triggers the **visual comparison page** for the selected document pair.
- **Colour Coding for Similarity:**
 - **Red Badge (High Match):** > 60% similarity
 - **Yellow Badge (Moderate Match):** 30% – 60% similarity
 - **Green Badge (Low Match):** < 30% similarity

The screenshot shows the DocuMatcher Summary Page. At the top left is the DocuMatcher logo. To its right is a search bar with placeholder text "Search...". On the far right, there is a user icon and the text "Our Team". The main content area is divided into two sections: a sidebar on the left and a large table on the right.

Sidebar: A vertical list of document names, many of which include the suffix "- Copy". The items are:

- Example (1)
- Example (10)
- Example (11)
- Example (12)
- Example (13)
- Example (13) - Copy
- Example (14)
- Example (14) - Copy
- Example (15) - Copy** (highlighted in blue)
- Example (16) - Copy
- Example (17) - Copy
- Example (2)
- Example (3)
- Example (3) - Copy
- Example (4)
- Example (5)
- Example (6)
- Example (7)
- Example (8)
- Example (9)

Content Area: A table displaying document comparisons. Each row contains a file icon, the document name, a file number, and a similarity percentage badge. The rows are color-coded by similarity:

File	#	Similarity (%)
Example (13) - Copy	679	100.0%
Example (13)	679	100.0%
Example (16) - Copy	679	100.0%
Example (14) - Copy	2106	98.1%
Example (17) - Copy	223	32.7%
Example (3) - Copy	23	00.0%
Example (10)	179	00.0%
Example (1)	76	00.0%
Example (2)	0	00.0%
Example (3)	0	00.0%
Example (4)	97	00.0%
Example (5)	97	00.0%
Example (6)	98	00.0%

List View

The **List View** provides a detailed, line-by-line breakdown of the comparison results.

- **How to Access:**

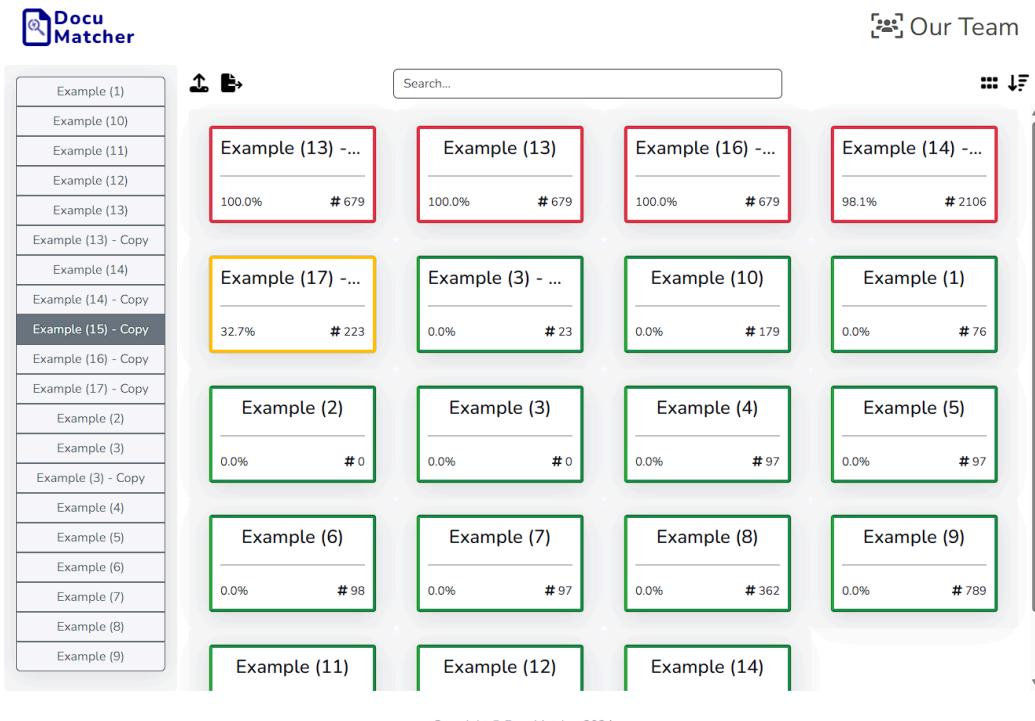
- Toggle to the **List View** by clicking the **List View button** in the toolbar or by selecting the option in the right-click **Context Menu** 

Grid View

The **Grid View** provides an easy-to-view, card-like structure for the comparison results.

- **How to Access:**

- Toggle to the **Grid View** by clicking the **Grid View button** in the toolbar or by selecting the option in the right-click **Context Menu** 



Sorting Options

Sorting helps users quickly rearrange the comparison results based on relevance.

- **How to Use:**

1. Click the **Sort Button** in the toolbar to toggle between different sorting modes.

- **Available Sort Modes:**

1. **Match % Descending:** Displays the highest similarity first. (Default) 
2. **Match % Ascending:** Displays the lowest similarity first 
3. **Filename A-Z:** Sorts files alphabetically from A to Z. 
4. **Filename Z-A:** Sorts files alphabetically from Z to A. 

Search Functionality

The **Search Bar** helps users find specific documents or results efficiently.

- **How to Use:**
 - Type in the search bar located in the **toolbar** at the top of the results page.
- **How It Works:**
 - Searches through **file names** and dynamically filters the results.
 - Works across both **list and grid views**.

A screenshot of a search interface. At the top is a search bar containing the text "Example (13)". Above the search bar are two icons: a magnifying glass and a document icon. To the right of the search bar are three more icons: a list icon, a document icon, and a refresh/circular arrow icon. Below the search bar is a list of results. The first result is "Example (13) - Copy" with ID # 679 and a similarity of 00.0%. The second result is "Example (13)" with ID # 679 and a similarity of 00.0%. The results are displayed in a grid view.

Percentage Filtering

The percentage filter allows users to **narrow down** the displayed results based on match similarity.

- **How to Use:**
 - Adjust the **slider** located in the **custom context menu** to set the minimum similarity percentage.
 - Matching documents with similarity below the set threshold will be **hidden**.
- **Live Filtering:**
 - As the slider value changes, the results update instantly to only show documents that meet the criteria.

A screenshot of a search interface showing a list of results. The results are:

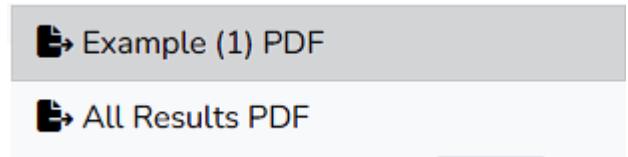
- "Example (15) - Copy" with ID # 679 and a similarity of 100.0% (highlighted with a red border).
- "Example (13) - Copy" with ID # 679 and a similarity of 100.0%.
- "Example (13)" with ID # 679 and a similarity of 100.0%.
- "Example (14) - Copy" with ID # 2106 and a similarity of 98.1%.
- "Example (17) - Copy" with ID # 223 and a similarity of 32.7% (highlighted with a yellow border).

Below the results is a slider control labeled "Min" with a value of 20 and a percentage indicator of 20%.

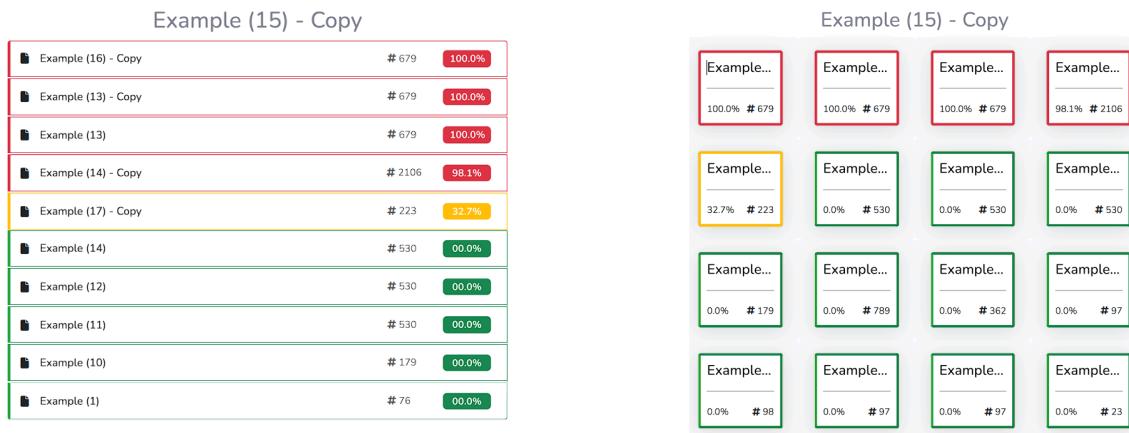
Both the Percentage and Search Function can work together to filter results simultaneously.

Export Options

DocuMatcher allows users to **export comparison results** for further analysis or documentation on the summary page.



- **Single PDF Export:**
 - Click the **Export PDF button** to export the selected document pair's comparison results as a **PDF**. 
 - **How to Export:**
 1. Select a document pair from the sidebar.
 2. Click the **Export PDF button** to generate and download the PDF.
- **All Results PDF Export:**
 - Use the **All Results PDF button** to generate a single PDF containing the comparison results for **all uploaded document pairs**.



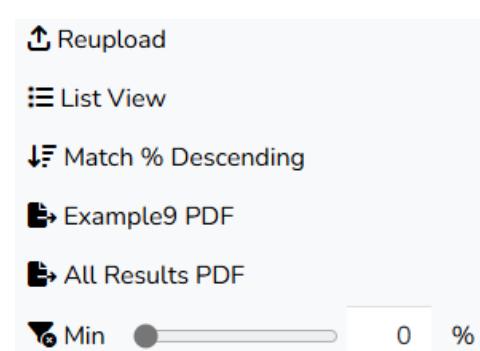
Reuploading Files

If users need to **modify or re-upload files**, they can click the **Reupload button** in the toolbar. 

Clicking the button takes users back to the **upload interface**, where they can select new files.

Context Menu Options

DocuMatcher offers a right-click **context menu** with additional features for ease of access.



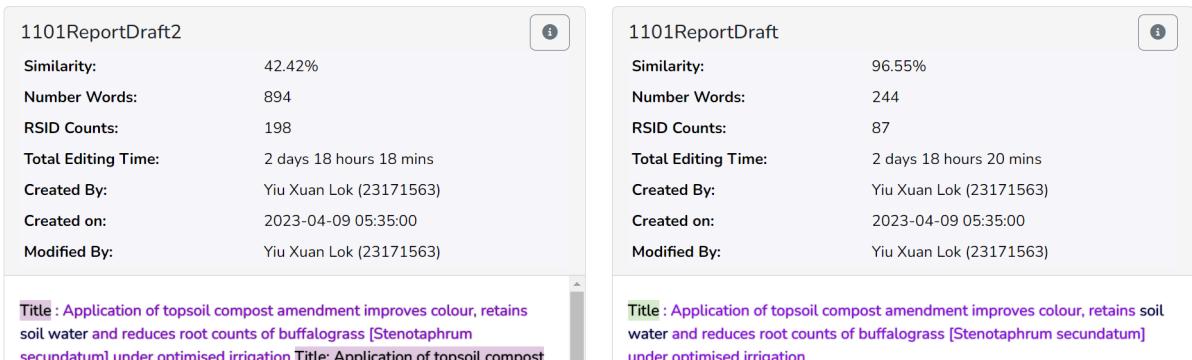
- **How to Access:**
 1. Right-click on the main content area or the sidebar to open the context menu.
- **Context Menu Actions:**
 1. **Switch Views:** Easily toggle between **list and grid views**.
 2. **Adjust Filters:** Modify the similarity percentage filter directly from the menu.
 3. **Adust Percentage Match Threshold**
 4. **Reupload Files:** Quickly access the reupload interface.

RSID Comparison Page (Visualise Page)

Docx Summary

Access the docx summary by clicking the  icon. It can be collapsed using the  button.

- **Similarity:** Displays the comparison from .docx_1 to .docx_2
- **Result of .docx_1:** Calculated as $\frac{100 \times commonCount}{totalRsidCountDocx1}$
- **Number of Words:** Shows the total word count in the specified .docx.
- **RSID Count:** Indicates the total RSID count in the specified .docx.
- **Total Editing Time:** Translates the time stored in the .docx file from minutes to days:hours:minutes.
- **Created By:** Displays the original creator of the .docx file.
- **Modified By:** Shows the individual who modified the .docx file.



The screenshot shows two document summaries side-by-side. Each summary includes a title, a collapse button, and a detailed list of metrics.

Document	Similarity	Number Words	RSID Counts	Total Editing Time	Created By	Created on	Modified By
1101ReportDraft2	42.42%	894	198	2 days 18 hours 18 mins	Yiu Xuan Lok (23171563)	2023-04-09 05:35:00	Yiu Xuan Lok (23171563)
1101ReportDraft	96.55%	244	87	2 days 18 hours 20 mins	Yiu Xuan Lok (23171563)	2023-04-09 05:35:00	Yiu Xuan Lok (23171563)

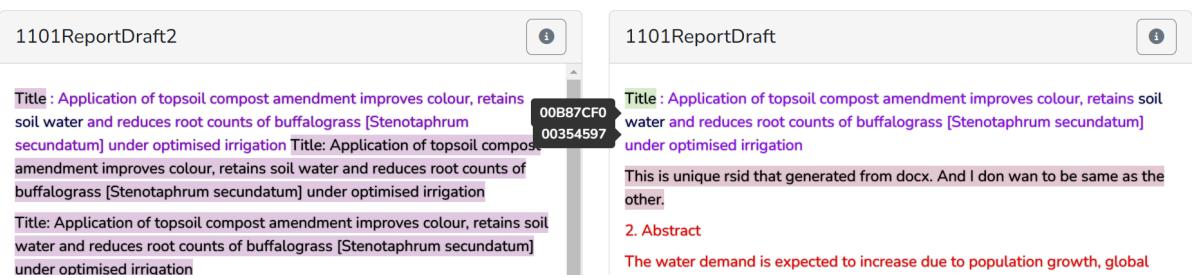
Below each summary, there is a tooltip for the 'Title' field:

1101ReportDraft2: Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

1101ReportDraft: Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

Hover RSID Text

Tooltips displaying the RSID from the corresponding section will appear when hovering over the text, allowing users to view relevant RSID details.



The screenshot shows the same two document summaries as above, but with a tooltip displayed over a specific RSID value in the first document's summary.

1101ReportDraft2: Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

1101ReportDraft: Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

A tooltip is shown for the RSID value **00B87CF0 00354597**, which is highlighted in black. The tooltip contains the following text:

This is unique rsid that generated from docx. And I don wan to be same as the other.
2. Abstract
The water demand is expected to increase due to population growth, global

Border Colour Green/Red

Text with a **red underlined** indicates sections with matched RSIDs in both .docx files. In contrast, **green underlined** text signifies content unique to that .docx file.

The screenshot shows two Microsoft Word documents side-by-side. Both documents have a title section with matching RSIDs (red underlined) and unique content (green underlined). The left document has a red border around its content area, while the right document has a green border. The right document also includes a context menu with various options like PDF export, font size adjustment, and text visibility controls.

1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss

Context Menu

The context menu can be accessed by right-clicking with the mouse and includes several useful features. The functions of these buttons will be explained in the following sections.

The screenshot shows two Microsoft Word documents. The right document has a context menu open over a red underlined section. The menu includes options like PDF export, font size adjustment, and text visibility controls. The menu is semi-transparent, allowing the underlying text to be seen.

1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss

Switch Button

Users can easily switch between docx1 and docx2 without needing to click the matching result again.

The screenshot shows two Microsoft Word documents. The right document has a context menu open with the 'Switch' option highlighted. A tooltip '003C3E6B' is visible near the cursor. The menu also includes options like PDF export, font size adjustment, and text visibility controls.

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss

1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

amendment improves colour, re

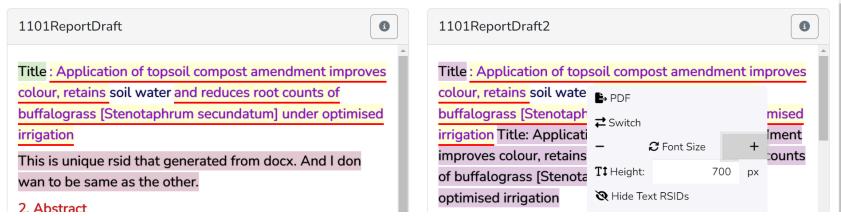
Title: Application of topsoil com

2. Abstract

Font Size Button Group

- Users can adjust the font size to either large or small using the + or - buttons.
- The font size can be reset by clicking the  button, with the default size set to 12px.
- Adjusting the font size on the comparison page enhances readability and allows users to fit more content on the screen, customising their viewing experience to suit their preferences.

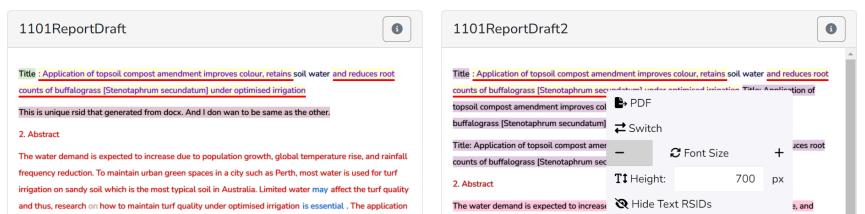
Large



The screenshot shows two document versions side-by-side. The left document, titled '1101ReportDraft', contains the following text:
Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation
This is unique rsid that generated from docx. And I don wan to be same as the other.
2. Abstract
The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential. The application

The right document, titled '1101ReportDraft2', shows the same content but with a larger font size applied to the entire body text. The font size is set to 700px. A 'Font Size' slider with '+' and '-' buttons is visible at the bottom of the right panel.

Small



The screenshot shows two document versions side-by-side. The left document, titled '1101ReportDraft', contains the same text as the first example. The right document, titled '1101ReportDraft2', shows the same content but with a smaller font size applied to the entire body text. The font size is set to 100px. A 'Font Size' slider with '+' and '-' buttons is visible at the bottom of the right panel.

Height

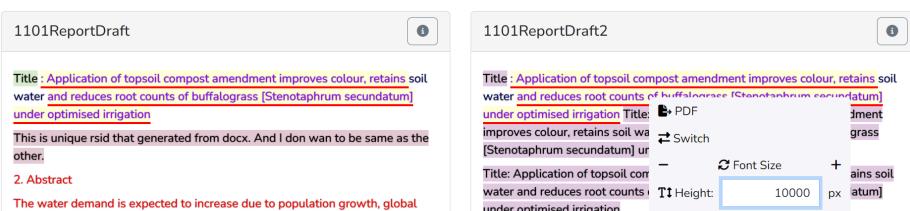
Users can manually adjust the height of the paragraph body content to a shorter or taller view. This feature is useful for eliminating scrolling in each document body.

Short



The screenshot shows two document versions side-by-side. The left document, titled '1101ReportDraft', contains the same text as the previous examples. The right document, titled '1101ReportDraft2', shows the same content but with a shorter height for the body text. The height is set to 100px. A 'Height' input field with a 'px' unit is visible at the bottom of the right panel.

Tall



The screenshot shows two document versions side-by-side. The left document, titled '1101ReportDraft', contains the same text as the previous examples. The right document, titled '1101ReportDraft2', shows the same content but with a tall height for the body text. The height is set to 10000px. A 'Height' input field with a 'px' unit is visible at the bottom of the right panel.

Hide RSID Colour

- Individual:** The *Hide/Show Individual RSID* button hides the background colour of the selected RSID on the visualisation page, reflecting changes in both documents.

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be based on the turf's greenness, soil water content, and root counts. This experiment was done by incorporating 10cm top soil amendment with irrigation twice a week in 50% replacement of evapotranspiration and analysing turf

1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be based on the turf's greenness, soil water content, and root counts. This experiment was done by incorporating 10cm top soil amendment with irrigation twice a week in 50% replacement of evapotranspiration and analysing turf

- Match:** The *Hide/Show Matching RSID* button hides the background colours for all matching RSIDs on the visualisation page, also reflecting in both documents.

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be based on the turf's greenness, soil water content, and root counts. This experiment was done by incorporating 10cm top soil amendment with irrigation twice a week in 50% replacement of evapotranspiration and analysing turf

1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be based on the turf's greenness, soil water content, and root counts. This experiment was done by incorporating 10cm top soil amendment with irrigation twice a week in 50% replacement of evapotranspiration and analysing turf

Change RSID Colour

- Users can change the colour for a selected RSID paragraph using the colour input.
- The HEX colour code will update and sync accordingly.
- If the text matches the RSID, the colour change will be reflected in other documents as well.

1101ReportDraft

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

This is unique rsid that generated from docx. And I don wan to be same as the other.

2. Abstract

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be based on the turf's greenness, soil water content, and root counts. This experiment was done by incorporating 10cm top soil amendment with irrigation twice a week in 50% replacement of evapotranspiration and analysing turf

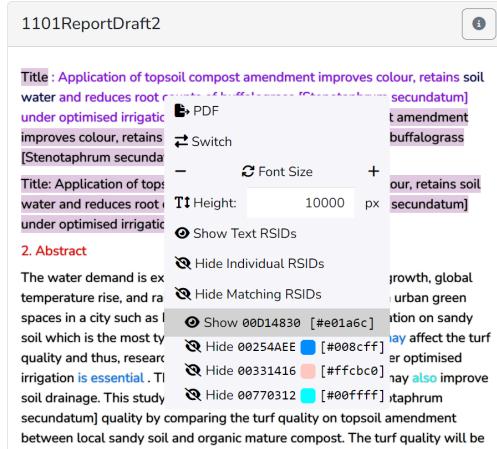
1101ReportDraft2

Title : Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation Title: Application of topsoil compost amendment improves colour, retains soil water and reduces root counts of buffalograss [Stenotaphrum secundatum] under optimised irrigation

The water demand is expected to increase due to population growth, global temperature rise, and rainfall frequency reduction. To maintain urban green spaces in a city such as Perth, most water is used for turf irrigation on sandy soil which is the most typical soil in Australia. Limited water may affect the turf quality and thus, research on how to maintain turf quality under optimised irrigation is essential . The application of topsoil amendment may also improve soil drainage. This study aimed to observe buffalograss [Stenotaphrum secundatum] quality by comparing the turf quality on topsoil amendment between local sandy soil and organic mature compost. The turf quality will be

Hide Single Colour

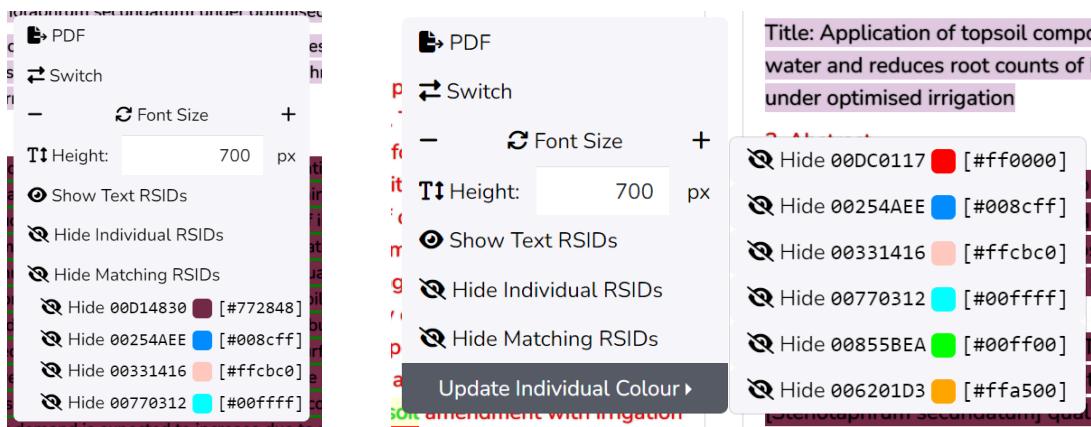
- Users can hide a single colour for the selected paragraph.
- If the text matches the RSID, the change will also reflect in other document colours.



Colour Input Display

Two views are available for colour input:

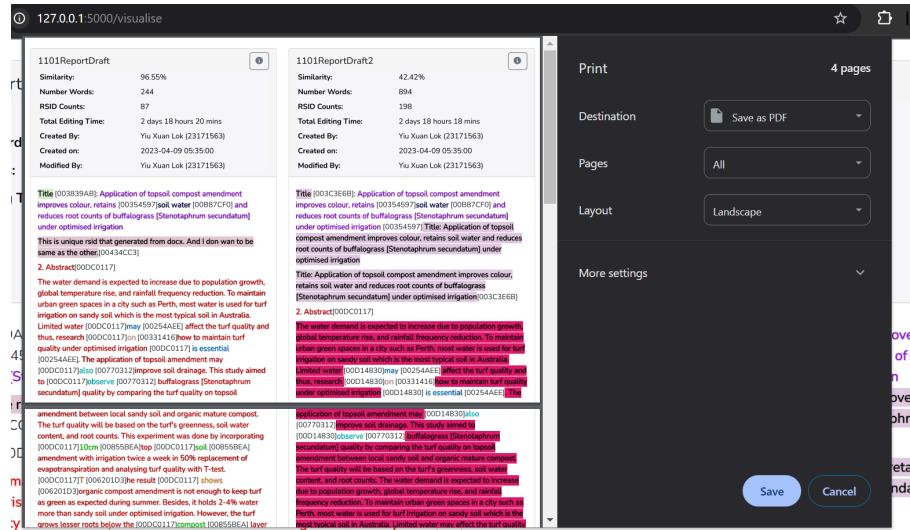
- If the RSID in the selected paragraph is less than 5, the colours are displayed below the context menu.
- For colour inputs greater than 5, they appear in a dropdown list to prevent exceeding the window size.



Exporting PDF

When generating the PDF with the button  in the Context Menu:

- The RSID text is displayed beside coloured sections when generating a PDF;
- Customised colours are updated and shown in the PDF;
- Summary details from each DOCX file are included in the PDF.



Performance

In order to measure the performance and real-world use case of the DocuMatcher platform, detailed below are distinct user test cases and stress-test results, which may help end-users make decisions about deploying and the expected performance.

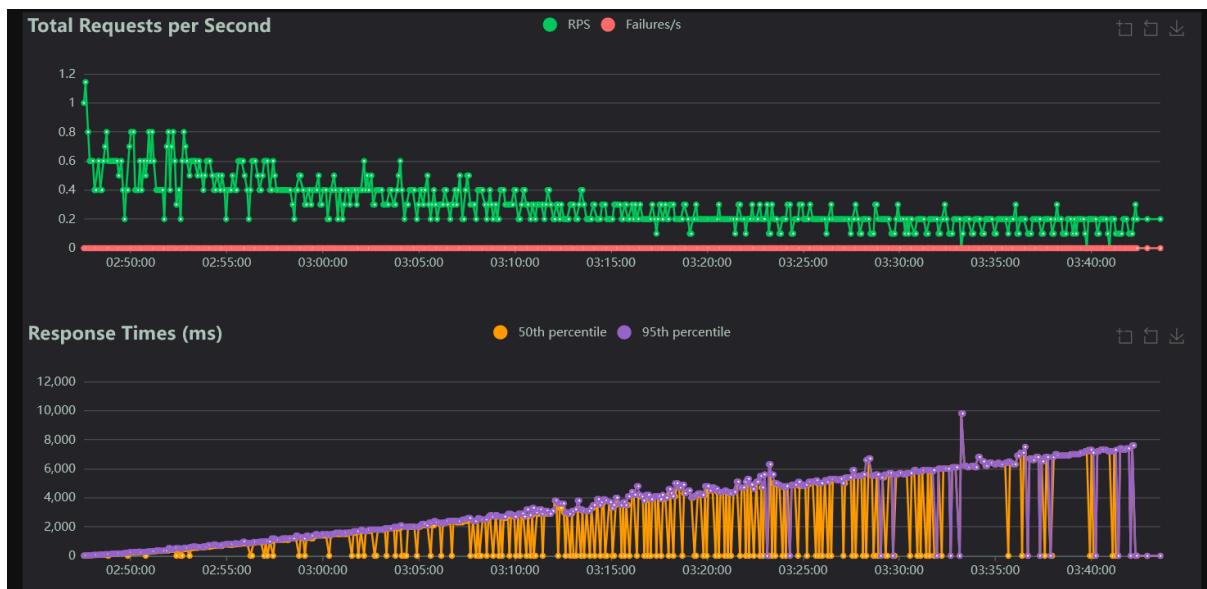
Test Case 1

Test samples

- 1 user
- 490 randomly generated .docx files
- each size: 34kb
- each contains ~1200 rsid count

Objective

The objective of this test was to evaluate the system's performance and responsiveness when uploading .docx files. The test began with 2 .docx files and incremented the upload count by 1 until reaching 490 files. The key metrics analysed were failure rate and response times (50th and 95th percentiles).



Failures:

There were no failures recorded during the test, indicated by the flat red line. The system maintained stability and reliability throughout the incremental uploads, with no significant failure rate.

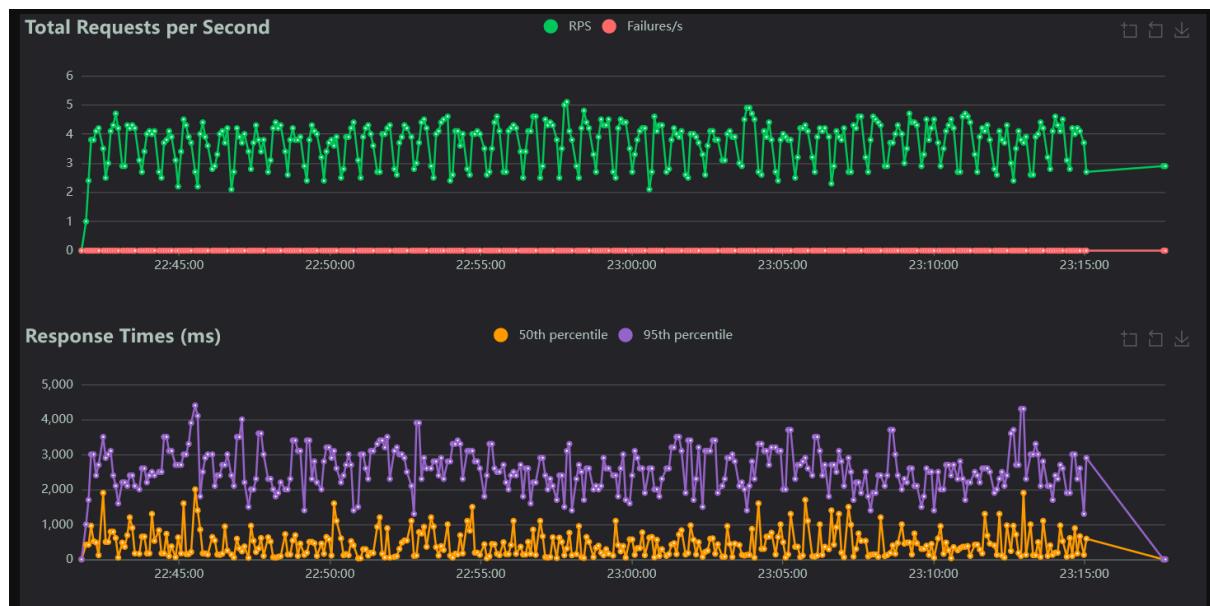
Response Times:

Trend: Both percentiles show a steady rise in response times, indicating that as more files were uploaded, the system took longer to process requests, and performance is linear.

Test Case 2

Test samples

- 10 users
- randomly upload 50 - 100 .docx per user
- each size 34kb
- each contains ~1200 rsid count



This load test was run for 30 minutes with 10 users randomly uploading between 50 and 100 .docx files. The upper chart shows that the request rate (RPS) stabilised between 3 to 5 requests per second, with no failures throughout the test. The lower chart indicates response times: the 50th percentile (median) response times fluctuated between 500 ms and 1,500 ms, while the 95th percentile occasionally spiked to over 3,000 ms, indicating some higher latency under load but still without errors.