User Manual
For SimiLabs

Table of Contents

Contents

Introduction

Welcome to the SimiLabs Document Upload System user manual. This document serves as a comprehensive guide to help you navigate the features and functionalities of our system, enabling you to efficiently manage and process your documents. In this manual, you will learn how to connect to the server, access various services, and utilize the system's capabilities to their fullest potential. Follow the outlined steps for a smooth setup and to ensure seamless interaction with the web interface and supporting applications.

A. System overview

This document processing and analysis system addresses the escalating societal need for data-driven insights, particularly within the education sector. As demands for informed decision-making, improved educational outcomes, and innovation continue to rise, universities face the challenge of extracting meaningful information from documents. The system's aim is to efficiently handle large volumes of documents while enabling users to customise analysis features based on their specific needs, thereby enhancing decision-making processes, and driving innovation within the education sector.

Once a document is uploaded and analysis functions are selected, the document enters a queue to prevent system crashes during heavy load periods. The system follows a modular design where all the analysis functions are hosted separately, which enables flexibility in adding or removing features based on user requirements and allows for load balancing by creating additional instances of specific services if necessary. The document then sequentially undergoes analysis by each selected service. Once all analysis functions are completed, the user is presented with an output report summarising the findings, which is saved to a database for later analytic purposes.

B. System requirements

To successfully use the system, please ensure you meet the following requirements regarding network access, operating systems, supported browsers, and JavaScript settings:

- Network Access: The system is hosted on an internal organization network and can only be accessed from:
    - The NWUWIFI wireless network.
    - A physical LAN port located on the campus.

- Operating System: The system is compatible with the following operating systems:

- Windows 10 or higher
- macOS 10.15 (Catalina) or higher
- Linux (any modern distribution such as Ubuntu 20.04+, Fedora, CentOS, etc.)

C. Supported Browsers: The system works with any modern web browser. It is recommended to use one of the following:
   a. Google Chrome (latest version)
   b. Mozilla Firefox (latest version)
   c. Microsoft Edge (latest version)
   d. Safari (latest version)

D. JavaScript Enabled: Ensure that JavaScript is enabled in the browser settings for optimal functionality.

Once you are set up, you will be able to explore the system's various features, including single and batch file uploads, report generation, and efficient document management.

System Setup

In this section, you will learn how to connect to the server and access the necessary services. Follow the steps outlined below to ensure a smooth setup and gain access to the web interface and other services. Before using the system, there are a few set-up steps to complete.

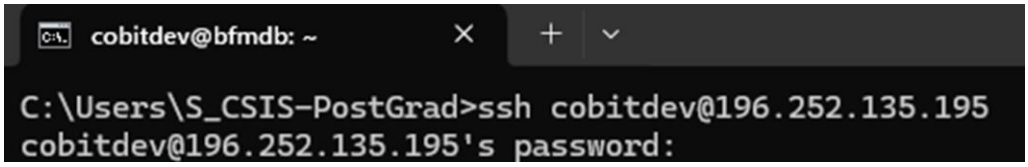
A. Server Login

To begin, you'll need to connect to the server via a terminal. Follow these steps:

1. Open your terminal application (Bash for macOS/Linux or PowerShell for Windows).
2. Connect to the server by running the following command:

```bash
ssh cobitdev@196.252.135.195
```

3. After running the command, you would be prompted to enter the password for the cobitdev account on the server.



Connection: ssh cobitdev@196.252.135.195

Password: iWillNeverBeEnough

B. Port Forwarding

After connecting to the server, you need to forward specific ports to your local machine. Open a new terminal and connect for each port you want to forward and run the following commands:

- **User Interface**:

```
cobitdev@bfmdb:~$ kubectl port-forward svc/ui-service 8082:80
Forwarding from 127.0.0.1:8082 -> 80
Forwarding from [::1]:8082 -> 80
```

Command: kubectl port-forward svc/ui-service 8082:80

- **API:**

```
cobitdev@bfmdb:~$ kubectl port-forward svc/gateway-service 8085:5000
Forwarding from 127.0.0.1:8085 -> 5000
Forwarding from [::1]:8085 -> 5000
```

Command: kubectl port-forward svc/gateway-service 8085:5000

- **MinIO:**

```
cobitdev@bfmdb:~$ kubectl port-forward pod/minio-0 9001:37563
Forwarding from 127.0.0.1:9001 -> 37563
Forwarding from [::1]:9001 -> 37563
```

Command: kubectl port-forward pod/minio-0 9001:"podport"

To check the port Minio is running on, use: kubectl logs minio-0. In the logs, look for the line under **WebUI. The port number will be listed, typically in the format:

```
WebUI: http://10.244.1.60:37563 http://127.0.0.1:37563
```

C. Connecting

After forwarding the necessary ports, you need to connect to each. Open a new terminal for each port you want to connect to and run the following commands:

- **User Interface:**

```
C:\Users\S_CSIS-PostGrad>ssh -L 8082:127.0.0.1:8082 cobitdev@196.252.135.195
cobitdev@196.252.135.195's password:
```

Command: ssh -L 8082:127.0.0.1:8082 cobitdev@196.252.135.195

Password: iWillNeverBeEnough

- **API:**

```
C:\Users\S_CSIS-PostGrad>ssh -L 8085:127.0.0.1:8085 cobitdev@196.252.135.195
cobitdev@196.252.135.195's password:
```

Command: ssh -L 8085:127.0.0.1:8085 cobitdev@196.252.135.195

Password: iWillNeverBeEnough

- **MinIO:**

```
C:\Users\S_CSIS-PostGrad>ssh -L 9001:127.0.0.1:9001 cobitdev@196.252.135.195
cobitdev@196.252.135.195's password:
```

Command: ssh -L 9001:127.0.0.1:9001 cobitdev@196.252.135.195

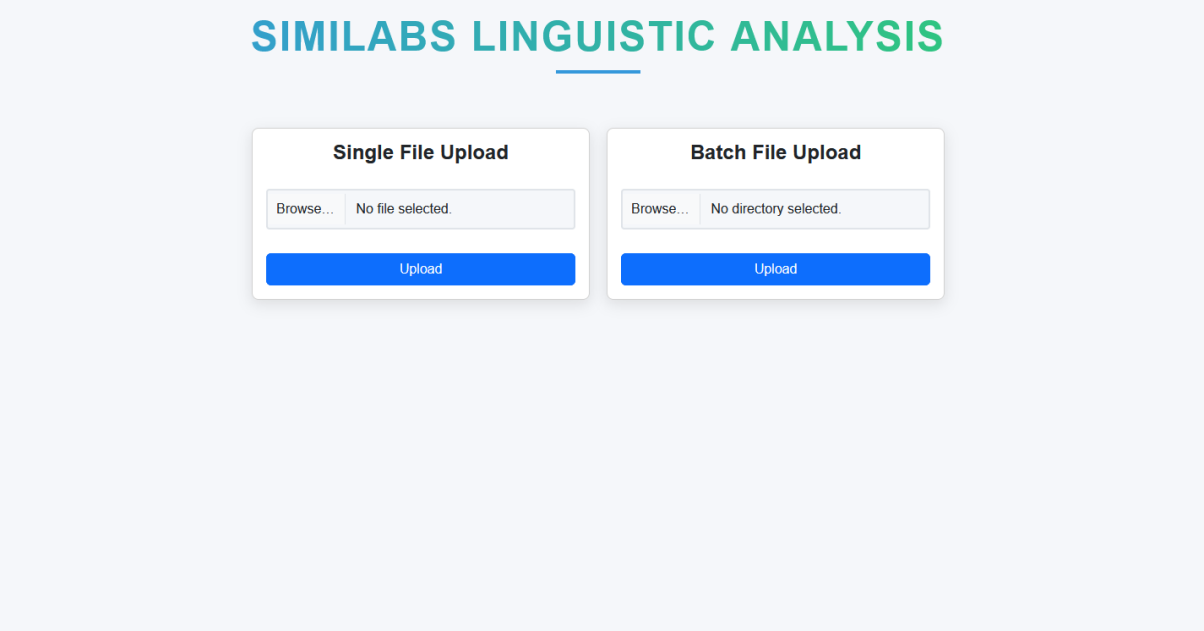
Password: iWillNeverBeEnough

D. System Access:

To get started, navigate to the following URLs in any web browser:

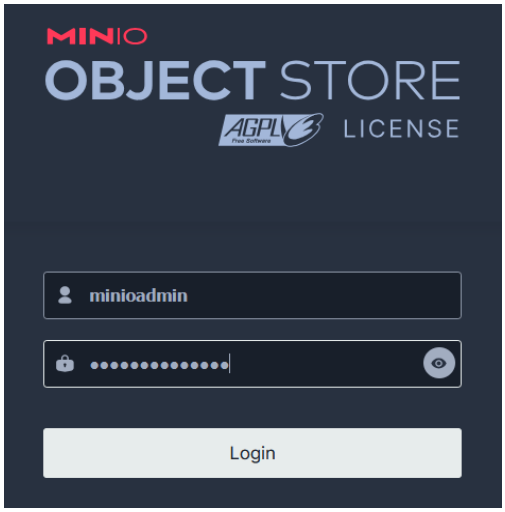
User Interface: http://127.0.0.1:8082/

If the setup is successful, you will be presented with the following interface:



MinIO: http://127.0.0.1:9001

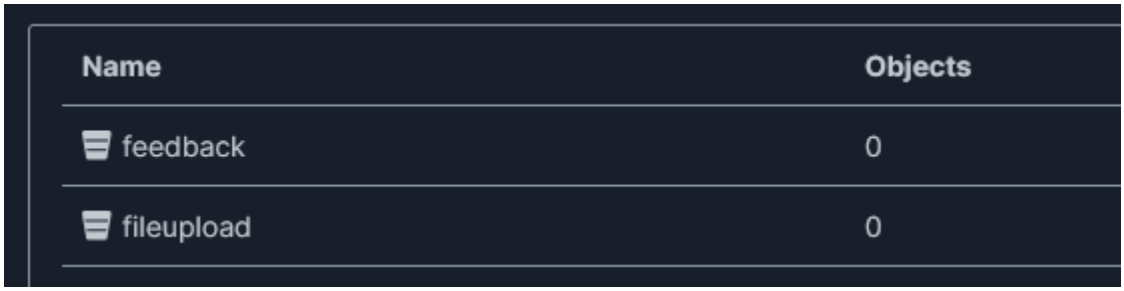
If the setup is successful, you will be presented with the following interface. Login into Minio using the following credentials:



Username: minioadmin

Password: similabs_minio

Once logged in, navigate to the **feedback** bucket in the MinIO interface. This bucket contains all the document reports.



Getting Started

Now that you have successfully set up the system, it's time to explore its features and functionality. This section will guide you through the various capabilities available within the platform, including how to interact with the MinIO interface, manage files, and utilize the feedback bucket.

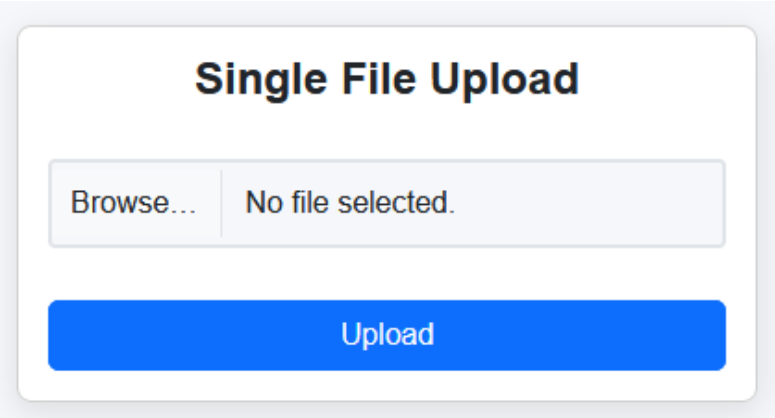
The system supports the following file types for upload and management:

- **PDF files**: .pdf
- **Word documents**: .docx

The system offers two convenient methods for uploading files: **Single Upload** and **Batch Upload**. Both methods are designed to be user-friendly, ensuring a seamless experience regardless of the volume of files you need to upload.

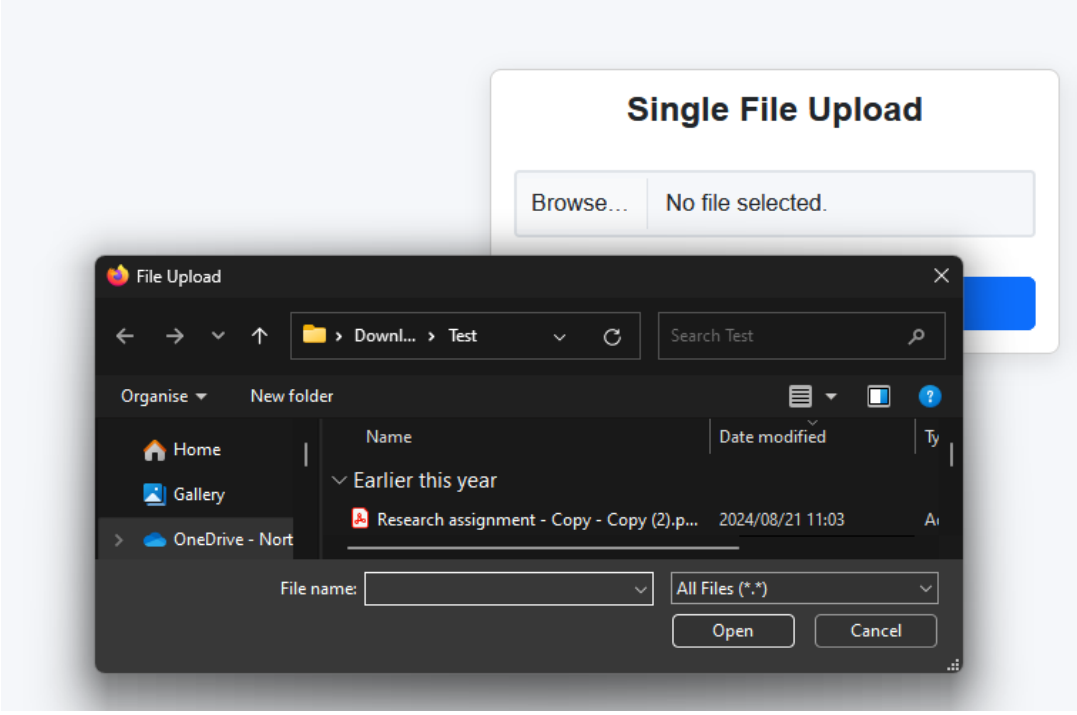
A. Single File Upload:

This option allows you to upload one file at a time. It is ideal for quick uploads and is straightforward to use. Simply select the file you wish to upload, and the system will handle the rest. This method is perfect for users who need to manage individual files without any bulk actions.



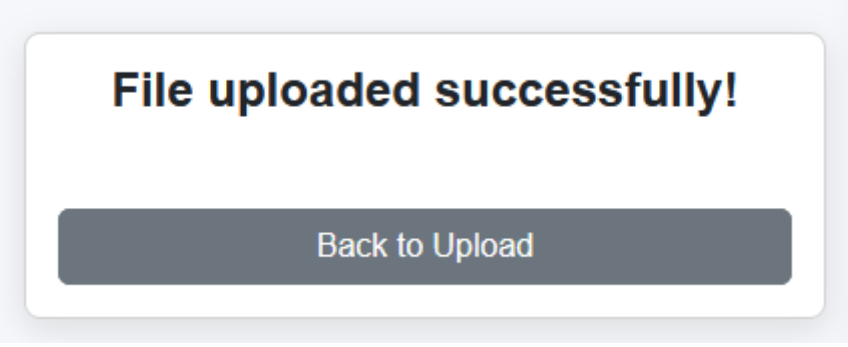
Browse for a File:

- Click the **Browse** button. This will open a file selection dialog where you can navigate to the file you wish to upload. Supported file types include .pdf and .docx.
- Once you locate your desired file, select it and click **Open**. The file path will be displayed next to the **Browse** button, indicating that your file is ready for upload.



Upload the File:

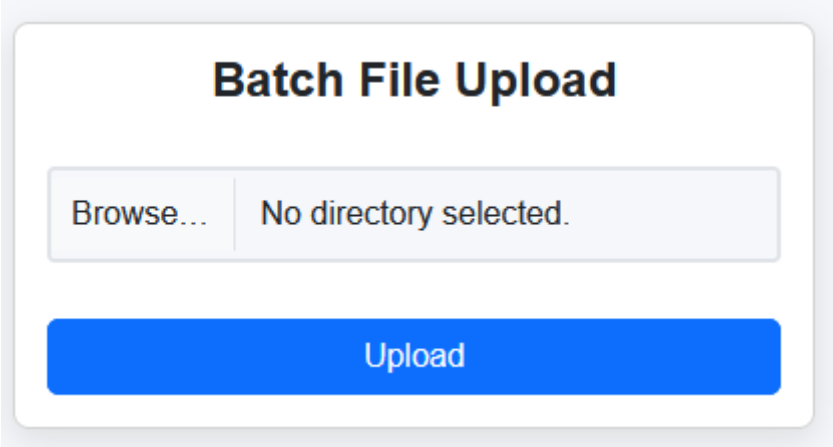
- After selecting the file, click the **Upload** button to initiate the upload process.
- If the file is successfully uploaded, you will see the following interface:



This interface confirms that your file has been uploaded and stored successfully and is queued to be processed.

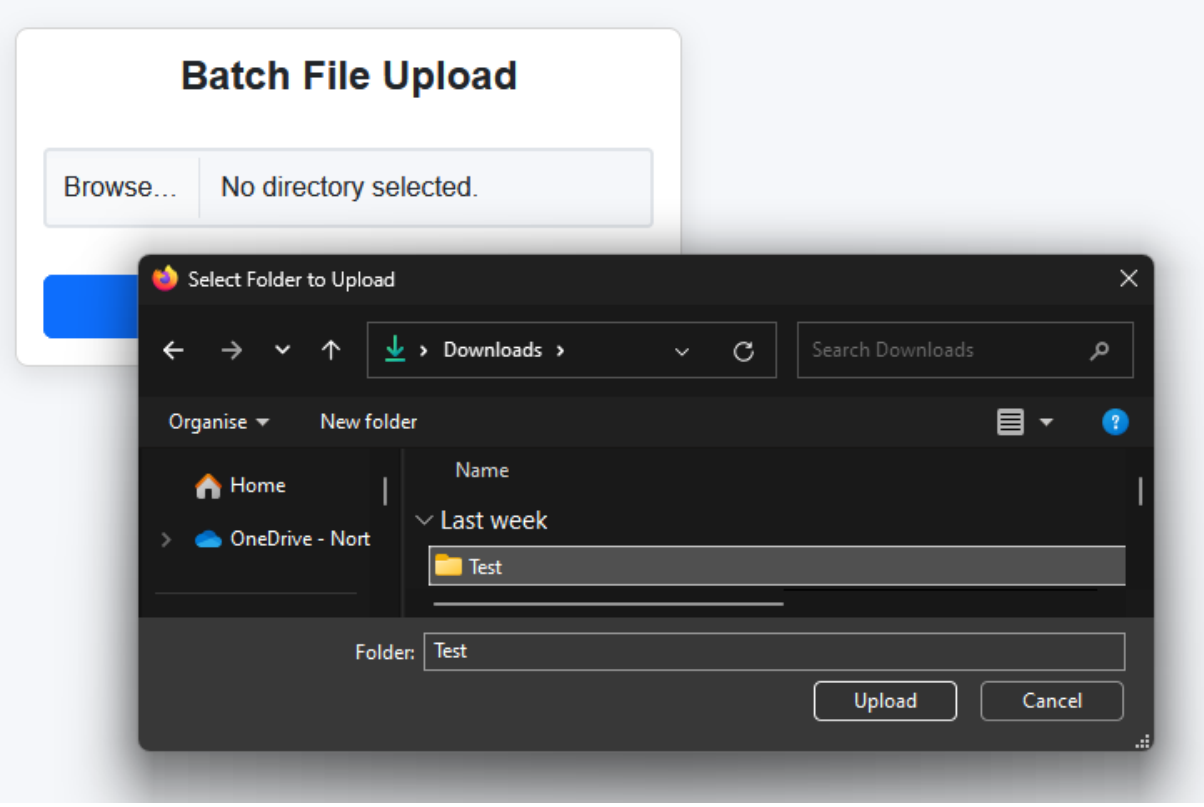
B. Batch File Upload:

The batch upload feature enables you to upload multiple files simultaneously. This is particularly useful for users who have many documents to process. By selecting several files at once, you can save time and streamline your workflow. The system will process each file in the batch, allowing for efficient management of multiple uploads.



Browse for a Folder:

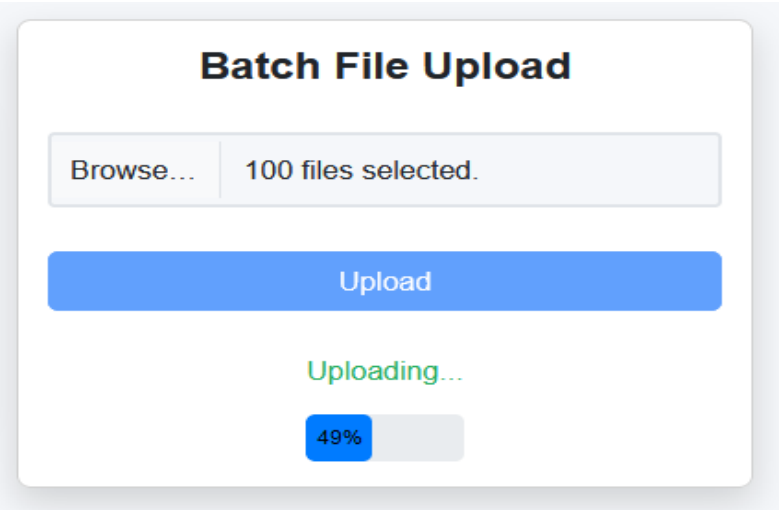
- Click the **Browse** button. This will open a folder selection dialog where you can navigate to the folder containing the files you wish to upload. Supported file types include .pdf and .docx.
- Once you locate your desired folder, select it and click **Open**. The folder path will be displayed next to the **Browse** button, indicating that your files are ready for upload.



Upload the Files:

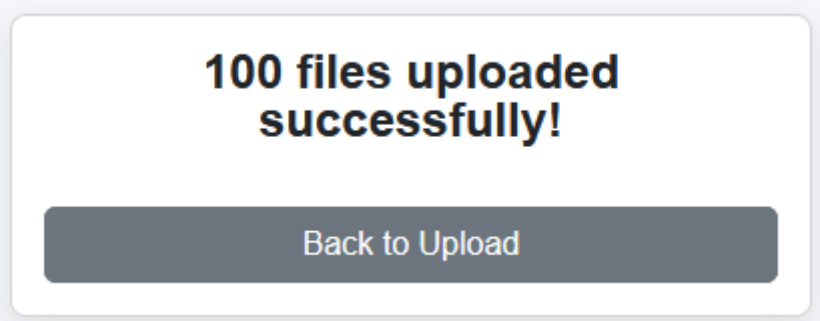
- After selecting the folder, click the **Upload** button to initiate the upload process. A progress bar will appear, indicating the status of the upload. Please note that this process may take some time, depending on the number and size of the files being uploaded.
- It is important to wait until the confirmation screen is displayed. If you navigate away or close the window before this confirmation appears, the upload process may not be successful.

Confirmation



of Upload:

- If the files are successfully uploaded, you will see the following interface:



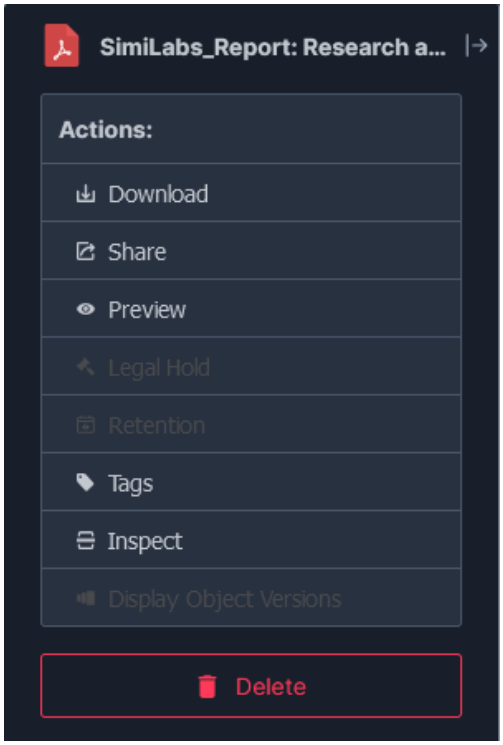
This interface confirms that your files have been uploaded and stored successfully and are queued to be processed.

C. Document Report

As the documents are processed, a PDF report is generated for each document and stored in the feedback bucket. The original filename remains unchanged; however, SimiLabs_Report: is prepended to the filename for easy identification.

Accessing Reports:

- You can view the generated reports in the feedback bucket. Selecting a specific report will allow you to either preview or download the report.
- If you wish to download reports for multiple documents, you can select several reports at once. This functionality enables you to easily obtain all the necessary reports for local use.



Report Contents:

- Each report presents the original file name and provides comprehensive results of the text analysis performed on the document. This includes key metrics and insights derived from the analysis.

SimiLabs Document Analysis Report

File Name: Research assignment - Copy - Copy (2).pdf

Results:

N-grams: [[['computer', 'science'], 2], [['paper', 'must'], 2], [['structure', 'rest'], 2], [['rest', 'paper'], 2], [['north', 'university'], 1], [['university', 'school'], 1], [['school', 'computer'], 1], [['science', 'information'], 1], [['information', 'systems'], 1], [['systems', 'itri'], 1], [['itri', 'databases'], 1], [['databases', 'ii'], 1], [['ii', 'research'], 1], [['research', 'ssignment'], 1], [['ssignment', 'sept'], 1], [['sept', 'marks'], 1], [['marks', 'choose'], 1], [['choose', 'one'], 1], [['one', 'two'], 1], [['two', 'microservices'], 1], [['microservices', 'software'], 1], [['software', 'pattern'], 1], [['pattern', 'worth'], 1], [['worth', 'discussing'], 1], [['discussing', 'research'], 1], [['research', 'paper'], 1], [['paper', 'api'], 1], [['api', 'gateway'], 1], [['gateway', 'give'], 1], [['give', 'critical'], 1], [['critical', 'review'], 1], [['review', 'th'], 1], [['th', 'pattern'], 1], [['pattern', 'addressing'], 1], [['addressing', 'following'], 1], [['following', 'aspects'], 1], [['aspects', 'context'], 1], [['context', 'pattern'], 1], [['pattern', 'problem'], 1], [['problem', 'addresses'], 1], [['addresses', 'solution'], 1], [['solution', 'provides'], 1], [['provides', 'benefits'], 1], [['benefits', 'example'], 1], [['example', 'paper'], 1], [['paper', 'microservices'], 1], [['microservices', 'resilience'], 1], [['resilience', 'fault'], 1], [['fault', 'retry'], 1], [['retry', 'circuit'], 1]]
Word Count: {'North': 1, '-West': 1, 'University': 1, 'School': 1, 'for': 9, 'Computer': 2, 'Science': 2, 'and': 14, 'Information': 1, 'Systems': 1, 'ITRI': 1, '623': 1, 'Databases': 1, 'II': 1, 'Research': 1, 'a': 24, 'ssignment': 1, '(Due': 1, '30': 1, 'Sept': 1, '2024': 1, '@': 1, '24h00)': 1, '100': 2, 'marks': 3, '----------': 1, '----------': 1, '----------': 1, 'Choose': 1, 'one': 1, '(or': 1, 'two': 1, 'related)': 1, 'microservices': 1, 'software': 1, 'pattern': 6, 's': 1, 'worth': 1, 'discussing': 1, 'in': 10, 'research': 7, 'paper': 9, '.': 11,

By utilizing the document report section, you can efficiently manage and access the reports generated from your uploaded documents, allowing for a detailed review of the text analysis outcomes.

4. Disconnecting

After completing your tasks within the system, it is important to properly disconnect and close your sessions. Follow these steps to ensure a clean exit:

Disconnect from Forwarded Ports:

- In each terminal where you have forwarded ports, press Ctrl + C to stop the port forwarding. This action will disconnect your session from the respective service.

Logout from the Server:

- If you are connected to a remote server via SSH, type the following command to log out:

- This will terminate your SSH session and return you to your local terminal.

Close Terminal Windows:

- After disconnecting from the server and stopping port forwarding, close any terminal windows that you no longer need. You can do this by clicking the close button on each terminal window or using the keyboard shortcut Alt + F4.

Close the Browser:

- If you have opened a web browser to interact with the system, close the browser window or tab after you have finished your tasks.

Troubleshooting

If you encounter any issues while using the system, refer to the following troubleshooting steps to resolve common problems.

A. Connection Issues

Problem: Unable to connect to the server via SSH.

- **Solution:**
    - Ensure that you are using the correct username and server IP address.
    - Check your network connection and ensure that it is stable.
    - Ensure you are connected to the organization network. The system can only be accessed via the NWUWIFI network or a physical LAN port on campus. If you are not connected to one of these networks, you will be unable to access the system.
    - Contact administration to verify that the server is running and accessible.

B. Port Forwarding Failures

Problem: Port forwarding commands do not work.

- **Solution**:
    - Ensure you are connected to the server before running the port forwarding commands.
    - Check for any typos in the command and correct them.
    - Make sure that the specified ports are not in use by another application on your local machine.

C. System Access Problems

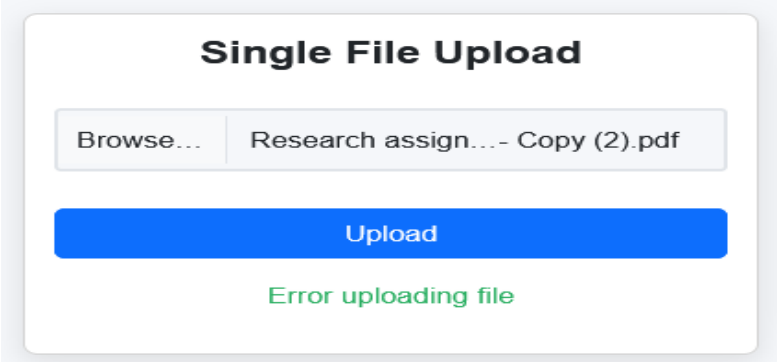
Problem: Unable to access the application via the forwarded port in the browser.

- **Solution:**
    - Verify that the port forwarding command was executed successfully and that it remains active. Check that you are using the correct port number as indicated during the port forwarding process.
    - Ensure that your browser's address bar is correctly formatted, using http://localhost:<port-number> where <port-number> corresponds to the forwarded port.

- Check for any firewall settings on your local machine that may be blocking access to the forwarded ports. Temporarily disable the firewall to see if it resolves the issue.
- If you receive a timeout or connection error, refresh the browser page and try again. If the issue persists, consider restarting the browser or clearing the browser cache.

D. File Upload Issues

Problem: Files are not uploading successfully.



- **Solution**:
  - Verify that the port forwarding and connection are still active.
  - Wait for the confirmation screen before closing the upload window, as premature closure may result in an incomplete upload.

E. Report Generation Problems

Problem: Reports are not being generated or do not contain all the required results.

Solution:

- Confirm that the uploaded file types are supported (.pdf or.docx).
- Ensure that the documents have been uploaded successfully and are being processed.
- Check the feedback bucket for any generated reports. If reports are missing, verify that the documents meet the criteria for analysis.
- If a file fails to generate a report or if the results were not successfully added, the document may contain components that are not supported by the application or may be saved in an unsupported file type. Please remove any abnormal components from the document and try again.

F. System Performance Issues

Problem: The system is slow or unresponsive.

- **Solution**:
    - Refresh the browser page to see if the issue is resolved.
    - Close any unnecessary browser tabs or applications running in the background to free up system resources.
    - Check for any large files being processed, as this may impact performance.

G. General Help

If the issue persists after following these troubleshooting steps, consult the technical documentation for additional guidance or reach out to the support team for assistance.