

Title: BotFolio

1. Introduction

Background

The process of creating a professional portfolio website manually from a resume is often time-consuming and requires technical expertise. Many users struggle with structuring their resumes into an appealing digital format, leading to inefficiencies in showcasing their professional credentials.

Objective

This project aims to automate the process of converting a resume (PDF) into a structured portfolio website using AI-driven resume parsing, cloud storage, and an agentic workflow to enable real-time modifications. The workflow will ensure seamless user onboarding, resume processing, and continuous integration and deployment (CI/CD) to GitHub for portfolio generation.

2. Project Overview

Scope

The project encompasses:

- User authentication

- Resume upload and parsing to extract unstructured content.
- Storing resume data in GCP and Snowflake.
- Leveraging an AI agent (Resume Agent) for modifications and enhancements.
- Automated CI/CD pipeline for portfolio creation using GitHub and Hugo templates.
- Managing workflows using Apache Airflow.

Stakeholders

- **End Users:** Individuals looking to create professional portfolio websites.
- **Developers:** Responsible for backend, frontend, and cloud integration.

3. Problem Statement

Current Challenges

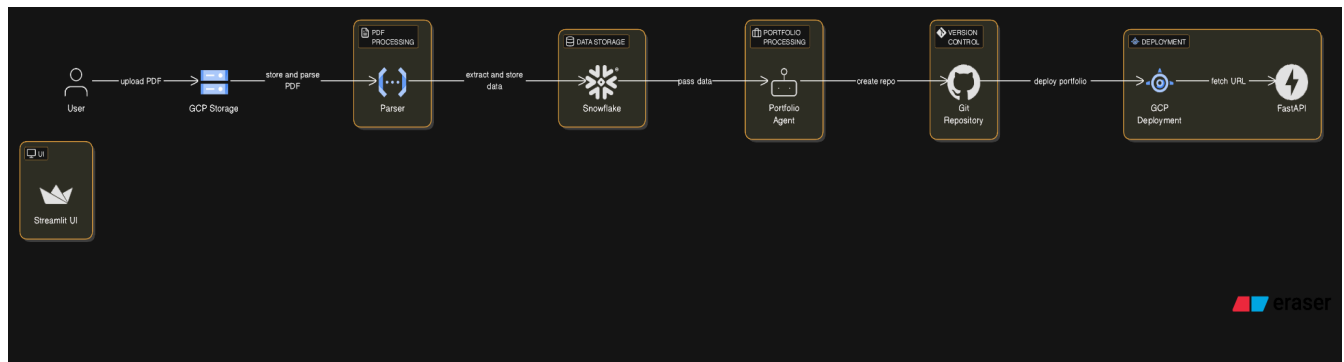
- Manual website creation from resumes is cumbersome and inefficient.
- Lack of structured data storage for resumes.

- Inability to modify the portfolio dynamically before deployment.

Opportunities

- AI-driven resume parsing enhances efficiency.
- Cloud-based storage and processing improve scalability.
- Automated deployment streamlines the website generation process

Architecture Diagram



4. Methodology

Data Sources

- User-uploaded resumes in PDF format.
- JSON output from LLM-based resume processing.

- Stored data in Snowflake for structured retrieval.

Technologies and Tools

- **Frontend:** Streamlit
- **Backend:** FastAPI
- **Cloud Storage:** GCP
- **Database:** Snowflake
- **AI Processing:** Gemini API, OpenAI API
- **Orchestration:** Apache Airflow
- **CI/CD:** GitHub Actions

Project Workflow:

1. **User Authentication:** Authentication for sign-up and login.
2. **Resume Upload:** PDFs uploaded to GCP.
3. **Parsing & AI Processing:** Extracting unstructured data and transforming it to JSON format.
4. **Storage:** JSON data stored in Snowflake.

5. **Agent-Based Modification:** Resume Agent manages edits before final deployment.
6. **Portfolio Generation:** YAML file creation and GitHub repository setup.
7. **CI/CD Execution:** Automated website deployment using GitHub Actions.

Data Processing and Transformation

- Parsing resumes into structured JSON.
- Extracting and categorizing information.
- Generating a dynamic YAML/TOML file (code generation) for portfolio setup.

5. Project Plan and Timeline

Milestones and Deliverables

Milestone	Expected Completion
Authentication & UI Setup	April 5, 2025
Resume Upload & Parsing	April 8, 2025

AI Processing & Storage	April 11, 2025
Agent Implementation	April 13, 2025
CI/CD and Deployment	April 16, 2025
Final Testing & Delivery	April 18, 2025

6. Resources and Team

Tanmay, Ankit, and Sathya will be working on this project, with the entire work distributed equally among them.

7. Risks and Mitigation Strategies

Identify Risks

- Potential failures in resume parsing accuracy.
- Deployment is halted if the generated code contains errors.

Mitigation Strategies

- Implement robust build validation checks.

8. Expected Outcomes and Benefits

Measurable Goals

- 90% accuracy in resume parsing.
- Deployment of portfolio websites within minutes.
- Seamless integration between all services.

Expected Benefits

- Saves time and effort for users.
- Ensures structured and modifiable resume-to-portfolio conversion.
- Enables automated portfolio updates and maintenance.

9. Conclusion

This project leverages AI-driven automation, cloud infrastructure, and continuous integration to transform static resumes into dynamic portfolios. By incorporating a Resume Agent and structured data storage, users can modify and refine their portfolios before final deployment. The project will serve as a

scalable and efficient solution for professionals seeking to enhance their digital presence.