AI-Powered Distributed Research Article Summarization

Team Compute-Crew:

Dheepankumar

Vaibhav Patel

Ali Hassan

WRIGHT STATE
UNIVERSITY

Target Market and Value Proposition

- **Target Market:**

  - Academics,
  - Students and
  - Professionals in various fields.

- **Value Propositions:**

  - Fast & Accurate AI Summaries.

  - Customizable User Summaries.

  - Real-time Research Processing at Scale.

  - Seamless Access Across Devices.

Conceptualization and Justification

- **Concept:**

  - Generates high-quality summaries with contextual understanding.
  - Option to edit summaries makes it interactive and personalized.
  - Distributed Architecture: Ensures fast processing and handles high traffic efficiently.

- **Justification:** Solves key industry challenges

  - Information overload.
 - Efficiency.
  - Accessibility.

Market Research on Current Offerings & Limitations

- **Existing Solutions:**

  - Semantic Scholar, Connected Papers, Research Rabbit, Elicit.

- **Limitations:**

  - Lack of AI-powered summarization.
  - No user-customizable summaries.
  - Centralized processing leading to slow performance during high traffic.
 - Limited ability to extract key findings based on user preferences.

Capabilities

- **Intelligent Research Paper Discovery:** Allows **keyword-based search** (e.g., "machine learning", "blockchain")

- **Easy-to-Browse Article List:** Interactive **search bar**, article **list view**, and **summary output**

- **PDF Access & Handling:** Extracts **direct PDF links** from selected articles.

- **Automated Text Extraction:** Parses and extracts **raw text** from downloaded PDFs

- **AI-Powered Summarization:** Uses **Hugging Face BART-large-CNN** model to generate summaries

System Architecture
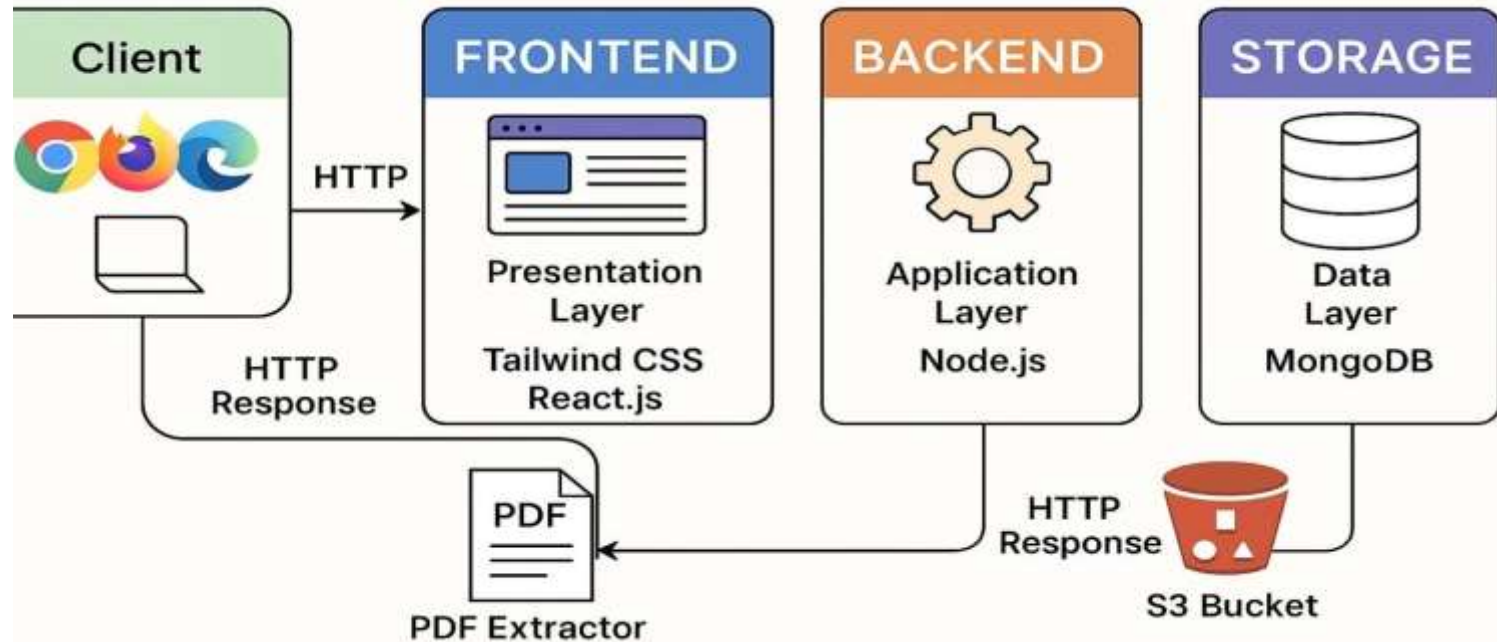
System Workflow
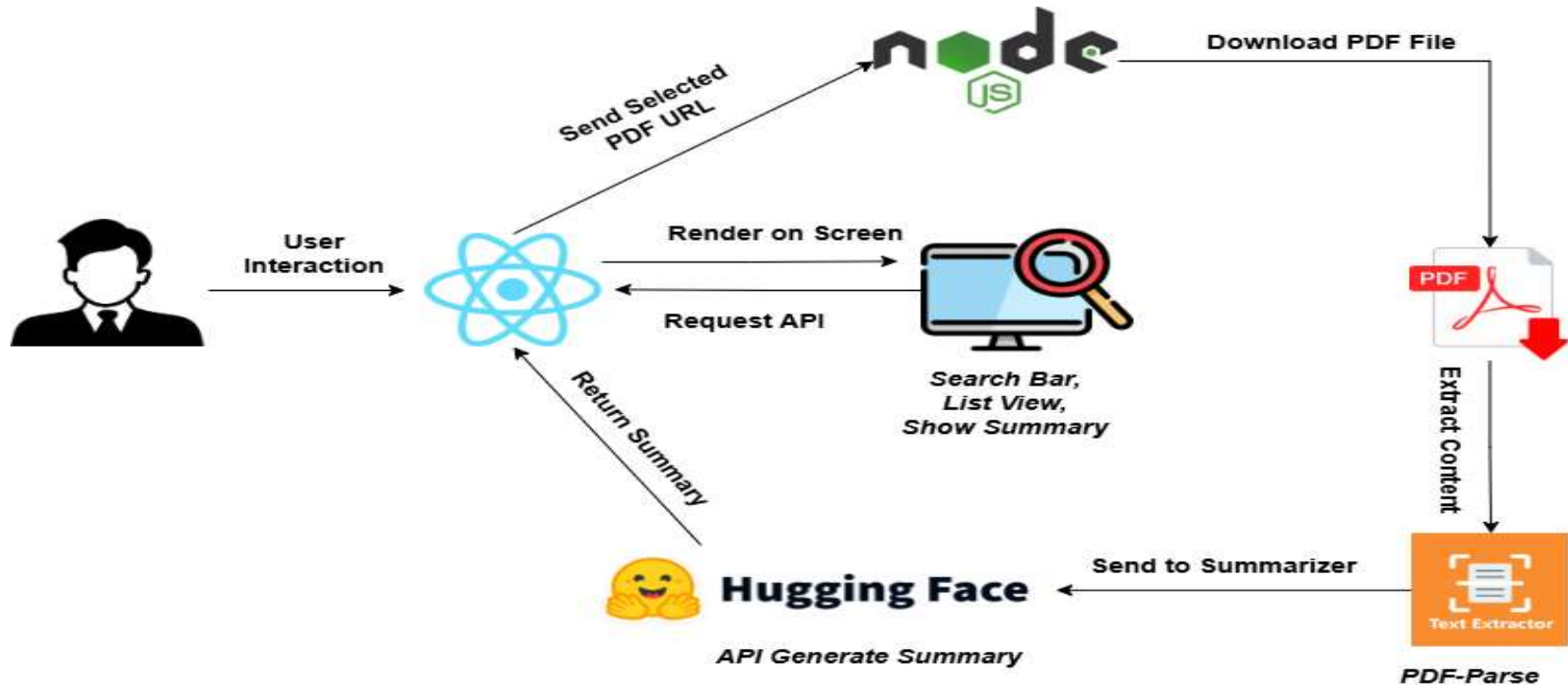
Key Winning Features

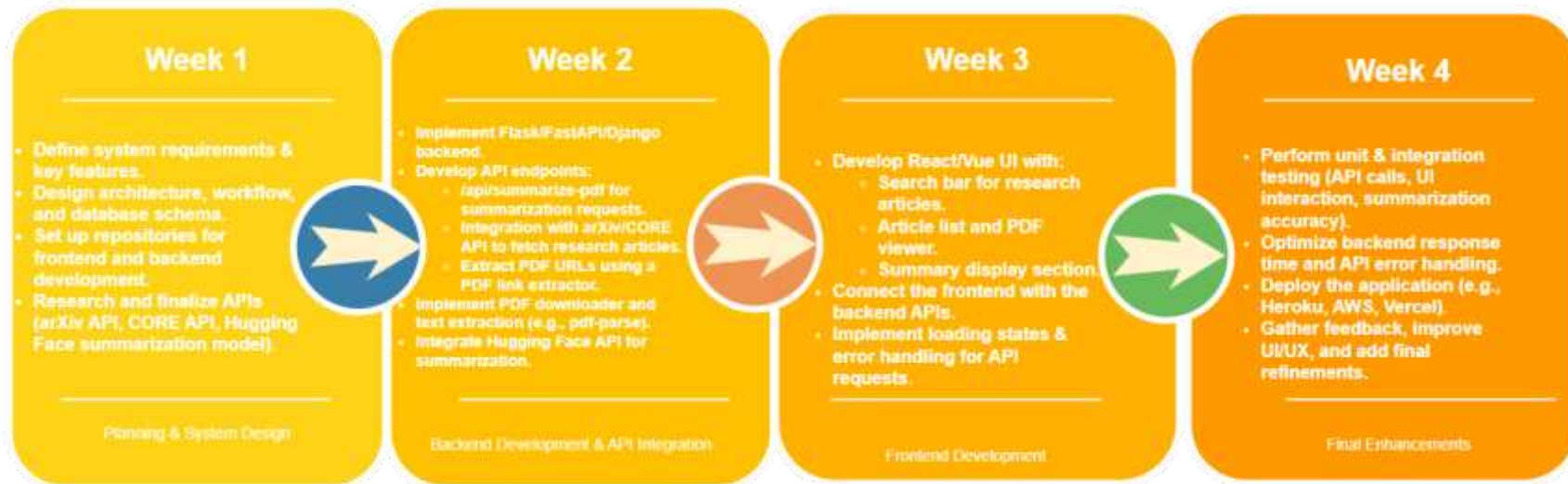| Smart Summarization:	Custom Summary Editing:	Real-time Research Processing at Scale:	High Performance with Distributed Computing:
• AI models like GPT-4, BART for high-quality contextual understanding.	• Users can refine AI-generated summaries for more personalized content.	• APIs (CORE/arXiv, Hugging Face).	• Apache Kafka (real-time data streaming). • Redis (caching for faster access). • Kubernetes & Docker (scalability & deployment).

Plan for Implementation Phase

Week 1

- Define system requirements & key features.
- Design architecture, workflow, and database schema.
- Set up repositories for frontend and backend development.
- Research and finalize APIs (arXiv API, CORE API, Hugging Face summarization model).

Planning & System Design

Week 2

- Implement Flask/FastAPI/Django backend.
- Develop API endpoints:
  - /api/summarize-pdf for summarization requests.
  - Integration with arXiv/CORE API to fetch research articles.
  - Extract PDF URLs using a PDF link extractor.
- Implement PDF downloader and text extraction (e.g., pdf-parse).
- Integrate Hugging Face API for summarization.

Backend Development & API Integration

Week 3

- Develop React/Vue UI with:
  - Search bar for research articles.
  - Article list and PDF viewer.
  - Summary display section.
- Connect the frontend with the backend APIs.
- Implement loading states & error handling for API requests.

Frontend Development

Week 4

- Perform unit & integration testing (API calls, UI interaction, summarization accuracy).
- Optimize backend response time and API error handling.
- Deploy the application (e.g., Heroku, AWS, Vercel).
- Gather feedback, improve UI/UX, and add final refinements.

Final Enhancements

Research articles

- S. Alotaibi and M. A. Alshahrani, "Literature Review of Automatic Text Summarization: Research Trend, Dataset and Method," *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1-6, 2019.

- S. Gupta and S. Gupta, "Survey on Automatic Text Summarization using NLP and Deep Learning," *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 1-7, 2022.

- Y. Zhang, Y. Li, and X. Wang, "Research and Application of Automatic Text Summarization," *2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 1-5, 2022.

THANK YOU!

WRIGHT STATE
UNIVERSITY