AI-Powered Distributed Research Article Summarization

Team Compute-Crew:
Dheepankumar
Vaibhav Patel
Ali Hassan

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.

The pain we are addressing

- **Information Overload:** Too many research papers, difficult to extract key insights.
- **Time Constraints:** Researchers lack time to read entire papers.
- **Lack of Personalization:** Generic abstracts do not always highlight the most relevant information.
- **Need for Distributed Computing:** Handling large-scale research data efficiently.

How we fix the pain

- **AI (NLP):**
  - Using NLP models like GPT-4, BART for generating concise research article summaries.
  - Custom Summary Option: Users can create their own summaries with AI assistance.
  - Real-Time News Aggregation: Uses APIs (Semantic Scholar, arXiv, PubMed) and Web Scraping.
  - Scalable Distributed Computing: Apache Kafka (real-time data streaming), Redis (caching), Kubernetes (auto-scaling), Docker (containerization).

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

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).

Business Model

- **Freemium Model:**
  - Free access with basic features.
- **Institutional Subscriptions:**
  - Universities and research organizations can subscribe for bulk access.
- **API Monetization:**
  - Paid API access for third-party integrations (academic platforms, research databases).

Demo Link

- https://www.loom.com/share/886fa511850d482a87eafb54615e110a?sid=27239036-caf9-4c4e-9cf1-348054328355

Team members and Contribution

- **Vaibhav Patel:**
  1. Expertise: Java, Express.js and Node.js
  2. Contribution: Distributed Servers, Backend API integration.
- **Dheepankumar:**
  1. Expertise: React.Js, Next.Js, Javascript and HTML5.
  2. Contribution: Frontend Development, Frontend API Integration.
- **Ali Hassan:**
  1. Expertise: Node.js, SQL, MongoDB, React.js, Deployment
  2. Contribution: Integrate Hugging face AI Model, Database Management and Deployment.

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