



Infosys Springboard Internship-6.0

*TextMorph – Advanced Text Summarization &
Paraphrasing*

Presented by
TEAM A



Team Members



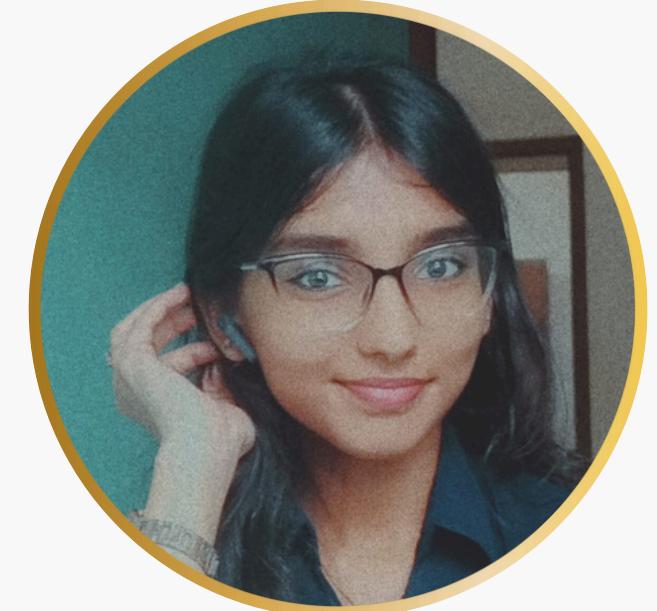
Tejo Thoshani

Member



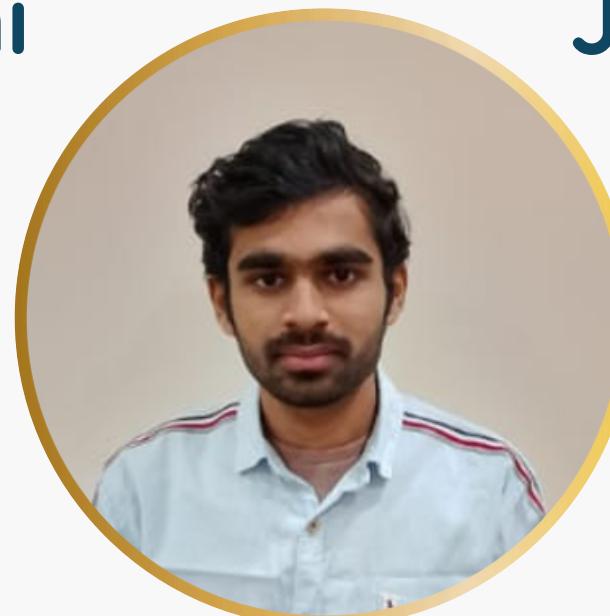
Jaagruthi Musinada

Member



Lithicka.G

Member



Vijayendra Cherupally

Member



Aakanksha Jadhav

Member



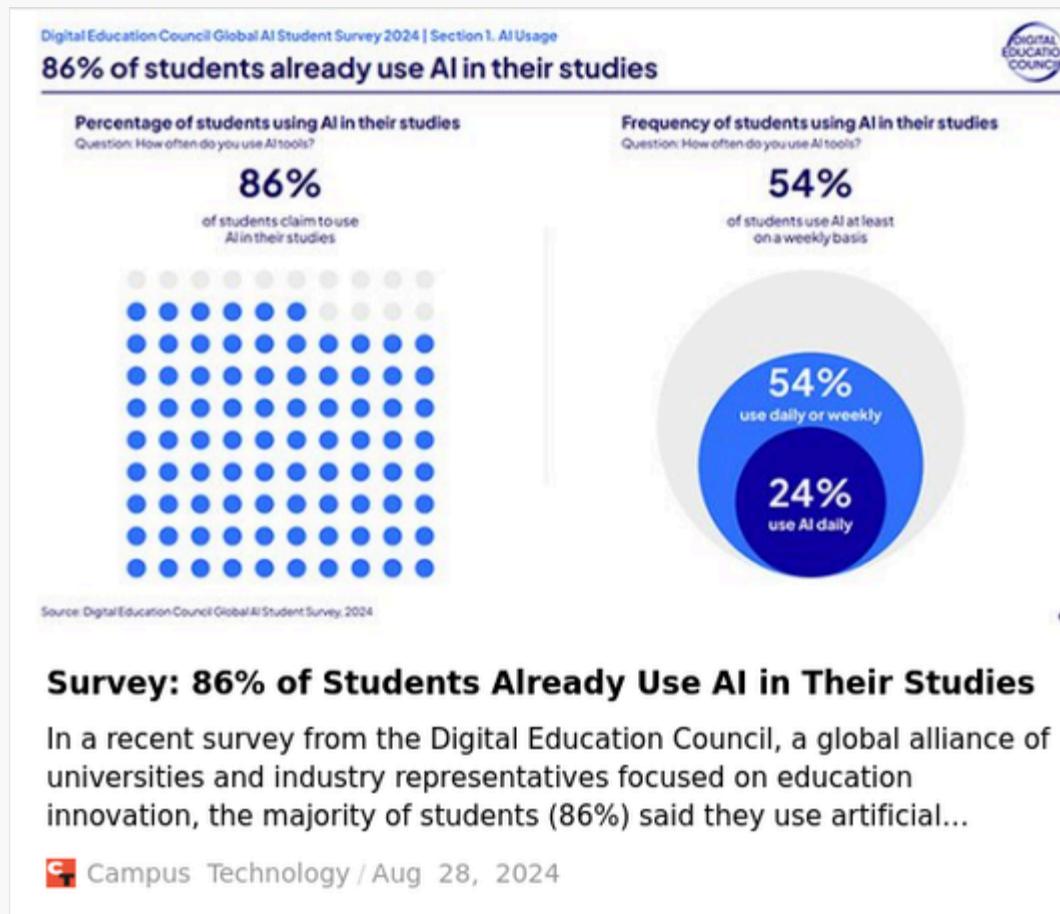
Problem Statement

Current scenario:

Students and readers struggle to understand and summarize long, complex texts

Manual paraphrasing is time-consuming and often inaccurate.

There is a growing need for an AI tool that simplifies text while preserving meaning.



Key Statistics:

- 86% of students already use AI in their studies.
- In the 2024 Global AI Student Survey (3,839 students, 16 countries), 86% said they use AI tools for coursework. Among the most common use-cases:
- 33% use AI to summarize documents
- 28% use AI to paraphrase a document



Objectives

- Develop an AI system capable of producing accurate, context-aware summaries and paraphrasing texts.
- Integrate Readability analysis to help users assess and refine text quality.
- Implement user accounts, secure storage, and a history dashboard for tracking past summaries, paraphrases, and readability metrics.
- Utilize advanced transformer models (FLAN-T5, BART, PEGASUS) combined with classical NLP tools (bigrams, similarity scoring, keyword extraction, visualizations).



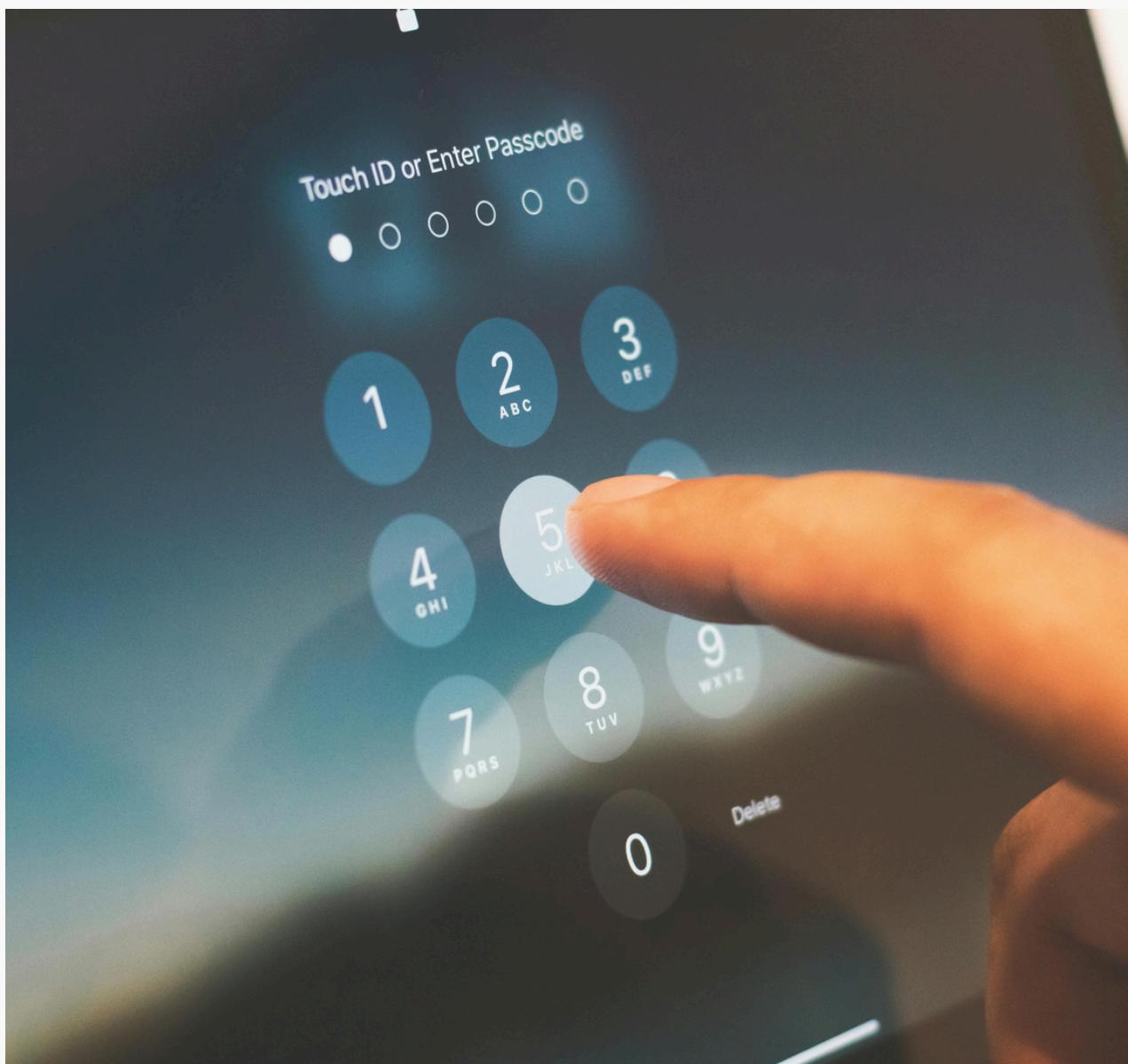


Solution Overview

TextMorph works as a complete NLP platform that includes:

- Full NLP platform with secure User/Admin authentication
- User Dashboard with:
 - 1)Advanced summarization and paraphrasing
 - 2)Readability analysis for understanding text complexity
 - 3)Data augmentation and model fine-tuning for improved performance
 - 4)Feedback system for ratings and comments
- Admin dashboard with :
 - 1)Analytics and Monitoring
 - 2)Complete activity and History tracking
 - 3>Docker Deployment

Authentication System

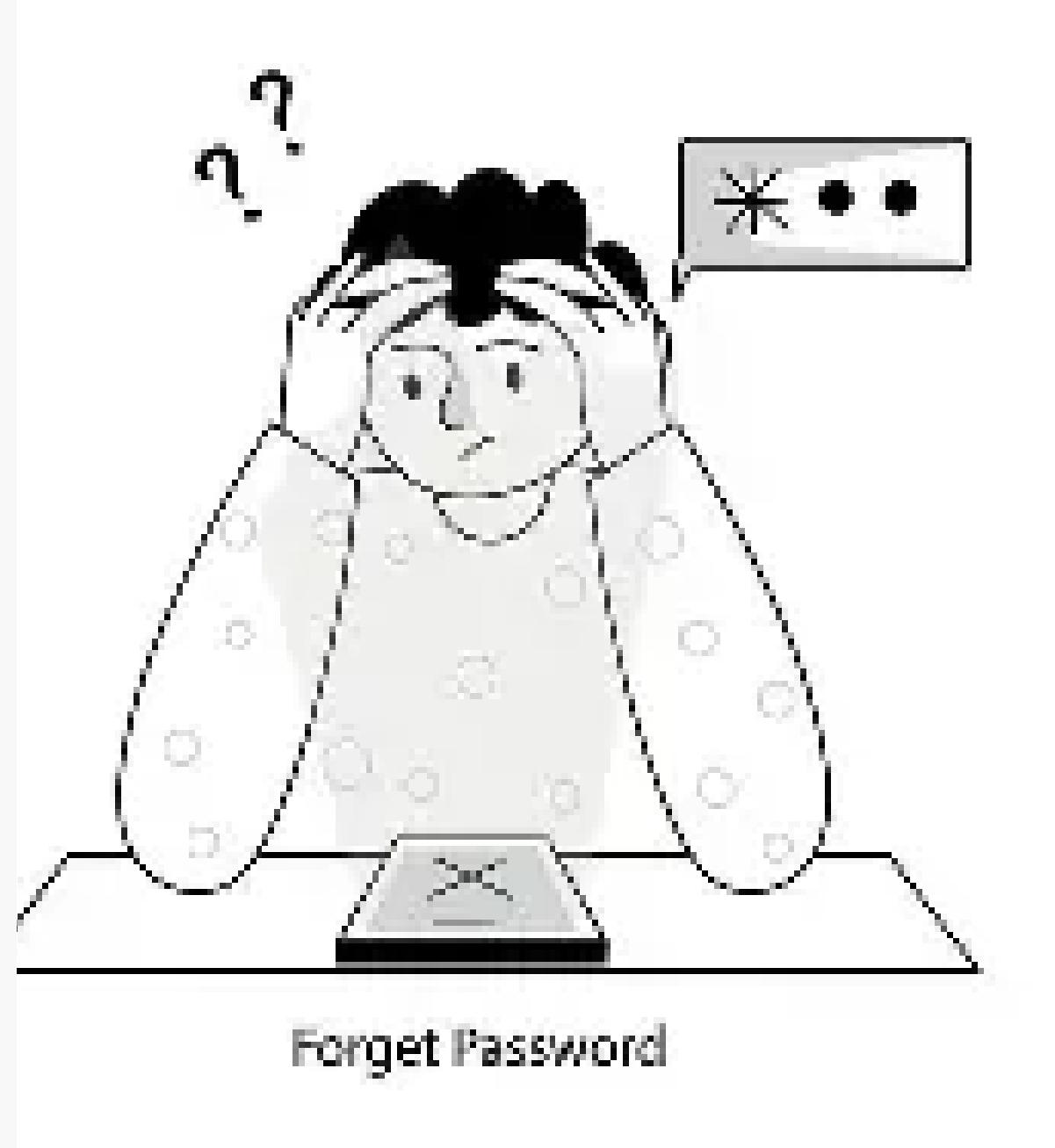


- Using JWT(JSON Web Token)-based login with role-based access control.
- Regular users can access the General user login(default), while only 2 admins are allowed to manage users and monitor analytics.
- All passwords are protected using Bcrypt Hashing (SHA256 Algorithm)

Password Recovery

Users can recover their accounts through a secure, multi-step verification process:

- A 6-digit OTP is sent to the user's registered Gmail via SMTP(Simple Mail Transfer Protocol) service
- Resetting passwords—require successful OTP verification for enhanced security.
- Security questions serve as a secondary recovery method if email access is unavailable.



Core Functionalities

Readability Analysis:

TextMorph supports text input and file uploads (.txt, .pdf).

It calculates important readability scores:

- Flesch Reading Ease – shows how easy the text is to read
- Flesch-Kincaid Grade Level – shows expected grade level
- Gunning Fog Index – measures writing complexity
- SMOG Score – shows the education level needed to understand the text

It also includes:

- Color-coded difficulty levels (easy, moderate, difficult)
- Simple graphs for comparing readability



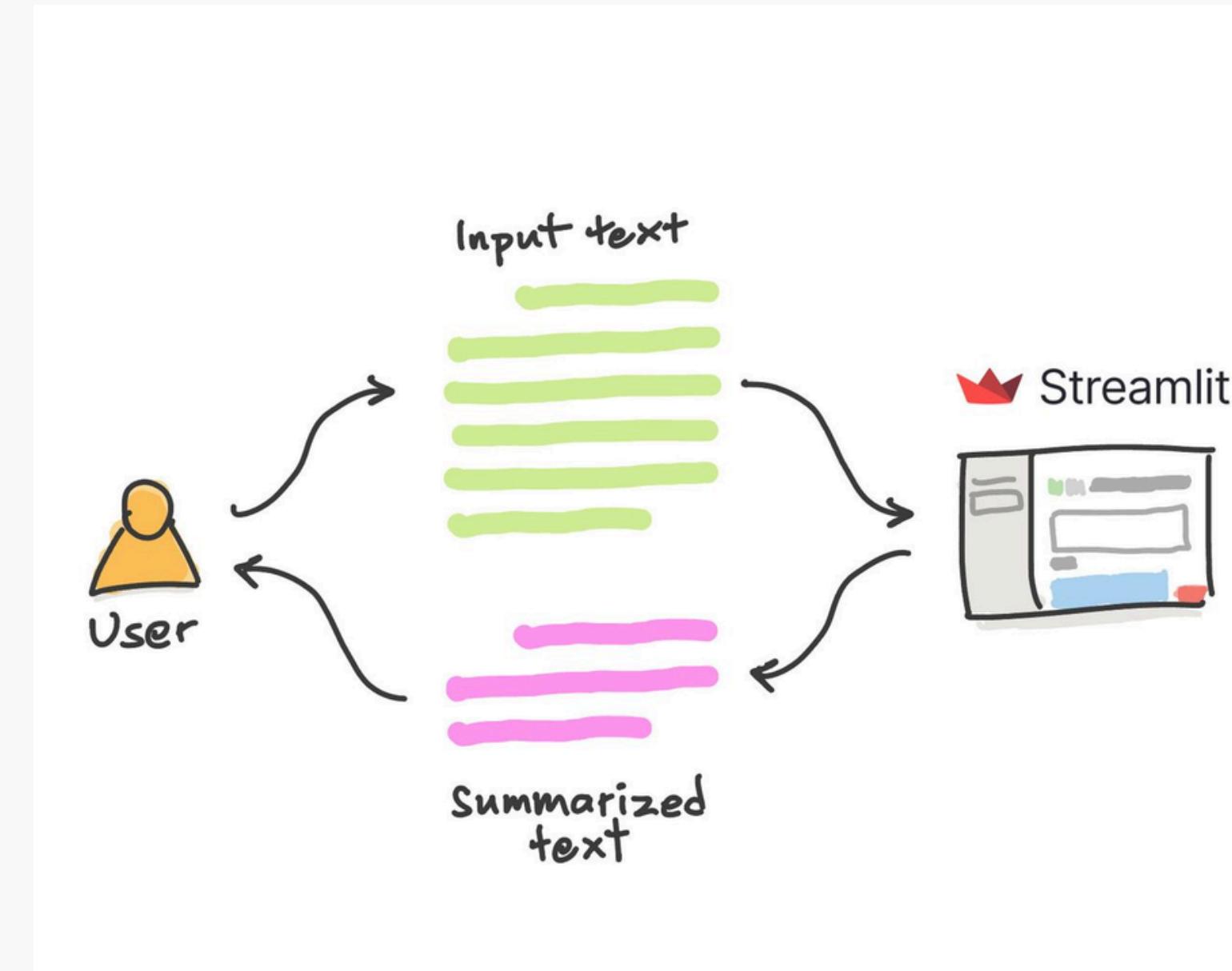
Summarization Engine

The system uses advanced transformer models for summarization:

- Flan T5 (From Google)- 250 million parameters
- BART (From META AI)- 139 million parameters
- PEGASUS (From Google) - 223 million parameters

We can modify:

- Summary length: Short(50-150), Medium(100-250), or Long(150-300)
- Readability metrics: ROUGE-1, ROUGE-2, ROUGE-L



Paraphrasing Engine

What is Paraphrasing?

Paraphrasing means to tell the meaning of something spoken or written using your own words.

My Robot. Its nose is a hover. Its belly is a computer. Its mouth is a pencil sharpener.

A smart tool that rewrites text clearly at both sentence and paragraph levels

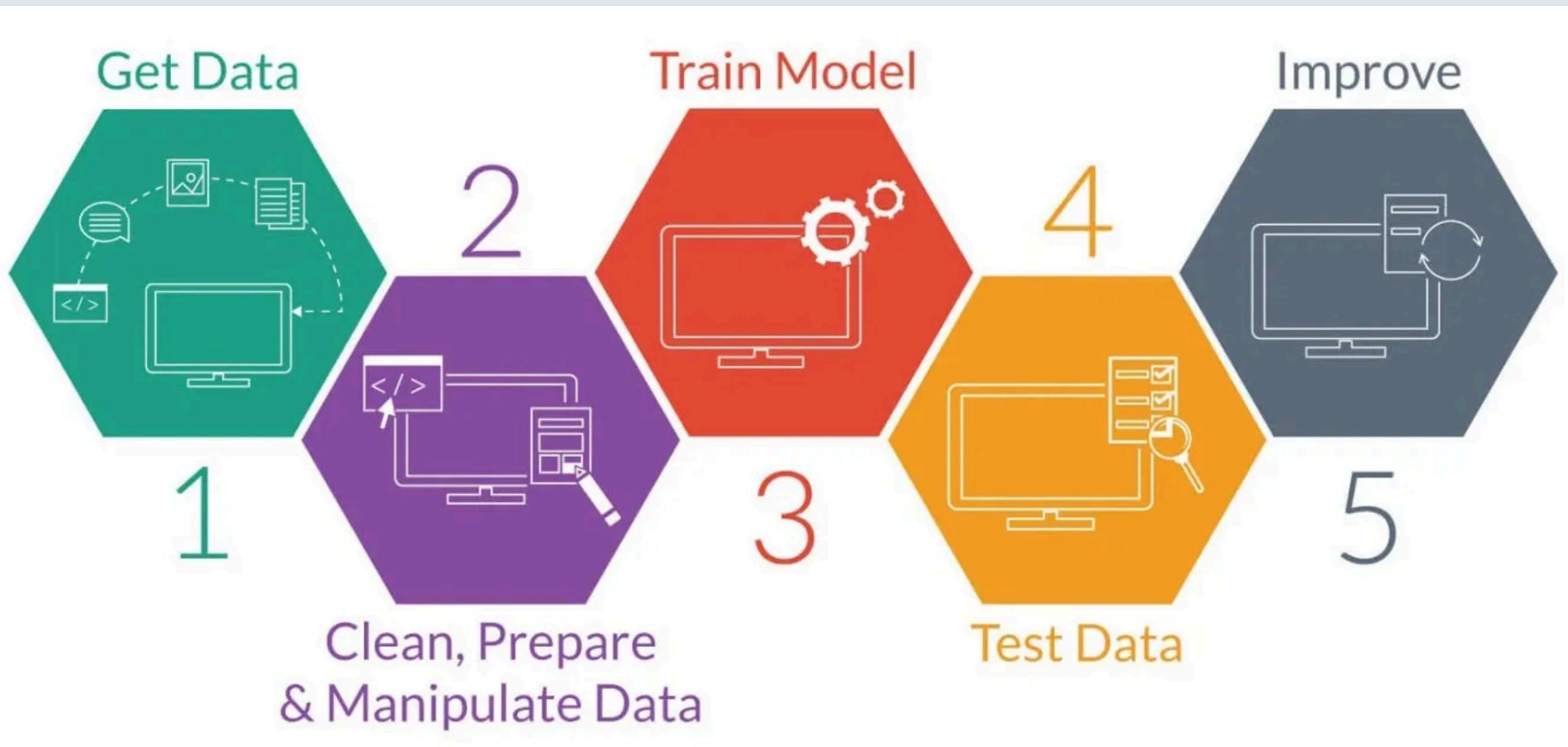
Model options:

- Flan T5 (From Google)- 250 million parameters
- BART (From META AI)- 139 million parameters

Target complexity: Beginner, Intermediate, Advanced, Expert

Paraphrasing Styles : Simplification, Formalization, Neutral

Data Augmentation and custom model training



TextMorph uses a diverse set of domain-specific and simplification datasets:

- Academic: ccdv/arxiv-summarization
- News: cnn_dailymail
- Medical: ccdv/pubmed-summarization
- Legal: billsum

Custom Trained model: Flan-T5

User feedback module

Key Features:

- Rating System: (1-5) star rating 
- Issue Selection: Tick boxes to report problems (like missing info or grammar mistakes).
- Comment Box: Extra suggestions or details about issues.
- Model Choice: Choose model (BART, Pegasus, T5) which made the output.
- Feedback History: All feedback is saved and shown in a table.
- Dashboard Metrics: Shows average rating, total feedback, and the most common problem.

Rate your experience ×

How's your overall experience today?



Describe your issue

We'd love to hear your suggestions

I would like to be contacted about my issue.

Submit

Admin Dashboard

Admin Controls

- **Manage Accounts:** Admins handle all user accounts.
- **Strict Control:** Only 2 admins allowed.
- **Transparency:** Every action is logged.
- **User Access:** Admins can add , remove users and promote users to admin.

Login to Admin Panel



Analytics and Insights

- Feature Usage: Shows if summary or paraphrase is used more
- Model Popularity: Highlights top models (T5, BART, PEGASUS)
- Active Users: Lists the most engaged users
- Ratings: Measure how well models perform
- Feedback Word Clouds: Show common user comments
- Future Quality: Insights help improve updates



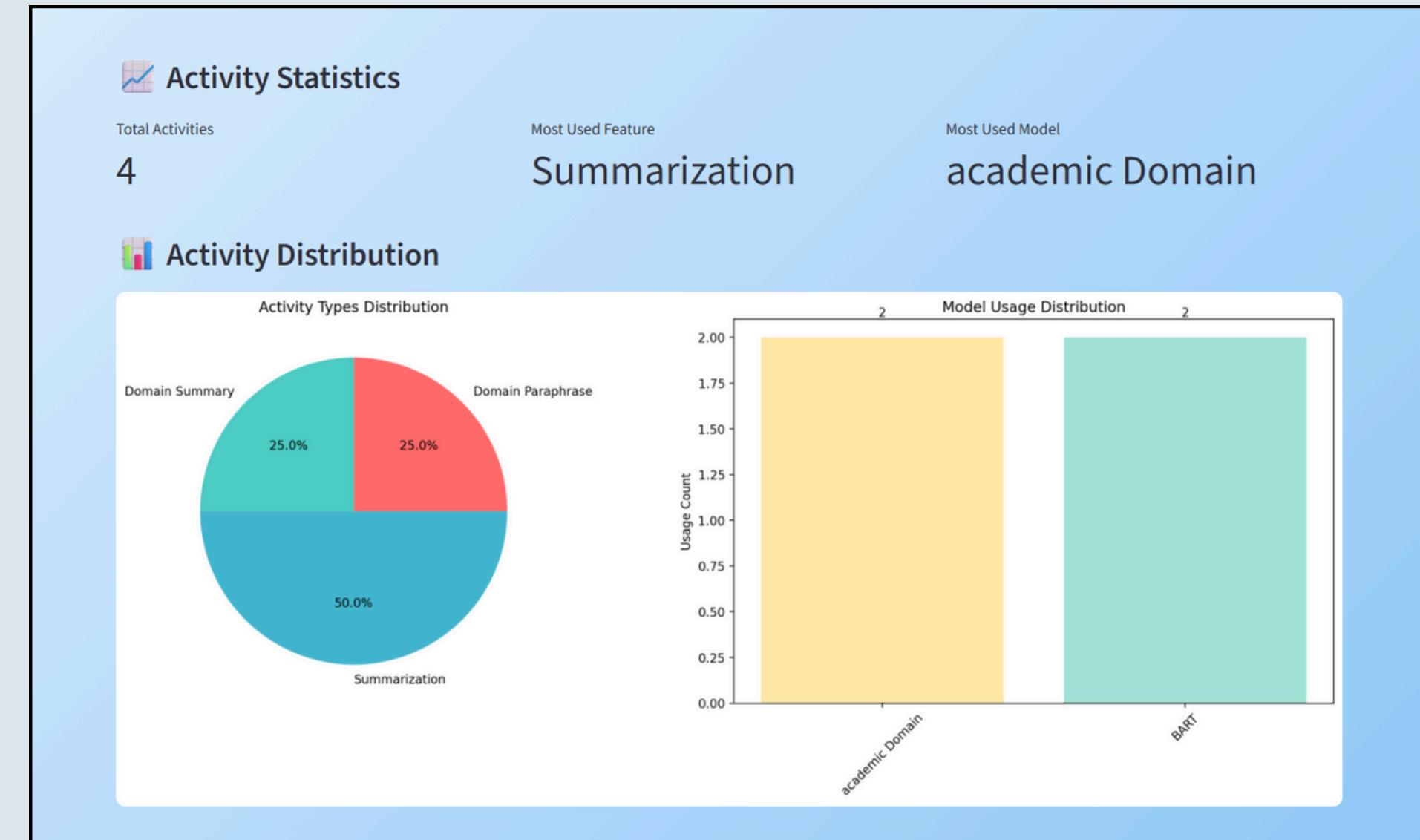
A word cloud visualization composed of various words in different colors, including blue, green, red, orange, yellow, and purple. The words are arranged in a roughly circular pattern. Some words are oriented vertically along the right edge. The colors represent different categories, as indicated by small colored squares next to some words.

cool presentations
live thoughts
fun easy inclusive share software
beautiful reflection
interactive brainstorm knowledge
ideas ice breaker

exciting anonymous

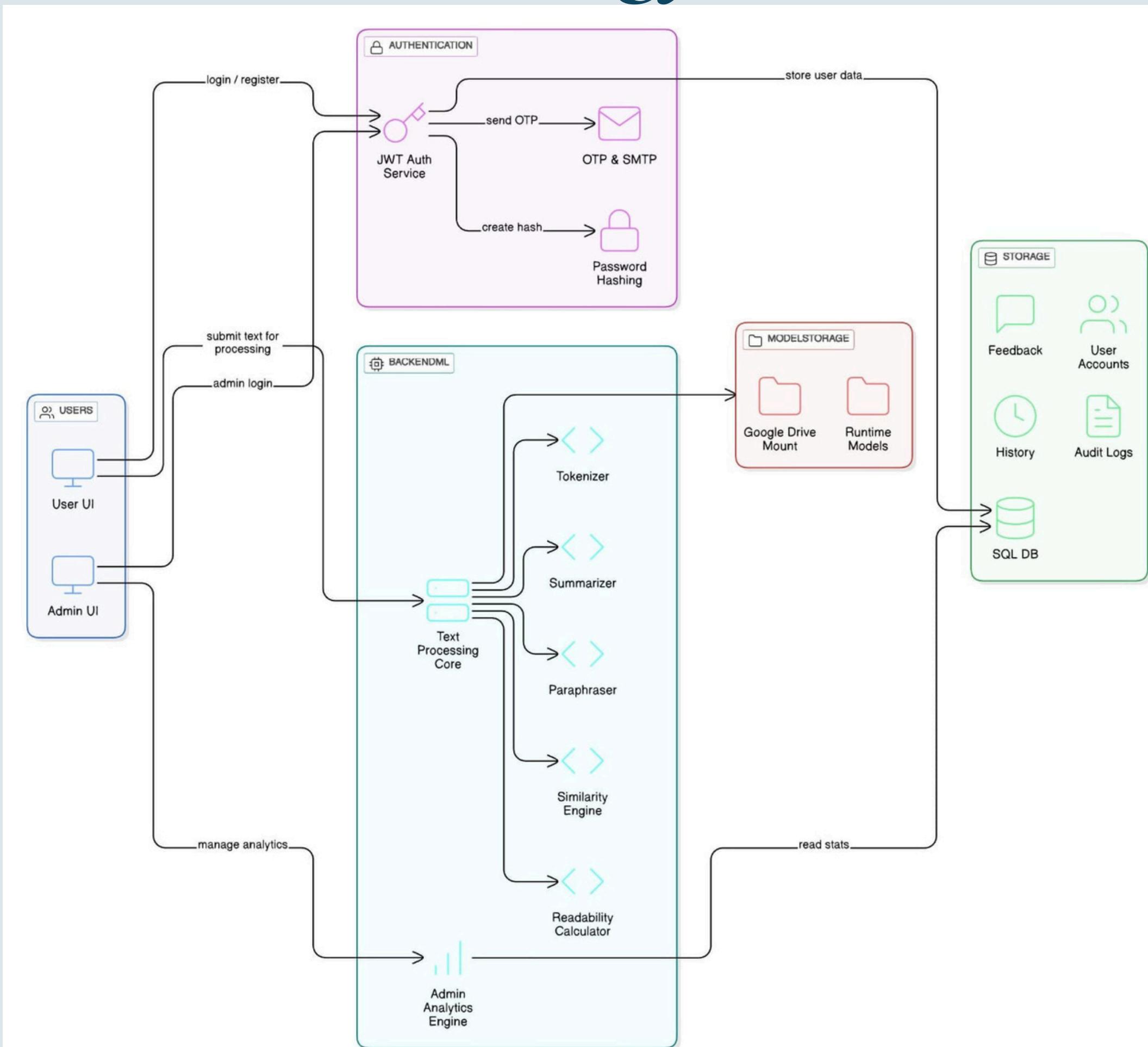
Global Activity Monitoring

- Complete Activity Log
 - Detailed Activity Information
- For each activity:
 - Timestamp
 - Activity Type (Summarization/Paraphrasing /Domain)
 - Model Metrics (FLAN-T5/BART/Pegasus)
 - Input Preview & Output Preview
 - Detailed Metrics (word counts, scores, settings)



Architecture & Technology

Architecture Diagram :



Tech Stack

Frontend:



Streamlit

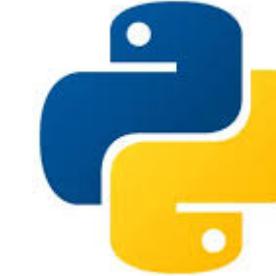
Databases:



Authentication:



Backend:



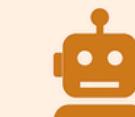
AI Models:



PEGASUS



FLAN-T5



BART



Deployment:



Conclusion

Why TextMorph Matters??

- The Cloud AI Problem is Real - Professionals worldwide face daily frustrations with blocked access, slow loading, and privacy concerns when relying on external AI services.
- We Built a Better Solution - TextMorph eliminates these pain points by providing a self-contained, locally-hosted platform that guarantees instant access, complete data privacy, and uninterrupted service.
- Enterprise-Grade, Accessible to All - With multiple AI models, domain-specific processing, and robust security, we deliver enterprise capabilities without enterprise complexity or costs.
- The Future of AI is Local - TextMorph represents the next evolution: powerful AI tools that work for you, not against you - ensuring reliability, privacy, and control where it matters most.



Thank you

Any Questions?

