

- GitHub: https://github.com/danielnashed/ReSearchGraph_1
- Website: <https://project2.d269jq01wtg7ba.amplifyapp.com/>

AI RESEARCH TRENDS DASHBOARD

Dashboard to Visualize Latest AI Trends

TABLE OF CONTENTS

1	Problem Statement
2	Goals
3	Demo
4	System Overview
5	Future Improvements

PROBLEM STATEMENT

AI Research Field is Exploding

- More than 200 papers are released every single day in the field of GenAI
- Difficult to keep up with the latest trends in the field

GOALS

- Automate the retrieval and clustering of academic papers in Generative AI
 - ◆ Fetch & process daily stream of academic papers from arXiv.org
 - ◆ Group papers into existing clusters or new clusters automatically
 - ◆ Users can interactively explore research trends through a frontend interface.

DEMO

Website: <https://project2.d269jq01wtg7ba.amplifyapp.com/>

Login / Signup

Login

Signup

🔄 AI Research Trends

Start Retriever

Logout

🔄 AI Research Trends

Stop Retriever

Logout

🔄 AI Research Trends

**Multimodal Alignment
and Reasoning in
Vision Language
Models**
7 papers

**Theoretical Physics
and Cosmology**
5 papers

**Human-like AI
Interactions and
Content Generation**
4 papers

**Large Language Model
(LLM) Robustness and
Safety**
17 papers

Multimodal Alignment
and Reasoning in
Vision Language
Models
7 papers

Theoretical Physics
and Cosmology
3 papers

Human-like AI
Interactions and
Content Generation
4 papers

Large Language Model
(LLM) Robustness and
Safety
11 papers

Multimodal Alignment and Reasoning in Vision Language Models

This collection of papers focuses on advancing the field of multimodal alignment and reasoning in Vision Language Models (VLMs). The papers present various methods for aligning VLMs, including Re-Align, which leverages image retrieval to construct a dual-preference dataset, and rDPO, an extension of direct preference optimization that incorporates visual preference signals. Other papers explore the use of natural language to represent orientation in robotic systems, pre-training auto-regressive robotic models with 4D representations, and developing foundation models for multimodal AI agents. Additionally, the papers discuss the importance of multimodal perception and the challenges of fine-tuning large multimodal models for hateful meme detection. Overall, the collection highlights the need for more robust and effective multimodal alignment and reasoning techniques in VLMs, and presents various approaches to addressing these challenges.

- Re-Align: Aligning Vision Language Models via Retrieval-Augmented Direct Preference Optimization
- SoFar: Language-Grounded Orientation Bridges Spatial Reasoning and Object Manipulation
- Pre-training Auto-regressive Robotic Models with 4D Representations
- Magma: A Foundation Model for Multimodal AI Agents
- L4P: Low-Level 4D Vision Perception Unified
- Improved Fine-Tuning of Large Multimodal Models for Hateful Meme Detection
- SimpleVQA: Multimodal Factuality Evaluation for Multimodal Large Language Models

Multimodal Alignment
and Reasoning in
Vision Language
Models
7 papers

Theoretical Physics
and Cosmology
2 papers

Human-like AI
Interactions and
Content Generation
4 papers

Large Language Model
(LLM) Robustness and
Safety
12 papers

Human-like AI Interactions and Content Generation

This collection of papers focuses on the development of human-like AI interactions and content generation. The first paper, AV-Flow, presents a model that generates photo-realistic 4D talking avatars from text input, with synchronized speech, lip motion, and facial expressions. The second paper, Personalized Image Generation with Deep Generative Models, provides a decade-long survey of personalized image generation using deep generative models, including traditional GANs, text-to-image diffusion models, and multi-model autoregressive models. The third paper, Natural Language Generation from Visual Sequences, discusses the challenges and future directions of generating natural language from visual sequences, including multi-image vision-to-text settings. The fourth paper, Whose story is it? Personalizing story generation by inferring author styles, proposes a novel pipeline for personalized story generation by inferring author styles from their past work. Overall, these papers demonstrate advancements in human-like AI interactions and content generation, including personalized image and story generation, and highlight the importance of understanding intricate relationships between visual content and corresponding text.

- [AV-Flow: Transforming Text to Audio-Visual Human-like Interactions](#)
- [Personalized Image Generation with Deep Generative Models: A Decade Survey](#)
- [Natural Language Generation from Visual Sequences: Challenges and Future Directions](#)
- [Whose story is it? Personalizing story generation by inferring author styles](#)

Users



POST **/users/signup** Signup Route



POST **/users/login** Login Route



GET **/users/{user_id}** Get User Route



DELETE **/users/{user_id}** Delete User Route



Clusters



GET **/clusters/{user_id}** Get Clusters Route



Scheduler



POST **/scheduler/{user_id}** Get Clusters Route



Papers

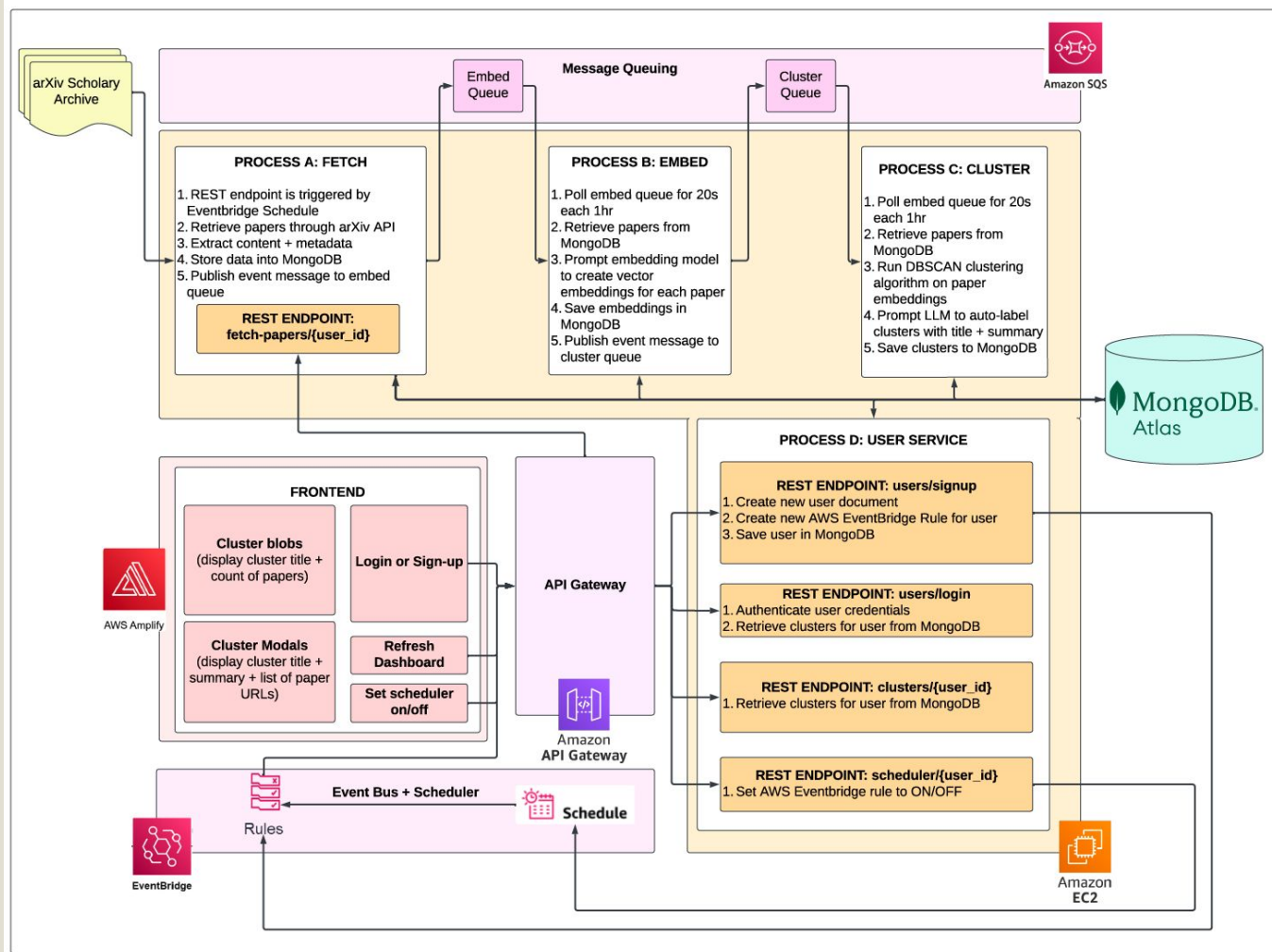


POST **/fetch-papers** Create Papers Route



SYSTEM OVERVIEW

SYSTEM DIAGRAM



TECH STACK

→ Backend

- ◆ **FastAPI** for web framework (python web framework)
- ◆ **MongoDB** for NoSQL database to store paper embeddings and metadata
- ◆ **AWS CloudFormation** as infra-as-code to provision cloud resources
- ◆ **AWS API Gateway** to manage API requests efficiently
- ◆ **AWS CloudWatch** for monitoring and logging system
- ◆ **AWS EventBridge** to schedule daily retrieval tasks
- ◆ **AWS SQS** to manage processing queues to handle high-load data streams
- ◆ **AWS ECS** for container orchestration
- ◆ **Docker** to build container images & **AWS ECR** to store images

→ Frontend

- ◆ **Next.js**: React framework optimized for SEO and fast rendering
- ◆ **TailwindCSS**: Lightweight CSS framework for styling.
- ◆ **daisyUI**: Prebuilt UI components for a modern user experience
- ◆ **AWS Amplify** for hosting frontend

→ AI/ML

- ◆ **AWS Bedrock**: Generates vector embeddings for research papers
- ◆ **Incremental-DBSCAN**: Dynamic clustering algorithm for research papers

KEY COMPONENTS

→ Fetching papers

- ◆ Retrieve research papers from the **arXiv API**
- ◆ Extract text and metadata & save in MongoDB

→ Embeddings

- ◆ Convert title + abstract of each paper into vector embeddings using **Cohere Embed** model
- ◆ Store embeddings in Paper Document inside MongoDB

→ Clustering

- ◆ Apply dynamic + incremental clustering algorithm **INCREMENTAL-DBSCAN** to group embeddings based on topic similarity.
 - Incremental aspect enables batch processing of papers. New batches either clustered into existing clusters or inserted into new clusters based on similarity threshold.
- ◆ Use **Llama 3.3 Instruct (70B)** to auto-label clusters + generate summary for each cluster

→ User Management

- ◆ User sign-up and login/authentication
- ◆ Creation of unique EventBridge rule + schedule for each user
- ◆ Toggling scheduler on/off for each user

**FUTURE
IMPROVEMENTS**

FUTURE IMPROVEMENT

- **Expand Data Sources:** fetch papers from Semantic Scholar and Google Scholar
- **Hybrid Clustering:** allow users to control initial clusters and labels for more fine-grained control over clustering
- **Context Management:** implement memory model to track important research trends over time
- **Graph Database:** build a graph using Neo4j to capture semantic relationships between papers (authors names, institution names, journal names, publication year, # of citations, # of references)