Master Project — Multi-Agent Climate System

Abstract

The Multi-Agent Climate System is an advanced platform for collecting, processing, analyzing, and visualizing climate narratives using modern NLP techniques, transformer models, and a multi-agent architecture.

Motivation and Impact

The system enables the understanding of opinions, emotions, themes, and contradictions within global narratives about climate change.

General System Architecture

It includes a Python backend with FastAPI, MongoDB as the database, FAISS as the vector store, orchestration with APScheduler and Prefect, NLP processing with spaCy/sentence-transformers, modeling with BERTopic and BERT, and an interactive Streamlit dashboard.

Complete Data Flow

1. Collection → RAW

2. Preprocessing → CLEAN

3. Modeling → CURATED

4. Multi-Agent → Insights

5. Visualization → User

Phase and Submodule Description

PHASE 1 — Collection

- Scraping, APIs, RSS, storage.

PHASE 2 — Preprocessing

- Cleaning, normalization, embeddings.

PHASE 3 — NLP Modeling

- Classification, topic modeling, transformers.

PHASE 4 — Multi-Agent System

- Agents: collector, analytical, clustering, summarization, ethical, contradiction, historical, explorer.

PHASE 5 — Visualization

- Dashboard with narratives, topics, emotions, semantic networks.

Analysis API

FastAPI provides endpoints for queries, summaries, clustering, and access to the vector database.

Dashboard

Built in Streamlit with interactive charts, timelines, word clouds, and narrative maps.

Database

MongoDB structured in RAW, CLEAN, and CURATED layers. FAISS for semantic searches.

Implementation Plan

1. Create repository structure

2. Implement scraping

3. Embeddings and cleaning

4. Train models

5. Create agents

6. Integrate FAISS

7. Build FastAPI

8. Design dashboard

9. Document

Risks and Mitigations

• API rate limits — use of a scheduler

• Model biases — ethical agent

• Data volume — use of FAISS and MongoDB