



Enhancing PPM with ML and LLMs to Further Assist Transfer Students

Faculty Advisors: Dr. Hui Yang, Dr. Arno Puder

Student Researchers: Mark Kim (graduated), Jay Lodha, Shreyas Raghuraman,
Divya Panchal, Dhvanil Bhagat

Department of Computer Science
San Francisco State University

Proposed tasks, Fall 2023.

LLM-driven Machine Learning for Transfer Student Advising

Major Tasks	Data Required	Candidate ML Techniques
Course equivalence evaluation	<ul style="list-style-type: none">• California ASSIT• PPM data• Description & syllabi of courses from different institutions	<ul style="list-style-type: none">• Information extraction (e.g., learning objectives)• Topic modeling (e.g., list of topics)• Semantics-based similarity ranking using LLM embeddings• LLM specific<ul style="list-style-type: none">◦ Prompt engineering◦ In-context learning◦ Fine-tuning LLMs
Multi-year pathway generation	<ul style="list-style-type: none">• The above data• Recommended curriculum roadmaps from different institutions• Course-taking data	<ul style="list-style-type: none">• All the above for course equivalence evaluation• Term-by-term pathway generation using LLMs
chatbot Engine for students	<ul style="list-style-type: none">• All the above data• Commonly asked questions and answers from students• New questions and vetted answers	<ul style="list-style-type: none">• Question-intent classification (e.g., concerning course equivalence or a 3-year pathway)• Semantic search of similar and previously asked questions• Retrieval of relevant data• Response generation using LLM

Spring 2024 -

Summer 2024 -

Summer 2025 -

Course Equivalency Evaluation: Overview

Two courses: {meta info, catalog description}

↓
Assist.org (to do)

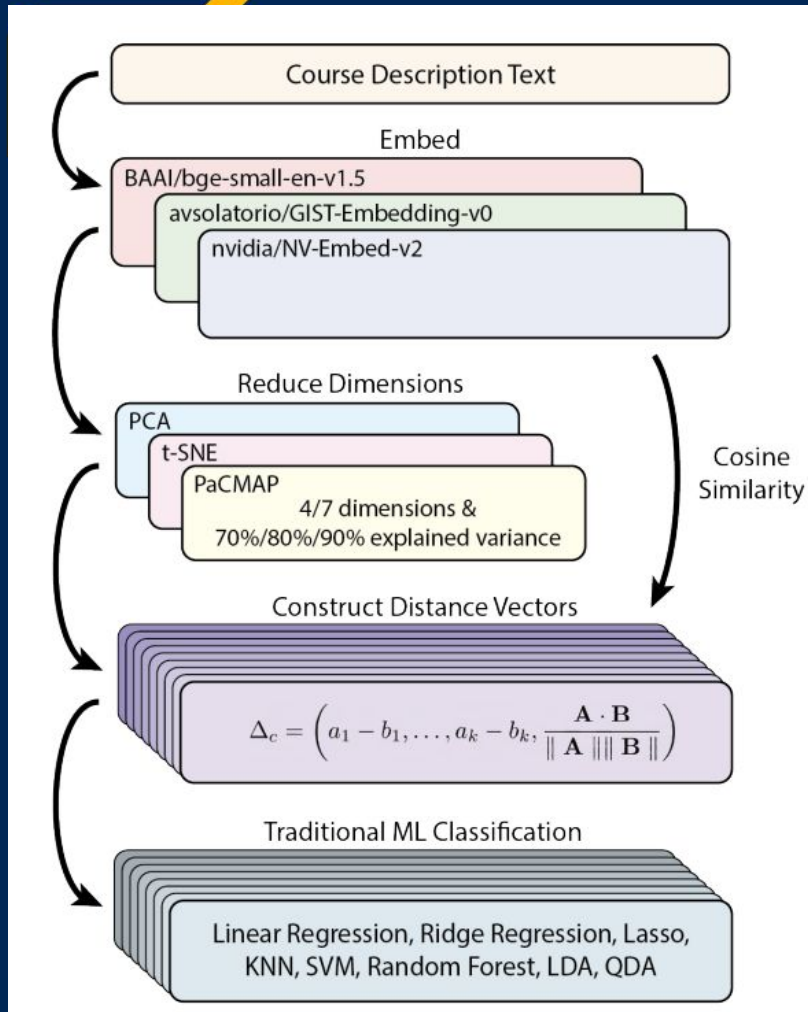
↓
C-ID Data (done)



↓
Fine-tuned LLM Embedding Model (done)
Pre-trained ML Classification Models

↓
Yes or No

Course Equivalency Evaluation (1/2): Embedding-based Ensemble Classification



- Fine-tuned a 33-million embedding model (bge-1.5)
 - Integrated PPM-CID data
 - SFSU's POLARIS HPC cluster



EDM 2025 Best Poster Paper Award



Mark Kim successfully defended his Master's Thesis in August 2025.

Course Equivalency Evaluation (2/2): Optimization, Containerization, AWS Microservices

- **Optimizations:**
 - Enable ensemble-based equivalence prediction
 - Quantize the fine-tuned embedding models
- **Containerization**
- **AWS Microservices**
 - Embedding model API
 - Ensemble-classification API
 - C-ID data lookup API

Model	Size (MB)	Mean Time(ms)	Accuracy
Original	419	241.26	97.5
ONNX	416	263.49	97.2
ONNX INT8	105	132.08	97.1
CT2 Float32	418	210.65	97.2
CT2 Float16	209	179.47	97.5
CT2 Int16	210	145.01	97.5
CT2 Int8	107	138.53	97.4

Flexible Pathway Generation

Given:

- Degree requirements & roadmap (B.Sc. in Computer Science)
- Student's input:
 - Number of terms to complete (11 terms)
 - Areas of interest



Term 1 (10u): ['CSC101', 'GE_AREA_4_1', 'MATH226']
Term 2 (11u): ['GE_AREA_1_1', 'MATH227', 'PHYS220_222']
Term 3 (11u): ['CSC215', 'GE_AREA_1A_1', 'PHYS230_232']
Term 4 (12u): ['CSC220', 'CSC230', 'GE_AREA_3_1', 'GE_AREA_4_2']
Term 5 (12u): ['CSC256', 'CSC340', 'GE_AREA_5B_1', 'MATH324']
Term 6 (9u): ['GE_AREA_1_2', 'GE_AREA_6_1', 'MATH225']
Term 7 (12u): ['CSC300GW', 'CSC317', 'CSC510', 'GE_AREA_3_2']
Term 8 (12u): ['CSC413', 'CSC415', 'GE_AREA_4UD_1',
'GE_AREA_5UD_1']
Term 9 (9u): ['CSC665', 'CSC676', 'GE_AREA_3UD_1']
Term 10 (12u): ['CSC620', 'CSC621', 'SFSTATE_GE_1', 'USCG']
Term 11 (10u): ['CSC648', 'CSC671', 'SFSTATE_GE_2']

Flexible Pathway Generation

Algorithm: MCTS → CP-SAT

Monte Carlo Tree Search (MCTS)

- Mimic a single-player game
- Balance exploration & exploitation

Challenges:

- Hyperparameter tuning: time-consuming
- Generation time: 15-20 minutes

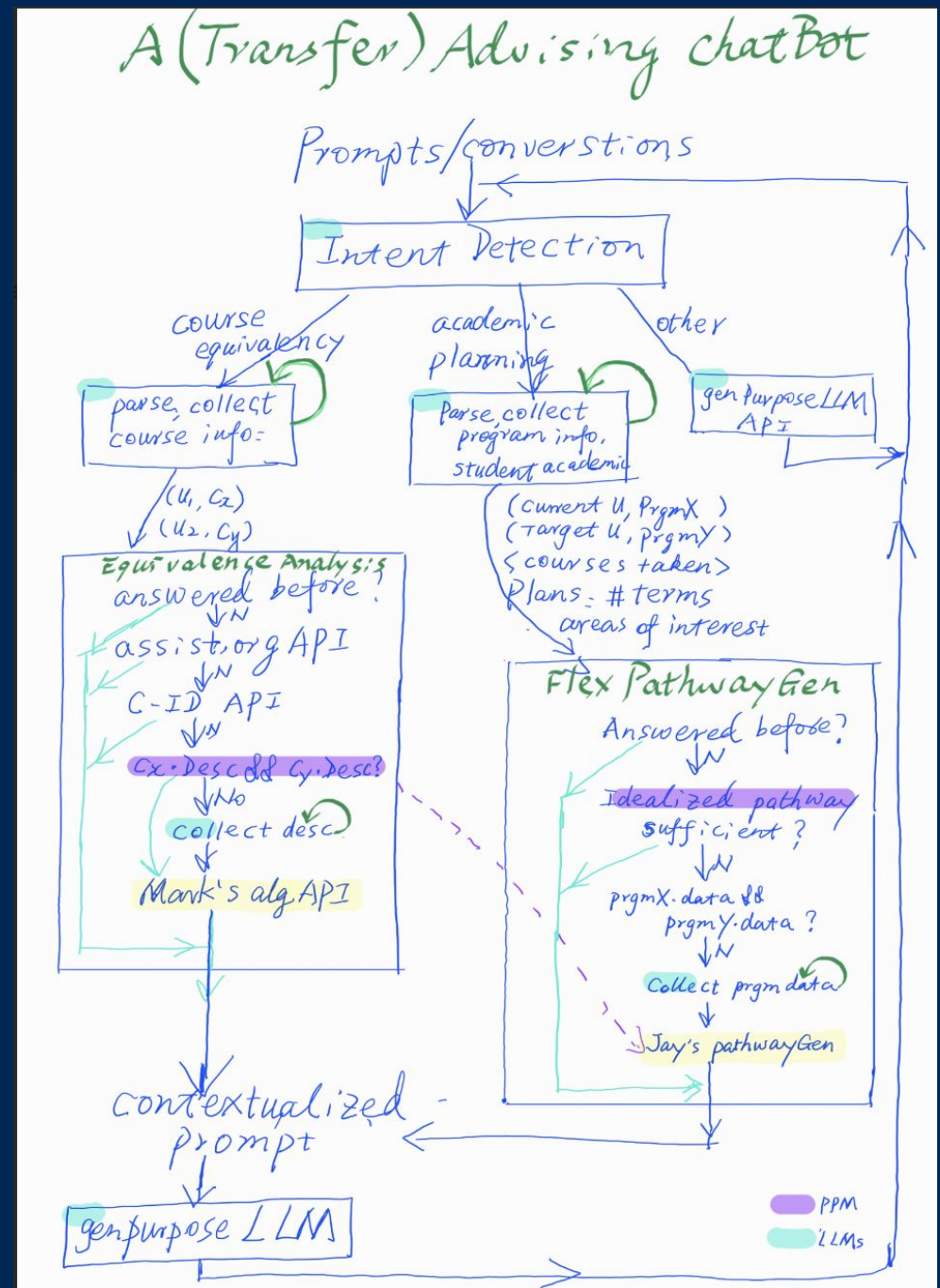
Constraint Programming, Satisfiability (CP-SAT) Solver using ILP

- Curricular constraints
- Curricular objectives
- Reference pathways

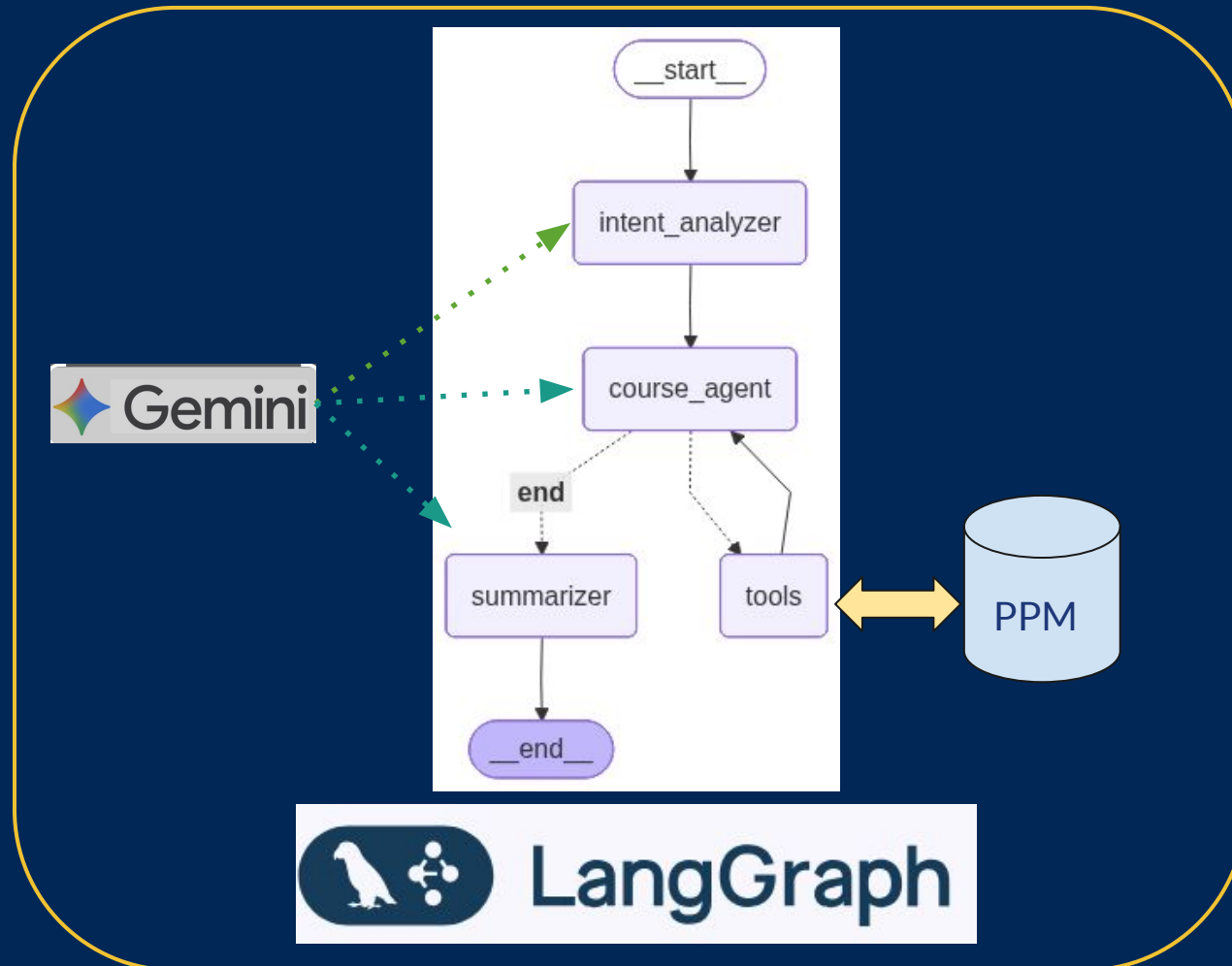
Advantages:

- Fast: a few seconds
- Easy to control: multiple feasible or optimal solutions
- Readily repurposed for pathway validation

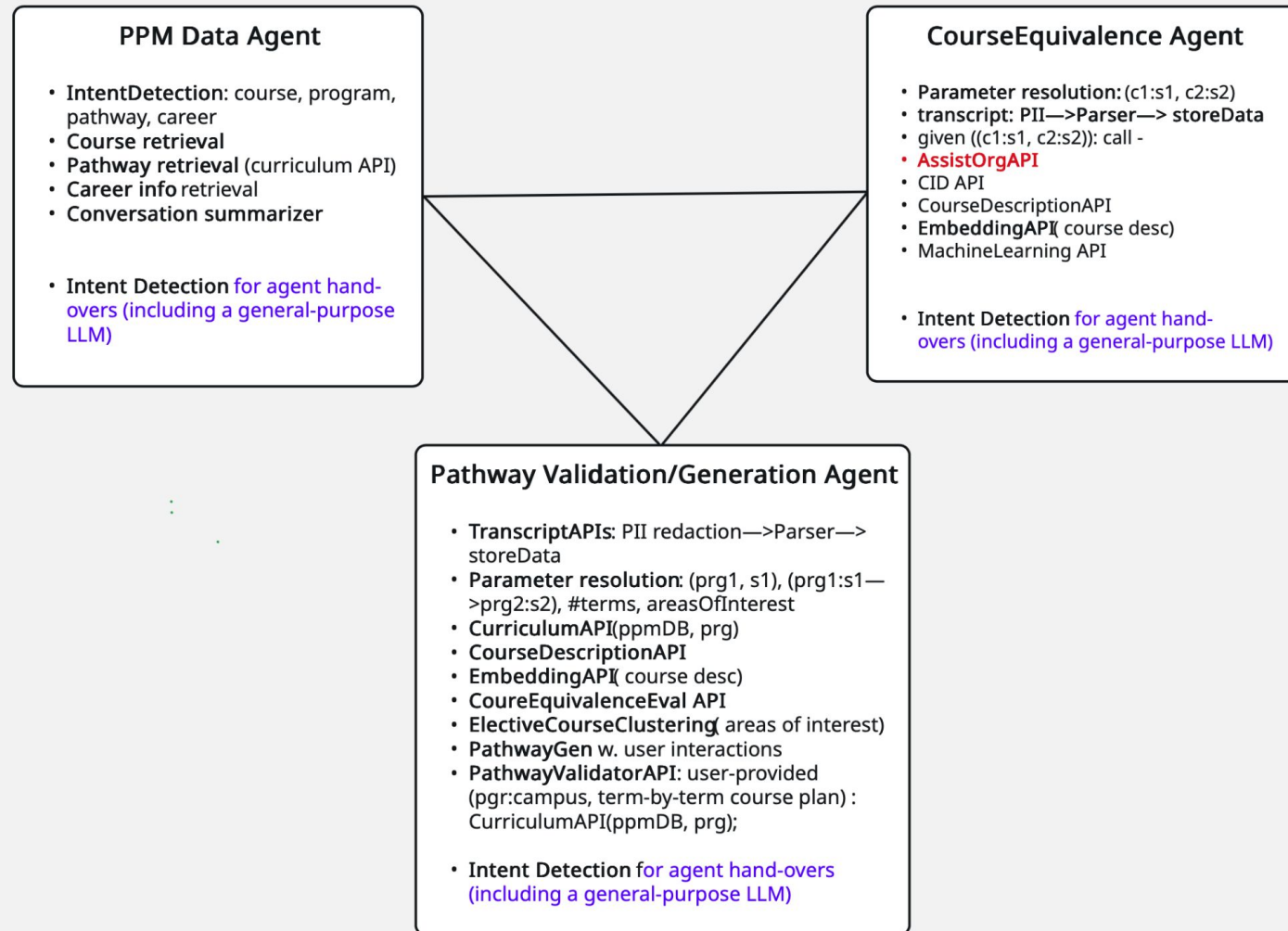
March 2025: A Transfer Advising Chatbot



New Addition: PPM Data Chatbot

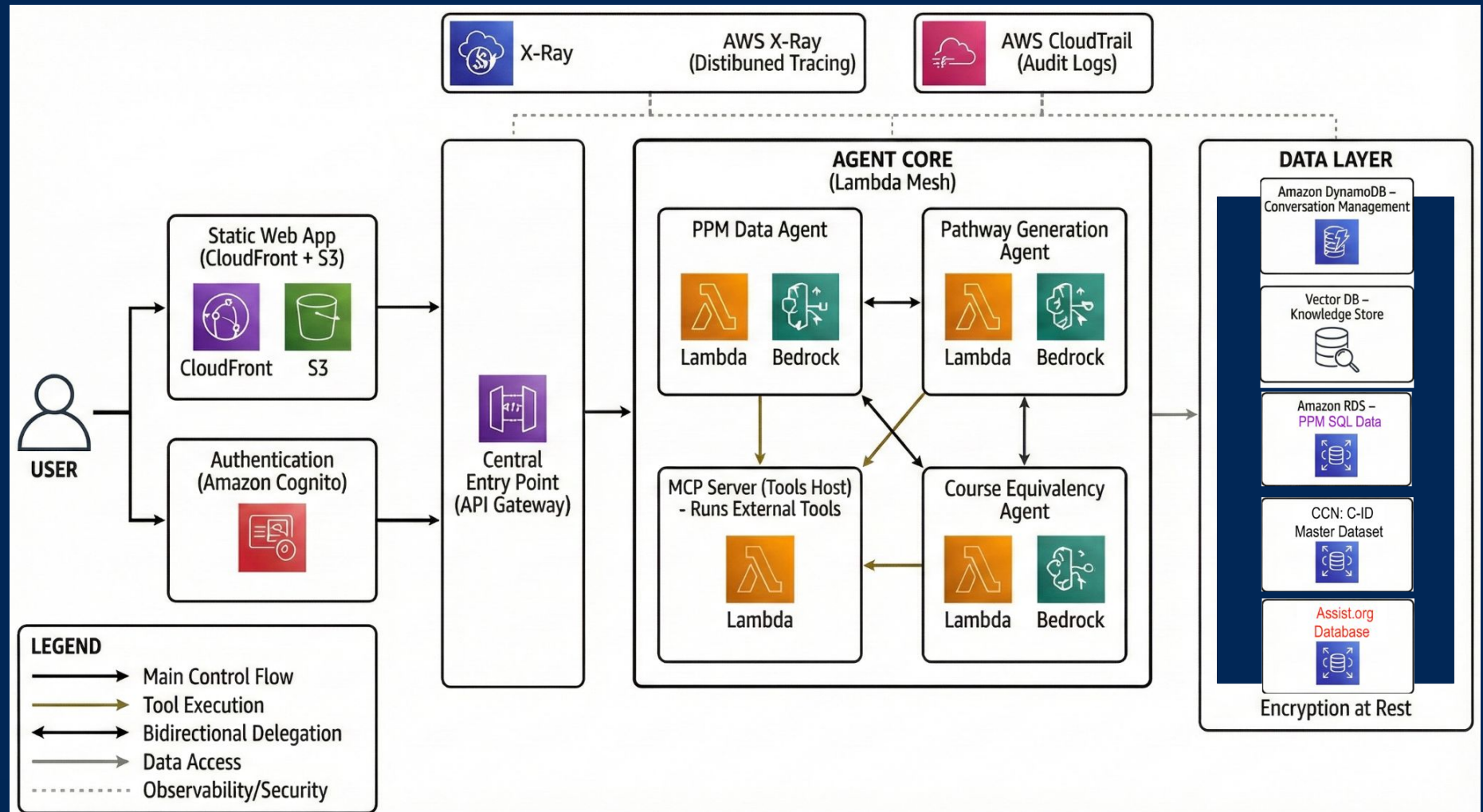


Put it all together: A Multi-agent Architecture



AWS-based Development and Deployment using Strands Agents

Strands Agents:
an open source AI
agents SDK



Live Demos

- **Dhvanil Bhagat:** PPM data chatbot
 - Main tech stack: PPM 1.0 database, LangGraph, Gemini 2.5
- **Divya Panchal:**
 - Transcript PII redaction and parsing using a local LLM
 - C-ID based course equivalency evaluation
 - Embedding-based course equivalency evaluation
 - Student's interest-based elective course clustering
- **Jay Lodha:** flexible pathway generation using CP-SAT
- **Shreyas Raghuraman:** a preliminary multi-agent Chatbot built with Strands agents SDK

Major Pending Tasks

Missing/Incomplete Data Sources

- assist.org database
- PPM 2.0 production database
 - A full list of elective for a given program

SF State

- assist.org APIs
- WIP: PPM 2.0 Data APIs
- Pathway generation
 - bridge a CC program with a CSU one
 - conduct user studies
 - prepare a manuscript for EDM 2026
- Continue developing/deploying the multi-agent prototype
- Conduct user studies



Thank You!

Questions?