

**Skillpath AI** 

AI-powered platform for personalized  
learning and seamless career transitions.

Sangram, Divya, Anusree

## PROBLEM STATEMENT

---



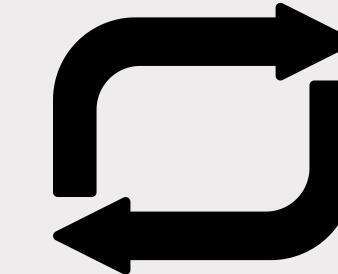
### FRAGMENTED LEARNING LANDSCAPE

**60%** of learners struggle to find courses aligned to their career goals



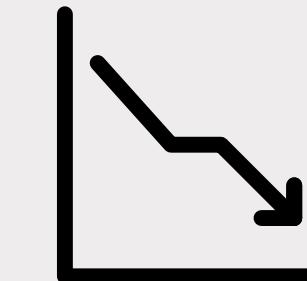
### WASTED TIME AND EFFORT

Learners spend **5–10 hours** manually researching courses, delaying their upskilling journey.



### MISALIGNED UPSKILLING FOR CAREER SWITCHERS

**70%** of career changers lack clarity on skills needed for their new roles.



### LACK OF PERSONALIZED GUIDANCE

Most platforms offer course catalogs, but lack personalized, skill-gap-based learning paths.

## OUR SOLUTION – SKILLPATH AI

---

- We set out to simplify and personalize the learning and career transition journey with SkillPath AI, an AI-powered platform that dynamically maps individual skills to career-specific learning paths.
- Using LLMs, retrieval-augmented generation (RAG), and Snowflake Cortex for semantic skill extraction, we deliver a solution that is both precise and adaptable.

SkillPath AI provides two pathways:



### Career Transition Assistant

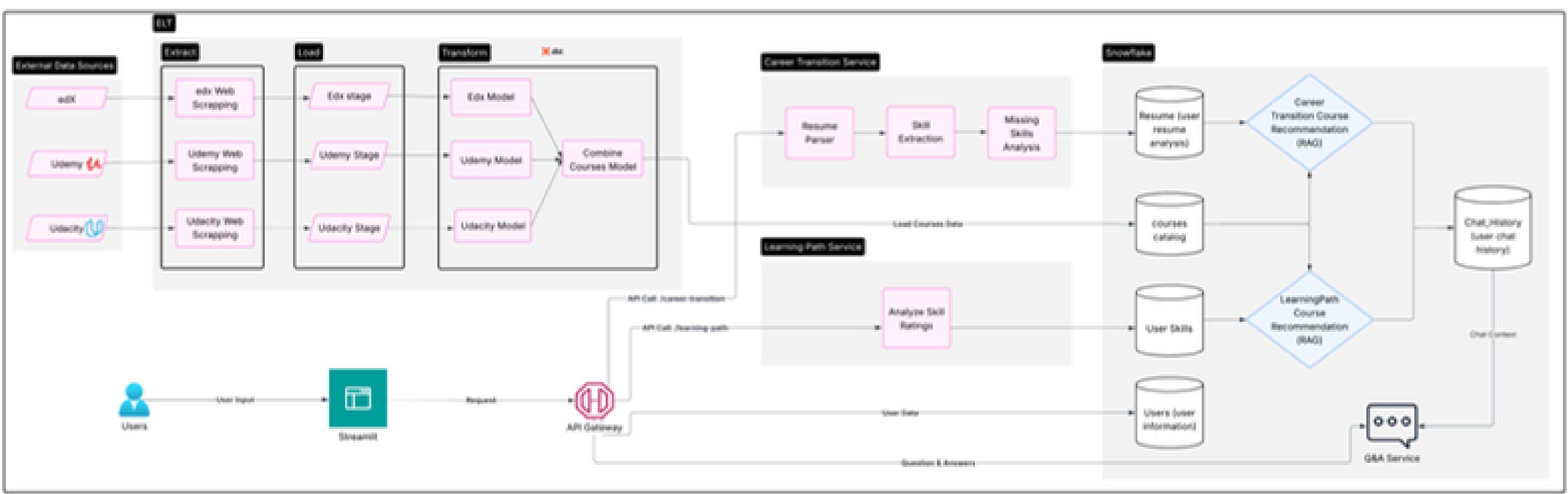
- Analyzes uploaded resumes, detects transferable skills and gaps.
- Builds a targeted 3–6 month upskilling roadmap toward a chosen role.



### Learning Path Assistant

- Users self-rate their proficiency on critical skills for their chosen target role.
- SkillPath AI generates a personalized, beginner-to-advanced learning path based on these self-assessments.

# HOW SKILLPATH AI WORKS: SYSTEM ARCHITECTURE



# **DEMO**

# TECH STACK

## FRONTEND



STREAMLIT

## BACKEND



FastAPI

## DATABASE



snowflake

## DATA ENGINEERING



dbt

## AI/ML LAYER



SNOWFLAKE CORTEX

## CONTAINERIZATION



docker

## VERSION CONTROL



git

# PLATFORM FEATURES

---

## PERSONALIZED LEARNING PATH GENERATION

- Creates structured course sequences based on user resumes or skill self-ratings.

## CAREER TRANSITION GUIDANCE

- Builds targeted upskilling plans by identifying transferable skills and skill gaps.

## SKILL GAP DETECTION AND ANALYSIS

- Highlights missing skills compared to industry-standard role requirements.

## GUARDRAILS FOR RESPONSIBLE AI INTERACTIONS

- Ensures AI responses remain focused on career development and upskilling.

## SESSION HISTORY AND RESUMABLE CONVERSATIONS

- Saves chat sessions, allowing users to resume their learning journeys anytime.

## CLOUD-READY, DOCKERIZED DEPLOYMENT

- Containerized Streamlit and FastAPI application for scalable cloud hosting.

## CHALLENGES & LESSONS LEARNED

---

### DATA CHALLENGES

- Initial datasets (Udemy, Coursera, edX, Udacity, etc) had missing skills, vague descriptions, and outdated metadata.
- API sources had limited daily calls and often required paid access.
- Solution: Enriched datasets using **Cortex COMPLETE**, built custom scrapers, and standardized data with dbt.
- Web scraping was complex due to different site structures but we created a structured way for scraping.

### AI/RAG PIPELINE CHALLENGES

- Considered Pinecone and FAISS but found them complex due to infrastructure overhead.
- Chose **Snowflake Cortex Search** for seamless integration — **blending semantic, keyword search, and metadata filtering** inside Snowflake.
- Faced challenges fine-tuning prompts for accurate skill extraction and course retrieval.

# DATA CHALLENGES

Udemy Developers

Overview Udemy Instructor API Documentation (v1.0)

Methods

- DELETE api-course-question-detail
- GET api-course-question-list
- GET api-course-question-replies-list
- GET api-messagethread-messages-list
- GET api-messagethreads-list
- GET api-taught-courses-list
- GET api-taught-courses-questions-list
- GET api-taught-courses-review-list
- POST api-course-question-replies-list

Over the last few years, we've noticed that instructors were using our internal Udemy APIs to build their own tools to help with their workflows. Since our internal APIs were not supported for external use, we decided to create a supported set of APIs that instructors can rely upon to build their tools. We've seen interesting applications for these APIs so far. One instructor built a tool to manage his Q&A workflow with his teaching assistants. Another instructor built an analytics dashboard for Q&A and course reviews. What will you build? Currently, the API is in maintenance mode and we do not have plans to add new features in the next few months. However, if you notice any bugs or have any feature requests, please do send them to our instructor support team ([instructorsupport@udemy.com](mailto:instructorsupport@udemy.com)) so we can address them or take them into consideration.

## Getting Started

The Udemy Instructor API exposes several functionalities of Udemy to help instructors build client applications and integrations with Udemy. It is organized around REST. Our API is designed to have predictable, resource-oriented URLs and to use HTTP response codes to indicate API errors. We use built-in HTTP features, like HTTP authentication and HTTP verbs, which can be understood by off-the-shelf HTTP clients. We only accept [https](https://) calls to the API. All responses will be returned in JSON format, including errors.

Udemy Instructor API is currently at version 1 and the root endpoint is <https://www.udemy.com/instructor->

```
571
572 | SELECT * FROM PLURALSIGHT_COURSES LIMIT 5;
573
```

↳ Results ⚡ Chart

	A LINK	A TYPE	A NAME	A INSTRUCTURE	A LEVEL	A DURATION	⌚ DATE CREATED	⭐ RATING
1	<a href="https://www.pluralsight.com/courses">https://www.pluralsight.com/courses</a>	courses	Oracle Database 12c Fundamentals	by Tim Warner	Beginner	3h 42m	0021-10-01	357.0
2	<a href="https://www.pluralsight.com/courses">https://www.pluralsight.com/courses</a>	courses	Oracle Database 12c Disassembling	by Tim Warner	Intermediate	3h 26m	0021-10-01	65.0
3	<a href="https://www.pluralsight.com/courses">https://www.pluralsight.com/courses</a>	courses	Oracle Database 12c: Instant Client	by Tim Warner	Beginner	2h 47m	0021-10-01	152.0
4	<a href="https://www.pluralsight.com/courses">https://www.pluralsight.com/courses</a>	courses	Microsoft Azure Solutions /	by John Savill	Advanced	1h 10m	0020-11-04	35.0
5	<a href="https://www.pluralsight.com/courses">https://www.pluralsight.com/courses</a>	courses	Microsoft Azure Solutions /	by Aaron Sampson	Advanced	1h 17m	0020-11-17	35.0

```

219  SELECT
220    ID,
221    COURSE_NAME,
222    DESCRIPTION,
223    LEVEL,
224    SNOWFLAKE.CORTEX.COMPLETE(
225      'mistral-large2',
226      'Given the following course title, description, and level, extract the most relevant technical or professional skills. \
227      If the description lacks detail, use the title to infer skills. \
228      Return only a concise, comma-separated list of skills (e.g., Python, SQL, Agile, Data Analysis). \
229      Do not include full sentences or general goals. \
230      Course Info: ' || COURSE_NAME || ', ' || DESCRIPTION || ', Level: ' || COALESCE(LEVEL, 'Not Specified')
231      ) AS INFERRED_SKILLS
232    FROM SKILLPATH_DB.RAW_DATA.STG_UDACITY_COURSES
233    WHERE (SKILLS IS NULL OR TRIM(SKILLS) = '')
234      AND DESCRIPTION IS NOT NULL;
235
236

```

↳ Results

↗ Chart

ID	COURSE_NAME	DESCRIPTION	LEVEL	INFERRED_SKILLS
bba42c0b-ee3e-4adf-b1cd-7f29bccf5a	Introduction to Javascript Courses	In this course, you will learn to build web applications using JavaScript.	Beginner	JavaScript, Data Types, Conditionals, Loops, Functions
e85636f3-7d56-4410-a15a-9b80bd23	Preparing and Modeling Data with Pandas	This course is a crucial step in Python data science. You will learn how to prepare and model data using Pandas.	Fluency	Data Modeling, Data Cleaning, Data Transformation, Pandas
e10c7da0-331a-4641-9245-5774db8b	Building Generative AI Solutions	This lesson covers vector databases and generative AI models.	Advanced	Vector Databases, AI, Long-term Memory, Retrieval Models
0b69f057-5672-4a63-83a1-07b35f719	Embedded Systems and Robotics	This course will teach you the basics of embedded systems and robotics.	Fluency	Embedded Systems, Electrical Circuits, Digital Logic, C
7e1f395d-e493-4d46-a2d9-3d1d33e6	Contrastive Language-Image Processing	In this lesson, you will learn about contrastive learning and its applications.	Fluency	CLIP, Machine Learning, Natural Language Processing
5e285ab9-ab28-4989-b4af-071d42ae7	UX Fundamentals & Design Research	Learn the fundamentals of User Experience design and research.	Beginner	UX Design, Design Research, User Experience, User-Centred Design

## ENRICHED DATASETS USING CORTEX COMPLETE

## FUTURE SCOPE & ROADMAP

---

### SCALING DATA & LEARNING SOURCES

- Add platforms: LinkedIn Learning, Skillshare, FutureLearn.
- Build a structured skill graph.
- Automate course updates and enrichment.

### ENHANCING PLATFORM FEATURES

- Track user progress, certifications, and goals.
- Add gamification: badges, milestones, streaks.
- Expand chatbot to cover career advice and portfolio building.

### IMPLEMENTING MODEL CONTEXT PROTOCOL (MCP)

- Dynamic Contextual Awareness
- Structured Input-Output Alignment

## REFERENCES

---

### **Official Documentation & Technical References**

Snowflake Official Documentation

Snowflake Cortex Search Documentation

FastAPI Documentation

Docker Documentation

OpenAI (ChatGPT) Documentation

Lucidchart — Diagram and architecture visualization

Cursor AI — AI coding assistant

Claude AI — Research reference

Perplexity AI — Quick search and research assistant

### **Course Datasets and Sources**

Udemy Courses Meta-Data (Kaggle).

### **Official Websites for Scraping:**

edX Official Course Catalog.

Udacity Official Course Catalog.

### **Snowflake Marketplace**

Chmura Economics & Analytics – Job Postings US

# Thank you