# AI Project

AI-Driven Career Guidance and Learning Roadmap Platform

## AI Course Project

**Team Members:**

Hanae OUAZZANI-AMRI 119276 - Data & Knowledge Graph
Oubahmane Mohamed Omar 120119 - AI/NLP and RAG
Akoujan Ali 118588 - Backend APIs Integration
Ramou Nassim 120068 - Frontend

UIR - 4th Year
November 16, 2025

# Contents

# 1 Introduction

Choosing a suitable career path has become increasingly complex for students and early learners in today's rapidly evolving job market. Job roles transform quickly, skill requirements vary significantly across companies and industries, and standard career advice often lacks the personalization needed to guide individual learners effectively. Many students feel overwhelmed, uncertain about what skills they possess, what competencies they need to develop, and how to structure their learning journey over time.

**SkillAtlas** addresses these challenges by providing an AI-driven, personalized career analysis and learning roadmap platform. By combining Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Neo4j knowledge graphs, and skill extraction models, SkillAtlas guides both experienced users with existing CVs and complete beginners toward meaningful career decisions backed by data-driven insights.

The platform serves as an intelligent career advisor, offering personalized recommendations, skill gap analysis, and time-aware learning roadmaps that adapt to each user's unique background and goals.

# 2 Problem Statement

Students and early-career professionals face several critical challenges in their career development journey:

- **Career Uncertainty:** Many students do not know which careers match their interests, strengths, or academic background, leading to uninformed decisions and potential career dissatisfaction.
- **Inconsistent Information:** Job descriptions across companies are inconsistent and difficult to interpret, making it hard to understand actual role requirements and expectations.
- **Lack of Beginner Guidance:** Fresh learners with no CV or professional experience receive little to no personalized guidance, leaving them without direction in their skill development.
- **Unclear Learning Paths:** There is no clear, structured path showing what skills to learn, in what order, with what priority, and realistic time estimates for skill acquisition.
- **Rapid Industry Evolution:** The fast-changing technology industry makes traditional academic advising insufficient, as curricula often lag behind industry needs.
- **Skill Gap Visibility:** Even students with some experience struggle to identify specific skill gaps between their current competencies and target roles.

SkillAtlas addresses these issues through personalized analysis powered by AI reasoning, structured knowledge modeling using graph databases, and intelligent recommendation systems that consider both user context and industry requirements.

# 3 Objectives

The primary objectives of SkillAtlas are to:

1. **Provide Personalized Career Recommendations:** Deliver tailored career suggestions based on user interests, identified strengths, or existing skills extracted from CVs.
2. **Extract and Analyze Skills:** Utilize advanced NLP models to automatically extract technical and soft skills from user-uploaded CVs with high accuracy.
3. **Build Comprehensive Knowledge Graphs:** Construct and maintain a Neo4j-based knowledge graph representing relationships between roles, skills, tools, technologies, and learning resources.
4. **Perform Intelligent Skill Gap Analysis:** Compare user competencies against target role requirements to identify specific missing skills and competency levels.
5. **Generate Time-Aware Learning Roadmaps:** Create realistic, personalized learning paths with time estimates (e.g., 6 months, 1 year, 2 years) based on skill complexity and user availability.
6. **Provide Interactive Visualizations:** Offer intuitive visual representations of knowledge graphs linking roles to required skills, tools, and career progression paths.
7. **Enable Conversational AI Assistance:** Implement a RAG-based chat assistant that answers career-related questions using grounded, reliable information from the knowledge base.

# 4 Target Users

SkillAtlas is designed to serve two primary user groups with distinct needs and contexts:

## 4.1 User Group 1: Students with a CV

This group includes students and early professionals who already have some experience through internships, projects, or coursework and possess a CV documenting their skills and achievements. These users want to:

- Understand how close they are to specific target roles (e.g., Data Engineer, ML Engineer, Full-Stack Developer)
- Identify exact skill gaps preventing them from qualifying for desired positions
- Receive actionable recommendations to improve their CV and skill profile
- Get validated assessments of their current competency levels
- Plan their next learning steps efficiently based on their existing foundation

### 4.2 User Group 2: Students without a CV / Beginners

This group consists of early learners, pre-university students, career changers, or undecided individuals who lack professional experience but have interests or aptitudes in certain areas. These users benefit from:

- Interactive questionnaires that help identify their interests and strengths
- Exploratory career suggestions that match their profile
- Beginner-friendly learning roadmaps starting from foundational concepts
- Guidance on building their first projects and experiences
- Understanding of different career paths and their requirements from scratch

Both user groups receive personalized, AI-powered guidance adapted to their unique starting points and career aspirations.

## 5 System Overview and User Interaction Flow

SkillAtlas operates through two primary modes tailored to different user types, with a common backend infrastructure supporting both pathways.

### 5.1 Mode 1: CV-Based Career Analysis

For users with existing CVs and experience:

1. **Input:** User uploads a CV (PDF or text format) through the web interface.
2. **Processing:**
    - NLP model (using transformer-based architectures) parses the CV and extracts skills, technologies, tools, and experience indicators
    - Skills are normalized and categorized (technical skills, soft skills, tools, frameworks)
    - User's skill profile is created and stored in MongoDB
3. **Role Selection:** User selects a target career role from the platform's role catalog.
4. **Analysis:** System queries Neo4j knowledge graph to retrieve all skills required for the chosen role and compares them against the user's extracted skills.
5. **Output:** User receives a comprehensive analysis including:
    - **Skill Gap Table:** Clear visualization of possessed skills vs. required skills
    - **Role Summary:** Overview of the target role, responsibilities, and career prospects
    - **Prerequisite Skills:** Foundation skills needed before advancing
    - **Personalized Roadmap:** Time-sequenced learning path with milestones
    - **Knowledge Graph Visualization:** Interactive graph showing skill relationships

### 5.2 Mode 2: Interest-Based Career Discovery

For beginners without CVs:

1. **Input:** User answers interactive questions about interests, academic background, strengths, and career preferences through a conversational interface.
2. **Processing:**
   - AI assistant (powered by LLM) analyzes responses and identifies career alignment patterns
   - System uses semantic matching to find careers that fit the user's profile
   - Interest vectors are embedded and compared against role embeddings
3. **Recommendations:** Platform suggests 3-5 career paths ranked by fit score with detailed explanations.
4. **Exploration:** User can explore recommended roles, view requirements, and ask questions through the AI chat assistant.
5. **Output:** Once a role is selected, user receives:
   - Complete beginner roadmap starting from foundational concepts
   - Estimated learning timeline with milestones
   - Recommended resources and learning materials
   - Project ideas to build practical experience

### 5.3 Common Features

Both modes provide access to:

- **AI Chat Assistant:** RAG-powered conversational interface for asking career questions, getting clarifications, and exploring alternative paths
- **Knowledge Graph Explorer:** Interactive visualization of skills, roles, and their relationships
- **Progress Tracking:** Ability to mark completed skills and update learning progress
- **Roadmap Customization:** Adjust timelines based on available study hours per week

## 6 Technical Architecture

### 6.1 System Components

SkillAtlas is built on a modular architecture integrating multiple AI and database technologies:

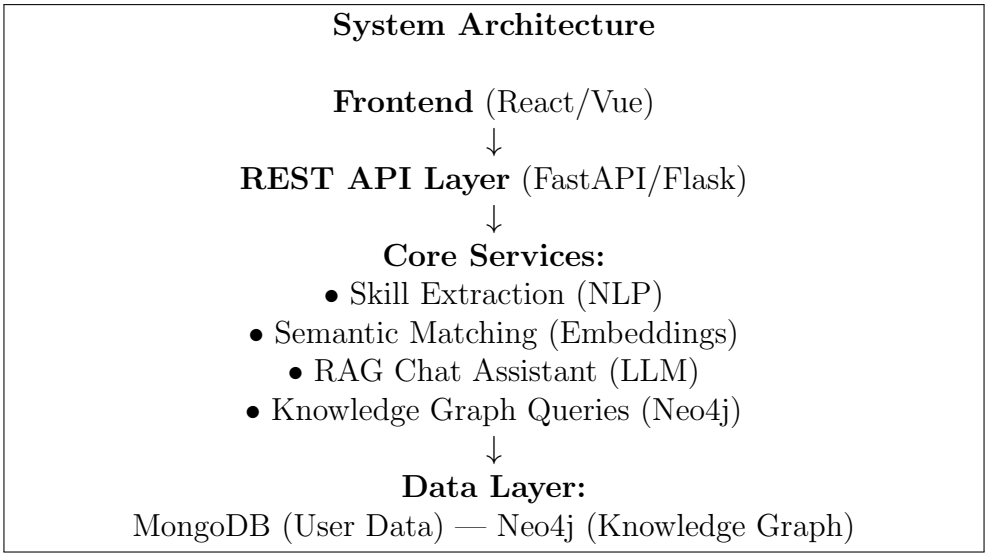


Figure 1: High-Level System Architecture

## 6.2 AI and Machine Learning Components

### 6.2.1 Large Language Models (LLMs)

LLMs serve multiple critical functions in SkillAtlas:

- **Conversational Interface:** Powers the AI chat assistant, interpreting user questions and generating natural, contextual responses
- **Career Matching:** Analyzes user interests and background to recommend suitable career paths
- **Roadmap Generation:** Generates personalized, time-aware learning sequences with explanations for each milestone
- **Skill Normalization:** Standardizes extracted skills to canonical forms (e.g., "React.js" → "React", "ML" → "Machine Learning")

**Implementation:** Uses models like GPT-4, Claude, or open-source alternatives (Llama, Mistral) via API or local deployment.

### 6.2.2 Embeddings

Embeddings transform text into dense vector representations for semantic similarity:

- **Skill Matching:** Converts skills into embeddings to find semantic similarities (e.g., "Python programming" is similar to "Python development")
- **Role Comparison:** Embeds role descriptions to match users with careers based on interest descriptions
- **Document Retrieval:** Powers the RAG system by embedding knowledge base documents for efficient semantic search

**Models Used:** OpenAI text-embedding-ada-002, Sentence-BERT, or open-source alternatives like all-MiniLM-L6-v2.

### 6.2.3 Retrieval-Augmented Generation (RAG)

RAG enhances LLM responses with accurate, grounded information:

- **Knowledge Base:** Stores career guides, skill descriptions, industry insights, and learning resources in MongoDB with vector embeddings
- **Query Processing:** When a user asks a question, the system:
  1. Embeds the user query
  2. Retrieves top-k most relevant documents from the knowledge base using vector similarity search
  3. Passes retrieved context to the LLM along with the query
  4. LLM generates a response grounded in the retrieved information
- **Benefits:** Reduces hallucination, provides cited answers, and keeps information up-to-date without retraining models

### 6.2.4 NLP Skill Extraction

Specialized models extract skills from unstructured CV text:

- **Named Entity Recognition (NER):** Fine-tuned transformers (BERT, RoBERTa) trained to identify skill entities in text
- **Keyword Extraction:** TF-IDF and custom models to extract technical terms
- **Classification:** Categorizes extracted skills into types (programming languages, frameworks, soft skills, tools)

## 6.3 Knowledge Graph (Neo4j)

Neo4j stores the structured knowledge domain:

- **Nodes:** Roles, Skills, Tools, Technologies, Learning Resources
- **Relationships:** REQUIRES (Role→Skill), USES (Role→Tool), PREREQUISITE_OF (Skill→Skill), LEADS_TO (Role→Role)
- **Properties:** Skill difficulty, average learning time, proficiency levels
- **Queries:** Cypher queries traverse the graph to find skill gaps, prerequisites, and career progressions

**Example Query:** Find all skills required for "Data Engineer" role with prerequisites:

```
MATCH (r:Role {name:"Data Engineer"})-[:REQUIRES]->(s:Skill)
OPTIONAL MATCH (s)-[:PREREQUISITE_OF]->(prereq:Skill)
RETURN s, prereq
```

### 6.4 Database Architecture

- **MongoDB:** Stores user profiles, CV data, extracted skills, chat history, and RAG document vectors
- **Neo4j:** Maintains the knowledge graph of domain relationships

## 7 Input and Output Specifications

### 7.1 System Inputs

1. **CV Upload:** PDF or TXT files containing resume information (for experienced users)
2. **Interest Questionnaire Responses:** Text responses to career interest and strength questions (for beginners)
3. **Target Role Selection:** User's chosen career goal from available options
4. **Chat Queries:** Natural language questions about careers, skills, and learning paths
5. **Time Preferences:** Available study hours per week for roadmap customization

### 7.2 System Outputs

1. **Extracted Skills Profile:** Structured list of user's current skills with categories
2. **Career Recommendations:** Ranked list of suitable careers with fit scores and explanations
3. **Skill Gap Analysis:**
   - Visual table comparing possessed vs. required skills
   - Gap percentage and priority rankings
   - Skill categories breakdown
4. **Personalized Learning Roadmap:**
   - Phase-by-phase learning plan (Foundation → Intermediate → Advanced)
   - Time estimates for each skill/milestone
   - Recommended learning resources and projects
   - Prerequisite chains and skill dependencies
5. **Knowledge Graph Visualization:** Interactive graph showing:
   - Selected role at center
   - Required skills as connected nodes
   - Tools and technologies used
   - Related career paths
6. **AI Assistant Responses:** Context-aware answers to user questions with citations

# 8 Use Cases

Figure 2 illustrates the primary use cases for SkillAtlas, showing interactions between different user types and system functionalities.

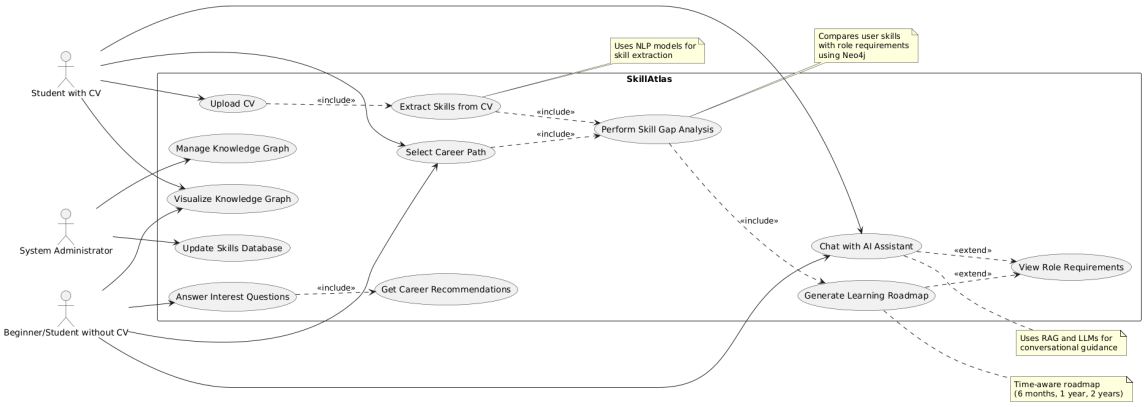


Figure 2: SkillAtlas Use Case Diagram

**Key Use Cases:**

1. **CV-Based Skill Analysis:** Experienced user uploads CV, system extracts skills, performs gap analysis, and generates roadmap
2. **Interest-Based Career Discovery:** Beginner answers questions, receives career recommendations, explores roles, and gets beginner roadmap
3. **Interactive Career Guidance:** Any user chats with AI assistant to ask questions, clarify doubts, and explore alternative paths
4. **Knowledge Graph Exploration:** Users visualize skill relationships, role requirements, and career progression paths
5. **Knowledge Base Management:** Administrators update role definitions, add new skills, and maintain the knowledge graph

# 9 Team Organization and Responsibilities

The project follows a modular development approach with four specialized tracks:

## 9.1 Track 1: Data & Knowledge Graphs

**Responsibilities:**

- Design and implement Neo4j knowledge graph schema
- Define nodes (Roles, Skills, Tools) and relationships (REQUIRES, USES, PREREQUISITE_OF)
- Populate graph with career data from industry sources

- Develop Cypher queries for skill gap analysis and prerequisite chains
- Create graph visualization endpoints
- Ensure data consistency and graph integrity

## 9.2 Track 2: AI/NLP and RAG

**Responsibilities:**

- Develop and fine-tune skill extraction models for CV parsing
- Implement embedding pipeline for semantic matching
- Build RAG system architecture with vector database integration
- Configure and optimize LLM prompts for career recommendations
- Develop semantic search for knowledge retrieval
- Implement skill normalization and categorization logic
- Tune model parameters for accuracy and performance

## 9.3 Track 3: Backend Integration

**Responsibilities:**

- Design and implement REST API architecture (FastAPI/Flask)
- Develop endpoints for CV upload, skill extraction, and roadmap generation
- Integrate MongoDB for user profile and document storage
- Connect Neo4j knowledge graph to API layer
- Implement authentication and user session management
- Build chat API endpoints for AI assistant interaction
- Ensure API security, error handling, and performance optimization
- Design data flow between frontend, AI modules, and databases

## 9.4 Track 4: Frontend development

**Responsibilities:**

- Design and implement user interface (React/Vue.js)
- Create CV upload and skill visualization components
- Build interactive chat interface for AI assistant
- Implement knowledge graph visualization using D3.js or similar

- Design roadmap display with timeline and milestone tracking
- Develop responsive layouts for mobile and desktop
- Ensure smooth user experience and intuitive navigation
- Integrate frontend with backend APIs

# 10 Academic Contribution and Innovation

SkillAtlas demonstrates several key academic and practical contributions:

1. **Applied AI System Design:** Showcases integration of multiple AI technologies (LLMs, embeddings, NER) in a cohesive system addressing real-world problems
2. **Retrieval-Augmented Generation:** Implements RAG architecture to ground LLM responses in accurate, domain-specific knowledge
3. **Knowledge Graph Applications:** Demonstrates graph database usage for modeling complex domain relationships and enabling intelligent querying
4. **Semantic Matching:** Applies embedding-based similarity for skill comparison and career recommendation
5. **Modular Architecture:** Exemplifies best practices in dividing complex systems into manageable, specialized components
6. **User-Centric AI:** Focuses on practical, personalized AI assistance rather than generic automation

# 11 Implementation Timeline

| Phase | Activities |
|---|---|
| Phase 1 | Data collection, knowledge graph design, initial schema |
| Phase 2 | NLP model training/fine-tuning, embedding pipeline setup |
| Phase 3 | Backend API development, database integration |
| Phase 4 | Frontend UI development, visualization components |
| Phase 5 | System integration, testing, and optimization |
| Phase 6 | Documentation, presentation, and final delivery |

Table 1: Project Implementation Phases

# 12 Conclusion

SkillAtlas represents a comprehensive AI-driven solution to the complex problem of career guidance and skill development planning. By combining Large Language Models, Retrieval-Augmented Generation, knowledge graphs, and intelligent skill extraction, the platform provides personalized, actionable guidance to both experienced students and complete beginners.

The modular architecture ensures each team member can focus on their area of expertise while contributing to a cohesive, production-quality system. The project aligns with academic best practices in AI system design, demonstrating practical applications of cutting-edge technologies to solve real-world problems in education and career development.

Through SkillAtlas, students gain clarity on their career paths, understand their skill gaps, and receive realistic roadmaps for achieving their professional goals—transforming career planning from an overwhelming challenge into a guided, data-driven journey.