UNIVERSITY OF PLYMOUTH

PUSL3190 Computing Project

Project Initiation Document (PID)

AI Partner for Career Development:
A Multi-Agent System for Personalised Job Seeking

Supervisor: Mr Diluka Wijesinghe

Name: Jeyaseelan Jakshigan
Plymouth Index Number: 10953354
Degree Program: BSc (Hons) Computer Science

Table of Contents

Tables

Diagrams

Introduction

The IT sector of Sri Lanka continues to grow every year, yet many graduates struggle to align their skills and job applications with current industry demands. Candidates are forced to use a scattered set of disconnected tools, such as Canva to build a CV, Jobscan to check for keywords, LinkedIn to find job listings, and the Google interview platform to practice generic interview questions. This current state of affairs is not only cumbersome but also dampens the job search process. Recent advances in multi-agent AI systems enable the design of stateful, coordinated agents capable of executing multi-step workflows. This project applies these advances in an academic context to explore how agent orchestration and semantic reasoning can improve personalised career guidance, as well as the Agent-to-Agent (A2A) protocol for structured inter-agent communication.

Problem Statement

Current digital career tools suffer from four key limitations:

1. Fragmentation: Feedback from CV analysis is not reused during interview preparation or skill planning.
2. Semantic Limitations: Many tools rely on keyword matching rather than understanding relationships between skills and roles.
3. Stateless Interaction: User progress and historical feedback are lost between sessions.
4. Limited Contextual Awareness: Tools rarely incorporate up-to-date, local labour-market data.

As a result, candidates receive repetitive and generic advice that does not evolve with their goals or performance.

Scope and Limitations

The project will deliver a functional prototype that demonstrates coordinated multi-agent behaviour for career development tasks. Core functionality focuses on text-based CV analysis, semantic matching, and career guidance. Advanced components such as real-time speech-based interview interaction and enhanced security guardrails are treated as stretch objectives. The system is designed for candidate use only and is not intended for production deployment. Real-time market data relies on public job listings, which may undergo structural changes over time.

Expected Impact and Stakeholders

The AI Partner for Career Development is designed to transform the job-seeking process from a fragmented task into a cohesive, context-aware journey. By integrating Multi-Agent Systems (MAS), the project aims to reduce "application fatigue" and improve the "Semantic Match Score" between candidates and industry requirements. The primary direct stakeholders are final-year undergraduates, recent graduates, and interns who utilise the platform for personalised CV optimisation and low-latency interview coaching. The indirect stakeholders include University Career Guidance Units (CGUs) seeking scalable support tools, IT recruiters looking for industry-ready talent, and academic examiners evaluating the project's technical rigour. According to Ransbotham et al. (2025), such agentic systems move beyond simple automation to provide stateful, long-term career

planning. Ultimately, this project seeks to bridge the national "skills gap" identified by SLASSCOM (2024), providing a high-impact, AI-driven blueprint for modern career architecture.

Business Case

2.1 Business Need

Manual job preparation is time-consuming and inefficient, where candidates often tailor CVs through trial and error without clear insight into how employers evaluate applications. This leads to misaligned skill development and reduced confidence due to continuous rejections.
An integrated system that combines CV analysis, job matching, and interview preparation can significantly reduce preparation effort and also improve the relevance of feedback. By utilising the A2A Protocol, this system solves the isolation problem, allowing agents to fetch external market data without any user intervention.

2.2 Business Objectives

The project aims to:
Increase Efficiency: Reduce the time required for CV tailoring and interview preparation by 70% through automated, personalised feedback generated by the CV Optimiser Agent.
Improve Alignment: Improve the "Semantic Match Score" between candidate skills and job requirements by atleast 25% using GraphRAG analysis rather than keyword matching.
Validate Architecture: Evaluate the robustness of a multi-agent architecture by analysing system behaviour under partial agent failure conditions, which aims to maintain 99% session integrity.

Project Objectives

3.1 Functional Deliverables

ADK Orchestrator: A coordinator agent implemented using Google ADK (Python SDK) to manage workflows (SequentialAgent, RouterAgent) and maintain session state via the Interactions API.
Specialist Agent Mesh: A set of specialised agents for CV critique, semantic matching, career guidance, and market-trend analysis.
A2A Communication Layer: An implementation of the Agent-to-Agent (A2A) protocol enabling structured task delegation between agents (e.g., on_message_send for task handoffs).
GraphRAG Engine: A skill reasoning component utilising the ESCO ontology stored in Neo4j to infer implicit skills (e.g., "React" implies "JavaScript").
Native Audio Interview Coach: An interactive agent using Gemini 2.5 Flash Native Audio via WebSocket for real-time, low-latency speech-to-speech mock interviews.

3.2 Quality Criteria

Performance: Text-based analysis responses returned within 10 seconds. Audio latency must be under 500ms to ensure natural conversation.

Reliability: Agent failures do not terminate the entire workflow, and the ADK Coordinator must handle exceptions and offer fallback options.
Privacy: Personally identifiable information (PII) is removed client-side using Transformers.js before cloud processing.

3.3 Success Metrics

Match Improvement: Quantitative improvement in cosine similarity scores between baseline and optimised CV–job description pairs across a controlled test set.
Usability: Positive feedback from student testers regarding the naturalness of the Interview Coach's voice and the relevance of the CV suggestions.
Workflow Completion: Successful execution of multi-agent workflows (e.g., CV Analysis -> Market Lookup -> Interview Prep) without loss of state.

Literature Review

Recent research indicates a significant shift from stateless chatbots to stateful multi-agent architectures capable of long-term planning and coordination (Ransbotham et al., 2025). Frameworks such as Google ADK and LangGraph have emerged as the standard for defining these "state machines" where agents possess distinct roles and persistent memory. In terms of information retrieval, traditional keyword-based matching is increasingly replaced by embedding-based approaches like Sentence-BERT (SBERT), which report higher accuracy in aligning CVs with job descriptions (Sankar et al., 2021). However, newer studies on GraphRAG (Graph-based Retrieval Augmented Generation) suggest that vector-only search often fails to capture complex professional hierarchies, proposing knowledge graphs as a superior method for mapping "skill trees" and career trajectories (Edge et al., 2024).

A critical emerging theme in recent years is the transition from human-centric Identity and Access Management (IAM) to Agent-centric Security. As multi-agent systems proliferate, the industry has identified a "Looming Authorisation Crisis" where traditional human-session protocols fail to secure autonomous, ephemeral agents (ISACA, 2025). Recent research from the Cloud Security Alliance (2025) emphasises that AI agents often operate as Non-Human Identities (NHIs), which now vastly outnumber human users in the enterprise. This necessitates a shift toward Zero Trust Architecture for agents, where every request is authenticated and authorised through robust Machine-to-Machine (M2M) authentication and just-in-time (JIT) credentials rather than static API tokens (CyberArk, 2026). Key strategies highlighted in the recent 2026 playbooks include Zero Standing Privileges (ZSP), ensuring agents only possess the permissions necessary for the duration of a specific task (Prefactor, 2025).

Despite these advancements, a critical analysis of existing literature reveals a persistent "Orchestration Gap." While individual studies explore CV analysis, job matching, or interview bots in isolation (Swain & Malik, 2025), there is a lack of a holistic, candidate-facing pipeline where these components share a persistent user context. Most academic prototypes are designed as "black-box" optimisation tools for recruiters rather than collaborative, stateful "co-pilots" for job seekers (Zhang et al., 2024). Furthermore, while advancements in real-time multi-modal feedback and agentic

learning loops exist, they are rarely integrated into a closed-loop agentic workflow that continuously adapts system behaviour based on user performance and contextual requirements, particularly in candidate-centric career development scenarios (Cai et al., 2025).

The implications of this review for the current project are clear: while the individual technical components for an AI career co-pilot are maturing, the primary research frontier lies in the seamless integration and stateful management of these tools. This project addresses the identified gaps by utilising LangGraph to manage a "Coordinator Agent" that maintains context across fragmented tasks, from CV optimisation and local market scraping to real-time speech-based interview coaching. By implementing the A2A (Agent-to-Agent) Protocol and incorporating an "Adversarial Debate" pattern for iterative refinement (Chan et al., 2024), this project moves beyond isolated tools toward a fully integrated, candidate-centric ecosystem that bridges the gap between theoretical research and practical career preparation.

The data for the comparative analysis in the table below is synthesised from a cross-evaluation of the industry standards and empirical research discussed throughout this review. Specifically, the architectural distinction between monolithic LLMs and the proposed Multi-Agent System (MAS) is grounded in the shift toward stateful, role-based workflows identified by Ransbotham et al. (2025). The transition from simple semantic similarity to GraphRAG-driven "Skill Trees" is supported by the structural advantages of knowledge graphs documented by Edge et al. (2024) and the matching accuracy benchmarks of Sankar et al. (2021). Furthermore, the security parameters for Agent IAM and Zero Trust are based on the emerging risk frameworks established by ISACA (2025) and CyberArk (2026) regarding the management of Non-Human Identities. Finally, the advancements in interviewing capabilities, specifically the move toward native, low-latency audio, directly address the "Orchestration Gap" and the lack of closed-loop feedback loops identified in the critical analyses of Zhang et al. (2024).

Table 1: Comparison of Existing Career AI Approaches

Feature	**Traditional RAG Tools**	**Modern LLM Chatbots**	**Proposed AI Partner**
Architecture	Single-task scripts	Monolithic LLM	Multi-Agent (MAS)
Context	Stateless (per session)	Short-term memory	Persistent Candidate Profile
Reasoning	Keyword/Vector Match	Semantic Similarity	GraphRAG (Skill Trees)

Interviewing	Text-only feedback	Generic prompts	Native Audio (Low-latency)
Security	Static API Keys	User-based Auth	Agent IAM (Zero Trust)

Method of Approach

5.1 Research Design

The project follows a Design Science Research (DSR) methodology, iteratively designing, implementing, and evaluating the proposed system. Development follows Agile principles with two-week sprints, allowing for regular refinement of agent behaviours based on testing.

The development will commence with a backend-first approach, prioritising the engineering of specialised agents for CV critique, generation, and market analysis using the Gemini API. This will be followed by the implementation of a multi-modal interview agent leveraging Gemini Flash's native audio capabilities. A centralised RAG architecture will be established using a vector database to provide the agents with domain-specific knowledge. Integration will be managed by a central Orchestrator utilising the A2A protocol for seamless agent communication, secured via an Agent-specific IAM framework. The project will culminate in a Next.js frontend integration, with a rigorous testing phase employing Jest, Mocha, and Pytest to ensure functional and performance benchmarks are met.

5.2 Data Sources and Collection

Ontology: ESCO v1.1 (European Skills, Competences, Qualifications and Occupations) imported into a graph database to serve as the "Skill Brain."

Job Descriptions Data: Public job descriptions scraped from Sri Lankan job portals (topjobs.lk) to train the Market Trends agent.

Test Data: Synthetic CVs generated to represent various student profiles (Intern, Junior SE) for controlled evaluation.

Privacy Considerations: To ensure data security, the system implements Client-Side Redaction using Transformers.js. This strips all Personally Identifiable Information (PII), such as names, phone numbers, and emails, locally before any data is transmitted to cloud-based LLM APIs.

5.3 Tools and Technologies

Orchestration: Google Agent Development Kit (ADK) for defining agent hierarchy and managing the "Interactions API".

Communication: Agent-to-Agent (A2A) protocol for standardised task exchange.

Semantic Intelligence: Sentence-BERT (all-MiniLM-L6-v2) for vector similarity and Neo4j for Knowledge Graph storage.

Audio Model: Gemini 2.5 Flash Native Audio (via Live API) for the interview coach.

Runtime: Vertex AI Agent Engine for scalable deployment.

Agent IAM: Identity management for Non-Human Identities (NHI) using Short-Lived JWT Tokens and

Machine-to-Machine (M2M) authentication to secure inter-agent communication.

5.4 Evaluation Strategy

Baseline Comparison: Baseline semantic similarity scores are compared against post-optimisation scores across a minimum of 20 CV–job description pairs.

A2A Tracing: System logs will be analysed to verify that the Coordinator Agent correctly delegates tasks via A2A (e.g., verifying that the Resume Agent successfully requests data from the Market Agent).

Security Audit: Verification that Just-In-Time (JIT) credentials correctly expire and that PII redaction successfully prevents data leakage in logs.

User Feedback: Qualitative feedback is collected from student testers via a "Net Promoter Score" survey.

Conceptual Diagram

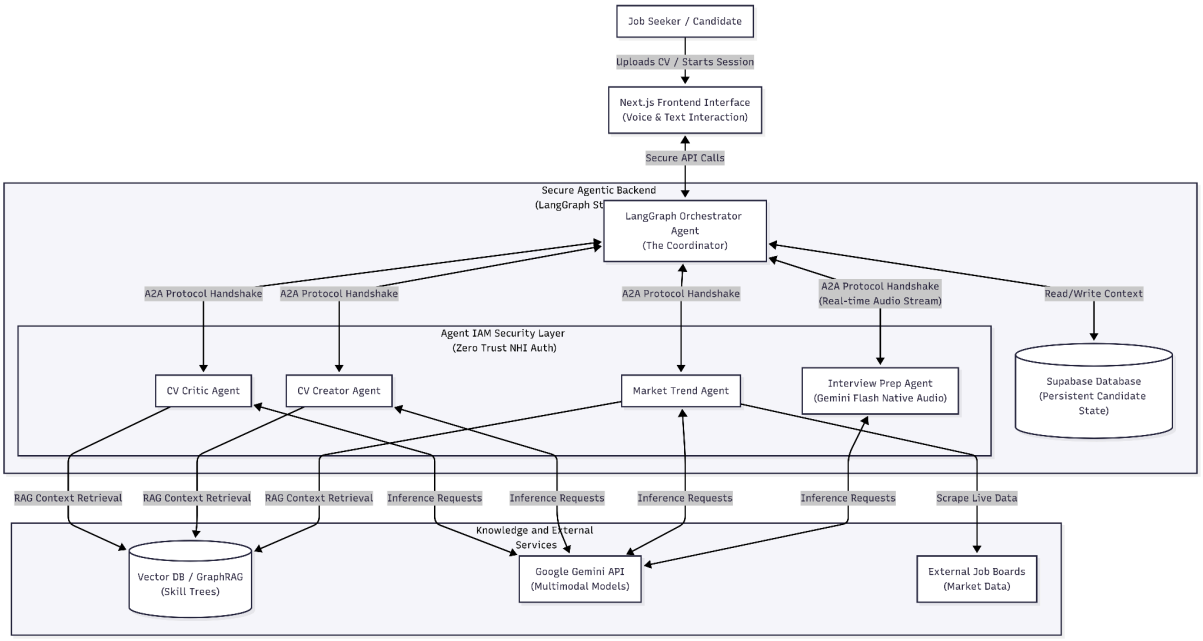

Diagram 1: Conceptual Diagram

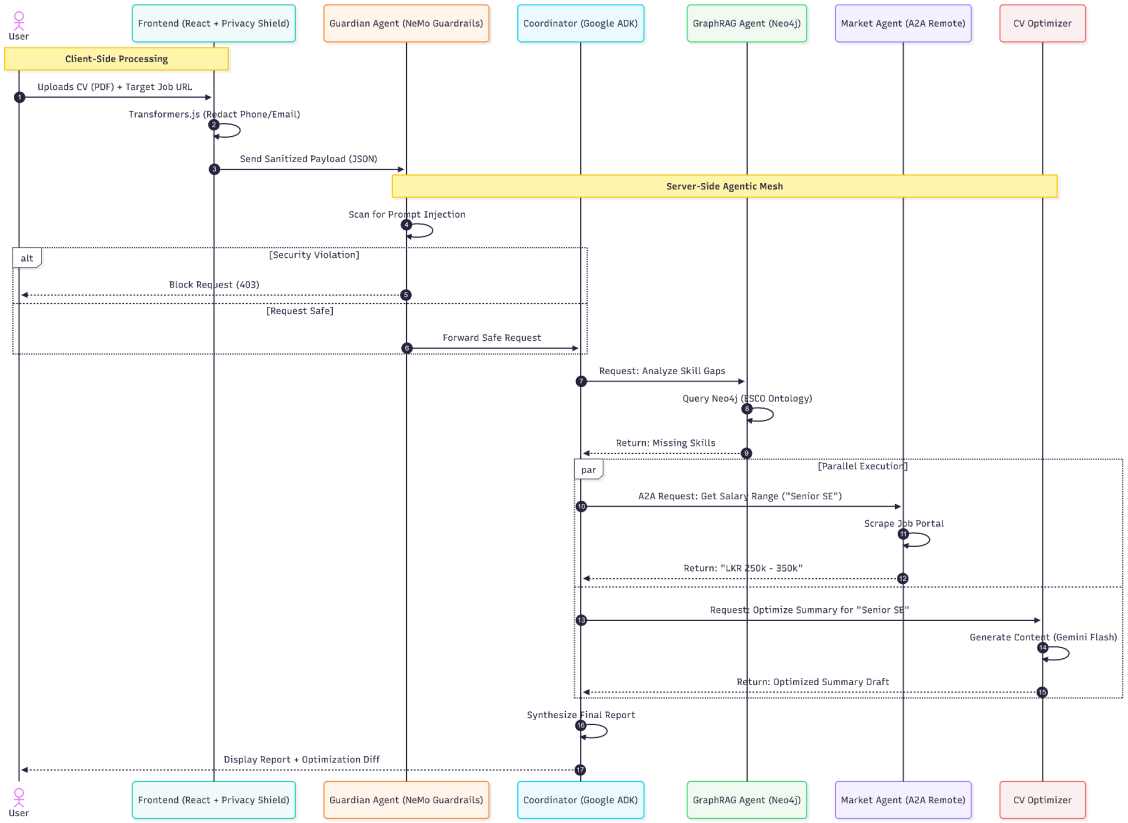

Diagram 2: Sequence Diagram

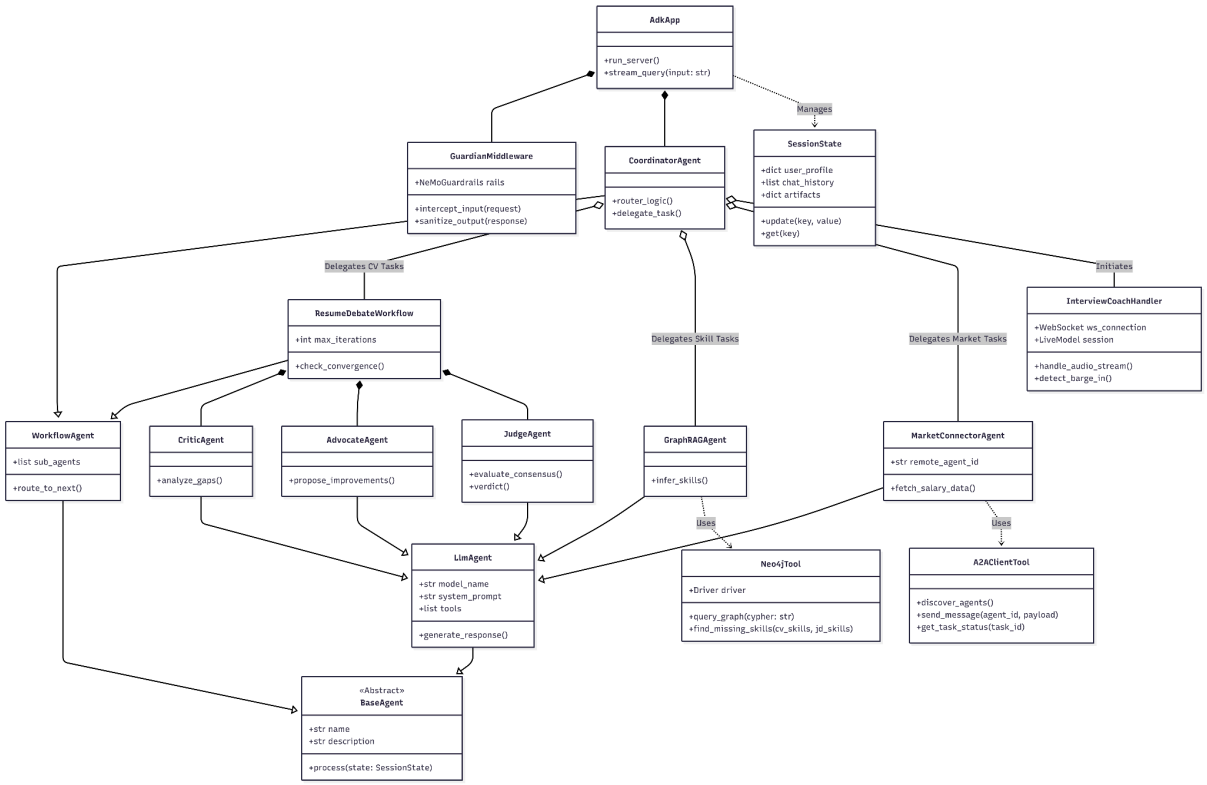

Diagram 3: Class Diagram

Initial Project Plan

Table 2: Project timeline Table

SDLC Phase	Task / Deliverable	Start Date	End Date
1. Requirement Analysis	Define problem, gather requirements, and finalise proposal	01 Oct 2025	31 Oct 2025
2. System Design	UI mockups(Wireframes) Design architecture, database schema	01 Nov 2025 16 Nov 2025	15 Nov 2025 30 Nov 2025
3. Implementation – Phase 1	Develop Core CV Agents (Critique, Matcher)	01 Dec 2025	15 Dec 2025
4. Implementation – Phase 2	Develop Interview Coach and Market Trends agents	16 Dec 2025	30 Jan 2026
5. Integration and Testing	Integrate all modules, perform testing	01 Feb 2026	15 Mar 2026
6. Deployment and Maintenance	System Deployment Prepare final report	16 Mar 2026	30 Apr 2026

Risk Analysis

Table 3: Risk Analysis Table

Risk	Likelihood	Impact	Mitigation Strategy
Agent Orchestration Complexity: Agents may get stuck in infinite loops or fail to hand off tasks correctly via A2A.	High	High	Use ADK's LoopAgent primitives with strict max_iterations limits. Implement a "Supervisor" logic to kill stuck processes.
Scope Creep (Frameworks): Rapid updates to Google ADK or Gemini models (e.g., API deprecations) may break code.	Medium	Medium	Freeze core library versions early in development. Use the Vertex AI Agent Engine, which abstracts some infrastructure changes.

Latency (Audio): Real-time audio via WebSockets depends on network stability.	Medium	High	Use Gemini 2.5 Flash Native Audio, which is optimised for streaming. Implement a "Push-to-Talk" fallback if network latency exceeds 500ms.
Data Privacy: Potential leakage of PII to the LLM.	Low	High	Implement Client-Side Redaction using transformers.js to strip emails/phones before the API call is made.

Reference List

A2A Project (2025) *Agent-to-Agent (A2A) Protocol Specification*. Available at: https://github.com/a2aproject/A2A (Accessed: 15 January 2025).

Ransbotham, S., Kiron, D., Khodabandeh, S., Iyer, S. & Das, A. (2025) *The Emerging Agentic Enterprise: How Leaders Must Navigate a New Age of AI. MIT Sloan Management Review & Boston Consulting Group* report

Chan, Y., Zhang, H. and Liu, K. (2024). 'Adversarial Debate and Multi-Agent Critique: Improving LLM Outputs through Iterative Refinement'. *arXiv preprint*. arXiv:2405.08922.

Cai, S., Zhang, J., Bao, K., Gao, C., Wang, Q., Feng, F. and He, X. (2025) *Agentic feedback loop modeling improves recommendation and user simulation.* arXiv preprint arXiv:2410.20027v2. Available at: https://arxiv.org/abs/2410.20027v2 (Accessed: 15 January 2025).

Cloud Security Alliance (2025). *Securing Non-Human Identities (NHI) in the Age of AI Agents*. Seattle: CSA.

CyberArk (2026). *The Identity Security Playbook: From Humans to Agents*. Newton, MA: CyberArk Research.

Dharmendra, T. A., Meenakshi, K. and Bhargava, D. (2025). 'NLP-Powered Resume Screening and Ranking System'. *2025 3rd International Conference on Disruptive Technologies (ICDT)*, pp. 1361-1366.

Edge, D., Trinh, H. and Cheng, S. (2024). 'From RAG to GraphRAG: Utilizing knowledge graphs for complex professional hierarchies'. *Microsoft Research Blog*.

ISACA (2025). 'The Looming Authorisation Crisis: Multi-Agent Systems and Zero Trust'. *ISACA Journal*, 4.

Mandal, S., Gupta, A. and Rao, P. (2023). 'AI Mock-Interview Systems: A survey of text-based vs. multimodal approaches'. *Computers in Human Behavior*, 142.

Masters, J., Thompson, R. and Khalil, S. (2025). 'The Manager Agent: Orchestrating Human-AI Teams'. *Journal of Autonomous Agents and Multi-Agent Systems*, 39(1).

Sankar, V., Raman, N. and Iyer, S. (2021). 'A comparative study of keyword-based and semantic matching for CV alignment'. *International Journal of Computer Applications*, 178(41).

SLASSCOM (2024). *Employability Skills Report 2024: Sri Lanka IT-BPM Industry*. Colombo: SLASSCOM.

Swain, D. and Malik, P. (2025). 'Holistic Integration of AI Career Tools: Theoretical Frameworks and Practical Gaps'. *Education and Information Technologies*, 30(2).

Wu, Q., Bansal, G. and Wang, C. (2023). 'AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation Framework'. *arXiv preprint*. arXiv:2308.08155.

Xenova (2025). *Transformers.js Documentation*. Hugging Face.

Zeng, J. et al. (2025). 'Standardizing Agentic Workflows: From Proprietary APIs to Open Protocols'. *Journal of AI Research (JAIR)*, 72.

Zhang, R. et al. (2024). 'User-Centric AI in Career Architecture: Bridging the Gap between Recruiters and Candidates'. *ACM Transactions on Interactive Intelligent Systems (TiiS)*.