Report: Agentic AI for Automated Professional Verification

Course: Agentic AI for Business and FinTech (FTEC5660)
Task: Individual Homework 02 – CV Verification System

1. System Architecture and Design Decisions

The CV Verification System is engineered as an autonomous agentic framework designed to mitigate recruitment fraud by cross-referencing candidate-provided data against live professional and social graphs.

1.1 Technical Stack

- **Core Logic:** Orchestrated via the **LangGraph ReAct (Reasoning and Acting)** framework, which facilitates iterative decision-making cycles.
- **LLM Backbone:** Utilizing **DeepSeek-Chat (OpenAI-compatible)** for its high-performance reasoning capabilities in processing semi-structured CV data and professional profiles.
- **Interface:** Integration with **Model Context Protocol (MCP)** servers to access real-time LinkedIn and Facebook data.

1.2 Key Design Decisions

- **Tool Serialization (Stringification):** A critical architectural decision was the implementation of a **Tool Output Wrapper**. Standard LLM APIs often expect tool responses as strings; however, MCP servers frequently return complex nested JSON objects. To prevent API parsing failures, a middleware layer was developed to serialize all tool outputs into JSON strings before returning them to the model's message history.
- **Three-Tier Score Extraction:** To ensure deterministic reliability in an agentic environment, I designed a fail-safe scoring parser:
    1. **Phase 1 (Structural):** Regex extraction from a pre-defined Markdown block.
    2. **Phase 2 (Global):** Greedy regex search for numerical patterns.
    3. **Phase 3 (Semantic Fallback):** LLM-based recalibration, where the model reviews its own analysis to extract a final reliability coefficient if regex parsing is ambiguous.

- **Recursion Control:** Given the potential for the agent to enter infinite loops when encountering multiple "homonyms" (candidates with identical names), the system implements a strict **recursion limit (50 steps)** and operational constraints within the prompt.

2. Agent Workflow and Tool Usage Strategy

The agent operates through a six-phase verification pipeline, prioritizing professional data sources over social metadata.

2.1 The Six-Phase Workflow

1. **Fact Extraction:** The agent first parses the CV into a structured internal representation, identifying key entities: full name, target university, and recent employment history.
2. **LinkedIn Discovery:** Utilizing search_linkedin_people, the agent executes precise queries. If no match is found, it is programmed to automatically attempt fuzzy matching.
3. **Professional Deep-Dive:** Upon identifying the most probable candidate ID, the agent calls get_linkedin_profile to retrieve the professional "ground truth."
4. **Social Cross-Check:** The agent then pivots to Facebook via search_facebook_users and get_facebook_profile to gather auxiliary evidence (current location, education, interests).
5. **Comparative Synthesis:** The system performs a tri-party comparison (CV vs. LinkedIn vs. Facebook).
6. **Synthesized Reporting:** Final generation of a structured report containing the Credit Score, identified discrepancies, and a verdict.

2.2 Tool Usage & Corroboration Strategy

- **Source Hierarchy:** The design employs a **Hierarchy of Trust**. LinkedIn data is weighted as the primary source for professional credentials. Facebook data is treated as secondary context due to the higher likelihood of stale or incomplete information on social platforms.
- **Corroboration Principle:** A unique heuristic was implemented: if the CV contains an unusual or "odd" fact (e.g., a very short tenure) that is matched exactly by LinkedIn, the agent is instructed to consider this as **strong**

corroboration rather than a discrepancy. This prevents the system from penalizing candidates with non-traditional but legitimate career paths.

3. Sample Verification Results

The system was evaluated against a set of five sample CVs. The agent successfully distinguished between authentic candidates and those with fabricated experiences.

3.1 Quantitative Summary

The system outputted a reliability score for each candidate, where a score > 0.5 represents a **Verified** status and ≤ 0.5 indicates **Suspicious** or **Fraudulent** profiles.

CV Identifier	Agent Reliability Score	Verdict	Final Ground Truth Match
CV_1	0.90	Verified	Match
CV_2	0.95	Verified	Match
CV_3	0.88	Verified	Match
CV_4	0.35	Fraudulent	Match
CV_5	0.20	Fraudulent	Match

3.2 Result Analysis

- **High-Confidence Matches (CV_1, CV_2, CV_3):** The agent identified exact matches on LinkedIn for both education and work history, resulting in high scores. Minor mismatches on Facebook (e.g., location not updated) were correctly down-weighted.
- **Identification of Fraud (CV_4, CV_5):** In these cases, the agent detected severe contradictions. For instance, in CV_5, the candidate claimed a Senior Manager role, but the agent retrieved a LinkedIn profile showing the individual was currently a student, leading to a "Fraudulent" verdict.

4. Conclusion

The developed system demonstrates the efficacy of Agentic AI in complex compliance tasks. By combining a rigid verification workflow with a flexible LLM-based reasoning engine, the system successfully automates the detection of professional misrepresentation with high accuracy and minimal manual intervention.