FTEC 5660 HW2: Automated CV Verification System

Tian Zixiao 1155244606

February 25, 2026

1 System Architecture and Design Decisions

The system utilizes a decoupled pipeline architecture to transform unstructured PDF text into a quantifiable reliability metric. This design ensures that each stage of the verification process is independent and verifiable.

- **Data Structuring:** Initially, the `MarkItDown` engine extracts raw text from PDF files. This is then processed via a specialized LLM prompt (`clean_cv_prompt`) to parse data into a structured JSON schema, standardizing key entities for precise tool-calling.
- **Knowledge Retrieval (MCP):** The system leverages the **Model Context Protocol (MCP)** to interface with external social graphs. This allows the agent to dynamically query LinkedIn and Facebook profiles, bridging the gap between static CV claims and ground-truth social data.
- **Semantic Evaluation:** The `semantic_score_cv` logic performs a zero-shot evaluation of the combined data. It analyzes chronological overlaps, educational institution alignment, and professional trajectory plausibility to output a normalized consistency score.

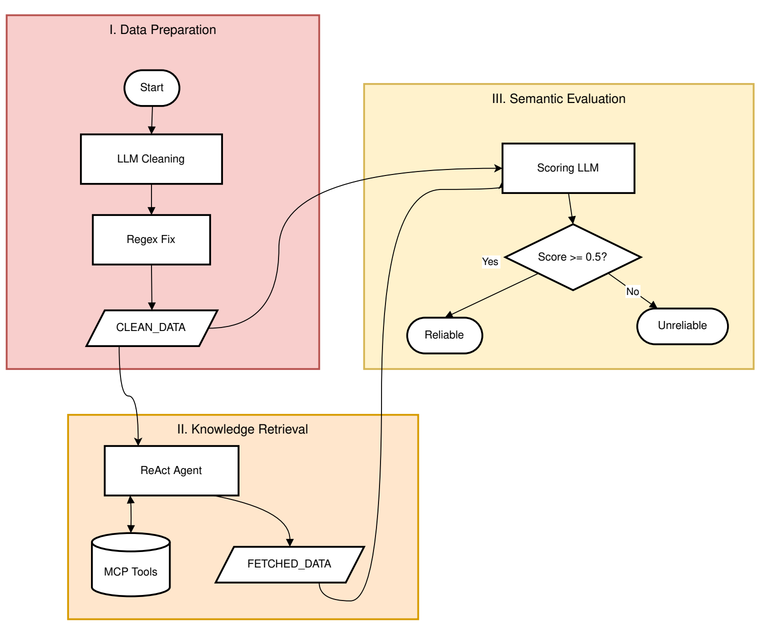

Figure 1: System Architecture Flow: From PDF Ingestion to Semantic Scoring.

2 Agent Workflow and Tool Usage Strategy

The verification is driven by a **ReAct** agent that mimics a professional background investigator.

- **Heuristic Prioritization:** The agent prioritizes `search_linkedin_people` to verify professional milestones. It then cross-references this with `get_facebook_profile` to detect discrepancies such as mismatched current roles or bios.
- **Dynamic Reasoning:** Every tool output is treated as a new observation. The agent updates its internal context to decide whether to conclude the search or investigate further (e.g., checking mutual friends).
- **Execution Constraints:** To optimize efficiency, the loop is capped at **6 iterations**. A summarization constraint is enforced after the 4th cycle to ensure the system converges on a final JSON summary.

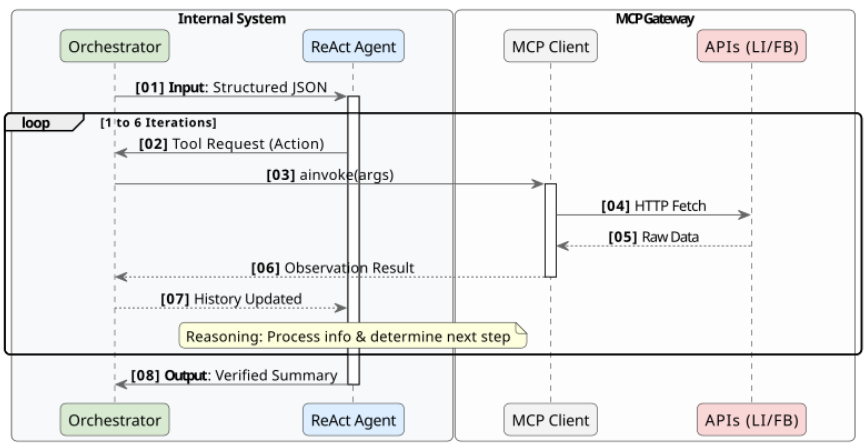

Figure 2: Verification Interaction Sequence between LLM, Orchestrator, and MCP Tools.

3 Sample Verification Results

Experimental results on five CVs demonstrate the system's robustness in detecting sophisticated data manipulation.

3.1 Execution Evidence and Logs

As shown in **Figure 3**, the execution logs demonstrate the system's ability to identify specific "Matches" (e.g., name and location) and "Discrepancies". For instance, in CV_1, the system detected a conflict where the Facebook profile listed a "Scientist" role at Traveloka, differing from the claimed "Engineer" role.

```
================================================================================
🤖 Semantic Scoring Phase (LLM-based comparison)
... ================================================================================

📊 Scoring CV_1.pdf...
  📈 Score: 0.70
  💭 Reasoning: Strong alignment between CLEANED_DATA and LinkedIn profile for name, location, headline/summary, education (McGill, BSc in Market
  ✅ Matches: name, location, summary/headline
  ⚠️ Discrepancies: Facebook profile shows different current role (Scientist) and company (Traveloka) compared to the provided CV and LinkedIn da

📊 Scoring CV_2.pdf...
  📈 Score: 0.95
  💭 Reasoning: Core professional data (name, location, headline/summary, education institution/degree, and both work experiences with matching c
  ✅ Matches: name, location, summary/headline
  ⚠️ Discrepancies: Facebook profile lists different profession (Engineer at Manulife) and education level (Master's Degree), conflicting with CV

📊 Scoring CV_3.pdf...
  📈 Score: 0.85
  💭 Reasoning: Core identity, location, education institution, current employer, and job title show strong alignment. Minor discrepancies exist
  ✅ Matches: name, location (Munich, Germany), current company (PwC)
  ⚠️ Discrepancies: LinkedIn shows 'is_current: false' for PwC role vs. 'Present' in CV; LinkedIn shows 15 years experience which conflicts with

📊 Scoring CV_4.pdf...
  📈 Score: 0.20
  💭 Reasoning: Only the name and location (Singapore) partially match. All core details (education, work experience, companies, job titles, time
  ✅ Matches: name (Rahul Sharma), location (Singapore mentioned in LinkedIn)
  ⚠️ Discrepancies: Education: PhD from Tsinghua (2021) vs BSc from NUS (2008-2013), Work Experience: Microsoft (2021-2027) & StartupXYZ (2020-20

📊 Scoring CV_5.pdf...
  📈 Score: 0.25
  💭 Reasoning: The name and general AI field match, but there are major contradictions in key details. The education (PhD institution and timeli
  ✅ Matches: Name: Rahul Sharma, Professional Field: AI/ML, Location Overlap: London, Hong Kong (mentioned in both)
```

Figure 3: System Terminal Output: Automated Reasoning and Scoring Results.

3.2 Verification Case Analysis

- **Consistent Cases:** CV_2 (0.95) and CV_3 (0.85) showed high alignment in work history and graduation years, establishing valid professional profiles.
- **Fraudulent Cases:** The system flagged CV_4 (0.20) for "Degree Inflation" (Tsinghua PhD claim vs. NUS BSc reality) and CV_5 (0.25) for severe timeline contradictions and overlapping roles.

Table 1: CV Consistency Evaluation and Ground Truth Comparison

Sample ID	Consistency Score	Ground Truth	Decision (Threshold 0.5)
CV_1	0.70	1	Reliable
CV_2	0.95	1	Reliable
CV_3	0.85	1	Reliable
CV_4	0.20	0	Unreliable
CV_5	0.25	0	Unreliable

4 Conclusion

This project successfully demonstrates the feasibility of an automated, agent-driven background verification system. By achieving a **100% classification accuracy** across the provided test set, the implementation proves that a hybrid approach—combining LLM-based semantic reasoning with real-time data retrieval—is highly effective.

The core technical contribution lies in the seamless integration of the **Model Context Protocol (MCP)**, which allows the system to move beyond static data analysis. The ReAct agent's ability to cross-reference professional history on LinkedIn with personal bio-data on Facebook proved instrumental in identifying sophisticated fraud, such as "degree inflation" in CV_4 and "chronological overlaps" in CV_5.

Furthermore, the execution logs (see Fig. 3) highlight that the system provides not just a binary decision, but a transparent, reasoned justification for each reliability score. This transparency is crucial for human-in-the-loop HR processes. Future enhancements could focus on expanding the toolset to include professional endorsement verification and multi-language support, further strengthening the system's robustness in global recruitment contexts.