

Background Investigation Specialist Agent: Architecture and Evaluation Report

1. System Architecture and Design Decisions

Our background investigation system is designed as an automated, tool-calling agent capable of cross-referencing candidate resumes (CVs) against digital footprints.

- **Core Intelligence (LLM):** The system is powered by the gemini-2.5-pro model. This model was selected for its advanced reasoning capabilities and reliable function-calling execution, which are critical for processing unstructured CV text and orchestrating multiple verification searches.
+2
- **Orchestration Framework:** We utilize LangChain to build the agent logic. Specifically, we use `create_tool_calling_agent` and `AgentExecutor` to bind our Large Language Model to external tools and manage the execution state.
- **External Data Integration (MCP):** The system connects to a MultiServerMCPClient (Model Context Protocol) to securely access real-time social graph data. This server exposes crucial investigative tools, including `search_linkedin_people`, `get_linkedin_profile`, `search_facebook_users`, and `get_facebook_profile`.
+4
- **Resilience and Error Handling:** Recognizing that external APIs (like the simulated social graph) are prone to rate limits (e.g., 503 Service Unavailable errors encountered during testing), the system implements an asynchronous retry loop (`max_retries=3`) with a 10-second backoff (`asyncio.sleep(10)`) to ensure robust evaluation.
+1

2. Agent Workflow and Tool Usage Strategy

The agent operates on a strict, prompt-engineered 5-phase workflow to ensure consistent and objective scoring.

Phase 1: Resume Extraction

The agent parses the unstructured CV text to extract key entities: the candidate's Full Name,

Location, Work History, Education, and Core Skills.

Phase 2: LinkedIn Validation (Primary Verification)

LinkedIn serves as the primary ground truth.

- **Tool Usage:** The agent executes search_linkedin_people using the exact full name, prioritizing results based on location or industry.
- **Scoring Matrix:** The agent evaluates the top profiles using a weighted scoring system (e.g., +4 for company match, +3 for university match, +2 for exact city match).
- **Anomaly Handling:** The agent is instructed to treat platform glitches gracefully; for example, a missing end_year (null) combined with is_current: false is classified as a MINOR discrepancy rather than a fabricated role.

Phase 3: Facebook Validation (Secondary Verification)

Facebook is used to cross-reference personal and geographical data.

- **Tool Usage:** The agent uses search_facebook_users filtered by location, followed by get_facebook_profile on likely matches to corroborate details found on LinkedIn.
- **Strategy:** Because not all users have public Facebook profiles, the absence of a profile is treated lightly (maximum deduction of 0.05).

Phase 4: Discrepancy Analysis

The agent catalogs inconsistencies and categorizes them:

- **MINOR:** Typographical errors, slight date shifts, or platform data anomalies.
- **MAJOR:** Unverified education, entirely missing employers, or fabricated roles.

Phase 5: Final Evaluation

The agent calculates a final confidence score between 0.0 and 1.0. It starts with a baseline (e.g., 0.75 if education and one primary role are verified) and applies specific modifiers (penalties for discrepancies, bonuses for perfect multi-platform alignment). The absolute final output is formatted strictly as FINAL_SCORE: <score>.

3. Sample Verification Results

The system was evaluated against 5 sample CVs, using a binary decision threshold of 0.5 (Score > 0.5 = Valid, Score ≤ 0.5 = Invalid). The agent achieved a 100% accuracy rate against the ground truth labels [1, 1, 1, 0, 0].

Here are the final verification results:

- **CV_1.pdf:** Validated. The CV showed near-perfect alignment with the LinkedIn profile, with only a minor platform glitch regarding an employment end date.
 - **Final Score: 0.68**
- **CV_2.pdf:** Validated. A perfect match across all parameters on LinkedIn.
 - **Final Score: 0.75**
- **CV_3.pdf:** Validated. Strong match on education and primary role, with a minor date format discrepancy.
 - **Final Score: 0.68**
- **CV_4.pdf:** Rejected. Major discrepancies found. The identified LinkedIn profile contradicted every significant claim regarding education and employment history.
 - **Final Score: 0.0**
- **CV_5.pdf:** Rejected. While the candidate's identity and PhD were verified via LinkedIn and Facebook, the entire employment history and core skills listed on the CV appeared entirely fabricated.
 - **Final Score: 0.25**