The system employs a **Multi-Stage Hybrid Verification Architecture** designed to automate the background check process by simulating human investigative logic. The workflow consists of five core stages: structured extraction, logical pre-screening, professional verification, identity fallback, and final penalty application. The specific logic is detailed below:

Stage 1: Structured Information Extraction The system initiates by invoking the MarkItDown tool to ingest the PDF CV and convert it into plain text. Subsequently, it utilizes the Gemini Large Language Model (LLM) to perform semantic analysis on the unstructured text. This process precisely extracts three key entities: **"Name"**, **"Target Company"**, and **"Location"**. These entities are transformed into a standardized JSON format, serving as the foundational anchors for all subsequent search and cross-referencing operations.

Stage 2: Hard Fail Logic Flagging (Pre-Screening) Before triggering external API calls, the system executes a rule-based logic scan to detect fatal anomalies that violate common sense. The system incorporates specific keyword detection mechanisms: first, identifying future dates (e.g., "2027"), which violate temporal logic; and second, detecting known fictitious placeholders (e.g., "StartupXYZ"). Upon detecting such anomalies, the system does not immediately terminate execution (to ensure the testing process completes) but instead sets an internal hard_fail_flag to True, flagging the profile for mandatory penalization in the final stage.

Stage 3: Primary Verification — Professional Background Check This is the core phase for assessing the candidate's professional validity. The system constructs a search query in the format of "{Name} {Company}" and invokes the LinkedIn people search tool. To prevent the target profile from being buried by common names, the system adopts a **"Deep Scan" strategy**, setting the search result limit to 15 and enabling fuzzy matching. It then retrieves detailed profiles for each candidate and cross-references their "Experience" list to verify if it contains the target company claimed in the CV. A successful match results in a high confidence score of **0.95**.

Stage 4: Secondary Verification — Identity Existence Fallback If the LinkedIn verification fails to find a professional match (resulting in a confidence score below 0.5), the system activates the Facebook fallback mechanism. This addresses scenarios where a candidate may be a real person lacking a comprehensive professional online presence. The system queries the Facebook user search tool using the candidate's name and validates whether the user's city or country matches the location extracted from the CV. If both the name and location align, the system infers that the candidate's **"physical identity"** is valid despite the professional ambiguity. Consequently, a "passing score" of

0.55 is assigned (just above the 0.5 decision threshold) to prevent misclassifying a real person as a completely fictitious entity.

Stage 5: Final Penalty Application This is the system's final line of defense, designed to handle the anomalies flagged in Stage 2. regardless of the search results generated in the intermediate steps (even if a real person with the same name was found), the system checks the status of the hard_fail_flag. If the flag is set to True—indicating the presence of fatal errors like "2027" or "StartupXYZ"—the system ignores any interim high scores and forcibly overrides the final result to a penalty score of **0.05**. This mechanism ensures that any CV containing logical impossibilities is decisively rejected.