FTEC5660 — Individual Homework 02 (Part 1) Report

Name: <Shiliang Chen>

Student ID: <1155245855>

Course: Agentic AI for Business and FinTech (FTEC5660)

> *Notes:*
>
> - *This report corresponds to the runnable implementation `homework2_part1_submission.py`*

1. System Architecture & Design Decisions

1.1 High-level Architecture

Runtime pipeline

- **Input**: `CV_1.pdf` … `CV_5.pdf` (download from Google Drive folder if not present)
- **PDF → text**: `MarkItDown`
- **LLM extraction**: extract candidate attributes from CV text (name, location hint, current job, education)
- **Evidence collection (MCP)**:
  - LinkedIn: `search_linkedin_people` → `get_linkedin_profile`
  - Facebook: `search_facebook_users` → `get_facebook_profile`
- **Deterministic identity scoring (heuristics)**:
  - compute identity anchors (education/location/email/current job) and confidence
  - handle multi-location CV strings (e.g., `Beijing | Hong Kong`)
  - down-weight generic titles (e.g., "Engineer") to avoid false matches
- **LLM adjudication**:
  - synthesize discrepancies/improvements using CV text + social evidence + identity signals
  - output JSON only
- **Score post-processing (guardrails)**:
  - avoid returning exactly 0.5
  - cap scores when identity is not confirmed / major discrepancies exist
- **Output**: `llm_results.txt` (includes `score_raw` and `score_final` + debug fields)

1.2 Key design decisions (why)

A) Hybrid: heuristics for identity + LLM for reasoning

Pure “LLM-only” scoring tended to be unstable when social profiles were ambiguous or tools were flaky. The final design uses:

- **Heuristics** for identity confidence and disambiguation (repeatable, debuggable)
- **LLM** to generate structured discrepancies/improvements and a base score

B) Conservative scoring to avoid “4/5 too high”

An earlier failure mode was a **too-generous default** (e.g., “evidence insufficient” → 0.55) combined with threshold=0.5, which caused negatives to be predicted positive.

Mitigation implemented in code:

- If **major discrepancy** exists → cap `score_final ≤ 0.45`
- If **social identity not confirmed** and confidence is low → cap `score_final ≤ 0.49`
- If identity is partially supported but still uncertain → cap into a cautious band (≤ 0.60)

C) Dual-channel identity confirmation (LinkedIn + Facebook)

If LinkedIn identity confidence is low but Facebook matches strongly (or vice versa), the system uses the **more confident channel** as `social_identity_*`. This prevents false negatives when one platform is missing/outdated.

D) MCP robustness

Because the MCP endpoint is served behind ngrok and can be unstable, the implementation:

- disables proxy env usage in httpx (`trust_env=False`)
- retries network calls with exponential backoff
- probes both `http://` and `https://` scheme variants when loading tools

2. Agent workflow & tool usage strategy

2.1 Workflow overview (per CV)

- **(1) Extract** candidate attributes from CV text via a constrained JSON prompt.
- **(2) LinkedIn evidence**
    - search by name (+ optional location)
    - pull multiple candidate profiles (top-K) and re-rank using:
        - headline/location weak signals
        - current job similarity (company/title)
        - education token overlap
        - down-weight generic titles to reduce false matches

- **(3) Facebook evidence**
  - search users by name, fetch top profiles, re-rank similarly (location/current job/education)
- **(4) Identity scoring**
  - compute per-platform identity confidence and `social_identity_*` (max across channels)
- **(5) Adjudication**
  - ask LLM to output final JSON: `score`, ids, `discrepancies`, `improvements`
  - apply deterministic post-processing to obtain `score_final`

2.2 Tool usage strategy

- **LinkedIn**
  - `search_linkedin_people(q, location?, fuzzy=True)` to get a candidate set
  - `get_linkedin_profile(person_id)` to validate education / timeline / current job
- **Facebook**
  - `search_facebook_users(q, fuzzy=True)` to find candidates
  - `get_facebook_profile(user_id)` as supplementary evidence (often incomplete/outdated)

2.3 Failure modes & mitigation

- **Tool call failure / timeouts**: continue and record limitations in `improvements`; scoring becomes conservative due to identity uncertainty.
- **Same-name collisions**: require stronger anchors (education+location, or education+current job) to confirm identity; otherwise mark for manual review.
- **Outdated profiles**: treat timeline mismatches as "improvements" unless identity is strongly confirmed and the time periods overlap.

3. Sample verification results (final result on sample CV)

This section lists the final results for **CV_1 … CV_5** (from the latest `llm_results.txt`).

CV_1.pdf

- **score_final**: 0.65
- **Matched**: LinkedIn `person_id=9`, Facebook `user_id=213`
- **Discrepancies**: none
- **Improvements (examples)**:
  - LinkedIn headline differs from CV title (possible different self-description / profile not updated)
  - Facebook current company/title differs from CV (low-confidence FB match → manual review)

CV_2.pdf

- **score_final**: 0.75
- **Matched**: LinkedIn `person_id=47`, Facebook `user_id=180`
- **Discrepancies**: none
- **Improvements**: none

CV_3.pdf

- **score_final**: 0.80
- **Matched**: LinkedIn `person_id=97`, Facebook `user_id=97`
- **Discrepancies**: none
- **Improvements**: none

CV_4.pdf

- **score_final**: 0.49
- **Matched**: LinkedIn `person_id=null`, Facebook `user_id=823`
- **Discrepancies**: none (identity not confirmed)
- **Improvements (examples)**:
  - Facebook current company/title do not match CV; education appears lower than CV claim → needs manual identity review
  - LinkedIn: unable to select a profile due to ambiguity / low confidence

CV_5.pdf

- **score_final**: 0.45
- **Matched**: LinkedIn `person_id=95`, Facebook `user_id=765`
- **Discrepancies**:
  - **major**: education level mismatch (CV: PhD vs social: MSc)
  - minor: current title mismatch (Senior Engineer vs Engineer)
  - minor: role/title wording mismatch on DataForge (Senior Analyst vs Analyst)
- **Improvements (examples)**:
  - Facebook current company/title differs from CV (low-confidence FB match → manual review)