

CV VERIFICATION SYSTEM USING MCP-BASED SOCIAL GRAPH TOOLS

RUIQIAN LUO(1155244663)

1. INTRODUCTION

This project builds an agentic AI system for CV verification using an MCP (Model Context Protocol) server. The goal is to automatically verify candidate CV claims against public social media profiles (LinkedIn and Facebook), detect inconsistencies, and output a reliability score for each CV. The system processes a batch of CV PDFs, extracts candidate claims (e.g., name, roles, locations), queries the MCP SocialGraph tools, retrieves matching social profiles, and produces a confidence score in $[0,1]$ indicating whether each CV is likely valid.

2. METHOD

- (1) Load CV PDFs and convert them to text.
- (2) Extract candidate claims (mainly name / roles / locations).
- (3) Search LinkedIn and Facebook using MCP tools.
- (4) Retrieve the best matched profile from each platform.
- (5) Use an LLM to judge consistency and return a score + short reason.
- (6) Apply a small modification (e.g., exact match bonus, location support bonus).

Design choice:

- **LinkedIn = primary evidence** (professional identity, industry, experience, location)
- **Facebook = secondary evidence** (identity/location consistency)

This avoids over-penalizing a CV when Facebook is noisy but LinkedIn strongly supports the claims.

3. MCP TOOLS USED

I tested all 6 MCP tools, and the main verification pipeline uses the four core ones:

- `search_linkedin_people`
- `get_linkedin_profile`
- `search_facebook_users`
- `get_facebook_profile`

The remaining tools (`get_linkedin_interactions`, `get_facebook_mutual_friends`) were tested but not heavily used in final scoring.

4. RESULT

An example run of my final system produced:

scores = $[0.635, 0.815, 0.41, 0.455, 0.41]$

Using threshold 0.5, the predicted labels are:

$[1, 1, 0, 0, 0]$

For the provided example ground truth:

[1, 1, 1, 0, 0]

This gives:

$$\text{Final Score} = \frac{4}{5} = 0.8$$

5. DISCUSSION

The system works well when LinkedIn provides strong professional evidence and the candidate name matches clearly. It performs worse when:

- CV text extraction is hard to identify,
- names are similar,
- claims include many unsupported roles/locations,
- social profiles are incomplete or inconsistent.

Possible improvements:

- better claim extraction (education, dates, company names),
- stronger candidate re-ranking,
- more structured use of interaction / social graph consistency signals.

6. CONCLUSION

I implemented a working MCP-based CV verification system that:

- processes CV PDFs,
- searches LinkedIn/Facebook profiles through the provided MCP server,
- compares CV claims with profile evidence,
- outputs a confidence score for each CV.