

QA Test Report AI Powered CV Analysis Tool

Candidate: Gabriel Samuel
Date: 3/02/2025
Project: AI Powered Resume Parsing Validation
Objective: Identify and report all discrepancies between the JSON output and the original PDF resume to ensure data accuracy, completeness, and consistency.

Summary of Findings

After a detailed and critical analysis, I have identified a total of **26 discrepancies** across multiple categories, including name inconsistencies, phone number errors, missing training dates, job description truncations, skill duplication, and formatting issues

Detailed Discrepancy Table

Category	Error Type	JSON Output	PDF Document	Recommendation
Personal Information	Name Spelling Error	"name": "John Chigwarira"	"John Chigwarira"	Implement string similarity checks for name validation.
Contact Details	Phone Number Mismatch	("+40) 724344620	+40 724334620	Apply digit validation algorithms to detect minor number differences.
Email Formatting	Capitalization Difference	"emails": ["chigwarira@yahoo.com"]	"Chigwarira@yahoo.com"	Normalize email case formatting to prevent case sensitivity issues.
Location	"Home" missing	Sibiu, Romania (Home)	Sibiu, Romania	Add "Home" location detail in JSON
Location Information	Incorrect Country Code	"country_code": "RM"	"Romania"	Use ISO 3166 standard country codes (RO instead of RM).
Work Experience	Incorrect Employment Status	"is_current": true (DK Recruitment, Dec 2023 – Nov 2024)	"Dec 2023 – Nov 2024" (Job should have ended)	Compare employment end dates with the current date for automatic updates.

Category	Error Type	JSON Output	PDF Document	Recommendation
Work Experience	Inconsistent Location for CCPorter	"locations": [{"country": "Poland"}, {"country": "Romania"}]	"Remote, Poland"	Ensure remote roles correctly indicate the primary country of operation.
Work Experience	Work Experience	"description": "Communicate with customers... Maintain a working knowledge..."	"Communicate with customers... Use data to help customers achieve goals."	AI should ensure full extraction of job descriptions without truncation.
Education	Typographical Error	"degree": "Collage Graduate"	"College Graduate"	Apply spell-checking logic for academic terms.
Education	Website Formatting Error	"website": "https://www.ulbsibiu.ro/](https://www.ulbsibiu.ro/)"	"https://www.ulbsibiu.ro/"	Fix improperly formatted URLs in JSON extraction.
Education And Training		"institution": "Colegiul Tehnic Energetic",	"High School Graduate: Informatics Colegiul Tehnic Energetic"	
Skills	Missing Categorization	"soft_skills": ["Problem-solving", "Negotiation", "Research"]	"General Skills: Proactive, Creative, Sociable"	Improve AI categorization accuracy for soft and hard skills.
Skills	Skill Duplication	"soft_skills": ["Communication", "Interpersonal Skills", "Communication"]	"Communication, Teamwork, Interpersonal Skills"	Remove duplicate skills in AI extraction to avoid redundancy.
Certifications & Training	Missing Training Dates	"courses": ["MediaData training", "KMS Lighthouse training"]	"MediaData training (Oct 2022 – Nov 2022)"	AI should extract and map training program dates explicitly.

Category	Error Type	JSON Output	PDF Document	Recommendation
Languages	Different Proficiency Levels in English	"proficiency": "C2" (English)	"Spoken Production: C1, Writing: C1"	Extract detailed subcategories instead of generalized proficiency levels.
Languages	Incorrect Arabic Proficiency Extraction	"proficiency": "A2"	"Spoken Production: A1, Writing: A1"	Ensure granular extraction of language skills.
Languages (German)	JSON simplified proficiency levels	B1	B1 Listening, B1 Speaking, B1 Reading, B1 Writing	Preserve full proficiency details for accuracy
Metadata Structure	Data Ordering Inconsistency	"experiences" listed before "education"	"Education" appears before "Work Experience"	Ensure consistent data ordering to match standard resume format.
General skills	The JSON file introduces new skills that are not explicitly listed in the PDF under General Skills.	the JSON file includes additional skills that seem to be classified as Hard Skills, such as: Communication, Interpersonal skills, Teamwork, Initiative, Problem-solving, Negotiation, Relationship building, Assessment, Adaptability, Attention to detail, Research, Collaboration.	The PDF file lists only General Skills, including: Proactive, creative, adaptable, capable of active listening, result-driven, detail-oriented, effective communicator, patient, computer-literate, sociable.	Implement a validation mechanism to filter only explicitly mentioned skills
General Skills	Inconsistency	formats the first alphabets of each skill to be Upper case	Alphabets are all in lowercase except the first skill	

Category	Error Type	JSON Output	PDF Document	Recommendation
General Skills	Inconsistency	"Problem-solving",	"problem-solving abilities"	
Sales Consultant/Customer Retention (CCPorter)	JSON added an extra location (Romania)	Feb 2020 – Dec 2022 (Remote, Poland)	Feb 2020 – Dec 2022 (Remote, Poland & Romania)	Improve location extraction to avoid incorrect additional locations
Honours And Awards	Typographical Error	"title": "Graduate of the training course for the sales department of REPORTER Sp. z o.o.",	"Graduate of the training course for the sales department of CCPORTER Sp. z o.o."	
Formatting Issue	Extra Space in Some Extracted Values	"job_title": "External Recruiter"	"External Recruiter"	Remove leading/trailing spaces from extracted data.
Training Program	Typographical Error	"title": "MediaData training course",	"MediaData training coruse"	
Training Program	Typographical Error	"title": "Sales executive training course for the sales department of REPORTER Sp. z o.o.",	Sales executive training course for the sales department of CCPORTER Sp. z o.o.	