



Why Can't I Get A Job In AI Testing?

Introduction:

As companies race to put AI into every workflow, they are running head-first into failures, inconsistencies, and unpredictable behaviour. The speed of adoption has massively outpaced the development of testing standards, evaluation roles, or even basic organisational understanding of how these systems behave. The result is an overwhelming need for AI testing — and a complete lack of clarity about who is supposed to do it.

Look on any job site and you will see hundreds of roles that are, in reality, testing jobs. They appear under dozens of different titles, including AI Safety Analyst, prompt tester, annotator and many more.

These are all testing roles. But they are not being filled.

And it is not because people like you can't do the work. It is not because there is a shortage of available positions. It is not because the talent pool is too small. It is not because the work is too advanced.

The truth is far deeper - and far, far stupider.

The industry has created a situation where the work exists, the demand exists, the failures are everywhere, and yet the hiring pipeline is fundamentally broken. Companies don't know how to describe the roles they need. Recruiters don't know how to filter for the skills required. Applicants don't know which job titles map to the work they can already do. And the people who understand the testing problem are rarely the ones writing the job descriptions.

Section 1. Hiring is Broken

1.1 Job Titles

The job titles are incoherent. Companies are advertising AI testing roles under dozens of unrelated names, and none of them match what the work actually is. Without a standard job title or description, hiring is broken before it begins. These titles also do not fit the filters used by most job boards, and many job boards do not even list AI as an industry category. This means the roles are effectively hidden from both applicants and recruiters. These titles include:

- AI Safety Analyst
- LLM Evaluator
- AI Red Team Specialist
- AI Product Specialist
- AI Content Reviewer
- Model Evaluation Contractor
- AI Annotator
- AI Quality Analyst
- AI Operations Analyst
- AI UX Researcher
- Prompt Engineer
- AI Behaviour Analyst

All these jobs are essentially for the same role. The work is real, but the naming is chaotic. Job boards cannot classify it, recruiters cannot filter for it, and applicants cannot search for it. The hiring pipeline collapses before anyone even applies.

1.2 Recruiters

Recruiters don't understand the work. Recruiters are filtering for keywords they recognise, not the skills the job requires. They look for degrees, frameworks, and years of experience in a field that did not exist. They reject candidates who can do the work because:

- they don't recognise the terminology
- they don't understand evaluation or behavioural testing

- they assume AI testing is software QA
- they filter out anyone without a formal AI background

The gatekeepers don't understand the gate!

1.3 Companies

Companies don't know how to describe the role Hiring managers know they need AI testing, but they don't know how to define it. They write job descriptions that mix unrelated responsibilities, contradictory requirements, and vague expectations. This leads to:

- roles that sound like three jobs combined
- responsibilities that don't match the title
- requirements that don't match the work
- descriptions written by people who have never tested an AI.

The result is confusion on both sides.

1.4 Skills

The skills invisible AI testing requires logical thinking, attention to detail, an enquiring mind, and a modest amount of patience. These are the skills that uncover reasoning failures, behavioural inconsistencies, and subtle model breakdowns. But organisations do not recognise these as "real" skills because they are still treating AI like a traditional IT system. Developer teams fall back on familiar noise like Java, JSON, Python, and framework experience, even though none of these reveal how an AI actually behaves.

This hides qualified candidates because:

- the skills that matter are not the ones companies search for
- logical reasoning is undervalued compared to technical keywords
- attention to detail is ignored in favour of tool familiarity
- enquiring thinking is not listed as a requirement
- patience and persistence are not considered technical skills
- developer teams are still using IT testing methods that do not work on AI

A person with clear thinking and behavioural insight will test an AI far deeper than someone with ten years of Python experience. But the hiring filters are built to detect the Python, not the insight.

1.5 The Profession Does Not Exist Yet

The work is real, but the category doesn't exist. AI testing is essential, but it is not formally defined as a profession. There are no standard titles, no standard qualifications, and no standard hiring pathways. There are no junior roles that can grow into the senior-sounding roles being advertised.

This creates a paradox:

- the work exists
- the demand exists
- the failures are everywhere
- the job category does not officially exist

You can be qualified for a job that the industry is still arguing over naming and is still yet to even define!

1.6 Whose Responsibility Is It?

At the moment within most companies, everyone assumes someone else is responsible. Inside organisations, AI testing is done half-heartedly because nobody knows who is supposed to do it. There is no clear indication of who should test what. Is it the developers, the end users, the QA team, or someone else entirely? Each group probably assumes another group is handling the testing, and as a result the work is fragmented, inconsistent, or simply not done at all.

This leads to:

- developers assuming QA will catch the issues
- QA assuming developers understand the model better
- end users assuming the system was already validated
- managers assuming the teams have it covered
- leadership assuming testing is a solved problem.

With no ownership, no structure, and no defined responsibilities, AI testing falls into a responsibility vacuum where everyone touches it but nobody owns it. That's partly why you get incoherent job descriptions asking for unicorn employees who will never be hired.

1.7 The Stupidity At The Core

Part of the stupidity is the complete lack of methodology around AI testing. Companies are deploying AI into critical workflows without any agreed-upon way to evaluate behaviour, measure reasoning, or document failures. INQUISITOR LABS is one of the first to propose a behavioural testing methodology - the LLM INQUISITOR METHODOLOGY - to plug this gap and give organisations a structured way to test, document, and understand model behaviour. Companies are crying out for standards, but inertia is great, confusion is everywhere, and panic is beginning to spread as failures increase. The industry will have to change, because without a methodology, the entire testing pipeline collapses.

This leaves organisations with:

- no shared language for testing
- no standard expectations
- no consistent documentation
- no behavioural evaluation framework
- no way to compare results across teams
- no way to prove liability or defend decisions

The stupidity is not the lack of talent or the lack of jobs. It is the lack of a defined, recognised, behavioural testing methodology. LLM INQUISITOR provides the structure the industry has been missing, and the change that companies are now being forced to adopt.

Conclusion

We can't really give anyone practical advice, because the truth is that the industry itself is not ready. We can end on a positive note, but it has to be an honest one: the industry will have to adapt or die. With AI still in the infancy of what is possible, the AI testing industry will have to rise to the challenge. The need is real, the failures are increasing, and the pressure is building. Change is coming whether the industry wants it or not.

But that still does not get you a job. The roles that are easy to come by are the low-paid online testing mills, whose work is soul-destroying and largely meaningless in terms of

valid test results. They do not build real skills, they do not reflect real evaluation work, and they do not prepare anyone for the behavioural testing the industry actually needs.

Large companies are probably the most corporately inert. With the jobs market in the UK and USA, as well as many other countries, largely broken, the disconnect between candidates and hiring requirements is already severe. The lack of joined-up thinking in AI testing makes it even worse. Companies want testing, but they do not know how to hire for it. Candidates can do the work, but they cannot match the keywords. Recruiters do not understand the roles, and the roles do not match the work.

All we can promise is that things will change. The industry will be forced to evolve, and new pathways will open. You may have to find an indirect route into the field, because the direct route does not exist yet. Not much help, I know, but at least you know the truth.

It's not you. It's them.

Good luck!

Links:

Inquisitor Labs Homepage:

<https://assimilatedhuman.github.io/inquisitor-labs/index.html>

-Document Ends-