

Implementation Consultant (AI SaaS)

evaluagent | Remote, US-based | Full-time

Location	Remote, US East Coast (Eastern Time preferred; Central considered)
Reports to	Delivery Team Lead
Work authorization	Authorized to work in the US without sponsorship

About the role

evaluagent is scaling its US delivery function as our customer base grows, and getting customers live and to value quickly matters more with every new account. This is a new role, created to expand our delivery team and deepen the expertise we bring to every engagement. We are looking for an Implementation Consultant focused on getting customers successfully onto evaluagent and, specifically, on helping them make the shift from legacy quality practices to AI-powered quality. You will report to our Delivery Team Lead.

Most of the customers you will work with have spent years building scorecards and line items designed for a human to read and apply judgment to. They run calibration sessions, argue about edge cases, and carry a hard-won sense of what a good interaction looks like in their world, whether that is a regulated collections call or a first-line support chat. Automated quality works differently. Getting repeatable, accurate, and defensible results from an LLM depends on how the requirement is expressed to it. The core of this role is translating a customer's existing quality framework, their business goals, and the way they think about a "good" interaction into prompts and configurations that produce reliable outcomes in our platform and, critically, then teaching the customer to do this for themselves.

You will be the person who can sit with a head of quality or a contact center operations lead, understand what they are trying to measure and why, and turn that into something the product delivers consistently. You will run training and working sessions, scope proof-of-concept and onboarding engagements alongside the wider team and carry the customer's point of view back to product and engineering when it needs to shape how a feature is built.

You will also need to defend how we do this. Customers come to us from manual QA, or from older automated tools built on keyword spotting and rigid rules. They often expect a machine to behave like a deterministic rulebook: the same input, the same output, every time. Our approach is different, and it gives better results. An LLM evaluates an interaction the way a good human assessor does, against intent and context rather than the presence of a phrase. You need to be fluent in why that is

more accurate and more defensible than brittle rules, and honest about where variation comes from and how we keep it under control, through clear requirements, golden datasets, and calibration. Holding that line with a skeptical head of quality, and bringing them with you, is central to the job.

This is not a project management role. The sequencing, prioritization, and resourcing of delivery work is owned elsewhere on the team. It is also not an engineering role. You will not be writing production code. You sit between the two: someone who understands the underlying technology well enough to be credible, who represents the user and the customer, and whose craft is in getting the best possible outcome out of the tool through clear thinking, clear writing, and rigorous attention to detail.

This is a US-based, remote role. Our sales, proof-of-concept, and delivery activity is increasingly US-centric, and we want this person close to those customers for time-zone coverage and day-to-day responsiveness, with some overlap into UK working hours expected given the wider team is UK-based.

What you will own

Customer implementation and enablement

- Lead customer engagements across the full range of delivery work: new-customer onboarding, proof-of-concept engagements for both prospects and existing customers, expansion and upgrade PoCs, migration onboardings, evaluations, and re-onboardings, from scoping through to a customer who is live, confident, and self-sufficient.
- Run customer training sessions, Q&A, and working sessions with quality, CX, and operations stakeholders.
- Define and teach the process by which a customer builds their own business practices and goals into the platform, rather than building it for them.

Translating quality into the platform

- Translate customers' legacy, human-oriented quality frameworks (scorecards, line items, evaluation criteria) into LLM-oriented prompts and configurations that deliver consistent, accurate, and defensible results.
- Build reusable assets (prompt patterns, templates, best-practice guidance) that raise the quality and speed of every future engagement.

Cross-functional partnership and advocacy

- Work closely with Technical Success, Sales, and Customer Success to scope engagements, agree what is and is not in scope before work begins, and hand customers on cleanly. Getting these handoffs right is a big part of the role's success.
- Act as an informed customer advocate to product and engineering, challenging the approach where the user or customer perspective needs to be represented.

What you bring

Required

- Excellent written and verbal communication and genuinely high attention to detail. This role lives and dies on the precision of how requirements are expressed, so this is non-negotiable.
- A structured, analytical thinker who can take an ambiguous business goal and break it down into clear, testable logic.
- Fluency with AI and large language models: a working understanding of how LLM-based evaluation produces its results, sound judgment about how to express a requirement so the model behaves reliably, and the ability to reason and speak credibly about accuracy, variation, and defensibility. This is about conceptual fluency and prompting judgment, not writing production code.
- Comfort with technology and a willingness to get into the detail of how an AI-powered product produces its results. You should be the kind of person who wants to understand why something works, not just that it does.
- Experience working directly with customers in an implementation, onboarding, solutions, or consulting capacity.
- Confidence running training and working sessions and holding the room with senior customer stakeholders.
- A bias toward enabling customers to succeed on their own rather than doing the work for them.
- A genuine collaborator who works well across Technical Success, Sales, Customer Success, and Product, and understands that getting these relationships right is central to delivering for the customer.
- Someone who aims for great rather than settling for what exists today. We are at a stage where we want people who care, who hold a high bar, and who actively help us get better.

Bonus

- Experience at a CCaaS, workforce engagement, quality management, or adjacent contact center technology vendor. Domain knowledge of contact center quality is a strong advantage, though we will train the right candidate.
- Hands-on familiarity with writing prompts for large language models, or a demonstrable interest in how to get reliable outputs from them.
- Experience helping customers move from a manual or legacy process to an automated one.
- Familiarity with integrations and data flows between contact center platforms, useful for scoping, though deep technical configuration is not the core of this role.

What success looks like

We measure this role on the quality of outcomes, not on activity or throughput.

First 90 days

- Ramped on the platform and the team's delivery methodology, and running your first customer engagements with support.
- A clear grasp of how our most common customer types think about quality, and how their legacy frameworks translate into LLM-driven configurations.
- First reusable prompt patterns and templates started, drawn from the engagements you have run.
- Strong working relationships established with Technical Success, Sales, and Customer Success.

6 to 12 months

- Customers go live with quality configurations that produce consistent, accurate, and defensible results they trust and can stand behind.
- Customers can build and adjust their own evaluations after working with you, rather than depending on us to do it for them.
- Skeptical quality leaders are brought along and become confident advocates for the AI-powered approach.
- The reusable assets and patterns you build make every later engagement faster and better, for you and for the rest of the team.

Why this role matters

evaluagent's US base is growing, and the hardest part of every engagement is the same: moving a customer from a quality program built for human assessors to one that runs reliably on an LLM. Done well, automated quality is faster, more consistent, and more defensible than what came before. Done without a guide, customers stall, lose trust in the results, and fall back on old habits.

This role is that guide. You turn a customer's hard-won sense of what good looks like into configurations that hold up, you teach them to run it themselves, and you bring skeptical quality leaders with you rather than around them. That is how a defensible, AI-powered approach becomes a reason customers choose us and stay.

It also makes the delivery team stronger. The patterns and assets you build raise the standard of every engagement that follows, so customer outcomes depend on a repeatable craft rather than on any single person.

About evaluagent

evaluagent is a contact center quality and performance platform that helps customer experience teams understand every conversation, coach their agents, and prove the impact of quality on business outcomes. We are at the leading edge of applying large language models to quality management, moving our customers from manual, human-scored evaluations to automated, AI-powered quality that is faster, more consistent, and more defensible.

We work with mid-market and enterprise organizations, and we are a remote-first company with team members across the US and UK.