

A large, dark purple circular graphic on the left side of the page, partially overlapping the job title.

JOB TITLE:

AI Trust & Benchmarking Intern

Company: Avaly.ai

Industry: AI Security, Cybersecurity

Job Type: Full-time Internship (Remote)

Start Date: Rolling Applications – Apply Soon


About Avaly.ai

Avaly.ai is pioneering the future of cybersecurity in the AI-first world. Our mission is to protect individuals, organizations, and AI agents from emerging security risks. Founded by experts in AI and cybersecurity, we deliver trusted, resilient protection through advanced algorithms, holistic risk frameworks, and ethical AI design. Our goal is to ensure that organizations can confidently deploy AI securely and responsibly.

Our Culture

We're looking for passionate interns who think beyond the classroom—trailblazers who've explored open-source tools, contributed to GitHub projects, published academic or technical research, or built something unique on their own. Whether it's a personal AI project, a hackathon win, or an innovative class assignment, we value hands-on curiosity and initiative.

Why Join Avaly.ai?

A large, dark purple circular graphic on the right side of the page, partially overlapping the 'Why Join Avaly.ai?' section.

At Avaly.ai, you'll have the opportunity to work alongside top experts in AI and cybersecurity, gaining hands-on experience through real-world projects that matter. You'll be immersed in the most current industry practices in AI security, working with cutting-edge tools and technologies that go far beyond what's typically covered in the classroom. This internship is designed to help you build a strong, practical portfolio that will set you apart and prepare you for your next big opportunity. You'll also have direct access to leadership and mentors who are deeply invested in your growth, and contribute to projects aligned with the latest real-world threats, frameworks, and AI security standards.

Key Responsibilities:

- Assist in developing and executing trust scoring frameworks for benchmarking various Large Language Models (LLMs) and AI platforms.
- Support assessments on AI trustworthiness dimensions including fairness, transparency, explainability, privacy, and robustness.
- Collect, manage, and analyze data from benchmarking tests, compiling results and generating meaningful reports
- Conduct research and literature reviews on AI trustworthiness metrics, industry standards, and best practices
- Collaborate closely with Avaly.ai's team to refine assessment methodologies and reporting templates.
- Document benchmarking processes, assessment criteria, and results for internal use and client-facing reports.
- Help automate benchmarking processes where applicable, streamlining the scoring and assessment workflow.

Required Qualifications:

- Currently pursuing or recently completed a Bachelor's or Master's degree in Computer Science, Data Science, Information Systems, Business Analytics, or related fields.
- Strong analytical and quantitative skills with experience in data collection and analysis.
- Familiarity with AI and machine learning fundamentals, especially Large Language Models (e.g., GPT-based systems).
- Basic programming skills (Python preferred) for data manipulation, analysis, and visualization.
- Excellent research and documentation skills.
- Good understanding of ethics, bias, fairness, and transparency issues in AI.
- Ability to communicate technical findings clearly to both technical and non-technical audiences.

Nice-to-Have Qualifications:

- Experience or coursework related to AI ethics, trustworthy AI, and AI governance frameworks
- Familiarity with AI evaluation metrics and benchmarking tools.
- Experience with data visualization tools such as Tableau, Power BI, or similar.
- Previous internship or research experience related to AI, trustworthiness assessment, or technology benchmarking.
- Exposure to regulatory frameworks and standards (e.g., EU AI Act, NIST AI Risk Management Framework).
- Experience or familiarity with automated testing and evaluation frameworks.

Ready to shape the future of AI Security?

To apply, please include the job title you're applying for in the subject line of your email and attach your resume.

We also encourage you to include a short cover letter or personal note that highlights why you're passionate about securing intelligent systems and what makes you a strong candidate for the role. If available, feel free to share your GitHub profile, LinkedIn, or links to relevant projects. We're excited to hear from you!

✉ **Send an email to careers@avalay.ai**

