



Salesforce Trusted AI Resources

Salesforce is committed to the responsible development of its AI Services, prioritizing safety, transparency, accuracy, privacy and security. Salesforce makes tools and guidance available to assist customers in safety testing and output monitoring as set forth in the Trust and Compliance Documentation. We are providing you with materials that are Documentation (as that term is defined in your Main Services Agreement with Salesforce), as well as more general information that you may find useful about Salesforce's approach to AI in the Section below entitled Informational Resources.

Documentation

- [ISO 42001 Certification](#)
 - ISO/IEC 42001 is the world's first international standard for Artificial Intelligence Management Systems (AIMS), providing a structured framework for organizations to establish, implement, maintain, and continually improve responsible AI practices. It is designed for entities that develop or use AI-based products and services, ensuring ethical, transparent, and accountable AI operations.
- [Salesforce-owned Model Cards](#)
 - Model cards provide documentation on how trained machine learning (ML) and artificial intelligence (AI) models perform. They provide information about how our models work, including their inputs, what they output, the conditions under which models work best, and ethical considerations in their use.
- Relevant [Salesforce Help](#) articles. More information about specific AI services and features is available at the applicable Help & Training Documentation pages.
 - [Agentforce Testing Center](#)
 - [Einstein and Data Usage](#)
 - [Einstein Global Model Opt Out](#)
 - [Large Language Model Support](#)
 - [Manage Salesforce Access to Customer Data](#)
 - [Manage Model Provider Access](#)
 - [Monitor Agent Guardrails](#)
 - [Monitor Your Agent](#)
 - [Slack Privacy Principles](#)
 - [Trust Layer](#)



Informational Resources

- [Office of Ethical and Humane Use webpage](#)
 - Information about Salesforce's approach to responsible AI practices.
- Trusted AI and Agents Impact Reports
 - Salesforce's Trusted AI and Agents Impact Report details progress on trusted AI initiatives across the company.
 - [FY26 Trusted AI and Agents Impact Report](#)
 - [FY25 Trusted AI and Agents Impact Report](#)
- [Salesforce Stakeholder Impact Reports](#)
 - Salesforce's Stakeholder Impact Reports provide information on the company's performance and impact across environmental, social, and governance (ESG) metrics. These reports serve as a resource for customers and investors to understand the company's commitment to responsible business practices.
- [Generative AI Benchmark for CRM](#)
 - Salesforce has developed the world's first agentic benchmark for CRM to assess the efficacy of AI agents and models for business applications. This benchmark evaluates agents across enterprise use cases on key metrics including accuracy, cost, speed, trust & safety and sustainability based on real CRM data and expert evaluations. .
- [Trusted Agentic AI Trailhead](#)
 - Learn how Agentforce uses safeguards and responsible AI principles to create ethical AI on Trailhead, Salesforce's free e-learning platform.
- [Office of Ethical and Humane Use blog content](#)
 - Read responsible AI thought leadership from Salesforce's Office of Ethical and Humane Use.