Privacy Engineering Domains

Data Scientist (including AI)

By the 2024-2025 IAPP Privacy Engineering Section Advisory Board

"I turn data into valuable insights that drive business strategies and decision-making. However, I often work with sensitive personal information, making privacy a crucial element in my role. I need to ensure that I'm balancing the utility of data with strong privacy practices to protect individuals' rights and build trust in our data-driven solutions."

- Data Scientist

Tasks

Data analysis and modeling: Extract insights only using necessary, proportionate data, ensuring privacy compliance throughout analysis and modelling.

Privacy-preserving techniques: Apply privacy-enhancing technologies like differential privacy, anonymization, aggregation and federated learning to protect data.

Privacy impact assessments: Conduct assessments during the planning and design phases to evaluate potential privacy impacts and identify necessary mitigations.

Govern data use and provenance: Process data for its intended purpose, manage its lifecycle and track consent and provenance to ensure ethical reuse.

Ensure fairness and protect sensitive data: Identify and address bias risks in AI models and safeguard against unintended inference of sensitive data.

Collaboration: Work closely with privacy engineers, legal and compliance teams to align data activities with privacy policies and standards.

Professional profile

Technical competencies: Proficiency in statistical analysis, machine learning, data anonymization, encryption and data lifecycle management

Areas of experience: Programming, data science, algorithm development, artificial intelligence, data engineering and cloud-based analytics

AI lifecycle experience: Active across all stages: planning, design, training, evaluation, implementation, deployment, online learning, post-deployment training and maintenance

Privacy tools: Familiarity with privacy-preserving technologies, such as federated learning, homomorphic encryption and synthetic data generation

Privacy certifications: Certifications like the Certified Information Privacy Technologist or other data protection credentials to enhance privacy expertise

In the organization

Reports to: Chief data officer, head of AI or chief technology officer

Cross-functional collaboration: Works with: privacy engineers, UX designers, legal teams and product managers to ensure privacy is maintained throughout the AI development process

Key stakeholders: AI product, business operations, product development and marketing teams

Strategic drivers

Privacy by design: Embed privacy principles in every step of the data analysis process, from data collection to the deployment of models.

Transparency and accountability: Maintain transparency in data use and establish accountability mechanisms to uphold privacy commitments.

Ethical data usage: Ensure data models, including AI, are fair, transparent and respectful of individual privacy and societal norms.

Regulatory adherence: Stay compliant with evolving privacy laws and standards to avoid legal repercussions and enhance business reputation.

Tools and resources

Privacy-preserving technologies: Pretty Good Privacy, Privacy Preserving Machine Learning, TensorFlow Privacy, Diffprivlib and Microsoft SEAL for privacy-preserving techniques

Guidance and standards: ISO/TR 31700, NIST Privacy Framework and the European Union Agency for Cybersecurity guidelines for data protection best practices

Privacy certifications: Certified Information Privacy Technologist and other certifications to deepen privacy expertise

Getting it right means

Effective data minimization: Collect and only use necessary data to achieve project goals — for example, data required to train or run DataStage models.

Successful integration of privacy-preserving technologies: Effectively use techniques like differential privacy, federated learning, and secure multi-party computation to protect data.

Transparency and accountability: Ensure AI systems are explainable and their data usage is transparent to stakeholders and end-users.

Trust and compliance: Achieve high levels of user trust through transparent data practices and maintain a record free of privacy violations.

High data utility: Extract actionable insights from data without compromising privacy, ensuring that all analyses align with ethical standards and regulations.

Bias mitigation and fairness: Maintain fair and unbiased AI models and mechanisms that continuously monitor and correct and deviations.

Published July 2025.

iapp.org