



Reimagining modern digital trust and online protection (DT&OP)

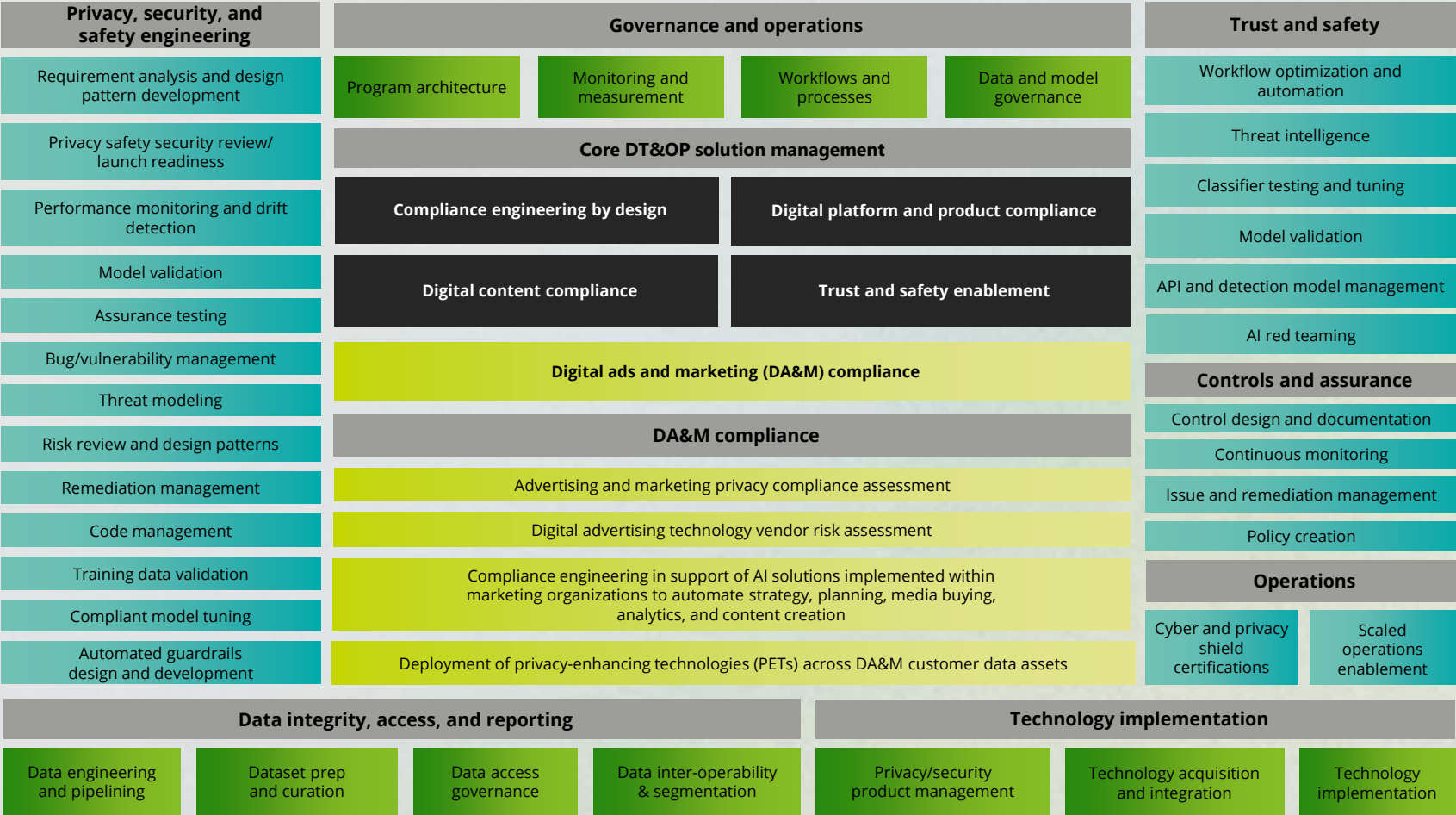
The rapid evolution of artificial intelligence (AI) is fundamentally reshaping how organizations approach engineering, trust and safety, and marketing—positioning these as powerful enablers and new sources of competitive advantage.

Organizations are seeing the rapid adoption of AI across digital platforms, unlocking machine learning-driven anomaly detection, predictive analytics, and AI assistants that can transform processes from manual and reactive to automated, proactive, and scaled. Against this backdrop, businesses are leveraging dynamic trust and safety frameworks to enhance operations, improve accuracy, and enable increased scalability and cost efficiency—driving meaningful value and operational excellence.

However, this transformation also brings significant new threats and complex regulatory demands. The proliferation of AI introduces online harms such as misinformation, deepfakes, algorithmic bias, and automated fraud—leading to stricter regulations and new compliance standards such as the Digital Services Act and Online Safety Act. To keep pace, organizations need to innovate in threat mitigation, compliance automation, and data-driven governance to proactively reduce harm, maintain auditability, and build lasting trust in a dynamic digital environment.

The future of DT&OP: Embracing AI

A blueprint for an AI-powered future: This is a model of the DT&OP function and the underpinning capabilities/processes that provide a baseline to reimagine and map to an AI-driven model. The legend below identifies the new and evolved AI-enhanced services.



LEGEND



Partial AI uplift with heavy human FTE



DT&OP-specific AI uplift



New DT&OP services

Connect to accelerate

Contact our leaders to dive deeper into the blueprint and reimagine what’s possible for your organization.



Mark Nicholson
Principal
Cyber AI GTM Leader
Deloitte & Touche LLP
manicholson@deloitte.com



Naresh Persaud
Principal
AI Transform Leader
Deloitte & Touche LLP
napersaud@deloitte.com



Tanneasha Gordon
Principal
Digital Trust & Privacy Leader
Deloitte & Touche LLP
tagordon@deloitte.com



Esther Choi
Partner
US DT&OP Leader
Deloitte & Touche LLP
eschoi@deloitte.com



Steve Ruzzini
Senior Manager
Cyber AI GTM
Activation Lead
Deloitte & Touche LLP
sruzzini@deloitte.com

DT&OP for AI: Safeguarding autonomous agents

AI agents require governance and compliance to facilitate safe, lawful, and policy-aligned operations.

Area	Focus	Foundational priorities
Governance	Establishing oversight, accountability, and risk management for AI agents.	Life cycle oversight: Implement processes for onboarding, monitoring, updating, and decommissioning AI agents.
		Governance roadmap: Develop a roadmap for adaptive AI agent governance.
Trust engineering	Embedding security, privacy, and compliance into the design and deployment of AI-enabled products.	Continuous monitoring: Routinely validate and assess AI model performance using defined guardrails and controls.
		Compliant model training: Train and tune AI models using compliant, bias-reducing data and methods.
Trust and safety	Protecting users and enabling safe, fair, and compliant AI-generated content.	Adversarial testing: Stress testing models with intentionally malicious or adversarial prompts, seeking jailbreaks, deception, or misuse.
		Transparent decision tracking: Maintain clear records of AI moderation for audits, appeals, and compliance.
DA&M	Facilitating ethical, secure, and compliant use of AI in customer engagement and brand activities.	AdTech vendor compliance testing: Assess AI-powered AdTech vendors for compliance and third-party risk.
		IP detection automation: Use automated systems to identify intellectual property and copyrighted material.

AI for DT&OP: Specific uplifts to achieve greater autonomous execution

Explore how harnessing AI and smarter solutions can reduce manual effort and accelerate results across key functions.

DT&OP capability/function	AI uplift approach	Potential resulting impact
Privacy, security, and safety engineering	Agentic AI automates regulatory monitoring, risk assessment, and compliance reporting —alerting teams early to changes and facilitating continuous alignment with evolving regulations.	Reduced manual effort, increased accuracy, and accelerated compliance cycles
Trust and safety	Increase the use of AI/ML systems for automated detection, triage, and monitoring of harmful content , evolving threats, and control effectiveness in near real-time.	Simplified trust and safety processes, reduced harm through enhanced risk detection , enabled continuous learning and adaptation, and reduced manual effort without sacrificing speed or coverage
Controls and assurance	Leverage AI-driven tools for evidence gathering, risk assessment, mitigation recommendations, and automated transparency reporting to enhance compliance and response protocols.	
Operations	Use agentic AI to identify and link audit evidence, combine document generation for audit packages, and enable automatic updates of control narratives.	Reduced manual effort by 30-40% focusing on end-to-end campaign*
DA&M compliance	Generative AI (GenAI) builds risk frameworks, automates compliance analytics, and monitors ad policies. AI creates, optimizes, and labels GenAI-driven marketing content for compliance.	Reduced manual effort and increased efficiency to execute

*Percentages are estimates based on recent project delivery for 10-15 organizations, ranging from 12-week implementation to 3-year operate engagements.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

As used in this document, “Deloitte” means Deloitte & Touche LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Copyright © 2025 Deloitte Development LLC. All rights reserved.