Navigating the AI Frontier in Toxicology: Trends, Trust, and Transformation

The integration of Artificial Intelligence (AI) into toxicology marks a profound paradigm shift in chemical safety science. In 2026, AI is no longer limited to automating traditional workflows; it is actively redefining how we assess risk, interpret complex biological data, and inform regulatory decision-making. This presentation explores the convergence of AI and New Approach Methodologies (NAMs), emphasizing a move toward a "*trustblazing*" future where predictive science is auditable, updateable, and human-centric.

A significant hurdle remains that toxicological evidence is often "text-locked" within unstructured PDF reports and narrative dossiers. We examine how Natural Language Processing (NLP) and Retrieval-Augmented Generation (RAG) are being utilized as core infrastructure to convert these records into machine-readable data streams, enabling semi-automated literature triage and Adverse Outcome Pathway (AOP) assembly. However, moving from simple text generation to high-stakes regulatory application requires Agentic AI systems, i.e., specialized sub-agents capable of multi-step scientific workflows. We propose an Evidence-based Agent Stack that protocolizes tasks such as risk-of-bias appraisal and mechanistic integration, ensuring mandatory provenance and versioning.

To bridge the gap between innovation and adoption, we introduce the e-validation framework. This model operationalizes the TREAT principles (Trustworthiness, Reproducibility, Explainability, Applicability, Transparency) through AI-powered modules for virtual study simulation and post-validation surveillance. By addressing ethical considerations, including bias audits and participatory governance, and adopting a "co-pilot" model where AI augments rather than replaces human judgment, the field can transition toward more inclusive and accurate safety assessments. Ultimately, the future of toxicology lies in reinventing the discipline as an adaptive, transparent science that aligns global policy momentum with the technical capabilities of the AI frontier.