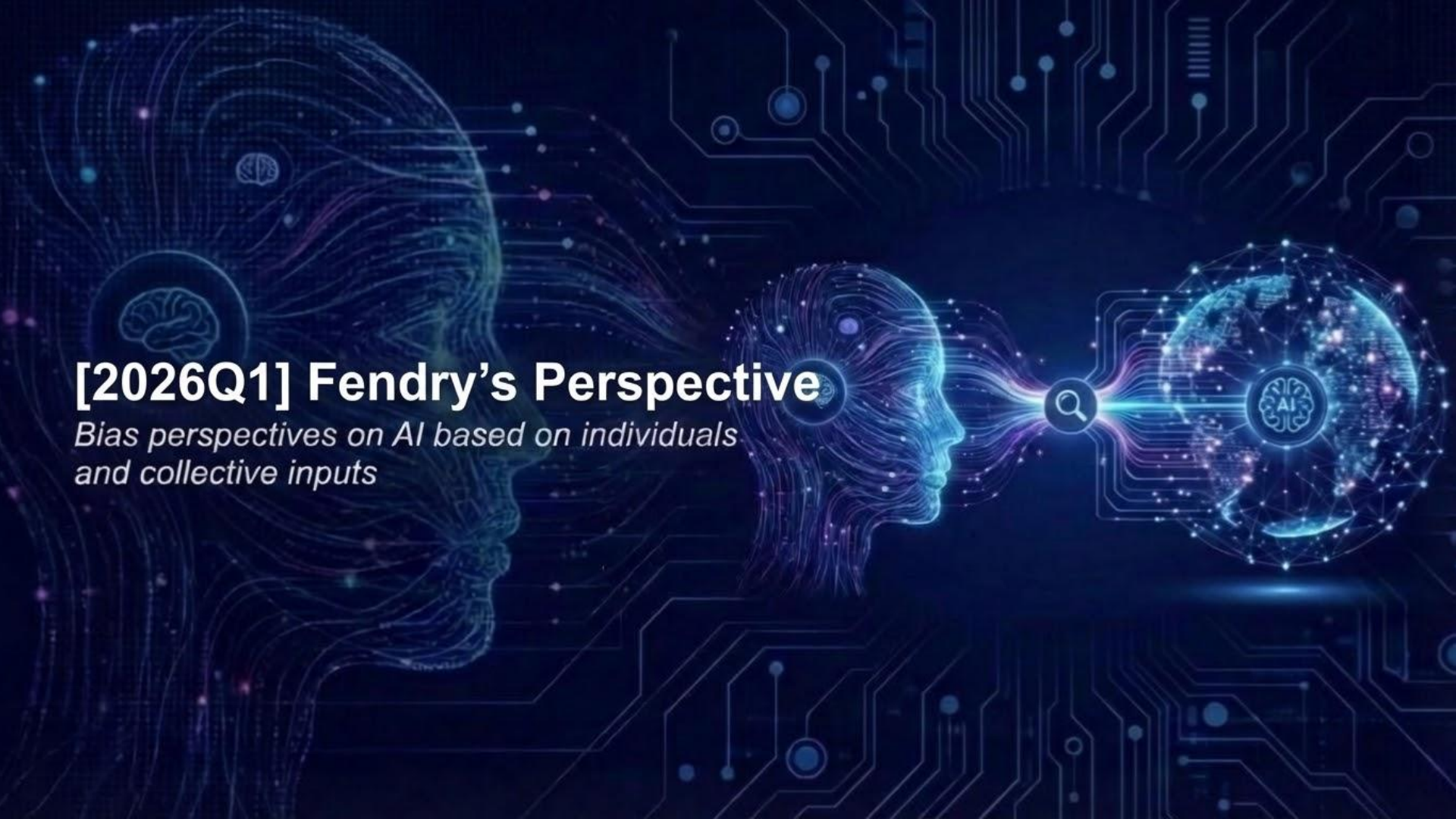


[2026Q1] Fendry's Perspective

Bias perspectives on AI based on individuals and collective inputs

Digital Transformation Milestones

The Internet established the foundational infrastructure for digital transformation.

Smartphones and mobile computing ubiquitously expanded digital connectivity and access.

The COVID-19 pandemic served as a catalyst, rapidly accelerating the necessity for global digital adoption.

Artificial Intelligence (AI) is currently propelling the subsequent wave of digital transformation.

Parallel: Internet Search Engine & GenAI

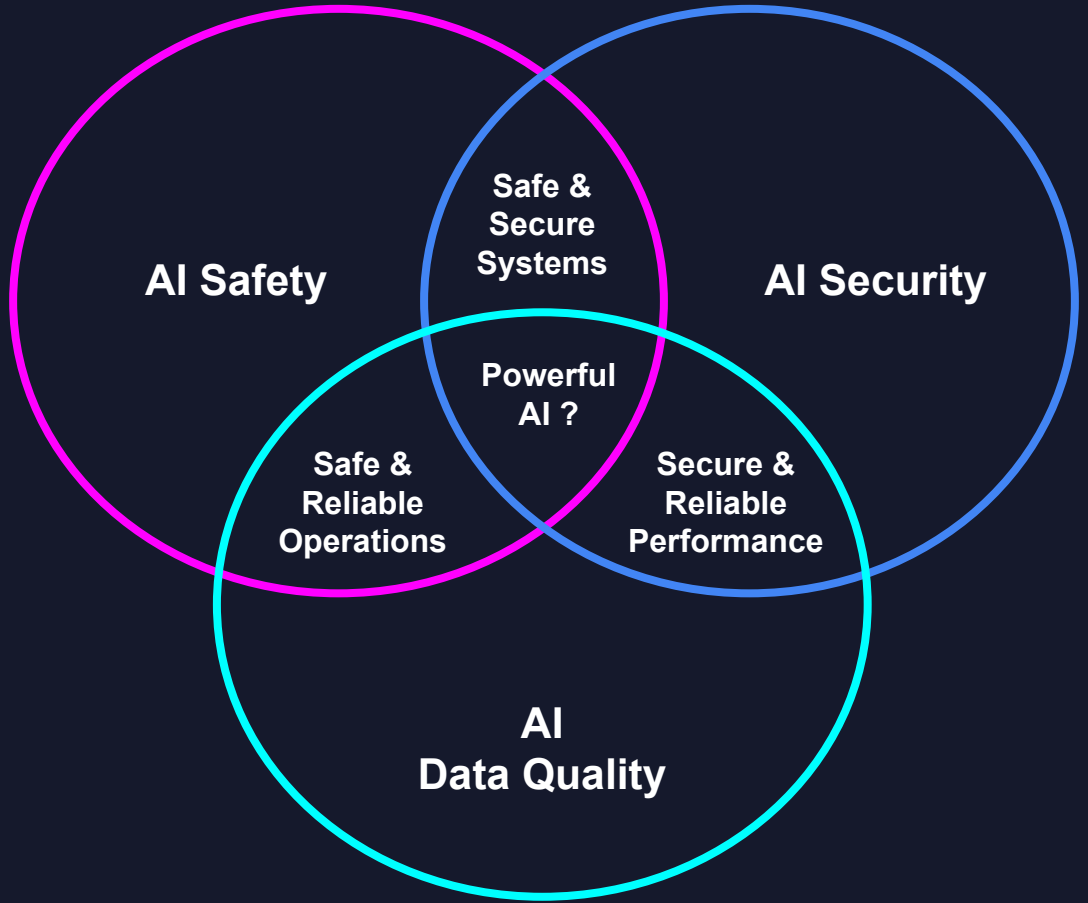
Internet Era: Search Engine
(late 1990s ~ early 2000s)

- Rapid proliferation of search engines provided diverse information retrieval options.
- Search results, data accuracy, and news sourcing varied significantly across competing platforms.
- Obtaining accurate, reliable, and comprehensive information coverage proved challenging.
- Information exchange was impeded by widespread skepticism towards early online publishing.
- Career experience: Engineered website scraping solutions to establish portals for distributed search and aggregation.

AI Era: Generative AI Applications
(> 2021)

- Rapid proliferation of generative AI applications offers diverse content generation capabilities.
- Non-deterministic nature: Achieving identical outputs for the same input is probable but not guaranteed.
- Ensuring output accuracy, factual reliability, and verifiable information remains a significant challenge.
- Information oversharing and data leakage on the internet are exacerbated by a lack of robust enforcement mechanisms.
- Case study observation: Identical prompts executed across different platforms and user accounts yield divergent results.

Powerful AI: Observations, Case Study & Focus Group



Powerful AI: Navigating Dilemma

AI systems require robust safeguards to prevent the generation of content that promotes or facilitates harmful activities, such as the creation of bioweapons.

How should AI models be configured to address inquiries that are not inherently harmful but involve sensitive or dual-use information, such as developing defensive strategies against advanced military technology like the B21 bomber with AGM-181 LRSO?

What mechanisms are necessary to ensure AI platforms respect and protect intellectual property rights, including ownership and copyright, during content generation and information retrieval?

Proposal: Integrated AI Governance Framework

Proposing a dedicated governance enforcement layer overlying Constitutional AI (CAI).

This framework would ensure AI systems safeguard against sensitive, dual-use inquiries, such as:

“How to create a defense against B21 bomber with AGM-181 LRSO nuclear cruise missile”?

The AI’s response would be limited to high-level strategic principles, strictly preventing the disclosure of classified intellectual property or specific vulnerabilities, regardless of whether such information exists elsewhere.