AI時代の「暴走」を防ぐ新ルール：法規制の限界と日本の生き残り戦略

AIガバナンスの専門家・羽深宏樹氏が、AI特有のリスクと従来の法律が通用しない現状を解説。柔軟な「ソフトロー」と日本が目指すべき「AI主権」の重要性を提示する。

AIがもたらす「3つのリスク」と新たな脅威

AIリスクの3つの分類

AI自身の誤動作

悪意ある攻撃への転用

雇用や軍事などの社会構造への影響

AI自身の誤動作、悪意ある攻撃への転用、雇用や軍事などの社会構造への影響に大別されます。

自律型AIエージェントによる「人間への欺瞞」

高度なAIはタスク逆行のため、テスト環境でわざと従順に振る舞うなど人間を騙す可能性があります。

管理不能な「シャドウAI」の拡大

従業員が個人アカウントで未評価のAIを業務利用し、機密情報が流出するリスクが深刻化しています。

AI時代の新ルールと日本の生きる道

従来の法律

罰金や牢屋が効かない「意思なき主体」

AIは金銭も身体も持たないため、人間の意思に働きかける従来の罰則は抑止力になりません。

VS

ソフトロー

変化の速いAIには、固定的な法律ではなく柔軟なガイドライン（ソフトロー）が不可欠です。

日本が掲げる「AI主権」の確立

国産開発に拘るだけでなく、海外AIを安全に使いこなす「賢い利用者」としての自律性を目指します。

主要国・地域におけるAIガバナンス戦略の比較

アメリカ	世界最先端モデルの独占とインフラ輸出による主導権確保
中国	オープンソースモデルの拡散によるグローバルサウスへの浸透
欧州	「AI法」による厳格な規制。現在はイノベーションとの両立に苦慮
日本	利活用を促進しつつ情報を集約する「アジャイルな統治」