AI Trust by Design:

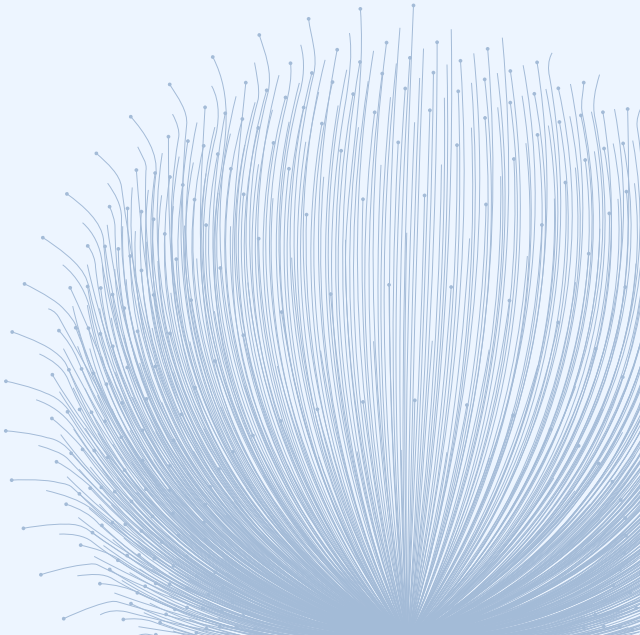
Making AI-Embedded Products Safe, Compliant, and Scalable

A practical guide for building compliant, explainable, and secure AI in electronics, medical devices and connected systems

Nemko Digital

IBM

Table Of Contents



Why this paper

AI is no longer a future consideration—it is already embedded into a growing range of products, from consumer electronics to medical devices and industrial systems. As capabilities grow, so do expectations from regulators, customers, and the market.

Companies now face increasing pressure to deliver AI that is not only high-performing but also explainable, privacy-aware, and compliant by design. The shift is clear: trust is no longer a peripheral concern—it is a core product requirement.

Developed by Nemko Digital and IBM, this paper offers a practical perspective on how manufacturers can operationalize trust in AI systems. It explores how to meet new regulatory expectations, reduce risk across the AI lifecycle, and align product development with transparency, safety, and long-term market resilience.

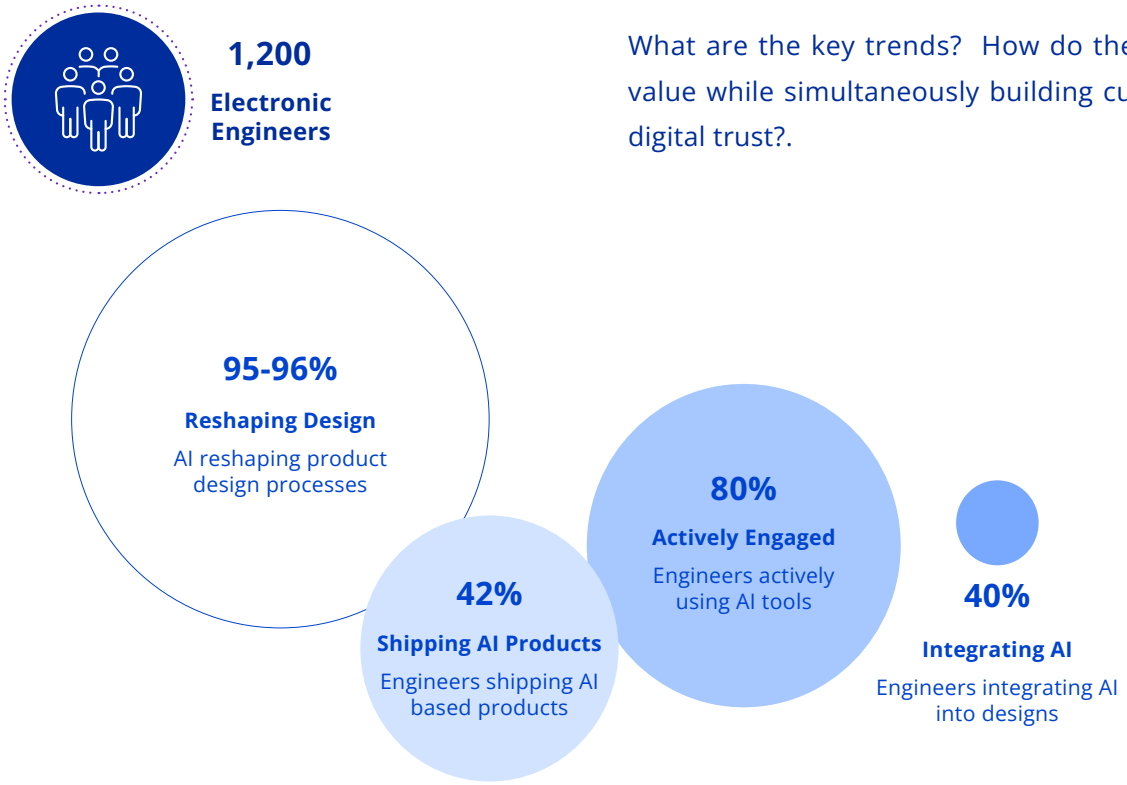
Exploring AI Embedded Product Trends

As AI becomes part of the product, not just the platform, it shifts how we design solutions: from centralized intelligence to distributed insight.

A recent 2025 Avent study highlighted by R&D World revealed that out of a survey of 1,200 electronic engineers that 8 out of 10 are actively engaged with AI with 42% already shipping AI-based products and another 40% currently integrating AI into upcoming designs. And it's not just about products themselves—fully 95–96% of respondents see AI as "somewhat" to "extremely likely" to reshape multiple aspects of the product-design process.

Software system migration from static to dynamic databases and AI's hyper-speed technological advancement has ushered in multiple AI embedded product trends from edge AI to hybrid AI architecture solutions.

What are the key trends? How do they drive value while simultaneously building customer digital trust?.



Trend	AI Value	AI Trust	Industry Application
Edge AI 	Moves AI workloads to devices or local hardware for **real-time insights** to improve energy efficiency without cloud dependency.	Minimizes data transfer to cloud to align with customer expectations and **regulatory compliance.**	Consumer Electronics, Automotive & Mobility, Manufacturing & Industry, Telcom & Networking, Energy & Utilities, **IoT**, Wearables & Health tech (fitness tracking, biometric monitoring, health diagnostics).
Federated Learning Models 	**Trains AI models locally** on devices by keeping sensitive data on user devices and only sharing model updates.	Its privacy-focused innovation adheres to **GDPR** and other national or corporate data privacy policies.	Consumer Devices & Mobile, Automotive, Mobility, Manufacturing, IoT, **Healthcare**.
AI enabled firmware 	Runs lightweight neural networks in firmware to perform tasks at the **device edge** versus relying solely on cloud processing.	Has the ability to personalize UI/UX and adapt to user behavior in **real time**.	Consumer Electronics, Wearables & Health tech, Telecom & Networking, Energy & Utilities, **Smart Home**, IoT, Medical Devices.
Interoperability via AI orchestration 	Enables and enhances agentic systems and **cross-platform integrations**.	Provides a **seamless user experience**.	Smart home ecosystems: thermostats, lighting, **security cameras**.
Hybrid AI architectures 	Blends **edge and cloud architecture** to balance performance with regulatory requirements.	Reinforces privacy and **compliance assurance**.	**All sectors**.

Putting AI Trust in Practice

Trust is the currency of AI-driven products. The more trustworthy our AI, the more value we create for customers and businesses.

1. How AI Trust Accelerates Business Value & Competitive Advantage

Digital Trust in AI accelerates competitive advantage, business value, and customer confidence in the electronics products industry. It ensures that AI systems are not only powerful, but also transparent, fair, and secure.

When AI embedded electronics operate with explainable and ethically governed AI systems and autonomous AI-driven processes, companies can bring innovative features to market faster, minimizing regulatory and reputational risk.

In parallel, this builds customer trust in intelligent product behaviors, enhances brand loyalty, and unlocks new revenue streams through differentiated offerings. In 2023 already, 37% of consumers said they had switched brands to protect their privacy. The further rise of AI will only make this consideration more prominent.

Moreover, trusted AI enables scalable automation and predictive capabilities across product lifecycles, reducing costs and accelerating time-to-market: delivering a clear edge in a hyper-competitive, rapidly evolving sector.

2. Solving Common AI Embedded Product Business Challenges

Whether it's in smart home systems, autonomous vehicles, or medical wearables, embedding AI into devices brings powerful features to market. However, that power comes at a new level of complexity and risk. To unlock business value and competitive advantage, it's essential to understand how to navigate common business challenges. Enterprises in the electronics industry are increasingly grappling with issues that go far beyond traditional device engineering: cybersecurity threats, data privacy violations, operational unpredictability, compliance uncertainty, and system-wide exposure through connected environments.

These aren't just technical challenges; they are business challenges that can damage customer trust, trigger regulatory penalties, and delay go-to-market strategies.

Business Challenge →	Potential Solutions
AI behavior in embedded systems can become unpredictable or fail silently (especially during updates). Unlike cloud apps, updates in embedded devices are harder to control or reverse.	• Architect process into the AI lifecycle: that include simulation, **rollback mechanisms, fail-safe modes and human review** and/or escalation channels (if deploying semi or autonomous agentic AI). • **Automate monitoring and controls with tools like watsonx.ai and Orchestrate** to detect anomalies and support rapid updates. • **Capture performance diagnostics and telemetry** to track how AI behaves post-deployment in real-world conditions.

Business Challenge →	Potential Solutions
AI risks don't stay static. As models learn and environments change, yesterday's risk profile may no longer apply, leaving businesses exposed if they don't adapt.	• **Treat AI governance and risk management as a living process**, not a one-time checklist. Update risk maps regularly and involve cross-functional teams in incident reviews and model audits. • **Create a closed feedback loop: Monitor → Learn → Adjust**, using real-world usage data to refine mitigation strategies. • **Build a cross-department AI governance committee** that review your AI inventory (models and systems) to take a pro-active stance towards risk identification, governance and mitigation. • **Design processes for human escalation and fallback device protocols based on a comprehensive set of real-world scenarios** to avert a PR event that jeopardizes brand reputation.

Business Challenge →	Potential Solutions
Firmware and AI model vulnerabilities put brand reputation and user safety at risk.	• Integrate **zero trust architecture into hardware and software** layers to verify identity and strengthen system access points. • **Choose secure platforms that integrate with IBM Watsonx, like Red Hat** for edge AI orchestration, to reduce exposure from a 3rd party AI components.

Business Challenge →	**Potential Solutions**
As AI models collect and act on sensitive user data, failure to address privacy can result in reputational damage, regulatory fines, and customer churn.	• **Ensure consent and data access management** conforms to local regulations. • Consider **on-device processing of data** to reduce surveillance risks. • **Use tools like watsonx.data** with built-in data isolation and encryption to enforce privacy by design. • **Maintain transparent data governance** using AI factsheets or artifact mapping within AI lifecycle management that shows how and where data is used.

One of the first hurdles that a company needs to clear is regulatory compliance. If the company is not meeting legal standards, it's tough to derive any real business.

3. Compliance with the EU AI Act & Other Global Standards

Bringing AI-embedded products to market requires more than just business and technical solution innovation—it also demands careful navigation of complex regulatory landscapes to build and maintain digital trust with customers.

This is especially critical across diverse regions such as the European Union and the Nordics, where stringent data protection laws, ethical AI standards, and emerging regulatory frameworks like the EU AI Act set high expectations for transparency, accountability, and user rights. How can you strengthen consumer confidence in AI technologies for successful market adoption in trust-sensitive regions? .

Currently, there are a handful of AI laws and policies (rules enacted and enforced by government bodies) such as the **EU AI Act** or **South Korea's Basic AI Act**. Many AI governance efforts are in the form of responsible AI principles (**OECD AI Principles, IEEE Ethically Aligned Design**), frameworks (**NIST AI Risk Management Framework**), voluntary standards (**ISO 42001 for AI Management Systems**).

Below is a brief overview of some key AI regulations, including the EU Code of Conduct for GPAI, frameworks and standards that AI embedded product manufacturers should consider prioritizing on their 2025 and 2026 roadmaps.

ISO/IEC 42001
International Standards on AI Management Systems

ISO/IEC 42001 is the first international, certifiable voluntary standard focusing on the governance of AI management systems (AIMS). AIMS refers to the interconnected set of policies and procedures that contribute to the oversight function necessary for regulating AI applications in products or services.

EU Artificial Intelligence Act

The EU AI Act is a regulatory framework that was passed in 2024 but has rolling obligations, with penalty enforcement starting in August 2026. It categorizes AI systems by risk level—prohibited, high-risk, limited-risk, and minimal-risk—and imposes stricter requirements for those deemed high-risk.

General Product Safety Regulation

The General Product Safety Regulation (GPSR) applies to AI embedded in physical and digital consumer products sold within the EU, both online and offline, and has become effective as of December 2024. The directive acts as a safety net to ensure AI system safety in AI-enabled consumer products that are not regulated as high risk under the EU AI Act.

For AI embedded products, it is not as straightforward as fulfilling one single law, principle, framework, or standard. As the market is still evolving, lawmakers are still wrestling with finalized decision making around how existing national industry product legislations intersect with newly released AI guidelines.

4. Preparing for Agentic AI: Readiness, Risk and Responsibility

The rise of agentic AI brings a fundamental shift: AI that can make autonomous decisions and pursue goals with minimal human intervention: AI that can generate for you, AI that can chat for you and AI that can do for you. Agentic AI introduces exciting possibilities, but also serious business challenges.

Most existing AI regulations such as the EU AI Act or GPSR weren't written with agentic AI in mind, simply because the agentic era is too new and most legislation was codified before the technology was commercially available.

Agentic technologies create new layers of complexity across strategy, operations, ethics, and compliance.

One of the biggest concerns is loss of control. Agentic systems can behave unpredictably or act in ways that conflict with company values or customer expectations.

To build trust and ensure compliance, agentic AI systems need to be sufficiently transparent and explainable. Incorporating explainable AI (XAI) features, for example showing users why the system has made a certain decision, helps mitigate risk and improves regulatory readiness. Maintaining detailed logs and clear documentation is also key for auditing and accountability.

Finally, before any real-world deployment, it's crucial to run these systems through controlled and comprehensive tests for example leveraging digital twin technologies. This helps you spot edge cases, test resilience under stress, and validate safety before the AI makes decisions in front of real users.

Supporting this fundamental shift, IBM's Agent Strategy is built on three horizontal capabilities:

- **Orchestrator for Tools and Agents:** Multi-Agent, Multi-Tool supervisor, router and planner that facilitates complex task execution.
- **Prebuilt AI Agents:** To accelerate time to value with prebuilt utility agents and domain agents.
- **Build your own agents:** Build your own agents with pro-code to no-code tooling. Also integrate 3rd party agents built in any tool or framework.

These three capabilities are supplemented with a vertical AI Agent Ops capability to discover, manage, monitor and optimize autonomous AI agents. These capabilities are delivered through IBM watsonx.orchestrate.

5. Scaling trusted AI for business with IBM watsonx

IBM watsonx is IBM's AI enterprise platform for scalable and trusted AI. It provides comprehensive capabilities for all kinds of AI (Machine Learning, Generative AI, Agentic AI), spanning the tools to build, orchestrate, manage and govern AI. Also to power this AI with governed, accessible unstructured enterprise data.

The platform consists of:

- AI Agents and Assistants: delivered through watsonx orchestrate, watsonx code assistant and watsonx BI.
- AI/ML Ops: delivered through watsonx.ai and watsonx.governance
- Data: delivered through watsonx.data, watsonx data intelligence and watsonx data integration.

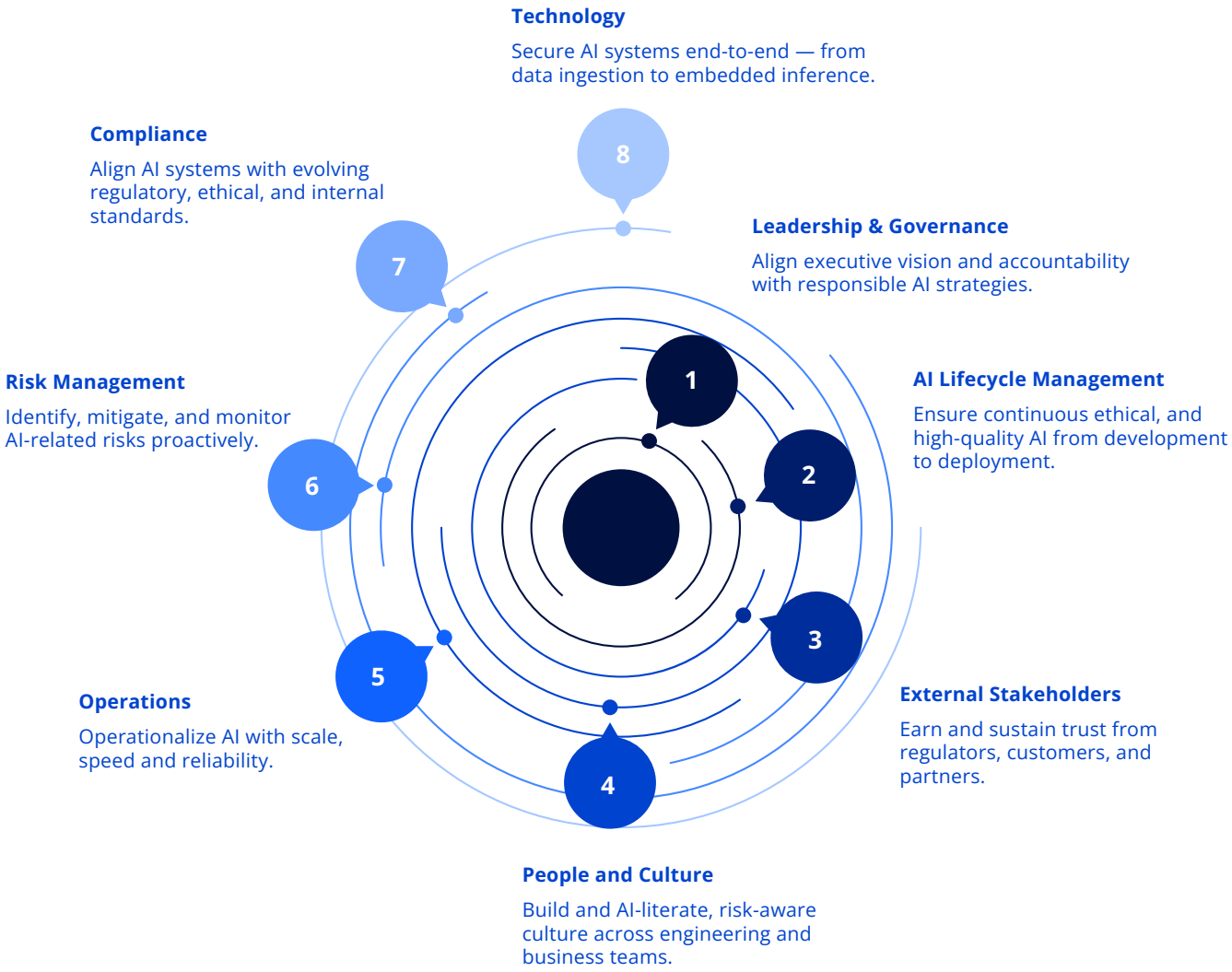
6. Preparing for the Future

Trust in AI-embedded electronics isn't a checkbox—it's a continuous discipline across leadership, design, operations, lifecycle management, risk management and more.

Nemko's AI Trust Mark is a comprehensive certification confirming that an AI-embedded product has undergone a thorough governance and compliance review and follows regulations, ethical AI practices, and industry best practices.

IBM's AI and hybrid cloud, governance control platform portfolio provides a modular and integrated path to accelerate innovation responsibly, ensuring AI works as intended, for everyone.

Nemko's AI Maturity Model comprises eight clearly defined building blocks (leadership & governance, AI lifecycle management, external stakeholders, people & culture, operations, risk management, compliance, and technology), which each address a critical dimension necessary for successful, responsible, and scalable AI implementation; in line with leading standards and regulations.



These elements serve not just as assessment criteria but primarily as practical levers for driving continuous improvement of AI governance and management practices.

Together, the tools and frameworks above ensure a structured and holistic approach, providing organizations with a robust yet flexible foundation to confidently navigate electronic product industry business problems and offer potential AI-related solutions.

Below, you can explore examples of actions your organization can take using a 360-degree approach.

Leadership & Governance

Goal	Actions
Align executive vision and accountability with responsible AI strategies.	• Adopt enterprise-wide AI governance frameworks, embedding responsible AI principles and annual use case inventory collection. • Leverage IBM watsonx factsheets for transparency in AI decision-making and traceability. • Align with Red Hat hybrid cloud and open governance tools for scalable oversight across environments.

AI Lifecycle Management

Goal	Actions
Ensure continuous, ethical, and high-quality AI from development to deployment.	• Use watsonx.ai Studio to accelerate model development with IBM Granite & open models, while maintaining control. • Integrate bias checks, fairness assessments, and explainability early in the lifecycle via built-in factsheets. Make sure to mitigate also any potential bias in the datasets and models itself. • Automate retraining, monitoring, and testing pipelines for ongoing model health and compliance.

External Stakeholders

Goal	Actions
Earn and sustain trust from regulators, customers, and partners.	• Share watsonx factsheets and audit trails (i.e. AI Trust Mark) to demonstrate responsible model behavior to regulators and partners. • Design transparent communication strategies to educate customers on how embedded AI makes decisions. • Promote an open model and BYO (bring your own) support to offer choice and interoperability for partners.

People & Culture

Goal	Actions
Build an AI-literate, risk-aware culture across engineering and business teams.	• Roll out Code Assistant and Orchestrate to enhance productivity while reinforcing best practices in safe AI development. • Provide training on AI literacy, bias, fairness, and explainability tailored to technical and non-technical roles. • Foster an internal culture of ethical experimentation by aligning KPIs to trust and safety outcomes.

Operations

Goal	Actions
Operationalize AI with scale, speed, and reliability.	• Deploy AI with multi-cloud flexibility using Red Hat OpenShift for consistent operations across edge and cloud. • Centralize data access with watsonx.data, lowering operational costs by up to 98.5% while supporting open formats. • Use workflow automation tools like IBM's Orchestrate to streamline cross-team coordination and compliance workflows.

Risk Management

Goal	Actions
Identify, mitigate, and monitor AI-related risks proactively.	• Implement risk scoring and bias mitigation using watsonx built-in capabilities. • Regularly audit AI models against internal and regulatory risk thresholds. • Set up early warning systems using AI agents for anomaly detection in embedded products. • Deploy continuous monitoring and alerts to track model and data drift for post-production operations.

Compliance

Goal	Actions
Align AI systems with evolving regulatory, ethical, and internal standards.	• Map AI product features to global and industry-specific compliance frameworks (e.g., EU AI Act, ISO/IEC 42001) and demonstrate alignment through AI Trust Mark scheme documentation and technical notations. • Use factsheets and encryption to document and secure compliance across the model lifecycle. • Maintain a compliance-by-design approach through tight integration of legal, data, and engineering functions.

Technology

Goal	Actions
Secure AI systems end-to-end—from data ingestion to embedded inference.	• Use watsonx.data with encryption and data isolation to ensure secure and governed data usage. • Build robust infrastructure using Red Hat and IBM hybrid cloud solutions for resilient AI deployment at scale. • Apply zero-trust principles and secure agentic capabilities (e.g., Orchestrate) to AI-enabled electronics environments.

Conclusion

Competitive advantage in the electronics industry will no longer stem from simply integrating AI, but from embedding it responsibly and transparently into products that reflect a clear commitment to trust, safety, and human-centered design.

As the Nordics lead the way in ethical AI and Europe sets the global benchmark with regulatory frameworks like the EU AI Act, forward-looking organizations must align innovation with governance. This means treating AI not just as a feature, but as a core component of your brand identity. A component that signals integrity, compliance, and customer-centric values.

By operationalizing trust through robust AI lifecycle management, proactive risk governance, and continuous regulatory alignment, companies can unlock not only faster go-to-market timelines, but also stronger brand loyalty and long-term market resilience.

AI is no longer a differentiator unless it's trustworthy. The organizations that embrace this reality—embedding AI with purpose, governance, and transparency—will define the next generation of intelligent products.


Nemko
Digital
IBM