The webinar summit focused on how to build and scale trustworthy AI in electronics, balancing innovation with compliance and safety. The central theme was “guardrails as enablers of innovation,” with practical guidance on:

- Interpreting and preparing for the EU AI Act
- Bridging management-system standards and product-level assurance
- Managing emerging risks from agentic AI
- Translating trust-by-design into operational practices and ROI
- Real-world case studies in public policy tracking, journalism, and healthcare

Key Points and Segment Summaries

Introduction: The Urgency of AI Trust in Electronics

Bas Overtoom from Nemko Digital opened the session by highlighting the unprecedented speed at which AI is being integrated into consumer electronics. He noted that concepts that seemed futuristic just months prior, such as AI-enabled home appliances, are now commercially available. This rapid proliferation underscores the urgent need for a collaborative dialogue on AI trust. He framed the conversation around the idea that trust is not an obstacle but a necessary component for responsible innovation, setting the stage for the day's key theme: implementing guardrails to build trustworthy AI from the ground up.

Government Perspective: Norway's National AI Strategy

State Secretary Mariana Williamson provided a comprehensive overview of Norway's ambitious strategy to become a leading nation in the development and use of ethical and safe AI. She acknowledged the public's skepticism, citing a Norwegian study where only 30% of respondents had high trust in AI systems. To address this, the government is pursuing a five-track approach:

1. Investing in National AI Infrastructure: Developing open and free Norwegian and Sami language models and expanding high-performance computing capacity.

2. Competence Building: Allocating over 1 billion Norwegian kroner to AI research centers, including the Norwegian Centre of Trustworthy AI, and promoting digital skills from primary school to lifelong learning.

3. AI in the Public Sector: Setting a goal for 100% of public agencies to adopt AI by 2030 to improve services and efficiency, while maintaining public trust through transparency.

4. International Cooperation: Working closely with the EU, OECD, and Nordic countries to promote a unified approach to safe and ethical AI.

5. Regulatory Framework: Implementing the EU AI Act into Norwegian law to ensure a level playing field for businesses and establish a cornerstone for trust and innovation. She announced the establishment of AI Norway, a national competence hub, and a Regulatory AI Sandbox to help companies, especially SMEs, develop and train AI systems in a controlled environment.

Industry Context and Future Trends

Dr. Pepijn van der Laan from Nemko Digital set the scene for the industry, discussing the increasing complexity of modern products and the corresponding need for trust. He introduced the seven principles for trustworthy AI as defined by the EU, connecting them to real-world challenges and news events. The presentation from Fisita focused on the practical aspects of building and implementing trustworthy AI, moving from theory to application. A speaker from IBM discussed the future of "Agentic AI," exploring how advanced AI systems will transform workforces and business operations through examples like generative AI for employee training, which drastically reduces training time and improves job performance.

Important Concepts and Insights

- **Guardrails as Innovation Enablers:** The recurring insight was that regulations like the AI Act should not be viewed as restrictive. Instead, they provide a clear and predictable framework—or "guardrails"—that allows companies to innovate more confidently and quickly, knowing their products will meet safety and ethical standards.

- **Augmented Intelligence vs. Artificial Intelligence:** The IBM representative made a crucial distinction, expressing a preference for the term "Augmented Intelligence." This reframes the technology's purpose as one that augments and enhances human intelligence rather than replacing it. This philosophy champions a "human-in-the-loop" approach, where technology serves as a collaborative partner to automate tasks and provide insights, but the ultimate decision-making remains with humans.

- **The Role of Procurement in AI Trust:** A key discussion point was the evolving responsibility of procurement organizations. As companies deploy AI systems developed by third-party providers, their procurement teams must become adept at understanding the associated risks and ensuring that the technology they acquire is compliant and trustworthy. This shifts a significant part of the responsibility to the deployer of the AI, not just the developer.

- **Human-AI Teaming:** A concept from the research community was introduced, suggesting that the future of AI in business lies in "Human-AI Teaming." This goes beyond simple human oversight and involves a deeper, collaborative operational model where human experts are integrated at every stage to validate and guide the outcomes produced by AI systems.

Segment Summaries and Key Insights

Opening context: Why AI trust, why now (Bas Overtoom, Nemko Digital)

- AI adoption in electronics is accelerating—from vision-based air conditioners to connected appliances and edge AI—making trust a prerequisite for scale.
- Key theme: Guardrails should be built in from the start, evolving from POC to production.
- Program focused on policy, industry practices, future trends (agentic AI), standards, and practical implementation.

Government keynote: Norway's roadmap for trustworthy AI (State Secretary Mariana Williamson)

- Ambition: Make Norway a front-runner in ethical and safe AI and "the most digitalized country in the world."
- Five tracks:
    1. **Infrastructure:** Open Norwegian and Sámi language models trained on local data; HPC; new data center strategy to enable AI scale.
    2. **Competence:** Talent pipeline and lifelong learning; > NOK 1B to six AI research centers (incl. Norwegian Centre of Trustworthy AI); a Practical Guide for Responsible AI; creation of "AI Norway" as a national hub and regulatory sandbox.
    3. **Public sector:** Goal for 100% of agencies to adopt AI by 2030; transparency and responsibility to preserve high public trust.
    4. **International collaboration:** Close work with EU, OECD, Nordics; participation in G20 AI task force.
    5. **Regulation:** Implement the EU AI Act in Norwegian law, emphasizing that early preparation enhances products, safety, and trust. Certification and auditing (including work by IBM and Nemko) can accelerate adoption.

Industry landscape and trust-by-design (Nemko Digital)

- Electronics engineers report widespread use of AI "behind the scenes," and 82% already design products that embed AI.
- Hybrid architectures are rising: edge firmware, federated learning, and enterprise orchestration demand end-to-end trust.
- Regulations extend beyond the AI Act; many sector and product safety laws also apply.
- Maturity model: 8 organizational cornerstones and 38 capabilities to scale AI responsibly; moving early prevents costly "shadow AI" and rework.
- Agentic AI is adding two powerful capabilities beyond genAI: ability to act and multi-agent orchestration—introducing new risks such as error amplification among agents and harmful actions.
- Bottom line: Stop bolting on trust; build it into the foundation.

IBM: From productivity to ROI, and governing agentic AI (HP)

- Personal productivity (e.g., copilots) is ubiquitous but hard to tie to enterprise ROI unless time savings are reinvested into the business.
- IBM has operationalized ~6,500 AI systems internally, adding ~200 per quarter—showing that rigorous governance accelerates deployment.
- Many genAI pilots stall; an MIT study indicates only a small fraction make it to production, underscoring the need to prioritize use cases with clear ROI.
- Examples:
  - Predictive ML in insurance led to proactive care calls and a substantial reduction in long-term sickness risk—demonstrating classic ML value without genAI.
  - Udbetaling Danmark uses genAI to generate compliant call summaries for every contact—material efficiency gains in a high-volume setting.
- Agentic AI can turn every front end into a back end—agents plan, act, and coordinate across systems. The opportunity is huge, but so are risks.
- **Governance tools and practices:**
  - IBM AI Risk Atlas (open) catalogs risks across ML, genAI, and agentic AI.
  - watsonx.governance governs models at runtime, including agent-level monitoring and metrics such as "faithfulness" (validating generated output against cited sources).
  - IBM Ethics Board translates risk findings into runtime mitigations that are continuously integrated into platform capabilities.

EU AI Act and standards: What to expect and how to prepare (Jochen, IBM)

- **The AI Act follows the EU's New Legislative Framework:** law sets essential requirements; harmonized standards define "how," enabling innovation and presumption of conformity (Article 40).
- **Phasing:**
  - Prohibited practices in force (Feb 2, 2025).
  - GPAI (general-purpose AI) obligations in force (Aug 2, 2025), bridged by a Code of Practice until harmonized standards arrive.
  - High-risk AI systems obligations apply in 2026; timelines are tight and some standards may be finalized late.
- **Standardization:**
  - CEN/CENELEC JTC 21 coordinates EU standards aligned to essential requirements (risk management, data quality and governance, technical documentation/record keeping, transparency/user information, human oversight, accuracy, robustness and cybersecurity, QMS, conformity assessment).
  - Close alignment with ISO/IEC JTC 1 SC 42 is vital for global consistency ("Brussels effect").
- **ISO/IEC 42001 vs AI Act:**
  - ISO/IEC 42001 is an organizational AI management system standard.
  - The AI Act imposes product/system-level obligations—ISO/IEC 42001 helps but does not equal compliance.

- **GPAI Code of Practice:** interim transparency, safety/security measures; IBM's open-source Granite models publish required documentation.
- **Tools:** watsonx.governance mapped to AI Act obligations and evolving standards to support documentation, risk governance, evaluation, and monitoring.

Nemko AI Trustmark (Monica Fernandez, Nemko Digital)

- A product-level AI assurance scheme aligning the EU AI Act, ISO/IEC 42001, and NIST AI RMF.
- Developed with the Korean Standards Association (KSA), now expanding in Europe.
- Fills the practical gap between organizational standards and system-level obligations by focusing on AI-embedded products and services.
- Risk-based approach with ongoing alignment to evolving codes and standards.
- Process: risk categorization → assessment → validation/mark issuance → surveillance (1 year) → reassessment (2 years), with updates as guidance evolves.
- Benefits: concrete evidence for conformity assessment, market signaling of trust, and accelerated compliance readiness.

Practical implementations (Visito: Bruna and Morten)

- **EFTA regulatory tracking:**
  - AI curates and prioritizes EU acts and proposals affecting the EEA.
  - RAG over internal archives accelerates legal analysis and continuity.
- **Journalism "Jin" platform:**
  - Used by ~40 newspapers; ingests data from >140 municipalities.
  - Automates summarization, entity extraction, "newsworthiness" scoring, and feedback learning.
  - Reported outcomes: ~80% time savings on document triage, doubled story output in the covered domain, ~33% higher click-through rates.
- **"Gotti Bausen" project:**
  - Aggregates and structures municipal political documents; targeted users include journalists, researchers, municipalities, and the public.
  - Scaling to all 357 municipalities—roughly 40 million meeting documents per year—where AI is essential to surface signal from noise.
- **Healthcare decision support (POC):**
  - RAG-based tool for medical committees to retrieve similar cases and guidelines faster.
  - Strong focus on privacy (de-identification pipeline), security, topical chunking, vector search, and auditable source attribution.
  - Next steps: scale de-identification, performance monitoring, EU Act-compliant governance, and workflow integration.

Panel Highlights and Q&A Takeaways

- **Provider vs deployer duties:** For deployers (users), the AI Act can be manageable—start with transparency obligations and build maturity. Providers (builders) bear heavier responsibilities.
- **Don't overlook analytical AI:** Many Norwegian companies may see higher ROI in classic ML before genAI/agentic AI; focus on core process ROI, not just personal productivity.
- **Digital sovereignty and security:**
  - Trust evolves into target security: securing data pipelines, preventing shadow AI/agents, and hardening against prompt injection/abuse.
  - Security must be co-developed with governance; buyers should demand clear security controls from vendors.
- **Building on third-party components:**
  - Vet APIs, data flows, and hosting locations; when needed, bring models on-premises to control data movement.
  - Test components and the integrated system—compositional risks can emerge even if parts are "safe."
- **Two-sided trust:**
  - AI must "trust" inputs: data quality, provenance, and role-based access control are critical.
  - Threat models change with agentic systems and broader external access—security posture must adapt.
- **Accountability:**
  - Responsibility remains human—regulation is aimed at human accountability for deployment and use.
  - Agentic autonomy raises stakes; invest in runtime monitoring, guardrails, and auditable decision trails.
- **ROI beyond productivity:**
  - Move from "time saved" to core process transformation (e.g., proactive care in insurance; call summarization at scale; recruitment pipelines that fairly screen at volume).

Key Takeaways

- **Trust is a growth enabler:** early guardrails reduce technical debt, prevent shadow AI, and accelerate operationalization.
- **Prepare now for the AI Act:**
  - Map use cases to risk levels; begin documentation, data governance, and human oversight practices.
  - Track CEN/CENELEC JTC 21 standards; expect tight timelines—use interim codes and best practices.
- **Bridge org and product layers:**
  - ISO/IEC 42001 strengthens organizational readiness; product-level assurance (e.g., Nemko AI Trustmark) connects it to system compliance.

- **Govern the full AI spectrum:**
  - Use a unified governance approach for ML, genAI, and agentic systems; include runtime monitoring and source attribution checks.
- **Make security first-class:**
  - Treat AI security (data pipelines, access control, model endpoints, agent orchestration) as integral—not an afterthought.
- **Prioritize ROI-driven use cases:**
  - Start where value is clear and measurable; scale from pilots to production with governance in place.

Next Steps for Attendees

The webinar provided a clear and compelling case that trust is not a feature but the foundation of the future of AI in electronics and beyond. For attendees, the key takeaway is that passivity is not an option. Businesses should not wait for the AI Act to come into full force before taking action. The next steps involve proactively:

1. **Educating Teams:** Ensuring that development, compliance, and procurement teams understand the principles of trustworthy AI and the requirements of upcoming regulations.

2 . **Adopting a "Guardrails" Mindset:** Integrating risk management, human oversight, and data quality checks into the AI development lifecycle from the very beginning.

3 . **Leveraging National Resources:** Engaging with initiatives like AI Norway and its Regulatory AI Sandbox to test and validate AI solutions in a safe environment.

4. **Fostering Collaboration:** Continuing the dialogue between industry, academia, and government to collectively build a future where AI is a force for good, driving innovation, efficiency, and societal benefit in a manner that is safe, reliable, and deserving of public trust.

The event successfully underscored that building trustworthy AI is a shared responsibility and a competitive advantage that will define the leaders in the next wave of technological transformation.