WEBINAR RECORDED

PECB
Artificial Intelligence Professional

PECB

The AI Governance Illusion: Why ISO/IEC 42001 and CAIP Matter More Than Ever

JUNE 25 | 03:00 PM CEST

Dr. Roman Krepki

Senior Manager for Cyber-Security & Risk at Forvis Mazars

#GlobalLeadingVoices

Short CV: Dr. Roman Krepki

Dr. Roman Krepki | Senior Manager | Technology & Digital Cyber-Security, Data Protection, AI and Risk Management Systems

Senior Manager at Forvis Mazars

Cyber-Security, Data Protection, AI, Risk Management, Audit & Advisory

Forvis Mazars GmbH & Co. KG
Breitscheidstrasse 10; D-70174 Stuttgart

Roman.Krepki@forvismazars.com
Phn: +49 711 666-31814
Mobile: +49 15 120 333 140

Educational Background

- Diplom-Informatiker (Computer Science) – Technical University of Berlin
- Doctor of natural sciences (Dr. rer. nat.) – Computer Science, Artificial Intelligence and Neurophysiology

Professional Qualifications (selection)

- Certified ISO27001 Lead Auditor and Auditor Instructor (PECB)
- ISO27002:2022 Senior Lead Manager (PECB)
- Certified Data Protection Officer (CDPO – EU GDPR, HWK & PECB)
- Certified Business Continuity Professional (CBCP, DRII)
- Certified Information Security Systems Professional (ISC2 - CISSP)
- ISO42001 Certified AI Management System Lead Implementer (PECB)

Specialist Competences (selection)

- ISM, IT and Cyber Security
- IT Governance, Risk and Compliance (GRC)
- EU Data Protection Regulation (EU GDPR)
- Strategic Risk Management, Analysis & Assessment
- Business Impact Analysis (BIA)
- IT Audit / IT Security Audit
- Business Continuity Management (BCM)
- Automotive Information Security (TISAX)
- Artificial Intelligence, Neural Networks

Overview

1. **From ISMS towards AIMS**
   – *how ISO42001 extends traditional ISO27001*
2. **Is AI a Risk or a Security Tool?**
   – *the dual impact of modern AI Systems*
3. **AI Governance under the EU AI Act**
   – *what auditors and CISOs must prepare for now*
4. **Is Standalone-ISO27001 still Sufficient?**
   – *AI requires more beyond traditional ISMS*
5. **Prompt Injection, Data Poisoning & Model Leakage**
   – *new attack classes against AI systems*
6. **Shadow AI in the Enterprise**
   – *the new art of Shadow-IT*
7. **How an AI system should Actually be Audited**
   – *new controls, competences and approaches*
8. **AI Supply Chain Risks**
   – *who is really-accountable for the AI model?*
9. **"AI Compliance-Theater"?**
   – *between real governance and marketing*
10. **Humans Remain the Biggest Risk**
    – *…even in the Age of AI*
11. **"Errare Mechanicum Est" – Non Solum Humanum**
    – *what happens when AI makes a Mistake*
12. **AI – employed in Auditing**
    – *will the auditor be replaced?*
13. **Agentic AI & Autonomous Systems**
    – *the next governance challenge?*
14. **The Myth of Confidentuiality in an AI Model**
    – *how to protect data residing inside of an AI model*
15. **Zero-Trust meets AI**
    – *security architectures for AI-native enterprises*

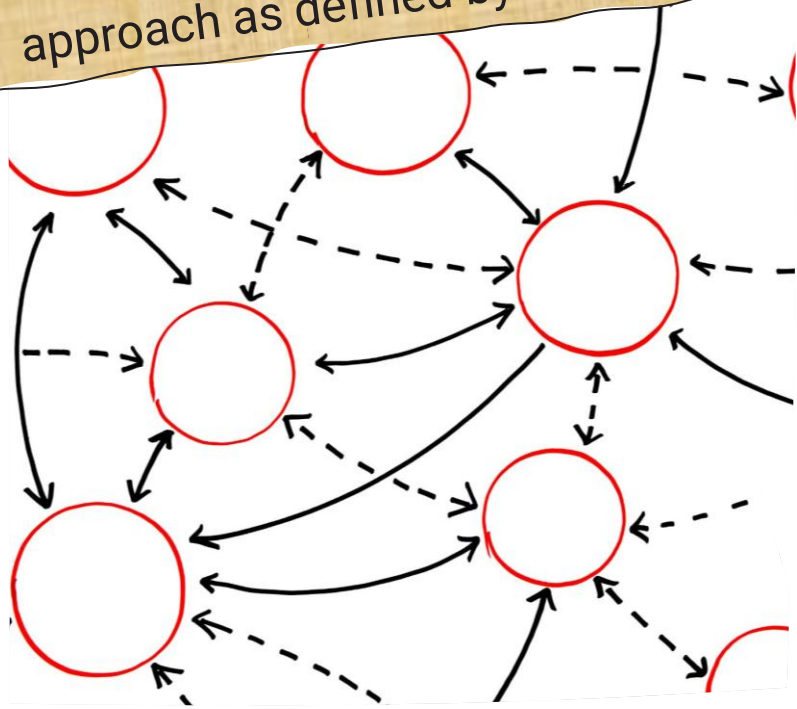
**ISO 42001 defines the framework –
CAIP enables the people to implement it.**

From ISO27001-based ISMS towards an ISO42001-based AIMS

How does ISO42001 extend the traditional ISO27001?

- **Scope Expansion**: From Information Assets to AI Systems & Lifecycle
  → data, models, training, deployment, monitoring
- **New Risk Dimensions Beyond Traditional Security**: From CIA to AI-Specific Risks
  → bias, explainability, robustness, unintended outcomes
- **Governance Extension**: From Security Controls to Ethical & Responsible AI
  → fairness, accountability, transparency, human oversight
- **Control Framework Enhancement**: New AI-Specific Controls & Processes
  → model validation, data governance, performance monitoring
- **Roles & Responsibilities**: Introduction of AI Accountability Structures
  → AI owners, model risk governance, cross-functional oversight
- **Lifecycle Integration**: From Static ISMS to Continuous AI Lifecycle Management
  → training–testing–deployment–retraining cycles
- **Regulatory Alignment**: Bridging ISO Standards with EU AI Act Requirements
  → risk classification, compliance evidence, auditability
- **From Security Compliance to Trustworthy AI**: From Information Security to AI Trustworthiness
  → reliability, safety, explainability as audit dimensions

Managing AI requires more than extending existing ISMS structures
– It demands a dedicated and lifecycle-oriented governance approach as defined by ISO 42001.



PECB

Is AI a Risk or a Security Tool?

The Dual-Impact of Modern AI Systems

AI is not just a new risk or a new tool!
– It fundamentally breaks the traditional separation between attacker and defender, forcing organizations to rethink governance altogether.

- **AI as a Force Multiplier for Cyber Attacks**
  → *automation, phishing, deepfakes, vulnerability discovery*
- **AI as an Enabler for Advanced Defense Capabilities**
  → *SOC automation, anomaly detection, threat intelligence*
- **Generative AI in the Enterprise**: Productivity vs. Data Exposure
  → *prompt leakage, uncontrolled data flows*
- **New Attack Surfaces**: AI Models, Pipelines and APIs
  → *model exploitation, prompt injection, data poisoning*
- **AI-Augmented Decision Making**: Efficiency vs. Loss of Control
  → *overreliance, hallucinations, opaque decisions*
- **Dual-Use Dilemma**: The Same Technology for Offense and Defense
  → *attackers and defenders using identical capabilities*
- **Speed & Scale**: Acceleration of Both Risks and Responses
  → *real-time attacks vs. real-time detection*
- **Governance Challenge**: Managing Trade-offs Between Innovation and Security
  → *risk appetite, control frameworks, regulatory pressure*

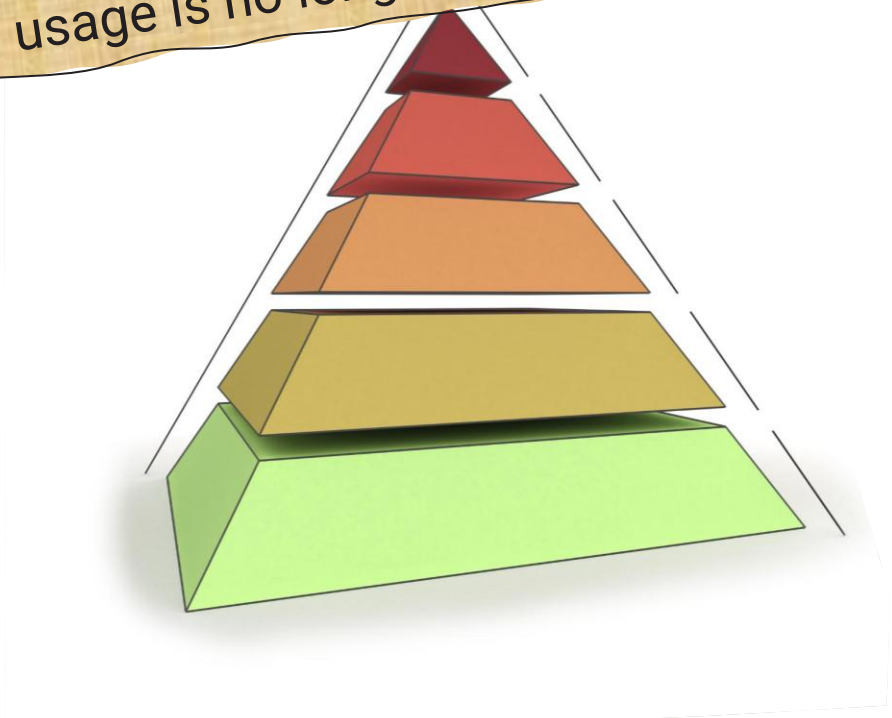
PECB

AI Governance under the EU AI Act

What do Auditors and CISOs must Prepare for Now?

- **Risk-Based Classification of AI Systems**
  → *prohibited, high-risk, limited-risk, minimal-risk*
- **Mandatory Requirements for High-Risk AI Systems**
  → *risk management, data governance, human oversight*
- **AI Governance & Accountability Structures**
  → *roles, responsibilities, three-lines-of-defense integration*
- **Documentation & Auditability Expectations**
  → *technical documentation, logging, traceability*
- **Continuous Monitoring & Post-Market Surveillance**
  → *performance tracking, incident reporting*
- **Integration with Existing Frameworks**: ISO 42001 / ISO 27001 / GDPR
  → *control alignment, governance synergies*
- **Third-Party & Supply Chain Obligations**
  → *provider vs. deployer roles, vendor risk, model transparency*
- **Enforcement, Liability & Regulatory Exposure**
  → *fines, accountability, reputational impact*

The EU AI Act transforms AI governance from a 'best practice' discussion into a regulatory obligation.
It creates the illusion of informal AI usage is no longer sustainable.

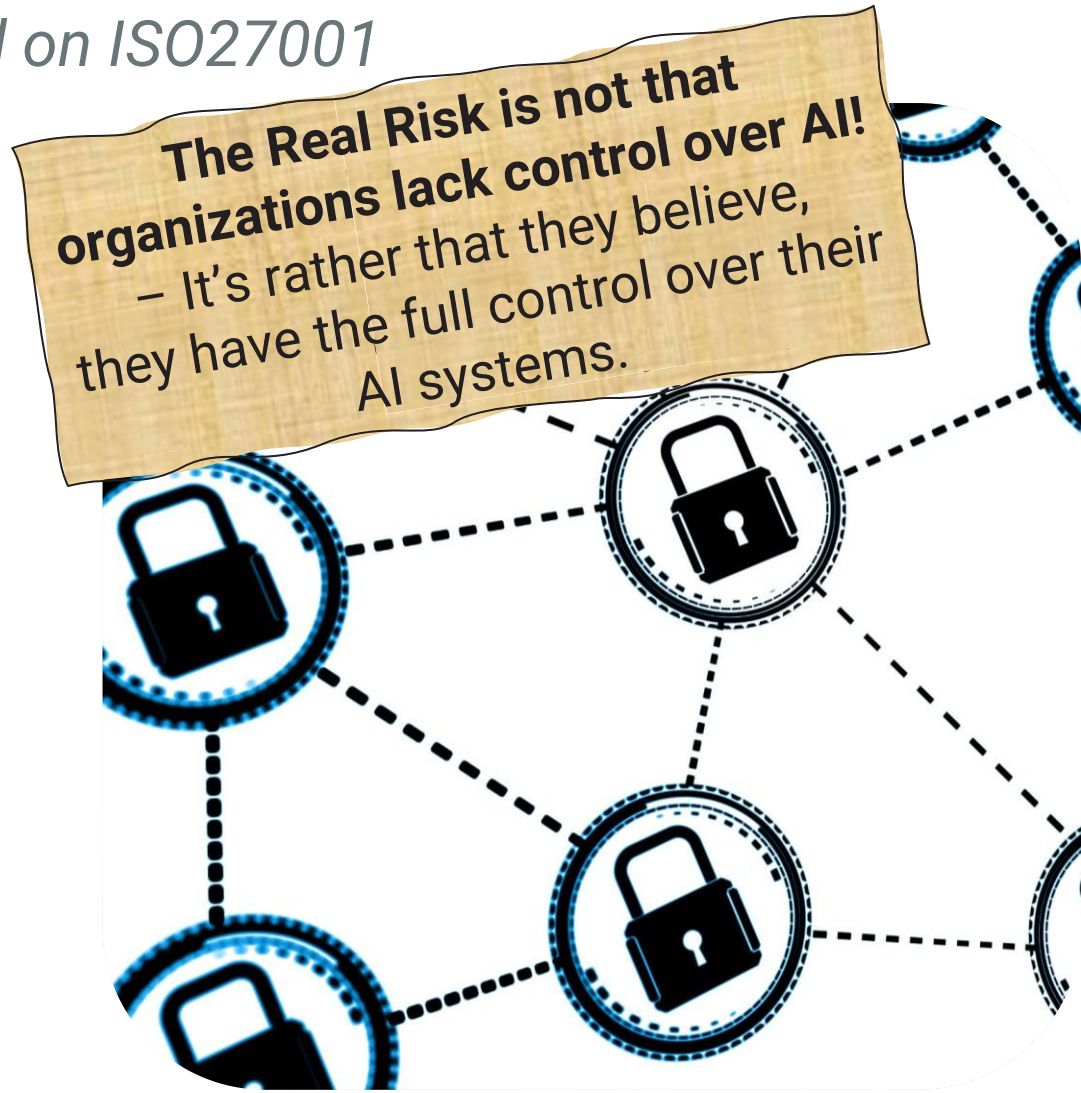


PECB

Is Standalone-ISO27001 still Sufficient?

AI Requires More beyond Traditional ISMS based on ISO27001

- **Limited Scope**: Focus on Information, Not AI Behavior
  → *no coverage of model logic, learning dynamics, decisions*
- **Missing AI-Specific Risk Dimensions**
  → *bias, hallucinations, explainability, model drift*
- **Lack of Controls for AI Lifecycle Management**
  → *no governance for training, validation, retraining*
- **Insufficient Data Governance for AI Use Cases**
  → *training data quality, lineage, representativeness*
- **No Governance for Automated Decision-Making**
  → *lack of oversight, accountability, escalation paths*
- **Inadequate Coverage of AI-Specific Attack Vectors**
  → *prompt injection, data poisoning, model manipulation*
- **Compliance Gap with Emerging Regulations (EU AI Act)**
  → *missing requirements for high-risk AI systems*
- **False Sense of Security**: "ISO 27001 Certified = AI Secure"
  → *governance illusion, incomplete risk coverage*

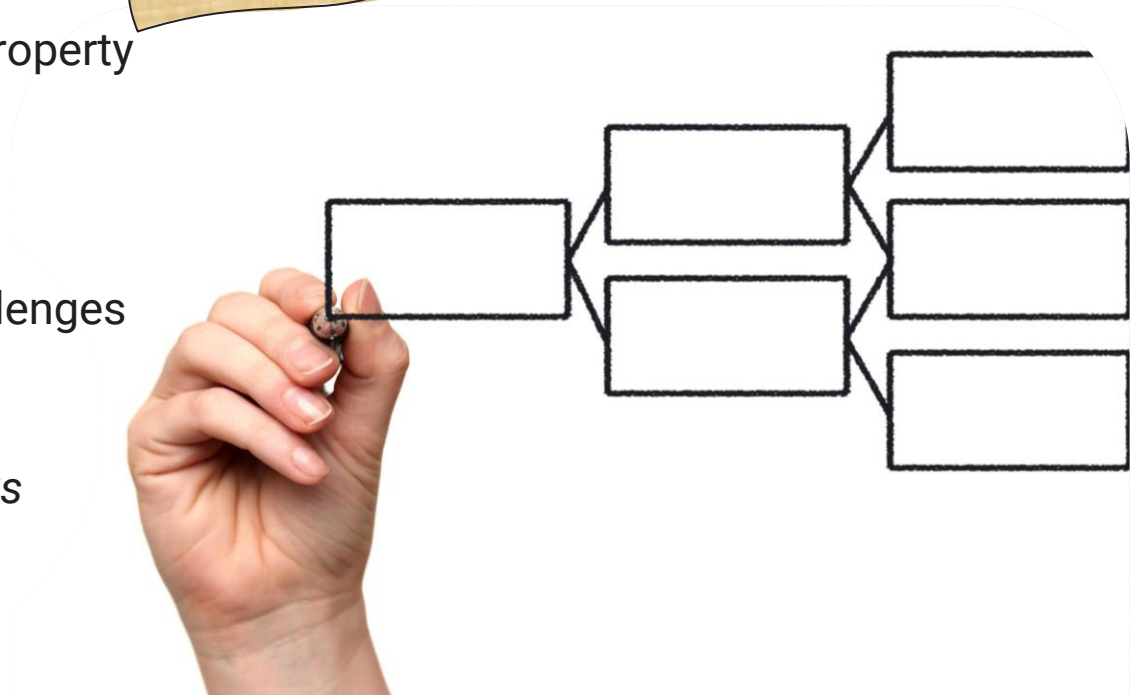


Prompt Injection, Data Poisoning & Model Leakage

Be Prepared for New Classes of Attacks against AI Systems!

- **New Attack Paradigm**: Targeting AI Logic Instead of Infrastructure
  → *manipulation of prompts, models, and outputs*
- **Prompt Injection**: Manipulating AI Behavior at Inference Time
  → *bypassing safeguards, extracting sensitive information*
- Data Poisoning: Corrupting Training and Fine-Tuning Data
  → *influencing model behavior before deployment*
- **Model Leakage**: Exposure of Sensitive Data and Intellectual Property
  → *training data extraction, model inversion risks*
- **Vulnerabilities in AI Pipelines and RAG Architectures**
  → *external data sources, embeddings, API integrations*
- **From Input Validation to Context Security**: New Defense Challenges
  → *lack of traditional boundaries and control points*
- **Detection & Monitoring Limitations in AI Environments**
  → *difficulty identifying manipulated outputs or hidden attacks*
- **Need for AI-Specific Security Controls and Governance**
  → *secure prompt design, data validation, model oversight*

The real challenge is that AI systems no longer fail only when systems break!
— It's when they behave exactly as designed, just under manipulated conditions.



Shadow AI in the Enterprise

Is Shadow-AI a New Kind of Shadow-IT in Your Enterprise?

- **From Shadow IT to Shadow AI**: Uncontrolled AI Usage by Employees
  → *use of ChatGPT, Copilots, external AI tools without governance*
- **Hidden Data Flows**: Sensitive Information in Prompts and Outputs
  → *data leakage, loss of confidentiality, regulatory risks*
- **Lack of Visibility and Control over AI Usage**
  → *missing inventories, undocumented AI use cases*
- **Circumventing Existing Security & Compliance Controls**
  → *bypassing DLP, access controls, approved processes*
- **Third-Party Risks through External AI Services**
  → *uncontrolled data sharing with providers, unclear processing*
- **Impact on Intellectual Property and Confidentiality**
  → *exposure of business logic, code, customer data*
- **Governance Gap**: Policies Not Adapted to AI Usage
  → *missing AI usage guidelines, unclear acceptable use)*
- **Need for Integrated AI Governance and Awareness Measures**
  → *policies, training, monitoring, technical controls)*

PECB

Why “AI Professional” Education matters Now

AI Governance requires not only frameworks like ISO 42001 – but also qualified professionals capable of implementing them in practice.

Why does AI Governance require Qualified Professionals?

- **Bridges the competence gap**: From ISO 27001/42001 Frameworks to Implementation
  → translating governance requirements into operational practice
- **Comprehensive AI Governance Knowledge**: Risk Management, Ethics, Compliance and Lifecycle Management
  → aligned with ISO 42001 and regulatory expectations
- **AI-Specific Risk & Control Expertise**: Bias, Explainability, Robustness and Security
  → understanding and managing AI-specific risk dimensions
- **Support for EU AI Act Readiness**: Governance, Documentation and Auditability Requirements
  → enabling compliance with emerging regulatory frameworks
- **Enhanced Auditor & CISO Capabilities**: Interdisciplinary Skills across Security, Data and AI
  → bridging IT audit with AI system understanding
- **Practical Application Focus**: Real-world Scenarios, Use Cases and Governance Implementation
  → enabling immediate application in organizations
- **Globally Recognized Certification (PECB)**: Professional Credibility and Structured Knowledge Validation
  → demonstrating competence in AI governance and risk
- **Foundation for Trustworthy AI Adoption**: Enabling Responsible, Transparent and Controlled AI Usage
  → building trust in AI-driven decisions and systems

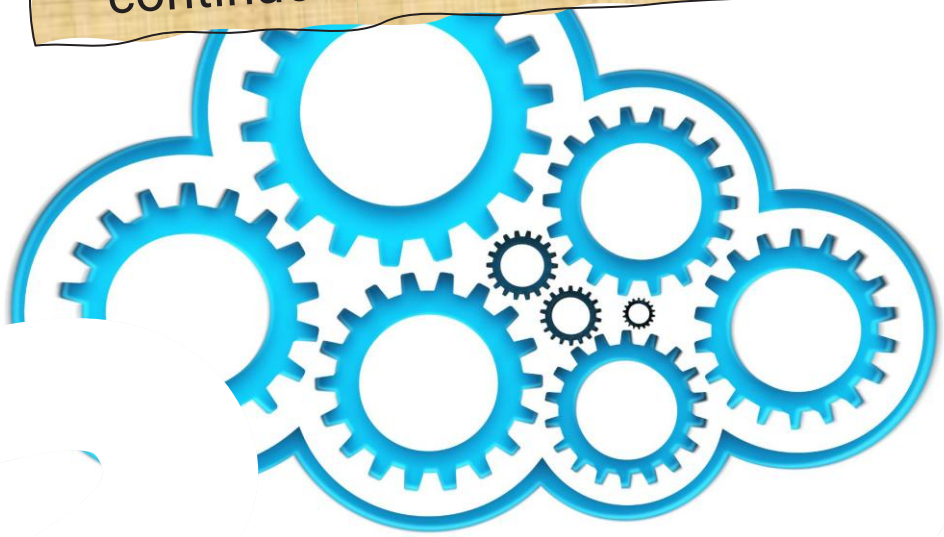
PECB

Auditing Approaches for AI Systems

How to Audit AI: New Controls, Auditor's Competences and Auditing Approaches?

- **From Control Testing to System Behavior Validation**
  → *evaluating outputs, decisions, and model performance*
- **AI-Specific Controls Beyond Traditional ISMS**
  → *model validation, bias testing, explainability checks*
- **Data-Centric Auditing**: Quality, Lineage and Representativeness
  → *training data governance as audit focus*
- **Lifecycle Auditing**: From Training to Deployment and Monitoring
  → *continuous assessment across AI lifecycle stages*
- **Evidence Challenges**: Limited Transparency and Black-Box Models
  → *explainability gaps, traceability limitations*
- **New Audit Techniques**: Testing Prompts, Outputs and Scenarios
  → *adversarial testing, simulation-based validation*
- **Expanded Auditor Competences**: Interdisciplinary Skill Sets
  → *AI, data science, risk, and regulatory knowledge*
- **Alignment with ISO 42001 and Regulatory Expectations**
  → *structured framework for AI governance audits*

Auditing AI is not just an extension of IT auditing!
— It requires a fundamentally new approach that focuses on behavior, data and continuous risk management.

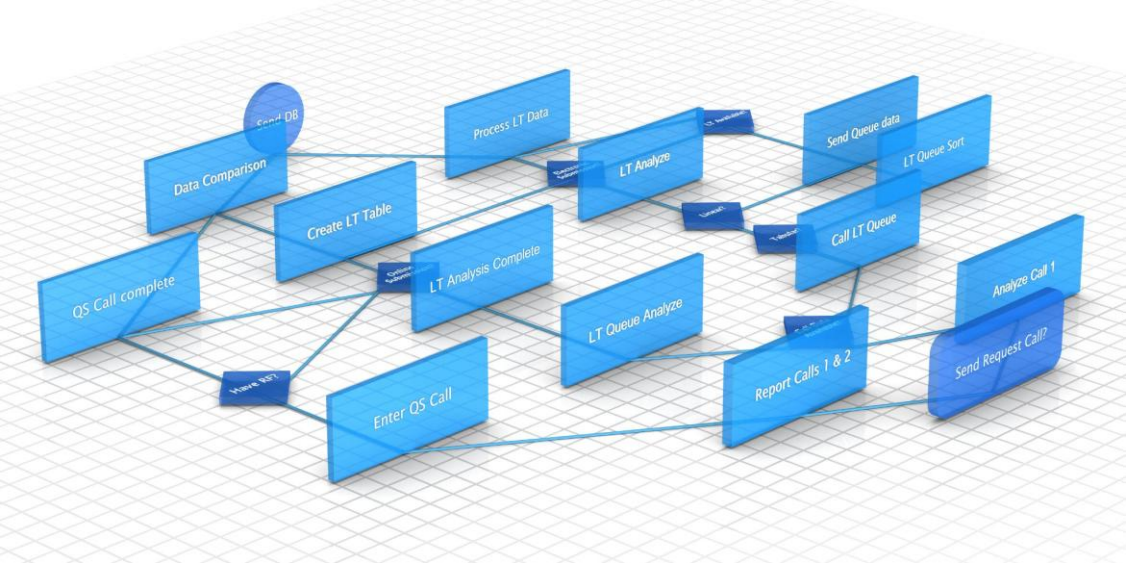


AI Supply Chain Risks

Who is Accountable and Responsible for the AI Model and AI-based Business Processes?

- **Complex AI Ecosystems**: Multiple Actors, Shared Responsibilities
  → *providers, deployers, integrators, data suppliers*
- **Blurred Accountability Across the AI Value Chain**
  → *unclear ownership of risks, decisions, and outcomes*
- **Dependence on Third-Party Models and Cloud Providers**
  → *external LLMs, APIs, managed AI services*
- **Lack of Transparency and Model Provenance**
  → *unknown training data, hidden model behavior*
- **Risks from Open-Source Models and Components**
  → *unverified code, hidden vulnerabilities, licensing risks*
- **Integration Risks in AI-Based Business Processes**
  → *embedding AI into critical workflows and decisions*
- **Vendor Risk Management Gaps for AI Systems**
  → *traditional TPRM not covering AI-specific risks*
- **Need for End-to-End AI Supply Chain Governance**
  → *contractual controls, transparency, continuous oversight*

In AI, responsibility cannot be outsourced!
— even if the model creation and operation is outsourced.

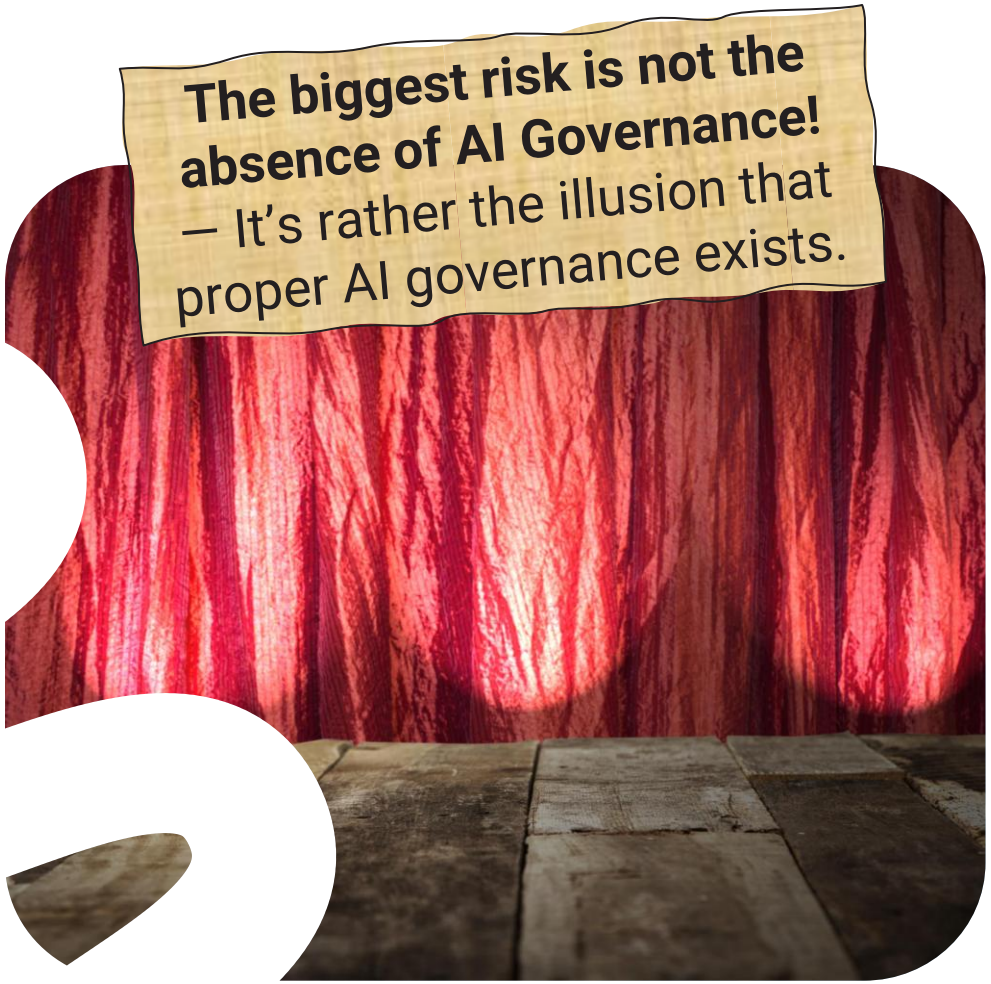


PECB

The “AI-Compliance Theater”

Between Real AI Governance and Marketing. What’s behind the “AI Inside” Label?

- **“AI Inside” as a Marketing Label vs. Real Governance**
  → *branding without underlying control structures*
- **Superficial Compliance vs. Effective Risk Management**
  → *policies exist, but are not operationalized*
- **Certification-Driven Approach vs. True Maturity**
  → *focus on passing audits instead of managing risks*
- **Lack of Measurable AI Governance Effectiveness**
  → *missing KPIs, unclear control performance*
- **Overreliance on Documentation Instead of System Understanding**
  → *paper-based compliance vs. real behavior assessment*
- **Misalignment Between Business Innovation and Governance**
  → *AI adoption outpaces control frameworks)*
- **Illusion of Control Through Existing Frameworks** (e.g., ISO27001)
  → *assuming current controls are sufficient for AI*
- **From Compliance Theater to Real AI Governance Maturity**
  → *embedding accountability, transparency, and continuous oversight*



PECB

The Human remains the Big Risk

… even in the Age of AI, since Made, Controlled and Governed by Humans

- **Human Error Amplified by AI Capabilities**
  → *incorrect prompts, misinterpretation of outputs, operational misuse*
- **Overreliance on AI and Automation Bias**
  → *blind trust in AI decisions, reduced critical thinking*
- **AI-Enhanced Social Engineering Attacks**
  → *deepfakes, voice cloning, highly personalized phishing*
- **Lack of Awareness and Misuse of AI Tools**
  → *untrained employees, unintended risk exposure*
- **Insufficient Human Oversight of AI Systems**
  → *lack of review, missing escalation mechanisms*
- **Misalignment Between Human Intent and AI Outcomes**
  → *unintended consequences, misunderstood system behavior*
- **Accountability Gaps in Human–AI Interaction**
  → *unclear responsibility for AI-driven decisions*
- **Need for Human-Centric AI Governance and Training**
  → *awareness, clear responsibilities, embedded oversight*

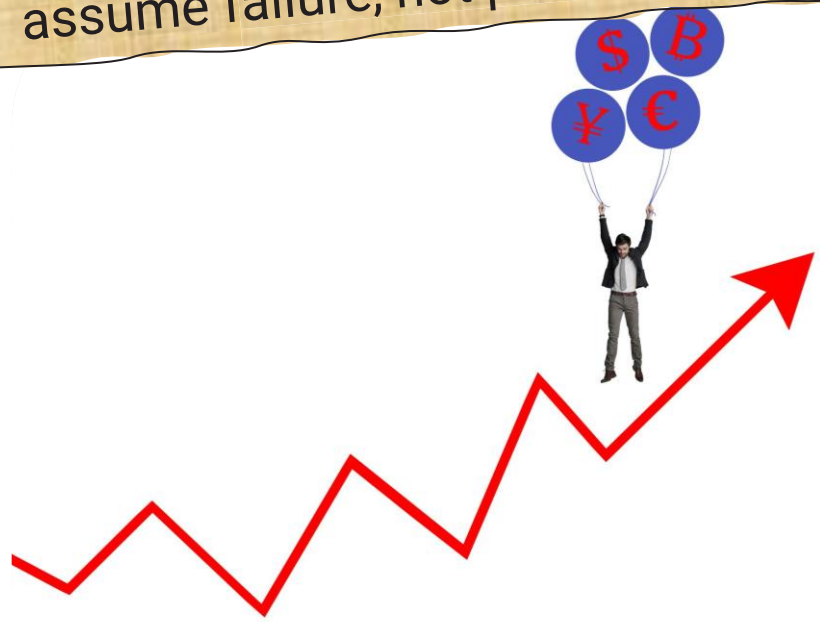
AI does not remove the human factor!
— It rather amplifies it, making governance of human behavior more critical than ever.

Errare Mechanicum Est – Non solum Humanum

What happens when AI Makes a Mistake, takes a Wrong Decision or Hallucinates?

- **AI Errors Are Systemic, Not Isolated**
  → *errors occur at scale across automated processes*
- **Hallucinations and Unreliable Outputs**
  → *plausible but incorrect or misleading results*
- **Lack of Explainability and Root Cause Analysis**
  → *difficulty understanding why decisions were made*
- **Cascading Effects in Automated Decision Chains**
  → *errors propagate across interconnected systems*
- **Impact on Business, Compliance and Reputation**
  → *financial loss, regulatory breaches, trust erosion*
- **Liability and Accountability for AI Decisions**
  → *unclear ownership of faulty outcomes*
- **Need for Incident Management and Escalation Processes**
  → *detection, reporting, and response to AI failures*
- **Designing for Failure: Resilient and Controlled AI Systems**
  → *safeguards, human-in-the-loop, fallback mechanisms*

AI does not fail like humans!
— It fails differently, at scale,
and often without warning.
That is why governance must
assume failure, not perfection.

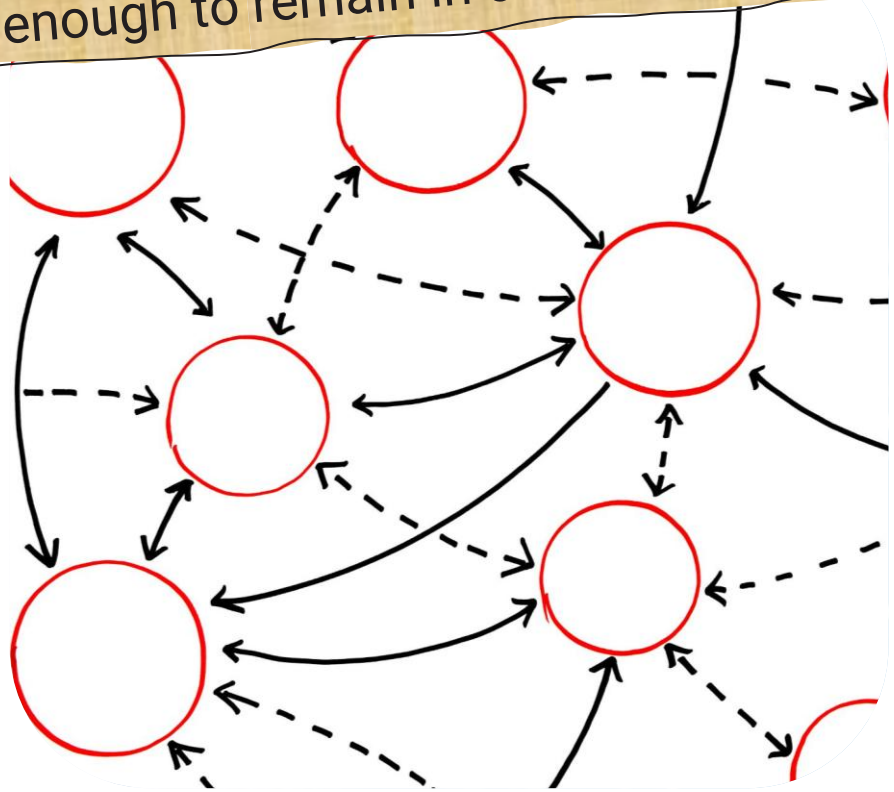


AI as a Tool for Auditing Human Behavior

Do Auditors have to Fear about their Jobs?

- **AI-Driven Audit Automation**: From Sampling to Full Data Coverage
  → *continuous auditing, anomaly detection across entire datasets*
- **AI as an Augmentation Tool** – It's not a Replacement for Auditors
  → *supporting analysis, not substituting professional judgment*
- **Automated Evidence Collection and Pattern Recognition**
  → *log analysis, behavioral insights, risk identification*
- **Shift from Manual Testing to Data-Driven Auditing**
  → *analytics-based assurance, real-time monitoring*
- **New Risks**: Bias, Overfitting and Incorrect AI Conclusions
  → *flawed models leading to incorrect audit results*
- **Need for Explainable and Audit-Ready AI Tools**
  → *transparency, traceability of AI-supported conclusions*
- **Changing Role of Auditors**: From Control Testing to Strategic Advisors
  → *interpretation, risk assessment, governance guidance*
- **Future Competences**: Combining Audit Expertise with AI Understanding
  → *data literacy, AI governance knowledge, critical thinking*

The real Question is not whether AI will replace Auditors!
– but whether Auditors will adapt fast enough to remain in control of AI.



Agentic AI and Autonomous Systems

Do Autonomous AI-based Agents Form the Next Governance Challenge?

- **From Assistive AI to Autonomous Decision-Making Systems**
  → *agents executing tasks independently without human intervention*
- **Delegation of Authority to AI Agents**
  → *automated actions impacting business processes and controls*
- **Loss of Transparency and Human Oversight**
  → *limited visibility into agent decisions and actions*
- **Dynamic and Self-Improving System Behavior**
  → *continuous learning, adaptation, and unpredictability*
- **Complex Interactions in Multi-Agent Ecosystems**
  → *autonomous agents interacting with other systems and agents*
- **New Risk Dimensions**: Control Loss and Unintended Actions
  → *cascading actions, misaligned objectives, emergent behavior*
- **Governance Challenges**: Defining Boundaries and Safeguards
  → *policies for autonomy levels, human-in-the-loop mechanisms*
- **Need for Advanced AI Governance Frameworks** (ISO 42001)
  → *structured oversight, accountability, and lifecycle control*

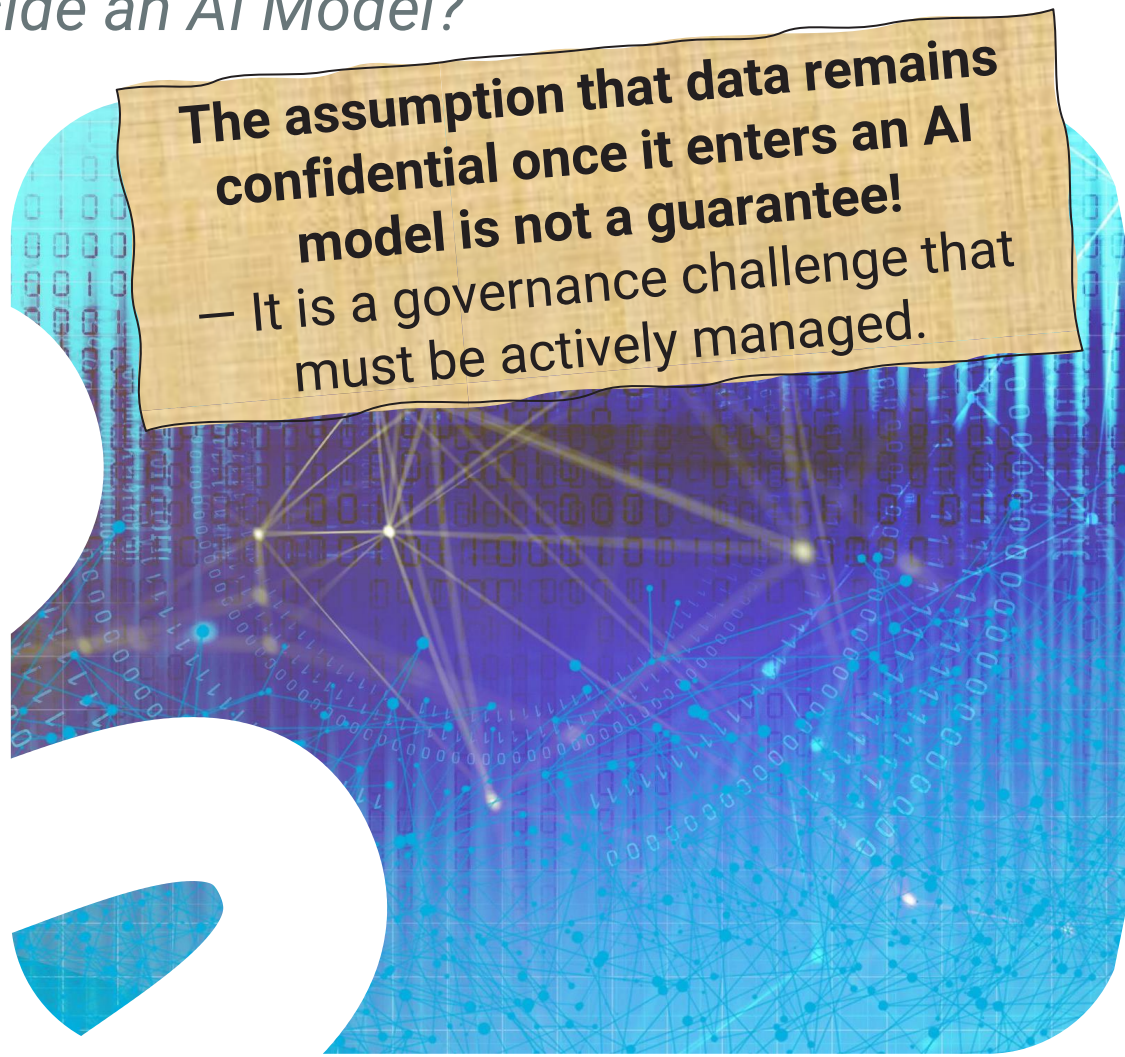
The real question is no longer whether AI can act autonomously!
— It's rather, whether organizations can still govern systems that act on their own.

PECB

The Myth of Confidentiality on an AI Model

Are we able to Protect Information Residing Inside an AI Model?

- **Training Data Becomes Embedded in Model Behavior**
  → *indirect memorization, latent data representation*
- **Risk of Data Extraction and Model Inversion Attacks**
  → *reconstruction of sensitive training data*
- **Leakage through Prompts and Model Outputs**
  → *unintended disclosure of confidential information*
- **Limited Control over Data Once Used for Training**
  → *inability to fully delete or trace data usage*
- **Exposure Risks in Fine-Tuning and RAG Architectures**
  → *embedding internal knowledge into AI responses*
- **Challenges in Enforcing Data Confidentiality Requirements**
  → *mismatch with classical confidentiality controls*
- **Dependence on Providers and Lack of Transparency**
  → *unclear data handling in external AI services*
- **Need for Technical Safeguards and Governance Controls**
  → *anonymization, access control, secure architectures*



PECB

Classical Zero-Trust Approaches meet AI

Are Security Architectures for an AI-native Enterprise Possible?

- **Zero Trust Principles Applied to AI Systems**
  → *"never trust, always verify" for users, data, models, and APIs*
- **New Trust Boundaries**: From Networks to AI Ecosystems
  → *models, pipelines, data flows, external services*
- **Identity & Access Management**: for Humans and Machines
  → *users, services, AI agents, API identities)*
- **Data Protection Across the AI Lifecycle**
  → *secure data handling from training to inference*
- **Securing AI Pipelines and Interfaces** (APIs, RAG, Integrations)
  → *end-to-end protection of AI components*
- **Continuous Verification of AI Behavior and Outputs**
  → *monitoring, validation, anomaly detection*
- **Challenges of Applying Zero Trust to Dynamic AI Systems**
  → *changing models, adaptive behavior, limited transparency*
- **Towards AI-Native Security Architectures and Governance**
  → *integration of Zero Trust with AI governance frameworks*

Zero Trust was designed for networks!
— Now it must evolve to handle systems that "think", "learn" and "act" on their own.

PECB

Summary

Are organizations able to govern AI?

- **AI Governance today is more illusion than reality in many organizations**
- **ISO 27001 alone is no longer sufficient to manage AI-related risks**
- **AI introduces fundamentally new risk classes beyond traditional IT security**
- **The EU AI Act transforms AI governance from optional to mandatory**
- **AI risks are amplified by scale, autonomy, and lack of transparency**
- **Governance gaps increasingly arise from third parties and Shadow AI usage**
- **Auditing AI requires new approaches, controls, and competencies**
- **ISO 42001 and qualified AI professionals, like those educated by the PECB's CAIP course, are key enablers for trustworthy AI adoption**

PECB

THANK YOU

✉ Roman.Krepki@forvismazars.com

in https://www.linkedin.com/in/roman-krepki-35ba512/