
DSX
Data Storage
Digital Conference

TLP:GREEN

Contains AI

Müll rein - Katastrophe raus

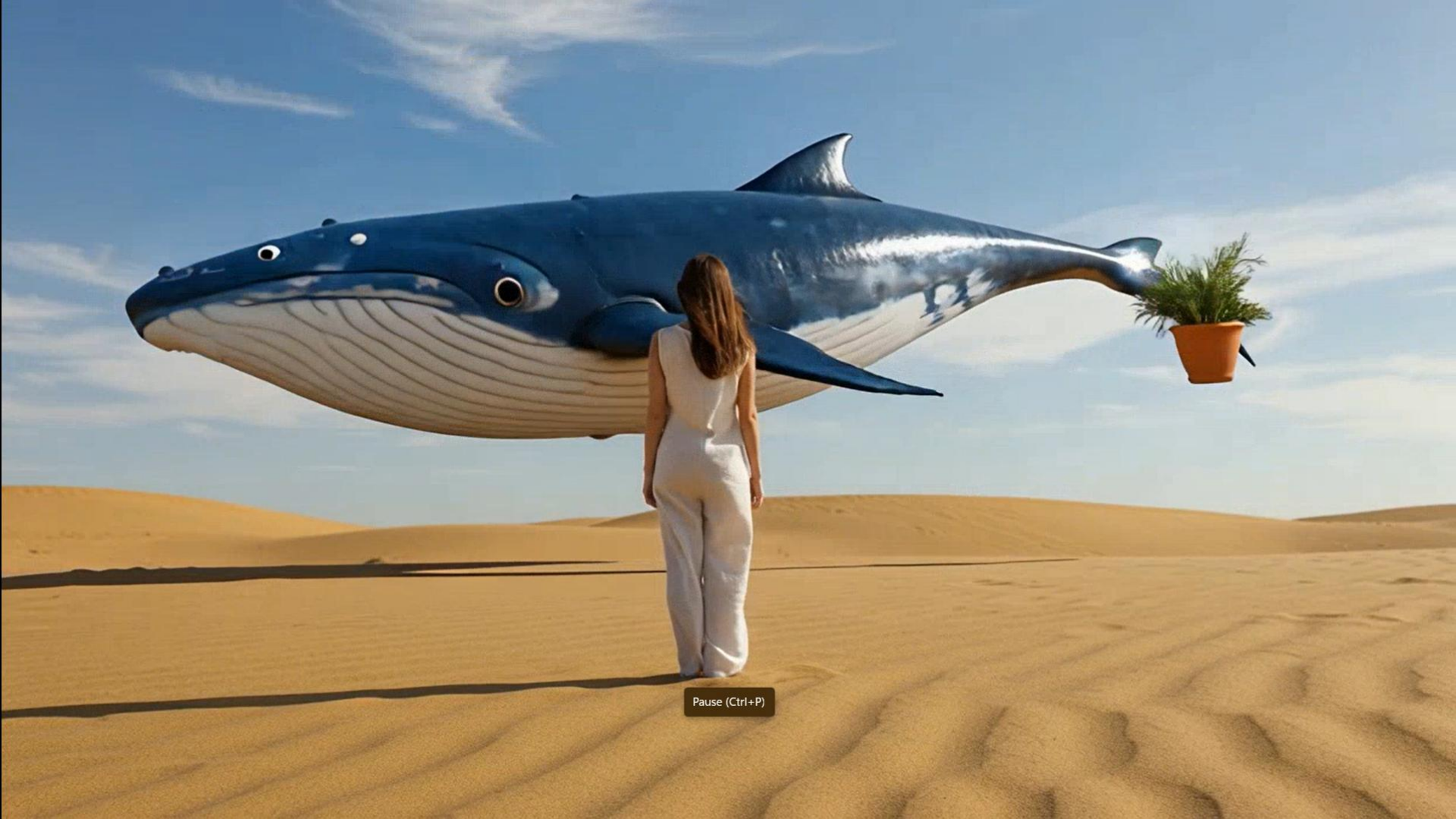
Warum KI sichere Daten braucht

Candid Wüest
Security Advocate @ xorlab
September 2025

DSX Data Storage Conference


Wunsch?


Pause (Ctrl+P)


Pause (Ctrl+P)

Was kann schief gehen mit KI?

New York City defends AI chatbot that advised entrepreneurs to break laws

Reuters

Air Canada found liable for chatbot's bad advice on plane tickets

Airline's cla
small claim

Jason

An AI-powered coding tool wiped out a software company's database, then apologized for a 'catastrophic failure on my part'

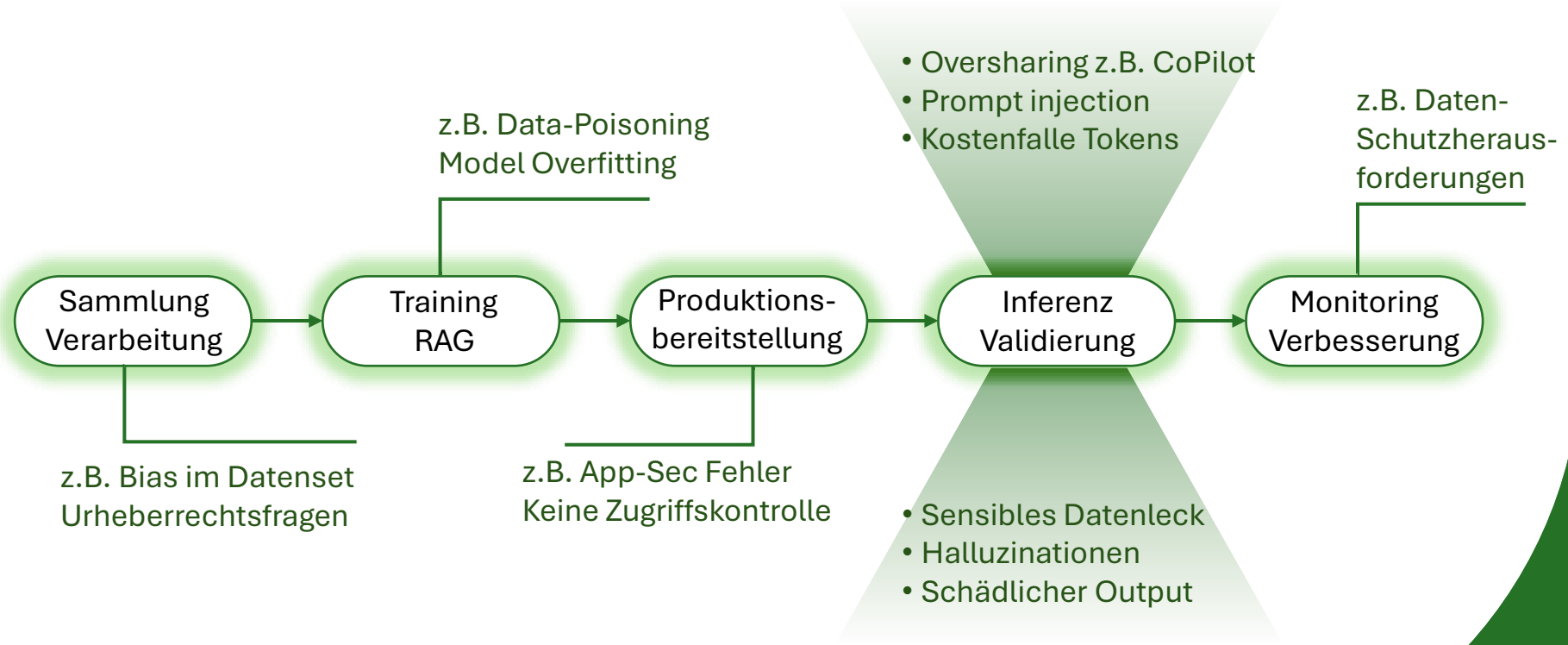
FORTUNE BY BEATRICE NOLAN TECH REPORTER July 23, 2025 at 7:22 AM EDT

KI funktionert nicht ohne (die richtigen) Daten

KI-Adaption – Realitätscheck

- ❖ Fragmentierte / verteilte Daten
- ❖ Mangelnde Datenqualität
- ❖ Ungeschützte KI Datenpipelines
- ❖ Angst und Unsicherheit bei Vorschriften

Datenprobleme im gesamten KI-Datenzyklus



Voraussetzung für Datenpipeline: Vertrauen

Transparenz

Woher kommen die Daten?
Lernen von meinen Fragen?

Kontrolle

Wer kann wann auf welche Daten wie zugreifen?

Schutz

Wie werden die sensiblen Daten gesichert?

Nachvollziehbarkeit

Wer hat was verändert?
Was war die Antwort?


Security ≠ Vertrauen ≠ Compliance

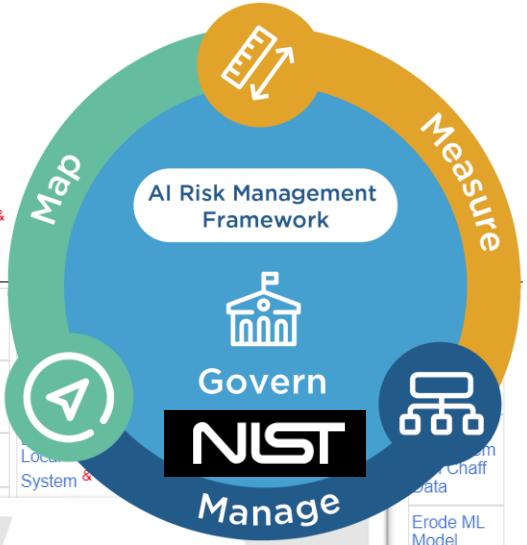
Secure by Design

- Datenzentrierte Sicherheit
- Robuste Zugriffskontrolle
- Sichere Infrastruktur & Betrieb
- KI Modellsicherheit: Schutz des Input/Outputs

KI Security & Risk Frameworks

MITRE ATLAS™

Reconnaissance &	Resource Development &	Initial Access &	ML Model Access	Execution &	Persistence &	Privilege Escalation &	Defense Evasion &	Credential Access &	Discovery &
5 techniques	7 techniques	6 techniques	4 techniques	3 techniques	3 techniques	3 techniques	3 techniques	1 technique	4 techniques
Search for Victim's Publicly Available Research Materials	Acquire Public ML Artifacts	ML Supply Chain Compromise	ML Model Inference API Access	User Execution &	Poison Training Data	LLM Prompt Injection	Evade ML Model	Unsecured Credentials &	Discover ML Model Ontology
Search for Publicly Available Adversarial Vulnerability Analysis	Obtain Capabilities &	Valid Accounts &	ML-Enabled Product or Service	Command and Scripting Interpreter &	Backdoor ML Model	LLM Plugin Compromise	LLM Prompt Injection		Discover ML Model Family
Search Victim-Owned Websites	Develop Capabilities &	Evade ML Model	Physical Environment Access	LLM Plugin Compromise	LLM Prompt Injection	LLM Jailbreak	LLM Jailbreak		Discover ML Artifacts
		Exploit Public-Facing Application							



AI TRiSM

Google's Secure AI Framework (SAIF)

OWASP Top 10 for LLM Applications

									Chaff Data
									Erode ML Model Integrity
									Cost Harvesting
									External Harms

Data Governance Strategie

Gut verwaltete, vertrauenswürdige Daten ermöglichen:

- ❖ Vertrauen der Stakeholder in KI
- ❖ Schnellere KI-Entwicklung
- ❖ Sicherere und zuverlässigere KI-Integration

Anwendung von Zero-Trust-Prinzipien auf KI-Pipelines:

- ❖ Kontinuierlicher authentifizierter Datenzugriff
- ❖ Überwachung von Modell-Inputs/Outputs
- ❖ Isolierung von Trainingsumgebungen

Fazit

- ❖ Ohne sichere Daten ist jede KI ein Risiko.
- ❖ Vertrauen ist die Grundlage der Datenpipeline.
- ❖ Nur eine ganzheitliche Data Governance Strategie macht KI nachhaltig nutzbar.

Ich wünsche mir ...

... mehr Daten-Sicherheit

Danke für die Aufmerksamkeit!

My LinkedIn

Candid Wüest
@candid.bsky.social
@MyLaocoon


DSX
Data Storage
Digital Conference