Building Trust in Technology:
A New Framework for Action

Date	10.09.2024
Place	T-Labs, Berlin
Author	Dr. Kim Nguyen, SVP Innovations & Carmen Dencker, Innovations Bundesdruckerei GmbH

bdr.

Agenda

1.Trust

2.Trust in Technology: A New Framework for Action

3.Two Examples:

- **Artificial Intelligence (in the Healthcare Sector)**

- **Cryptography**

4.Conclusion and Outlook

Trust

[...] Bevor wir in die Diskussion einsteigen, möchte ich gern über drei aus meiner Sicht zentrale Dinge reden, die es braucht, wenn es um KI und ihre Möglichkeiten geht:
Verständnis, Verantwortung und Vertrauen.

[...] Before we dive into the discussion, I would like to talk about three key elements that I believe are essential when it comes to AI and its potential:
Understanding, Responsibility, and Trust.

Frank-Walter Steinmeier, Federal President

Third Event in the Series "Forum Bellevue on the Transformation of Society," July 10, 2024, Berlin

Trust is the Foundation of Our Society

Trust enables people to be willing to take risks and show vulnerability

Trust is central to the interaction between people and systems

The increasing use of technology in various areas of life involves certain risks:

- Lack of transparency

- Lack of legal compliance

- Lack of traceability

- Lack of sovereignty

- Lack of security leading to dependencies

- Lack of fairness due to data bias
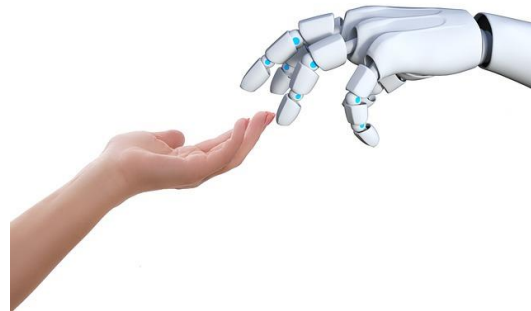
People's Trust in Technology Means

- Technology is fair, transparent, secure, and free from discrimination

- "I understand how the system arrives at its decisions"

- Appropriate measures for the protection of sensitive data are in place

- Sovereignty is ensured

- Legal compliance is ensured

Three dimensions of trustworthiness structure the analysis of risks and the implementation of trust-building measures

Capability
The system has the competence to correctly perform specific tasks

Goodwill
The system operates for the benefit of the users and society

Integrity
The development process is trustworthy, and sensitive data is protected

Trust in Technology: A New Framework for Action

Capability: The system works as it is supposed to

My expectations for functionality are met

I can rely on the system in unforeseen situations

Goodwill: The system aligns with societal values
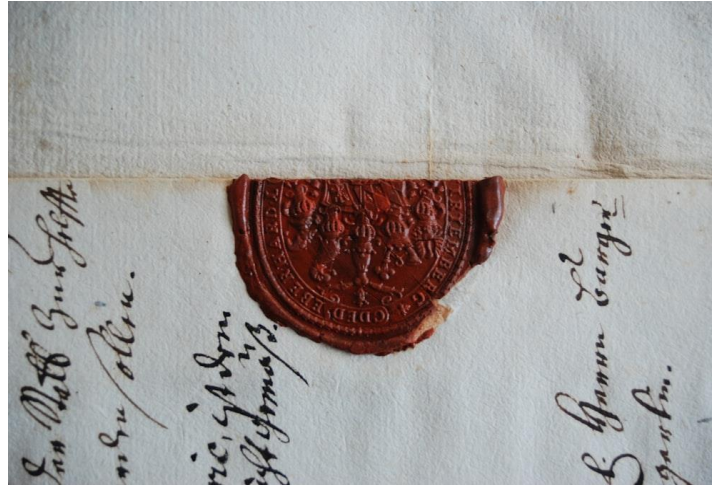
My data is protected

Legal compliance is maintained (e.g., copyrights, national regulations)

Results are transparent

Integrity: The product is trustworthy

The system adheres to national or international standards

The system is certified

Transparency & Accountability are ensured

Artificial Intelligence in the Healthcare Sector

Trust in AI systems is particularly critical in the healthcare sector

Sensitive Health Data

Critical Decisions

High ethical standards

What Could a Trustworthy AI in the Healthcare Sector Look Like? A Thought Experiment Based on Two Scenarios

Scenario 1: Reducing the quality and effort required for creating medical reports using AI

- Analysis of patient information such as diagnoses or medication plans

- Assisting the attending physician in creating a precise medical report for all healthcare professionals involved in the treatment

Trustworthiness in AI-Assisted Medical Report Creation

Capability

- The AI system has the technical capability to **accurately analyze and interpret** relevant patient data
- Generated medical reports consistently **meet high-quality standards**, regardless of the complexity of the medical cases

Benevolence

- **User Orientation**: The AI aims to minimize the workload for doctors and improve patient care through quick and accurate information delivery
- Doctors can **easily operate** the system and understand how it works

Integrity

- **Clear procedures** for the review and validation of AI-generated medical reports
- AI systems and their decision-making processes are **fully auditable**
- **Certification** of AI systems

What Could a Trustworthy AI in the Healthcare Sector Look Like? A Thought Experiment Based on Two Scenarios

Scenario 2: AI-Assisted Diagnostics to Improve Medical Decision-Making

- Implementation of an AI system capable of analyzing complex diagnostic data such as imaging, lab tests, and clinical symptoms

- Accelerates the diagnostic process through rapid analysis of large volumes of data

Vertrauenswürdigkeit in KI-gestützter Diagnostik

Capability

- **Diagnostic Precision**: The AI analyzes medical data such as imaging and lab tests with high accuracy
- **Reliable Detection**: Capable of recognizing complex patterns and identifying both rare and common conditions

Goodwill

- **Support in Clinical Practice**: Relieves medical staff by providing rapid and informed diagnostic insights
- **Intuitive user interface** facilitates the use and interpretation of diagnostic results by medical professionals

Integrity

- The AI is trained with **extensive, diversified datasets** to ensure broad coverage of medical conditions.
- **Verifiability and Compliance**: Adherence to medical standards and laws; capability for auditing and certification

Trust and AI

Data and technology Model

Governance / Traceability

Transparency and Certification

Trust in Cryptography

Cryptography enables secure digital interactions—and requires trust

Integrity, Confidentiality: Protection against unauthorized alterations of data and software, confidentiality of data

Authenticity: Origin of data can be verified independently

Free from backdoors

Trustworthiness and Cryptography

Capability

- Cryptography has the **technical capability** to implement the relevant use cases (encryption, signature)
- The **state of the art** is met regarding the **technical specifications** and the **runtime of the mechanisms** used
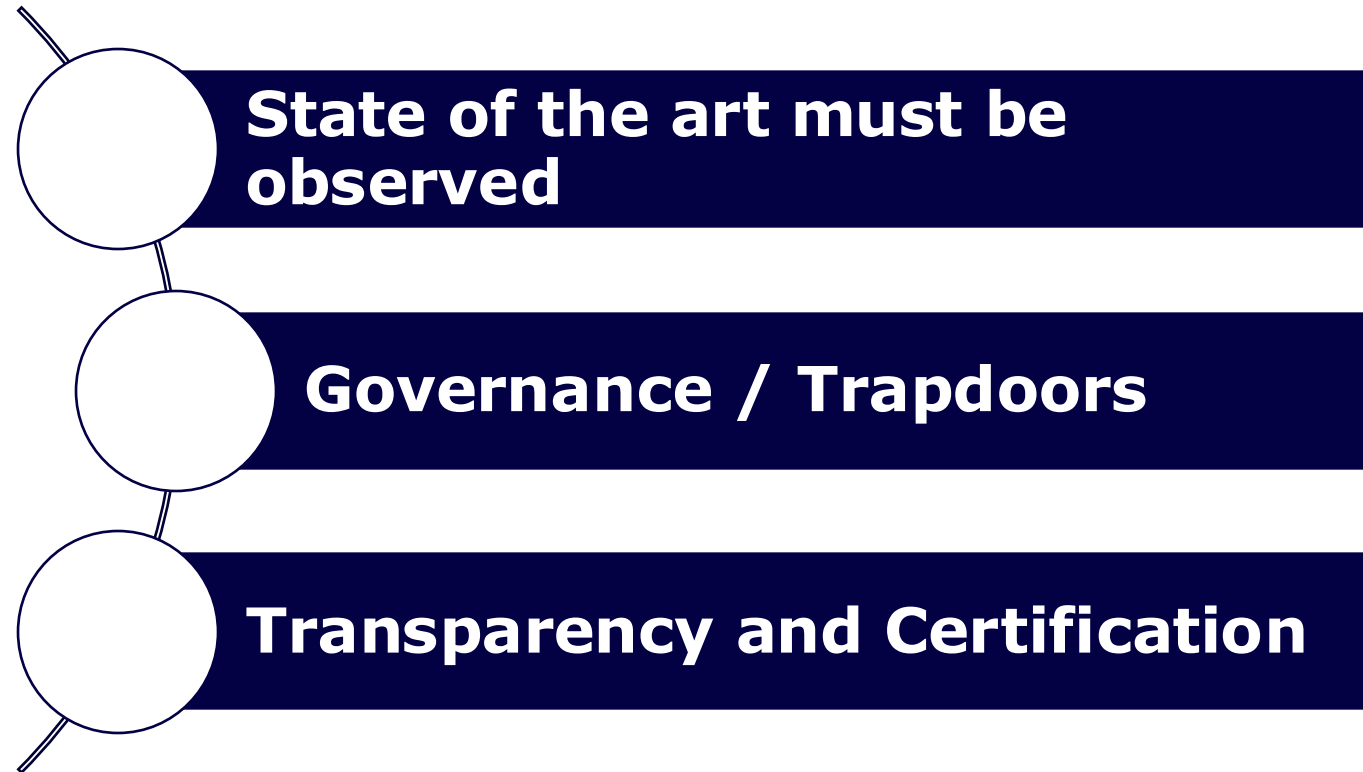
Goodwill

- **User Orientation**: Cryptography sustainably protects the integrity, authenticity, and confidentiality of user data without backdoors
- **User Sovereignty**: The user can independently and autonomously manage and use their key material

Integrity

- Clear procedures for the **review and validation of the mechanisms used**, ensuring they meet the state of the art and security standards
- Cryptographic mechanisms are **fully disclosed**, **auditable**, and **certified**

Trust in Cryptography

State of the art must be observed

Governance / Trapdoors

Transparency and Certification

Conclusion and Outlook

Conclusion and Outlook: Some Propositions...

Integrity: Functional Compliance through Testing & Certification

Integrity: (Formal) Security Certification, Open Source Approach

Benevolence: More of a Programmatic than a Technical Issue? Governance of the Company?

What questions arise for the sector of OPEN RAN along these dimensions?

Capability

Goodwill

Integrity

[…] Bevor wir in die Diskussion einsteigen, möchte ich gern über drei aus meiner Sicht zentrale Dinge reden, die es braucht, wenn es um KI und ihre Möglichkeiten (Ergänzung KN: und eigentlich Technologie insgesamt) geht: Verständnis, Verantwortung und Vertrauen.

[…] Before we dive into the discussion, I would like to talk about three key elements that I believe are essential when it comes to AI (*addition KN and actually technology in general*) and its potential: Understanding, Responsibility, and Trust.

Frank-Walter Steinmeier, Federal President

Third Event in the Series "Forum Bellevue on the Transformation of Society," July 10, 2024, Berlin

Outlook

- Trust is a key factor in gaining acceptance for technologies.

- Capability, goodwill and integrity are essential criteria for evaluation

- The whole Bundesdruckerei Gruppe is strongly guided by these values!

Technology and business is global – trust is local!

Dr. Kim Nguyen, SVP Innovations & Carmen Dencker, Innovations
Bundesdruckerei GmbH

Kim.nguyen@bdr.de