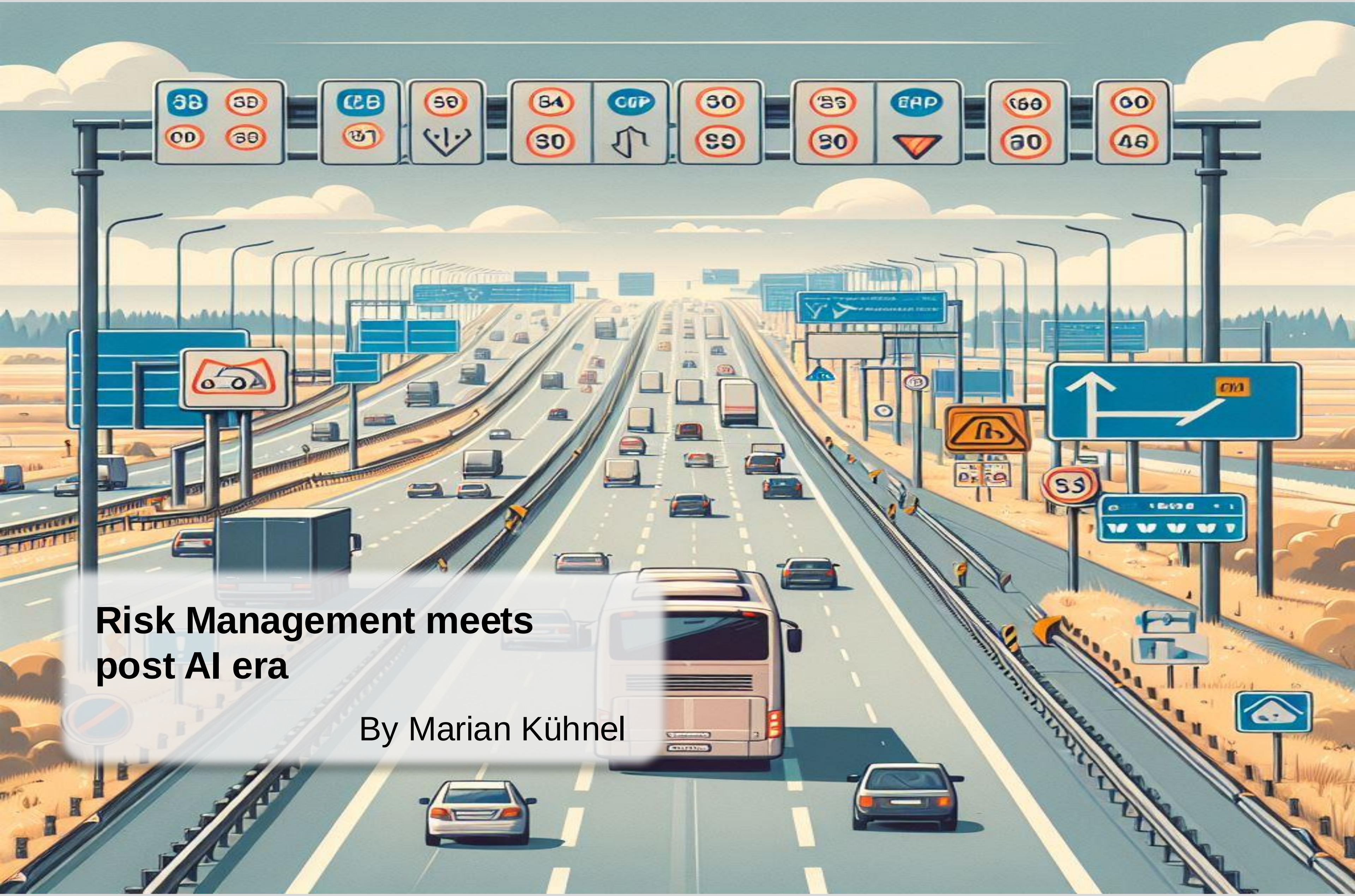
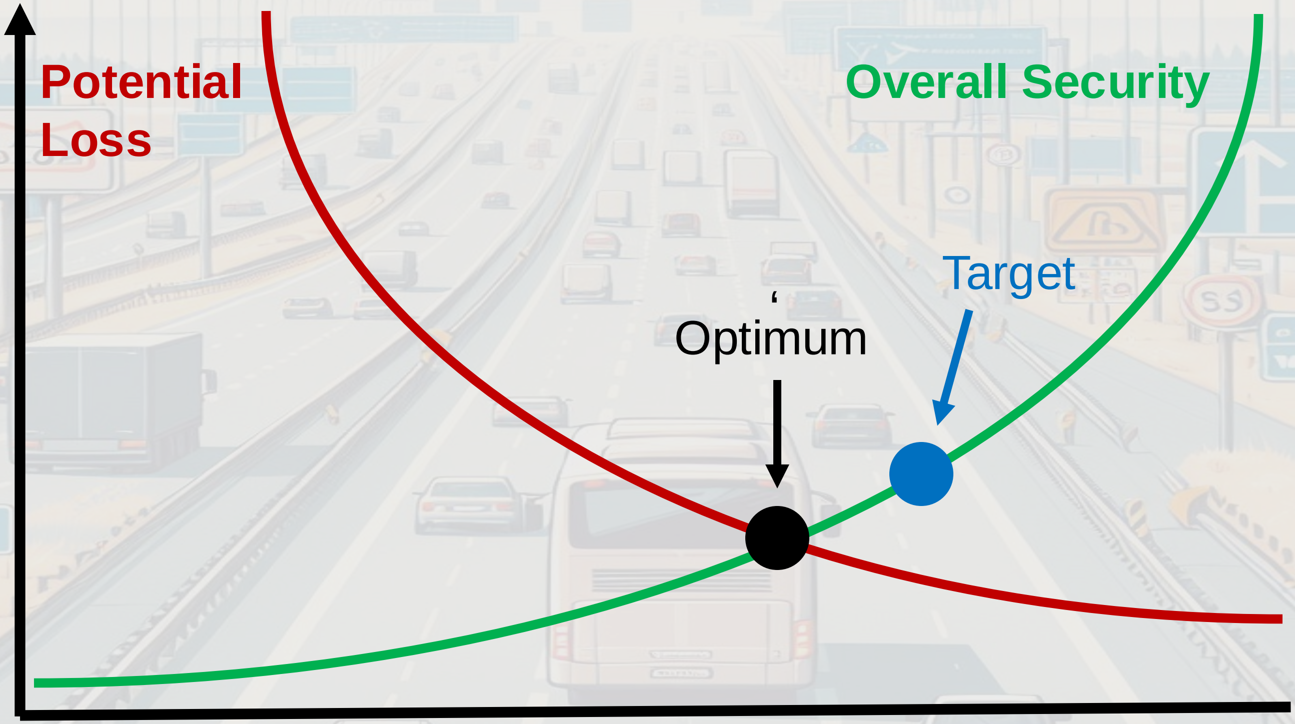


Risk Management meets post AI era

By Marian Kühnel

Risk Managment



Potential Loss

Overall Security

Target

‘

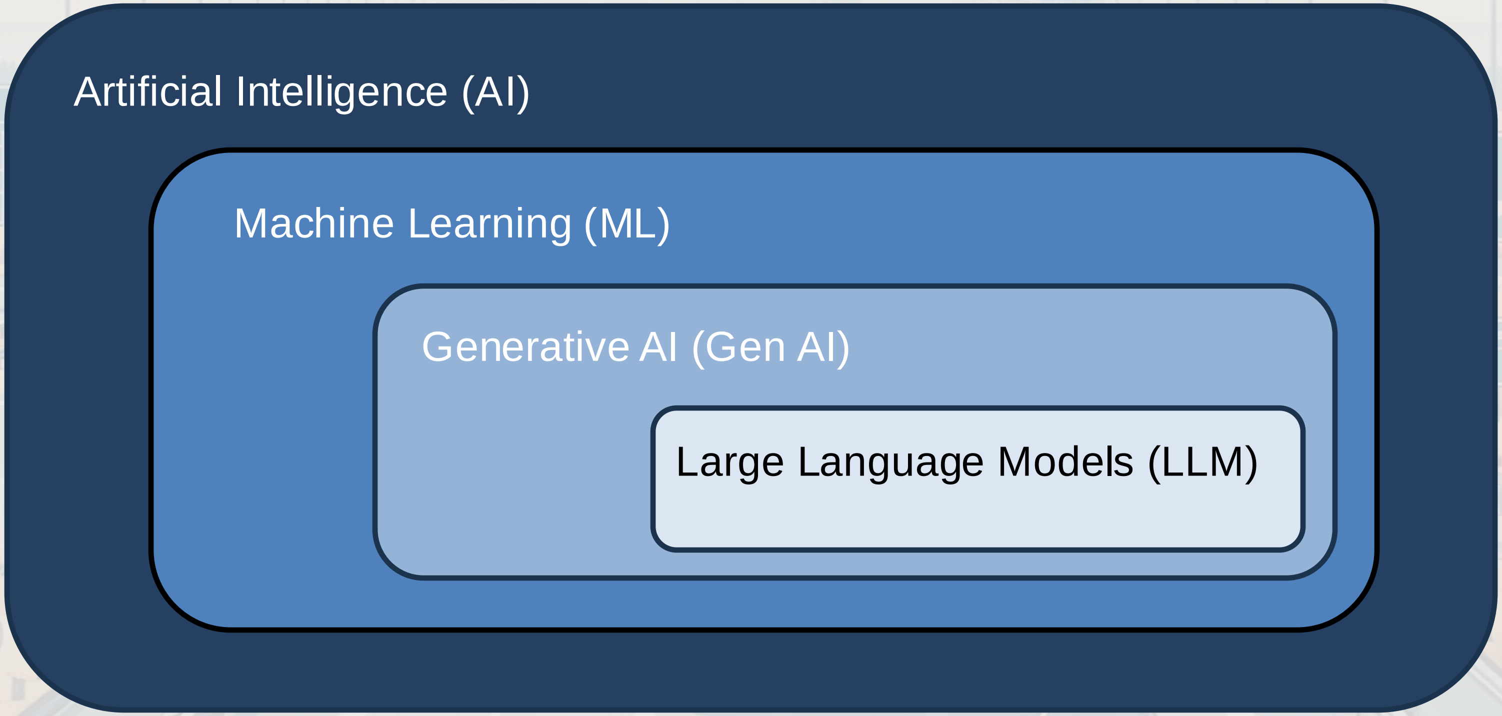
Optimum

AI Guardrails Principles

AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.

AI Guardrails Principles

ℹ **AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.**



Artificial Intelligence (AI)

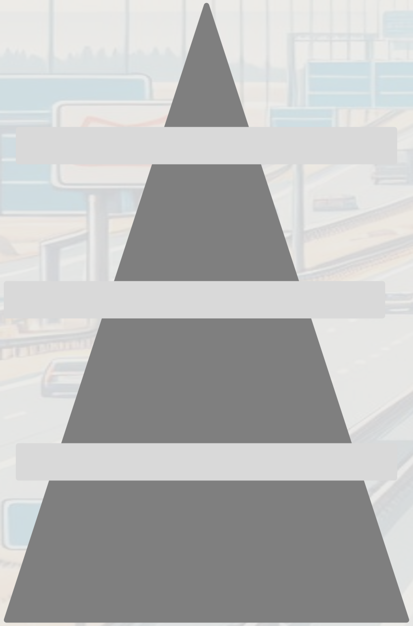
Machine Learning (ML)

Generative AI (Gen AI)

Large Language Models (LLM)

AI Guardrails Principles

AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.



Unaccetable risk - Secure scoring, AI-based manipulation, exploitation of vulnerabilities

High risk - AI safety components in critical infrastructures

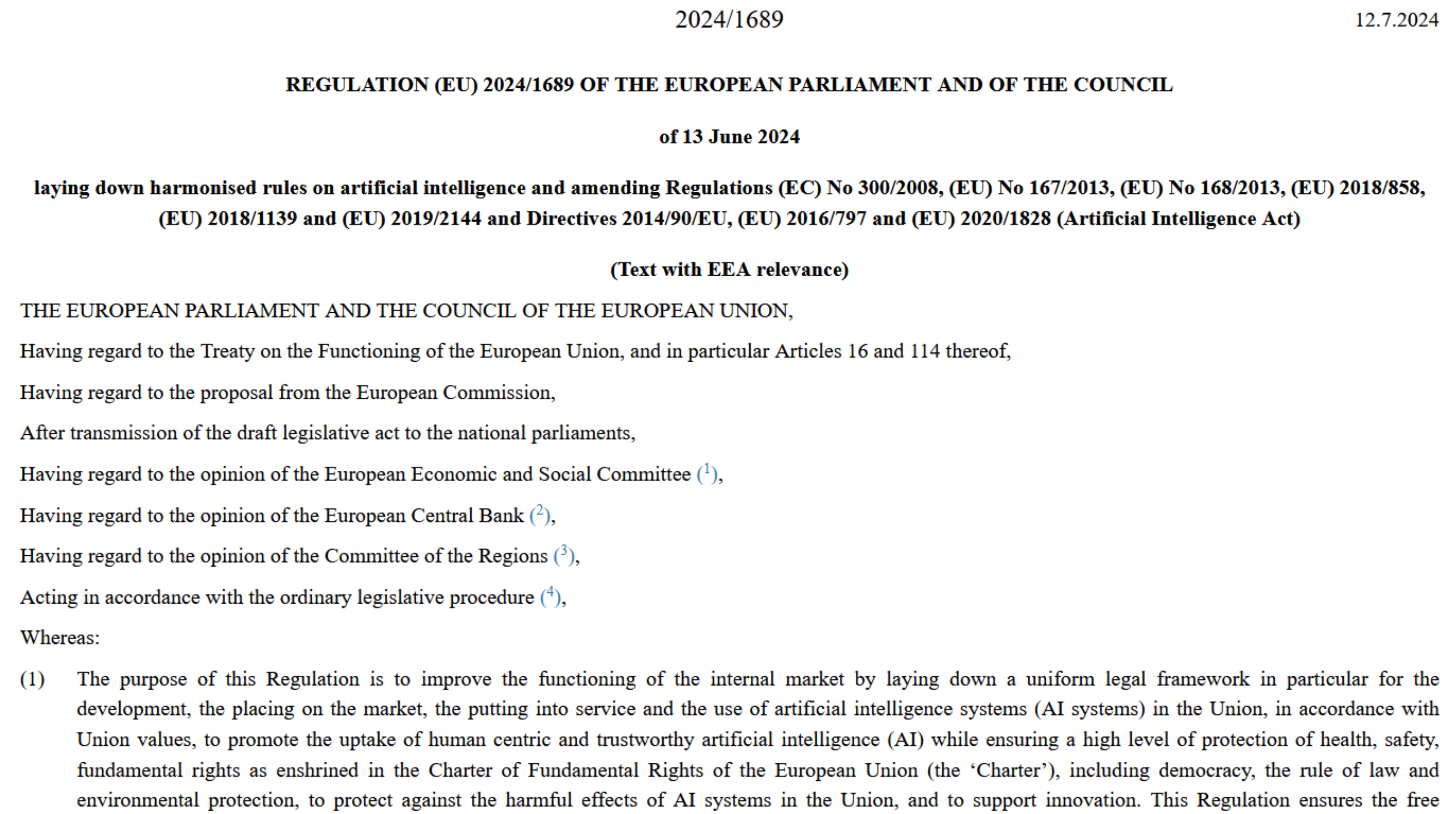
Limited risk - risks associated with a need for transparency around the use of AI

Minimal risk - AI-enabled pictures, spam filters

Source: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

AI Guardrails Principles

ℹ AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.



2024/1689 12.7.2024

REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 13 June 2024

laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Articles 16 and 114 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee (1),

Having regard to the opinion of the European Central Bank (2),

Having regard to the opinion of the Committee of the Regions (3),

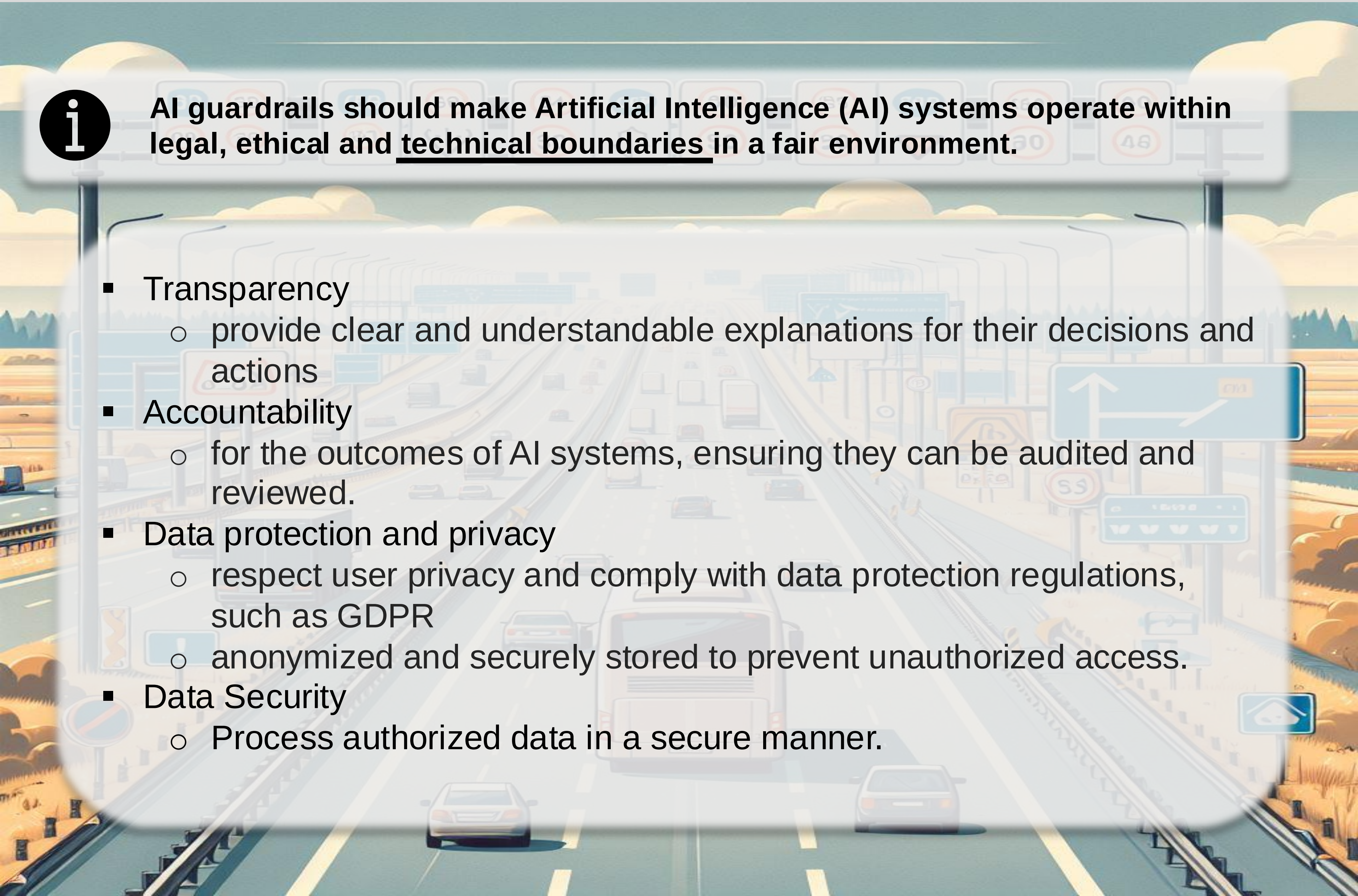
Acting in accordance with the ordinary legislative procedure (4),

Whereas:

(1) The purpose of this Regulation is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union, in accordance with Union values, to promote the uptake of human centric and trustworthy artificial intelligence (AI) while ensuring a high level of protection of health, safety, fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union (the 'Charter'), including democracy, the rule of law and environmental protection, to protect against the harmful effects of AI systems in the Union, and to support innovation. This Regulation ensures the free

Source: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL_202401689

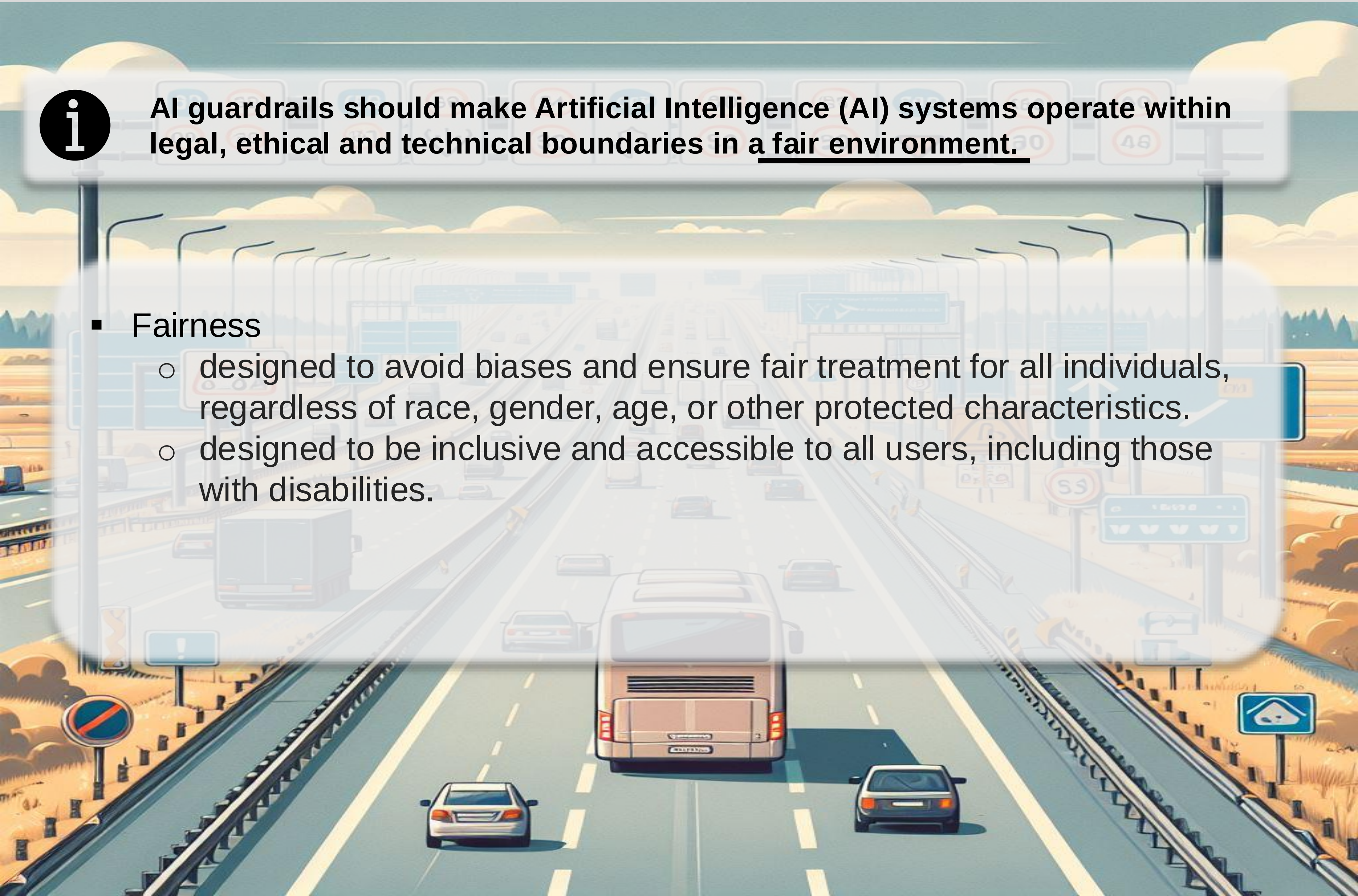
AI Guardrails Principles



AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.

- Transparency
  - provide clear and understandable explanations for their decisions and actions
- Accountability
  - for the outcomes of AI systems, ensuring they can be audited and reviewed.
- Data protection and privacy
  - respect user privacy and comply with data protection regulations, such as GDPR
  - anonymized and securely stored to prevent unauthorized access.
- Data Security
  - Process authorized data in a secure manner.

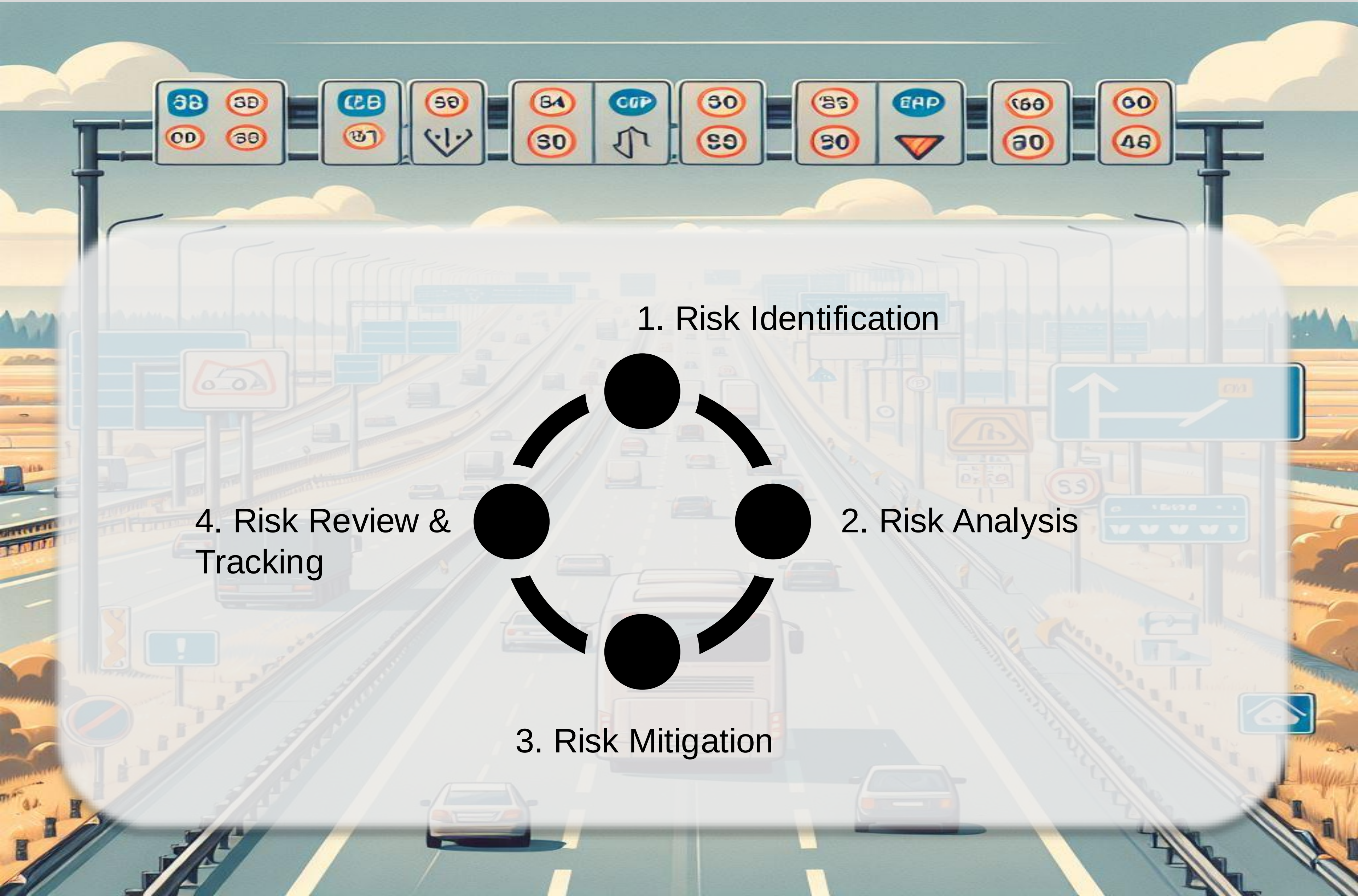
AI Guardrails Principles



AI guardrails should make Artificial Intelligence (AI) systems operate within legal, ethical and technical boundaries in a fair environment.

- Fairness
  - designed to avoid biases and ensure fair treatment for all individuals, regardless of race, gender, age, or other protected characteristics.
  - designed to be inclusive and accessible to all users, including those with disabilities.

Risk Management Lifecycle in Enterprises



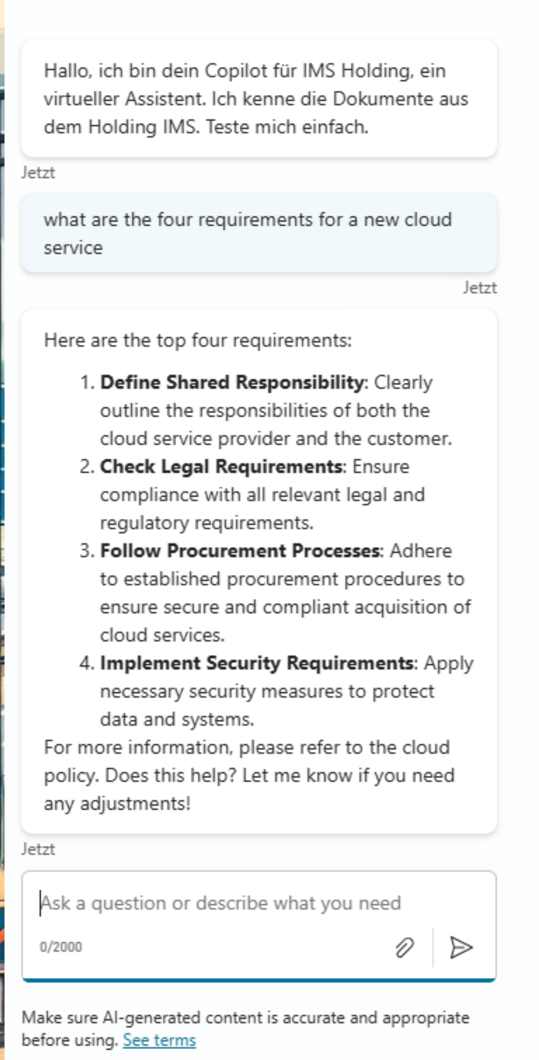
1. Risk Identification

2. Risk Analysis

3. Risk Mitigation

4. Risk Review & Tracking

1. Risk Identification



Hallo, ich bin dein Copilot für IMS Holding, ein virtueller Assistent. Ich kenne die Dokumente aus dem Holding IMS. Teste mich einfach.

Jetzt

what are the four requirements for a new cloud service

Jetzt

Here are the top four requirements:

1. **Define Shared Responsibility**: Clearly outline the responsibilities of both the cloud service provider and the customer.
2. **Check Legal Requirements**: Ensure compliance with all relevant legal and regulatory requirements.
3. **Follow Procurement Processes**: Adhere to established procurement procedures to ensure secure and compliant acquisition of cloud services.
4. **Implement Security Requirements**: Apply necessary security measures to protect data and systems.

For more information, please refer to the cloud policy. Does this help? Let me know if you need any adjustments!

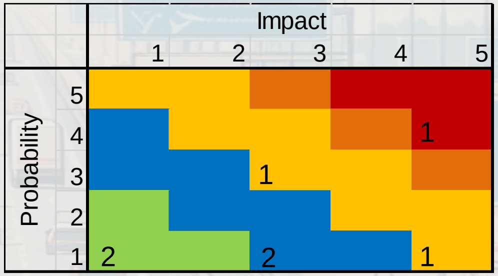
Jetzt

Make sure AI-generated content is accurate and appropriate before using. See terms

ISMS Chatbot

- Create targeted requirements
- Supervise the internal risk management process
- Lists contacts
- Compares documents such as a security concept to the internal ISMS

2. Risk Analysis



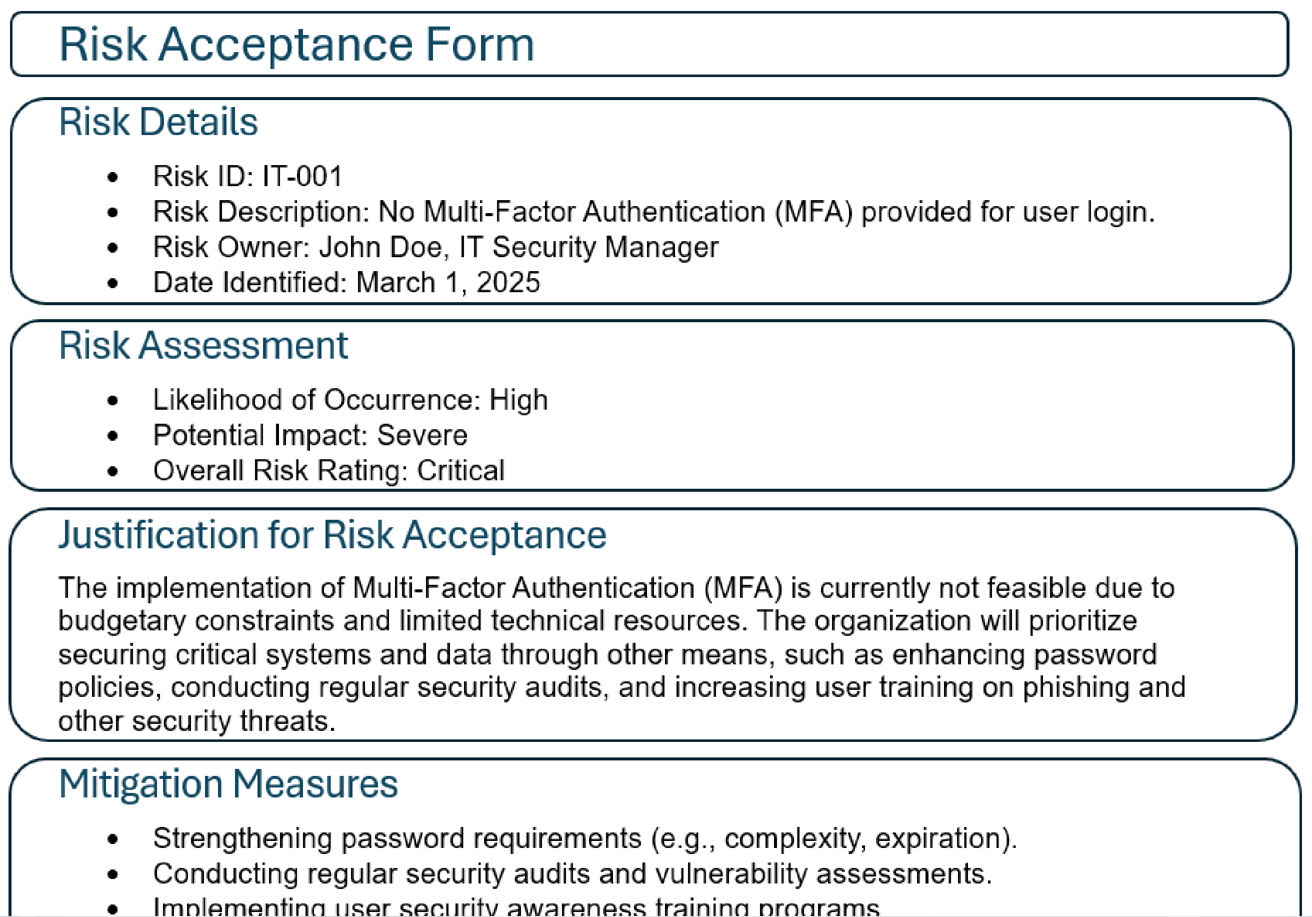
Probability \ Impact	1	2	3	4	5
5	🟨	🟨	🟧	🟥	🟥
4	🟦	🟨	🟨	🟧	🟥 1
3	🟦	🟦	🟨 1	🟨	🟧
2	🟩	🟦	🟦	🟨	🟨
1	🟩 2	🟩	🟦 2	🟦	🟨 1

Legend:

- 🔴 Very high risk
- 🟠 High risk
- 🟡 Medium risk
- 🔵 Low risk
- 🟢 Negligible risk

Analyse attack vectors, open vulnerabilities, and misconfigurations.

3. Risk Mitigation – Risk Acceptance



Risk Acceptance Form

Risk Details

- Risk ID: IT-001
- Risk Description: No Multi-Factor Authentication (MFA) provided for user login.
- Risk Owner: John Doe, IT Security Manager
- Date Identified: March 1, 2025

Risk Assessment

- Likelihood of Occurrence: High
- Potential Impact: Severe
- Overall Risk Rating: Critical

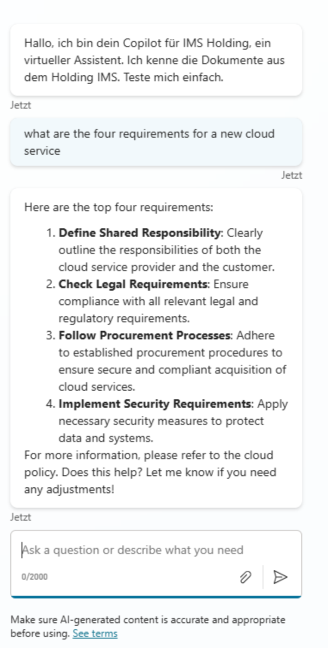
Justification for Risk Acceptance

The implementation of Multi-Factor Authentication (MFA) is currently not feasible due to budgetary constraints and limited technical resources. The organization will prioritize securing critical systems and data through other means, such as enhancing password policies, conducting regular security audits, and increasing user training on phishing and other security threats.

Mitigation Measures

- Strengthening password requirements (e.g., complexity, expiration).
- Conducting regular security audits and vulnerability assessments.
- Implementing user security awareness training programs

4. Risk Review & Tracking



Hallo, ich bin dein Copilot für IMS Holding, ein virtueller Assistent. Ich kenne die Dokumente aus dem Holding IMS. Teste mich einfach.

Jetzt

what are the four requirements for a new cloud service

Jetzt

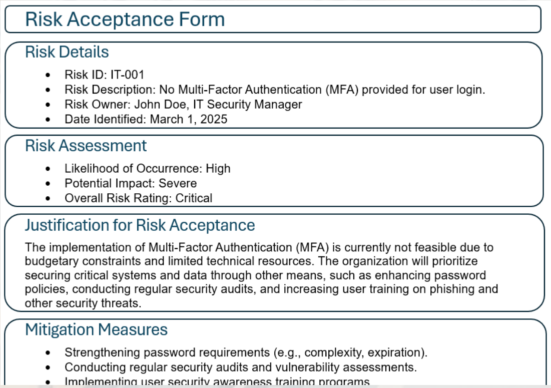
Here are the top four requirements:

1. **Define Shared Responsibility**: Clearly outline the responsibilities of both the cloud service provider and the customer.
2. **Check Legal Requirements**: Ensure compliance with all relevant legal and regulatory requirements.
3. **Follow Procurement Processes**: Adhere to established procurement procedures to ensure secure and compliant acquisition of cloud services.
4. **Implement Security Requirements**: Apply necessary security measures to protect data and systems.

For more information, please refer to the cloud policy. Does this help? Let me know if you need any adjustments!

Jetzt

Make sure AI-generated content is accurate and appropriate before using. See terms



Risk Acceptance Form

Risk Details

- Risk ID: IT-001
- Risk Description: No Multi-Factor Authentication (MFA) provided for user login.
- Risk Owner: John Doe, IT Security Manager
- Date Identified: March 1, 2025

Risk Assessment

- Likelihood of Occurrence: High
- Potential Impact: Severe
- Overall Risk Rating: Critical

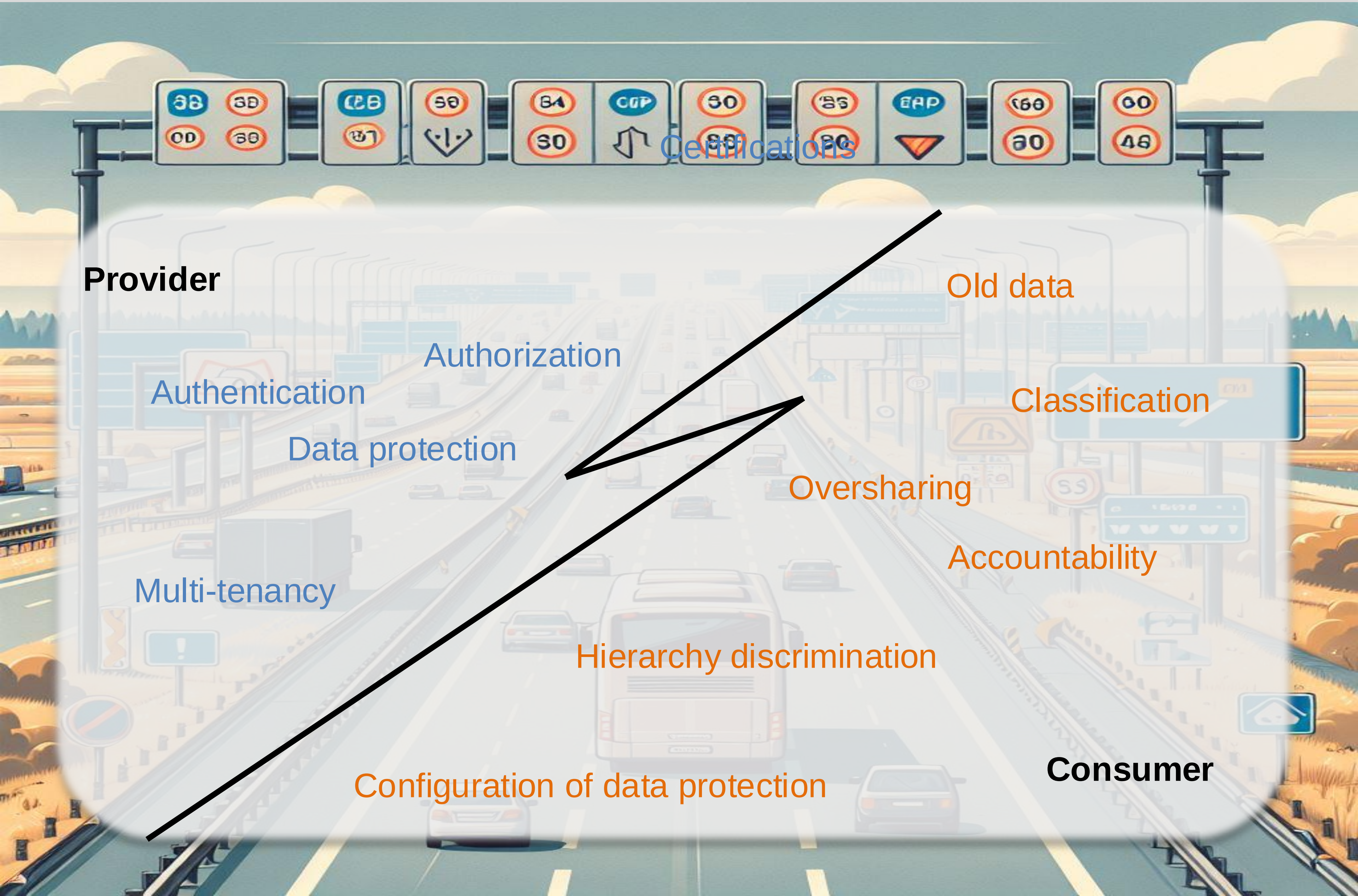
Justification for Risk Acceptance

The implementation of Multi-Factor Authentication (MFA) is currently not feasible due to budgetary constraints and limited technical resources. The organization will prioritize securing critical systems and data through other means, such as enhancing password policies, conducting regular security audits, and increasing user training on phishing and other security threats.

Mitigation Measures

- Strengthening password requirements (e.g., complexity, expiration).
- Conducting regular security audits and vulnerability assessments.
- Implementing user security awareness training programs.

AI Guardrails Principles



Certifications

Provider

Authorization

Authentication

Data protection

Multi-tenancy

Old data

Classification

Oversharing

Accountability

Hierarchy discrimination

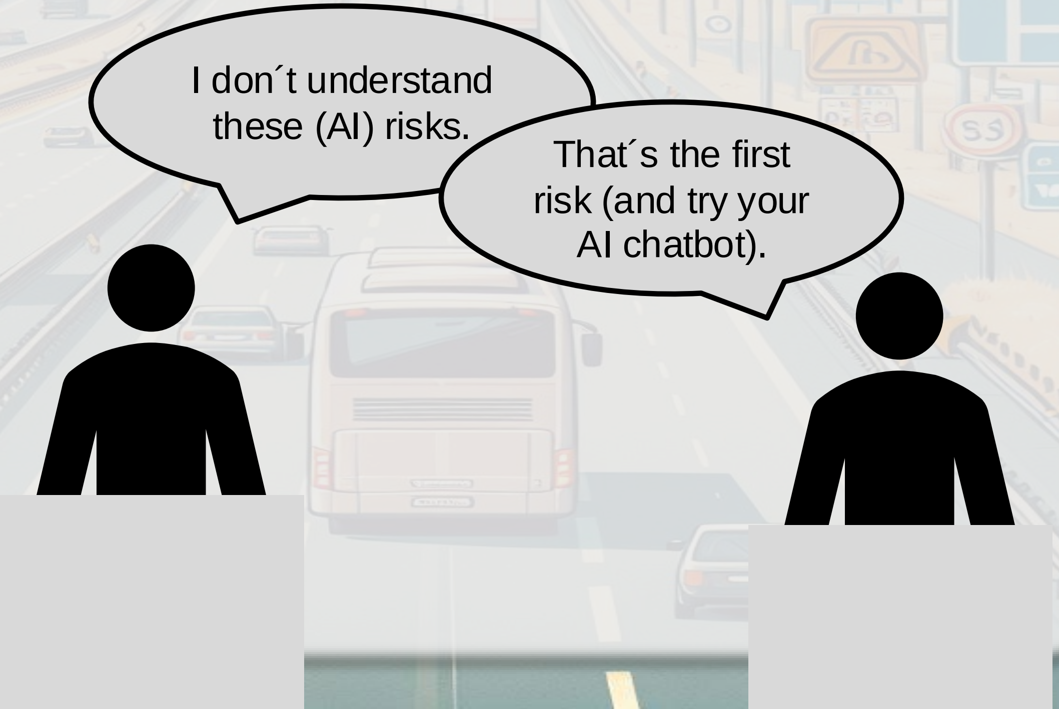
Configuration of data protection

Consumer

Thank you

Marian Kühnel
Information Security Architect
Business-IT & Cloud

ÖBB Group
+43 664 886 77170
Marian.kuehnel@oebb.at


I don´t understand these (AI) risks.
That´s the first risk (and try your AI chatbot).