

AI Ethics – A Primer to Guide your Organization on the Opportunities and Risks

Marcus Schueler, Head of Responsible AI, MHP – A Porsche
Company

I/MHP

A PORSCHE COMPANY

Barcelona

2024

SAPinsider

In This Session

You will:

- learn about risks and ethical issues associated with AI
- gain insights into the status of regulations, in particular the EU AI Act
- explore how and when companies should prepare for the AI Act
- engage in discussions about what the current development in the AI environment mean for you as an individual



What We'll Cover

- What is AI Ethics and why does it matter
- The EU AI Act – What it is and why it's a good idea
- How to best implement the AI Act
- The influence of AI on the personal sphere
- Wrap-Up

A large, stylized, glowing blue 'AI' text is positioned on the right side of the slide. It is set against a dark blue background with a complex network of glowing lines and nodes, resembling a circuit board or a neural network. The 'AI' text is the central focus of this graphic.

What is AI Ethics and why does it matter

- AI is like a powerful medication that bears huge benefits but also carries significant side-effects
- AI ethics focuses on this trade-off, aiming to enhance the positive effects of AI while addressing its associated risks
- AI ethics has a human-centric approach and considers humans always as an end and never solely as a means.
- Responsible AI is more than just mitigating risks

The “Classics” of AI Ethics

OECD Valued-based principles¹

- **Inclusive growth, sustainable development and well-being**

Trustworthy AI shall contribute to overall growth and prosperity for individuals, society, and planet

- **Human-centred values and fairness**

AI should respect law, human rights, democratic values, diversity and ensure a fair and just society

- **Transparency and explainability**

Ensure people understand when they are engaging with an AI System and can challenge outcomes

- **Robustness, security and safety**

AI must function in a robust, secure and safe way throughout their lifetimes be continually assessed

1. <https://oecd.ai/en/ai-principles>

At their core, the requirements are repeated in the respective directives

SAP's Guiding Principles for AI¹

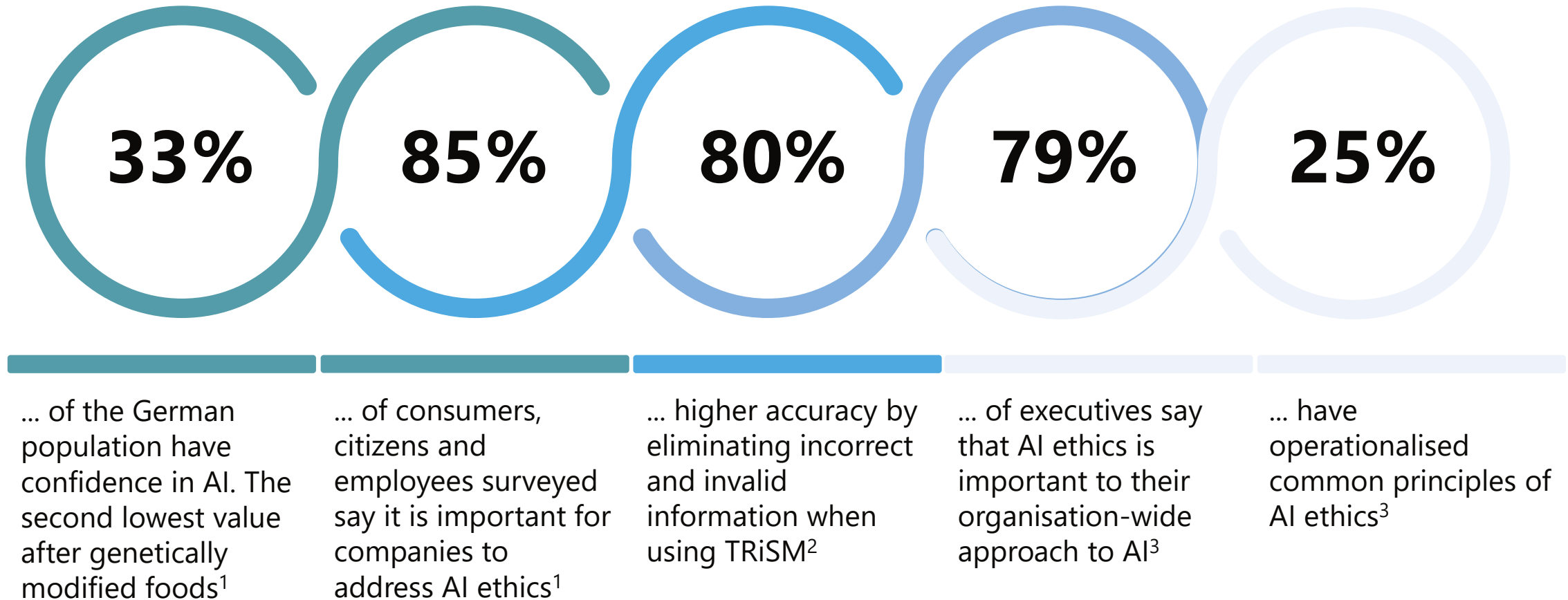
- ✓ We are driven by our values
- ✓ We design for people
- ✓ We enable business beyond bias
- ✓ We strive for transparency and integrity in all that we do
- ✓ We uphold quality and safety standards
- ✓ We place data protection and privacy at our core
- ✓ We engage with the wider societal challenges of AI

AI Ethics Guidelines for Trustworthy AI (HLEG EU)²

- ✓ Human agency and oversight
- ✓ Technical robustness and safety
- ✓ Privacy and data governance
- ✓ Transparency
- ✓ Diversity, non-discrimination, and fairness
- ✓ Environmental and societal well-being
- ✓ Accountability

1. <https://www.sap.com/products/artificial-intelligence/ai-ethics.html> 2. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> ; HLEG High Level Expert Group

Some of the reasons why AI Ethics is important



¹ Edelman Trust Barometer 2024 | ³Gartner Top Strategic Technology Trends /Trust, Risk and Security Management |

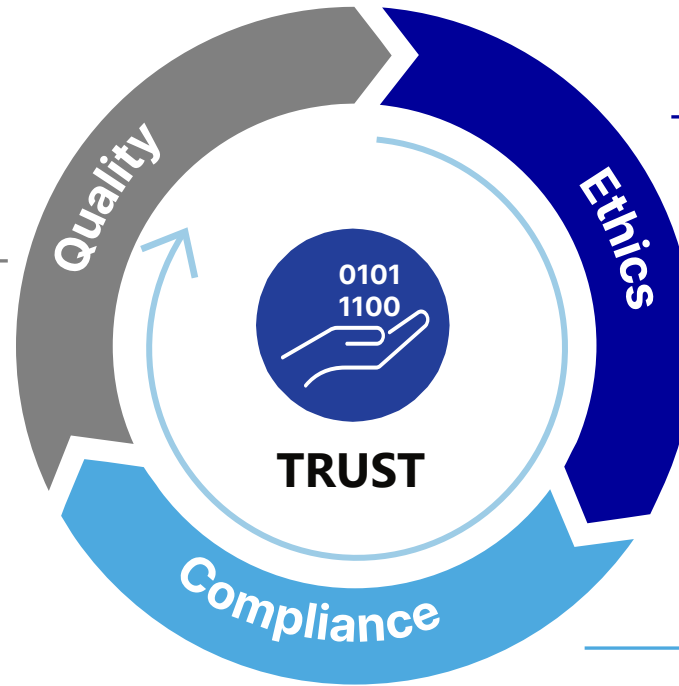
⁴IBM - The CEO's Guide to Generative AI

Responsible AI is more than just AI Ethics and becomes the Headline

1

Standards:

Responsible AI is increasingly being standardized. VDE and DIN have developed proposals. The new ISO/IEC 42001:2023 will prevail



2

Best Practice:

In the field of AI ethics, there is a broadly shared international understanding of the core elements

3

International regulations:

The main regulations will come from the EU, USA and China and will provide more legal certainty. The G7, for example, is also looking for overarching solutions

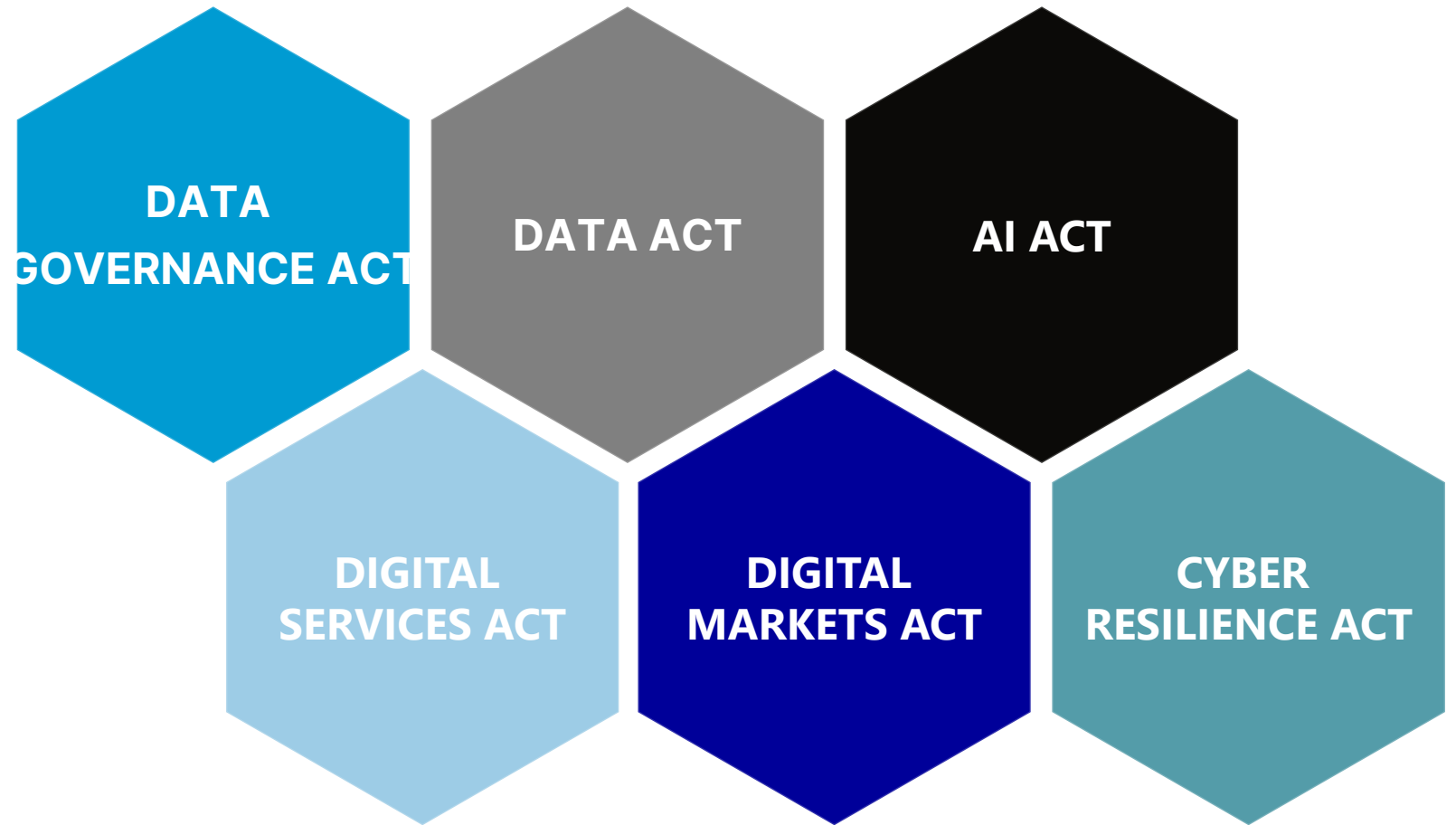
The EU AI Act – What it is and why it's a good idea

- **Context of the EU AI Act**
- **Content and Structure of the EU AI Act**
- **Benefits and Impact of the EU AI Act**

Regulations as part of the EU DIGITAL STRATEGY

“Shaping Europe's **digital future**” formulates three main objectives:

- Technology at the **service of people**
- A **fair** and **competitive** economy
- An **open, democratic and sustainable** society



What is the purpose of the AI Act?



*“The **purpose** of this Regulation is to improve the **functioning of the internal market** and promoting the uptake of **human centric and trustworthy artificial intelligence**, while ensuring a **high level of protection of health, safety, fundamental rights** enshrined in the Charter, including **democracy, rule of law and environmental protection** against **harmful effects** of artificial intelligence systems in the Union and **supporting innovation**”¹*

1) Artificial Intelligence Act Article 1: Subject Matter (1)

What is an AI System?

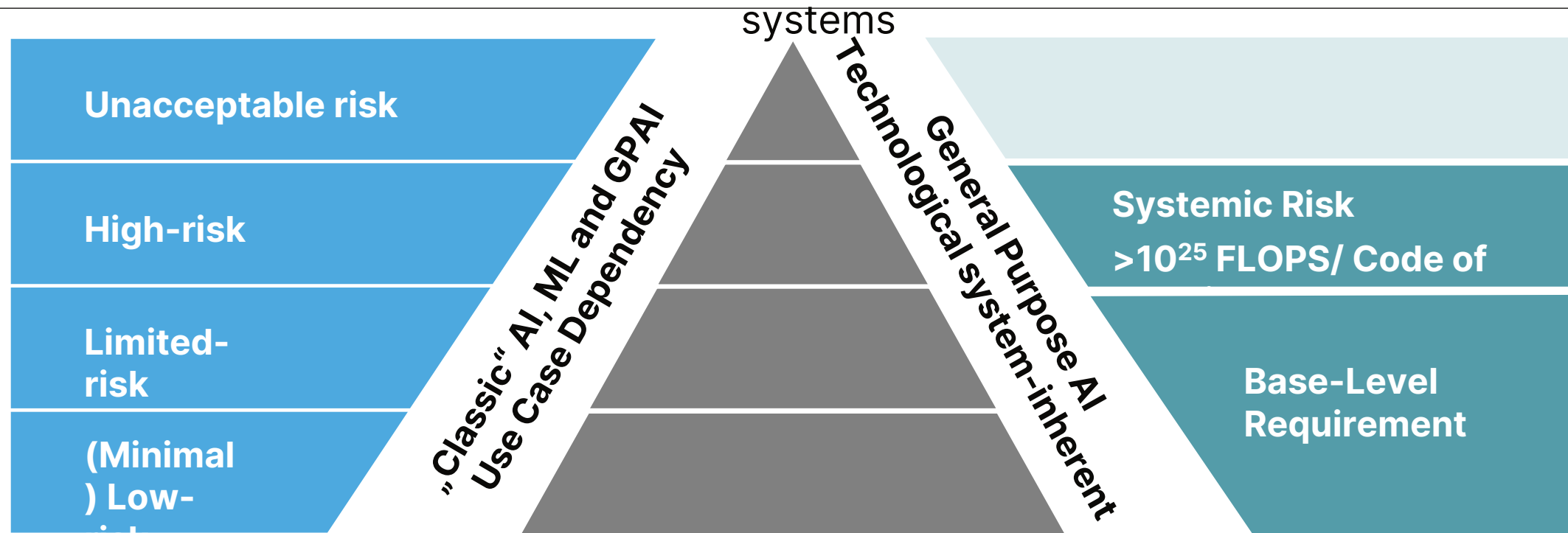


*“**AI system**’ means a **machine-based system** designed to operate with **varying levels of autonomy** and that may **exhibit adaptiveness after deployment** and that, for **explicit or implicit objectives**, infers, from the **input it receives**, how to generate outputs such as **predictions, content, recommendations, or decisions** that can influence physical or virtual environments”¹*

1) Artificial Intelligence Act Article 3: Definitions (1)

What are the basic principles of the AI ACT

The AI Act takes a **Use-Case risk-based approach** to the regulation of AI



An **exception** applies to **GPAI**, which is **always subject** to similar requirements as a **high-risk system** if certain **system-critical criteria** are met

Criteria for the Use Case Risk Classification for Prohibited Systems

Use Case Verification



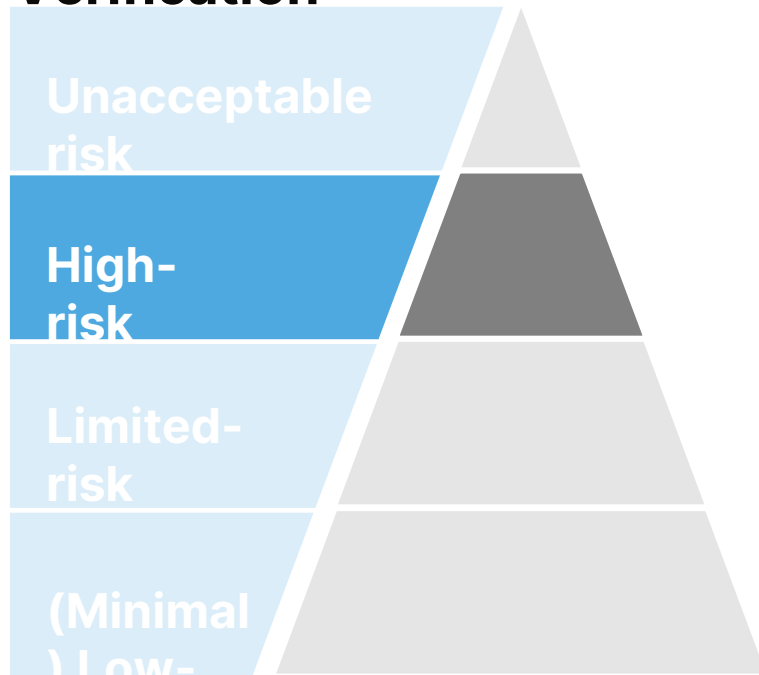
AI systems with **unacceptable risks are forbidden**. This includes, for example:

- **Real-time biometric systems** that can be used for any type of surveillance (exception: crime prevention)
- Algorithms for the **social evaluation of people** (social scoring)
- **Manipulative systems** that exploit the vulnerabilities of certain individuals to influence their behavior in a way that can lead to physical or psychological harm.

Penalties: up to 35 million EUR or up to **7% of the worldwide total annual turnover** in the previous fiscal year whichever amount is higher

Criteria for the Use Case Risk Classification for High-risk Systems

Use Case Verification



High-risk AI systems include, for example:

- AI systems used as **a safety component of a product**
- **Biometric identification** and **categorization** of natural persons
- Administration and operation of **critical infrastructures**
- **Education** and **vocational** training (access and assessments)
- **Employment** and **employee management** (recruiting processes)

Penalties: up to 15 million EUR or up to 3% of the worldwide total annual sales in the previous fiscal year, depending on which amount is higher

- Administration of **justice and democratic** processes

Criteria for the Use Case Risk Classification for Limited-risk Systems

Use Case Verification



Transparency obligations apply for some **limited-risk** systems:

- AI that **interacts directly with humans** falls into the "limited risk" category
- This includes, for example, **emotion recognition** systems and AI-generated or modified content that **resembles real people, objects, places or events** and could be **mistaken for real** ("deepfakes")

AI Ethics Requirement fulfillment via Code of Conduct for every use case

For every AI Use



Case

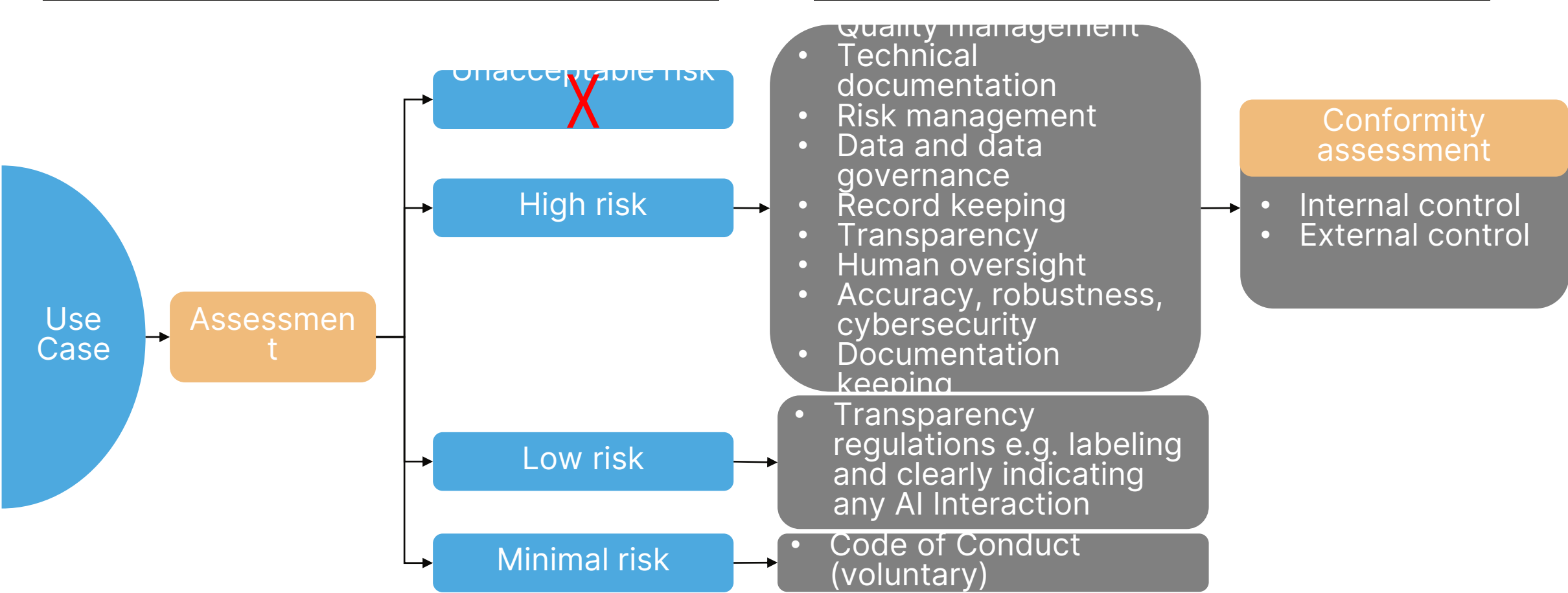


- The EU encourage the voluntary drawing up of codes of conduct, including related governance mechanisms
- applicable elements of EU **ethic guidelines for trustworthy AI**
 - impact of AI systems on **environmental sustainability**
 - promoting **AI literacy**
 - **inclusive** and **diverse design** of AI systems
 - preventing **negative impact** on **vulnerable persons** or **groups of persons**
-

Risk Evaluation logic and required measures

Risk Evaluation

Necessary measures

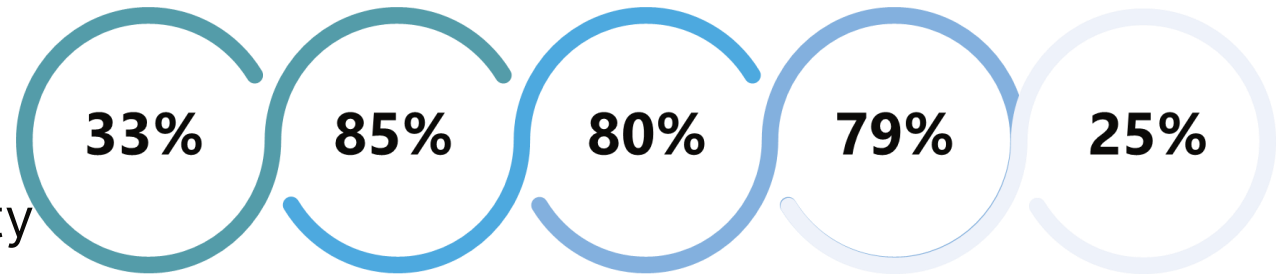


Sound like a lot of work.

Is it worth it?

Apart from the obvious, the purpose of the AI Act to protect and support people, what other reasons are there?

- Isn't that already enough?
- Legal certainty and planning reliability
- More security for investments
- Building trust of employees, consumers and society
- Competitive advantages and fostering of innovations
- Significantly higher quality
- Pioneer for a global standard (> 1000 AI policy initiatives from 69 countries, territories

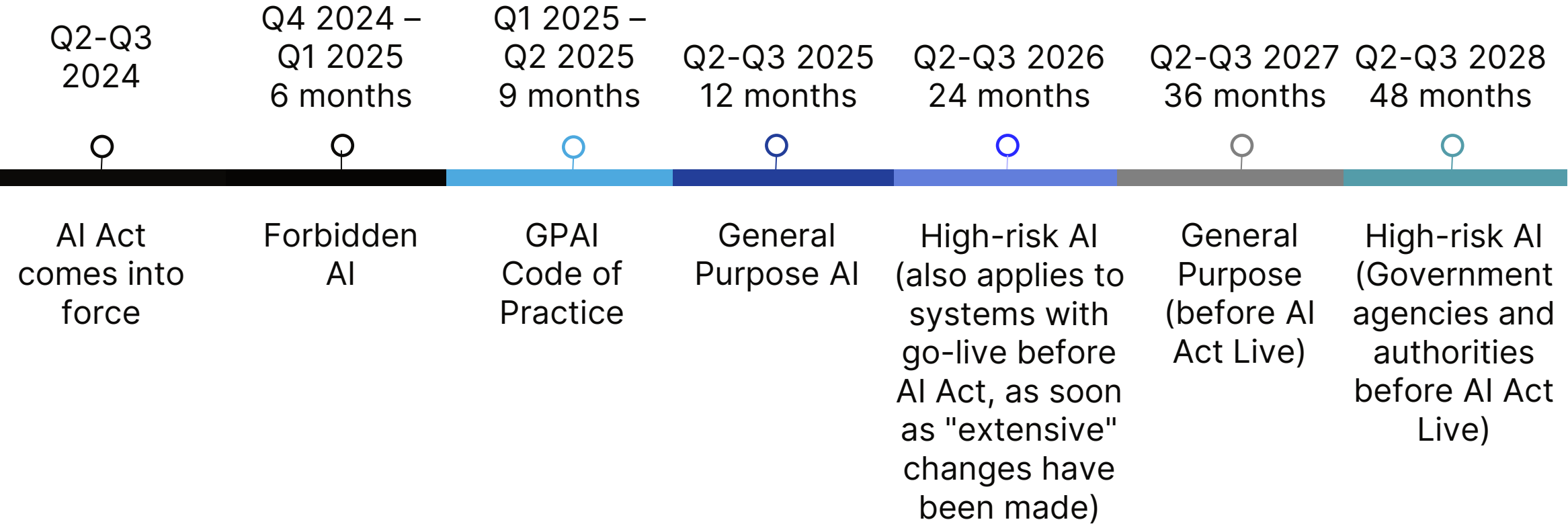


1) <https://cdm.cai/en/dashboards/overview>

How to best implement the AI Act

- **What does the timeline look like**
- **Which components should be brought together?**
- **How should the project be set up?**

EU AI-Act Timeline of transition Periods



Which components should be brought together?



Requirements

- Conformity for High Risk
- Transparency for Limited Risk
- Code of Conduct for Low Risk
- Requirements for GPAI



Governance

- Quality management
- Risk management
- Responsibilities and guidelines
- Operations and support



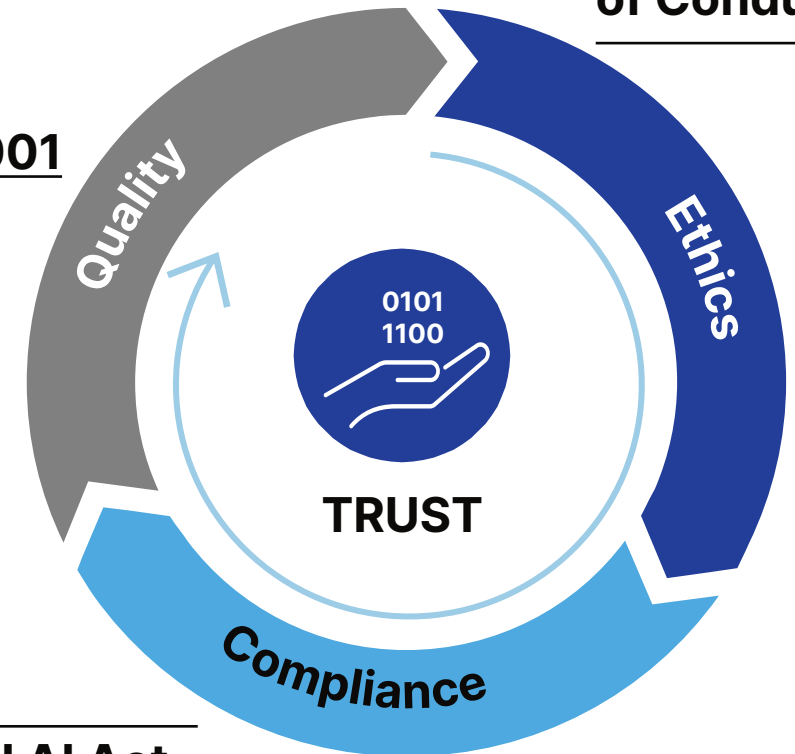
Systems

- Internal governance systems
- External governance and assessment systems

ISO 42001

Code
of Conduct

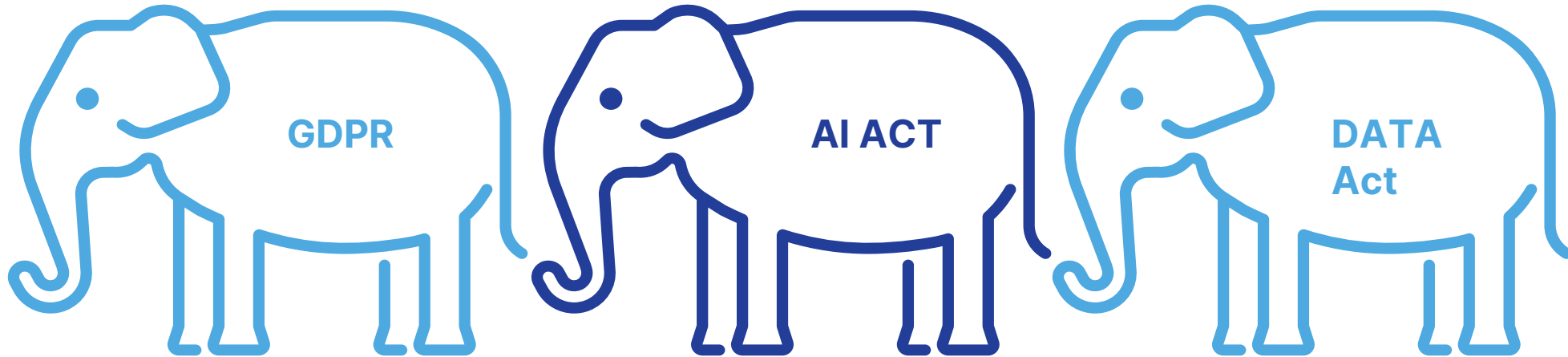
EU AI Act



How should the project be set up?

Cut it into manageable pieces

PROGRAM ALIGNMENT AND INTEGRATION WITH FURTHER
REGULATIONS



ASSESSMENT

CODE OF
CONDUCT

ISO 42001

(DELTA) AI ACT

SYSTEM ARCHITECTURE, SELECTION AND IMPLEMENTATION

CHANGE MANAGEMENT

The influence of AI on the personal sphere

- The spectrum of ethical issues
- What can you Do about it?



The spectrum of ethical issues

Discrimination

Discrimination is the most common problem and can affect almost everyone. Some Examples:

- **LLMs** produce **gender bias**, as well as **homophobia** and **racial stereotyping**¹
- Lensa app creates cartoonishly **pornified female avatars** while male avatars got to be astronauts, explorers, and inventors²
- AI **skin cancer** diagnoses risk being **less accurate** for **dark skin** application processes³
- Amazon **recruiting tool** that showed **bias against women**⁴

1) <https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>

2) <https://www.technologyinsider.com/2023/12/1064751/the-viral-ai-avatar-app-lensa-undressed-me-without-my-consent/>

3) <https://www.theguardian.com/society/2021/nov/09/ai-skin-cancer-diagnoses-risk-being-less-accurate-for-dark-skin-study>

4) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/>

Beyond Discrimination

Like AI, ethical issues will permeate every part of our lives Some Examples:

- **Fake audio** of one of the top candidates for **election** in Slovakia⁵
- “**Deepfaked Podcast** Has Joe Rogan Interview Undead **Steve Jobs**”⁶
- **TikTok**...well, where to start⁷
- **Autonomous driving** – the real life trolley problem ⁸
- Using AI in **emergency room triage**⁹

5) <https://edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs/index.html>

6) <https://futurism.com/the-byte/deepfaked-podcast-joe-rogan-interview-steve-jobs>

7) <https://en.wikipedia.org/wiki/TikTok>

8) <https://core.ac.uk/download/pdf/231922494.pdf>

9) <https://www.atlas-digitale-gesundheitswirtschaft.de/blog/2024/02/25/using-ai-in-emergency-room-triage/#>

What can you Do about it?

Protect yourself, your children and get involved

- **Keep up-to-date** with AI technology and how it impacts your life
- **Educate yourself** and your Family and **encourage critical thinking**
- Be **cautious** with **personal data** and **teach your children** about the importance
- Contribute to or participate in ethical **AI research initiatives or NGOs**
- Become familiar with the concept of **Responsible AI**

However, AI is a **great opportunity** and has **great potential** to make all our **lives better** in many ways

Wrap Up

~~You will:~~ I hope you have

- learned about risks and ethical issues associated with AI
- gained insights into the status of regulations, in particular the EU AI Act.
- explored how and when companies should prepare for the AI Act.
- engaged in discussions about what the current development in the AI environment mean for you as an individual

Where to Find More Information

<https://oecd.ai/en/>

OECD Landig page AI

<https://www.sap.com/products/artificial-intelligence/ai-ethics.html>

SAP Responsible Overview

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>

Latest version of EU AI Act

<https://artificialintelligenceact.eu>

The EU Artificial Intelligence Act / Up-to-date developments and analyses of the EU AI Act

<https://algorithmwatch.org/en/>

AlgorithmWatch is a human rights organization based in Berlin and Zurich. We fight for a world where algorithms and Artificial Intelligence (AI) do not weaken justice, democracy, and sustainability but strengthen them.

<https://www.turing.ac.uk/research/research-programmes/artificial-intelligence>

the UK's national institute for data science and artificial intelligence (excellent research)

<https://www.responsible.ai/>

The Responsible AI Institute is a global non-profit dedicated to equipping organizations and AI professionals with tools and knowledge to create, procure and deploy AI systems that are safe and trustworthy

or drop me a note!

Key Points to Take Home

- **Responsible AI** will be a **core topic** to bring **opportunities and risks** into the **right balance**
- **Familiarize** yourself or your company with the **AI Act** in the **short term** if you have not already done so
- The **timeline** for implementation **is tight**. Think about the GDPR
- **Proceed in steps** - Code of Conduct, ISO 42001...) that you parallelize sensibly
- AI will affect **every area** of our **life** and so will its side effects
- Actively **engage** with the topic of AI in order to **take advantage of its opportunities** and **mitigate its issues**

Thank you! Any Questions?

Marcus Schueler

<https://www.linkedin.com/in/marcusschueler/>



Please remember to
complete your session
evaluation.

SAPinsider



SAPinsider.org

PO Box 982Hampstead, NH 03841
Copyright © 2024 Wellesley Information Services.
All rights reserved.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies. Wellesley Information Services is neither owned nor controlled by SAP SE.

**SAPinsider
comprises the
largest and fastest
growing SAP
membership group
with more than
800,000 members
worldwide.**
