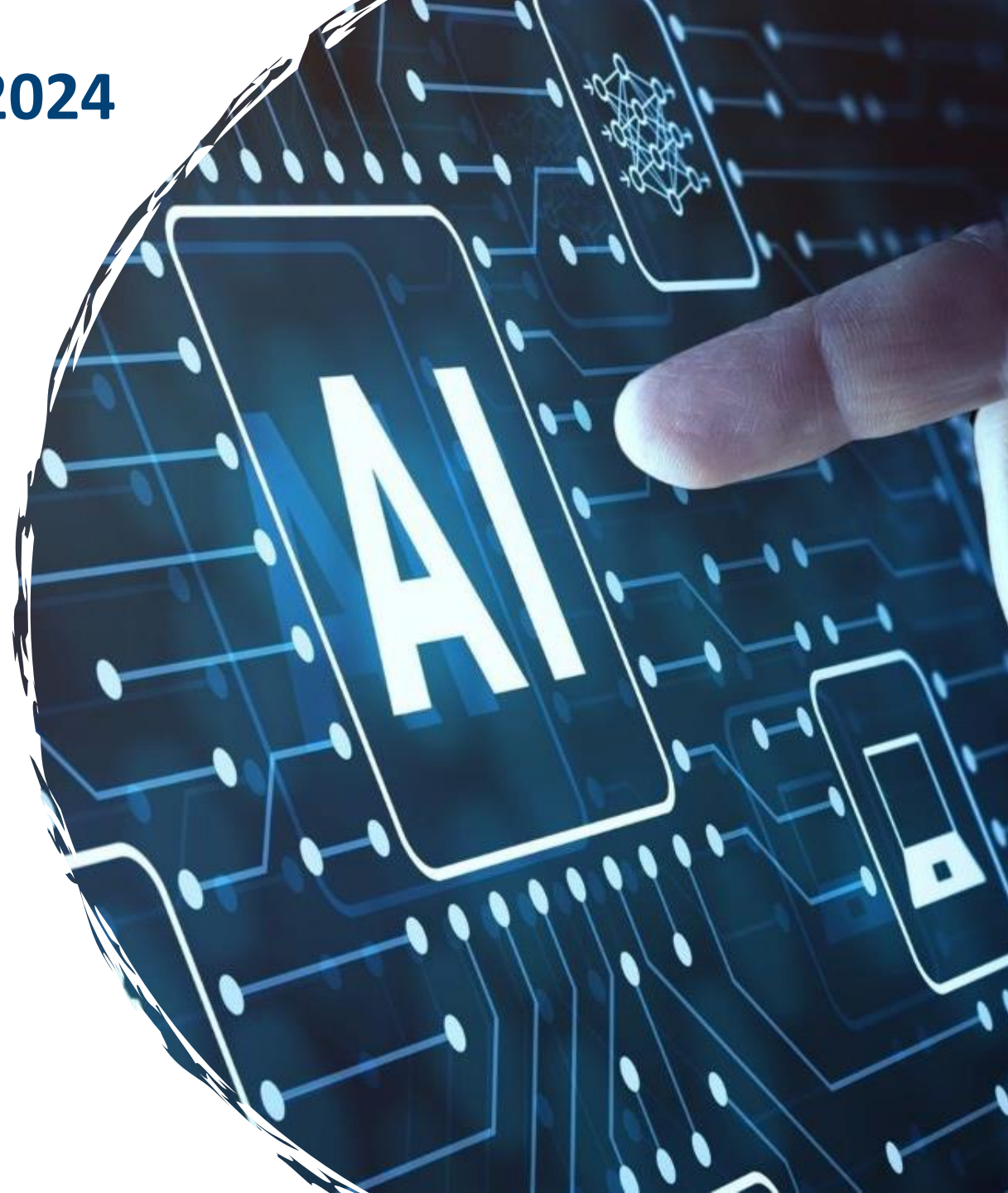


Future Conformity Assessments for AI Safety and Security

Presented by: Dr. Xavier Valero

Director AI & Advanced Analytics










About DEKRA

Safety and security experts in 6 regions and 60 countries



SERVICES

-  Vehicle Inspection
-  Claims & Expertise
-  Digital & Product Solutions
-  Industrial Inspection
-  Advisory & Training Services
-  Audit
-  Temp Work



Future Conformity Assessments for AI Safety and Security

Agenda



EU AI Act Conformity Assessment Obligations

3



Conformity Assessment Options for AI

9



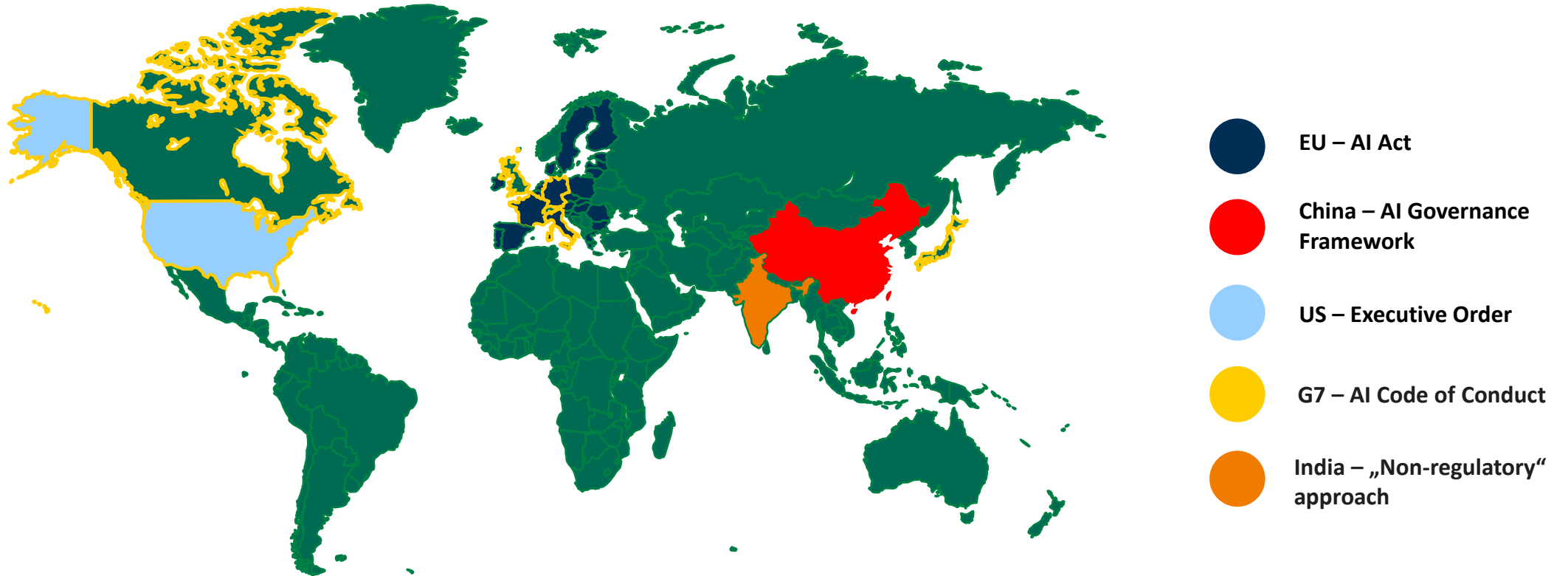
Conclusion

18



Regulatory Approaches to AI

Overview of representative regulatory approaches *



- Non-exhaustive list, other regulations are being drafted in other countries like Singapore, Australia, Canada, etc.

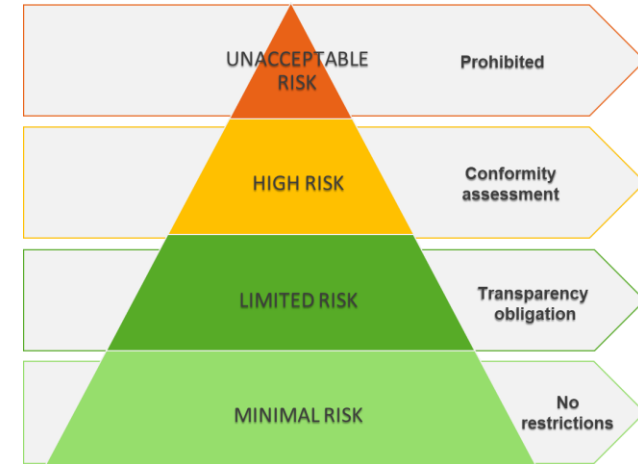
EU AI Act

Risk-based regulation

- ▶ Conformity Assessment procedures are mandatory for high-risk AI providers

OBLIGATIONS TO PROVIDERS OF HIGH – RISK AI SYSTEMS

- ▶ Compliance High-risk AI systems requirements
- ▶ Implement AI Quality Management System
- ▶ Draw up technical documentation and keep logs
- ▶ **Undergo Conformity Assessment procedure and take required corrective actions.**
- ▶ Register the AI system in the EU database



High Risk AI System requirements:

- Risk Management System
- Data governance
- Transparency to users
- Human oversight
- Accuracy, robustness and cybersecurity
- Technical documentation
- Record-keeping

EU AIA Conformity Assessment Obligations

► Conformity Assessment obligations depend on the type of use case and applicative legislation.

Annex III - High-risk AI systems:

1. Biometric person identification
2. Critical infrastructure management and operation
3. Education and vocational training...
4. Employment, workers management and access to self-employment
5. Access to essential private and public services and benefits
6. Law enforcement
7. Migration and border control management...
8. Administration of justice and democratic processes

Biometric person identification:



Annex III,
point 1

No harmonized
standards applied

Notified body CA

Harmonized standards
applied

Reminder:



Annex III,
point 2-8

Internal control CA

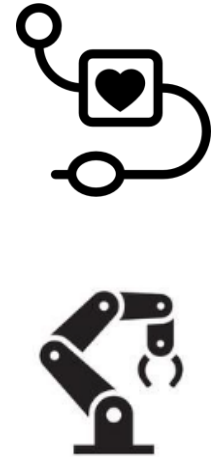
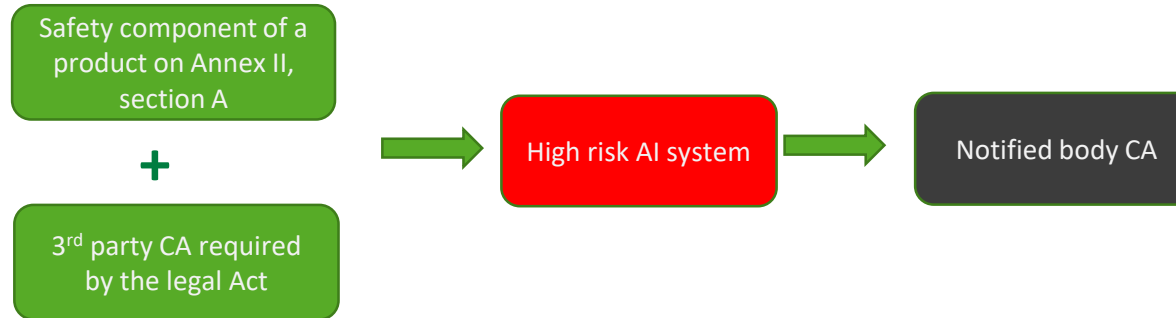
EU AIA Conformity Assessment Obligations

► Conformity Assessment obligations depend on the type of use case and applicative legislation.

Annex II – Section A – New Legislative Framework

1. Regulation on Machinery Products
2. Toy Safety Directive
3. Recreational Craft Directive
4. Lift Directive
5. ATEX Certification Directive
6. Radio Equipment Directive
7. Pressure Equipment Directive
8. Cableways Directive
9. PPE Regulation
10. Gas Appliances Regulation
11. Medical Devices Regulation
12. In Vitro Diagnostics Regulation

Annex II, section A- High-Risk AI systems



Annex II – Section B – Other legislations

1. Civil Aviation Security
2. Type Approval Regulation
3. Marine Equipment Directive
4. Rail System Interoperability Directive

Annex II, section B



- EU AI Act not directly applicable
- EU AI Act requirements shall be taken in consideration in applicable regulations

EU AIA Conformity Assessment Obligations

- ▶ Conformity Assessment obligations for providers of General Purpose AI.

Obligations for providers of GPAI models

Transparency requirements:

- Drawing up technical documentation
- Complying with EU copyright law
- Provide information about the data used for training



Internal control CA

Obligations for providers of GPAI models with systemic risk*

- Perform model evaluation in accordance with standardized protocols and tools, including **adversarial testing**
- **Assess** and mitigate possible systemic risks
- **Report** to authorities information about serious incidents and possible corrective measures
- Ensure an adequate level of cybersecurity protection



Internal control CA

Notified body CA

* GPAI model with systemic risk':
compute used for its training is greater than 10²⁵ FLOPs (floating point operations)

Future Conformity Assessments for AI Safety and Security

Agenda



EU AI Act Conformity Assessment Obligations

3



Conformity Assessment Options for AI

9



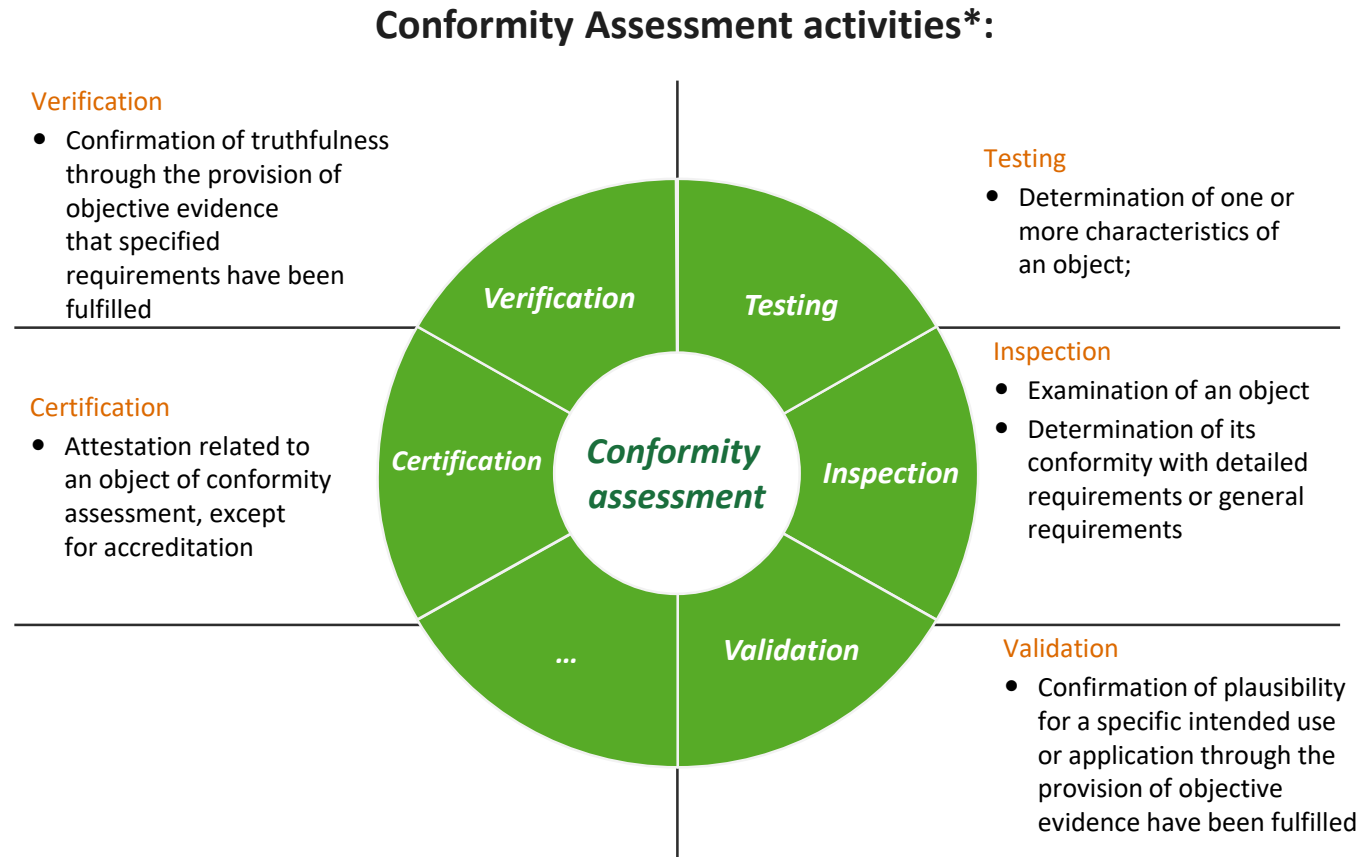
Conclusion

18



What is a Conformity Assessment?

- ▶ Demonstration that specified requirements relating to a product, process, system, person or body are fulfilled.
- ▶ Can be conducted by internal controls or third-party assessments.

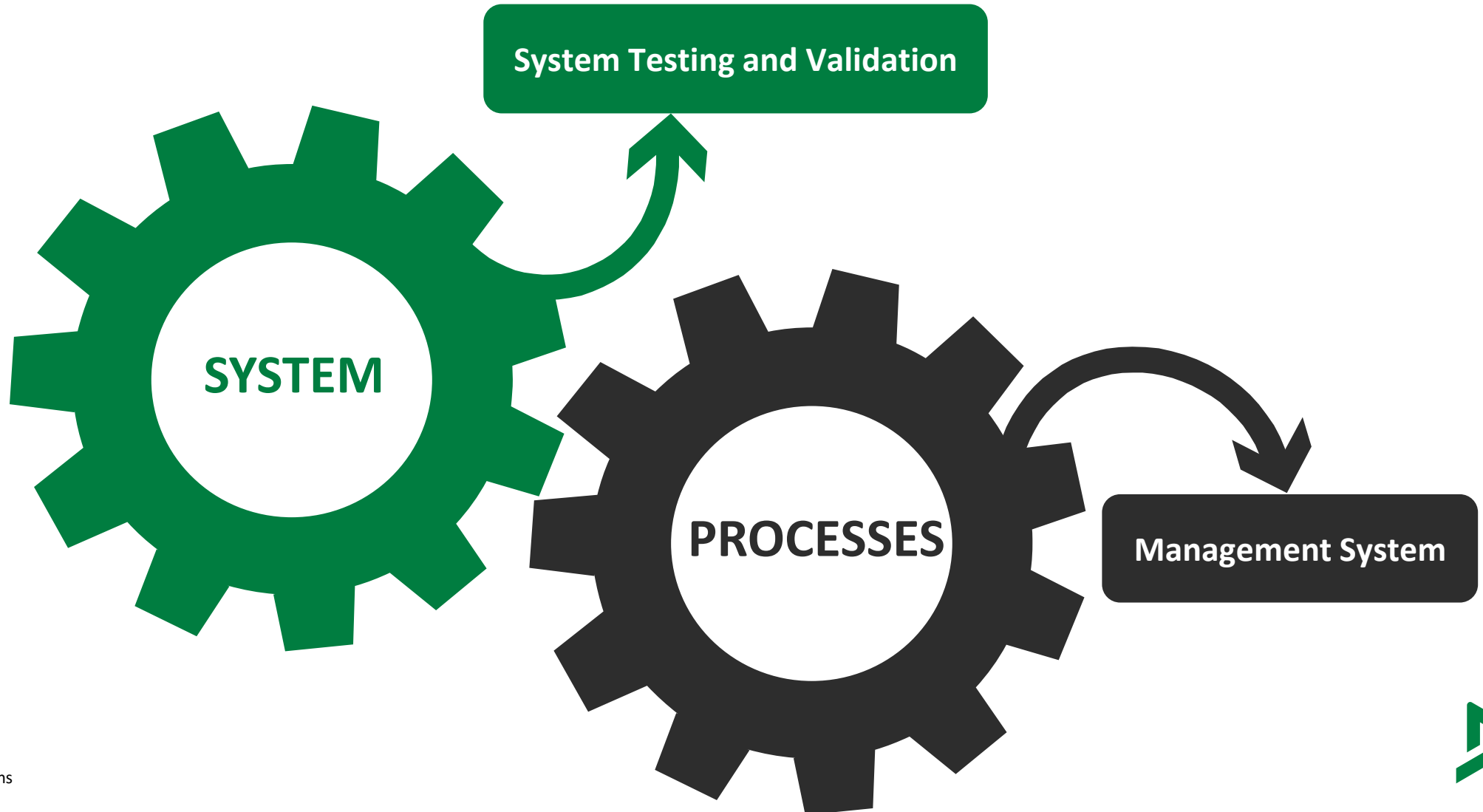


Source: ISO 17001

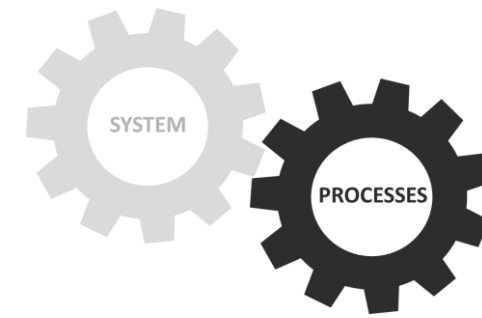


Building Safe and Secure AI

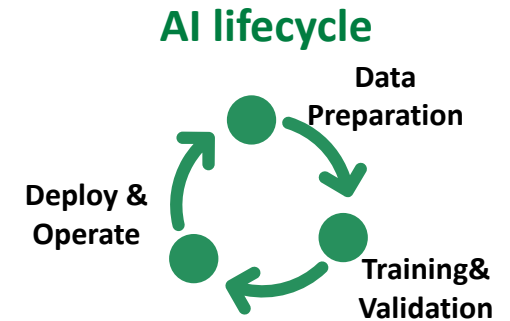
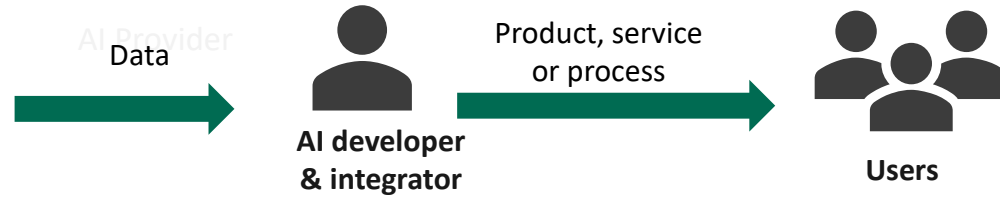
- ▶ Ensuring trustworthy AI and compliance with EU AI Act may be summarized on two aspects.



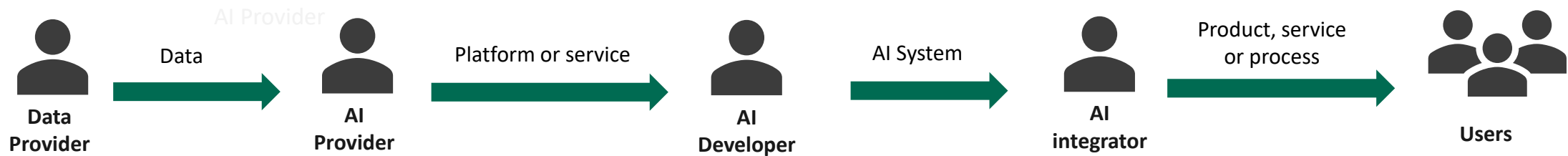
The AI lifecycle and AI Stakeholders Evolution



Traditional AI

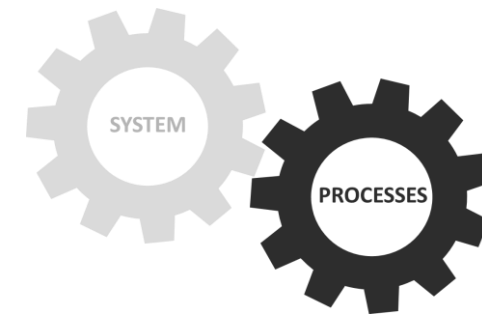


Modern AI



ISO42001: AI Management Systems

- ▶ AI Management Systems guarantee quality along complete AI lifecycle



ISO 42001: KEY COMPONENTS

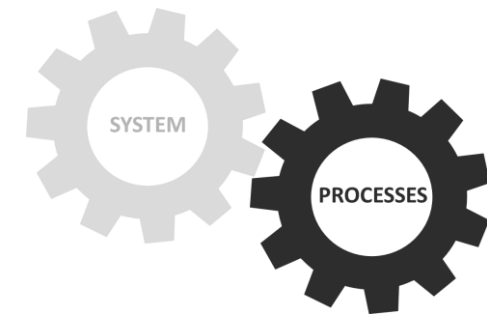
- 1** AI Policy incl. strategy and governance specifications
- 2** Documentation of all AI resources (people, data, systems, etc.) and responsibilities.
- 3** Documentation of AI lifecycle processes
- 4** Completion of a risk assessment and definition of mitigation controls.
- 5** Internal and external audit for conformity assessment.



- ▶ Covers most EU-AI Act requirements for high risk AI



ISO42001: AI Management Systems



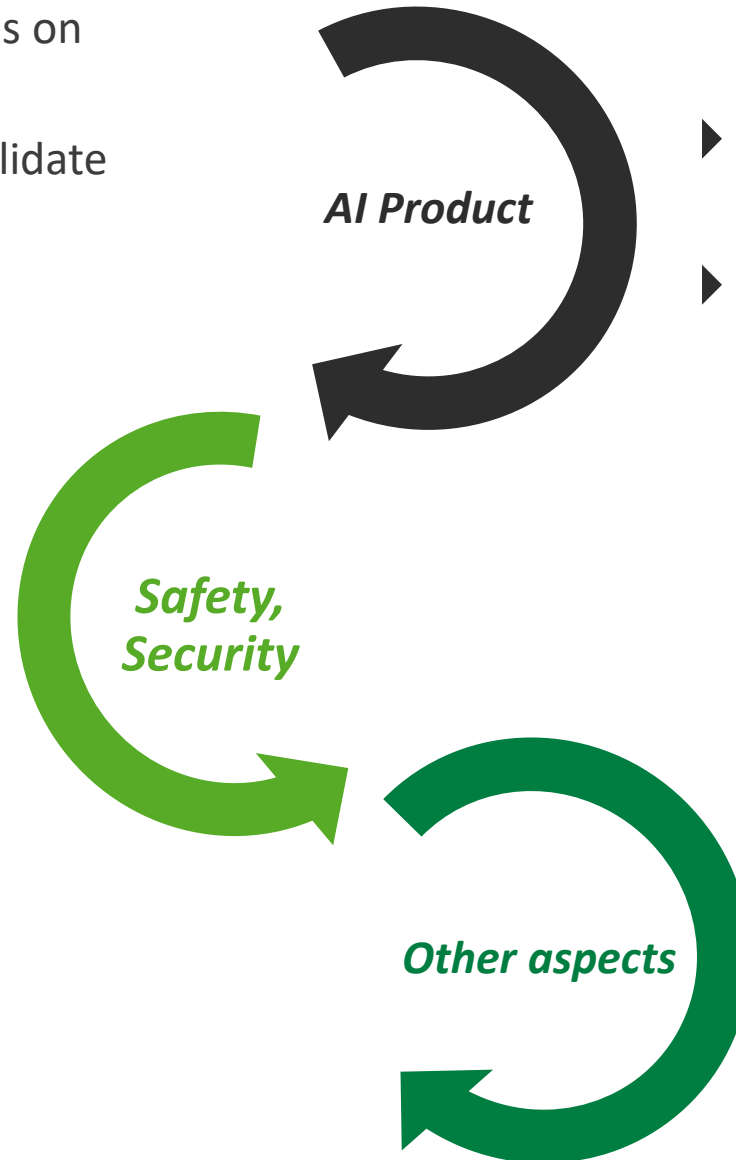
- ▶ Pivotal standard for Conformity Assessment in AI.
- ▶ Applicability beyond high-risk AI systems
- ▶ Will become standard for AI quality management system



AI System Testing

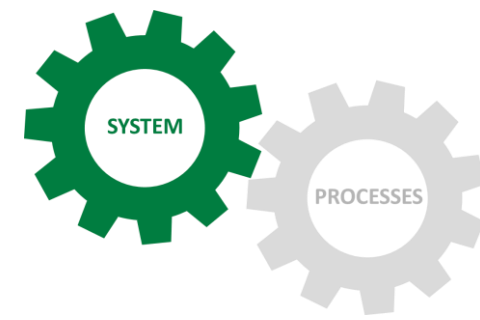
- ▶ Focus on the measurement of concrete properties on one specific object of conformity assessment.
- ▶ **Access to both AI model and data** to train and validate the model is **required**.
- ▶ Highly **dependent on the product and sector**.

- ▶ *Model Robustness (ISO24029)*
- ▶ *Trustworthiness (ISO24028)*
- ▶ *Cybersecurity (ISO27090)*

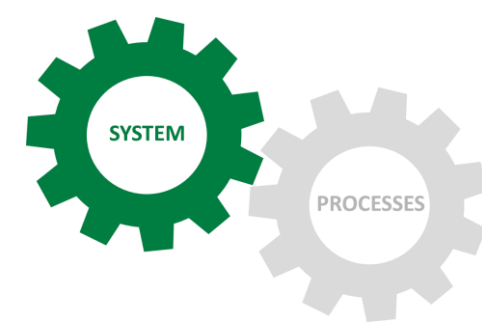


- ▶ *SW and systems engineering: Guidelines testing of AI-based systems (ISO 29119-11)*
- ▶ *AI quality product (ISO 25010)*

- ▶ *Data Quality – Part 2 – Metrics (ISO5259-2)*
- ▶ *Bias (ISO24027)*



AI System Inspection, Validation and Verification



▶ Additional activities for AI System Conformity Assessment

Inspection

Assessments checking conformance with requirements on:

- **DATA QUALITY (ISO 5259)**
 - Data Quality
 - Data Management
 - Data Governance
- **DATA PROTECTION (ISO 27701)**
- **COMPLIANCE**
- **CONTROLS**

Validation&Veritication:

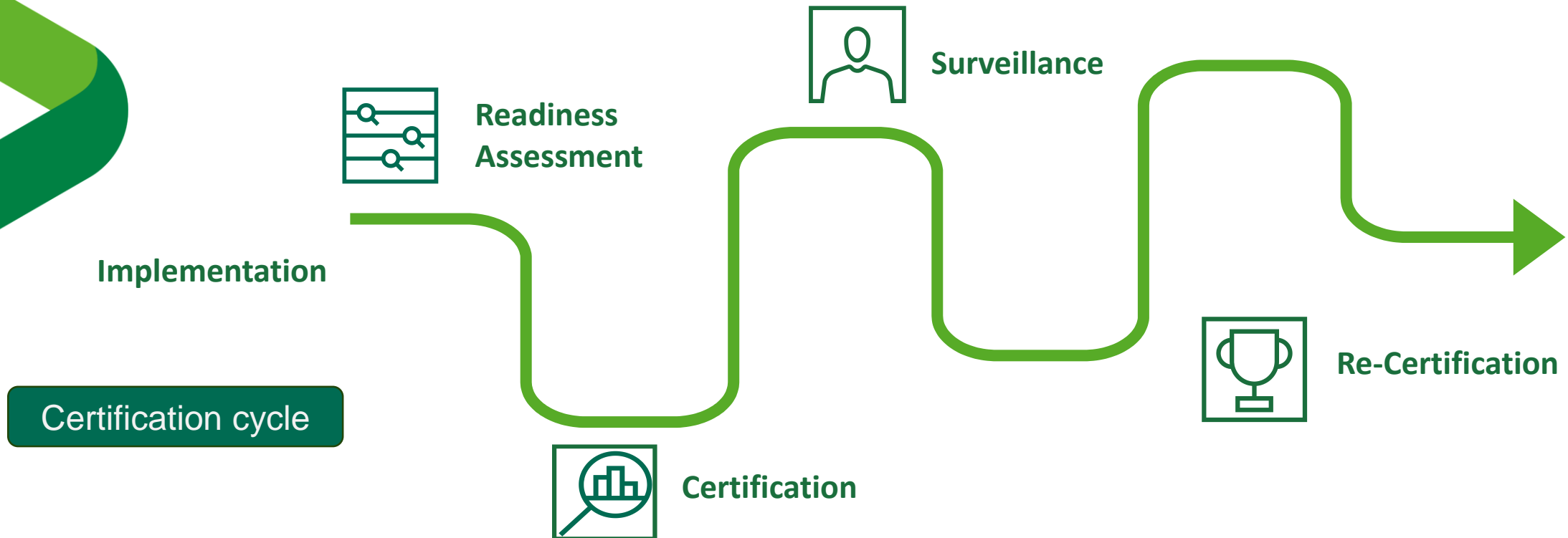
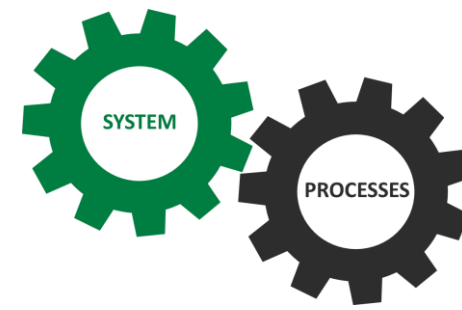
Assessment of AI algorithms aspects such as:

- **EXPLAINABILITY AND TRACEABILITY:**
 - Explainability and interpretability of AI (ISO6254)
 - Explicability & transparency of AI processing (ETSI SAI GR003)
- **FUNCTIONALITY:**
 - Prediction accuracy (ISO4213)
 - Uncertainty estimation (DIN 92005)
 - Formal verification model training (ISO17847)



Certification journey

- ▶ Allows a comprehensive conformity assessment of a complex product, process or device by integrating various conformity assessment activities.
- ▶ Certification can only be performed by a third party.
- ▶ Certification schemes to be developed, considering both Management Systems and Product Testing and Validations



Future Conformity Assessments for AI Safety and Security

Agenda



EU AI Act Conformity Assessment Obligations

3



Conformity Assessment Options for AI

9



Conclusion

18



Conclusions

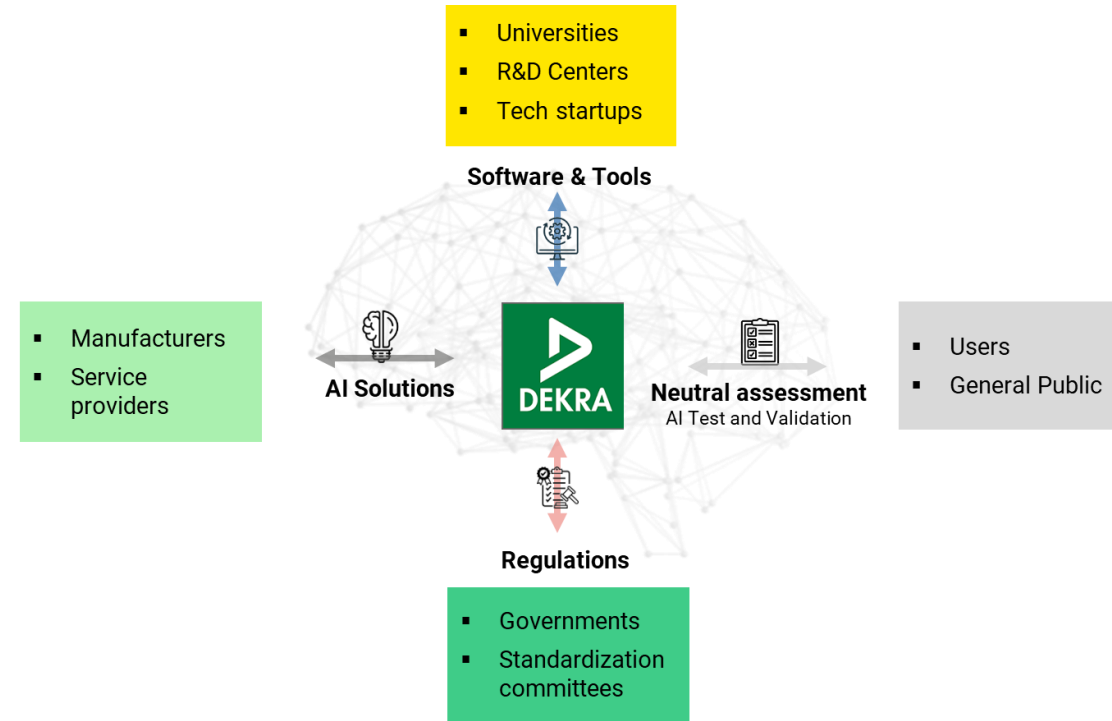
Future Conformity Assessments for AI Safety and Security



- ▶ EU AI Act imposes Conformity Assessment obligation to high-risk AI systems
- ▶ **Compliance must be addressed** from a double perspective:
 - ▶ **Processes** are widely covered in the ISO 42001- AIMS
 - ▶ **AI systems** is of higher complexity
 - Covering multiple AI aspects
 - Highly sectorial
 - Quickly evolving technology
- ▶ **Harmonized standards** and certification schemes required to operationalize Conformity Assessment obligations



Conformity Assessments guarantee AI system quality and provide trust to users, also for non-high risk AI



Thank you!

Dr. Xavier Valero
Director AI & Advanced Analytics

 xavier.valero@dekra.com

 [linkedin.com/in/xvalero](https://www.linkedin.com/in/xvalero)

Digital & Product Solutions

innovating safety & security