AI in Automotive Systems: Aligning with ISO/PAS 8800

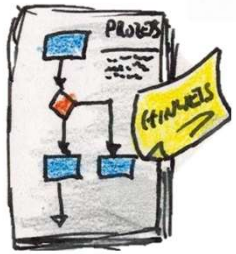
S. Keller / 21.10.2025

What makes our vehicle safe?

Laws require safety!

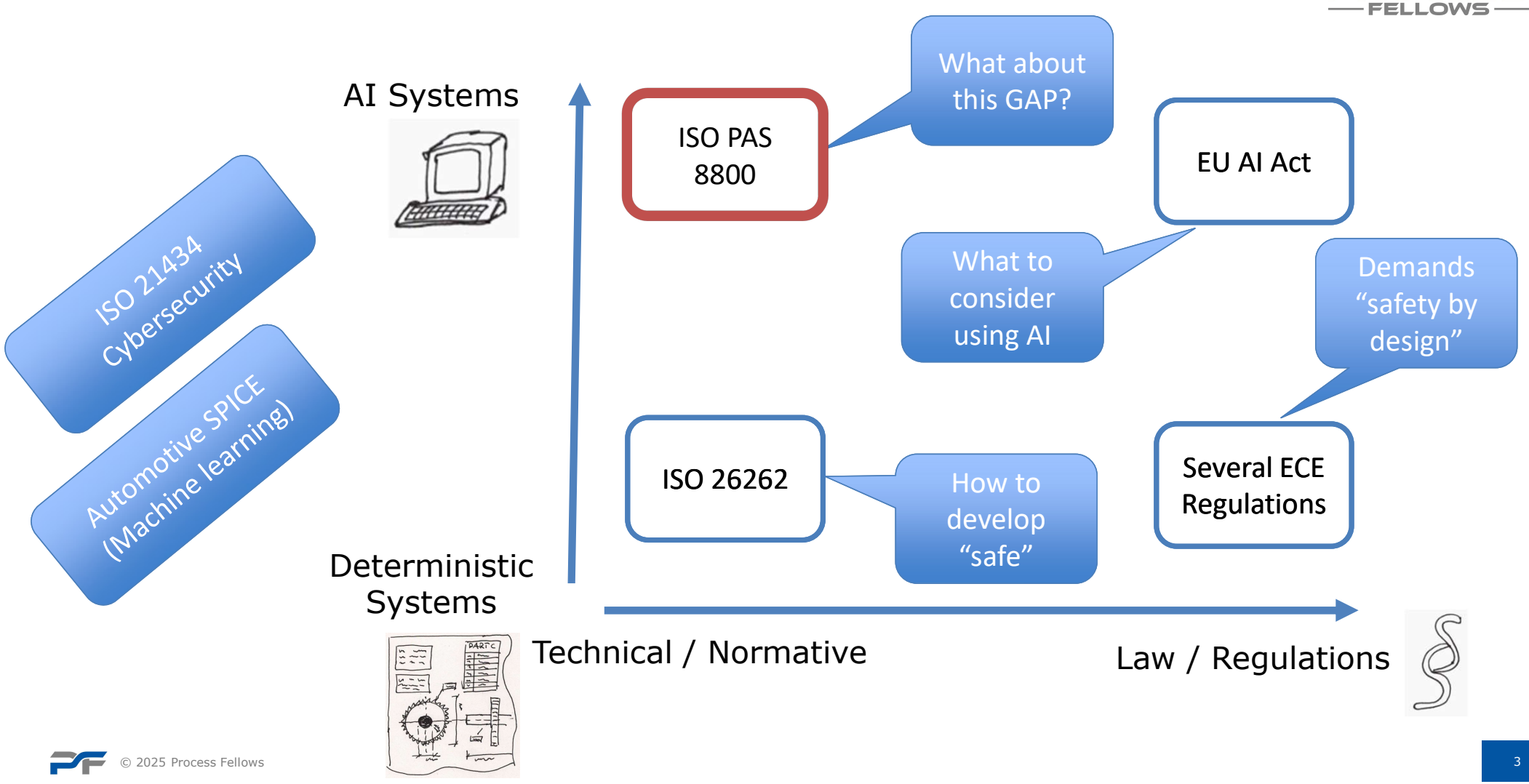
Is this enough in modern times?

Are processes enough?



Cybersecurity vs. safety?

What's relevant for safety combined with AI?



Why ISO PAS 8800? Gaps to the existing standards

Limitations of ISO 26262:

- **Assumes deterministic behavior**
  → Not suitable for learning systems with dynamic or data-driven behavior
- **No concept of training data or data quality**
  → Ignores data lifecycle despite its critical role in AI performance
- **Relies on specification-based verification**
  → AI functions can't be fully specified or verified in traditional ways
- **Lacks roles, metrics and methods for AI safety**
  → No support for handling uncertainty, model changes or explainability

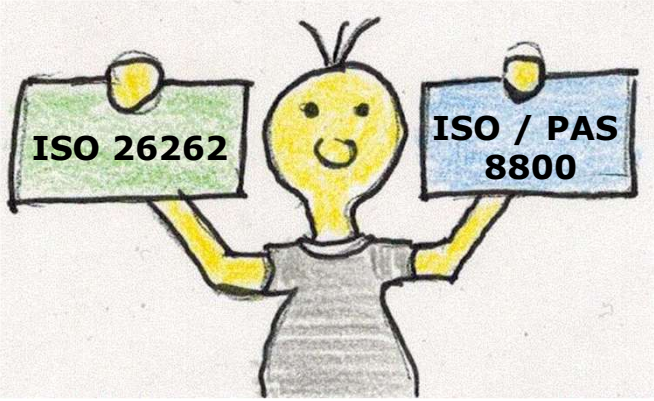
How ISO/PAS 8800 addresses these gaps

- **Considers non-deterministic and learning systems**
  → Provides guidance for AI-specific risks and behaviors
- **Introduces data and model lifecycle management**
  → Covers requirements for data quality, traceability and updates
- **Expands verification & validation approaches**
  → Supports assurance cases, monitoring and continuous evaluation
- **Defines new roles, metrics and practices for AI safety**
  → Establishes responsibilities like AI Safety Manager or Data Governance Lead

Bridging the gap between traditional safety and AI

Purpose and Scope

- **Provides guidance** for the use of **AI in safety-relevant automotive systems**
- Addresses **non-deterministic AI**, especially **machine learning** components
- Does **not replace** existing standards like ISO 26262 – it **complements** them

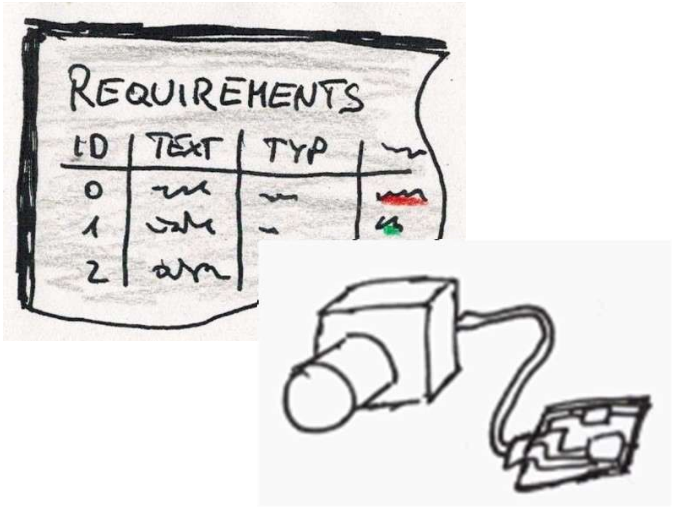


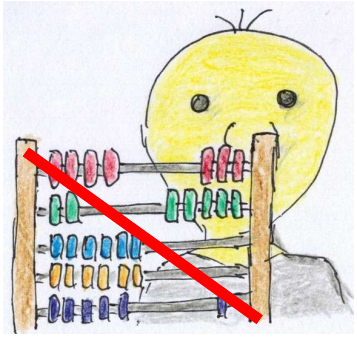
Why it matters

- AI challenges **the assumptions of traditional safety standards**
- Safety now depends on **data quality**, **model behavior** and **runtime environments**
- ISO/PAS 8800 introduces **governance, roles and risk management concepts** tailored for AI
- It forms the **missing link** between **functional safety** and **AI governance**

What makes AI safety so different?

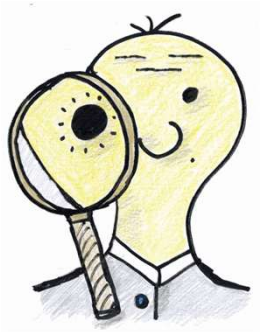
Classical approach



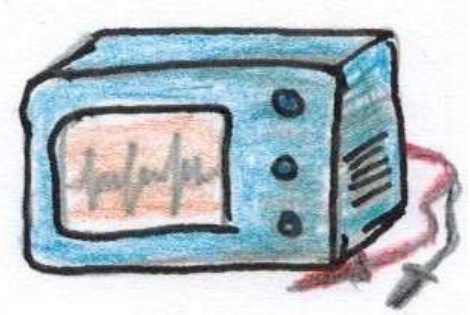


Nondeterminism

Data

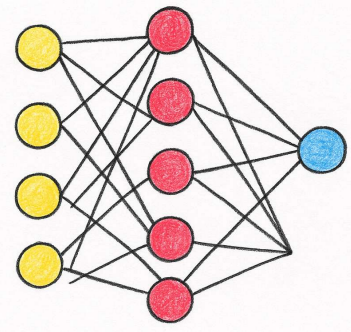


Learning instead of coding



New verification approaches

Classical Software vs. AI Systems – A Different Logic



Classical Software System

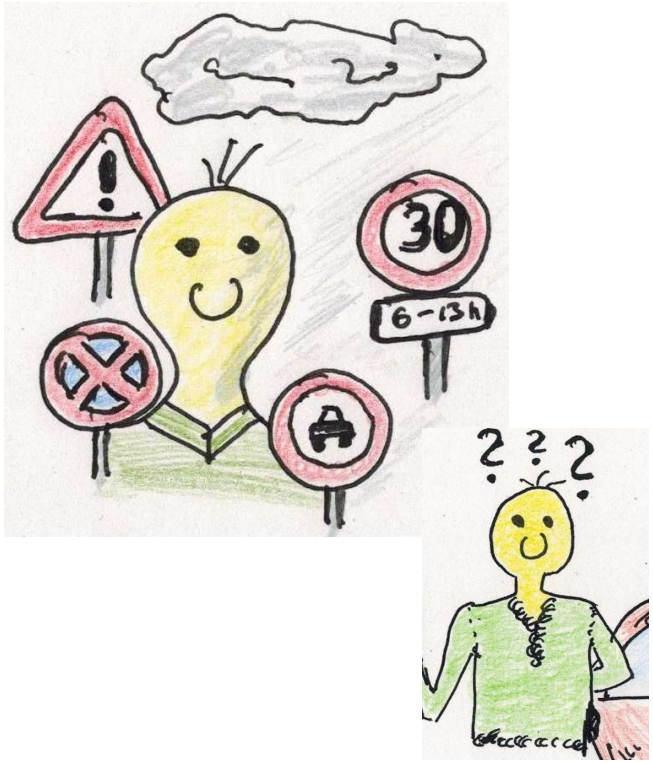
- Behavior **defined by code**
- Fully **deterministic** logic
- Verification through **requirements coverage**
- **Input/output** mapping is traceable

AI-based System

- Behavior **learned from data**
- Often **non-deterministic** and probabilistic
- Evaluation via metrics (e.g. **accuracy, confidence**)
- Internal logic often opaque (“**black box**”)

Non-Determinism & Black Boxes – A Safety Challenge

- Same input may lead to **different outputs**
- **Behavior influenced** by training data, model state, and randomness
- **Hard to explain** internal decision logic (“Why did it choose that?”)
- Safety requires traceability and justification – **AI often lacks both**
- **Black box behavior** impairs trust, debugging, and certification

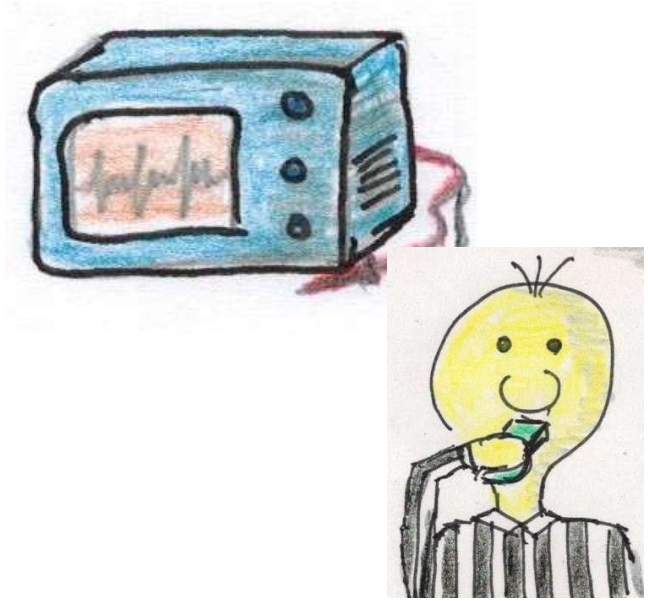


Bias, Drift & Garbage – When Data Becomes a Risk

- Training **data** defines system behavior – it becomes **part of the product**
- **Poor data = poor safety**, even if the model is well designed
- Bias: Data is **not representative** (e.g. weather, demographics, rare cases)
- Drift: **The world changes** – data from last year may no longer be valid
- **Garbage in → Garbage out**: No QA = unknown risks in deployment
- Data needs **governance, traceability and monitoring** – like code

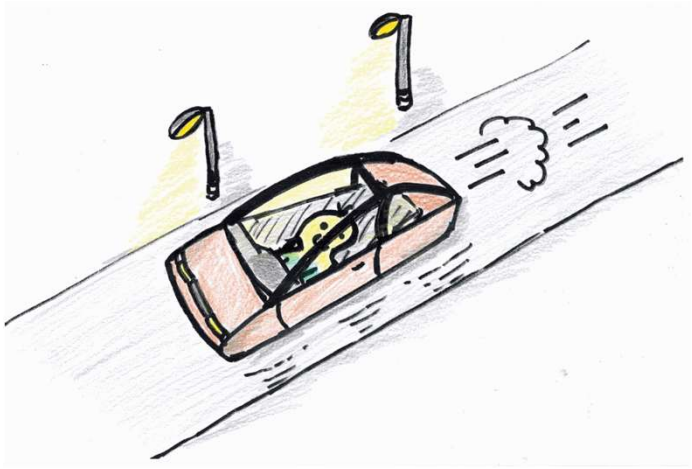
Verification & Validation – When Test Coverage Isn't Enough

- AI behavior **can't be fully specified** – classic requirements don't capture it
- No **fixed output for each input** → how to test what's "correct"?
- Traditional test coverage (e.g. MC/DC) **doesn't apply** to learned logic
- Need for statistical validation, robustness testing, uncertainty metrics
- V&V must include **model behavior, data quality, and training process**
- Safety arguments shift from "Did we test enough?" to "**Do we understand what it does**?"



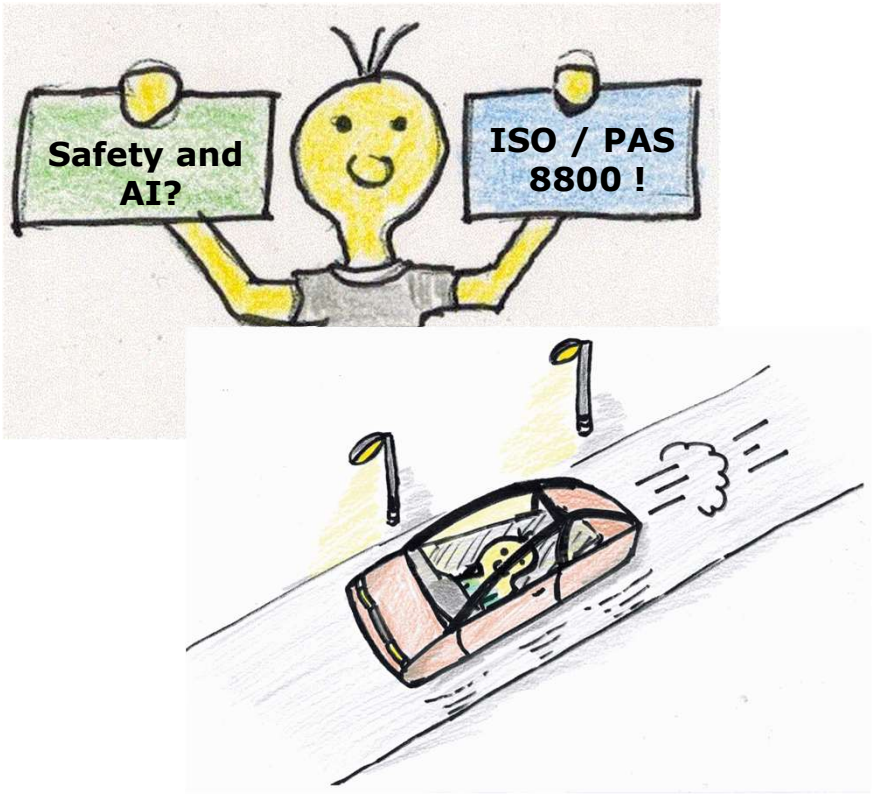
AI After SOP – Staying Safe in the Real World

- Operational **context changes** – AI models may become less reliable over time
- **Monitoring** needed: performance, data drift, unexpected behaviors
- No runtime learning → but **feedback loops** via retraining & updates
- Safety must cover **update strategies**: re-validation, rollback, traceability
- **PAS 8800** stresses need for post-deployment lifecycle management
- **Organizational roles**: who owns AI performance after release?



What we are talking about – ISO / PAS 8800

- Published in 2024 by ISO TC 22/SC 32/WG 8 – **same committee as ISO 26262**
- Focus: **Safety-related use of AI** in road vehicles
- Scope: **Machine learning functions** for perception, decision & control
- Addresses **limitations of ISO 26262 & SOTIF** in AI-heavy systems
- Covers the **entire lifecycle** – from concept to post-deployment updates
- Builds toward **future standardization** – currently a PAS (specification)

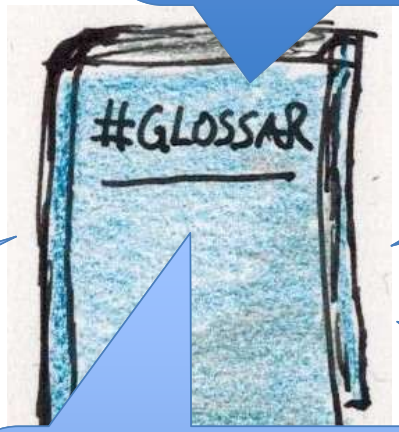


Key Terms in ISO/PAS 8800

AI Use Case
A specific safety-relevant application of AI in the automotive context (e.g., lane detection, object recognition)

Trustworthy AI Function
An AI function that demonstrably behaves reliably and safely – based on training coverage, evaluation metrics, etc.

AI Safety Goal
A safety requirement tailored to AI-specific risks (e.g., failure to handle unknown inputs)

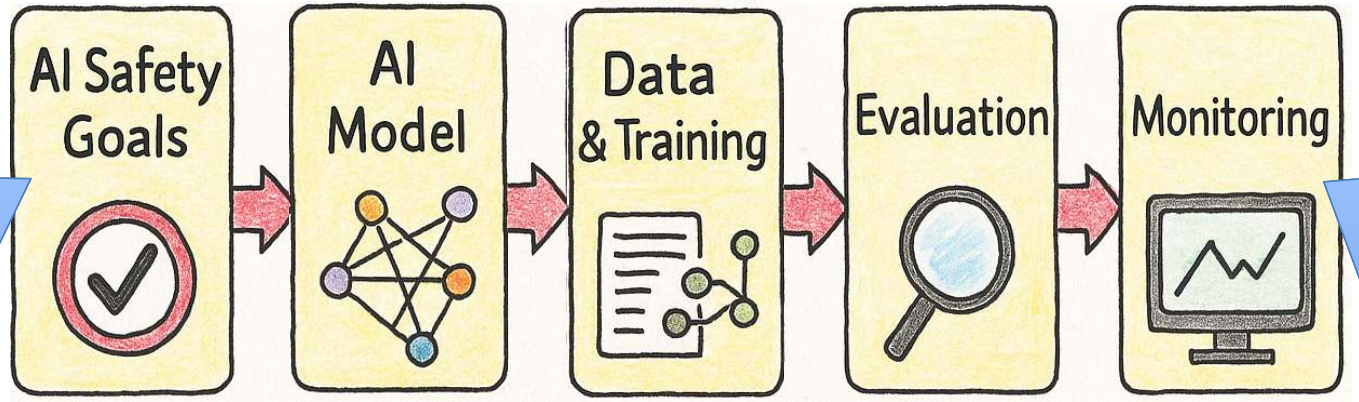


AI Item
A system or subsystem that includes one or more AI functions; based on ISO 26262's "Item", but adapted for AI

Confidence Measure
A quantifiable or qualitative indicator of trust in the AI model's behavior – such as robustness tests or uncertainty analysis

Data Strategy
A structured plan for handling data across the AI lifecycle – includes data sourcing, curation, testing and monitoring

ISO/PAS 8800 – Mapping Lifecycle Phases to Standard Sections



Clause 9
Identification of safety-relevant AI use cases and derivation of AI Safety Goals from system context and potential hazards

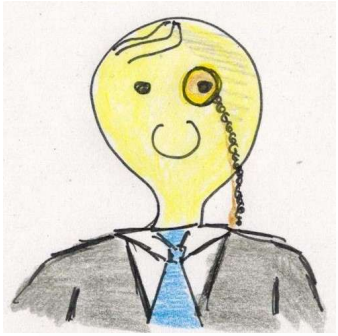
Clause 10
Definition of the AI model including architecture, configuration management and model versioning

Clause 11
Requirements for data quality, data collection, labeling, training dataset strategy and data governance activities

Clause 12 + 13
Verification of AI performance, robustness testing, uncertainty analysis, and definition of confidence metrics.

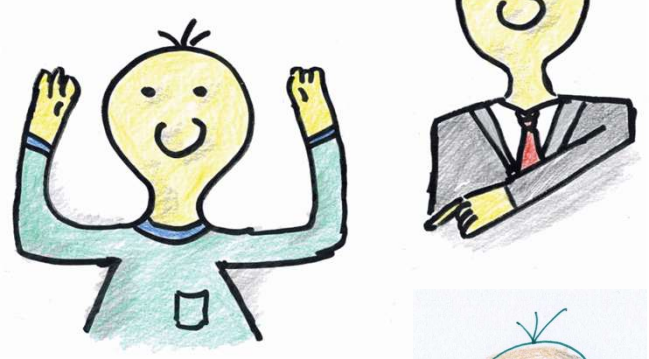
Clause 14/Annex B
Operational monitoring, model drift detection, and update strategies after deployment

Established Roles (ISO 26262)

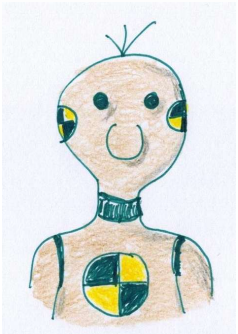


Safety Manager

Project Lead



Architect / Developer



Tester

- **Safety Manager**: This role **remains essential** for traditional software systems. In AI systems, coordination with the AI Safety Manager is key.
- **Project Lead**: Still relevant, but now **responsible for integrating AI-specific roles** into planning and risk management.
- **Architect/Developer**: **Roles expanded to cover ML** models, data dependencies, and lifecycle planning for AI systems.
- **Tester**: Existing role, now **extended to include AI-specific challenges** such as coverage of data-driven behavior and black-box explainability.

New Roles and Responsibilities

AI Safety Manager

Data Governance Lead

Operational Monitoring Owner

- **AI Safety Manager:** Owns **AI safety planning** and lifecycle-wide argumentation.
- **Data Governance Lead:** Ensures **data** quality, traceability, and compliance across all phases.
- **Operational Monitoring Owner:** Oversees **post-deployment** behavior, drift detection, and updates.

Why Traditional Safety Arguments Don't Work for AI

Classical approach

Relies on **deterministic** behavior

Uses **requirements-based traceability** (Req → Code → Test)

Emphasizes **full explainability** of logic and control flow

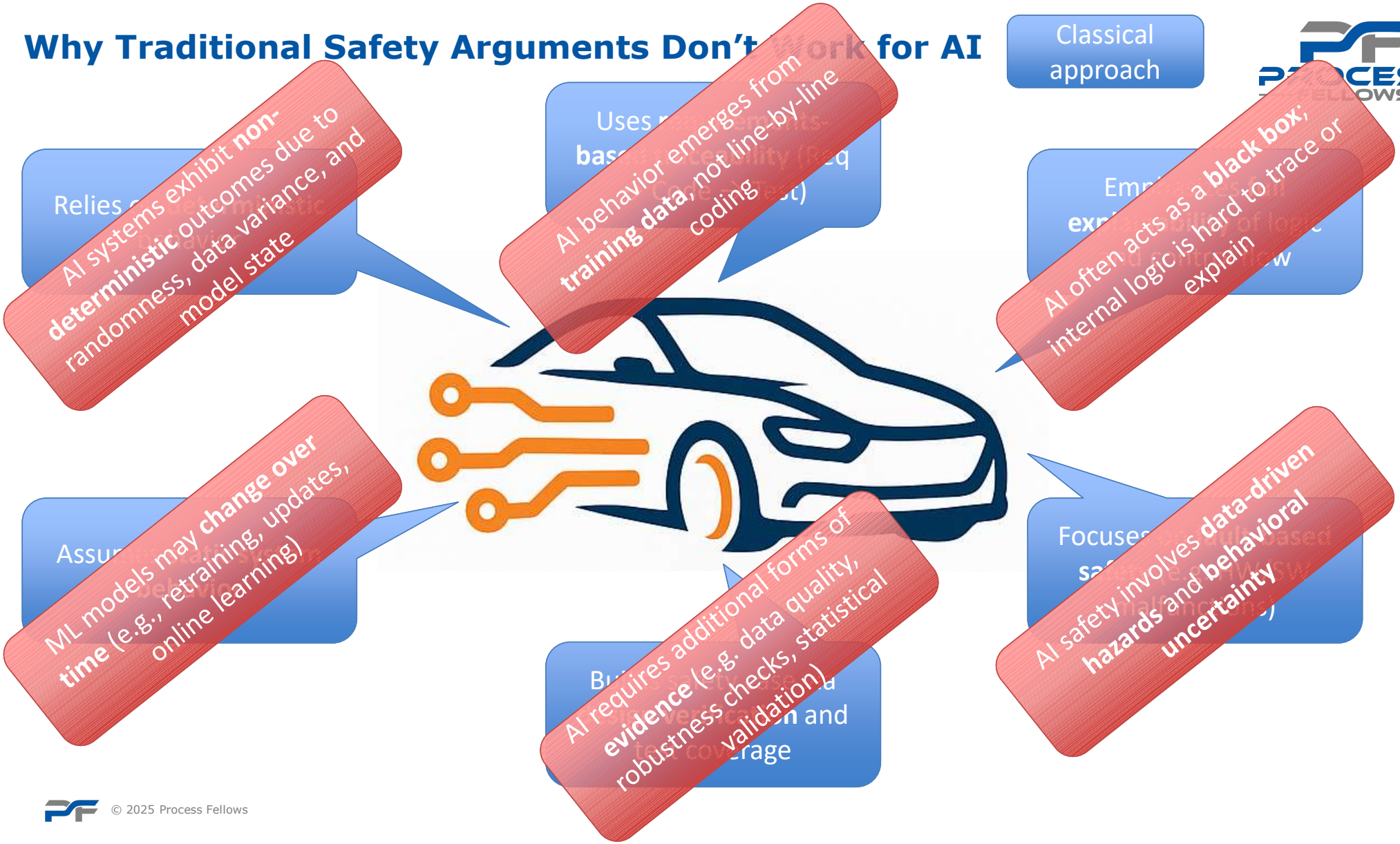
Assumes **static system behavior**

Builds safety case via **design verification** and test coverage

Focuses on **fault-based safety** (e.g. HW/SW malfunctions)

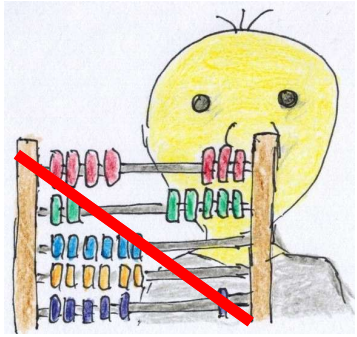
Why Traditional Safety Arguments Don't Work for AI

Classical approach



How the ISO/PAS 8800 handles those topics

Non-determinism



Clause 13.4 – AI-specific robustness and uncertainty analysis:
AI systems' non-deterministic behavior is addressed through robustness testing, statistical validation, and AI-specific safety analyses such as metamorphic or combinatorial testing.

Clause 11 – Dataset lifecycle management
Training data are treated as safety-relevant artefacts. The PAS introduces a dataset lifecycle covering collection, labeling, validation, and governance to ensure data integrity and traceability.

Behavior emerging from Training Data

Black Box Behavior

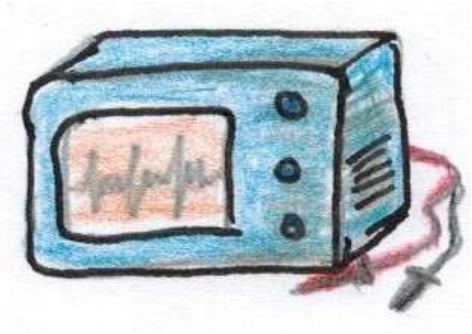
Clause 8.4 – Assurance argumentation and confidence metrics
Instead of direct introspection, the PAS requires quantitative assurance arguments with measurable confidence levels, uncertainty bounds, and explainability evidence..

How the ISO/PAS 8800 handles those topics

Changes over Time (drift)

Annex G – Monitoring and drift management
Post-deployment monitoring must detect, quantify, and mitigate data or concept drift. Continuous evidence collection supports model updates and re-evaluation.

Clause 8.3 – AI safety assurance case
The PAS expands the notion of safety evidence beyond testing. It includes data quality records, model-training documentation, evaluation results, and lifecycle work products as part of the assurance case.

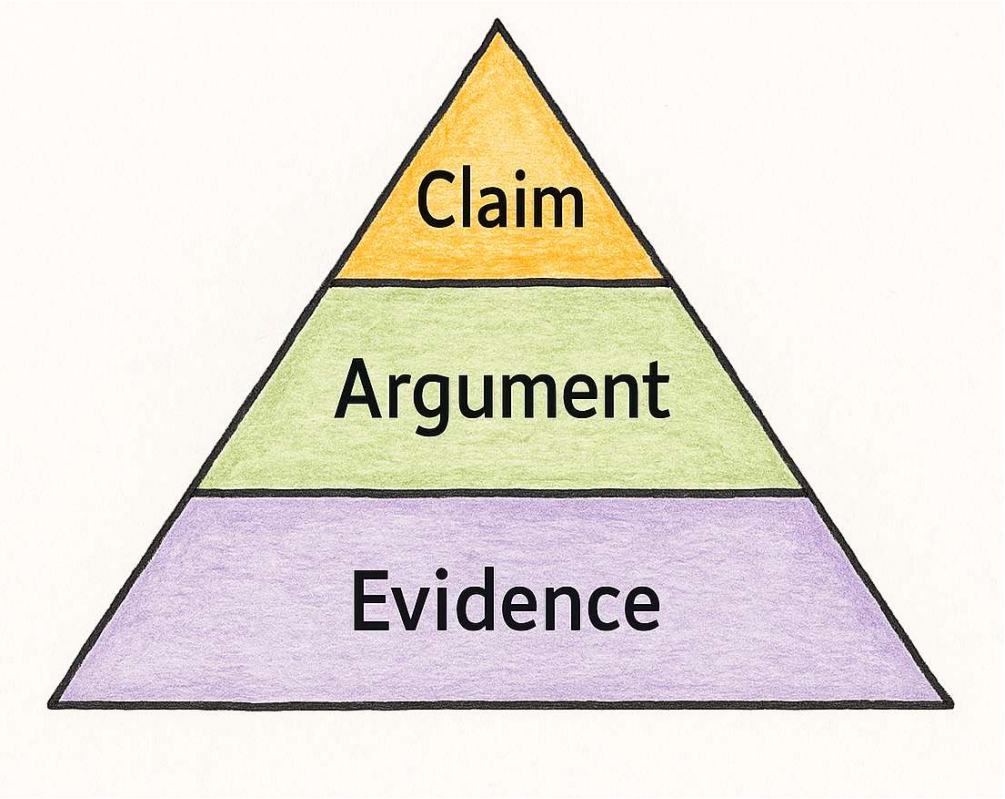


Additional evidence

Data driven Hazards

Clause 8.4(b) – Statistical reasoning for input space coverage
The PAS introduces statistical safety arguments to demonstrate adequate coverage of the input space, including edge-case behavior and spurious correlations.

Assurance Argumentation



- A claim is a true-false statement about a property of the AI system and its limitations, including associated uncertainties.

  Reference: ISO/PAS 8800:2024, Clause 3.3.3 (Definitions of Claim) Clause 8.5.1 (Context of Assurance Argument)

- An assurance argument is a reasoned, auditable artefact linking claims, arguments, and evidence to justify confidence that safety requirements are achieved.

  Reference: ISO/PAS 8800:2024, Clause 3.3.2 (Assurance Argument), Clause 8.5.1–8.5.2 (Structuring Assurance Arguments & Evidence Categories)

- Evidence shall consist of relevant work products generated during the AI safety lifecycle, supporting the assurance claims.

  Reference: ISO/PAS 8800:2024, Clause 8.3.2 and 8.5.2 (Categories of Evidence), Clause 8.8 (Work Products)

Core deliverables required by ISO/PAS 8800

AI Safety Plan – defines scope, responsibilities, methods and activities

AI Safety Case Summary – structured argument integrating all lifecycle evidence

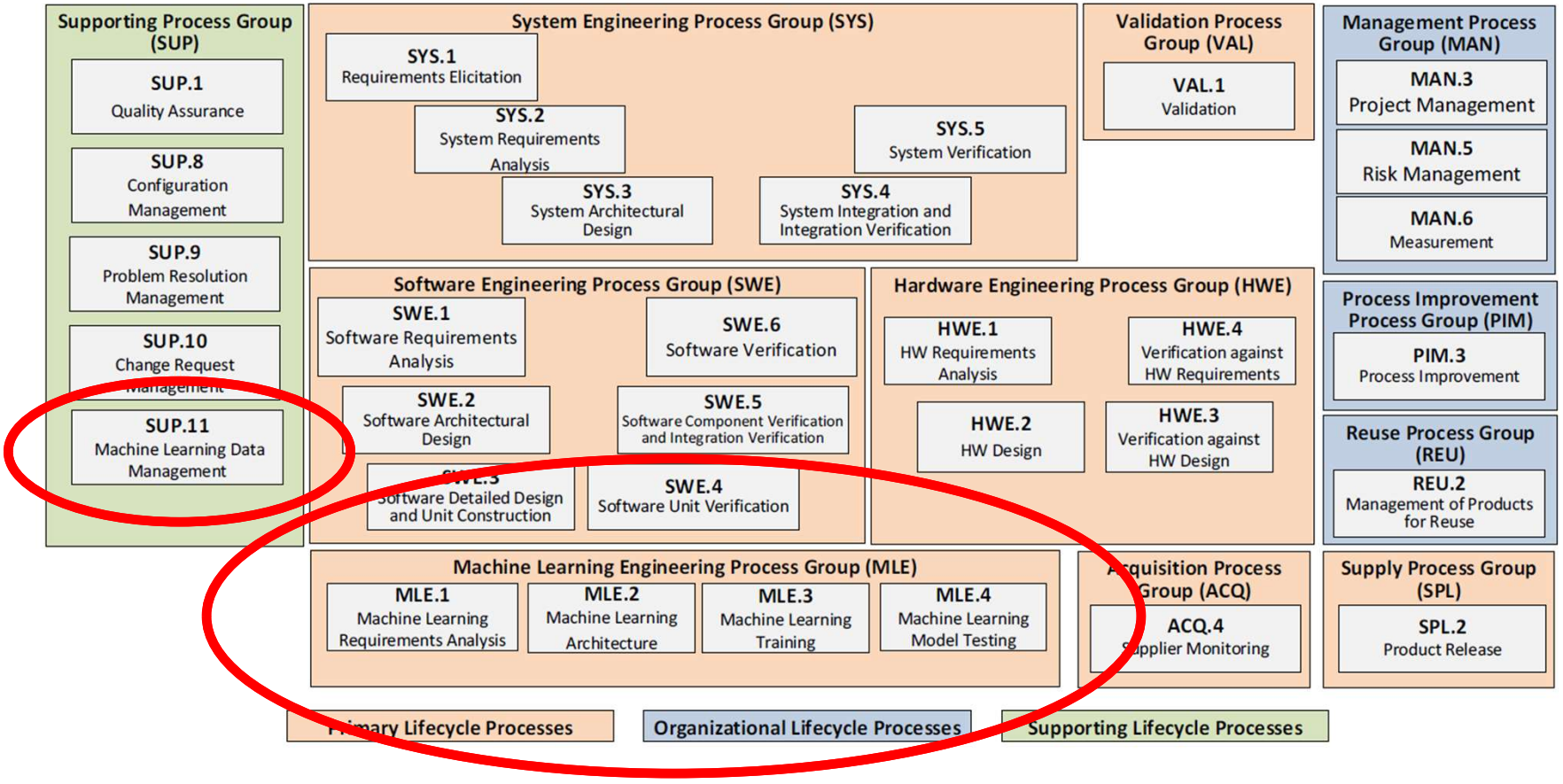
Dataset Lifecycle Report – covers collection, labeling, quality control and maintenance

Model Definition & Training Report – describes AI model architecture, configuration, training data, and versioning

Monitoring and Drift Management Plan – specifies post-deployment monitoring, detection, and re-evaluation activities

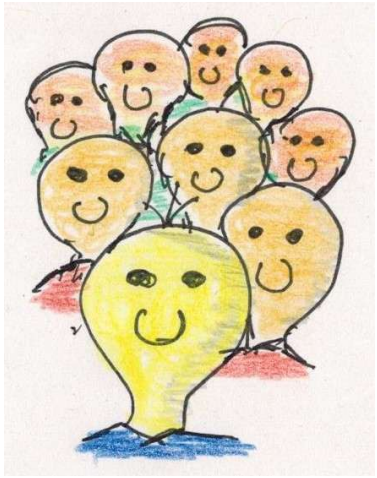
Model Evaluation & Robustness Report – summarizes verification and validation evidence

Applying ISO/PAS 8800 in Practice – Challenge your process world



[Find additional Information at Spice4Cars]

Applying ISO/PAS 8800 in Practice – Roles and work products



Define and apply new roles
Align AI-safety roles (AI Safety Manager, Data Governance Lead) with existing ASPICE roles (Project Lead, Quality Manager)

Create the first work products:

- AI Safety Plan aligned with classical Safety Plan
- Dataset Lifecycle Template (possibly aligned with ASPICE "Machine Learning Data Management" process)
- Monitoring & Model Drift Plan

AI Safety in Automotive – Where to Start and How to Proceed

Key Takeaways

- **ISO/PAS 8800** bridges the gap between *AI innovation* and *functional safety*.
- The **AI lifecycle** (Goals → Data → Model → Evaluation → Monitoring) complements ISO 26262
- Success depends on **data governance, traceability, and explainability** — not just model performance.

Next Steps for Organizations

- Assess current **AI maturity** and identify missing safety work products.
- Integrate ISO/PAS 8800 activities into **ASPICE roles and processes**.
- Start small: pilot an **AI Safety Plan** and **Dataset Lifecycle Template** on one project.

Thank you for your attention

If you'd like to learn more about **how to implement ISO/PAS 8800 and integrate AI safety** into your ASPICE processes, **feel free to contact us** at Process Fellows. Additional insights and resources are available at:

Spice4cars.com