Why ISO/PAS 8800 is pivotal across all technology sectors including aerospace

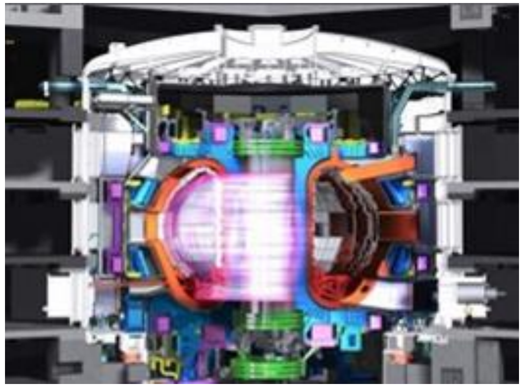
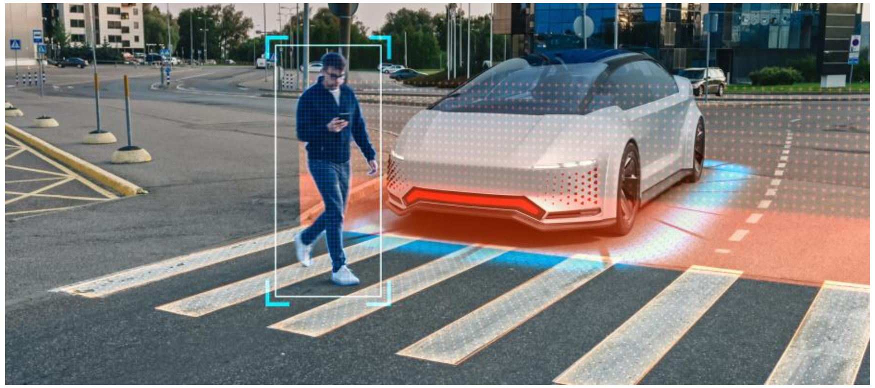
European Space Agency – Software Product Assurance Conference 2025

Jeff Joyce, Ehsan Ghahremani, Jonathan Groves, Laure Millet

Presentation Abstract

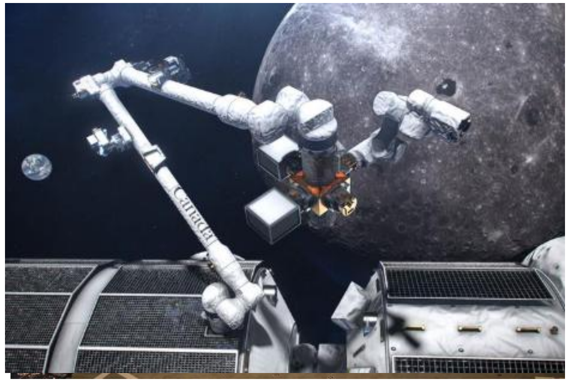
ISO/PAS 8800 is a publicly available specification for the use of “AI systems”. Although developed for the automotive industry, most of its technical content is useful across a broad range of technical domains including aerospace. This specification is pivotal because it addresses the unique challenges of managing safety risk for AI systems in a way that is not simply a patch-like accommodation of AI within a conventional approach. Many technical sectors are actively seeking to close the wide gap between functional safety standards and the use of AI in high assurance systems. However, ISO/PAS 8800 stands out as a significant step forward. It could be used outside the automotive industry as an interim measure and inspiration for future standards and other forms of published guidance.

Critical Systems Labs






Critical Systems Labs
Innovating Safely

How is AI/ML different?

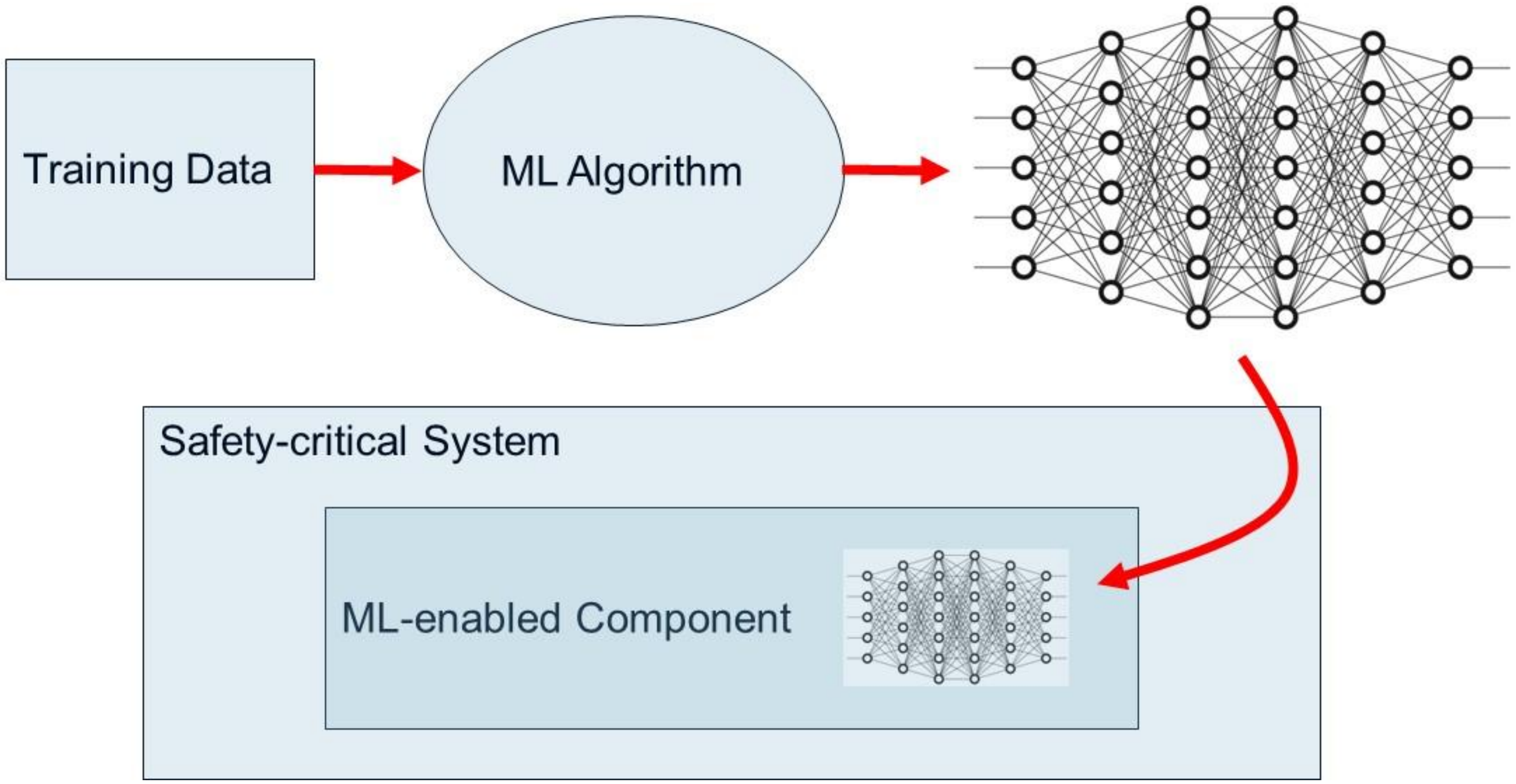
Non AI/ML

- Functional safety requirements are ***specified***
- Design, implementation and V&V is based on ***specified requirements***

AI/ML

- ML Models are generated by ML algorithms from training data
- Confidence depends heavily on adequacy of testing (in the ***absence of specified requirements***)

ML Development

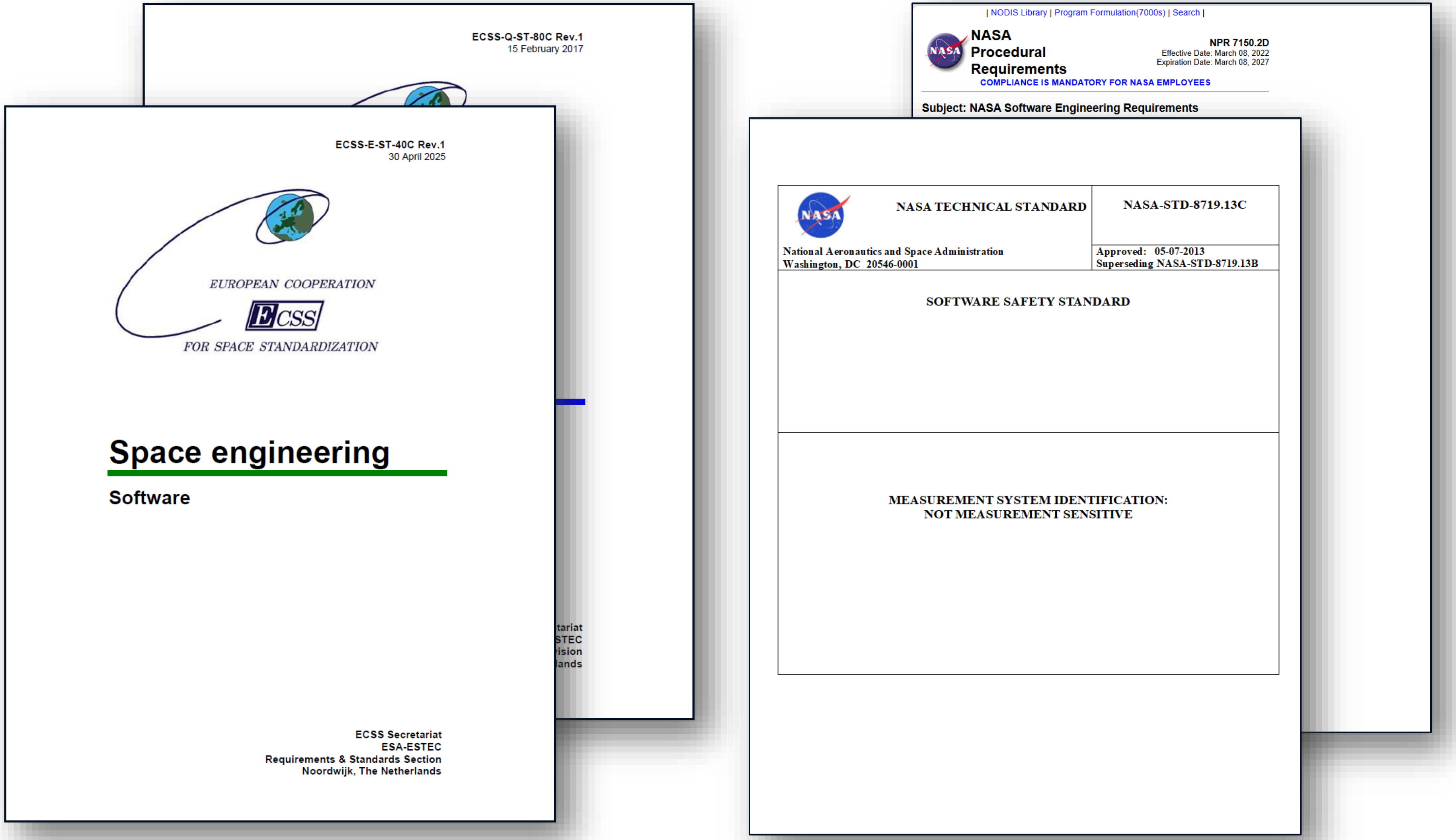


Other considerations

- With AI/ML expect the pace of software changes (e.g., ML models updates) to accelerate from years to months or even weeks
- Conservative risk adverse mindset of traditional safety assurance is at odds with “move fast and break things” culture of AI/ML

Critical Systems Labs
Innovating Safely

Established guidance based on correct implementation of functional safety requirements



Critical Systems Labs
Innovating Safely

ECSS-E-ST-40C Rev.1
30 April 2025

EUROPEAN COOPERATION
ECSS
FOR SPACE STANDARDIZATION

Space engineering

Software

ECSS Secretariat
ESA-ESTEC
Requirements & Standards Section
Noordwijk, The Netherlands

5.4.3 Software architectural design

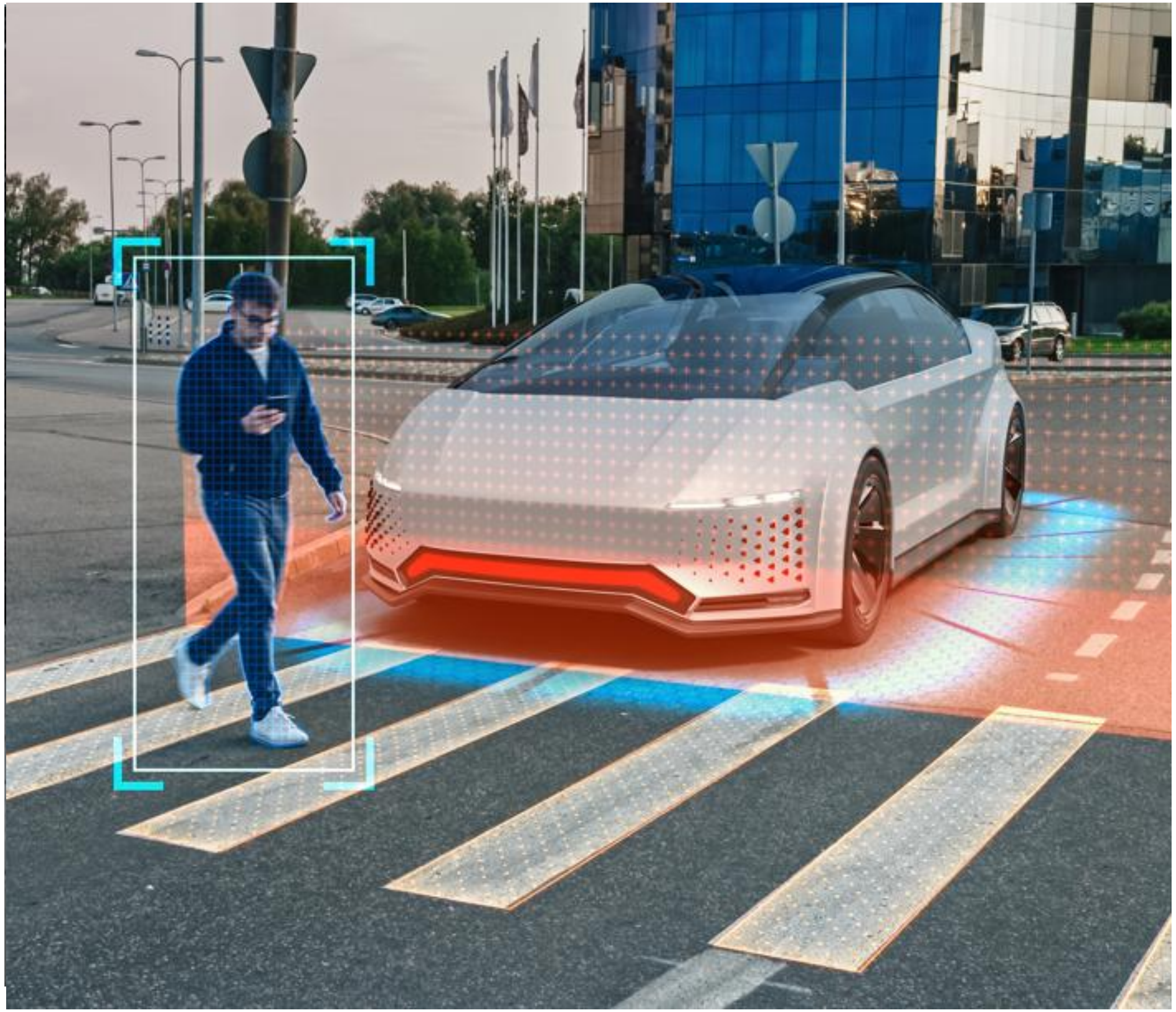
5.4.3.1 Transformation of software requirements into a software architecture

ECSS-E-ST-40_0860059

a. The supplier shall transform the requirements for the software item into an architecture that:

1. describes its top–level structure;
2. identifies the software components, ensuring that all the requirements for the software item are allocated to its software components and later refined to facilitate detailed design;
3. covers as a minimum hierarchy, dependency, interfaces and operational usage for the software components;
4. documents the process, data and control aspects of the product;

… and with a similar requirements centric approach to design, implementation, verification and validation



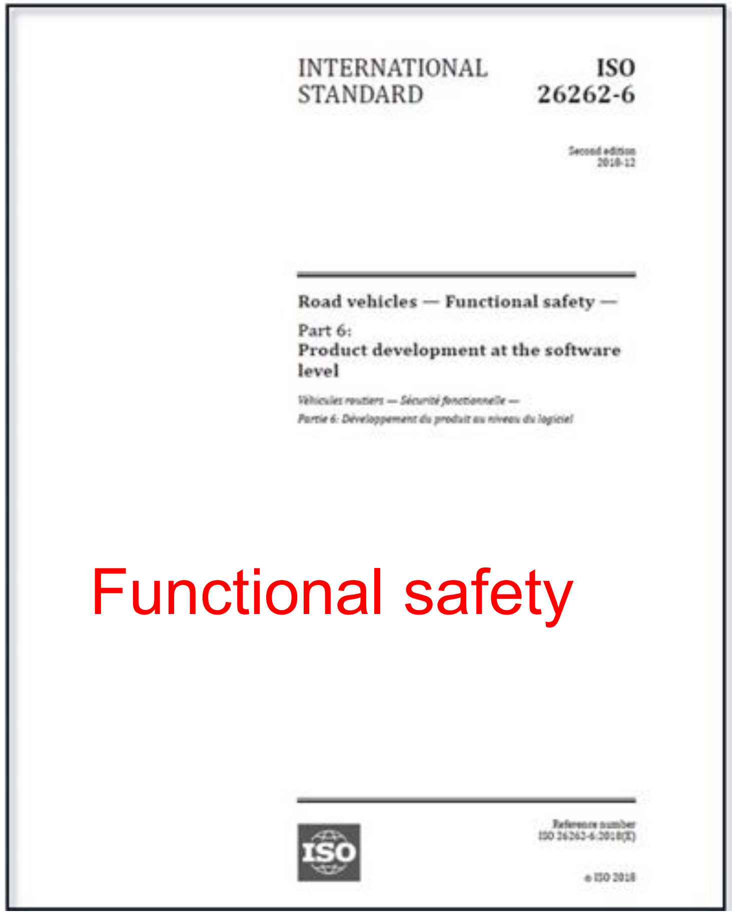
Perception

Localization

Motion Planning

...

ISO Standards for Automotive Safety


INTERNATIONAL STANDARD

ISO 26262-6

Second edition 2018-12

Road vehicles — Functional safety —

Part 6:

Product development at the software level

Véhicules routiers — Sécurité fonctionnelle —

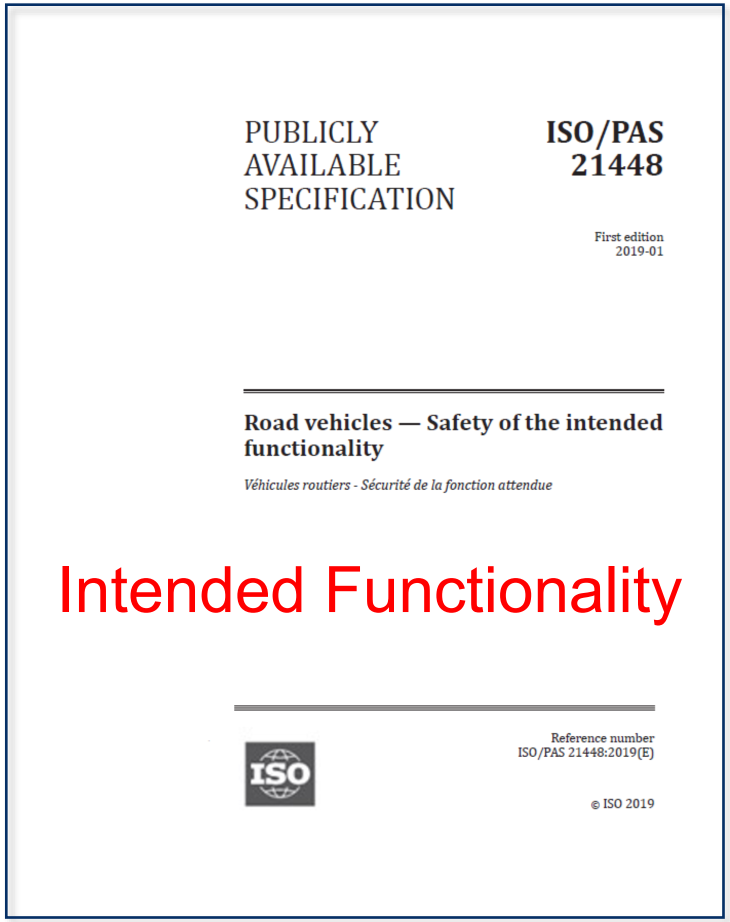
Partie 6: Développement du produit au niveau du logiciel

Reference number ISO 26262-6:2018(E)

© ISO 2018

Functional safety

2011, 2018


PUBLICLY AVAILABLE SPECIFICATION

ISO/PAS 21448

First edition 2019-01

Road vehicles — Safety of the intended functionality

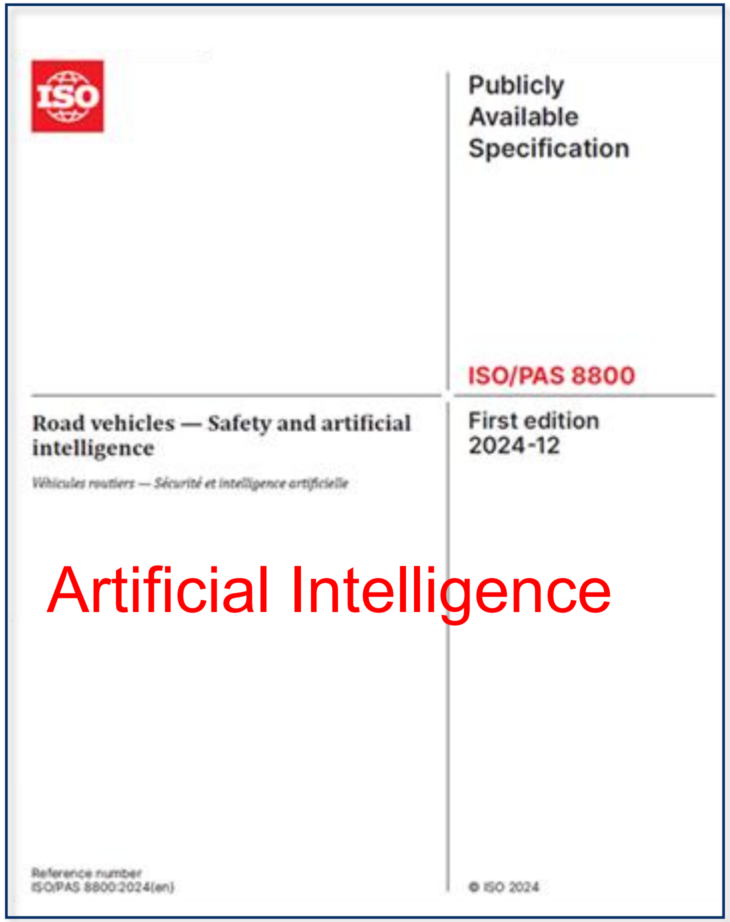
Véhicules routiers - Sécurité de la fonction attendue

Reference number ISO/PAS 21448:2019(E)

© ISO 2019

Intended Functionality

2019, 2022


Publicly Available Specification

ISO/PAS 8800

First edition 2024-12

Road vehicles — Safety and artificial intelligence

Véhicules routiers — Sécurité et intelligence artificielle

Reference number ISO/PAS 8800:2024(en)

© ISO 2024

Artificial Intelligence

2024

ISO 26262
Functional Safety


INTERNATIONAL STANDARD

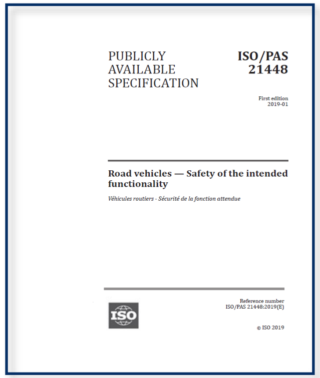
ISO 26262-6

Road vehicles — Functional safety —
Part 6:
Product development at the software level

Is the specified functionality adequately safe?

Is the specified functionality correctly implemented?

ISO 21448
Safety of the Intended Functional

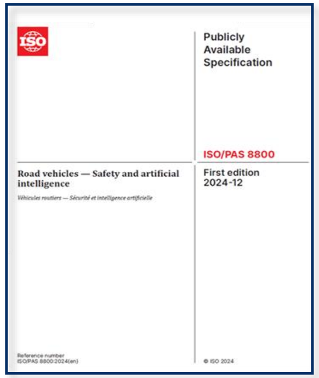

PUBLICLY AVAILABLE SPECIFICATION

ISO/PAS 21448

Road vehicles — Safety of the intended functionality

Does the system adequately address the functionality insufficiencies of system elements?

ISO/PAS 8800
AI Safety


Publicly Available Specification

ISO/PAS 8800

Road vehicles — Safety and artificial intelligence

First edition
2024-12

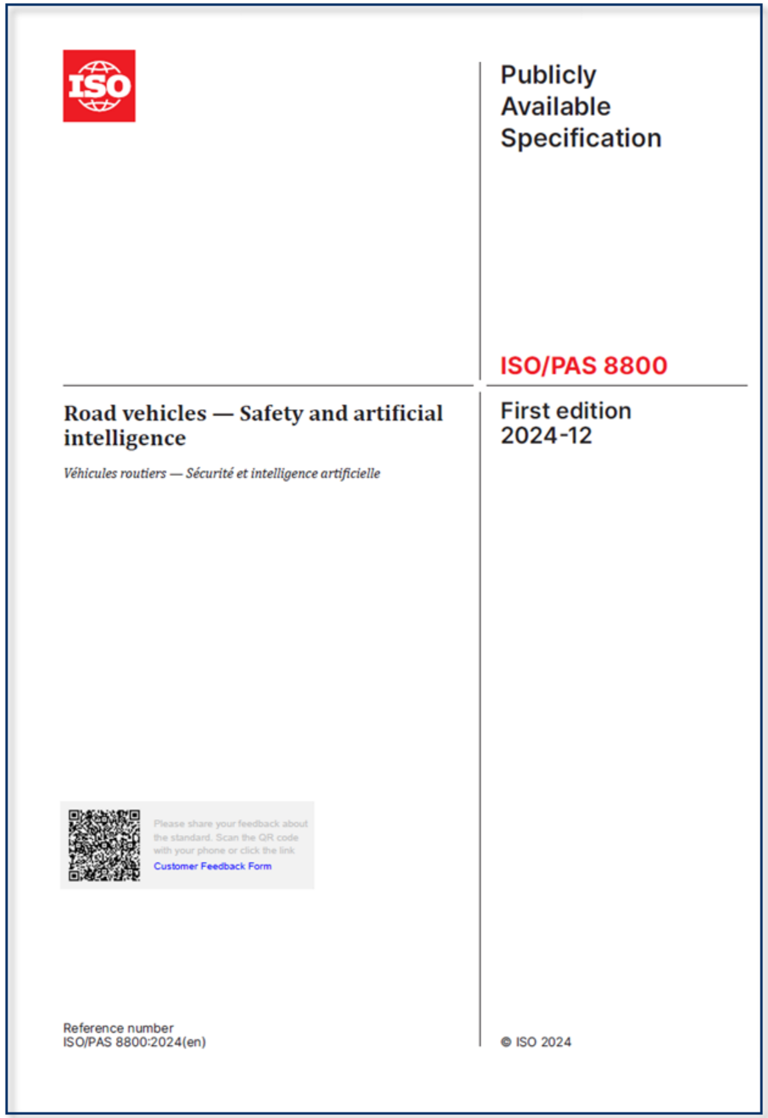
Can decisions made by the AI system be trusted?

Critical Systems Labs
Innovating Safely

ISO/PAS 8800

"… it is not possible to provide detailed requirements on the process or product characteristics required to achieve an acceptably low level of residual risk associated with the use of AI systems"

"... this document focuses on the principles that support the creation of a project-specific assurance argument for the safety of the AI elements within on-board vehicle systems"



Publicly Available Specification

ISO/PAS 8800

Road vehicles — Safety and artificial intelligence

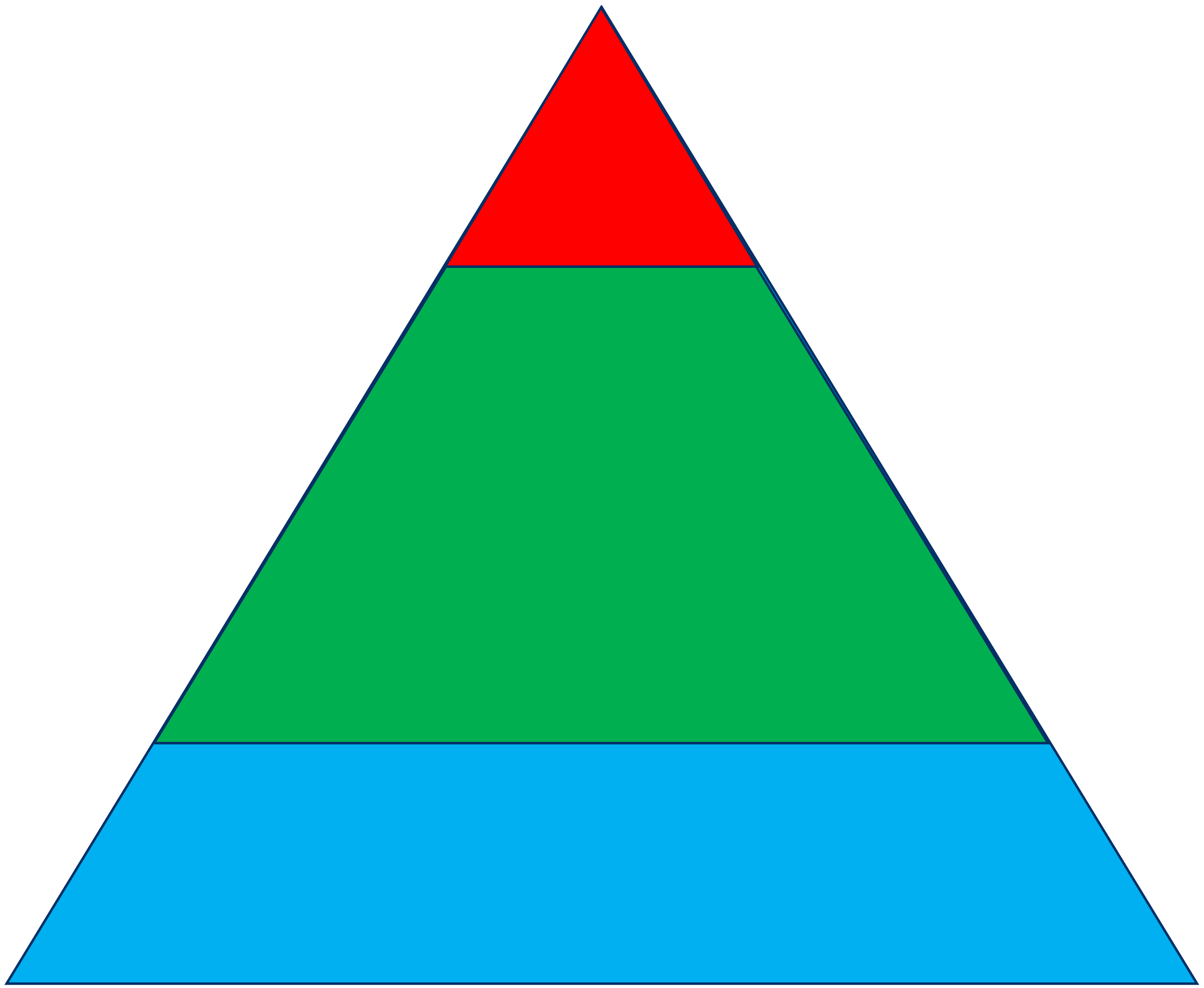
Véhicules routiers — Sécurité et intelligence artificielle

First edition
2024-12

Please share your feedback about the standard. Scan the QR code with your phone or click the link
Customer Feedback Form

Reference number
ISO/PAS 8800:2024(en)

© ISO 2024

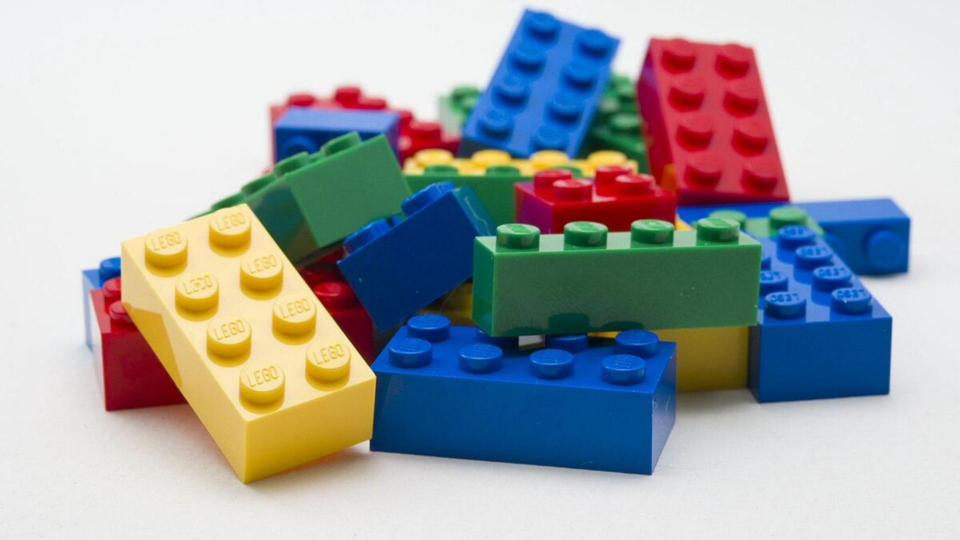


Claim

Argumentation

Evidence

“Lego Blocks” for Making Arguments



C####
A claim forming part of the argument.

[**C**]LAIM

D####
A potential doubt within the argument.

[**D**]EFEATER

S####
An outline of the argumentation strategy used for sub-arguments.

[**S**]TRATEGY

IR####
A rule outlining how to combine siblings in support of the parent.

[**I**]NFERENCE

E####
Description of supporting data or evidence.

[**E**]VIDENCE

A####
A statement assumed to be true, without further argumentation.

[**A**]SSUMPTION

X####
Contextual information needed to increase understanding.

CONTE[**X**]T

Res

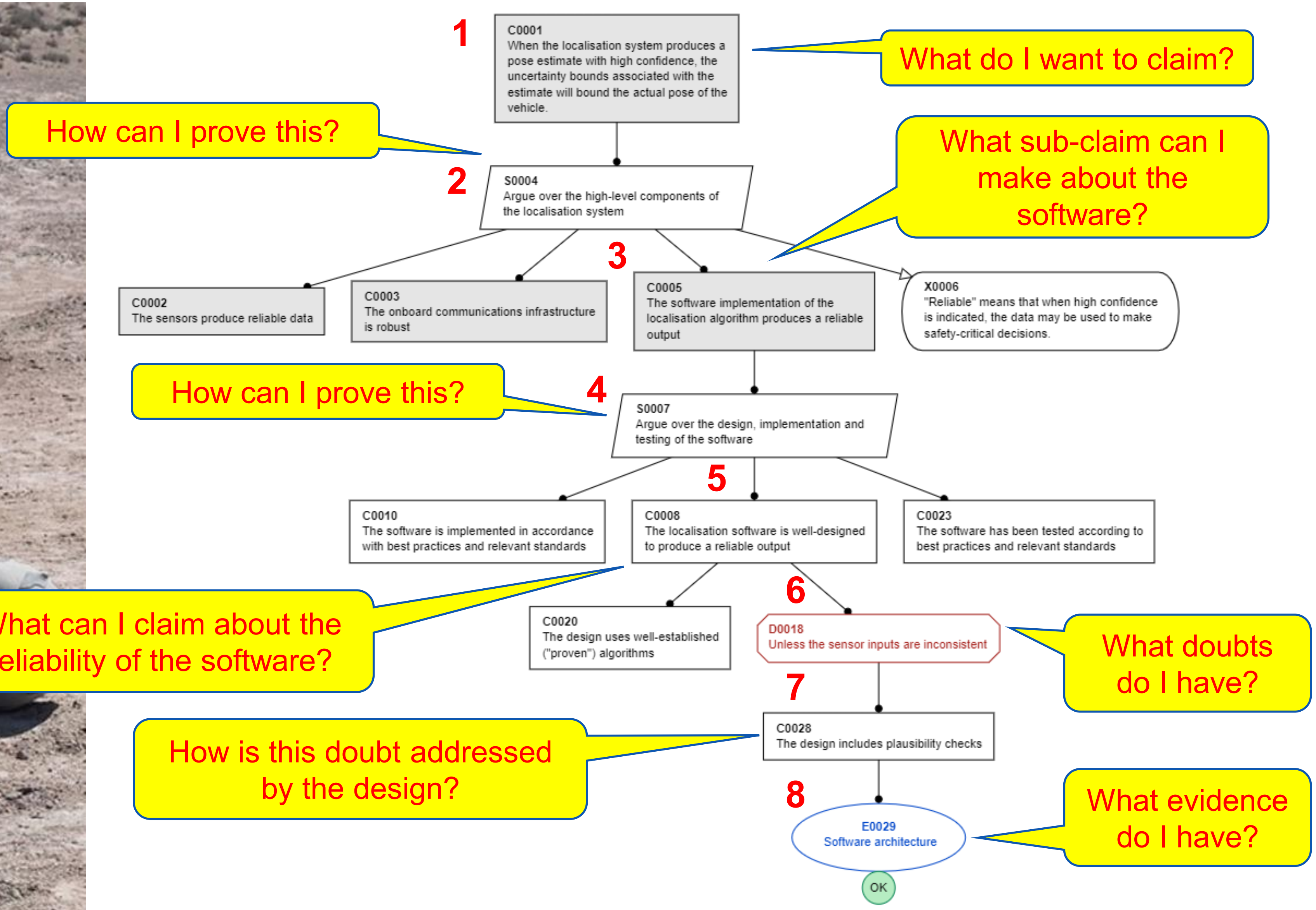
[**R**]ESIDUAL

OK

COMPLE[**T**]E

UND

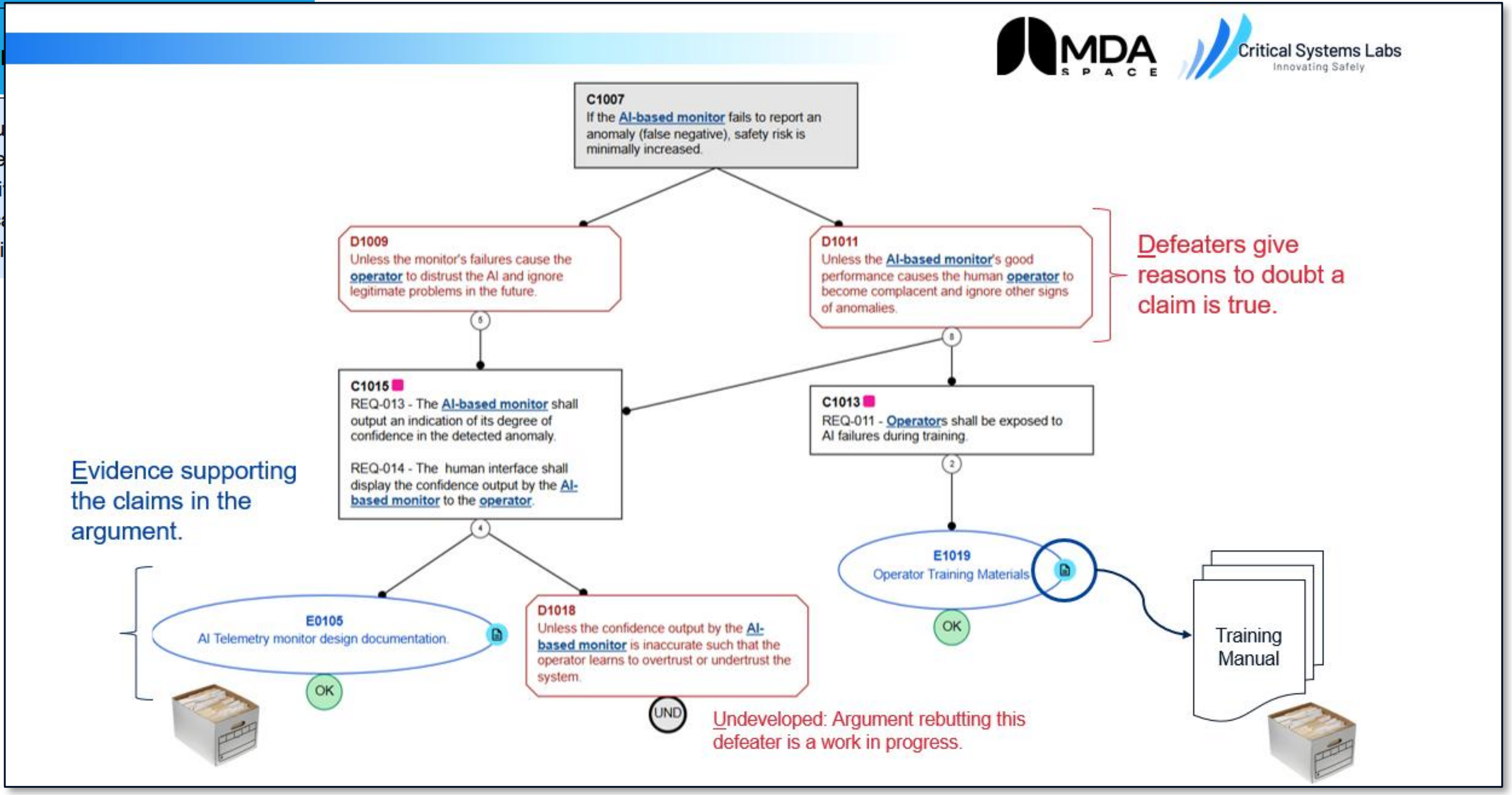
[**U**]NDEVELOPED


1
C0001
When the localisation system produces a pose estimate with high confidence, the uncertainty bounds associated with the estimate will bound the actual pose of the vehicle.
What do I want to claim?
How can I prove this?
2
S0004
Argue over the high-level components of the localisation system
What sub-claim can I make about the software?
C0002
The sensors produce reliable data
C0003
The onboard communications infrastructure is robust
3
C0005
The software implementation of the localisation algorithm produces a reliable output
X0006
"Reliable" means that when high confidence is indicated, the data may be used to make safety-critical decisions.
How can I prove this?
4
S0007
Argue over the design, implementation and testing of the software
5
C0010
The software is implemented in accordance with best practices and relevant standards
C0008
The localisation software is well-designed to produce a reliable output
C0023
The software has been tested according to best practices and relevant standards
What can I claim about the reliability of the software?
C0020
The design uses well-established ("proven") algorithms
6
D0018
Unless the sensor inputs are inconsistent
What doubts do I have?
7
C0028
The design includes plausibility checks
How is this doubt addressed by the design?
8
E0029
Software architecture
OK
What evidence do I have?

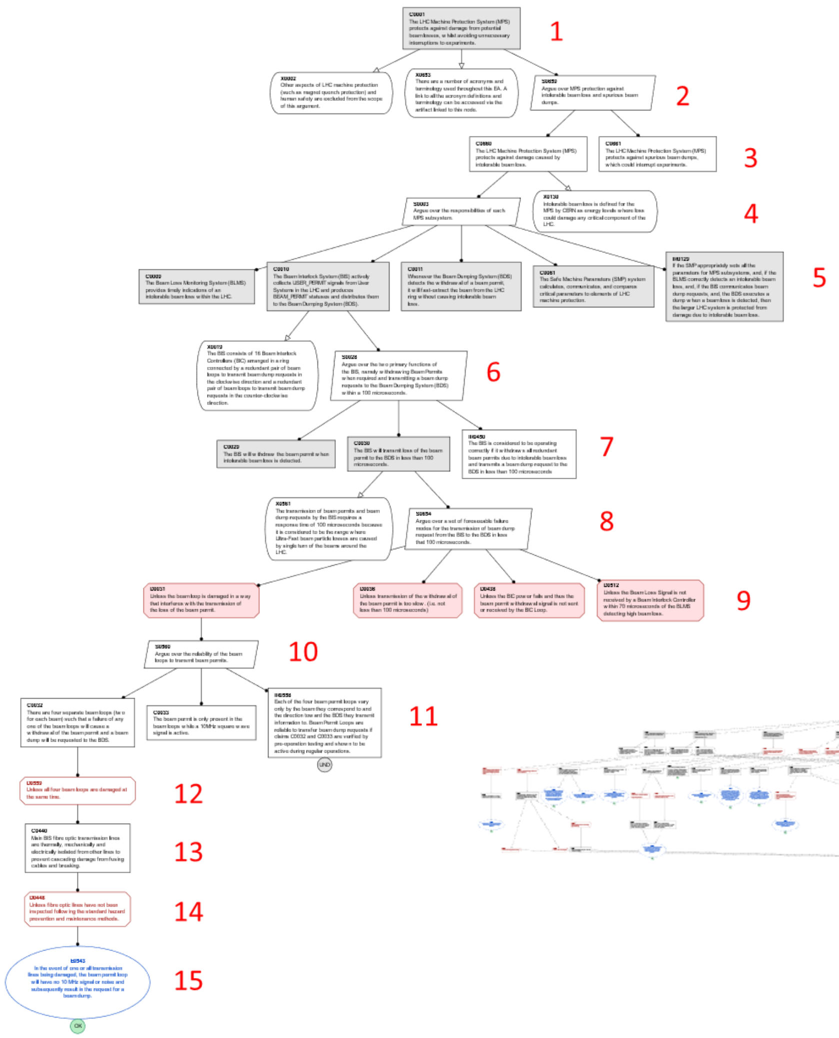
Critical Systems Labs
Innovating Safely

Yesterday in Session 3

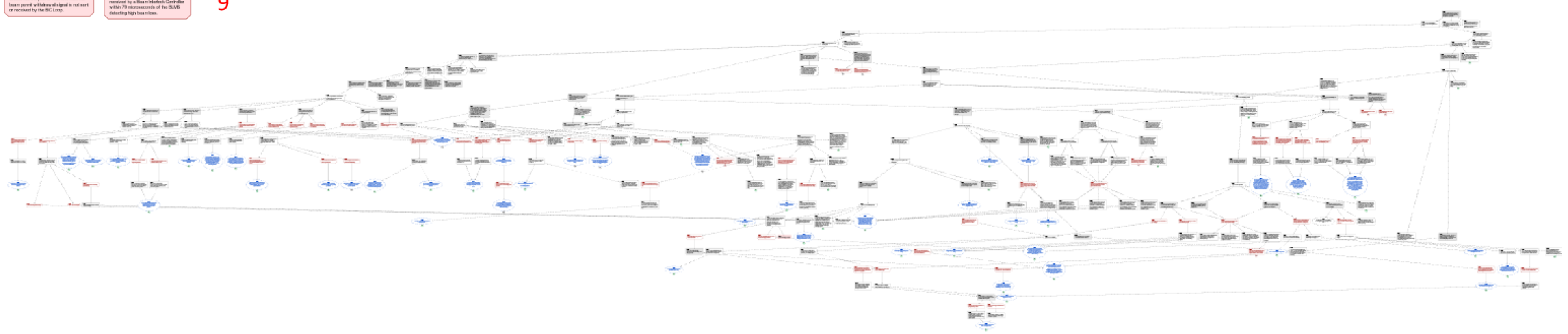
Session 3: Assurance & Artificial Intelligence			
Session 3	**14:00**	**14:10**	**Session O**
Session 3	15:40	16:00	Product Assu AI-based Te Health Moni Safety-Critic Roboti



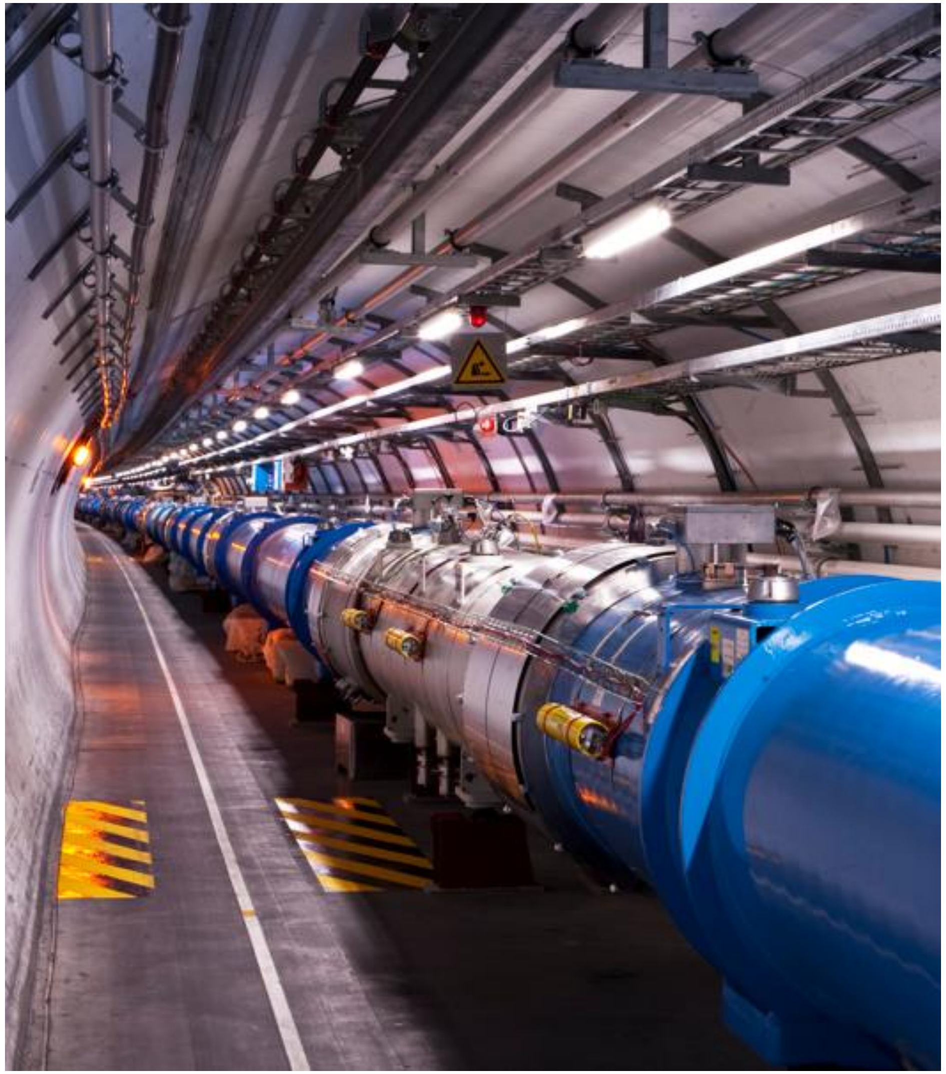
Scaling up to Complex Systems

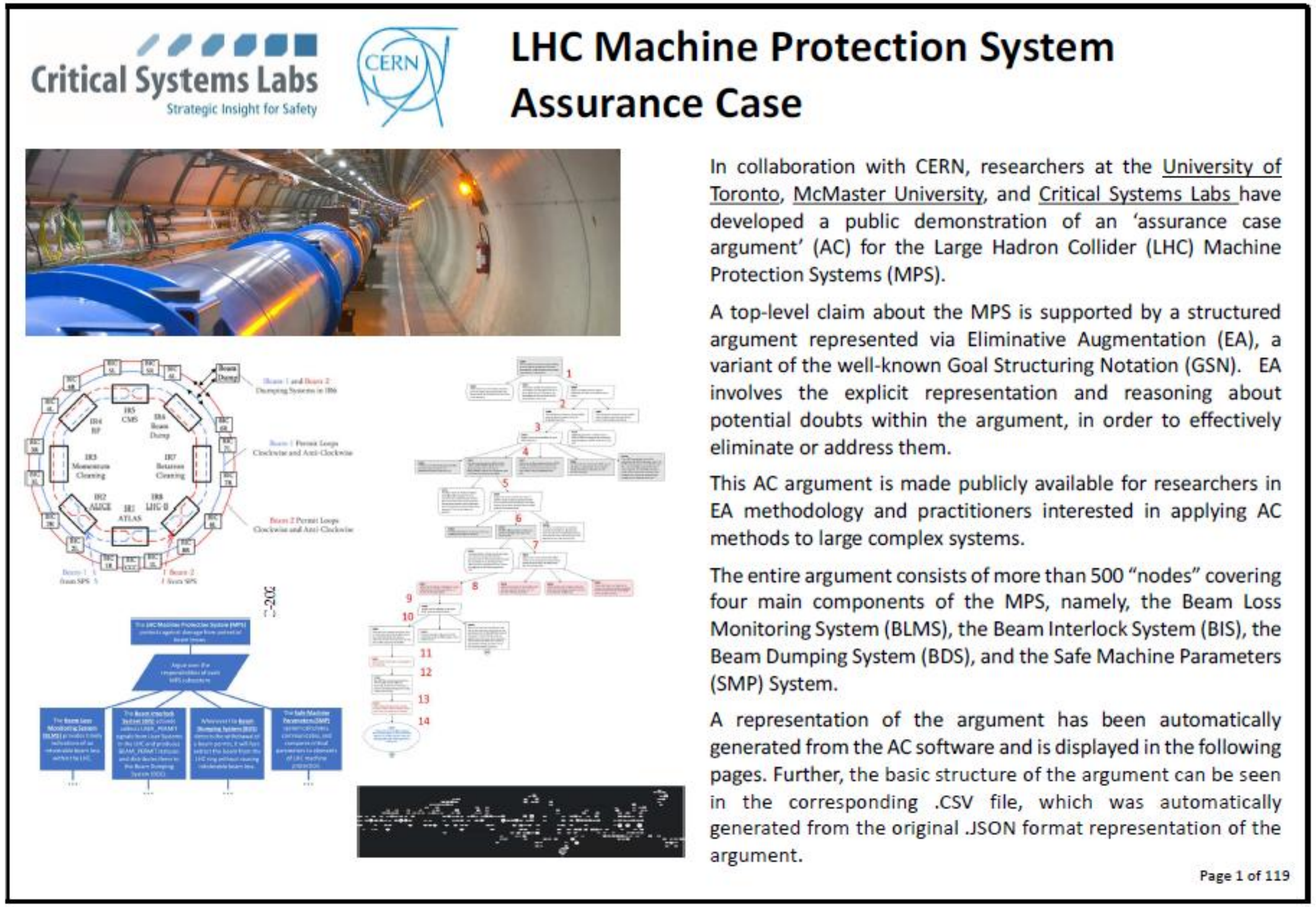


X 100 =



Real-world Example – CERN LHC MPS




Critical Systems Labs
Strategic Insight for Safety

CERN

LHC Machine Protection System Assurance Case

In collaboration with CERN, researchers at the University of Toronto, McMaster University, and Critical Systems Labs have developed a public demonstration of an 'assurance case argument' (AC) for the Large Hadron Collider (LHC) Machine Protection Systems (MPS).

A top-level claim about the MPS is supported by a structured argument represented via Eliminative Augmentation (EA), a variant of the well-known Goal Structuring Notation (GSN). EA involves the explicit representation and reasoning about potential doubts within the argument, in order to effectively eliminate or address them.

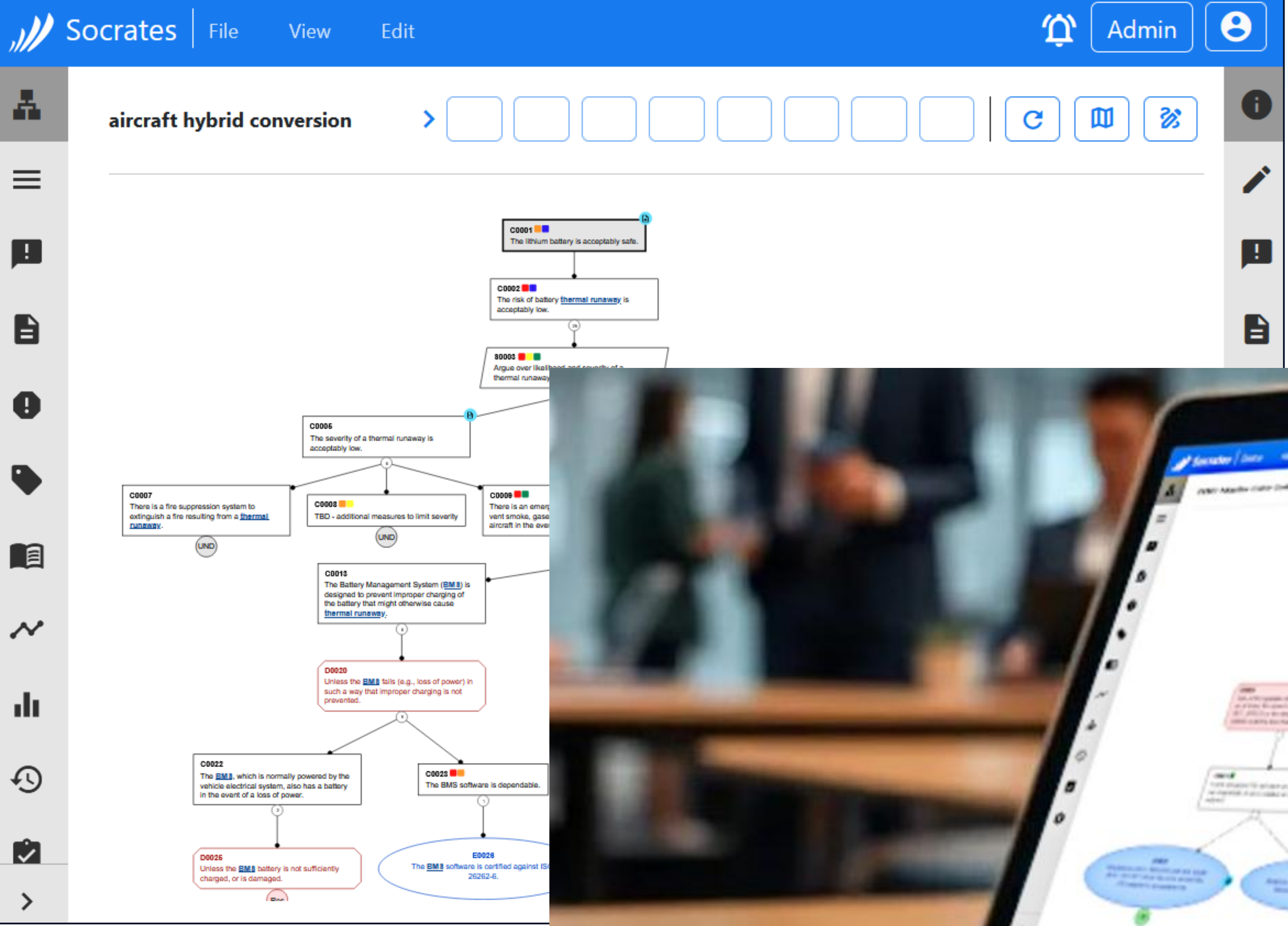
This AC argument is made publicly available for researchers in EA methodology and practitioners interested in applying AC methods to large complex systems.

The entire argument consists of more than 500 "nodes" covering four main components of the MPS, namely, the Beam Loss Monitoring System (BLMS), the Beam Interlock System (BIS), the Beam Dumping System (BDS), and the Safe Machine Parameters (SMP) System.

A representation of the argument has been automatically generated from the AC software and is displayed in the following pages. Further, the basic structure of the argument can be seen in the corresponding .CSV file, which was automatically generated from the original .JSON format representation of the argument.

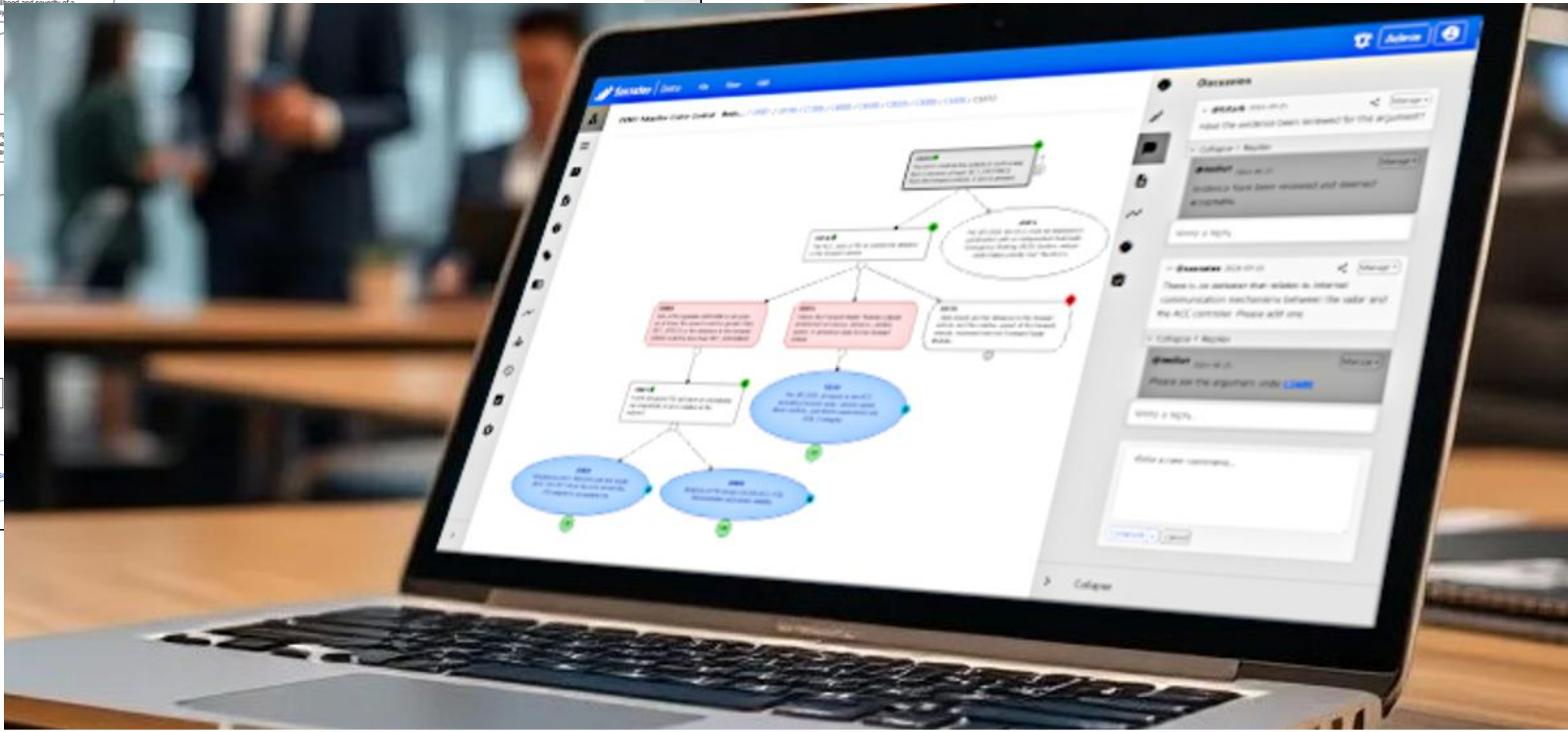
Page 1 of 119

https://cds.cern.ch/record/2854725



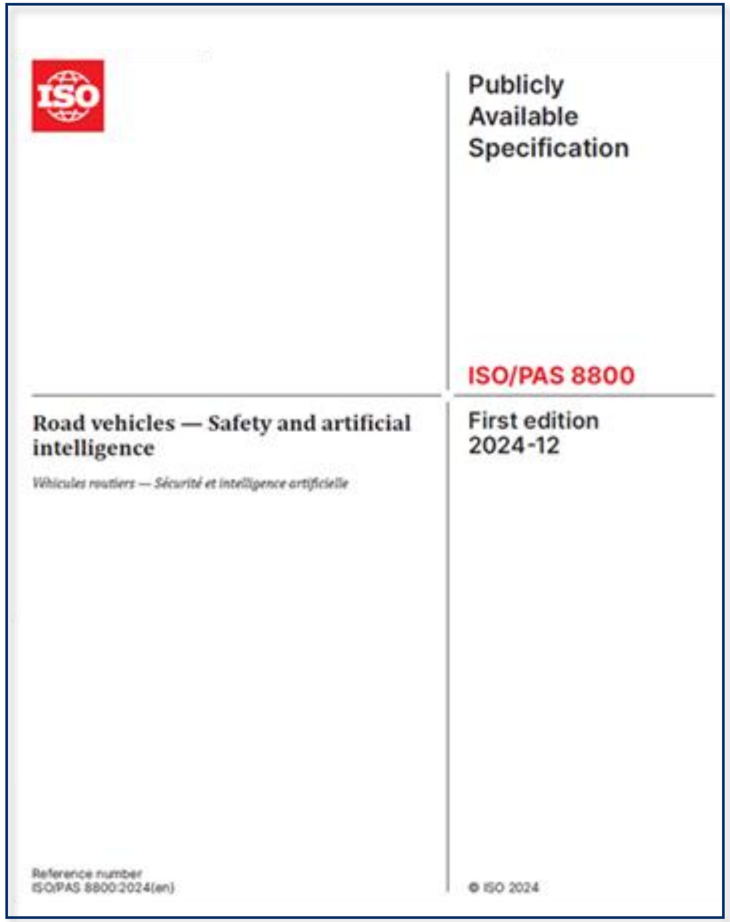
Critical Systems Labs
Innovating Safely

Socrates
Assurance Case Editor



Critical Systems Labs
Innovating Safely

Useful Across Multiple Technical Domains


Publicly Available Specification

ISO/PAS 8800

Road vehicles — Safety and artificial intelligence

Véhicules routiers — Sécurité et intelligence artificielle

First edition
2024-12

Reference number
ISO/PAS 8800:2024(en)

Applicable to “AI systems” with relatively little content unique to automotive, for example …

ISO/PAS 8800:2024(en)

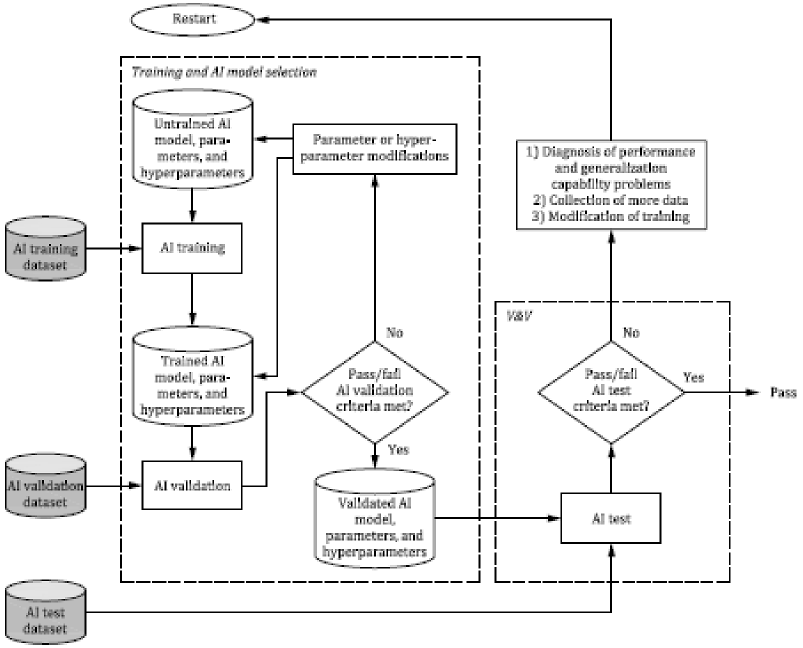
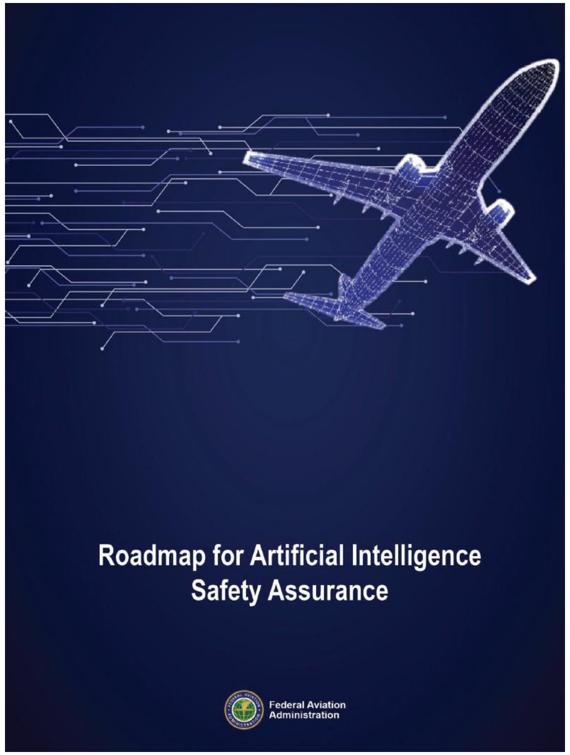


Figure 11-2 — An example supervised learning flow

Other High Assurance Technical Domains



Aerospace

EUROPEAN STANDARD
NORME EUROPÉENNE
EUROPÄISCHE NORM

EN 50716

November 2023

ICS 35.240.60

Supersedes EN 50128:2011; EN 50128:2011/AC:2014; EN 50657:2017; EN 50128:2011/A1:2020; EN 50128:2011/A2:2020; EN 50657:2017/A1:2023

English Version

Railway Applications - Requirements for software development

Applications ferroviaires - Exigences pour le développement de logiciels

Sektorübergreifende Software-Norm für Eisenbahnen

This European Standard was approved by CENELEC on 2023-10-30. CENELEC members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the CEN-CENELEC Management Centre or to any CENELEC member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CENELEC member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

CENELEC members are the national electrotechnical committees of Austria, Belgium, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Republic of North Macedonia, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Türkiye and the United Kingdom.

CENELEC
European Committee for Electrotechnical Standardization
Comité Européen de Normalisation Electrotechnique
Europäisches Komitee für Elektrotechnische Normung

CEN-CENELEC Management Centre: Rue de la Science 23, B-1040 Brussels

Ref. No. EN 50716:2023 E

Rail

Nuclear

Conclusion

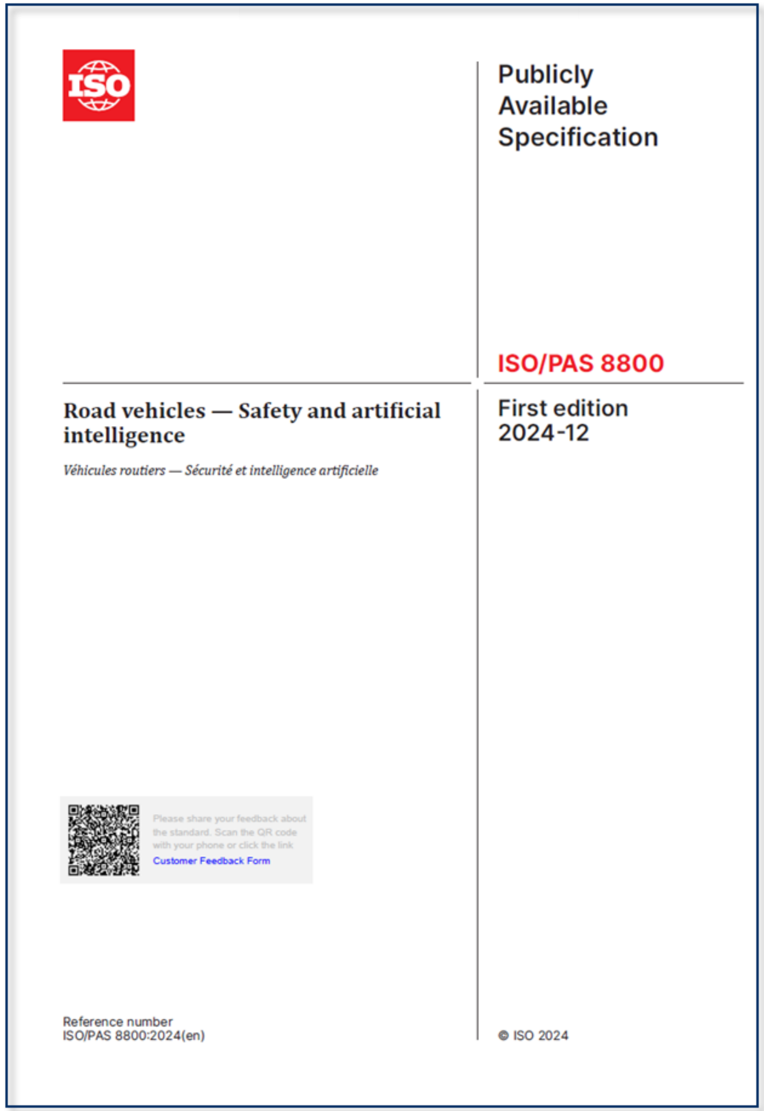
Dictionary

piv·ot·al

/ˈpivədl/

adjective

of crucial importance in relation to the development or success of something else.


Publicly Available Specification

ISO/PAS 8800

Road vehicles — Safety and artificial intelligence

Véhicules routiers — Sécurité et intelligence artificielle

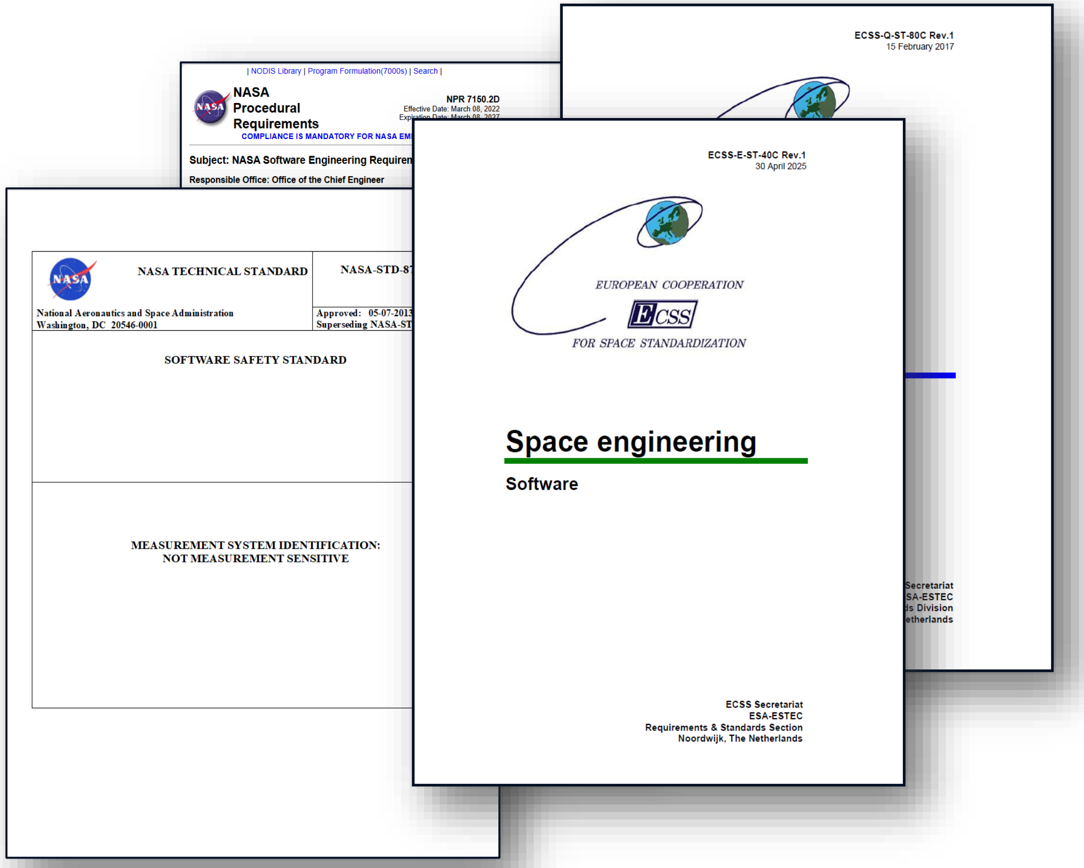
First edition
2024-12

Please share your feedback about the standard. Scan the QR code with your phone or click the link
Customer Feedback Form

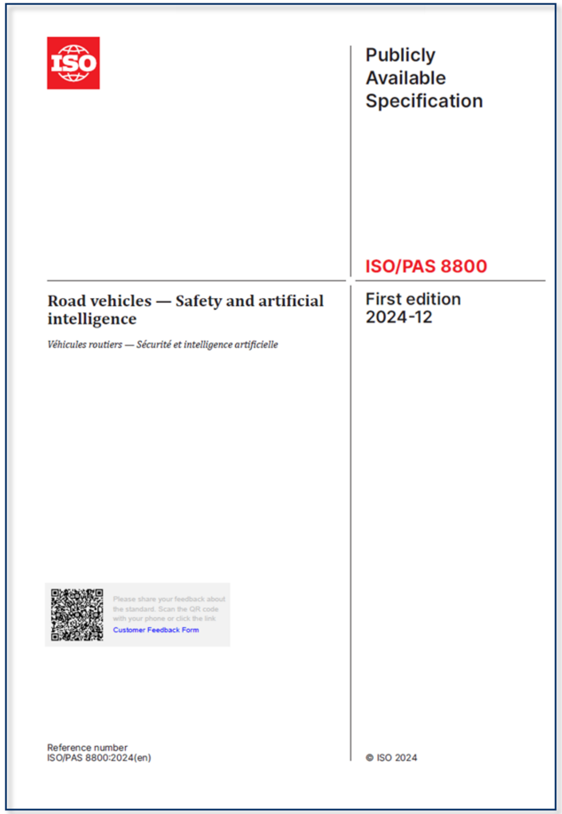
Reference number
ISO/PAS 8800:2024(en)

© ISO 2024

Recommendation



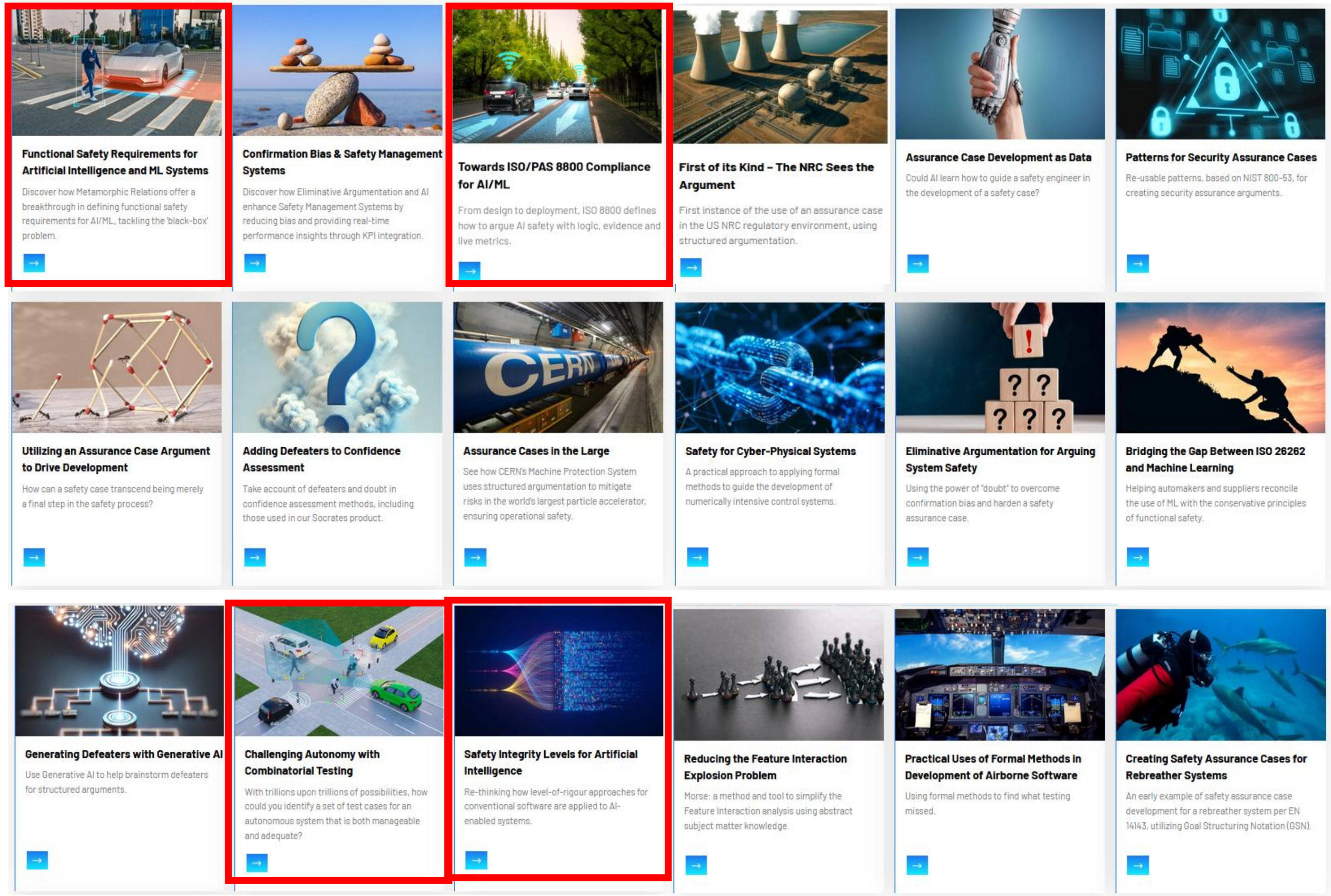
+



= interim solution

Critical Systems Labs
Innovating Safely

For more on ISO/PAS 8800 and AI Safety



Publications and whitepapers available on the "perspectives" page of our website at www.cslabs.com

Critical Systems Labs

Innovating Safely

Jeff Joyce

jeff.joyce@cslabs.com