FIRST SWISS

Functional Safety and Cyber Security Certification Body

Functional Safety

OT & IT Cyber Security

Artificial Intelligence

CERTX

CERTX

Unconditional reliability where it is needed

We are:

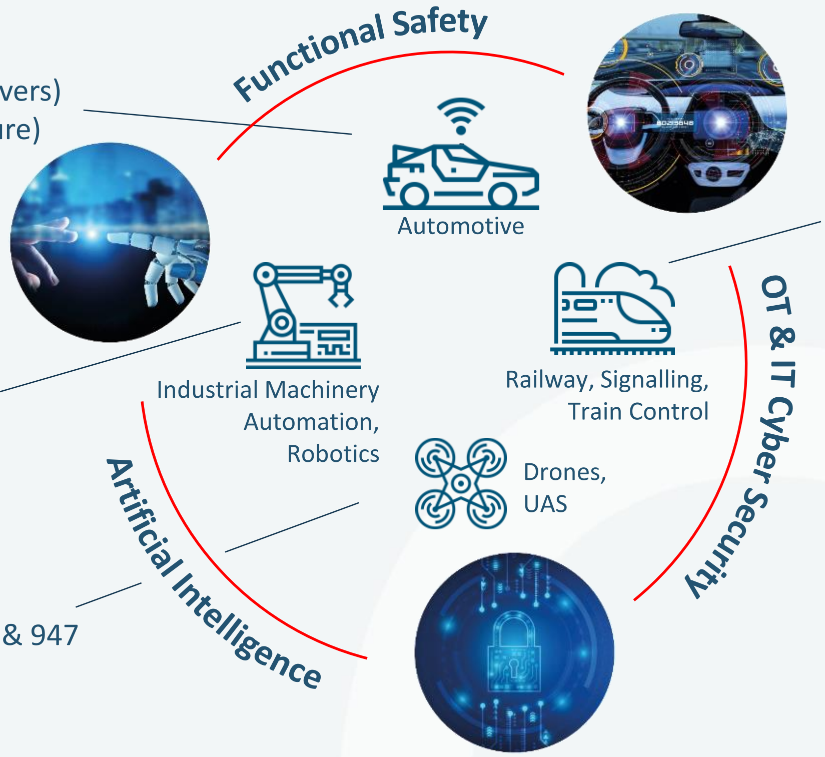
SWISS ACCREDITATION
sas.admin.ch SCESp 0129

A Certification Body for Functional Safety and Cyber Security accredited by Swiss Accreditation Service (SAS) with international validity

Experts in Functional Safety, Artificial Intelligence and Cybersecurity with Swiss DNA and Quality. Innovations at heart, pragmatic in style.

Co-authors of the standards for future automated systems, autonomous mobility, artificial intelligence, and cyber security

Expertise in all Safety and Security Critical Domains



ISO 26262
ISO 21448 (SOTIF)
ISO 19014 (Earthmovers)
ISO 25119 (Agriculture)
ASPICE
UN-R 157

EC 2006/42
CE

EN ISO 13849
EN IEC 61508
EN IEC 61496
EN IEC 62061
EN ISO 61800

EC 2019/945 & 947
EN 4709

EN 50126 / 50128 / 50129
EC 2009/402 (CSM)
cen EN 50657

ISA/IEC 62443
ISO 21434
UN-R 155
UN-R 156

ISO 27001/2
GDPR

....and general quality QMS: ISO 9001

CertX Services

TRAINING & CERTIFICATION of ENGINEERS and MANAGERS

INSPECTIONS of products and processes for e.g.:

- ISO 26262 confirmation reviews, assessments
- ISA (Independent Safety Assessments) for railway

CERTIFICATION of PRODUCTS

incl. EU machine directive & CE **Conformity** (Notified body No. 2948)

Test services for the homologation of cars in the EU & CH

CERTIFY CORPORATE ORGANIZATIONS and PROCESSES

CertX provides comprehensive safety and security services

4 Challenges

The Challenge of Automation: Functional Safety of Electronic Controls

Effects of a systematic S/W failure of the fuel injection pump:

Toyota "Unintended Acceleration" Has Killed 89

A 2005 Toyota Prius, which was in an accident, is seen at a police station in Harrison, New York, Wednesday,

Source: The ASSOCIATED PRESS

… Toyota had to recall 10M cars for repair and paid 1.6Bn$ of compensation for damages.

only noticed if it is too late – because it is missing

The Challenge of Connectivity: Operational Cybersecurity

Hijacking cars

Source: Wired

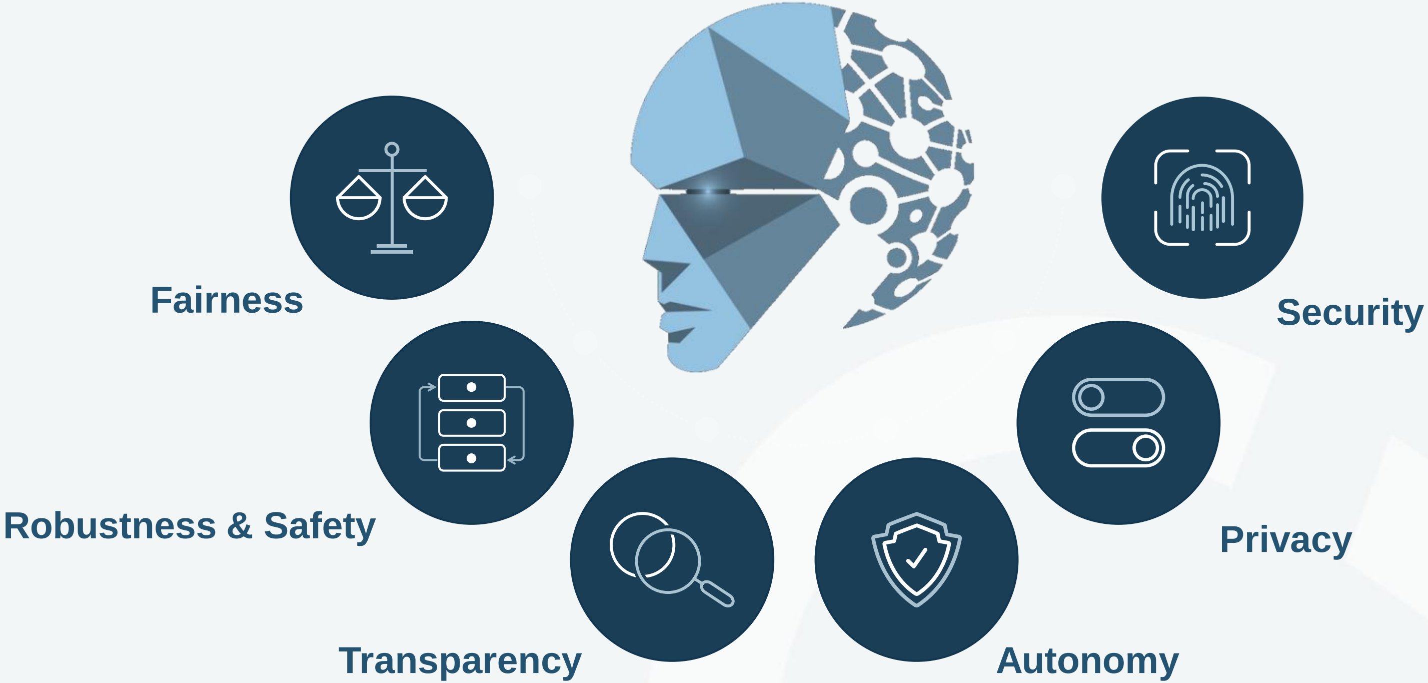
Chrysler had to recall 1.4 million vehicles after a pair of hackers demonstrated that they could remotely hijack a Jeep's digital systems over the Internet.

A 19-year-old German teen claimed hacked into dozens of Teslas worldwide early 2022.
The security flaw can unlock doors, honk horns and start the cars

```
$ ssh tesla1@cid
tesla1@cid$ sudo bash
root@cid# ./carkill.sh
root@cid#
```

may compromise (public) safety with serious consequences

AI Applications to be verified in 6 dimensions



Fairness

Robustness & Safety

Transparency

Autonomy

Privacy

Security

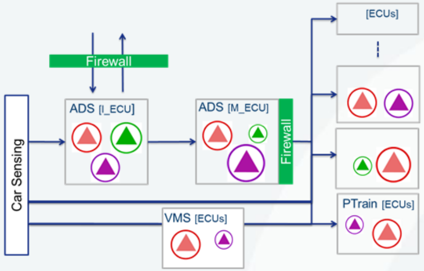
Last not Least: Performance

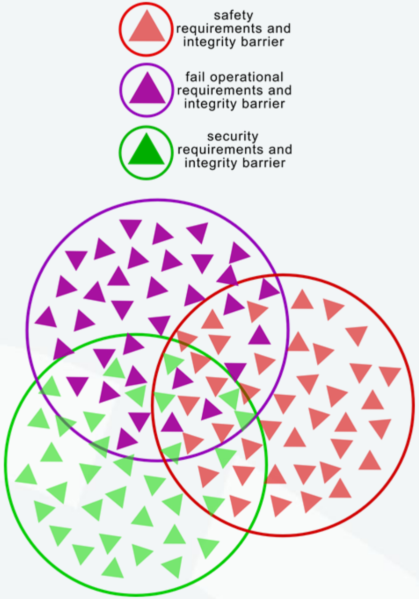
ADS systems must be at the same time:

- Fail Safe: In case of failure, systems can achieve the safe state. Most systems are Fail-safe.
- Fail operational: in case of failure, systems remain operational.

Challenges:

- How to ensure Fail-safe and Fail operational ?
- Normally fail-safe and fail operational goals are in contradiction
- The Safety Mechanism has a different goal – detecting fault and react. It does not cover the fail operational aspect
- for fail operational we need full or nearly full functionality in the case of a fault.

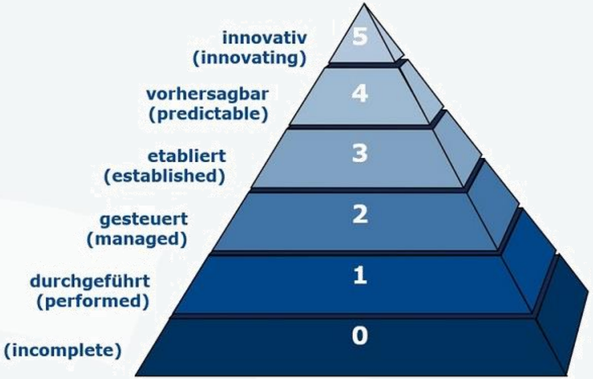



safety requirements and integrity barrier
fail operational requirements and integrity barrier
security requirements and integrity barrier

4 Steps to a successful integration of functional safety, cybersecurity, SOTIF, and AI

Step 1: Complex development: Reliable Foundation Processes

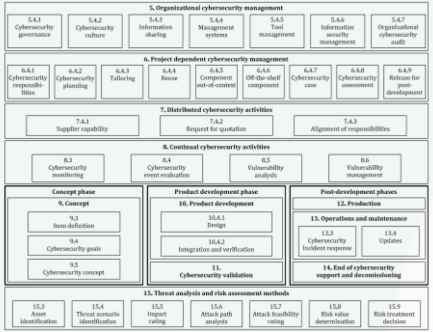
- The cyber-physical systems for autonomous mobility are one of the most complex.
- The development requires stable processes for success
- A thoughtfully implemented V-cycle development process is the basis for the integration challenge
- As a first step for a development process of an autonomous system development, an assessment of processes e.g. according to ASPICE or CMMI can be a helpful move:
    - Assessment of all crucial processes for maturity
    - Helps to identify if “living” processes are in place, rather than dispensable descriptions that are not applied



The VDA maturity pyramid

Step 2: Ensure safety and security

- Functional safety acc. to ISO 26262 and automotive security acc. to ISO 21434 are both based on the V-model and can be mapped to the ASPICE processes (see e.g., ASPICE CS for cybersecurity aspects)
- Companies often hesitate to set up an integrated development process that can cover safety- and non-safety and security developments
- However, with more standards to be followed, process complexity has to be minimized by joint processes that can realistically be followed by the development teams



5. Organizational cybersecurity management
5.4.1 Cybersecurity governance
5.4.2 Cybersecurity culture
5.4.3 Information sharing
5.4.4 Management systems
5.4.5 Tool management
5.4.6 Information security management
5.4.7 Organizational cybersecurity audit

6. Project dependent cybersecurity management
6.4.1 Cybersecurity responsibilities
6.4.2 Cybersecurity planning
6.4.3 Tailoring
6.4.4 Reuse
6.4.5 Component out-of-context
6.4.6 Off-the-shelf component
6.4.7 Cybersecurity case
6.4.8 Cybersecurity assessment
6.4.9 Release for post-development

7. Distributed cybersecurity activities
7.4.1 Supplier capability
7.4.2 Request for quotation
7.4.3 Alignment of responsibilities

8. Continual cybersecurity activities
8.3 Cybersecurity monitoring
8.4 Cybersecurity event evaluation
8.5 Vulnerability analysis
8.6 Vulnerability management

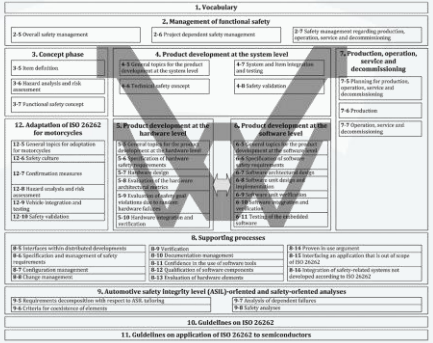
Concept phase
9. Concept
9.3 Item definition
9.4 Cybersecurity goals
9.5 Cybersecurity concept

Product development phase
10. Product development
10.4.1 Design
10.4.2 Integration and verification
11. Cybersecurity validation

Post-development phases
12. Production
13. Operations and maintenance
13.3 Cybersecurity incident response
13.4 Updates
14. End of cybersecurity support and decomissioning

15. Threat analysis and risk assessment methods
15.3 Asset identification
15.4 Threat scenario identification
15.5 Impact rating
15.6 Attack path analysis
15.7 Attack feasibility rating
15.8 Risk value determination
15.9 Risk treatment decision

ISO 21434



1. Vocabulary

2. Management of functional safety
2-5 Overall safety management
2-6 Project dependent safety management
2-7 Safety management regarding production, operation, service and decommissioning

3. Concept phase
3-5 Item definition
3-6 Hazard analysis and risk assessment
3-7 Functional safety concept

4. Product development at the system level
4-5 General topics for the product development at the system level
4-6 Technical safety concept
4-7 System and item integration and testing
4-8 Safety validation

7. Production, operation, service and decommissioning
7-5 Planning for production, operation, service and decommissioning
7-6 Production
7-7 Operation, service and decommissioning

12. Adaptation of ISO 26262 for motorcycles
12-5 General topics for adaptation for motorcycles
12-6 Safety culture
12-7 Confirmation measures
12-8 Hazard analysis and risk assessment
12-9 Vehicle integration and testing
12-10 Safety validation

5. Product development at the hardware level
5-5 General topics for the product development at the hardware level
5-6 Specification of hardware safety requirements
5-7 Hardware design
5-8 Evaluation of the hardware architectural metrics
5-9 Evaluation of safety goal violations due to random hardware failures
5-10 Hardware integration and verification

6. Product development at the software level
6-5 General topics for the product development at the software level
6-6 Specification of software safety requirements
6-7 Software architectural design
6-8 Software unit design and implementation
6-9 Software unit verification
6-10 Software integration and verification
6-11 Testing of the embedded software

8. Supporting processes
8-5 Interfaces within distributed developments
8-6 Specification and management of safety requirements
8-7 Configuration management
8-8 Change management
8-9 Verification
8-10 Documentation management
8-11 Confidence in the use of software tools
8-12 Qualification of software components
8-13 Evaluation of hardware elements
8-14 Proven in use argument
8-15 Interfacing an application that is out of scope of ISO 26262
8-16 Integration of safety-related systems not developed according to ISO 26262

9. Automotive safety integrity level (ASIL)-oriented and safety-oriented analyses
9-5 Requirements decomposition with respect to ASIL tailoring
9-6 Criteria for coexistence of elements
9-7 Analysis of dependent failures
9-8 Safety analyses

10. Guidelines on ISO 26262

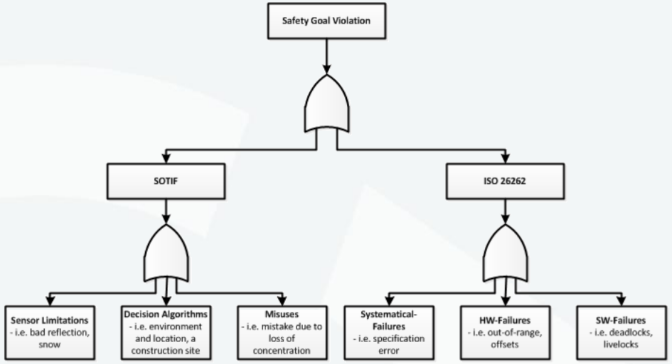
11. Guidelines on application of ISO 26262 to semiconductors

ISO 26262

Step 3: ISO 21448 (SOTIF) as precondition for AI development

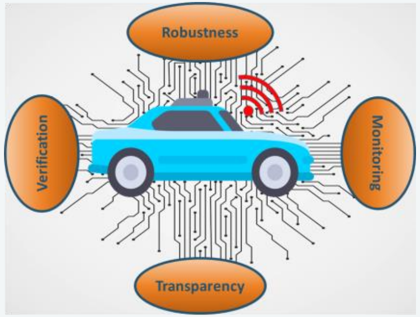
- With the publication of ISO 21448 last year, another process is available to provide guidance for autonomous vehicle developments.
- SOTIF can be categorized as system level standard, focusing on performance- and completeness requirements for advanced driver assistance- and autonomous systems

- Even if general risks caused by using AI are considered in the standard (e.g., necessary awareness of training data bias), no further guidance is provided for the implementation of AI
- For safe AI developments, national standards like VDE AR-2842-61-3 can provide guidance, however a recognized international standard is still missing.

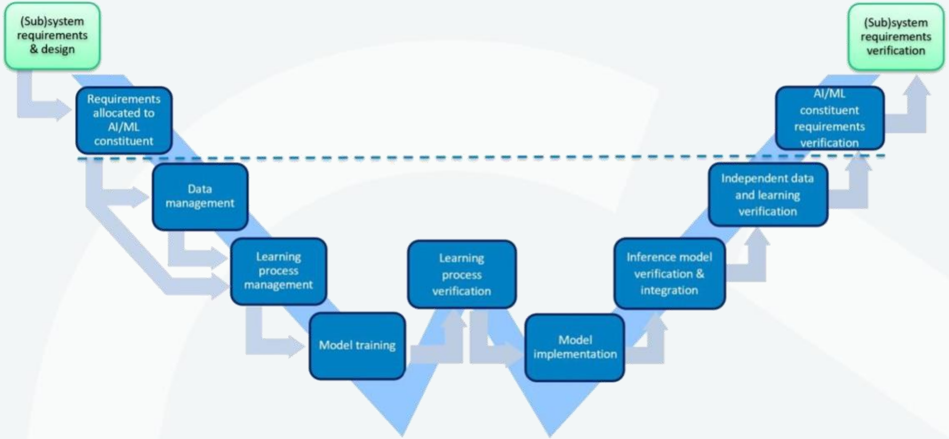

Safety Goal Violation
SOTIF
ISO 26262
Sensor Limitations - i.e. bad reflection, snow
Decision Algorithms - i.e. environment and location, a construction site
Misuses - i.e. mistake due to loss of concentration
Systematical-Failures - i.e. specification error
HW-Failures - i.e. out-of-range, offsets
SW-Failures - i.e. deadlocks, livelocks

Step 4: AI integration - from a V to W

- AI systems are treated as system element in a vehicle level context and are handled as black box (to some extent)
- The top-level SOTIF efforts to evaluate system-level reliability are sometimes criticised as “mile collecting”, putting too much emphasis on system reliability as a statistical parameter
- The development of a process to reach trustworthy AI must consider more parameters and is therefore handled in a separate verification/validation loop

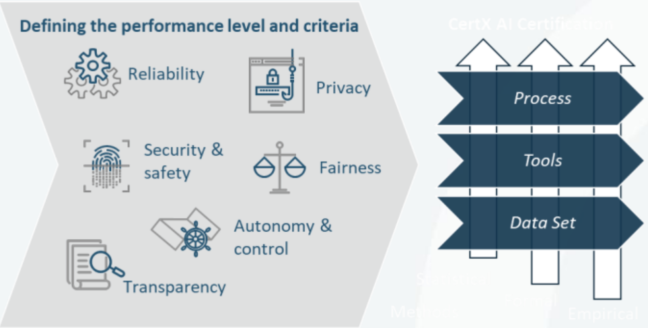


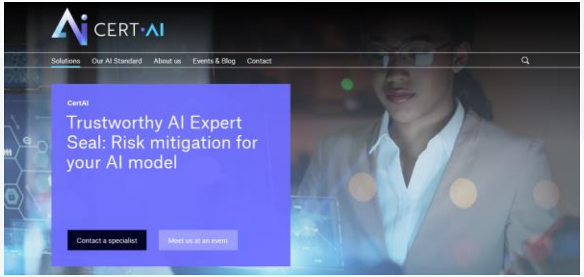
Around **40%** of the methods presented in **ISO-26262-6** do not apply to the Machine Learning (ML) models and algorithms.


(Sub)system requirements & design
Requirements allocated to AI/ML constituent
Data management
Learning process management
Model training
Learning process verification
Model implementation
Inference model verification & integration
Independent data and learning verification
AI/ML constituent requirements verification
(Sub)system requirements verification

Specific risks related to AI systems

- Like for software- and hardware development, system engineering activities have to define requirements for AI systems, e.g., Intended purpose, operational situations and performance
- Those outputs have to be defined following the approach of ISO 21448 (SOTIF) on system level
- Today, AI for safety critical applications is still equipped with a deterministic “guiderail”, limiting the effect of possible failures
- Schemes to evaluate trustworthiness of AI are currently being rolled out e.g., by CertAI





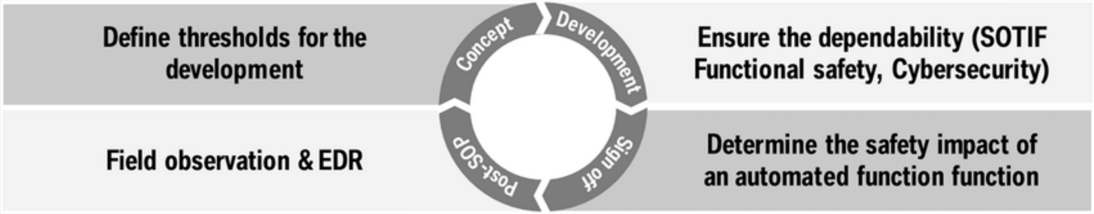
CertAI

Trustworthy AI Expert Seal: Risk mitigation for your AI model

Check out: www.CertAI.com

Positive Risk Balance: A constant trade-off

- No matter if an integrated process is defined from scratch or if it is integrated to an existing process landscape – it needs to be ensured, that the product must meet its performance targets while security and safety is a necessary "must".
- It can be helpful to keep a dependability - or trustworthiness model in mind:
  - Safety can jeopardize system availability (SOTIF), which becomes more critical when systems are intended to operate fully autonomous
  - Known contradictions between security and safety create the need to plan updateability of software components and limit re-verification effort over the lifecycle
  - Security- and safety features can decelerate communication and hence compromise performance
  - AI, if trained incorrectly, can generate failures during operation that cannot be detected during testing


Define thresholds for the development
Concept
Development
Ensure the dependability (SOTIF Functional safety, Cybersecurity)
Field observation & EDR
Post-SOP
Sign off
Determine the safety impact of an automated function function

From: Positive risk balance: a comprehensive framework to ensure vehicle safety,
Nina Kaufmann, Felix Fahrenkrog, Ludwig Drees and Florian Raisch, Springer Verlag, 2022

AWARD Horizon 2020
Real World Application Coming Soon

www.certx.com

Warehouse

Hub-to-hub

Airport

Port

This project has received funding from the European Union's Horizon 2020 research and innovation
programme under grant agreement No 101006817.

The content of this presentation reflects only the author's view. Neither the European Commission nor the CINEA is responsible for any use that may be made of the information it contains.

Putting it together

- A landscape of shattered development processes can hardly reach the required overall system optimum
- Only a thoughtful integration of safety-, security-, SOTIF, and machine learning development processes into an overall V-model system engineering approach can result in the intended vehicle properties
- Sufficient time needs to assumed for process roll-out and staff adaption
  - The early involvement of experts can prevent companies from common drawbacks

CERTX

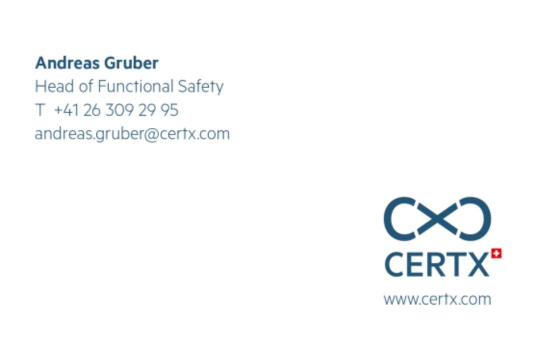
We are your safety and security partner

We look at your product or process

- with an **unbiased** and **independent** mindset,
- a **safety and security** focus,
- with a **unique skill and experience** from all domains
- in a **systematic**, holistic fashion and
- in **confidentiality**

Your Contact Points at CertX

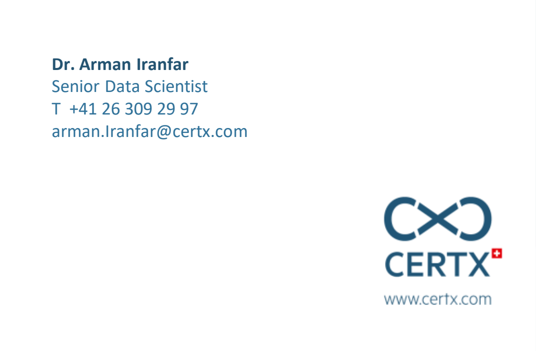
- Swiss chair of ISO 26262 standardization working group (edition 3)
- Swiss chair of ISO 21448 standardization Working group (SOTIF)
- Personal and Standardized Certification: ISO26262 (ISCN/ECQA)
- In-depth experience in automotive safety engineering and safety management process modeling
- Former product safety officer for a tier 1



Andreas Gruber
Head of Functional Safety
T +41 26 309 29 95
andreas.gruber@certx.com

www.certx.com

- Swiss delegate in ISO/IEC JTC1/SC42 - Artificial Intelligence
- In-depth experience in AI assessment and data governance
- Expert in AI reliability, transparency, security, and safety
- Former senior researcher at EPFL developing AI solutions for H2020 European projects (DeepHealth, MANGO)



Dr. Arman Iranfar
Senior Data Scientist
T +41 26 309 29 97
arman.Iranfar@certx.com

www.certx.com

The information contained in this presentation is the property of CertX AG.

This presentation and extracts thereof may only be duplicated or forwarded to third parties following explicit written approval by CertX AG.

All product names used in this documentation are trademarks or otherwise protected by law, even if not specifically indicated.

Passage Du Cardinal 13b
1700 Fribourg, Switzerland

www.certx.com