Considerations on safety assurance for AI

Dr David Ward, Global Head of Functional Safety

2025

What is “assurance”?

- ISO/IEC/IEEE 15026-1:2019, “Systems and software engineering — Systems and software assurance”
- Assurance: grounds for justified confidence that a claim has been or will be achieved
- Assurance case: reasoned, auditable artefact created that supports the contention that its top-level claim (or set of claims) is satisfied, including systematic [arguments] and its underlying evidence and explicit assumptions that support the claim(s)
  - Note 1 to entry: An assurance case contains the following and their relationships:
    - one or more claims about properties
    - arguments that logically link the evidence and any assumptions to the claim(s)
    - a body of evidence and possibly assumptions supporting these arguments for the claim(s)
    - justification of the choice of top-level claim and the method of reasoning

What do standards require?

- Standards for functional safety and cybersecurity specify a number of activities to give assurance in the robustness of the product, including independent reviews
- These activities include the following (but precise application varies by standard)
  - Verification reviews
  - Confirmation reviews
  - Validation
  - Assurance case
  - Management system (engineering process) audit
  - Product audit
  - Product assessment
  - Release for production

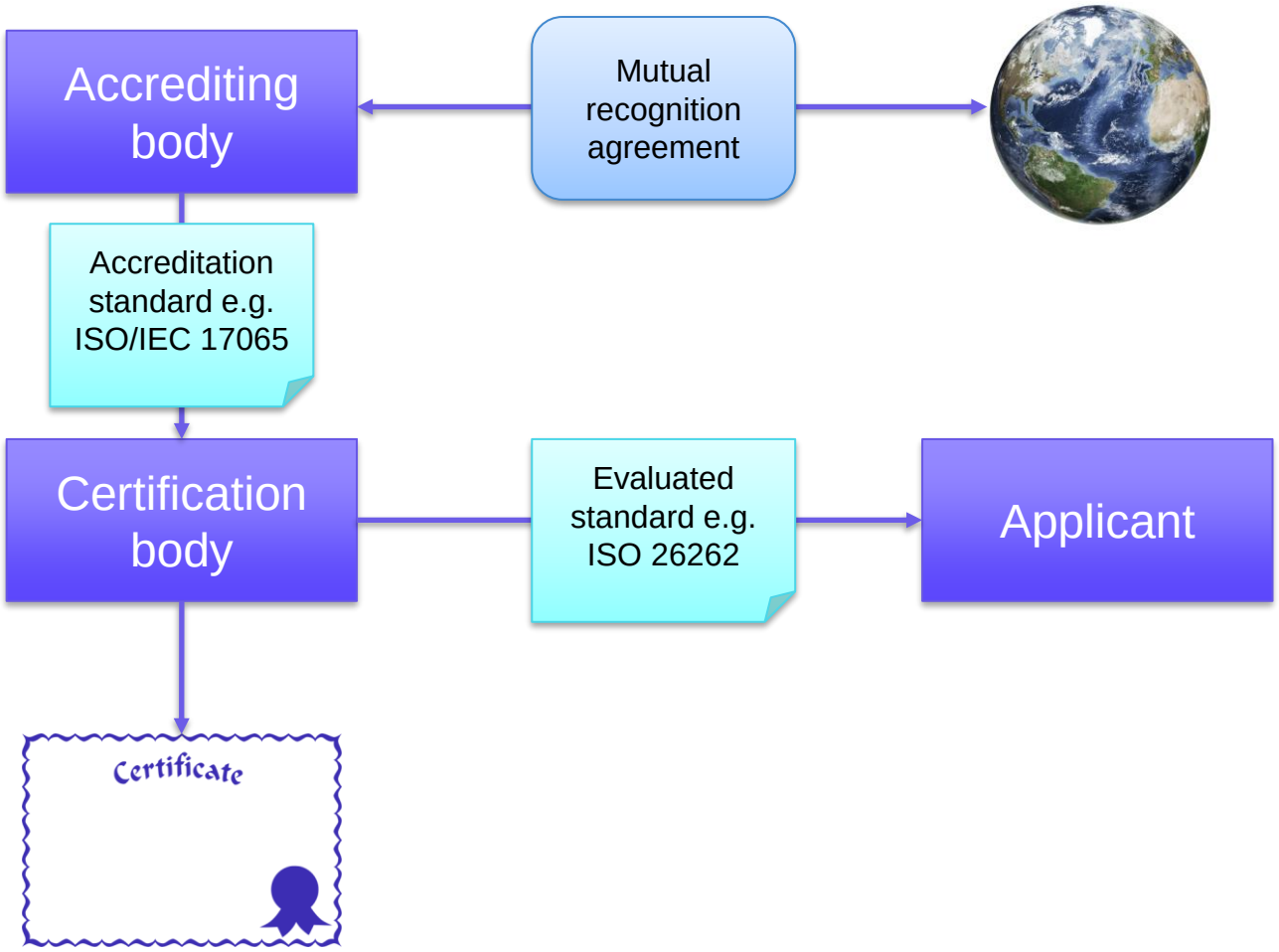
Application to AI – ISO/PAS 8800

- Clause 7 – AI safety management
  - Calls up ISO 26262 Part 2 including the confirmation measures
    - Confirmation review of safety plan, AI system validation report, AI safety analyses, AI safety assurance argument
    - AI safety audit and AI safety assessment
- Clause 8 – assurance arguments for AI systems
  - Seen as part of the ISO 26262 safety case
  - Objectives
    - Demonstrate the safety requirements allocated to the AI system are fulfilled
    - Evaluate whether the assurance argument reflects the actual residual risk of the AI system violating its safety requirements
  - Significant content around evidence and structure of the assurance argument

What about certification?

- Certification is a more formal activity of declaring that a product or process is in conformity with something, usually a standard
- Certification is not a requirement of ISO 26262 etc. and they are not “certification” standards that have to be met in the same formal way as e.g. DO-178C
- “Anyone” can issue a certificate, but the gold standard is for these to be accredited by a national body (e.g. UKAS in the UK, DAkkS in DE)
- Accreditation brings impartiality requirements that are much more onerous than “I3” independence in ISO 26262
  - In particular a certification body cannot provide conformity assessment on a product or process it (or its related bodies) has also provided consultancy on

Certification relationships



Care points with certification

- ISO 26262, ISO/SAE 21434, etc. are engineering standards and certification against them is not mandatory
  - Any certification schemes are commercial products of the bodies offering them, are not “official” and do not follow a common approach
- Certification can be seen as attractive in the supply chain so that providers of components can demonstrate that a competent independent party has reviewed a product or process
- A certificate does not guarantee a product is correct for its application
- A certificate does not remove the need to follow a safety manual, integrator guidance, etc.
- A certificate does not guarantee a product is free from issues, nor grant immunity from marketplace actions

Questions to ask a Certification Body

- A certification body will have a “scheme document” that is required to be public and should be made available on request
- Who is the accrediting body and which accreditation standard has been used?
  - The appropriate standards are
    - ISO/IEC 17021 – for xSMS audit alone
    - ISO/IEC 17065 – for product certification (also covers process audit)
- How is impartiality managed?
- What are the surveillance requirements?
- What is the scope of a certificate? Also ask to see the associated report
  - What Parts or Clauses of a standard have been examined?
  - What process or product variants does it apply to?
  - What are the obligations on the end user? e.g. requirements in a safety manual

Dr David Ward
MA PhD CEng CPhys MInstP MSAE
Global Head of Functional Safety

D +44 24 7635 5430 **T** +44 24 7635 5000
E david.ward@horiba-mira.com

HORIBA MIRA Ltd.
Watling Street
Nuneaton
Warwickshire
CV10 0TU
United Kingdom

www.horiba-mira.com