Observatoire ARGA

INTERNATIONAL ANALYTICAL REPORT

Study No. 17

International Standards for Conformity Assessment, Certification and Accreditation

AI Assurance: Reliability of Evidence, Accreditation Status and Limits of Mutual Recognition

Organisation: Observatoire ARGA

English

21 rue de l’Aviation, 64600 Anglet, France

info@argaobservatory.org | www.argaobservatory.org

21 October 2026

Contents

Annotations — 3
Scientific grounds and research design — 4
1. Scope of AI Assurance and the research question — 6
2. Terminology and levels of assessment — 8
3. The object of assurance and the claims map — 10
4. The evidence architecture of AI Assurance — 13
5. Institutional quality infrastructure — 15
6. ISO/IEC 42001:2023: AI management systems — 17
7. ISO/IEC 42006:2025: quality of AIMS certification bodies — 19
8. ISO/IEC 17021-1:2015 and certification audit methodology — 21
9. ISO/IEC 17065:2012: certification of products, processes and services — 23
10. Laboratory testing, ISO/IEC 17025 and technical uncertainty — 25
11. ISO/IEC 17029:2019 and verification of specific claims — 27
12. ISO/IEC 17011 and accreditation of conformity assessment bodies — 28
13. The global mutual recognition system and the 2026 transition — 31
14. The European accreditation framework: Regulation 765/2008 and EA MLA — 32
15. The EU AI Act: the 2026 version and application timetable — 35
16. Classification of high-risk systems and selection of the assessment route — 37
17. Annex VI: internal control without a notified body — 40
18. Annex VII: notified body involvement and technical documentation — 41
19. Notified bodies: designation, competence and oversight — 43
20. Harmonised standards, common specifications and presumptions — 44
21. EU declarations, certificates and CE marking — 46
22. Technical documentation, quality systems and traceability — 48
23. Technical test design and the limits of statistics — 49
24. Bias, fundamental rights and impact assessment — 52
25. Robustness, cybersecurity and external dependencies — 53
26. Generative AI, general-purpose models and assessment issues — 56
27. Continuous monitoring, substantial modifications and withdrawal — 58
28. The United Kingdom: the regulated AI Assurance market and UKAS — 60
29. The United States: NIST, voluntary frameworks and sectoral context — 62
30. Singapore and AI Verify tools — 64
31. The OECD, due diligence and non-governmental standards — 66
32. Procuring AI Assurance and checking counterparties — 69
33. Confidentiality, personal data and model access — 71
34. Sectoral scenarios: banking, sanctions and international cooperation — 73
35. Case study: a legal assistant and the limits of certification — 74
Checking the adequacy of the legal assistant's provisional file — 76
36. The proposed institutional programme of Observatoire ARGA — 77

37. Conclusions and limits of international mutual recognition — 79
Appendix 1. Certificate of approval and result AI Insurance — 80
Appendix 2. Authentication protocol and certificate area — 82
Appendix 3. Map of the control route on AI Act (ed. 2026) — 84
Appendix 4. Functional map of ISO/IEC standards and EU sources — 86
Appendix 5. Model test protocol for sanctions matches — 87
Appendix 6. Statistical uncertainty: replicable training calculations — 88
Appendix 7. Evidence register for limited independent analysis — 89
Appendix 8. Example of limited expertise — 91
Appendix 9. Contractual conditions for external evaluation of IA — 92
Appendix 10. Form of assessment of the independence of the perpetrator — 93
Appendix 11. Certificate monitoring programme and re-evaluation triggers — 95
Appendix 12. Draft institutional proposals — 96
Appendix 13. Draft Universal Passport AI Insurance — 98
Appendix 14. Decision-making matrix on certificates and certificates — 99
Appendix 15. Working Protocol on Legal Qualifications under EU AI Act — 100
Appendix 16. Independence and quality control matrix of the assessment team — 101
Appendix 17. Control protocol for AI sanctions screening systems — 102
Appendix 18. Technical protocol for a generic legal assistant — 103
Appendix 19. Contract annex on reproducibility and storage of evidence — 104
Appendix 20. Documentary monitoring programme after AI Insurance — 105
Index of Practical Recommendations — 106
References and Supporting Sources — 108

Annotations

The report examines what evidence makes it possible to substantiate a specific property of AI and what legal meaning the confirmation has. AI Assurance is considered as a set of research and organizational procedures with different subjects: management process, system technical property, individual statement or evaluator competence. The central result is a methodology for verifying the consistency between the promise, the object actually studied and the permissible conclusion. It combines metrology, statistics, quality management theory and comparative legal analysis. [M1–M2]

The normative part distinguishes between auditing, testing, verification, certification, accreditation and notification, and also considers the revision of AI Act 2026 and conformity assessment standards. The scientific-methodological part reveals how to determine the measured property, take into account uncertainty, check the rule of decision and assess the stability of classification of evidence. For practical use, a breakdown of the legal assistant's conditional file has been developed:

the management system certificate is compared with a technical test, a statement on the quality of links and the limits of the intended use. [1–8; 21–28; 32; M1–M2]

The main conclusion: the suitability of the certification is determined by the substantive relationship between the criterion, method and solution. The process organization document does not automatically confirm the quality of each response; a successful test of a single function does not establish the legality of all applications. The ARGA is offering a research programme to verify the adequacy of evidence, including independent qualification of materials, a rule for addressing contradictions, and pilot verification of the reproducibility of an expert's conclusion.

Area, method and restrictions

The report examines the state of publicly available regulatory and institutional sources as of October 21, 2026. EU standards are analysed by consolidated act of 27 July 2026 and official article charts; ISO/IEC professional standards - by published descriptions of their subject matter and status. Full paid texts of the standards were not obtained; therefore, no table in this report gives an author's methodology for literal reproduction of a standard item. The legal consequences of a specific certificate are determined taking into account the system category and applicable laws. [1-3,21-31]

Content models, conditional numerical examples, questionnaires and forms are methodological tools developed within the framework of this study. They do not indicate that ARGA has actually been accredited, that contracts have been concluded with certification bodies, or that the designated models have actually been tested. The findings should be used as a basis for consultation after internal agreement and updating of the relevant right at the date of submission.

Scientific grounds and research design

The scientific task of assurance is to justify the transition from observation to solution. Metrology helps determine the measured property and procedure; statistics evaluates the uncertainty of observations; quality control theory links verification to the reproduced process; institutional analysis investigates the competence and independence of the evaluator. The legal methodology establishes the legal value of the result. Their functions are divided so that the formal status of the document does not replace the verification of its contents. [M1–M2]

Basis	Research task	Results in the file
Metrology	Identify property, method and conditions of observation	The passport of the parameter being measured
Statistics	Check the denominator, dependency and uncertainty	Calculation and limits of statistical output
Quality management	To determine whether the declared procedure is in operation	Relationship of the document to the execution records
Institutional analysis	Checking the competence and influence of interests	Founded area of trust in the evaluator
Comparative law	Establish applicability of status and procedure	Separate legal opinion

Operational determination of sufficiency

In the proposed methodology, evidence is sufficient for this assertion if the object and criteria are defined, the materials are suitable for verification, substantial contradictions are resolved, and the language of the conclusion does not go beyond the area studied. This is a copyrighted working definition for the analysis of the case files. It does not replace the special rules of a formal conformity scheme and does not make sufficiency a universal numerical indicator.

The definition of the parameter precedes the measurement. "94 per cent of correct references" can mean the proportion of references among all references, the proportion of answers without error, or the proportion of files without an incorrect substantive reference. The conversion of one indicator to another without baseline data is not permitted. The passport includes the unit of analysis, the denominator, the rules of significance and the conditions under which the case is recognized as uncertain.

Rule of decision and uncertainty

JCGM 106 reviews the relationship between uncertainty in measurement and conformity assessment [M2]. For the proposed study, this is the basis for requiring clear rules of decision before the result is obtained. The rule should explain how the assessment and threshold are compared, what the risks of mistaken admission and rejection are taken into account, and when the outcome of the "data insufficient" is applied. The specific threshold is determined by the appointment and powers of the owner of the procedure.

Statistical uncertainty does not cover all the defects in the study. Furthermore, they check the completeness of the input data, the suitability of the benchmark, the environment and the independence of the observations. A narrow interval around a displaced estimate does not make the evidence sufficient. If there is a significant measurement defect, the method is first corrected or the statement narrowed rather than the number of repetitions of the same inappropriate procedure increased.

Pilot research issues

The pilot should check whether a structured passport of approval improves the reproducibility of an expert qualification compared to the usual reading of a set of documents. The unit of analysis is the file; the results are missing a substantial replacement of the subject, unfounded positive conclusions and reasoned disagreements among experts. The working time is an additional characteristic and does not replace the quality of the qualification.

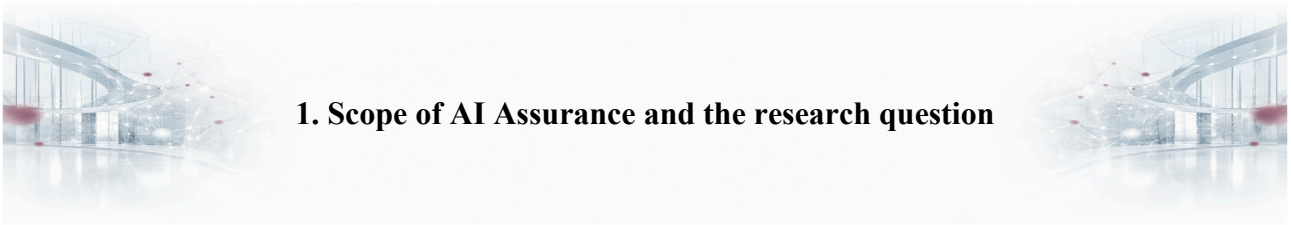
For comparative design, sets of comparable complexity, distribution of work order and concealment of the prepared reference breakdown from evaluators are offered. Re-reading the same set can have a learning effect; it should be taken into account or used by different, pre-matched sets. The reference group preserves the basis of its decision and allows for several reasoned answers where the question is really ambiguous.

Conditions for scientific verification of the result

The consistency of the conclusion to a change in the composition of the file, the existence of a contradictory application and the concealment of advertisements by the supplier is checked. If the qualification changes after the logo has been removed with unchanged content data, the impact of

the reputational signal is required to be analyzed. For a positive outcome, it is important that another specialist restore the basis of a decision without an explanation from the author.

Pilot results do not apply to all certification procedures: the transfer is limited by the composition of documents, the training of experts and the tasks investigated in the program. The present report presents a draft audit and a conditional review. Proof of the advantages of the proposed methodology will require the description of the study, registration of deviations and independent evaluation of the materials obtained.



1. Scope of AI Assurance and the research question

1.1. Guarantee as a valid claim

The certificate under investigation is presented as a verifiable claim for a particular object. Instead of "helper is reliable", it is necessary to specify the function, version, composition of the input material and the criterion: for example, the proportion of files in which all significant references are supported by primary documents. Separately determine what action the addressee of the result is entitled to take. This allows you to test whether the area of commercial promise matches the area of the research being done.

Reasonable trust is built from the verified thesis to the proof, not from the outwardly convincing document to the thesis. It should be determined to whom the result is addressed: to the customer when choosing a supplier, to the regulator, to the group auditor, to the buyer of services or to the potential affected person. The necessary details are being changed for each addressee. Even an impeccable formal management system certificate can tell little about the behavior of a single model on a rare set of input data.

1.2. Four limits of assurance

First, the technical boundary is fixed: whether the model, preprocessing rules, search database, interface, decision maker, and external API are included. Then the functional boundary: which solutions are covered and in what cases the result can be used. The third is the normative boundary: whether a voluntary scheme or statutory assessment is applied. The fourth is temporary: which version and data was checked and what changes made it obsolete.

An error occurs when the conclusion of one level is extended to the other. Testing the stability of a document's recognition does not prove the legitimacy of using biometrics; certification of information storage procedures does not prove the absence of unlawful grounds for processing. So the volume map is part of the result and should be available with the output.

1.3. Which is proof and what the marketing signal is.

A copy of the promotional brochure, the word "ISO compliant" and the certificate of course are not equal to the certificate. The original document is assessed by issuing authority, identifier, status, subject matter, scheme and traceability to the accreditation register. The results of the tests are

examined separately: input data, marking rules, code, version, expert qualification and uncertainty are important even when there is a recognition mark.

For public use, it is advisable to introduce a rule: each mention of a certification is accompanied by a reference to the evidence available for verification and a narrow wording of its value. Otherwise, the reader may incorrectly conclude that the certification authority has certified the whole product or compliance with all requirements of EU law.

Practical checklist

1. Name one measurable statement and one who will rely on results.

2. Determine the version, context of use, period of observation and exclusion.

3. Divide the proof of procedure, proof of result and proof of legal characterisations.

Illustrative example

The Legal Assistant correctly refers to documents in 94 of the 100 responses checked: this number only refers to the sample, the valuation method and the version, and not to the "legal accuracy" of any future responses.

Operational definition of sufficient assurance

The degree of sufficiency of assurance must be determined not by the prestige of the performer but by the consequences of the erroneous conclusion. When using AI to sort out open news, the depth of the investigation is quite different than when deciding whether to close a bank account or report a person's identity to the police. In a sensitive case, it is necessary to check not only the average quality of the model, but also the mechanism of refusal of response, availability of secondary checks, resilience to poorly documented names and the ability to recover all significant inputs and actions. The decision to apply the result should record what remaining doubt has been accepted and who has the authority to accept it. [39,44]

The use of an "approval passport" is useful for the expert opinion. It links the subject matter of the verification, the person concerned, the risk category, the criteria applied, the verified version identifier, the test environment, the selection limitations, the status of the assessor's independence and the timing of the validity. Such a structure protects against an implicit shift from the phrase "a set of tests" to "a safe system". A positive result without a clear list of exceptions is incomplete: the recipient is unable to understand what questions the expert has not studied at all.

A particular problem arises in combining individual estimates. An API cybersecurity report, a corporate management certificate, and a recognition quality test do not automatically form a single conclusion about the legitimacy of the final decision. For a unified conclusion, a clear logic of transition between evidence and independent verification of assumptions about the interaction of components is necessary. This also applies to the results of subcontractors, who may have different observation times and different versions of the facilities.

Practical case analysis 17.01

Recommendation: Clarify the subject of the assurance

Hypothetical example: The client asks for confirmation of the safety of the entire product following the results of the quote check.

Analysis: The quote test relates to one characteristic and does not cover safety of action.

Practical steps: The head formulates separate statements about quotes, data and actions; assigns evidence for each.

Outcome criterion: The conclusion contains proven properties and excluded areas.



2. Terminology and levels of assessment

2.1. Self-assessment, audit, testing and independent audit

Self-assessment can be useful for continuous monitoring, but the customer is both the source of the data and the person who is interested in the result. The audit shall verify that the criteria are met on the basis of a set of evidence. The test is a technical experiment with a pre-known procedure. Verification assesses the truth of the formulated statement; validation - the validity of the statement with respect to the intended use. These actions may be combined, but they are not synonymous. [27-28,39]

A conditional model accuracy test does not reproduce an integrated risk management audit. In turn, verifying the existence of a monitoring procedure does not replace the actual measurement of false matches. The working case must contain individual conclusions on the operational processes, technical characteristics, legality and independence of the procedure.

2.2. Certification and accreditation

The certificate reports the result of evaluating a particular object according to a particular pattern. Accreditation assesses the competence and impartiality of the body itself conducting testing, inspection, validation or certification. For example, the accreditation of a management system certification body for the applicable field does not automatically accredit each of its additional services. [23-25]

The competence of the specialist, recognition of professional qualifications, membership of the association and notification for a specific EU procedure require separate grounds. For this reason, it is wrong to replace the audit of the accreditation area with "what company logo is on the website". Scope and status are always more important than formalization.

2.3. Notification and legal presumption

Notification means the authorisation of a body for certain functions under the EU's special legal regime. It is not the same as voluntary accreditation: Articles 28-31 of AI Act provide for a separate role of competent authorities and a procedure for reporting to the Commission. The existence of an accreditation can be a basis for a notification, but the absence of accreditation does not always completely exclude the possibility of providing alternative proofs of competence. [9-12]

The presumption of conformity to the harmonised requirement does not arise from the publication of ISO as such. It is required to comply with the requirements of the special standard and, in the

European mechanism at issue, to publish a reference to the harmonised standard in the Official log of the EU in the volume covering the relevant requirement. [7]

Practical checklist

1. Indicate the type of document: test report, auditor's opinion, certificate or accreditation.

2. Name the body that gave recognition and the programme in which it operated.

3. To check whether there is an unacceptable transition from voluntary to mandatory legal status.

Matrix A. Type of result and its legal meaning

View	Which confirms	Doesn't prove anything.
Internal audit	Monitoring of certain criteria and control of their own processes	Independent State recognition
Laboratory test	Measured properties by method and in the designated area	Legality of any use
ISO/IEC 42001	IA management system in the declared area	Quality and safety of all selected products
CAB accreditation	Competency of the evaluator in a specific area	Mandatory suitability of any document issued by him
EU AI Act Notification	Official competence of the body under the designated procedure	Authority outside the object of notification
EU Declaration of Conformity	Responsible declaration by special regime supplier	External certification by internal route

Register of types of evidence and prohibition of substitution of status

Each case should include four separate fields: type of activity performed, status of its performer, legal basis for the application of the result and specific field of actual verification. For example, a laboratory may be accredited to test a certain class of properties, but its conclusion about another property does not become accredited because of the logo on the first page. Similarly, a management system certificate covers a specific organizational outline, not all legal entity products without exception. [23-28]

If the employer requires "independent certification of artificial intelligence", the perpetrator must translate this formula into the language of the audited requirements prior to the conclusion of the contract. It should be asked whether management processes, metrically determined product quality, a special legal conclusion or independent confirmation of the supplier's claim are necessary. Only after the answer can you choose the body, the accreditation area, the scheme and the appropriate criteria. Otherwise, the customer pays for a document that may be formally flawless but is useless for its actual purpose.

In the manual, it is recommended to keep the original terminology of the party and place its qualification next to it. Thus, the "trustworthy AI certificate" should be described as a private commercial approval pending receipt of the scheme number, criteria and competent authority. It's not an accusation of dishonesty: mistakes often arise from the simplicity of marketing language. But for legal use, simplification is not acceptable.

Practical case analysis 17.02

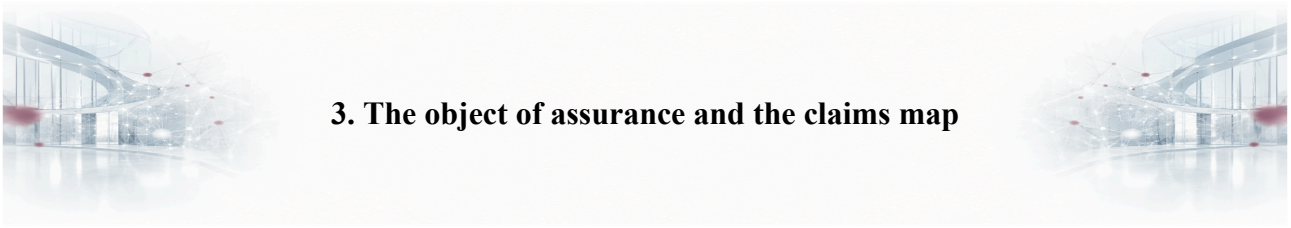
Recommendation: Segregate test and certification

Hypothetical example: The supplier calls the internal test a certificate of conformity.

Analysis: The name does not confirm the procedure and authority of the person issuing it.

Practical steps: The buyer asks for a scheme, criteria, executor and area; describes the document in actual content.

Outcome criterion: The register indicates the valid type of document without an extended name.



3. The object of assurance and the claims map

3.1. Management system and product

ISO/IEC 42001:2023 establishes an IE organizational management framework. The object can be an organization's management system that implements AI, develops it, or provides AI services. The certificate of such a system indicates the conformity of the management system in the claimed field. It doesn't guarantee the accuracy of every model, and it doesn't turn into a security certificate for any vendor algorithm. [21,48]

A particular product will require separate evidence that is focused on its properties: a version description, functional designation, realistic scenarios, test results and a legal map. Technical product testing may support the management system, but it is not a complete replacement; similarly, a high-quality corporate process does not correct a critical error in a particular configuration.

3.2. The fragmentation of an object in the I.I. chain

The fundamental model provider provides access through the API, the integrator adds search through the internal archive, the customer sets the access rules, and the employees use the result to evaluate the client. In practice, each link controls only part of the system's behaviour. The general phrase "certified AI" erases this difference and creates an unacceptable expansion of responsibility.

For each element of the chain, it is necessary to identify the controlling person, the version used, the proof of control, the dependence on third parties and the conditions for re-verification. If the external model changes without notice, the previous integration test may lose its validity even without updating the user interface.

3.3. Commercial approval and audited boundary

The phrase "ethical", "fair", "safe" and "compatible to all international standards" is rarely suitable for assurances without a decomposite. The terms used, the method of measurement, the thresholds

for acceptability and the risk approval document should be requested. In the absence of such information, the conclusion may be limited to indicating that the original statement is not verifiable.

The individual risk is a certificate whose area is too broad or incomplete: AI services does not disclose whether model training, client support, medical data processing or subcontracting processes are included. The exact area specified in the existing annex to the certificate should be checked.

Practical checklist

1. Create a map of components and external relationships with change owners.

2. Place advertisements for approvals and verifiable indicators.

3. Set the limits for the applicability of the result to each version and subject area.

Division of responsibilities in the multilevel system

The area of assurance should be given through an architectural scheme and a register of relations with suppliers. A hybrid system can have a general model provider, a cloud execution provider, a corporate search owner, a rule integrator, an end-operator, and a significant decision maker. Everyone has separate control capabilities. If a person only owns the search subsystem, he has no right to confirm the immutability of the external model. If an external service stores requests, the organization's internal certificate does not prove the actual storage of those records. [21,22,49]

An annex to the contract should contain versions of models and components, data transition points, subcontractors lists, API change notification obligations, access to logs, emergency shutdown conditions and re-checking requirements. For models that run on constantly updated subscriptions, it is particularly important to distinguish a fixed version from services with unlimited parameter changes. A declaration of the unchanged area of a certificate in the event of a genuinely unmanageable update presents an evidence risk.

The practical map of the statements is made in two dimensions: lines show the properties of the system, columns show elements of its chain. The cell is filled in only if a specific responsible person is present and a evidence is available. An empty cell does not mean no risk; it indicates a control gap. This allows you to identify at an early stage where additional powers or restrictions on planned use are needed.

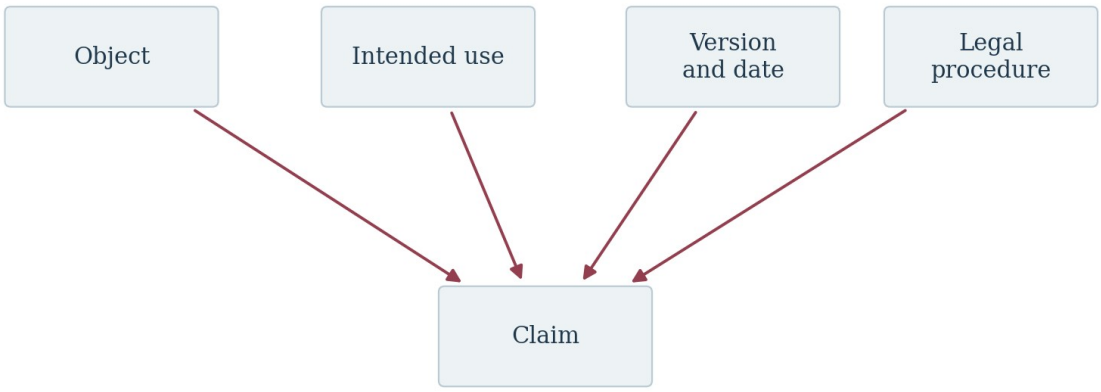


Figure 17.1. The boundaries of a conformity claim. Four separate boundaries from Section 1.2. A change to any one may require reassessment of the assurance provided.

Practical case analysis 17.03

Recommendation: Create a passport for each claim

Hypothetical example: The 95 per cent statement of accuracy does not contain a denominator.

Analysis: These may be links, answers or files; their shares are not comparable.

Practical steps: The analyst requests the unit of account, the original quantities, the version and the scope.

Outcome criterion: The indicator can be re-calculated and linked to a validated appointment.



4. The evidence architecture of AI Assurance

4.1. From criteria to chain of confirmation

For each statement, a justification map is drawn up: precise formulation, object and version, source of criterion, method of observation, source data, rule of decision and permissible field of application. The card also includes a circumstance in which an affirmation must be rejected or narrowed. The documentation of this connection makes it possible to make a substantive review: the specialist checks not the number of applications, but the suitability of each evidence for a specific conclusion.

Different evidence can only be complementary if they have an explanatory relationship with the statement. For an application to perform the procedure, records of actual use are required; for an application to answer quality - a test; for an application to legal status - a document of the competent authority and a verification of its field. Three documents that relate to one supplier's report do not constitute three independent confirmations. The origin of the data and the overall relationships are shown in the map.

4.2. Quality of input materials and reproducibility

The reproducibility of a technical evidence requires the initial answers, task identifiers, system configuration and evaluation procedures. The researcher must restore the denominator and explain each exception. If 94 of the 100 assignments in the report were successful, the fate of the other six would have to be determined: errors, failures, technical failures or exceptions to the protocol lead to different conclusions. Recalculations of stored materials precede the evaluation of the transportability of the result. [39–41]

If there's a restriction on access to source code or learning data, it shouldn't be hidden. Independent functional testing or process evaluation can be carried out with sufficient border transparency. However, the conclusion is bound to state explicitly which properties have not been tested due to lack of access.

4.3. Conclusions and level of confidence

Five statuses are offered for the analysis of a separate claim: confirmed in the specified field; partially confirmed by enumerating boundaries; denied by an observation inconsistent with the criterion; insufficient evidence; not investigated. Partial confirmation is not used as an indefinite compromise: it must be clear which part is established and which remains open. For a scheme with its own categories of solutions, its terminology is used, and the research map explains the reasons for this decision.

In a legally sensitive sphere, it is desirable to record the competence of the signatory. A technician can evaluate the reproducibility of the procedure, but a legal characterisation is required to determine the legal applicability of a particular regimen. The signature of one person under both types of conclusions must be accompanied by an explanation of the necessary competence.

Practical checklist

1. Segregate monitoring from interpretation and recommendation.

2. For each significant negative conclusion, note the amount of material tested.

3. Do not "fully correspond" if a limited sample or an inaccessible component has been used.

Matrix B. Selection of tools for proof

Control issue	Main type of evidence	Warning
Is there a management procedure?	Documents, designated roles, decision log	Document may not be executed
Does the model work on a given hull?	Version test protocol, reference, error matrix	Do not extend to all data
Is the independent evaluator competent?	Accreditation roster, field, staff	Checking this service
Are there rights of the person concerned?	Description of appeal, real case, regulatory analysis	Not enough interface notice
Is the quality of change preserved?	Version register, regression tests, API contract	One reference test becomes obsolete

Evidentiary file and rebuttal architecture

Sufficiency testing involves trying to find a contradiction between the statement and the materials. The researcher compares the claimed function with the certificate area, re-computes the key metric, checks the excluded cases, and chooses a difficult example that can change the conclusion. The discovery of a contradiction is not resolved by a vote or a reference to the reputation of the supplier. It is resolved by new evidence, narrowing of the statement or fixing the inadequacy of the available information.

Negative evidence should be documented. If the log does not contain the requested event, the expert should know the storage time, the full download, the timing system, the filters and the likelihood of loss. Only then can it be distinguished "no event" from "no trace of the event". Similarly, the absence of open source incidents does not exclude incidents that have not been published or described in other terms.

Two-step verification is recommended for substantial conclusions: the author forms a conclusion and a list of references to the original artifacts; an independent reviewer attempts to repeat the transition from materials to conclusions without the author's involvement. If the reviewer needs oral explanation, there is a gap in the text or in the registry. When working with sensitive data, repeatability can be achieved through pseudonymized control records, control amounts, and documented access procedures without spreading a complete closed array.

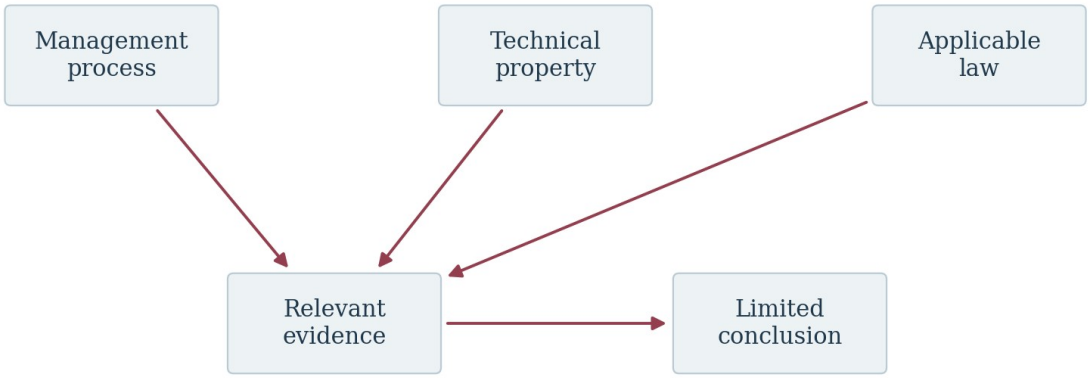


Figure 17.2. What different forms of evidence establish. Management certification, technical testing and legal classification address different objects. They are not interchangeable.

Practical case analysis 17.04

Recommendation: Check the sufficiency of the entire chain of evidence

Hypothetical example: The management certificate is accompanied by a different version test and an undated description of the process.

Analysis: The evidence does not belong to the same facility; its sum does not confirm the current assembly.

Practical steps: The expert associates each statement with the version and date; he forms gaps by requesting additional material.

Outcome criterion: For untraceable properties, the conclusion is that there is insufficient data.

5. Institutional quality infrastructure

5.1. The chain of participants - criteria - decision

The standardisation organisation develops the document, but usually does not issue a certificate based on it. The conformity assessment body shall carry out the verification procedure provided for; the accreditation body shall confirm the competence of this evaluator in a particular field; the regulatory body shall determine the mandatory consequences of individual procedures. Depending on the sector, several roles may be allocated differently, but their mixing is unacceptable. [19,21-28]

When checking the provider, it is necessary to trace the chain from the verified endorsement to the real entity that has the right or ability to confirm it. International standards and accreditation are useful only within such a chain, not as a universal symbol of integrity.

5.2. Objectiveity and competing interests

The evaluator's relationship with the developer can influence the choice of tests and thresholds, even if there is no open financial conflict. The problem is participating in the development of the system, promising a result before the start of work, a reward for providing a positive report, and depending on one client's permanent contracts. The impartiality management system should not only be declared, but also include a threat assessment and measures taken. [12,22-24]

Notification and accreditation procedures are built around independent competence; external research verification must be based on a comparable principle, without giving formal status. For ARGA, this means separate conflict of interest checks when drawing conclusions that are planned to be submitted to regulators.

5.3. Monitoring the status of public statements

An accredited laboratory, non-accredited consulting, staff training and voluntary evaluation services can be simultaneously available on the organization's website. A reasonable user cannot consider them to be one form of official recognition. You should ask for the accreditation number, the name of the area, the standards, the list of addresses and the date of the last update.

Even if several companies are part of the same group, the accreditation of a particular legal entity is not automatically transferred to affiliated organizations. Similarly, the logo of an international agreement on mutual recognition should not be accompanied by conclusions outside the scope of accreditation.

Practical checklist

1. To verify a legal entity, not just a trade name.

2. Check the separate services and areas of accreditation.

3. Record the result of the status check at a specific date.

II. International standards

Three lines of control and organizational independence

The independence of the verifier must be analyzed throughout the chain: method selection, data acquisition, thresholds, interpretation of exceptions, approval of conclusions and publication. An external contractor is not automatically independent if its remuneration depends on receiving a positive score or if the subject under review may prohibit disclosure of material negative findings. For such situations, a pre-agreed condition is required that the professional conclusion cannot be edited by the customer. [22,23,28]

It's practically useful to differentiate between the process owner's internal control, internal audit of risk functions and external confirmation on specially agreed criteria. No line is a universal replacement for others: the developer knows the technology better, but has an interest; the outside authority has distance, but may not get enough access; the supervisory authority assesses issues of legality that are not always the subject of the certificate. Therefore, the duty matrix should contain a separate line to escalate differences between the lines.

When selecting an expert, it is reasonable to obtain declarations on financial, labour, family and contract relations, information about previous consulting services, decision-making structure and the procedure for dismissing the examiner. Not every possible conflict requires automatic refusal, but every significant conflict must be assessed in terms of the threat of impartiality and the availability of safeguards. A declaration signed only when a contract is concluded, without re-checking it during the project, is not sufficient.

Practical case analysis 17.05

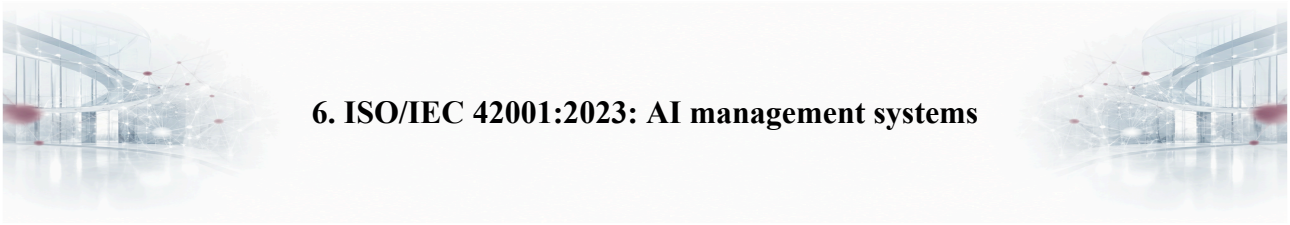
Recommendation: Check each member of the quality chain

Hypothetical example: The certification authority refers to the accreditor but does not specify the area of recognition.

Analysis: The existence of an organization does not confirm specific competence.

Practical steps: The responsible person checks the authority, area, time frame and scheme with the primary registry information.

Outcome criterion: The status chain to be tested for the correct type of assessment is retained.



6. ISO/IEC 42001:2023: AI management systems

6.1. Subject matter of the standard and legal status

ISO/IEC 42001:2023 describes the creation, maintenance and continuous improvement of an artificial intelligence management system. It is aimed at organizations involved in the development, delivery and use of AI and serves as a framework for allocating roles, processes and management measures. According to the open ISO record, its subject matter is management system, not evaluation of a particular model or transferring control rights to the regulator. [21]

Management system standards are applied on a voluntary basis if a special rule or contract does not make their enforcement mandatory in a specific respect. Even after the standard has been implemented, certification remains a separate decision; an organization can run a workable management system without a certificate, and the certificate does not eliminate its own legal obligations. [48]

6.2. Practical evaluation of AIMS area

The subject of certification can be reasonably compared to business processes: inventorying AI systems, changing models, purchasing external services, data control, unit responsibility, employee training, complaint management and incident documentation. In this case, the laboratory evaluation of a specific model response goes beyond simply verifying that processes are provided.

If "IT solutions" are certified, it is not enough to consider all processes of risky use included. It's important to check the literal text of the certificate area, the functions of the excluded branches and the responsibility of the cloud provider. When purchasing one product, a customer may need not only a company certificate but also product test protocols.

6.3. That certificate 42001 should not promise

AIMS certificate cannot be mechanically inferred to comply with the GDPR, the EU AI Act, sanctions monitoring rules or specific requirements of a medical device. It is part of the compliance evidence and can confirm organizational maturity, but legal characterisations and compliance with a particular system require other materials. [1,17,21]

Correct statement in the contract: "the management system of the organization is certified in the declared field". Uncorrect: "All decisions of a company are approved by an international regulator and can be applied without further verification". The rule applies to purchases, to the publication of conclusions, and to the handling of claims about misleading marketing.

Practical checklist

1. Read the literal area of certification and exceptions.

2. Link AIMS to actual processes of design, implementation and operation.

3. Separately request product tests and applicable regulatory map.

Audit of AIMS certificate of contents

For ISO/IEC 42001, the key question is not how many models the organization uses, but what management system processes are covered by the certification area. In a single legal entity, development, sale and operation may be distributed between units and territories. Therefore, when purchasing a certificate, you should ask for an annex with the area, the list of areas covered, information about the issuing authority and the current status of the certificate. A public logo cannot be used to determine whether a particular service falls within AIMS area. [21,23,48]

Evidence of the functioning of the management system can be an AI registry, liability policy, impact assessment records, model change logs, unacceptable risk decisions, monitoring procedures, and proof of the error-correcting mechanism. But the availability of these documents does not prove the effectiveness of each procedure. For independent analysis, it's important to select a sample of real solutions: who detected the deviation, who authorized it, whether the affected persons were notified, whether the criteria changed after the adverse outcomes.

A particular risk relates to the term "certified model". It is legitimate to indicate that the organization has certified AIMS in a particular field; it is unlawful to disseminate such a conclusion without any separate reason to the accuracy, safety, absence of discrimination or legal characterisation of each product. In ARGA reports, it is advisable to keep the original formulation of the certificate and make a short note about its limits.

Practical case analysis 17.06

Recommendation: Segregate Management System from Response Quality

Hypothetical example: The supplier shows AIMS certificate after a false link has been discovered.

Analysis: The organizational process itself does not confirm the absence of product errors.

Practical steps: The assessor takes the certificate into account as a separate certificate and assigns technical verification of the references.

Outcome criterion: The technical conclusion is based on a specific version of the test.

7. ISO/IEC 42006:2025: quality of AIMS certification bodies

7.1. Special Supplement to ISO/IEC 17021-1

In July 2025, ISO/IEC 42006, dedicated to the audit and certification bodies of AI management systems, was published. The ISO Open Card describes it as an additional requirement to ISO/IEC 17021-1, necessary because of specific AI-risks, data quality, ethics, transparency and competence. The use of 42006 does not cancel out the basic requirements for impartiality and processes of the certification body. [22-23]

The practical value of the new revision is in a standardized language for the evaluation of the "verifier". It is no longer enough for the client to ask whether the auditor has been trained in general management systems. It is necessary to understand whether an evaluator is able to understand AI architecture, organizational contours, data boundaries and the role of humans in decision making.

7.2. Competence, team composition and supervision

AI Assurance requires a combination of knowledge about system management, model evaluation, data protection and industry use. The presence of a specialist with a technical certificate does not mean the ability to qualify EU law; the presence of a lawyer does not replace the statistical verification of the sample. The evaluation body must determine the necessary competencies and ensure their practical presence, not a conditional signature on the final report.

Accreditation checks are not only about the employee resume, but also how the body appoints teams, monitors independence, provides access to expert assistance and makes decisions. The certification report requires a documented link between the auditors' observations and the opinion of the competent decision-maker.

7.3. Transferability limits 42006

ISO explicitly indicates that 42006 can be used within broader product, process or service assessment schemes based on ISO/IEC 17065, but this does not make a management certificate a product certificate. Such a scheme requires independent criteria, subject matter and procedure. [22,25]

In particular, AIMS verification of the biometrics developer does not allow to state that the special Annex VII procedure has already been completed. A separate reporting regime, roles and access to evidence is applied. The error "42006 = the right to issue AI Act certificates" should be excluded in all instructions and marketing texts.

Practical checklist

1. Request the applicable accreditation area of the body, not just a reference to 42006.

2. Assess AI specific competence of the team and the authority's authority's authority to sign the document.

3. Check the type of certificate with the type of declared object.

Competence of AIMS bodies and separation of decisions

ISO/IEC 42006:2025 addresses requirements to bodies that audit and certify AIMS by supplementing ISO/IEC 17021-1; on an open ISO record it is an independent published standard, not a project. Competence assessment should analyze not only the auditor's overall experience in management systems, but also the team's ability to work with systemic AI risk, data, software, exploitation, human rights and corporate responsibility. It's the combination of competencies that reduces the likelihood of a formal audit without analyzing the actual impacts. [22,23]

In a complex infrastructure, the certifier may engage subject matter experts, but their participation should not erase responsibility for the audit decision. It is important for the customer to determine who evaluates the technical reports, who is responsible for impartiality, who decides on certification and how complaints are handled. The professional biography of the consultant or the publication of his/her own training programme is not sufficient proof of the status of the certification body.

It is particularly useful to distinguish the area of competence from the area of commercial offer. The agency can advertise a broad portfolio of AI Assurance, while its accreditation extends to the narrow AIMS scheme. Results outside the scope of this agreement may remain services, but the customer must understand their status and not describe them as an accredited certification. This distinction should be reflected in the procurement documents and in the final report.

Practical case analysis 17.07

Recommendation: Check the competence of the certification team

Hypothetical example: An audit of a complex language tool is conducted by a team without substantive experience.

Analysis: A formal appointment does not guarantee an understanding of the process under study.

Practical steps: The requisitioner requests a competency matrix and a review of contentious technical issues.

Outcome criterion: Substantive issues have been addressed by a specialist with proven competence.

8. ISO/IEC 17021-1:2015 and certification audit methodology

8.1. Procedures, impartiality and decision

ISO/IEC 17021-1:2015 is addressed to the bodies that conduct audits and certification of management systems. The open description confirms the emphasis on competence, consistency and impartiality. In AIMS-sphere, it's a common layer standard to which AI-specific 42006 is added. [23]

When preparing for a verification, it is important to distinguish between the independence of the fact-gatherer and the independence of the certification decision. Functional division reduces the risk when the conclusion is accepted by a person interested in a commercial relationship with the client. It is also important for the customer to complain and check the status of the certificate after initial assessment.

8.2. Observation, sampling and inconsistencies

The certification audit uses a selection of processes and evidence: it's impossible to control every minute of an organization's decisions. For this reason, the description of the sampling method should take into account the critical nature of the process, the number of sites, the frequency of changes and the history of incidents. A positive result for a sample does not mean a mathematical guarantee that there are no unknown violations.

For AI-systems, the inconsistency can be organizational (no designated risk owner), procedural (changes are not approved), evidence-based (no required log) or technical (the system gives a dangerous result). The report proposes that these species be separated instead of being grouped into one category, "AI risk".

8.3. Integrated management systems

Companies can combine AI management with information security and quality. This integration reduces the duplication of roles and reporting, but does not automatically transfer certification from one standard to another. Each programme has its own criteria and a documentary trail.

When you buy, it's important to ask who's in charge of the interface between AI and information security: delimitation of access rights, staff training, logging, and responding to the model finding. The phrase "we already have ISO 27001" does not prove that the consequences of an erroneous financial or legal opinion have been assessed.

Practical checklist

1. Describe the audit programme and independence of decision-making.

2. Select high-risk processes for the extended sample.

3. Document the inconsistency and empirical error of the model.

Certification cycle and verification

ISO/IEC 17021-1 concerns the body that certifies management systems and provides a systematic approach to impartiality and competence. In the technical part of the report, it is advisable to study AIMS operational cycle rather than replace it with a one-off policy demonstration. When auditing, information about initial checks, corrective actions, supervisory assessments and re-certification within the applicable scheme is useful; specific periodicities should be checked according to the rules of the certifier and the scheme, rather than assuming the same for everyone. [23,22]

For companies that simultaneously certify quality, information security and AI management, an integrated audit of overall processes is possible, but the results must be distributed according to requirements. The general document management procedure does not in itself indicate a full assessment of the errors of the generative model, and the information security standard does not replace the evaluation of the consequences of an algorithmic solution. The joint use of evidence is only permitted with a documented coverage map.

If significant inconsistencies are found, the conclusion shall not be reduced to "corrected". The cause, the scope of the affected systems, the temporary measures, the effectiveness check and the residual impact on users should be established. For comparative research purposes, ARGA

recommends keeping a history of changes in status, rather than one photo of the certificate at the time of purchase.

Practical case analysis 17.08

Recommendation: Check the actual implementation of the audit programme

Hypothetical example: The plan included interviews with operators, but only management responses were included in the file.

Analysis: The process described is not supported by independent performance observations.

Practical steps: The reviewer compares the program, a sample of the recordings and the actions performed; records the uncovered sections.

Outcome criterion: Each essential step has a record of performance or a reasonable limitation.

9. ISO/IEC 17065:2012: certification of products, processes and services

9.1. Variety of product scheme

ISO/IEC 17065:2012 refers to bodies that certify products, processes and services. This is not the entire management system of the company, but a subject defined by the scheme: for example, a specific version of the system, type of function, available configuration or test technology. The scheme should contain measurable requirements, rules of decision-making and supervision. [25]

This is especially important for AI Assurance because a typical product updates faster than corporate procedures. If the certificate applies only to a fixed model, its marketing use after a significant update without a new assessment can create the false impression of continuous confirmation.

9.2. Certification Scheme and Owner

A standard that establishes the competence of a body should not be confused with a technical standard that sets requirements for a product. ISO/IEC 17065 does not in itself define the overall "passive assessment" of the quality of the IE. Such an assessment arises from a particular scheme and the criteria set out therein. Two programs of good faith can use different properties and thresholds, and their certificates are not necessarily equivalent.

The customer should ask for the text of the scheme, the basis of the thresholds chosen, the list of permissible deviations, the rules on new versions and the procedure for complaints. In AI space, we need data definitions, rejection scenarios, and the role of human control. Without them, a public sign remains too abstract.

9.3. Status of revision 2026

According to the open ISO record for October 2026, the new edition of ISO/IEC FDIS 17065 is available; the current published edition remains ISO/IEC 17065:2012. The text of the draft could not be issued for mandatory requirements. [25-26]

Contracts and opinions must record the year of the revision and the date of its verification. If the revised standard is subsequently published, the organization determines the transition period and requirements of its accreditation program separately. It is also not possible to automatically declare previous certificates invalid.

Practical checklist

1. Set the item and version of the product as reflected in the certificate.

2. Get the evaluation scheme, its criteria, and the basis for the retesting.

3. Not refer to FDIS as a published standard.

Product scheme criteria and the problem of promised quality

Certification of a product, process, or service is built around a scheme that should define the object, set of requirements, permissible assessment methods, decision-making and supervision. ISO/IEC 17065:2012 sets requirements for the activities of relevant bodies; however, the new edition being discussed in 2026 is at the FDIS stage and should not be considered a standard that has entered into force. Before it is published, the reference in the existing schemes must be verified by a specific regulatory edition. [25,26]

For AI, the importance of a scheme is particularly great: the same model can have high quality on a mass set of tests and an unacceptable number of errors in a rare but legally critical subgroup. Therefore, a product evaluation program should determine the origin of the test scenarios, critical thresholds, how to construct the statistical interval, the role of the human operator and the consequences of substantial updating. The broad formula of "ethical verification" without such elements is poorly reproducible.

If the product is delivered as a service, it is necessary to describe separately the change in operating model and the availability of control versions. A positive certificate obtained in an old configuration is not required to cover a new one; the issue is settled according to the appropriate scheme. The buyer must obtain the conditions for suspension, extension and withdrawal of the confirmation. This is particularly important where the supplier periodically replaces the basic model without changing the name of the commercial service.

Practical case analysis 17.09

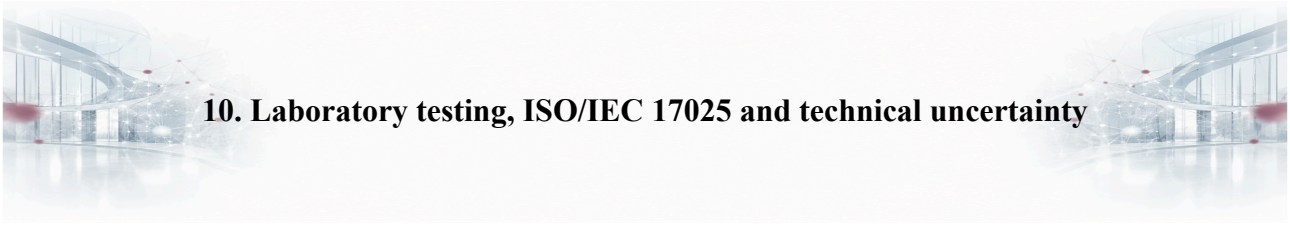
Recommendation: Specify the item of the product scheme

Hypothetical example: The certificate is part of the search module, and the commercial is extended to the entire assistant.

Analysis: The composition of the product is wider than the certified facility.

Practical steps: The expert compares the area of the scheme with the product architecture and corrects the external statement.

Outcome criterion: Untested modules are clearly excluded from the statement.



10. Laboratory testing, ISO/IEC 17025 and technical uncertainty

10.1. The competence of the laboratory is not equal to the universality of the method

ISO/IEC 17025:2017 sets requirements for the competence, impartiality and robust operation of test and calibration laboratories. For AI evaluation, this is a useful infrastructure where a particular test method exists. But lab certification in one field does not give it universal authority to evaluate any model, and the accreditation area can describe limited types of testing. [27]

A legally sensitive customer must determine whether the accreditation covers exactly the desired indicator. For example, accreditation of cybersecurity measurements does not imply an assessment of the fairness of personnel decisions or the accuracy of legal translations without additional grounds.

10.2. Test, benchmark and representativeness

The test result depends on the set of input tasks, the way to form the correct answer, the repetition and independence of the test data from the learners. If the test body is designed by the developer and the model is well known, the high quality can reflect the finishing. It is necessary to distinguish the validating developer sample from the external, pre-closed control sample.

When comparing models, the same rules for query, tools, time limits and evaluation criteria should be provided. It is particularly advisable to use manual correction of responses for one provider, and a fully automatic counting for another. This design violates comparability even without intent to distort the outcome.

10.3. Uncertainty and limit of conclusion

When you look at 100 tasks, you can't predict with high accuracy the likelihood of rare failures in millions of future cases. Statistical intervals should be reported with sample size, success determination and calculation methodology. The accuracy of the number "96.8 per cent" can be elusive if the gauges themselves differ in determining the correct answer.

Systematic errors are considered separately: disproportionately easy questions, invisible data misses, and the exclusion of uncomfortable scenarios. Increasing the sample reduces the random error, but does not correct the systematic nonrepresentativeness. Therefore, independent methodological control must precede the issuance of the final result.

Practical checklist

1. Check the laboratory accreditation area with a specific test.

2. Record the origin of the reference set and possible overlap with training.

3. Describe uncertainty and excluded groups of scenarios.

Tests and admissibility of the finding of rare events

An accredited laboratory demonstrates competence in a particular field, but even an impeccable protocol does not automatically transfer the properties of a tested sample to all future exploitation. For AI, the origin of data, the relevance of allocating inputs, the markup procedure and the rules for handling ambiguous cases are critical. ISO/IEC 17025 is used to assess the competence of a laboratory, but the measurement itself must be justified for a specific task. [27,39]

In sanction screening tasks, high aggregate accuracy can hide the unsatisfactory detection of rare true matches or an unbearable burden of false anxiety. It is advisable to construct the check on layers: common names, transliteration, alternative order of name components, incomplete dates of birth, pseudonyms and misspelled entries. In each layer, the size of the sample, the rule of truth determination and the independence of the markers are fixed. The results of the different layers cannot be confused without an explanation in one general percentage.

With a small number of observations, it is necessary to give an interval of uncertainty and warn of the impossibility of accurately evaluating rare errors. For example, the absence of errors in a hundred tests does not prove zero probability of error in real use. Applications should use reproducible learning calculations and clearly separate them from empirical data from real systems.

Practical case analysis 17.10

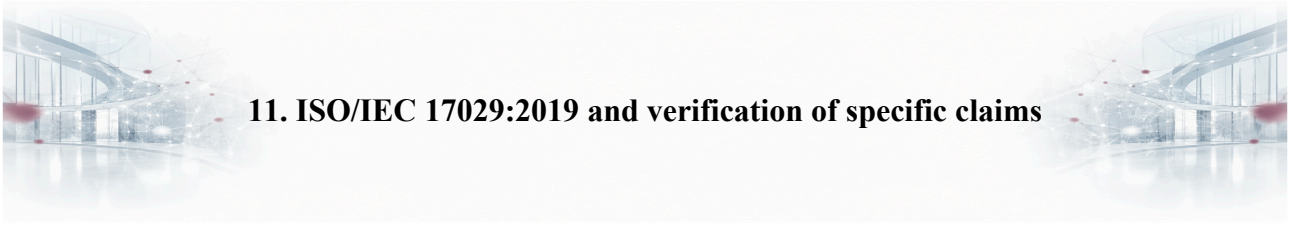
Recommendation: Account for uncertainty at the threshold

Hypothetical example: The pilot's working criterion requires at least 95 successful files out of 100; 94 received.

Analysis: The threshold has not been reached; rounding and subsequent change of rules conceals the outcome.

Practical steps: The analyst applies the established rule, shows the uncertainty and directs the system to refinement.

Outcome criterion: The decision meets the initial criterion without rounding in favour of admission.



11. ISO/IEC 17029:2019 and verification of specific claims

11.1. Certification, validation and approval language

The ISO/IEC 17029:2019 open record distinguishes between verification of the authenticity of the statement and confirmation of the veracity of the statement with respect to its intended future use (validation). The standard is addressed to conformity assessment bodies working on specially established programmes. It cannot be inferred that any independent software verification is already formal verification. [28]

For AI Association, it is rational to formulate individual cryims: "the percentage of unauthorized requests to secure data does not exceed the protocol value", "the model is tested for specific classes of vulnerability", "the selection of sanctions matches is formed without the inclusion of teaching examples". The earlier the declaration, the easier it is to set its limits and repeat the procedure.

11.2. Programmes as a source of claims

ISO/IEC 17029 does not replace the subject-matter program that determines the type of evidence, threshold and content of the conclusion. A report that promises to "verify transparency" should first determine what transparency means: accessibility of documentation, user's right to explain, reproducibility of input data, or open model architecture. These are different properties with different actors and legal consequences.

In the case of complex organizational complaints, it is helpful to combine documentary and empirical checks: check whether a notification process exists, and determine in control scenarios whether it is being implemented. But such a comprehensive evaluation must be stated directly in the program, otherwise the reader mistakenly attributes the result to nonexistent coverage.

11.3. Independence of reporting of results

It is essential to verify who chooses the wording of the public statement. If the client orders "confirmation that the product is ethical" but does not permit the investigation of complaints by affected users, the underlying statement should be narrowed down or withdrawn. The approval must be based on the volume of material being studied, not the marketing purpose.

For ARGA, it's reasonable to separate the preparation of a proof note from a formal verification statement: until there's a confirmed field of accreditation and program, let's just allow a research report with a description of the work and limitations. Institutional status cannot be created by the style of the document.

Practical checklist

1. Formalize the original statement without valuable slogans.

2. Tie it to the test program and the measurable properties.

3. To indicate whether the result gives an independent assurance or only an expert assessment.

Verification of declarations and suitability for future application

Confirmation that the stated numerical characteristic has been achieved on a particular set is different from an assessment of whether that characteristic is of sufficient value for actual use. In the first situation, the statement of past measurement is checked; in the second, a justification of suitability for the given purpose is required, where possible users, context of exploitation and severity of errors are important. This practical demarcation is compatible with the publicly described ISO/IEC 17029 guideline, although it does not replace reading the full text of the standard and the requirements of a particular program. [28]

The "model detects 99 per cent of dangerous cases" is not subject to serious verification without the identification of the hazardous case, the composition of the recruitment, the way the share is formed, the number of observations, the exceptions and the period. If the evaluator receives test data from the developer, the risk of leakage of examples into the learning environment and the possibility of independent reproduction should be assessed separately. Where direct transmission of the source array is not possible, independent control case generation and test journaling in a controlled environment are used.

It is useful to leave two outcomes in the report: whether the original thesis is confirmed on the tested materials and whether it is permissible to use it for a specific external decision. The second result may be negative even with the positive one. Such a division prevents a broad interpretation of the technically correct protocol.

Practical case analysis 17.11

Recommendation: Check the audited statement separately

Hypothetical example: The company states that all the sources of the reply were verified by a person.

Analysis: The signature on the outcome document does not show the verification of each reference.

Practical steps: The vic selects the answers, compares the links with the log and interviews the checkers.

Outcome criterion: The conclusion is limited to a proven procedure and a demonstrated percentage of evidence.

12. ISO/IEC 17011 and accreditation of conformity assessment bodies

12.1. Who and for what are accredited

ISO/IEC 17011:2017 covers the competence, consistency and impartiality of accreditation bodies that assess other conformity assessment bodies. Accreditation refers to the ability of an institution to carry out a specific category of work in a particular area, but does not transfer the right to all possible services of the institution. [24]

For the client, the difference between accrediting a laboratory, accrediting a certification body and formal recognition of individual schemes is of particular importance. Even a single national

accreditation body can guide these areas through different programmes and make different decisions on the fields of application.

12.2. Area and status as at date of circulation

The verification does not begin with a scanned copy of the evidence, but with the current register of the accrediting body. Usually you need to compare the number, full legal name, branch, area, standard and status: valid, limited, suspended, terminated. A leaked date doesn't always cover the issue when changing a scheme, but it can't be ignored.

It should be assessed whether the action extends to specific sites or remote units. If the test is carried out by a subcontractor, it is necessary to establish who has confirmed his competence and under whose responsibility the report is signed. Otherwise, the chain of evidence breaks down between the name of the certification group and the actual perpetrator.

12.3. Proportional use of accreditation

Accreditation increases trust in the verification process, but it does not rule out substantial errors in a particular report. The reviewer still evaluates the question, the design of the test and the logical power of the conclusion. If the body is accredited but its program does not verify the required AI characteristics, having a status does not correct the defect of the subject.

The lack of accreditation also does not mean that the research analysis is useless or illegal. It's important to be honest about the status of the work. ARGA's specific tasks allow for expert reviews of regulatory regimes and analysis of evidence submitted until it is issued for official certification.

Practical checklist

1. Carry out a direct registry review of the accreditation body.

2. Compare the declared service with the accreditation annex.

3. Segregate the competence of the performer from the evidence of the quality of the result.

Audit of the area of accreditation as at a specific date

When accredited, the competence of the body performing certain tasks is evaluated, not the consumer quality of any product with its mark. Regulation 765/2008 defines the European national accreditation infrastructure; ISO/IEC 17011 describes the competence of accreditation bodies. In practice, the main evidence is in the current public status and domain, not in a random copy of the evidence. [19,24]

For each certificate, the purchaser must document the official name of the body, the legal person, the address and the branches, the current accreditation number, the accurately covered schemes, restrictions and temporary suspensions. It is particularly important to distinguish between the status at the date of issue of the certificate and the date of verification. If there is a discrepancy, the date of registration should be kept and an explanation from the accreditation body should be sought, not the seller's commercial letter as final proof.

If the recognition is based on an international agreement, this is not enough to conclude that the result must be recognised by the local regulator. It is necessary to compare the procedure category, the area of the MRA in force, the special requirements of domestic law and the scope of addresses of the document. For the legal file, the formula "is accredited in such an area; the legal effects in the chosen jurisdiction are set separately" is more accurate than the promise of automatic cross-border force.

Practical case analysis 17.12

Recommendation: Check accreditation area

Hypothetical example: The laboratory is accredited for one type of test and offers another.

Analysis: The general status does not automatically apply to the new methodology.

Practical steps: The buyer checks the application with the area and requires a separate justification for the new procedure.

Outcome criterion: The contract distinguishes between accreditation work and outside work.

13. The global mutual recognition system and the 2026 transition

13.1. New Participant after IAF and ILAC

Since January 1, 2026, Global Accreditation Cooperation Incorporated has joined the institutional functions of the IAF and ILAC and launched its own multilateral recognition agreement (MRA). This fact was confirmed by an official ILAC message and a press release of the transition period. To designate IAF and ILAC as operating independent centers of the new global system without transitioning would be an outdated description. [37-38]

At the same time, the documents and stamps of the previous mechanisms retain their meaning within the framework of the transitional rules. To verify the actual certificate, it is necessary to take into account the date of its issue, the area of recognition and the current rules for the use of the mark. It is not permissible to mechanically declare all documents with the previous abbreviation invalid from January 2026.

13.2. Subject to international recognition

The meaning of MRA is to trust the competence of accrediting bodies and the results of accredited conformity assessment bodies within the recognised areas. It's not a global treaty that reverses national laws on the admission of AI systems to the market. Even an internationally recognised test report does not exempt a supplier from the specific EU notification procedure if it is legally required. [19,35-38]

A separate category of problems arises when comparing different patterns. Two bodies may be members of the same accreditation community, but test different properties, different versions and at different thresholds. Mutual recognition of the evaluator's competence does not mean the equivalence of the measured object.

13.3. Practical verification of certificate from third country

First, the authenticity and status of the certificate are checked. Then it is the area of accreditation of the body, not just of the accreditation institution. The next step is whether the MRA covers the relevant job category and whether the recognition standard applies in the destination country. At the end, a substantive comparison of criteria is performed: technical characteristics, human rights, data requirements and evidence form.

For an international financial control project, the exported AIMS certificate may be a factor in the supplier's assessment, but it does not replace the local assessment of data processing, sanctions lists and complaint procedures. This combination of international quality infrastructure and national requirements best reflects the real potential of cross-border AI Assurance.

Practical checklist

1. Consider the new Global Accreditation Cooperation Incorporated MRA from 01.01.2026.

2. Segregate recognition of accredited report and legal authorization for use of product.

3. Check the area and transitional rules for old markings.

Transition 2026: Continuity and Limitation of Recognition

Since January 1, 2026, Global Accreditation Cooperation Incorporated (Global ACI) has combined the functions of the international IAF and ILAC systems and began work with the common multilateral agreement of the MRA. This transition is confirmed by the reports of accreditation organizations and does not mean that all previous markings will disappear at once: there are transitional conditions for their use. When checking a 2025 document, it's important not to invalidate it just because of the previous sign, but also not to consider the sign as sufficient proof of its current status. [37,38]

International recognition should be verified for a specific activity. The results of laboratory tests, management certificates and product certificates can fall into different recognition areas. An accreditation body's involvement in a global organization alone does not guarantee that each of its programs is included in a recognized volume. The verifier shall specify the area of the body applicable to this scheme and the specific document on the date of issue and the date of assessment.

A separate boundary is required for legal analysis: an international trust infrastructure facilitates the adoption of results, but does not override the mandatory requirements of domestic law, public-law permits or the competence of EU bodies. Under procurement conditions, special provisions for accreditation status and its renewal are recommended, and public statements prohibit the expression "recognized in all countries" without specifying a specific legal mechanism.

Practical case analysis 17.13

Recommendation: Not to equate mutual recognition with authorization of application

Hypothetical example: The foreign report is accepted by the partner as a technical evidence.

Analysis: The adoption of the report does not resolve the issue of national product clearance.

Practical steps: The lawyer checks separately the usage and requirements of the local addressee.

Outcome criterion: The confession record does not contain an unreasonable operational permit.

14. The European accreditation framework: Regulation 765/2008 and EA MLA

14.1. EU Special Legal Architecture

Regulation (EC) No. 765/2008 forms the legal basis for EU accreditation. According to the official materials of European co-operation for Accreditation, Article 11 concerns the recognition of the equivalence of services of national bodies that have successfully passed an assessment in the EA system and the results of organisations accredited by them in the relevant field. [19,35-36]

For the user, this means a stronger legal context of recognition than a simple contractual reference to a foreign certificate. However, the act does not erase boundaries of competence and is not a

universal license to use any AI model. The legal effect is assessed together with the relevant sectoral legislation.

14.2. EA MLA Make and zone verification

In 2025, EA announced a special MLA mark; the 2026 updates related to expanding the registration of the mark. EA's materials indicate that its use should be linked to a national accreditation mark and not be outside the recognised area. A single image of a trademark on an advertising page is not enough if you cannot identify the actual area where the document was issued. [35-36]

The verification shall include the register of the national authority, the area of accreditation of the CAB and the category of services covered by the EA MLA. It is crucial for an organizational decision that this analysis be kept in the case as a separate record. Otherwise, in a year's time, it will be impossible to recover the recognition that existed at the date of purchase.

14.3. Limitations of recognition with respect to AI Act

Even the European accredited ISO/IEC 42001 certification body does not become an AI Act notified body automatically. Notification is carried out according to Article 28-31 of the Regulation and specific types of evaluation; accreditation can be used in the evidence base of competence. [9-12]

Consequently, the contract with the "European accredited partner" does not replace the NANDO/applicable official list of authorized bodies, their areas and specific procedure. Especially serious is the risk of mixing management certification with biometric system technical assessment.

Practical checklist

1. Establish a legal basis for a national accreditation body.

2. Check the existing EA MLA area specified for the type of evaluation.

3. A.I. Act notify separately.

III. European legal mechanisms

The territorial features of the European system

In the European Union, accreditation is linked to the legal structure of Regulation (EU) No 765/2008, which operates national accreditation bodies and a mechanism for mutual trust in results. Voluntary international recognition and the European EA MLA system are institutionally linked, but the legal implications must be assessed according to the relevant act and field of assessment. The mention of the EA MLA sign without verifying a specific area proves nothing about the legal status of an AI system. [19,35,36]

In European procurement, at least three independent documents are important: a confirmation of the body's current accreditation area, an assessment result document and a legal justification that the scheme applied is suitable for the relevant regime. If AI Act provides for internal control, obtaining a non-mandatory certificate does not create an additional legal obligation; if the law requires the participation of a competent notified body, normal voluntary certification does not close that requirement.

Cross-border comparisons should not describe the "EU recognizes certificate X" but the audited chain: the national accreditation authority, the EA MLA area, the authority of the assessment body, the type of result and the relevant EU rule. Such a statement is suitable for contracts, regulatory consultations and evidence analysis. It also makes it possible to quickly detect cases of use of genuine but non-relevant certificates.

Practical case analysis 17.14

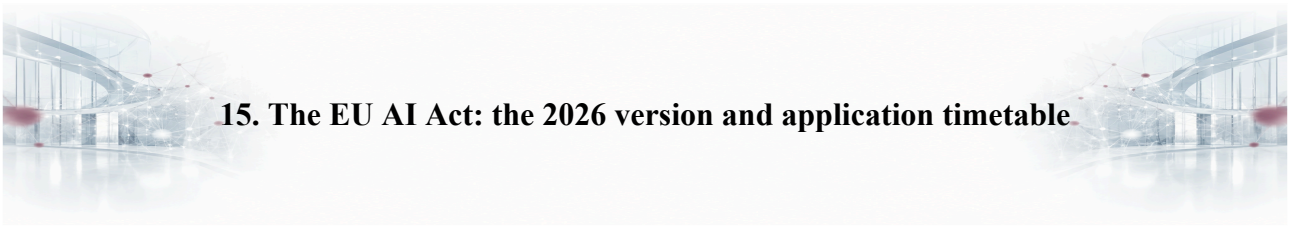
Recommendation: Check the specific area of European recognition

Hypothetical example: The supplier refers to the European network without showing the appropriate area of the agreement.

Analysis: Membership and coverage of a given result vary.

Practical steps: The expert shall verify the participant, the scope and the applicability to the type of document issued.

Outcome criterion: The reference to recognition is accompanied by a specific confirmation of coverage.



15. The EU AI Act: the 2026 version and application timetable

15.1. Consolidated text instead of historical presentations

In July 2026, Regulation (EU) 2026/1744 was adopted, amending certain provisions of AI Act, including the terms for applying requirements to high-risk systems. The official AI Act Service Desk shows the consolidated text for July 27, 2026, with changes to the rules; the general application date of August 2, 2026 is preserved, but should not apply to all duty groups. [1-3]

According to the updated Article 113, sections 1-3 of Chapter III for systems classified as high risk on the basis of Annex III apply from 2 December 2027; for systems under Article 6(1) and Annex I - from 2 August 2028. This calendar has a fundamental influence on the statement of readiness for compulsory conformity assessment and timing of the purchase of services.

15.2. Non-applicability of one time limit to the whole regulation

Some bans and general norms apply from February 2025, special provisions on GPAI and institutional chapters - from August 2025; individual updated bans will apply from December 2026. Therefore, the phrase "until 2028, AI has no obligations" is legally incorrect. GDPR, industry rules and treaty obligations also continue to operate. [1,3]

For working with a particular AI product, a date table is required: category, jurisdiction, role of person, presence of transitional rule, applicable norms and upcoming obligations. This table allows you to avoid two mistakes: to require a procedure that has not yet occurred as formally compulsory or to skip an already in place requirement.

15.3. Accreditation planning and late standardization

The date shift is linked, among other things, to delays in the readiness of harmonised standards and control bodies. It does not mean that the training of evaluators and developers is not needed until a new date. Organizations continue to design documentation, testing schemes, internal responsibility and accreditation infrastructure. [2-3]

For ARGA, it is advisable to use this period for analytical proposals on the composition of evidence files and the minimum language of certification restriction. Such proposals may be aimed at legal certainty but should not be presented as a binding standard approved by the European Commission.

Practical checklist

1. Check dates from 27.07.2026 instead of the original 2024 version.

2. Set the dates for Annex III and Annex I.

3. Do not conclude that there are no other existing obligations from the transfer of high-risk requirements.

Accuracy of dates and implications of AI Act reform

The date of publication of the act, the date of its entry into force and the date of application of the specific obligation are not identical. Regulation (EU) 2026/1744 amended AI Act from 27 July 2026. According to updated Article 113, the relevant provisions on high-risk systems in Annex III apply from 2 December 2027, and for the category of Annex I - from 2 August 2028. At the same time, individual sections, including some requirements for general purpose model suppliers, have different dates. The universal phrase "all AI Act has been in force since August 2026" misleads the addressee. [1-3]

Any table of duties should contain five characteristics: role of the party, territorial connection, specific category of system, article number and date of application. It is recommended to add the status of possible transitional rules and the subject matter of existing obligations. You can't build a consulting position on an outdated schedule, even if it's still in professional association presentations and consultant websites.

A real project can be subject to the GDPR, industry banking regulations and already applicable AI Act regulations at the same time, whereas some of the high-risk regime requirements have not yet begun to take effect. The postponement of one set of rules did not mean an exemption from the rest of the law. For ARGA, there is a fundamental separation between the current commitment, expected commitment and voluntary training measure; such statuses should be noted directly in each conclusion.

Practical case analysis 17.15

Recommendation: Maintain a calendar of responsibilities for each assignment

Hypothetical example: One product is used in several sectors with different regimes.

Analysis: One common date in the presentation may conceal the duties already in place.

Practical steps: The lawyer draws up a calendar for the role, purpose and applicable act; compares the date before external use.

Outcome criterion: Each duty has an entity, foundation and a verified date.

16. Classification of high-risk systems and selection of the assessment route

16.1. System category does not automatically define the route

Article 43 lays down different routes for systems falling within certain categories of Annex III and for systems constituting components or products in accordance with Annex I. It cannot be argued that every high-risk AI system requires the involvement of an outside notified body. For most categories listed in Annex III, paragraphs 2-8, the main route is the internal control procedure under Annex VI. [4-6]

For Annex III, paragraph 1, which concerns certain biometric applications, the choice between internal control and a notified body depends on the application of appropriate harmonised standards or general specifications. In the absence or incomplete application of these provisions, the

procedure provided for in Article 43 requires the involvement of a notified body taking into account special rules.

16.2. Interception with sectoral legislation

The Annex I systems follow the relevant procedures of the product legislation, including AI Act layer of requirements. After the 2026 amendments, the special rule of Article 43 (3) specifies that the very presence of a high-risk component of AI does not create an inevitable external assessment obligation if the applicable sectoral act provides for a legitimate internal pathway when the conditions are met. [1,4]

The qualification must take into account the product as a whole: a medical device, machine, transport system, digital component and its safety function may have different legal inputs. Combine the whole matrix to the question of "is there an IE needs an audit" means ignore the structure of EU harmonization law.

16.3. Pre-choice algorithm

First, the applicability of EU AI Act to the participant and the system is determined. It then checks whether or not it is prohibited or high-risk and on what grounds. The relevant sectoral law and the date of commencement of special duties shall be determined. Only after that Annex VI, Annex VII or the special procedure provided for in the sectoral legislation is selected. [1,4-6]

The result of the preliminary legal evaluation should be drawn up not as a final certificate, but as a map of the regulatory route with possible alternatives and open questions. If there is no material information, the decision remains conditional and requires verification by the manufacturer or competent authority.

Practical checklist

1. Clarify Annex I against Annex III rather than limit itself to the word "high-risk".

2. Check the specific biometric branch and the availability of applicable harmonized standards.

3. Set up a special sector route for a combined product.

Matrix C. Assessment routes in the EU

Scenario	Probable route	What else to find out?
Annex III, paragraphs 2-8	Annex VI, internal control [4-5]	Effective date; supplier role
Annex III, item 1, biometrics	Annex VI or VII under the terms of article 43	Availability and application of standards or general specifications
Annex I, sector article	Sectoral procedure with AI Act requirements	Product specific mode and option without third party
GPAI Model as such	Special chapter V	Are there systemic risks; do not mix with Annex VII

Scenario	Probable route	What else to find out?
Normal voluntary AI service	Selection of type of assignment by purpose	Industry law and data regime

Two-step decision on category and procedure

The first step is the system's function and impact profile. It should be established whether it is related to a product regulated in Annex I or to categories listed in Annex III, the exceptions should be taken into account and the actual supplier of the system should be indicated. The second stage is to define the assessment procedure in accordance with Article 43, including special rules for biometric category and product legislation. Missing the first stage often results in the wrong choice of expensive third-party procedures where the law requires internal control, or a reverse error. [1,4]

For Annex III systems in areas 2-8, Article 43 provides for internal control without a notified body. Annex III, paragraph 1, may be chosen depending on the application of harmonised standards or general specifications, and Annex VII is required under certain conditions. A separate route applies to systems covered by the sectoral acts of Section A of Annex I. Therefore, the simple question of "high risk or no risk?" does not define the procurement of the certification service.

The preliminary map must contain competing classification grounds and an explanation of why the alternative route has been rejected. This is particularly important for systems that combine biometrics, human resources solutions, and regulated equipment. In the final legal conclusion, it is permissible to indicate that the procedure has not yet been established on incomplete information. Replacing that conclusion with a commercial convenience creates a greater risk than a fairly marked gap.

Practical case analysis 17.16

Recommendation: To justify classification before selection of procedure

Hypothetical example: The system for preliminary searching of documents is called a high-risk one without a destination analysis.

Analysis: Technical complexity does not in itself determine the route.

Practical steps: The lawyer describes the actual application and verifies the classification criteria according to the sources of the report.

Outcome criterion: There is a reasoned decision in the file on the route and open questions.



17. Annex VI: internal control without a notified body

17.1. Contents of the internal procedure

Appendix VI provides for the verification by the quality management system supplier under Article 17, analysis of technical documentation for compliance with applicable requirements and comparison of development processes and post-market monitoring with the recorded documents. This procedure is called conformity assessment, despite the absence of a mandatory independent external evaluator on the relevant route. [5,17-18]

Internal quality control may involve an external consultant or laboratory, but the supplier's legal responsibility does not disappear. Contractual assistance does not change who makes the EU declaration and is responsible for the statements made therein.

17.2. Internal control and external evidence

The supplier may use laboratory testing, reviews and audits of third-party organizations to strengthen its own file. But these documents should be evaluated by subject matter. If a third-party appraiser has verified the quality of the data set, it cannot be considered as automatically verified incident management or information security. [5,18,27]

In the event of a subsequent supervision request, it is important to restore the evidence base: model version, date, conditions, results, internal inference and corrective action. For this purpose, it is reasonable to maintain a single matrix of requirements and evidence rather than a collection of discrete presentations from the supplier.

17.3. Limit of self-expression

Internal evaluation is not equal to recognition by a public authority. However, it does not automatically become less legal because of the absence of an independent organization's press, if the law chooses such a route. The legal adequacy of the procedure and the substantive conviction of the evidence in a particular dispute must be divided.

The ARGA approach should be form-neutral: check whether the procedure meets the required route and assess separately the risk of false self-confirmation. The imposition of a paid external certificate for each situation is not justified by the text of AI Act.

Practical checklist

1. Assess the consistency of documentation between Annex IV and Article 17 management.

2. Save the trail of actual checks, not just the final declaration.

3. Reflect the supplier's own responsibility.

Internal control as a full system of evidence

The internal procedure of Annex VI is not equal to the lack of evaluation. Responsibility for confirming compliance lies with the supplier; the relevant documentation should restore criteria, quality management system, verification results, technical specifications and the basis for declaring conformity. The customer who is only presented with a promotional presentation does not receive enough information to understand the completeness of the internal procedure. [5,17,18]

An external auditor may assist the supplier in examining evidence, but his or her voluntary report does not make Annex VI a compulsory third party certification procedure. Regulatory correspondence should indicate the legal status of each element. If a specialist has identified a discrepancy, it is necessary to distinguish between a methodological remark, a legal violation and a situation where the information is insufficient. Otherwise, internal control can become a set of formal markings without a suitable basis for resolution.

The ARGA practice programme proposes a preliminary audit of the pilot regulatory audit model: an independent team receives a system register, a demand map, a sample of technical materials and a gap management log. The result is a list of confirmed and unconfirmed theses, not an ARGA certificate. This makes it possible to position the research service fairly and not confuse it with the status of a notified body.

Practical case analysis 17.17

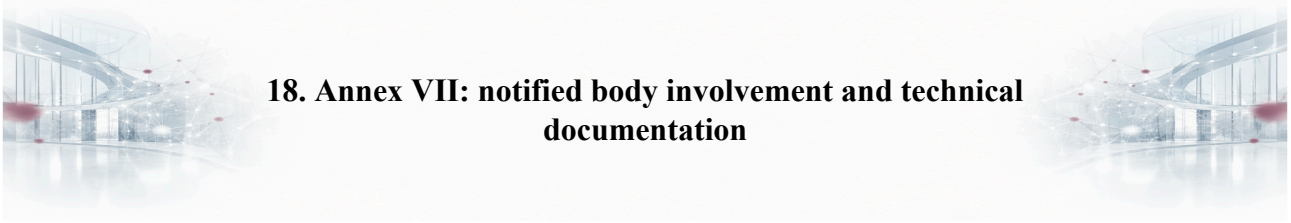
Recommendation: Document internal control

Hypothetical example: The supplier considers the absence of an external body to be unnecessary to maintain a file.

Analysis: The internal route does not eliminate the need to justify the procedure.

Practical steps: The person responsible shall collect the applicable requirements, tests, approval records and documentation versions.

Outcome criterion: The reviewer restores the basis for the internal decision on the file.



18. Annex VII: notified body involvement and technical documentation

18.1. Two parts of the procedure

Annex VII provides for the evaluation of the quality management system and the technical documentation of a particular system, including subsequent supervision. Not only are the formally accepted instructions important, but also the right of the competent evaluator to obtain the necessary information, request additional tests and justify refusal. [6,17-18]

The content of the procedure differs significantly from the voluntary certification of AIMS. When checking under Annex VII, the subject matter, the volume of access, the circle of applicants and the form of the result are determined by a special regulatory act. It's impossible to turn a typical report on the maturity of corporate processes into a document of this regime in retrospect.

18.2. Access to data, models and commercial secrets

Appendix VII provides for the necessary access to training, validation and test data; in exceptional cases - to the trained model and relevant parameters. At the same time, safeguards on intellectual property and commercial secrecy are taken into account. This illustrates why quality evaluation is not limited to reading advertising descriptions. [6]

For an external research contract, a similar principle is implemented through a coordinated access matrix: documents, APIs, anonymized corpus, logs and controlled environment. The obligation to provide more data than is justified for the legitimate purpose should be avoided, and an audit should not be declared legally impossible in the case of a reasonable technical alternative.

18.3. Decision, oversight and significant changes

Annex VII describes the reasoned decision and, with a positive result, an appropriate evidence of evaluation of the technical documentation; the quality management system continues to be monitored. Subsequent changes to the product or control system may require further evaluation. [6,13]

From a practical point of view, this means that there is a need for a version register. If the model has been updated in new sources or the decision-making module has been changed, the owner should assess whether the change affects previously confirmed properties. Without such analysis, the public reference to the original evidence may be incomplete.

Practical checklist

1. To distinguish quality system evaluation and technical documentation verification.

2. Record a legitimate model of access to closed data and parameters.

3. Establish review criteria with significant changes.

Control of documentation with the participation of a notified body

The Annex VII procedure includes the assessment of the quality management system and the assessment of technical documentation with the participation of a notified body in the specified volume. For AI, in practice, there are issues of access to source data, critical models, test infrastructure, and trade secrets. In the absence of a prepared access protocol, even the competent party may not receive sufficient material. Therefore, before the procedure begins, it is necessary to agree on a safe environment, a list of available artifacts, and copying limits. [6,12,18]

The protection of the secret and the right to independent evaluation are not mutually exclusive. Part of the verification may be done with controlled access, unidentified downloads, replicable tests and independent storage of control amounts. However, if the restrictions do not permit confirmation of the mandatory requirement, the fair result is not a "all reservation" certificate, but a request for additional data or the application of the prescribed waiver, suspension and re-evaluation procedures.

Separate control requires change after the original decision. If the algorithm, the composition of the learning data, the product's designation or human integration have been updated, it is necessary to determine the significance of the change and the terms of the new assessment. A register of such decisions is useful to include in the technical documentation package: it shows who has declared the change irrelevant, on what grounds and with what monitoring obligations.

Practical case analysis 17.18

Recommendation: Agree on the scope of external evaluation

Hypothetical example: The contract indicates the external authority but does not specify the versions and modifications of the system.

Analysis: It is not clear what object should be evaluated and what will happen after renewal.

Practical steps: The employer shall record the object, access, evidence and communication of the changes.

Outcome criterion: The outcome document is part of an agreed version and procedure.



19. Notified bodies: designation, competence and oversight

19.1. Nominating authority and independent competence

Each Member State shall designate or establish a competent notifying body responsible for the procedures for evaluating, appointing, reporting and monitoring conformity assessment bodies under AI Act. The State may entrust part of the assessment to the national accrediting body under Reg 765/2008, but this does not eliminate the special procedure for public recognition. [9]

The requirements of Article 31 concern organizational independence, confidentiality, resources, quality of processes and competence. In particular, the provision of advice that can challenge the impartiality of an assessment poses a special problem. Therefore, the model "we have developed a product and then as an independent body has certified it" cannot be considered neutral without an analysis of conflict of interest. [12]

19.2. Accreditation as evidence, not a synonym of notification

According to Article 29, the notification application shall be accompanied by the necessary confirmation of competence. Article 30 provides for the procedure for communicating to the Commission and the Member States; the existence of an accreditation certificate affects the process, but other conditions apply in the case of alternative documentary confirmation. Therefore, the issue is decided by the specific status of the notification, not by any foreign accreditation mark. [10-11]

When the candidate is checked, the customer must establish his authority with regard to the system and type of procedure. The general reference to "notified body" may refer to another sector act or to another type of product. For AI Act, it is the relevant area that needs to be documented.

19.3. Appeals and information

Article 44 provides for the rules of application of certificates of notified bodies and the availability of procedures for appealing against their decisions. Article 45 requires certain reports on issued, rejected, restricted, suspended and withdrawn documents. This shows that the certificate is not a permanent immunity certificate, but an element of ongoing public-law control. [13-14]

The interested party must be able to distinguish the current outcome from the historical record. For contracts, it is recommended that a separate requirement be immediately notified to the customer of the restriction, suspension or change in the area of confirmation.

Practical checklist

1. Checking the nominating authority and the exact area of recognition of the authority.

2. Find out if there's a consultative conflict of interest.

3. Clarify the status of the certificate and the mechanisms for notification of changes.

Competence of the notified body and issues of trust

AI Act notification is based on a special public legal procedure. It is required to establish the competent notifying authority, the declared and effectively provided volume of functions, as well as compliance with requirements for conformity assessment bodies. The ISO/IEC 42001 certificate and the accreditation of the standard laboratory do not automatically confirm the right of the notified body to act for another procedure. For a particular product, you need an official source of status and area, not just a letter from the service provider. [9-12]

When the appointment of an authority is based on previous powers in sectoral harmonisation, it is necessary to check transitional characteristics and limits of the new competence. For decisions relating to EU law, the assessment should take into account the 2026 text as amended, including the procedural provisions of Article 43, rather than being based on an old brochure. In case of doubt, the public data of the competent authority and the official text of the norm shall have priority. [1,4]

Independence also implies the ability to make adverse decisions. If the evaluation service provider's commercial model is focused solely on providing positive conclusions, confidence in the procedure is diminished. Procurement documents should require a mechanism for withdrawing an expert, rules for considering conflicts of interest, reviewing decisions and complaints. However, the result of the legal assessment should not replace the competence of the notified body itself.

Practical case analysis 17.19

Recommendation: Check the designation of the body for the appropriate object

Hypothetical example: The perpetrator shall accord powers in respect of another product.

Analysis: The status in the neighbouring sphere does not confirm the authority to carry out the required work.

Practical steps: The lawyer checks the appointment record and the area with the subject matter of the contract.

Outcome criterion: The discrepancy has been corrected prior to the procedure or another perpetrator has been selected.



20. Harmonised standards, common specifications and presumptions

20.1. International Standard and Harmonization of the EU

The publication of an international ISO/IEC standard does not automatically imply a presumption of conformity under AI Act. Article 40 associates such an effect with a harmonised standard or part thereof, the reference to which has been published in the Official log of the EU, and only within the requirements covered. [7]

The practical question consists of three parts: whether there is a link in the official publication, to what requirement it relates and whether the chosen implementation covers the specific function of the system. The certification tool should make this connection transparent; the general statement "is in line with ISO standards, so it is in line with AI Act" is excessive.

20.2. General specifications for article 41

Article 41 authorises the Commission, in the circumstances provided for, to adopt general specifications where harmonised standards are absent, inadequate or do not address the task required. Such specifications may give rise to a presumption provided for by law in the relevant part. Where an alternative solution is used, the supplier shall justify an equivalent level of compliance with the applicable requirements. [8]

Thus, general specifications are not an arbitrary list of "expert recommendations" but a special act in the established procedure. They can't be replaced by a private certification scheme just because that scheme uses the same words: transparency, security or human control.

20.3. Map of standards and claims

For purchase and legal conclusion purposes, a two-dimensional table is useful: lines - legal obligations, columns - technical norms, test methods, documentary artifacts and exceptions. Each cell is assigned "cover", "partly", "unconfirmed", "not applicable". This prevents an error when one common certificate is used to close the entire regulatory list.

The date of verification of the harmonized status should be further recorded. Standards are evolving, and the publication of a new edition of ISO and the publication of references to it in the Official log of the EU do not necessarily take place simultaneously. With the active standardization of 2026, this date is of practical importance.

Practical checklist

1. It is in the Official log of the EU that we search for a record of harmonization.

2. Specify the scope of the presumption on a specific demand.

3. Not to present a private scheme as the general specifications of article 41.

Standardization as evidence of compliance with a specific requirement

International designation of ISO/IEC and harmonisation of the standard in EU law describe different status. ISO develops and publishes standards; the presumption of compliance with the requirements of a special EU act depends on the conditions of the relevant standard, including the official publication of the reference to the harmonised standard in the Official log of the EU and the actual coverage of the requirement. Article 40 of AI Act does not automatically turn any AI standard into a legal presumption tool. [7,20,21]

The practical coverage table should link each mandatory requirement to a specific technical document, verified version of the standard, individual test and result. If the standardization does not close the issue of explainability, the relevant gap is not eliminated by reference to the general management system certification. The supplier may demonstrate compliance by other means, subject to applicable law, but the quality of the evidence and the permissible route of the procedure should be determined by law and not by marketing proposal.

The general specifications of Article 41 require separate analysis: their legal effect and conditions of application do not amount to the status of a voluntary standard. When drawing up a conclusion, it should be indicated whether the published harmonised standards, general specifications, private schemes or authorised methods are used. If the standard is still in the FDIS stage, a reference to it may have a predictive or methodical meaning, but its unpublished status should not be disguised.

Practical case analysis 17.20

Recommendation: Link the standard to a specific requirement

Hypothetical example: The conformity map contains one standard in contrast to all requirements.

Analysis: The general reference hides the unveiled properties and differences in the editing.

Practical steps: The expert shall set the requirements for the evidence and note the gaps in the coating.

Outcome criterion: Each essential requirement has a reasonable connection to the evidence.

21. EU declarations, certificates and CE marking

21.1. Miscellaneous documents in the same procedure

Article 47 of AI Act provides for an EU declaration of conformity to be drawn up by the supplier for the relevant high-risk system; it provides for the content of the declaration and the storage of the document. Article 48 regulates the CE marking, including the digital form and the indication of the identification number of the notified body where applicable. [15-16,55]

Neither the declaration nor the CE markings are synonymous with the ISO/IEC 42001 certificate. The declaration expresses the supplier's regulatory responsibility; a certificate issued by a notified body under a special procedure has other grounds; the CE sign indicates that it complies with the established regime, not that there is no risk in any application.

21.2. Examination of the substance of the document

The declaration must identify the model, supplier, applicable legislation and the grounds used for confirmation. If the document cannot be linked to the system version, it is of little use in the event of a dispute about the results. It is particularly risky when the declaration relates to a different model or modified package, but is presented as a general document for a new product. [15,55]

The verifier must compare the declaration with the technical file, the risk assessment logic and the post-release changes. If a notified body is present, its name, identifier and evidence relating to the product are checked. Not only is the existence of paper important, but it also has an active object.

21.3. Use of label in marketing

The supplier shall not transfer the CE marking value to properties not covered by the regulatory regime. For example, a mark attached to one product does not prove the accuracy of the legal assistant of the same company. Similarly, the certification body's accreditation mark is not a consumer quality marking of any model.

Internal verification of public statements should respond to: "what document", "to which object", "what regime", "to which date" and "in what configuration". These five questions reveal most of the unwarranted transitions from private evidence to a global security promise.

Practical checklist

1. Do not confuse the supplier's declaration, notified body certificate and AIMS certificate.

2. Check model, version, regulatory basis and validity period.

3. To verify the validity of the use of signs only in actual volume.

Chain from technical file to CE marking

The CE marking is a legal signal of the responsibility of the entity concerned to comply with the applicable requirements, rather than a sign of "unforgivable" model behaviour. The EU declaration of conformity, the assessment documents and the certificate of the notified body, if provided for, have different functions. The procurement system should establish who signed the declaration, which product and version is described, what acts are named and whether there is a prescribed documentation storage procedure. [15,16,55]

When combining several harmonisation regimes, it may be necessary to consider the requirements of different acts applicable to one system. It cannot be concluded that the presence of a CE marking on regulated equipment automatically certifies the quality of an individually integrated generator. All legally significant configurations and individual requirements for AI used are checked. Changes to the intended application may affect both the obligations of the supplier and the document confirming conformity.

For an international analytical report, a negative test is useful: What exactly remains unproven even when a document is presented? These can be non-binding ethical statements, accuracy in a particular subgroup, compliance with foreign national law or relevance of the version after a substantial update. A list of such gaps is needed not to discredit CE, but to accurately determine the legal and probative force of a particular result.

Practical case analysis 17.21

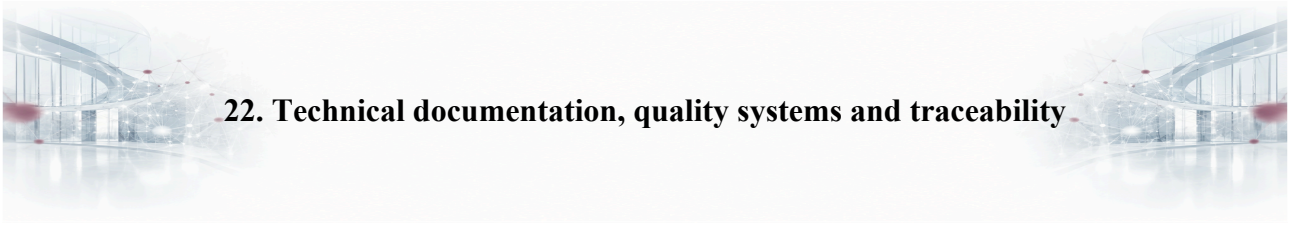
Recommendation: Agree on declaration and technical facility

Hypothetical example: The declaration describes version A, the actual delivery of version B.

Analysis: The matching trade name doesn't confirm the identity of the object.

Practical steps: The responsible person checks the identifiers, changes and documents prior to the external statement.

Outcome criterion: The markings and accompanying information refer to the actual object.



22. Technical documentation, quality systems and traceability

22.1. Documentary file architecture

Appendix IV of AI Act lists the contents of the technical documentation with regard to applicability: system designation, versions, interacting components, infrastructure used, development, testing and other information. Article 17 deals with the quality management system. These two blocks are interconnected, but not interchangeable: product documentation describes the object, and organizational procedures show how it is controlled throughout the life cycle. [17-18]

The composition of the case file should be fixed until the first test. If the evidence is collected only after the problem has been detected, there is a risk of selective preservation of favorable results. For each artifact, the source, date, responsible person, version and authorized disclosure mode are indicated.

22.2. Origin and independence of the record

Training data is required within the scope of applicable legal grounds and actual access. A reliable assessment requires not just the word "quality", but information about the origin, legality, suitability, representativeness and updating of the data. In some cases, commercial secrecy limits public disclosure, but that doesn't justify a lack of internal control and evaluation.

For the data of sanctions or police matches, an authoritative source of recording, extracted copy, normalization and algorithmic assumption of a person's identity must be separated. An error at one stage can lead to legal consequences that are not corrected by a high average metric at other stages.

22.3. Duration and protection of evidence

Article 18 of AI Act establishes a special rule for the storage of certain types of documents for a fixed period after placing on the market or commissioning of a high-risk system. However, the transfer of personal data, the retention period of a specific workflow and the auditor's access are also subject to independent data protection and confidentiality requirements. [45]

The permanent obligation to preserve regulatory technical documentation should be separated from the unjustified storage of personal data from the test body. Pseudonymisation, limited archives and verifiable copy-deleting rules are permitted while preserving the necessary evidence.

Practical checklist

1. Link each allegation to a specific element of the case file and its version.

2. Clear the control period of storage and personal data regime.

3. Trace the origin of information from external lists and bases.

IV. Technical and evidence-based verification

Storage time, version identification and evidence trail

The technical file should allow the system to be restored to the state for which a legal or engineering statement has been made. It's not enough for a file called the model. It is reasonable to include a description of the purpose, a component register, interface configuration, data origin, results handling rules, test environment information, and a correction table for known defects in the control package. When changing the external model, the exact identifier of the installed version and its availability conditions should be kept. [17,18,45]

Since documents can contain both commercial secrecy and personal data, separate access to primary materials and their evidence metadata is needed. The report may store control amounts and references to a closed repository instead of a full copy of the array; the reviewer must understand how to access and the possibility of independent re-verification. To comply with GDPR and industry regulations, they set the purpose of processing, the range of persons, the basis for transmission and the storage period for a specific type of record, without claiming that AI Act automatically allows any use of data.

A separate step is the qualification of a significant change. Not every replacement of a text label requires a completely new procedure, but changing the model, data or application area may violate the prerequisites of the previous verification. The decision on the need for a new evaluation should be documented by the responsible person with a comparison of the original and new risks and not left to the product owner or API supplier alone.

Practical case analysis 17.22

Recommendation: Maintain the traceability of the case file

Hypothetical example: The report contains links to files that are rewritten by the new version.

Analysis: The reviewer cannot restore the material used in the decision.

Practical steps: The custodian records the versions, the integrity control and the access time; it keeps the original records.

Outcome criterion: Each substantive allegation relates to an accessible, constant version of the material.



23. Technical test design and the limits of statistics

23.1. Test design before result

The test plan shall begin with the determination of the parameter to be measured. For the share of correct links, it is necessary to specify whether all links are considered to be only substantial links or a file without a single false substantial link. These parameters have different denominators and different meanings for the user. Then the set, the selection, the benchmark, the rules of refusal and passing, the method of assessing uncertainty and the solution at the threshold are recorded. The change in definition after viewing the result requires a new version of the protocol.

The configuration set, the independent final set, and the diagnostic scenarios have different functions. The first helps to develop the system and the criteria; the second allows for a predetermined evaluation; the third identifies ways to refuse. The results of the diagnostic set do not replace the assessment of the prevalence of errors. If the test is included in a formal conformity scheme, the rules of the scheme are established separately and have priority over the author's example of the report.

23.2. Precion, recall and legal price of errors

In the search for matches, false positive identification and false negative mismatch can have different consequences. The statement "accuracy 99 per cent" is poorly informed with a strong class imbalance: a model that always meets "no matches" can get a high overall share of the correct answers and is not suitable for the purpose. We need a error matrix, a threshold rule, and a characteristic of the basic prevalence of the class.

In law enforcement and finance, there is a requirement to record actions after the flag of the system: is it just a recommendation to an employee, temporary blocking, a message to an authority or a refusal to a client? Even the same probability of error can be acceptable or unacceptable depending on the nature and reversibility of the solution.

23.3. Intervals and zero observations

If a conditional independent sample of 300 cases does not reveal a single hazardous failure, it does not imply a proven zero risk. In the simplified binomial model, the upper 95 percent probability limit is 1 - 0.05^(1/300), that is, about 0.99%. For a rare event, this is an important limitation even in a seemingly impeccable trial.

The model of independent equally distributed tests may not be executed for a series of linked requests to a single generating system. The formal number should not then be published without reservation. For tests where one person may be seriously injured, statistics should be supplemented by stress scenarios and a legal analysis of the intervention procedure.

Rule of decision at border impact

In the conditional scheme, the criteria for admission are set in advance: the lower bound of Wilson's 95%-two-way interval for the success rate of the successful case files must be at least 90%. With 94 out of 100 successes, the interval is about 87.5-97.2%, so the data for the positive solution is not enough. This result does not prove that the true success rate is below 90 per cent. The interval does not automatically reflect the non-representativeness of the set, incorrect marking and dependence on the file; they are evaluated separately. The threshold and the rule are the conditions of the training scheme, not the requirements of ISO or AI Act. [M2]

Practical checklist

1. Publication of the error matrix and denominators, not one percent success.

2. Note the probability of selection of scenarios and possible systematic bias.

3. Clearly distinguish between model results and human outcomes.

Test report with predetermined thresholds

The threshold decision shall be checked in two planes. First, determine the correct calculation for the selected observational model. They then examine the consistency of the conclusion to justified marking changes, taking into account passes and excluding dependent repetitions. If a positive result only appears after removing uncomfortable observations, it requires separate justification. The plan for the additional data set is defined in advance: a sequential addition of cases to the first favorable result without a proper statistical design does not retain the meaning of the original procedure.

With multiple variations of the same name, repeats of the same client or tasks from the same file, the research has a cluster structure. The evidence can't be estimated just by the number of lines. The protocol shall specify an independent unit, the number of such units and the manner in which the investment is accounted for. If the initial report does not restore dependence on observations, an interval from it is not accepted as a sufficient basis for positive assurance without further verification.

Stress tests should explore the mechanism of failure: delay in receiving new information, conflict between registers, lack of a unique identifier, unexpected symbols, external API failure, and an incorrect human response to a system recommendation. One quality indicator does not adequately reflect the combination of these risks. The result is presented as a set of metrics, classified errors, and wording about what the study did not cover. If the results cannot be replicated from the preserved materials, the validity of the assurance is limited regardless of the beauty of the final percentage.

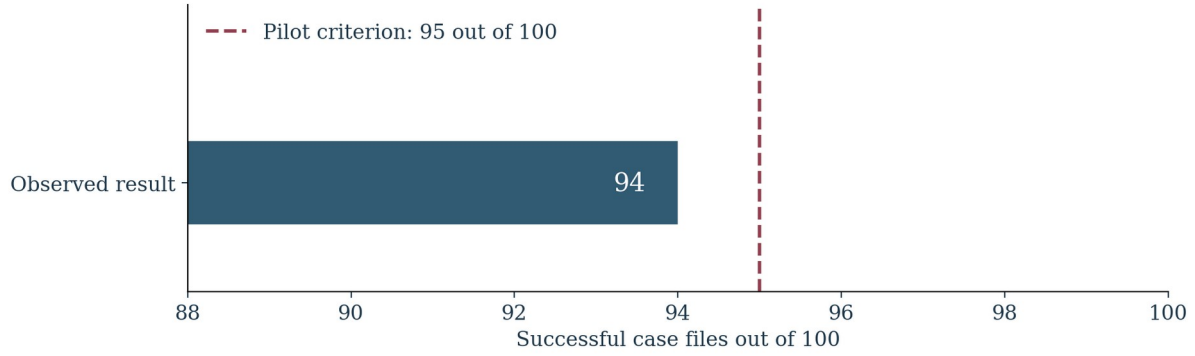


Figure 17.3. A borderline result and a predetermined criterion. Hypothetical example from Practical Case Analysis 17.10. The chart highlights a difference of one case file; the horizontal scale starts at 88. The threshold is not changed after testing.

Practical case analysis 17.23

Recommendation: To approve the test plan in advance

Hypothetical example: After a weak result, the most complex files were eliminated.

Analysis: The indicator refers to a different population and does not respond to the underlying question.

Practical steps: The statistician maintains the initial calculation and analyses the change in the set separately.

Outcome criterion: All exceptions and the influence of each on the decision are shown.

24. Bias, fundamental rights and impact assessment

24.1. Group differences require substantive interpretation

The difference in average metrics between groups may reflect the characteristics of the test data, system configuration, uneven quality of primary records or discriminatory effects. There is no universal measure of "justice". It is necessary to determine which groups are legally and effectively relevant, which characteristics are acceptable and which errors should be minimized. [29,39,44]

When considering credit or personnel decisions, it is particularly important to distinguish between statistical correlations and conclusions about legal discrimination. The latter requires checks on applicable law, context, possible justifications and procedural guarantees. The technical report is evidence of some of the circumstances, but it does not replace a qualified legal conclusion.

24.2. Impact assessment and certificate limits

ISO/IEC 42005:2025 is devoted to assessing the impact of AI systems on people, groups and society throughout the life cycle. The approach is useful for identifying potential negative consequences, but the open ISO record does not provide grounds to argue that the standard itself is a certificate of compliance with all human rights. [29]

Article 27 of AI Act provides for a special assessment of the impact on fundamental rights for certain participants and situations with their own conditions of application. Organizational assessment under international guidance and the legally required process should be considered separately, although information between them may be used in compliance with law. [51]

24.3. Engagement of affected parties

A reliable assessment of the impact is not possible only on the basis of the marketing instructions of the supplier. The practical use, complaints, possibility of human review and whether the persons concerned understand the reasons for the adverse decision must be taken into account. Where direct contact is not possible, this limitation should be indicated and the permissible indirect evidence used.

For ARGA, the principle of not automatically transferring a questionable record from one database to another is particularly important. A person must have a certain way of challenging, and quality checks must distinguish between identity error, interpretation error, and legal characterisation error.

Practical checklist

1. Select indicators that take into account the consequences of errors for different groups.

2. Separate technical statistics from the legal opinion on discrimination.

3. Check the reality of channels of complaint and human review.

Basic rights and the quality of the remedy

The assessment of bias requires an assessment of which individuals and rights affect the system. For employment decisions, access to banking services and international police exchanges, the same difference in metrics can have different consequences. Therefore, statistical tests should be supplemented by a map of the decisions taken, the degree of reversible harm, the notification procedure and the availability of effective human verification. The simple equality of averages does not preclude systematic error in rare but critical cases. [1,39,44,51]

When choosing comparison groups, the right to personal data protection and restrictions on sensitive information collection must be respected. The absence of admissible group data does not mean that there is no discrimination; the impossibility of measurement and alternative control methods available should be described. If the group is small, a narrow numerical result without intervals can create false confidence. For practical conclusions, it is important to distinguish between the observed discrepancy, legal characterisations and the conclusion that corrective action is necessary.

The complaints procedure should be in place. The existence of a "appeal" button does not prove an independent review if the same algorithm automatically adopts the original conclusion. It is necessary to check the availability of new data, the competence of the verifier, the timing, the log of corrections and the availability of a reasoned response. For ARGA, it is particularly important to check the ability to correct the original identification mark, not just to recalculate the rating on an incorrect original record.

Practical case analysis 17.24

Recommendation: Check the impact on affected persons

Hypothetical example: The system is more likely to reject incomplete digital applications.

Analysis: The problem may relate to the availability of the procedure and not only to the average accuracy.

Practical steps: The team analyzes the grounds for refusal and the available manual review; they study groups of comparable applications separately.

Outcome criterion: A meaningful review of the underlying information is available for disputed refusal.

25. Robustness, cybersecurity and external dependencies

25.1. Sustainability to reasonably predictable conditions

The stability of an AI system depends on the quality of input, language, document layout, noise, environmental changes and attempts to intentionally cause an error. Therefore, the test on "model" documents is not sufficient for a service that applies to multilingual names, damaged files and unstructured archives. We need a threat model and a documented range of conditions. [39-41]

When comparing sustainability, it is important to note not only the proportion of errors but also the nature of degradation: whether the system can reject the answer, transmit the result to a person,

show uncertainty or silently produce a false fact statement. For legal and financial purposes, a safe refusal is sometimes preferable to a confident invention.

25.2. Prompt injection and the tool chain

A generative assistant who reads external documents and calls in software tools can take hostile text from the source as a command. This risk is not only about the quality of the language model, but also the settings of permissions, the priorities of instructions, content filtering, journaling and human verification.

In the audit case, there should be a clear scheme of trusted and untrustworthy sources. The tests are conducted in an isolated environment with a consistent volume; the results of the attack test are described as identified configuration features, not as universal proof of ability to disrupt any system.

25.3. Supplier, cloud and update controls

With an external API, the main developer can change the model or infrastructure. The integrator must establish a contractual level of control: version notifications, a log of incompatible changes, disconnection mechanisms and a turnover plan. Without this, the initial report can only be valid for a short technical condition.

Particular attention is paid to cross-border data transmission, access rights to client data and restrictions on the use of input documents for training. Even a certified vendor management system does not allow to assume that a particular contract provides the necessary privacy regime.

Practical checklist

1. Form a threat model that takes into account linguistic and contextal variations.

2. To test the possibility of safe failure and restriction of the instruments.

3. It is contracted to determine the consequences of a hidden change in the external model.

Attacks on the model and dependence on infrastructure

For AI Assurance, the security of the interface is insufficient: you have to evaluate the data loading chain, the push-ups, the vector search, the agent tools, the external plugins and the access permissions. The vulnerability can arise when a user instruction is mixed with an unverified fragment of an external document. In that case, the technically correct model fulfils the false working context. The subject of the study is not abstract "protection from all attacks", but a pre-defined profile of threats and evidence of what attacks the system detects, limits and documents. [39,40,43]

Experiments should be repeated in realistic scenarios: the introduction of hidden instructions into a document, an attempt to extract client data, mistakenly granting of rights to create an email, contamination of the search index and disruption of communication with the external model. It is critical not to describe dangerous techniques as universal recipes, but to assess the presence of protective layers and controlled behavior of the system. The expert should specify the version used, the number of scenarios, and the remaining vulnerabilities, because changing the model and protecting it quickly changes the results.

An information security management certificate or a separate penetration check may be part of the evidence, but it does not replace a probability-based problem study. For sensitive legal transactions, it is important to limit the autonomous execution of an action: document preparation, human verification and sending must have different rights and logs. This is particularly important for agency tools with access to mail, files and financial transactions.

Practical case analysis 17.25

Recommendation: Check external relationships

Hypothetical example: After the upgrade of the libraries, the assistant stopped properly processing the investments.

Analysis: The previous tests do not cover the changed processing chain.

Practical steps: The engineer records the dependencies and performs regression on the saved and new permissible files.

Outcome criterion: The file contains a version of the relationship and the results of the re-check.



26. Generative AI, general-purpose models and assessment issues

26.1. From model to application

A universal language model can show high performance on a standard benchmark and at the same time make mistakes in a contract with an unusual legal context. Therefore, the evaluation of the basic model and the evaluation of the finished product should be divided. The second requires user documents, connected tools, rules for dealing with sources and the procedural consequences of a response. [39-41]

The suitability of the model for information search cannot be automatically transferred to legal advice, medical decisions or automatic refusal to the client. It is the application that determines the acceptable risk and the required evidentiary program.

26.2. GPAI in EU Law

AI Act sets out the independent obligations of general purpose model suppliers (GPAI), and for systemic risk models - an additional set of measures. These obligations are not legally identical to the rules for assessing the conformity of high-risk finished AI systems under Annex VI/VII. The presence of an external audit by the provider does not allow for any GPAI standard to be considered satisfied without verification. [1,49-50]

If the integrator uses the GPAI model as a component of a high-risk service, it is important to distinguish between the duties of the model supplier and the duties of the finished system supplier. The availability of information about the original model becomes a chain of responsibility, but it doesn't make all participants into one entity.

26.3. Instability and reproducibility

The parameters of generating, changing the model and changing the external search index can change the result of the same query. To validate this class of systems, reps, settings logs and periodic re-evaluations on independent sets are required. If the inputs are confidential, you can use artificial cases with specially marked errors without transferring real client files.

The publication should indicate that the estimates are fair for the given conditions and time. The statement "LLM does not hallucinate" is virtually unverifiable in the absolute sense; much more useful is the proportion of answers supported by a quote on a particular body, verified by an independent reviewer.

Practical checklist

1. Fix the basic GPAI model and the end-use application.

2. Measure quality in real configuration, including tools and sources.

3. Attach the conclusion to the exact version of the model and the re-evaluation rules.

Specificity of the generic model and the responsibility of the integrator

AI Assurance for a general purpose model is not the same as testing an application that uses such a model. The documentation of the basic provider may be appropriate for its role, but the specific integrator determines the sources of corporate search, employee access, instruction templates, and permissible external actions. His assurance should cover these distinct elements. The relevant responsibilities under AI Act vary for GPAI model suppliers and for end-system suppliers or operators; the mixing of roles distorts legal assessment. [1,49,50]

For a generative assistant, testing requires recording the prompt, search context, temperature and other essential parameters, model version, available tools, and the correct response benchmark. The metric "likes to the expert's response" does not always reveal an imaginary legal reference or omission of the procedural time limit. For sensitive files, it is reasonable to classify errors by consequences: incorrect citation, non-existent source, unconfirmed fact, disclosure of information, self-performing a forbidden action. Each class requires its own admission threshold.

When you use AI, you have to evaluate powers, not just text generation. Even the correct answer can be a risk if the agent has the right to send an email or change a file without confirmation. The 2026 Singapore AI agency management framework highlights the role of preliminary constraints and significant human control points. For ARGA, this requires separating the analytical function from the function of taking legally significant action. [43]

Practical case analysis 17.26

Recommendation: Evaluate the generic product in the context of use

Hypothetical example: The basic model answers general questions well, but it's wrong on the client's scans.

Analysis: The public evaluation of the model does not describe the quality of document extraction in the product.

Practical steps: The evaluator is testing the entire configuration with search, OCR and human verification.

Outcome criterion: The conclusion relates to the food assembly and the specific purpose.

27. Continuous monitoring, substantial modifications and withdrawal

27.1. Why isn't the certificate an indefinite guarantee?

For AI systems, data drift, external vendor updates, changing language environments, and user growth can affect properties after initial evaluation. AI Act provides for procedures for substantial changes and post-market monitoring; notified certificates may be restricted or revoked on the grounds provided. [4,13,52]

A separate management programme should link the change to a decision: leave the confirmation in force, carry out a limited re-test, extend the audit or initiate a new statutory assessment. The threshold cannot be determined solely by the commercial cost of repeated testing.

27.2. Events triggering an exceptional audit

The triggers are changes in model, data, purpose, user group, decision making rules, access to confidential information or regulatory classification. A serious incident requires an impact assessment, keeping logs and proper notification if required by applicable law. [52-53]

The organizational documentation should distinguish hot correction from substantial modification. Sometimes blocking a dangerous scenario doesn't require full certification, but every correction needs to be tested backwards to avoid creating a new hidden error.

27.3. Status of result and public update

If the authority has withdrawn the certificate, the supplier is obliged to correct the public statements and information provided to customers in accordance with the law and the contract. The current contract may provide for the obligation to report status limitations within a specified period and the right to independent review. When purchasing a critical AI service, these terms are no less important than the initial copy of the certificate.

Evidence of previous versions should be kept as historical metadata documents. Removing history after a failed evaluation destroys the chain of evidence and makes it difficult to consider potential complaints. At the same time, the storage of personal data must be kept to a minimum.

Practical checklist

1. Identify the legal and technical triggers for re-evaluation.

2. Describe actions in the case of restriction or withdrawal of confirmation.

3. Save the history of the findings without unjustified storage of client data.

V. EQUAL AND INSTITUTIONAL PRACTICES

Monitoring as a condition for maintaining evidentiary value

A one-time positive test quickly loses value when the data distribution, external provider, or workflow changes. The practical model should therefore distinguish between planned monitoring, re-checking after a significant change and investigating a specific incident. Triggers can be a deterioration in accuracy in the critical subgroup, a change in the data source, an increase in the load of false matches, user complaints, model updates, and a detected violation of access restrictions. [52,53]

Each trigger needs its own route: who receives the message, which artifact is saved first, who can stop using the system, which individuals are notified and when the re-evaluation is done. It is necessary to distinguish between the internal operating stop, the suspension of the certificate by the competent authority and the legal measures for special regulation. These are different status, and no one should be assumed without a decision.

It is recommended that the "detection - verification - correction - independent re-check - closure" chain be kept. The closure of the change of the instruction, without effectiveness checks, is not sufficient. Sensitive system suppliers are required to notify customers not only of a critical cyber

incident, but also of a significant degradation of metrics or the termination of the previous report. The public assurance should indicate the date of the last confirmed status.

Practical case analysis 17.27

Recommendation: Review evaluation after significant change

Hypothetical example: The supplier added the automatic dispatch of the notifications to the text assistant.

Analysis: The effects have changed, so the previous tolerance does not cover the new function.

Practical steps: The owner shall carry out an assessment of the change and a separate test of the authority before the shipment is engaged.

Outcome criterion: The new function has its own reasoned decision on admission.

28. The United Kingdom: the regulated AI Assurance market and UKAS

28.1. The policy of trusted external confirmation

The Trusted third-party AI assurance roadmap published by DSIT on 3 September 2025 looks at the professional market for external AI testing, expert competencies and quality assurance procedures. The document describes the possibilities, limitations and need for standardization, rather than introducing a universal mandatory certification for all AI systems in the UK. [32]

The approach is useful for comparative analysis: the state tries to improve the quality of assurance providers without mixing all industry tasks into one document. For the customer, this confirms the need to choose the type of assessment based on risk and purpose rather than the trendy name of the consulting product.

28.2. Accreditation ISO/IEC 42001

On 15 January 2026, UKAS announced the first BSI accreditation for the certification of AI management systems according to ISO/IEC 42001:2023. This is a concrete development of the market for accredited certification of AI management systems, but it is not a basis for asserting that all BSI-related systems have been licensed for legally sensitive applications. [33-34]

When choosing a UK body, the customer should check the area of accreditation and the applicable scheme. The British decision also does not replace the notification requirement when EU law is applied. Cross-border recognition requires a separate analysis of legal competence.

28.3. Conclusion from the British model

The UK differentiates between the development of the assurance services market, professional competence and accreditation infrastructure. This model opens up space for specialized expert work, especially in the legal assessment of AI systems in regulated areas, without immediately creating a single "AI Safety Performance" for the whole economy.

For ARGA, the quality of evidence studies, the drawing up of a legal map and the independent examination of the customer's claims are relevant. Pending receipt of any special accreditation,

these services should be described as analytical and expert, without using the term "accredited certification".

Practical checklist

1. Check UKAS approvals through the official registry and accreditation area.

2. Segregate the market strategy and the existing legal obligation.

3. Do not transfer British commercial recognition to the EU notification.

British ecosystem: evidence rather than abstract trust

The British model of AI Assurance is developed through a combination of industry requirements, professional accreditation infrastructure and published state road maps. DSIT's trusted third-party valuation documents are useful as a description of market development policy, but should not automatically be considered a single law requiring the certification of each AI system. The specific contract and the industry regime determine what type of evidence the customer really needs. [32,33,34]

UKAS is important as an accreditation body, but its competence for a separate programme is assessed by the area declared and in operation. Even the first AIMS recognition doesn't mean that we have global authority to issue certificates for all AI tools. When checking a British company, you need to confirm the accreditation status of the appraisal body, the allowable scheme, and whether the service is in line with the volume claimed.

Treaty discipline is particularly useful to international organizations. The customer may provide for access to independent testing, the obligation to keep logs, disclosure of conflicts of interest and review of status after updating the model, even if special legislation does not require a specific certificate. But in the ARGA report, such a contractual obligation should be defined as a condition of the parties, not an automatic consequence of British public law.

Practical case analysis 17.28

Recommendation: Check Local Suitability

Hypothetical example: The British report is offered as a universal confirmation to a foreign client.

Analysis: The procedure is useful as evidence, but its legal value depends on the addressee.

Practical steps: The lawyer compares the appointment and the criteria for the report to the local task.

Outcome criterion: Residual checks are listed before the result is adopted.



29. The United States: NIST, voluntary frameworks and sectoral context

29.1. Voluntary risk management framework

NIST AI Risk Management Framework 1.0 was published in January 2023 as a voluntary, inter-sectoral risk management tool. It helps organize risk identification, measurement and management processes, but is not a federal license or a nationwide certificate of compliance issued to all AI systems. [39]

The NIST AI 600-1 generative AI profile develops an approach to the inherent risks of generative models. In 2026, NIST also referred to work on the revision of AI RMF 1.0. Therefore, when using this framework in a contract, you need to fix the version and status of the document without giving the draft update as the final text. [40-41]

29.2. TEVV as an evidentiary function

NIST uses the concept of testing, evaluation, verification and validation (TEVV) for a set of verification processes. One program may require engineering testing, real-use monitoring and evaluation of management procedures. No one percent accuracy covers the whole set of tasks. [41]

American practice provides a convenient example of a functional approach: start with assessing the risks and goals of the system, identify appropriate types of testing, document the outcome and conditions for its applicability. But the official NIST material cannot be compared to a mandatory license for a specific commercial application in a particular state or sector.

29.3. Sectoral and contractual limitations

In the US, regulation may depend on industry, destination, and jurisdiction. Therefore, there is no short legal conclusion "the ISO certificate is valid in all of the United States". For the banking, human resources, medical and public sectors, special study of applicable norms and practices of supervisory bodies is required.

In cross-border projects, it is wise to use a common technical risk matrix and to conduct a separate legal characterisation for each country. Thus, it remains possible to reuse verified test evidence without false claims, and a single assessment has replaced all regional legal procedures.

Practical checklist

1. Call NIST AI RMF a voluntary frame, not a certificate.

2. Share the universal TEVV methodology and sector-specific commitments.

3. Check the relevance of the NIST before official use.

American risk assessment: the boundary of the voluntary methodology

NIST AI RMF 1.0 offers a voluntary approach to AI risk management, and AI generative profile complements it with subject-based risks. In itself, the use of the terms NIST is not the issuance of a NIST certificate. For an effective legal analysis, federal, state and industry obligations for a particular system should be checked separately. In the U.S., there isn't a single federal AI compliance certificate that replaces all types of testing, supervision and contractual verification. [39,40,41]

The RMF practical map can link management, context description, risk measurement and handling to management records and experimental tests. For example, controlling false matches in financial monitoring will require not only the level of risk, but also a threshold for activation, a review of false concerns, list quality, human decision-making, and an analysis of the consequences of denial of service. Working with only policy documents without checking real records is not complete.

For international research, it is important to avoid the "import of norms": the Federal American Institute's methodological recommendation does not create an obligation for a European or Singapore supplier without a special basis. It may be a useful general language of assessment, but the legal status must be given exactly. Contract inclusion of NIST in procurement requirements should also be separated from legislation.

Practical case analysis 17.29

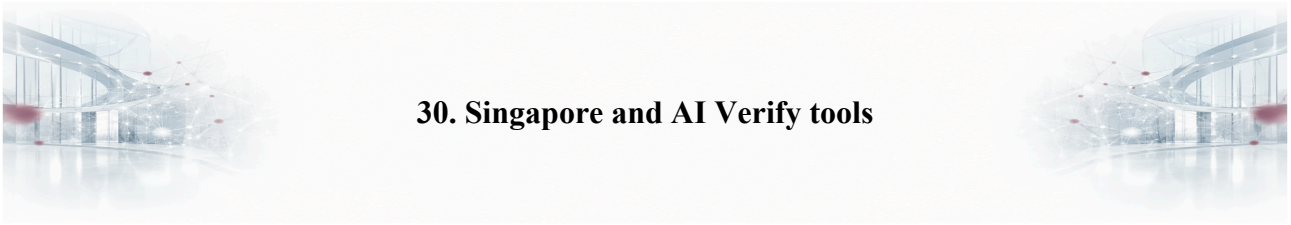
Recommendation: Do not convert the NIST frame into a certificate

Hypothetical example: The presentation says the product is NIST certified, but it's an internal risk profile.

Analysis: The methodological framework and the document issued by the authorized authority have different status.

Practical steps: The editor checks the basis and replaces the statement with an accurate description of the methodology used.

Outcome criterion: The public text is confirmed by an actual document.



30. Singapore and AI Verify tools

30.1. Test infrastructure without universal "licence"

IMDA has developed AI Verify as a framework for testing AI management and a set of software tools. Official information describes an orientation to the principles of transparency, reproducibility, security, fairness and other properties, as well as adaptation for generative AI. It's a useful practical guide for testing, but its use alone doesn't create a legal certificate of admission to all markets. [42]

The technical tool must be assessed on the available set of tests, the volume of input data and the qualifications of the person interpreting the result. Using the platform of well-known developers does not eliminate the obligation to explain why measurement answers a specific customer's legal question.

30.2. Agent Systems in Policy 2026

In January 2026, IMDA published material on a new approach to AI management, expanding attention to systems capable of self-calling tools and performing actions. For assurance, this means evaluating not only the quality of the text, but also the limits of authority, authorization, execution stop, log of actions and human responsibility. [43]

In an international comparison, such a timeline shows that evaluation schemes need to be updated along with product architecture. A repeat test of a generative response does not cover the consequences of automatically transferring funds, changing a record or sending a legally significant message.

30.3. Transportability of tests

When comparing the results of AI Verify with a European compliance file, specific methods and metrics should be analyzed, not matching terms. For example, an interface transparency test may not meet all the requirements of Annex IV technical documentation. A good engineering technique becomes a relevant proof only after it is compared with the compulsory test subject.

In ARGA projects, the logic of structured testing can be used as a methodological model, keeping the distinction between a private expert report recognised as a laboratory result and a legitimate European regulation.

Practical checklist

1. Appreciate AI Verify as a verification tool, not an automatic international certification.
2. For agency models, check real actions and limits of authority.
3. To carry abroad only proof of permissible volume, not legal status.

Singapore: test tool and limits of official recognition

AI Verify is a practical direction for developing testing infrastructure and methodological evaluation of responsible AI. The use of such a tool can enrich an independent protocol, but passing the test itself should not be submitted as a state certification or market authorisation without additional grounds. You need to establish the version of the tool, the selected principles, the verification actually performed and the status of the reporting organization. [42]

For agency systems, the IMDA framework, introduced in January 2026 and supplemented by practical cases in May of that year, offers guidelines for limiting human authority, life cycle control, and responsibility. In agency law applications, these topics are addressed in specific questions: can the agent access the client file of another case, send an external email, create a money transfer, change the original document or turn off the log? Every opportunity requires self-checking and verification rules. [43,56]

International comparison should differentiate between the guideline structure of management, private experimentation and mandatory regulatory procedure. A similar discipline is useful in preparing ARGA consultancy proposals: you can refer to Singapore's practical model as a source of organizational experience without attributing legal consequences that it does not create.

Practical case analysis 17.30

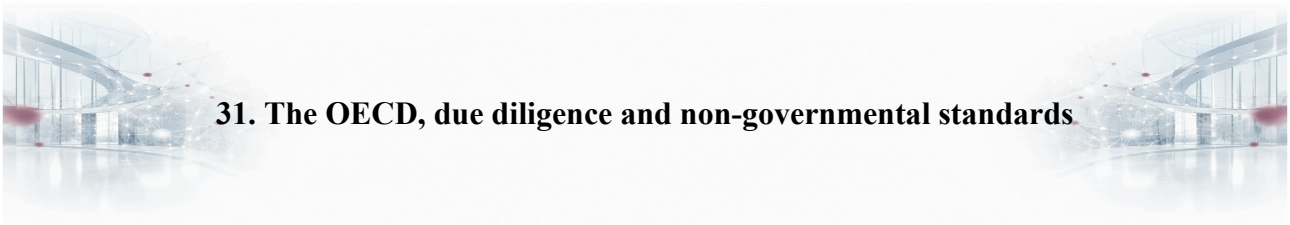
Recommendation: Show tool test boundaries

Hypothetical example: AI Verify set was completed with the data and declared the product fully legitimate.

Analysis: Technical verification does not replace legal characterisation and assessment of new conditions.

Practical steps: The appraiser indicates the data, functions and unverified properties; the lawyer evaluates the legal issue separately.

Outcome criterion: The publication distinguishes between the technical result and the legal conclusion.



31. The OECD, due diligence and non-governmental standards

31.1. OECD Guidelines 2026 as a source of procedures

In February 2026, the OECD published Due Digigence Guide for Responsible AI. The material offers a practical structure for identifying and eliminating adverse consequences when developing and applying AI. Such a guideline allows you to combine technical testing with verification of organizational responsibility and consequences for people. [44]

However, OECD recommendations are not automatically a national mandatory certification procedure. Their normative force arises from special legal reference, contractual inclusion or other established basis; in other cases they are used as a methodological guide and source of generally accepted practice.

31.2. Risk of declarative conformity with principles

Organizations often declare a commitment to transparency, equality or sustainability, but do not indicate which of these principles they follow. A good due diligence program requires a link between the risk claimed, the responsibility role, the evidence of management and the review mechanism.

For example, the existence of the "human in the loop" principle does not prove that a specialist can actually challenge the machine's evaluation. You need to see access rights, time to review, the ability to stop automatically and documents about the actual interference.

31.3. Complementarities among different sources

International guidelines, standards and laws can complement each other: the OECD helps identify adverse consequences, ISO/IEC 42001 forms organizational processes, ISO/IEC 17025 supports testing competence, and EU law imposes special legal obligations. But they do not fit into a single "superstandard" and do not eliminate procedural guarantees of applicable law. [1,21,27,44]

The expert opinion must specify separately which source makes the request and which only suggests a way to comply with it. This language improves the material's suitability for external consultation and prevents unreasonable legal conclusions from being drawn.

Practical checklist

1. Identify the legal requirement and the voluntary OECD guidance.

2. Document the actual application of the principles, not just their adoption.

3. Tie the methodological tool to the legitimate confirmable object.

Due diligence after assurance and mechanisms for remedial action

The OECD's responsible AI guidance of February 2026 places the assessment into a broader cycle of due diligence: policy consolidation, identification of real and potential impacts, prevention and mitigation, monitoring, reporting results, and participation in remediation of damage. This logic complements, but does not replace, certification and applicable law. Due diligence also focuses on the people who are responsible for the system's operation, not just the internal readiness of the corporation to audit. [44]

For sensitive algorithms, it's best to design an affordable way to challenge the inference beforehand. Even a perfectly formatted technical certificate does not answer the question of whether the victim will know about the mismatch, whether he will be able to submit corrective documents and who is responsible for updating the original register. Therefore, the assurance of reliability must contain a description of the error correction mechanism and the limits to its actual verification.

Within the ARGA study, it is advisable to keep a register of affected parties: customers, employees, counterparts, external sources, public analytics recipients. For each group, the risk, available evidence, procedure and independent point of escalation are determined. This approach increases the practical value of AI Assistance to legal and financial processes as compared to the general declaration of the responsible AI.

Practical case analysis 17.31

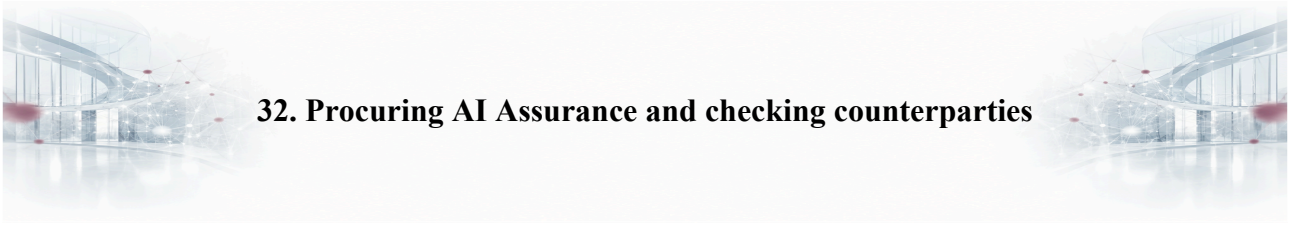
Recommendation: Link due diligence to the removal of harm

Hypothetical example: The company maintains a risk register, but clients' complaints are not transmitted to the owners of the measures.

Analysis: Recording risk without execution does not change the impact on people.

Practical steps: The coordinator links the risk, the turnover, the responsibility and the confirmation of correction.

Outcome criterion: The closed record shall contain a verifiable result for the person concerned.



32. Procuring AI Assurance and checking counterparties

32.1. Sound request for documents

The customer is advised to start with a map of goals: why do you need an external result - for procurement, regulatory filing, customer protection, board control or public trust. The purpose depends on whether you need a management system, a technical test, a specific certificate or a legal conclusion. The meeting party must submit a list of possible documents with the exact area of each.

The minimum package typically contains the valuer's legal name, accreditation or notification area if available, a scheme, a version of the standard, an example of an anonymized report, conflict of interest rules, confidentiality regime, additional testing costs and change handling plan.

32.2. Treaty management

The subject matter of the treaty should not be marked by the sole line "receipt of AI certification". The object, the criteria system, the independence, the data sources, the specific types of work, the format of the result, the limitations and the possibility of rejecting a positive conclusion need to be identified. Pre-payment for the conduct of the procedure should not imply a promise of a certificate.

In the event of subcontractors being involved, information transfer obligations and liability for the conclusion should be established. The customer shall have the right to know which essential types of tests have been passed on to third parties, including outside the jurisdiction of sensitive data storage.

32.3. Commercial and legal red flags

Attention should be drawn to the words "48-hour guaranteed certification", the lack of conditions for refusal, the reference to ISO without a year of revision, the refusal to name the accreditation authority, the inability to show the area of accreditation and the proposal to order a "GPDR certificate" without defining the scheme and legal status. It doesn't automatically prove dishonesty, but it does require additional evidence.

For ARGA, it is advisable to use a two-step procedure: a preliminary check of the status and scope of the documents, followed by a thorough check of the methods and suitability of the result for the client. It is only after these stages that a professional view of the adequacy of the package is possible.

Practical checklist

1. Request the name and status of the scheme before signing the treaty.

2. Provide for negative conclusion as an acceptable result of work.

3. Share the cost of the procedure and the promise of result.

Matrix D. Screening of typical false statements

The phrase	Problem	Correct check
ISO has accredited our model	ISO publishes the standard, not usually gives this status	Who actually gave the certificate, and what's the scheme?
42001 gives CE markings	The management certificate does not replace the EU AI Act mode	Was a special procedure applied?
Model certified worldwide	No universal scheme for any jurisdiction and function	On what subject and on what admission?
British CAB is a notified body of the EU	UKAS accreditation and EU notification vary	Document of the authorized State and region
300 tests zero errors = zero risk	Limited test volume	Upper limit and representativity limit

Procurement procedure: Certifier's review and confidence cost

When purchasing evaluation services, the price should be related to the subject matter, the competence of the performer and the availability of the materials, rather than to the promised "all IS" status. The technical task should be divided into a mandatory result, research hypotheses, scope and exclusions, expert qualifications, form of conclusions, permissible grounds for suspension and procedure of independent review. For normally required procedures, the status of the competent authority must be established in advance. [19,22,23]

A request for proposals may include two levels of cost: a basic public and documentary check and an extended test in a controlled environment. This eliminates the incentive to sell expensive universal "conformance assessment" without understanding the risk categories. In the comparative table of proposals, it is recommended to evaluate the availability of the methodology, accreditation status, independence, quality of the restrictions report and the contractual right of the customer to transmit the result to his or her lawyer or supervisory authority.

Cases of negative outcome should be envisaged. The contractor should not lose the right to compensation if a discrepancy is found; otherwise, the economic structure itself threatens impartiality. The customer must possess or have legal access to sufficient working materials for the other team to repeat critical conclusions. The commercial secrecy of the perpetrator shall not become a prohibition on independent verification of substantive legal claims.

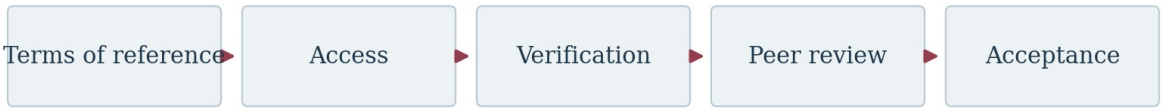


Figure 17.4. Acceptance of an AI Assurance service. The procurement and acceptance procedure in Section 32 pays for verifiable work, rather than a promise of a favourable result.

Practical case analysis 17.32

Recommendation: Buy a verifiable result

Hypothetical example: The offer promises a fixed-term certificate without a test description.

Analysis: The customer does not know what information he will receive and whether he will be able to use the conclusion.

Practical steps: The contract includes subject matter, sample, access, independence, review and admission criteria.

Outcome criterion: The reception is based on a set of evidence, not on the cover.

33. Confidentiality, personal data and model access

33.1. Legal purpose and minimum required volume

Technical audit may require access to logs, test records, images or fragments of training data. However, the requirement of competence and completeness of evidence does not override the principles of the applicable data protection right. For a specific project, the legal basis for processing, the distribution of roles between the parties, the allowable transfer and the need to preserve data are identified separately.

If the original case contains confidential legal records, an approach with artificial cases and controlled examples is preferred when it allows the required property to be verified. The use of real data can only be justified by a separate analysis of the necessity, protection measures and powers of the parties.

33.2. Technical alternative to full disclosure

The verifier can be provided with a remote controlled environment, an agreed API, a set of test functions or an anonymized log if these measures provide sufficient evidence. However, restrictions

on such access should be described in detail. It is not possible to declare a tested system "fully transparent" when the original models and parts of the data have not been investigated.

The evidentiary standard shall take into account the risk of leakage. Downloading the entire set of sensitive data only for the convenience of a third-party auditor can create an independent legal risk unmatched for verification purposes. So the contract should secure access to roles, logs of operations and secure destruction of excess copies.

33.3. Commercial secrecy and procedural protection

Participants can use trade secrets to restrict disclosure of certain information. The special regime of Annex VII AI Act provides for mechanisms for accessing the necessary information, taking into account the protection of intellectual property and commercial secrecy. It shows that secrecy and external verification should be balanced, not opposed. [6]

In the expert work, you can use an independent protected data room, action recording protocol, pre-determination of permissible selections and a separate privacy app. If critical evidence is not available, the outcome should be honest about this boundary.

Practical checklist

1. Identify the legal basis for data access and transmission.

2. Use artificial and impersonal test files whenever possible.

3. To record the extent to which restrictions on access affect the force of detention.

Legal access to data as a condition of audit quality

There is often a false alternative: either reveal the entire set of personal data to the auditor or refuse independent verification. Between these extremes are controlled viewing, supervised sampling, separation of metadata, pseudonymisation, synthetic test cases, and the creation of verifiable control amounts. The specific measure depends on the objective, the legal basis, the technical risk and the possibility of repeating the test. No general reference to the status of an auditor replaces the need to determine the permissible volume of processing and transmission. [18,39,44]

For legal cases, professional secrecy, client restrictions, storage regime and the possibility of transfer beyond the permitted infrastructure should be assessed separately. The agreement requires the processing purposes, the circle of persons, the conditions for removing working copies, the manner of notification of the incident and the procedure for confirming the destruction of temporary data. An independent expert should keep a record of the material and proof of deletion, but not make an additional copy of the sensitive archive for convenience.

If the customer cannot disclose critical data even in a controlled environment, the expert must limit the subject matter of the conclusion. The well-founded conclusion that "it was not possible to verify the frequency of errors in real records" was professionally preferable to the unjustified phrase on proven safety. In such situations, it is sometimes permissible to issue a conclusion about the methodology or organizational process without inferring the actual functioning of the system.

Practical case analysis 17.33

Recommendation: Segregate confidential access and public conclusion

Hypothetical example: The auditor needs a log with personal data, and the customer is ready to pass it on to all participants.

Analysis: Excess disclosure is not required to verify most of the findings.

Practical steps: Responsible provides minimal access roles and an impersonal review package.

Outcome criterion: The access log and the composition of the public package confirm the minimization.



34. Sectoral scenarios: banking, sanctions and international cooperation

34.1. Automated verification of sanctions matches

The sanctions screening system usually matches names, transliterations, dates of birth, documents and corporate connections to external sources. The legal risk depends not on the identification of the similarity itself, but on the following action of the organization: manual investigation, limitation of the operation or transfer of information to the competent authority. The verification should cover the origin of the sources, the quality of normalization and the rules for excluding mistaken identity.

In the certification scheme, the completeness of the identification of relevant records and the frequency of unreasonable matches should be checked separately. Otherwise, a high overall accuracy can mask a situation where the system regularly affects innocent customers and does not provide for effective correction.

34.2. AML/KYC and automatic client risk assessment

AML/KYC models often link the structure of transactions with the risk of the country, industry and customer behavior. A corporate AIMS certification document is not proof of the legality of a particular block or refusal of service. For a controversial episode, you need model rules, source data, log of operations, employee role and legal basis for the decision.

For ARGA, a limited independent verification model is useful: studying data sources, comparing error rates, and analyzing the client's ability to challenge a conclusion. In this case, the conclusion should not replace the powers of the financial regulator and should not contain unverified statements about the actual conduct of any bank.

34.3. International Search and Law Enforcement Analyst

An algorithmic search system for similar individuals, if used in international police cooperation, requires increased attention to the source of the record, the legality of access, the status of the warrant, identity identification and correction of misinformation. It is unacceptable to treat a match by name as a credible identity. Even a technically precise system does not prove the existence of a legal basis for detention.

This example is a conditional research scenario and does not claim that a particular international organization uses the algorithm described. For each real case, the applicability of AI Assurance

standards depends on evidence of the system, the type of data, and the procedural procedures available.

Practical checklist

1. To fix when an algorithmic flag leads to legal action.

2. Check identity by independent means, not just a match of a name.

3. Segregate the technical error from the legal characterisation and competence of the authorities.

Profile scenarios: sanctions, financial investigations and police data

For sanction matching, the target property includes not only the technical detection of a similar record, but also separate work with identifiers, dates, sources, transliterations and list updates. An error may result in an unreasonable service block or a significant match being missed. Therefore, the protocol should compare detection metrics and false alarms with the manual resolution procedure for uncertain cases. The ultimate responsibility for the applied decision is determined by the relevant law and internal distribution of powers, not by the fact of using a certified model. [39,44]

In international police cooperation, additional requirements are imposed for the origin and correction of source data. An independent assessment of a recognition tool does not confirm the legitimacy of placing a person in a specific database; the opposite is also true: a legitimate database does not prove the correctness of automatic identification. These two chains need to be examined separately. For ARGA's analytical unit, transparent references to materials, the processing of homonyms and the prohibition of assigning personality status only by model scoring are important.

In a financial investigation, it's useful to test not only the quality of connection detection, but also the reproducibility of the connection graph and the explainability of the source you choose. Automatic detection of an unusual transaction is a reason for professional verification, not proof of the wrong origin of the asset. AI Assurance report should prohibit the extensive reading of a probability estimate as an established legal fact.

Practical case analysis 17.34

Recommendation: Segregate the match signal and the client decision

Hypothetical example: The product marks a single name as a sanctions match.

Analysis: The measure of similarity does not establish identity or legal effect.

Practical steps: The specialist compares additional identifiers and documents the basis for the manual decision.

Outcome criterion: The final file differs between a machine signal and a reasoned decision.

35. Case study: a legal assistant and the limits of certification

35.1. Conditional system composition

The legal assistant draws up the documents, forms a summary and offers a draft response. The assurance file includes a supplier's AIMS certificate, a link testing report, a search index change log, and a description of human verification. The researcher examines these materials separately: they have different objects, dates and verifiable properties. It then determines which materials belong to the assembly actually used and which are characteristic of only one member of the chain.

When testing, the system must receive artificial files with ambiguous dates, conflicting documents, missing applications, multilingual names and false references within the files. The goal is to find out if it distinguishes the verified factual basis from the conjecture and whether it can reject the result in the absence of data.

35.2. Measurement programme

The measurement program has three independent outcomes: a file without an incorrect substantive reference; the completeness of mandatory facts; preservation of the evidence status of the data. In addition, there are cases of refusal and when the specialist cannot give the correct answer. The overall score does not unify these results without justification: a decrease in the number of false links may be accompanied by the loss of a substantive document. The decision must take into account every critical criterion.

It's wise to use artificial bodies and synthetic variations to access sensitive matters. The results on such a body are useful for limited quality demonstration, but do not provide a statistical guarantee of behaviour in all real cases; in conclusion this limit is indicated separately.

35.3. Legal language of opinion

With a positive limited result, the conclusion indicates the verified configuration, set of tasks, definition of the successful file, starting numbers, rule of decision, and completed constraints. If the rule is not followed due to uncertainty, it is reported what evidence is missing. The certificate is qualified by a separate conclusion. It does not automatically increase statistical confidence in the test result.

A "legal assistant is certified for international justice" is inadmissible unless a scheme, subject matter, authorized body and criteria are established. A signature by an external expert does not in itself create this status.

Practical checklist

1. Use an artificial shell with reference standards checked.

2. Include testing the origin of references and resistance to unreliable instructions.

3. To draw a conclusion without a promise of legal certainty.

Full audit cycle of the Legal Assistant

The legal assistant examination is organized as a promise check in a given field. First, the experts determine the essential provisions for each file and the permissible options for conclusion. Then the material available to the system is recorded, and the test is performed without adjusting the reference for the answers. Independent evaluators check the links and substantial omissions,

keeping the original disagreements. At the final stage, another specialist restores the calculation to the original markings and checks the language of the conclusion for data.

Monitoring includes comparing the primary document and interpreting its contents. The existing URL does not confirm the applicability of the quoted rule. The drafting, jurisdiction, date of validity and factual assumptions are checked. If there is a discrepancy between the comment and the original act, the reason for choosing the basis is recorded. If the question remains controversial, the success criterion allows for a reasoned indication of uncertainty; a confident wording without a reason is not considered an advantage.

It is recommended that the operations be divided into levels for subsequent operation. The assistant has the right to prepare a black chronology, but legal characterisations and referral to the addressee require the approval of the responsible specialist. If the source of a key fact cannot be determined, the system should show the gap rather than fill it with the most likely guess. Independent verification covers such a control architecture only within the real-life configuration.

Checking the adequacy of the legal assistant's provisional file

The supplier states that the legal assistant is "confirmed for the preparation of the documents without the need for a specialist to verify it". A management system certificate, a test of 100 files with 94 successful results and an internal description of the citation check are presented. It is the condition of the researcher to determine which part of the statement these materials support and which additional research is required.

evidence	What Can Confirm	Which requires separate verification
AIMS Certificate	Process in this area and period	The quality of the responses to a particular assembly
94 success stories out of 100 files	Observed percentage at specified marking	Selection, intervals, omissions and decision threshold
Description of the verification of quotations	Availability of a documented procedure	Actual performance and missing defects
Summary of materials	Selected elements of justification	Non-verification by a specialist

First, what is considered a success is specified: in this example, all relevant references to the file must have a proper primary basis. The reference marks are then restored and six unsuccessful cases verified. According to the conditional rule of section 23, the lower boundary of the interval should be at least 90%; at this volume it is about 87,5%. The grounds for the envisaged affirmative action are not sufficient. The management certificate doesn't change that calculation.

Even fulfilling the reference criterion would not support the entire original statement. The completeness of the legal analysis, the preservation of the actual observations and the consequences of the independent referral of the document remain separate issues. The reasoned conclusion is to narrow the assurance down to the investigated property, ask for missing data and determine the conditions for further evaluation. The human test is canceled only if there is independent legal and technical justification for such a decision.

Practical case analysis 17.35

Recommendation: Check the job approval without a lawyer

Hypothetical example: On 100 conditional files, the assistant has successfully processed 94, the supplier suggests cancelling the manual check.

Analysis: Six failures require qualification; the average share does not prove the safety of self-use.

Practical steps: The expert examines all six cases, checks for critical errors and assigns a limited control mode.

Outcome criterion: The decision about the regime is related to the gravity of the errors, not just the percentage.

36. The proposed institutional programme of Observatoire ARGA

36.1. Role of the independent think tank

The most realistic direction for ARGA is the legal and evidence-based examination of AI claims, selecting a regulatory route, and preparing analytical conclusions for institutional consultations. Such work does not require you to declare yourself an accreditation body or an EU notified body. The limits of the powers must be stated directly in the treaty and in the report itself.

It is recommended to keep a single register of international standards and regulatory changes, registering separately what is the current law, officially published standard, draft and expert proposal. In October 2026, such a registry is particularly important due to changes in the dates of AI Act and the transition to global accreditation infrastructure.

36.2. Three possible products with existing resources

The first product is the Regulatory Route Memorandum: a brief qualification of the area, subject, risk categories, applicable legislation and necessary documents. The second is Evidence Sufficiency Review: an analysis of documents and sources with a clearly limited conclusion about the sufficiency of evidence. Third-Party Assurance Due Diligence: verifying the authenticity and scope of the certificate, accreditation and applicability of the result to the client's activities.

Each product should contain a list of exceptions. For example, a study of legal documents without independent access to the software environment does not confirm the technical stability of the model. Such honesty limits commercial risk and increases the suitability of materials for professional use.

36.3. International consultations and conditions for submission

For external discussion, ARGA can offer narrow recommendations: a standard machine-readable passport of the certificate area, mandatory model change notification, documented differentiation of laboratory test and management system, as well as verifiable complaint procedure. Initiatives should be linked to the official agenda of the respective addressee. You can't declare a position sent or accepted until it's documented.

Internal harmonization shall be carried out according to the current ARGA procedure. An analytical publication may be an appendix to the position, but the independent report does not confirm either accreditation or appointment to AI Act system. The outcome of the institutional work should be assessed on the basis of verified outputs and official responses.

Practical checklist

1. Implement a single application passport and a standardized list of restrictions.

2. To verify the status of third-party certificates without attribution of authority to ARGA.

3. To link international consultancy positions with proven competence and relevant sources.

ARGA research model without false accreditation

The ARGA research programme can assess the quality of an assurance file's qualification without authorization from an accredited body. The pilot's main question is: can different experts on the same instruction identify the replacement of the subject and draw comparable conclusions about the adequacy of the evidence? The test uses depersonalized or specially prepared kits with known limitations. The result is measured by the missing material discrepancies and the reasons for expert disagreements, not by the volume of conclusions.

Before the pilot, the approval passport, the status dictionary, the list of critical qualification errors and the procedure for independent verification are approved. One set can contain both a correct management certificate and an unconfirmed promise of technical integrity. The procedure must preserve both conclusions. After the first step, the instruction is specified, and its new edition is checked on another set: the consistency of the answers to the examples already broken down does not confirm the tolerability of the skill.

International consultations should be linked to confirmed observations rather than general statements on "world certification". Potential proposals relate to assessment passports, evidence storage standards, protection against misidentification and mutual understanding of accreditation marks. For each external participation, an open channel, the subject of the consultation, the deadline, the result of the sending and documentary confirmation are required. The internal study should not be portrayed as the position of an international organization.

Practical case analysis 17.36

Recommendation: Pilot the expert qualification procedure

Hypothetical example: Two experts read the certificate in the same way, but in different ways assess its adequacy.

Analysis: Methodology needs operational rules and reproducibility testing.

Practical steps: The curator approves the passport of approval, independent evaluation and disaggregation of discrepancies on the new set.

Outcome criterion: The results of the pilot show the reasons for the discrepancies and the suitability of the instruction.

37. Conclusions and limits of international mutual recognition

37.1. Multilevel but not automatic trust

The global assurance architecture is evolving as a set of management standards, evaluator competence, special product schemes, national regulation and accreditation agreements. These mechanisms reinforce each other, but they don't form a universal document that allows AI to be used in any field without a new legal evaluation. International recognition of results is within the relevant area and does not eliminate the responsibilities of the receiving jurisdiction. [19,21-28,37-38]

The methodological conclusion is that the subject of the evidence, the quality of the method and the legal area of the result must be assessed separately. The transfer of the assurance is permitted only if the compatibility of these conditions has been confirmed. The document's external status may justify trust in the procedure in its field, but it does not extend the set of properties studied. This boundary must be maintained both in its full conclusion and in its brief public statement.

37.2. Proposals on transparency

For international exchange of evaluation results, it is appropriate to have a mandatory passport for each selected scheme: who has been verified, what has been verified, what version, what area of accreditation, on what criteria, when, by what methods and what properties are excluded. It is also desirable to standardize the notification of system suspension and significant change.

Separate discussion requires rules for statistical uncertainty disclosure, criteria for evaluating vulnerable groups and protecting confidential data in external testing. Without these elements, cross-border recognition can become a formal agreement on logos if the contents of the documents are not verified.

37.3. Final rule for ARGA

Before using the result, ARGA checks four grounds: the accuracy of the claimed property; the conformity of the evidence to this property; the fulfilment of the criterion under the established rule; the permissibility of using the conclusion in a particular jurisdiction and procedure. Unresolved space is identified by targeting. The conclusion may be to use the document in a narrow field, to request missing materials, to conduct a separate test or to reject an excessive application.

The next research stage is a pilot test of the reproducibility of qualifications and a breakdown of significant differences between experts. On its basis, you can clarify the passport of the certification area and propose requirements for uncertainty disclosure for international discussion. The legal provisions are updated on the date of specific use. The very fact of preparing this report does not imply a pilot or confirmation of the properties of the actual product.

Practical checklist

1. Not equate mutual recognition of accreditation with recognition of each AI product.

2. Require a regional passport and the conditions for any important assurance.

3. Maintain the distinction between an expert opinion, a certificate and a legal declaration.

PRACTICAL ANNEXES

Cross-border transferability limits and final conclusion test

The main legal outcome of the study is to differentiate trust in the evaluator's competence, evaluate a particular object and recognise the result in a specific legal procedure. The international agreement of accreditation bodies promotes comparability of results, but does not create an independent license for any AI to enter the market or be used in criminal proceedings. For each application, you need to find a specific norm, contract or accepted practice. [19,35-38]

Therefore, it is advisable to construct a final conclusion on a certificate in the form of six consecutive answers: who issued it; what area of authority is in effect; what object and version are verified; on what criteria; what properties are actually established; for which jurisdiction and procedure the result matters. If at least one important question is answered in the negative, the document should not be automatically declared invalid. An exact gap and the required additional confirmation should be identified.

For ARGA, a unified set of verification forms where each external recommendation can be verified by another specialist will be a promising outcome. The programme should allow for changes to the conclusions after new material has emerged, conclusion of outdated certificates and the availability of corrective human error data. So AI Assurance becomes not a decorative certification, but a process of responsible handling of uncertainty and the consequences of decisions.

Practical case analysis 17.37

Recommendation: Limit final recognition to a confirmed area

Hypothetical example: A set of evidence is sufficient for the internal preparation of the text, but it is offered for autonomous submission.

Analysis: The new area includes consequences that the tests did not verify.

Practical steps: The authorizer maintains a restricted regime; for extension, it specifies additional tests and legal checks.

Outcome criterion: The outcome lists the permissible use, exceptions and conditions for revision.



Appendix 1. Certificate of approval and result AI Insurance

The passport is required so that the next recipient of the report does not accept the certificate beyond its scope. Each project uses one self-approval record. If data protection, functional quality and risk management are confirmed for one product, at least three records with different criteria should be entered.

Code	Field	What needs to be specified
A01	Precise allegation	One verifiable sentence without absolute promises
A02	Type of result	Peer review / technical test / certification / accreditation / legal declaration
A03	Legal person	Registration and commercial name, branch if necessary
A04	Object	Model, component, ready system or management system
A05	Version ID	Model/code/API/Date of issue/check amount
A06	Appointment	Acceptable uses and user group
A07	Borders	Functions included and excluded, external services
A08	Normative source	Act, article, standard with year or private scheme
A09	Criterion	Threshold, method, error allowed, measurement method
A10	Method	Audit, surveillance, testing, document analysis
A11	Data	Origin, volume, sampling and marking principles
A12	Competences	Team, independence, industry experts
A13	Evaluation Officer	Full name CAB or researcher
A14	Accreditation	AB name, number, region, date and status; no in absence
A15	Notification	Country, number and area of special law; "not applicable" in the absence of
A16	Evidence	Register of attached materials and access restrictions
A17	Unverified	Questions outside the assignment; insufficient data; external components
A18	Result	Confirmed / partially / not confirmed / not evaluated
A19	Time limit	Date, current version, re-evaluation criteria

Code	Field	What needs to be specified
A20	Public language	Concise exact text allowed for publication

Recommended passport acceptance procedure: the lead analyst fills in fields A01-A11, the technical specialist checks A05-A11, the legal reviewer analyzes A08 and A15, the independent reviewer confirms A12-A14, the editor approves A20. Any significant change in the A05, A06 or A08 fields triggers a decision to reassess.

Note The A14 field should not be filled in with the word "ISO" in the box of the accreditation body. An institution that has indeed made a decision on accreditation should be named. [21,24]

Additional methodological clarification

When filling out a passport, one cannot leave the entry "normative basis" as the general word "ISO". It is necessary to specify the specific edition, the nature of the standard and the relationship with the stated result. For legal applicability, a date is also required: an existing certificate can be obtained before the current requirements change. Each parameter to which confirmation cannot be obtained is recorded as "not established" and not as "not applicable". These two categories are usually mixed in commercial questionnaires.

Appendix 2. Authentication protocol and certificate area

The protocol serves as an external confirmation of the information that the organization has presented in its advertising material. It does not verify the technical validity of the whole conclusion: it is a separate step. It is recommended to keep a dated copy of the registry result or search record, excluding personal data not necessary for the purpose.

Verification procedure

1. Receive a document from the party and mark the identifier, name, version of the standard and the claimed item.

2. The full legal name of the issuing authority and the signature or electronic identifier shall be established.

3. Determine status: voluntary certification, accredited assessment, notification, testing or advisory report.

4. If accreditation is applied for, open the official register of AB and check the matching of the name, number, status, site and area.

5. Check the system of mutual recognition in this area. In 2026, take into account Global Accreditation Cooperation Incorporated and the current rules of regional agreements.

6. If an AI Act notification is filed, request confirmation from the competent authority and the relevant field of procedures.

7. Compare the certificate object with a specific system or organizational function that the supplier is claiming.

8. Keep a documented decision: confirmed, found discrepancy, required clarification, status not verified.

Monitoring	Yes / No / Not specified	Document and commentary
Convergence of a legal entity	_____	_____
Current status	_____	_____
Standard version specified	_____	_____
Area includes required service	_____	_____
The signatory is competent	_____	_____
Accreditation confirmed	_____	_____
Notification confirmed if necessary	_____	_____
The result refers to the right version.	_____	_____
Date of validity/oversight specified	_____	_____
There is a complaint/change procedure	_____	_____

The material discrepancy should not be closed by one oral phrase from the supplier. If the register has a different status, an official explanation from the issuing party or accreditation body is required. In the absence of verification, the neutral phrase "copy provided by the supplier; no independent confirmation of existing status has been received" could be used.

Additional methodological clarification

The review of the registry shall be documented as a separate evidentiary act. It is preferable to save the official database URL, the date of the request, the search query line and the record found. If the official register provides only limited information, it is required to specify these limits rather than claim that the status is absent. If the register and the copy differ, the result is marked as disputed until the official explanation is given.

Appendix 3. Map of the control route on AI Act (ed. 2026)

The sequence of research issues is given below. It's a pre-selection scheme, not a replacement for the official legal characterisation of a particular product. References are made to the amended consolidated text of AI Act as of July 27, 2026. [1-8]

Analysis route

1. Determine whether the system and its participant fall within the territorial and subject area of the EU AI Act.

2. Check for specific prohibitions and obligations that may apply regardless of the high-risk classification.

3. Describe the real purpose of the system, not the marketing name of the technology.

4. Check the basis for high-risk: Article 6(1) and Annex I or Article 6(2) and Annex III; record special exceptions.

5. Verify the current date of duty: 2 December 2027 for the specified case Annex III or 2 August 2028 for the specified case Annex I; do not ignore other dates.

6. For Annex III, paragraphs 2-8, determine the provided Annex VI; for paragraph 1, check the selection of Annex VI/Annex VII and the standardisation conditions.

7. For Annex I, apply the industry assessment procedure and the special conditions for the integration of AI Act.

8. The status of the notified body, if legally necessary, and the required amount of technical documentation shall be determined separately.

9. Form a list of required declarations, designations, registrations and obligations after release.

10. Determine the re-evaluation point after a change in system, standard or legislative deadline.

Fact	What a participant has to prove	Basis
Category for Annex III	Specific point and purpose of the system	Art. 6, annex III [1]
Annex I category	Industry act and product class	Art. 6(1), annex I [1]
Assessment route	Why Anex VI, VII or Sectoral Procedure	Article 43 [4]
Effective date	Reference to amended article 113	Article 113 [3]
Harmonization	Publication of reference and scope of claims	Art. 40 [7]

Fact	What a participant has to prove	Basis
Third party evidence	Notification and organ area	Art. 28-31 [9-12]
Declaration and labelling	System responsibility and identifier	Art. 47-48 [15-16]

Additional methodological clarification

Hypothetical example: A personalised system may fall under Annex III, but its voluntary AIMS certification does not replace the statutory obligations of the provider. Another example: a medical device that includes AI as a safety component may require an analysis of the industry procedure and Annex I. In both cases, the date of performance of a specific duty and the status of the participant are indicated separately; the similarity of the algorithm does not make the routes the same. [1,4]

Appendix 4. Functional map of ISO/IEC standards and EU sources

The table distinguishes the objects of the standards from their official public descriptions. It's a synthesis of the author, not a literal matrix of obligations from paid texts. In real order, you should check the availability of a full regulatory edition and its inclusion in a specific evaluation scheme. [21-31]

Tool	Main object	Model use issue
ISO/IEC 42001:2023	Management I.I.	Is the organization systematically managing I.I.'s risks?
ISO/IEC 42006:2025	AIMS Certification Body	Is it competent and consistent to certify AIMS?
ISO/IEC 17021-1:2015	Management Systems Certification Authority	Is the process competent and impartial?
ISO/IEC 17065:2012	Product/process/service certification body	Is there an applicable product scheme?
ISO/IEC 17025:2017	Test laboratory	Is the specific measurement method reliable in a recognized area?
ISO/IEC 17029:2019	Validation/verification body	Is it true or appropriate for a certain statement to follow a pattern?
ISO/IEC 17011:2017	Accreditation body	Is the accreditation entity competent?
ISO/IEC 42005:2025	Impact assessment of the system	What are the consequences for people and society?
ISO/IEC 25059:2023	AI quality model	What quality properties can we describe and evaluate?
AI Act: Annex VI	Supplier's self-assessment by legitimate route	How have the mandatory internal audits been carried out?
AI Act: Annex VII	Notified body and technical file	How has the special legal assessment and oversight been carried out?

The legal presumption of conformity with AI Act should not be removed from any column above. The presumption requires the conditions of Articles 40 to 42 and the independent verification of the reference to the harmonised standard in the Official log of the EU. [7-8,47]

Additional methodological clarification

The register of regulatory criteria should use a separate "published current revision" / "final draft" / "not verified". In October 2026, the ISO/IEC FDIS 17065 and ISO/IEC FDIS 25059 designations do not represent a published standard; their significance is limited by the development forecast of schemes. The procurement must not replace the current criterion with a design document without clear agreement and explanation. [25,26,30,31]



Appendix 5. Model test protocol for sanctions matches

The example is completely conditional, does not relate to a particular financial organization and does not contain actual personal data. The goal is to show the structure of the test within a narrow AI Assurance framework: to separate the mathematical quality of the comparison from the operator's legally significant solution.

Protocol field	Conditional filling	Limitation
Appointment	Initial search for possible matches	Not a lockdown decision
System	Master X, version 2.4, settings 2026-10	The new version requires evaluation
Reference	200 marked artificial records	Not real bank statistics
Positive cases	80 confirmed matches	Base frequency is artificially high
TP	60 correctly identified matches	In the conditional hull
FP	20 miscalculated discrepancies	Could hurt a client.
FN	20 missed hits	Could lead to incomplete control
TN	100 correctly rejected discrepancies	In the conditional hull
Precision	60/(60+20) = 75%	Depends on the prevalence of class
Recall	60/(60+20) = 75%	Doesn't reflect the price of error
FPR	20/(20+100) = 16,7%	Relates to negative cases
Human control	Mandatory flag checks by the operator	Requires a separate audit
Legal solution	Not evaluated	A review of authority and basis is required

The difference between accuracy and recall remains, although in this curriculum they happen to be equal. For a real project, it is necessary to define a selection scheme, markers, language groups, thresholds and the mechanism for recovering an erroneous identity. Otherwise, even the correctly calculated fractions do not give a reliable conclusion on the suitability of the instrument.

Additional methodological clarification

For each incision, the evaluator should check whether the same initial personality is repeated in the training and test part. In the presence of such repetitions, a high metric may reflect memory rather than the ability to detect new cases. The rare-match test should be supplemented with a qualitative revision of each pass and a documented decision on the need for human review.

Appendix 6. Statistical uncertainty: replicable training calculations

The numbers are given only to illustrate the principles of statistical estimation. They are not the result of testing by an organization. The permissions of each model are indicated separately, as changing the selection scheme may make it unacceptable.

Educational background	Calculation	Conclusion and boundary
94 out of 100	94/100 = 94%; 95% Wilson = 87.5% - 97.2%	Doesn't prove quality >94% in total
0 malfunctions out of 300	1 - 0,05^(1/300) = 0.99%	One-sided upper limit 95 per cent with binomial conditions
0 malfunctions out of 1,000	1 - 0,05^(1/1000) = 0.30%	Rare risk does not become zero
FPR in Annex 5	20/(20+100) = 16,7%	The frequency of error in truly negative cases
Inter-group accuracy gap	(90/100) - (60/80) = 15 pp.	Group composition and confidence intervals need to be studied

Wilson's interval does not correct the non-representativeness of the sample and does not take into account the dependent series of requests. In the case of cluster data, checking the same client with variations in name and post-trained models, it is necessary to choose a different statistical model or to clearly limit the power of inference. No calculation indicator in itself is legal proof of non-discrimination or unconditional security.

Additional methodological clarification

All values in the annex are illustrative, not obtained during testing of real systems. When demonstrating the result, it is necessary to clearly state the assumption about the independence of the observations, the criteria for classification into the true category and the formula used. If these

prerequisites are violated, the usual numerical range does not have the claimed coverage. This is particularly significant for groups of related checks and repeated names.

Appendix 7. Evidence register for limited independent analysis

This sample is suitable as an annex to the technical assignment for expert work. It separates the document received from the interpretation. The column "necessarily required" prevents the false impression that the verification is complete with a incomplete package.

ID	Alleged evidence	Verification	Open issue
E01	Designation of the system	Comparison with actual functions	Are there hidden scenarios?
E02	Component and vendor scheme	Checking external dependence	Who's updating the API?
E03	Rules for changes in versions	History of actual issues	Is the model changing without notice?
E04	Technical file	Reconciliation with Annex IV when applicable	Are all sections relevant?
E05	Data quality assurance plan	Origin and markings	Is it possible to leak teaching examples?
E06	Test report	Meterics, sample, reference data	Is the standard independent?
E07	Appeal documents	Analysis of procedure and actual case files	Can a person review?
E08	AIMS Certificate	Certification area	Is a particular process included?
E09	CAB Accreditation Register	Number, status, area	Is it appropriate to do what you have tested?
E10	AI Act notification	Scope	Isn't this another procedure?
E11	Incident log	Verification of completeness and actions	Were there any unrecorded complaints?
E12	Agreement with external model	Access rights and notification of changes	Does the provider update the model without consent?

ID	Alleged evidence	Verification	Open issue
E13	Protocol on Conflict of Interest	Assessment of team independence	Was there a consultative relationship?
E14	Application log	Sample of real solutions	A man checking flags?
E15	Decision on sufficiency	Signatures, comments, limitations	Who confirms the public conclusion?

It is desirable to keep the original sensitive material in a closed storage with limited access, and to display only the code, type and date in the report's annex. Downloading a real client file into a public version is unacceptable without special legal grounds.

Additional methodological clarification

The registry should not only keep the name of the artifact, but also the chain of origin: who it came from, what system, how, at what time and with what filters. The materials provided by the verified party and the results of the independent re-test are indicated separately. The mixing of these categories in the final report makes it difficult to assess the independence of the conclusion.

Appendix 8. Example of limited expertise

MODEL MODEL The text below does not confirm the actual assessment of any real person, developer or service. It's a form for subsequent adaptation.

Subject

"Authorization of the validation of the accuracy of AI version assistant references [X] when using an artificial test shell [Y]».

Criteria

"Observation of the previously agreed method of reference verification, lack of fictional quotations and correct acknowledgement of insufficient information."

Methods

"Analysis of documentation submitted, implementation of agreed tests, independent review of responses by two evaluators and reconciliation of logs within the available access".

Limitations

"There was no analysis of internal training of the model, cloud supplier source code, actual client cases or conformity with the law of other jurisdictions."

Result

"The tested artificial hull has identified the indicators given in the Annex [N]. The observations obtained confirm only the characteristics and conditions specified in the protocol."

Status

"The conclusion is an independent expert study on an agreed assignment and is not an ISO certificate, an accreditation document or a document of a notified EU AI Act body."

Changes

"When changing the model version, the tools involved or the assignment, the results need to be revised according to the rules of assignment."

Signed

"F.I.O., professional role, competency description, declaration of independence, date and contact for authentication."

The final version should not leave the variables in square brackets without filling in. Unlike Word flags, these text fields do not create problem characters in another system, but should be filled in when the form is actually used.

Additional methodological clarification

In the formulation of a limited conclusion, the expert should not "compensate" the lack of access to broad terms of reliability. A validation reservation specifies which data were missing and which element of the result cannot be established. The professional phrase "cannot be confirmed" does not in itself constitute a breach by the supplier, but shows the limit of the evidentiary base.

Appendix 9. Contractual conditions for external evaluation of IA

The list is intended for negotiation and does not replace the legal analysis of a contract in the chosen jurisdiction. With regard to legally binding certification, the procedure and limits of the parties' freedom shall be determined by special rights and scheme rules.

Item	Proposed content
Subject	The service description indicates the system, version, separate criteria and applicable standards with the year of publication.
Type of result	The contract distinguishes analytical notes, test reports, certificates and assistance to the official notified procedure.
Lack of guarantee	Payment of the procedure does not imply an obligation on the evaluator to issue a positive document.
Data access	The Parties shall determine the minimum necessary set of documents and technical means, the confidentiality regime and the action log.
Competence	The executor shall report on the necessary specialists, the involvement of subcontractors and the status of accreditation in the field in force.
Independence	The evaluator discloses significant organizational and financial conflicts of interest; measures are envisaged to eliminate them.
Evidence	The report contains the source of each significant conclusion, limitation, unprotested components, and a version of the object.
Changes	A list of changes automatically starts the analysis of the need for re-evaluation is determined.

Item	Proposed content
Public statements	The Parties agree on the authorized wording of the result and the rules for the use of accreditation marks.
Suspension of status	The operator shall notify the accreditation/notification change if this affects the service or the valid evidence.
Complaints	The address, time frame and form of a reasoned reply to the declaration of actual error shall be determined.
Material protection	Storage, removal, return and exclusion for documents required by law shall be secured.
Applicable law	The law and the court/arbitration shall be established without limiting the mandatory requirements.
Responsibility	There's no confusion between the evaluator's responsibility for correctness and the developer's responsibility for their AI product.

It is recommended that the annexes "the matrix of criteria", "the rules of access", "the competence and independence", "the form of result", "changes and reassessments" be harmonized separately. This makes the assignment verifiable and reduces the risk of a dispute due to too general a commercial subject of "IIE verification".

Additional methodological clarification

The negative evaluation shall have the same contractual status as the positive evaluation. The acceptance of services is verified by the fact that the agreed methodology has been followed and the completeness of the reporting materials, not by the content of the conclusion desired for the customer. A different approach creates the valuer's financial interest and increases the risk of undisclosed restrictions.

Appendix 10. Form of assessment of the independence of the perpetrator

The assessment of independence is not limited to the formal absence of affinity. There are threats of self-testing, financial interest, client advocacy and dependence on future orders. The form below is used as a structured work tool; it is not an ISO normative form.

Threat	Control issue	Possible action
Self-checking	Did the team develop the test element?	Dismissal, separate independent team

Threat	Control issue	Possible action
Financial dependence	Does the reward depend on the positive certificate?	Prohibition of access-based certificate
Permanent client	Does the customer form a significant share of the turnover?	External review, additional safeguards
Common property	Is there a common owner or leader?	Disclosure and organizational separation
Consultations	Has the team made recommendations for the implementation of the monitoring being evaluated?	Another team and an independent decision
Data access	Can a client selectively remove unprofitable materials?	Independent evidentiary registry
Promise of result	Had there been any claims of imminent issuance of a positive document?	Correction of the treaty, waiver if necessary
Limited subject	Who determined which scenarios were not tested?	Public disclosure of exemptions
Public independence	Will the interests of the parties be properly disclosed?	Signed declaration of conflict

The decision on each threat must contain the name of the person responsible, proof of the action taken and the date of the review. The phrase "no conflicts" without the completion of the initial questionnaire does not provide reliable independence and makes it difficult to protect the opinion when a dispute arises.

Additional methodological clarification

It is recommended to update the form when the team, customer, subcontractor or payment order changes significantly. The declaration signed before the draft does not guarantee independence in a few months. Disagreements about the methodology and the right to a negative conclusion are best settled before the evaluator allows sensitive data.

Appendix 11. Certificate monitoring programme and re-evaluation triggers

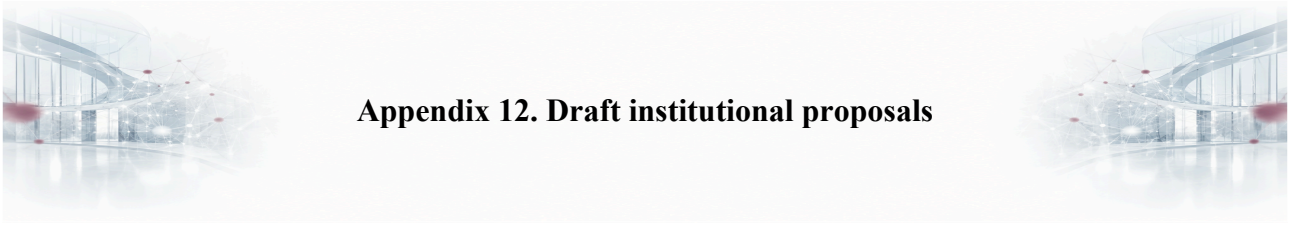
The following are examples of triggers that should be evaluated taking into account the specific scheme and applicable law. Automatic new certification is not required with each change in interface text; however, a small change in appearance can significantly affect legal risk.

Event	First control action	Possible outcome
New base model	Check API properties and status	Regressive test / new evaluation
Retraining on the new building	Check the origin of the data and drift	Rechecking quality and impact
Change of match threshold	Repeat error matrix	Clarifying admissibility
New State of Use	New legal characterisations	Local supplementary inspection
New Personal Data Class	Assess the basis, access and protection	Updating of documentation and procedures
Serious incident	Maintenance of logs, evaluation of notification requirements	Limitation of function and independent verification
Withdrawal of the accreditation of the evaluator	Confirm the legal effects of status	Request for alternative evidence
Update of harmonized standard	Check official publication and transition	Revision of the claim map
Changes in the objectives of the system	Check the regulatory route	New classification/quantity
New integration with external service	Check threats, cross-border data	Access and confidentiality tests

Each re-evaluation decision shall be entered in a brief record with date and argument. Having a historical positive certificate is useful as part of the evidence history, but it does not eliminate the obligation to consider new circumstances. [4,13,52-53]

Additional methodological clarification

The signal to verify may be not only the expiration of the certificate, but also a change in the covered activity, a reorganization of the legal entity, an update of the version of AI component and a new publication of the accreditation body. For each event, it is recorded whether a documentary confirmation or a new test of system properties is required.



Appendix 12. Draft institutional proposals

The following proposals are formulated for further internal revision. They are not addressed to a specific regulator and are not considered addressed, approved or included in official consultation.

Proposal 1. AI Insurance region passport

Develop a general machine-readable data structure: object, version, criteria, scheme, evaluator status, area of accreditation, applicable jurisdictions, verified properties, statistical restrictions, monitoring conditions and expiration date. Minimum composition does not require disclosure of secret source data; it sets a standard for transparent description of what is actually confirmed. Such a measure reduces the risk of marketing exaggeration and facilitates the procurement of evaluation services.

Proposal 2. Verification of mutual recognition by region

The international exchange should publish not only a reference to the multilateral accreditation agreement, but also the specific area of recognition on which the report is based. It is desirable to separate test accreditation, management certification, product certification and EU special procedure notification in the information manuals. Such detail is particularly needed after the transition to Global Accreditation Cooperation Incorporated in 2026. [35-38]

Proposal 3. Protection of persons in case of automatic matches

It is recommended to include in qualitative schemes for sensitive systems individual indicators of false identification, verifiable human retrieval capability, documented origin of records and available error-correcting procedures. The aggregate percentage of correct answers should not be used as the sole evidence of reliability in decisions affecting freedom of movement, access to banking services and procedural rights.

Modalities for a possible external position

Internal harmonization route

1. Choose a specific open consultation and set the official question, deadline and acceptable formats of submission.

2. Shorten the proposals to the subject of discussion and check the references to the current regulatory text.

3. Ensure a legal and technical independent review of the contested provisions.

4. Receive internal ARGA harmonization in accordance with the established procedure.

5. Submit material through the official channel and maintain confirmation of actual participation.

6. Enter the filing fact, number and response of the regulator in a closed log of documents; do not announce participation until confirmed.

BACKGROUND APPARATE

Additional methodological clarification

Draft institutional proposals should be kept separate from the actually directed consultation responses. In the public certificate, it is sufficient to state the status of "internal project", "harmonized submission" or "directed" if a receipt is available. This discipline protects an organization from accidentally creating the appearance of official participation or accreditation.



Appendix 13. Draft Universal Passport AI Insurance

The purpose of a passport is to allow a third-party lawyer, regulator or purchaser to understand the limits of the evaluation result without verbal explanation. The passport is filled in separately for each critical system or substantially independent subsystem. The declared object must have a reproducible identifier, and every evaluation must be tied to a specific version. The passport is not a state registration and does not replace the mandatory forms of special regulatory regimes. [18,22,39]

Mandatory requirements include the applicant's legal personality, the supplier's role, system designation, operation area, verified configuration, list of excluded functions, basic test scheme, basis for quality and independence of the evaluator, results of verification of key indicators, uncertainty intervals, known limitations, date and duration of the conclusion. For individual criteria, not only conformity should be indicated, but also the reason for the impossibility of verification, such as the unavailability of a closed database or a small number of rare events.

The passport must be accompanied by a table of evidence that can be traced from the conclusion to the working file. A closed-ended version is permitted: a public recording informs of the existence of evidence, its owner and the legal access regime. When significantly modified, the system receives a new version identifier and a record of the need for re-evaluation. For comparative analytics, the ARGA formalised passport allows you to disseminate current legislation, voluntary standards and private methodology.

Working table to the annex

Field	Description and confirmation
Subject and version	Name, ID, assembly date, configuration
The role of the complainant	Vendor/operator/integrator; jurisdiction
Declared property	Verification claim and its boundaries
Criterion	Article / standard / voluntary threshold
Document issued	Type of result, number of the authority
Independence	Conflicts of interest and measures to resolve them
Scope of recognition	Accreditation/notification/non-status

Field	Description and confirmation
Deficiencies	Exceptions, uncertainty, violations identified
Changes	Triggers of re-evaluation, responsible
Verification	Date, storage, signature of the

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 14. Decision-making matrix on certificates and certificates

It is recommended that due detection should not be limited to the binary "real/falsified". Practically, genuine documents are available, issued by the competent authority, but for a foreign area, an outdated version of a product, an unaccredited program or a legal entity that is not a provider of the verified service. It is equally important when the original document actually confirms a property, but in a commercial statement attributes it to another, broader effect.

The matrix below offers a process status that does not replace the legal decision of the competent authority. Category A means authenticity, status, domain and relevance are confirmed. Category B records a partial confirmation for a non-critical gap that requires disclosure. Category C denotes a substantial inconsistency or non-verification of a legally significant element. Category D requires the escalation of signs of forgery or deliberate misrepresentation; such signs are documented without public accusation until sufficient grounds are found.

For each statement, the date of registration check, the URL or official source, a copy of the document received and the result of the request for clarification are recorded. A distinction should be made between receiving a response from the supplier and an independent confirmation. Open materials do not use the real names of clients or case numbers without a valid reason. The decision on further use of the system shall be taken by a person authorised under the relevant legal regime.

Working table to the annex

Grade	What's established	Recommended action
A	Subject, status and area match	Accept limited output
B	There's a gap that doesn't change the main conclusion.	Open reservation, request clarification
C	Substantive unconfirmed condition	Do not invoke the legal effect
D	Evidence of significant lack of credibility	Record, refer for legal evaluation

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 15. Working Protocol on Legal Qualifications under EU AI Act

The purpose of the protocol is to ensure the reproducibility of the legal chain preceding the order for the valuation service. The lawyer records individual decisions about the status of the subject, the territory, the product's purpose, the grounds for high-risk classification and the applicable term. Only then will the evaluation procedure be determined. An error in the initial classification cannot be corrected by a qualitative technical audit performed for an unnecessary or insufficient procedure. [1-6]

The record shall include all grounds invoked by the party. If the product can be considered simultaneously within Annex I and Annex III, it should be clarified which route has priority and according to which norm. For paragraph 1 of Annex III, the conditions of Article 43 shall be considered separately, including the application of harmonised standards or general specifications. For paragraphs 2 to 8 of Annex III, internal control shall be specified. For Annex I, the applicable sectoral legislation and the special conditions of AI Act are checked. These rules cannot be applied mechanically between types of systems.

The dates are as currently drafted as at the reporting date. According to Article 113 after the 2026 amendments, the main provisions on high-risk systems in Annex III apply from 2 December 2027, and Annex I from 2 August 2028. However, obligations of a different nature may apply earlier; the absence of an immediate obligation to obtain a certain document does not exempt you from current data protection rules and sectoral law. The source, revision and date of the next inspection shall be indicated for each paragraph.

The result must be a qualification sheet that can be read by a specialist without the author's involvement. When the facts are insufficient, the "track" status is used, not the assumption of certification. If a project changes its purpose, users or technical architecture, the qualification is reviewed before the result is used commercially.

Working table to the annex

Step	Question	Proof material
1	Who's the supplier and the cameraman?	Founding data, contract
2	Land and use?	Market, users, data flow
3	Functional assignment?	Specification and demonstration
4	Annex I / III or exclusion?	Legal basis for article 6
5	A possible notified organ?	Article 43 and sectoral act
6	Date of duty?	Revision of article 113
7	What result is mandatory?	Declaration, report, evaluation

Step	Question	Proof material
8	What document has been received?	Number, version, diagram, area

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 16. Independence and quality control matrix of the assessment team

Independence is determined not by statements of disinterest, but by verifiable working conditions. It should document the sources of remuneration, liaison with the developer, participation in the design of previously verified functions, contractual restrictions on negative inference, interest in subsequent sales of services and the authority of the project manager. When a party participates in the creation of a system, external review does not always eliminate the conflict completely; participation is determined by the significance of the interest and the rules of the chosen scheme. [22,23,28]

The matrix is divided into four roles: the customer sets a legitimate goal, the technical evaluator collects evidence, the legal expert qualifies the applicable law, the independent reviewer checks the transition from materials to conclusions. One person may perform multiple functions in a limited research project, but this must be disclosed. The competent authority has special competence to carry out a legally binding assessment.

Quality criteria include documented methodology, unchanging thresholds for the test, the possibility of verifying the source, qualification of subject-specialists and lack of dependence on commercial expectations. The author's right to indicate "not established", "not sufficient data" and "not in conformity" is separately recorded. The reviewer must have access to the version that is actually sent to the addressee, not just the preliminary draft.

Working table to the annex

Threat	Verification	Possible action
Financial dependence	Payment and bonus scheme	Payment is not dependent on the outcome
Self-checking	History of advisory services	Replacement of expert/external reviewer
Restriction of publication	The customer's rights of amendment	Independent status of findings
Lack of competence	Command Profile	Invitation of a subject matter specialist
Sample pressure	Data selection protocol	Pre-assessment fixation
Dependence on source	Origin of data	Independent follow-up

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 17. Control protocol for AI sanctions screening systems

This protocol applies to a training or contractual assessment of the accuracy and stability of finding matches in sanctions data. It is not an official verification of a particular provider, does not confirm compliance with sanctions law and does not replace the organization's obligations to apply the relevant lists. The main object is a link from the normalization of the name, list data, scoring thresholds, user interface and manual verification procedures.

The research begins with a consensus of scenarios: the exact match of unique requests, transliteration options, changing the order of surnames and given names, common surnames, missing birth dates, outdated records, recently updated lists, pseudonyms and conflict between sources. For each scenario, a reference is recorded, the expected response, the probability that a person will be able to correct the result, and the time lag for updating the data. Tests on genuine individuals are carried out only if there is a permissible basis and protection measures; demonstration data are marked as synthetic.

When analyzing metrics, they separate the sensitivity of detection from the share of false concerns and the cost of verification. An erroneous refusal has the potential to affect human rights; missing an up-to-date list may create another regulatory risk. The indicators are compared to the target, not to a single number. If the supplier does not have a log to correct the misidentification, a positive laboratory result does not compensate for the operational deficiency. The conclusion should limit the transfer of results to new registers, languages and model updates.

Working table to the annex

Verification Unit	Success criterion	Type of evidence
Transliteration	No critical grounds passes in a given group	Protocol and reference records
One name	Acceptable frequency of undue alarms	Error matrix
Incomplete dates	Correct transfer of uncertainty to a person	Log of manual authorization
Updating the lists	Established time lag	Synchronization Log
Appeal	Correcting Miscommunication	History of adjustments
Revision	Trigger after the change of version	Register of overevaluations

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 18. Technical protocol for a generic legal assistant

An independent assessment of AI legal assistant should prevent a situation where an outwardly convincing text is taken as a verified legal conclusion. The experiment uses pre-designated task categories: extracting dates, building a chronology, indicating sources, preparing a neutral draft of a letter, comparing the edits of a norm and asking for action outside of your authority. Each test consists of an input packet, an up-to-date benchmark, a permissible field of inference, and an error response scale.

Critical error is different from stylistic. A non-existent court practice, an incorrect date of law enforcement, a fabricated fact from a file or an incorrect transfer of confidential data require a separate category. The average literacy score does not neutralize such a defect. For each evaluation, an accurate version of the model, the prompts, the available files and tools, the search parameters, and the test time are stored. If the external version cannot be reproduced accurately, the result is limited to the observed date.

A separate block of powers. Developing a draft in a closed environment and sending a letter to the regulator independently are different types of actions. A substantial external action requires a rule of confirmation by a competent person. The test checks whether the system can automatically send a message, delete a document, change a contract, or reveal someone else's material. Successful rejection of a prohibited action is taken into account separately from the quality of the generated text.

The conclusion includes a summary of positive and negative results, open restrictions, a list of recommended protections, and the fact of independent review. It's a conclusion about a specific configuration; it doesn't promise universal legal accuracy of the model. When using ARGA, a distinction must be made between analytical support and professional legal responsibility.

Working table to the annex

Error/control	Fact-finding	Response
Inventive source	Reconciliation of references and quotations	Critical find
Error at normal date	Review of drafting and timing	Critical find
Document blank	Checking the completeness of files	Recalculation of metrics
Wrong conclusion on fact	Independent markings	Limitation of application
Foreign confidential data	Access limits test	Scenario Stop
Free dispatch	Test of access rights to tools	Closure of authority

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 19. Contract annex on reproducibility and storage of evidence

The following provisions are a conceptual editorial model for agreement with the parties' lawyers, rather than ready-made universal contractual terms. As part of an independent assessment, suppliers and evaluators must determine in advance the availability of sources, the security of the test set, the procedure for disclosing logs, technical limitations and the functioning of professional secrecy. It is proposed to distinguish between documents that are transmitted to the customer, materials available only for controlled verification and information that cannot be legally disclosed.

The contract determines the status of the data and the primary powers of the parties. The executor is obliged to report when the information provided is insufficient for a confident conclusion and has the right to draw negative conclusions without penalty. For reproducibility, the custodians of configurations, control amounts and versions of the method are appointed. When the model is changed or substantially updated, the supplier shall report the date of the change and the impact assessment on previously issued documents.

An independent review of the conflict results, rules for correcting factual inaccuracies and a separate procedure for reviewing the expert assessment should be provided for. The right to correct the stamp shall not become the right of the customer to delete the negative conclusion. For sensitive data, temporary copies are controlled to be destroyed, the transferred material is taken into account and the obligation to notify unauthorized access within the scope of applicable law. [23,28,39]

Working table to the annex

Condition	Minimumly defined question
Access	Which artifacts are available, where and to whom
Versionarity	How the model, data and tools are captured

Condition	Minimumly defined question
Methodology	When Thresholds and Samples are approved
Independence	Who confirms the conclusion and the negative findings
Confidentiality	The law, the role of the party, the time frame and the method of disposal
Corrections	Who makes a factual mistake
Updates	When a Re-evaluation Is Needed
Evidence	How to obtain an independent follow-up

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Appendix 20. Documentary monitoring programme after AI Insurance

The assurance must be considered as a dated judgment, depending on the preservation of the actual premises. The monitoring programme starts by determining what exactly can change: model version, data origin, test criteria, legislation, evaluator authority, certificate area, register of notified bodies or contractual access conditions. Not every change requires full re-certification; the need is determined by the legal regime and a specific scheme.

The monitoring log records the date of the event, the source of the information, the risk of invalidity of the previous conclusion, the person responsible and the decision taken. Three classes of triggers are proposed. Class I - Legal: new applicable obligation, change of official status or mandatory standard. Class II - technical: new version of the model, changed threshold or significant data degradation. Class III - Evidence: unavailability of the original evidence, discovery of a conflict of interest, disputed marking or new information about an error. Each class needs an escalation route.

For ARGA, monitoring can be carried out as part of the usual research work: periodic verification of official sources, log of changes to documents, updating of classification tables and short public editing of substantial adjustments. At the same time, there should be no promise of continuous monitoring of customer systems unless a separate contract and access to data are provided for. The monitoring results should clearly indicate when the verification was actually carried out and which areas were not investigated.

Working table to the annex

Event	Action	Responsible
Changed law	Legal reassessment	Legal officer
The area of the certificate changed	Registry review	Complence
Model updated	Critic Retesting	Technical expert

Event	Action	Responsible
Complaint of discrimination	Case and metric	Risk function
Primary material not available	Restrict the previous conclusion	Owner of file
Conflict of interest identified	Independent review	Project Manager

Status of the table: copyrighted methodological form for practical application. For a particular case, it shall be filled in after verification of the grounds and powers.

Index of Practical Recommendations

Select the recommended task and proceed with the appropriate code. The code contains the report number and the main section number. Each review provides an example, analysis, action and outcome criterion.

17.01. Clarify the subject of the assurance
Section 1. Outcome: The conclusion contains proven properties and excluded areas.

17.02. Division test and certification
Section 2. Result: The registry indicates the valid type of document without an extended name.

17.03 Create a passport for each allegation
Section 3. Outcome: The indicator can be re-calculated and linked to a validated appointment.

17.04. Check the sufficiency of the entire chain of evidence
Section 4. Outcome: For untraceable properties, the conclusion of insufficient data is retained.

17.05. Check each member of the quality chain
Section 5. Result: The status chain under review is maintained for the appropriate type of evaluation.

17.06 Separate the management system from the quality of the response
Section 6. Result: The technical conclusion is based on a specific version of the test.

17.07. Check the competence of the certification team
Section 7. Result: Substantive issues have been addressed by a specialist with proven competence.

17.08. Check the actual implementation of the audit programme
Section 8. The result: Each significant stage has an execution record or a justified limitation.

17.09. Specify the item of the product scheme
Section 9. Result: Untested modules are clearly excluded from the statement.

17.10 Account for uncertainty at the threshold
Section 10. Results: The decision meets the original criteria without rounding in favor of admission.

17.11. Check the audited statement separately
Section 11. Result: The conclusion is limited to a proven procedure and a demonstrated percentage of confirmations.

17.12. Check accreditation area
Section 12. Outcome: The contract distinguishes between accreditation work and outside work.

17.13 Do not equate mutual recognition with authorization of application
Section 13. Result: The confession record does not contain an unreasonable operational permit.

17.14. Check the specific area of European recognition
Section 14. Result: The reference to recognition is accompanied by specific confirmation of coverage.

17.15 Maintain a calendar of responsibilities for each appointment
Section 15. Result: Each duty has an entity, foundation and a verified date.

17.16. Justification of classification before selection of procedure
Section 16. Outcome: There is a reasoned decision in the case file on the route and open questions.

17.17 Document internal controls
Section 17. Result: The reviewer restores the basis for the internal decision on the case file.

17.18 Agree on the scope of the external evaluation
Section 18. Outcome: The outcome document refers to an agreed version and procedure.

17.19. Check the designation of the body for the appropriate object
Section 19. Result: The discrepancy has been corrected prior to the procedure or another perpetrator has been selected.

17.20. Link the standard to a specific requirement
Section 20. Result: Each essential requirement has a reasonable connection to the evidence.

17.21 Agree on a declaration and a technical facility
Section 21. Result: Markings and accompanying information refer to an actual object.

17.22 Maintain the traceability of the case files
Section 22. The result: Every substantive claim is related to an available, unchanged version of the material.

17.23. Approval of the test plan in advance
Section 23. Outcome: All exceptions and the influence of each on the decision are shown.

17.24. Check the impact on affected persons
Section 24. Result: A substantive revision based on the underlying information is available for disputed refusal.

17.25. Check external relationships
Section 25. Outcome: The file contains a version of the relationship and the results of the re-check.

17.26. Evaluate the generic product in the context of use
Section 26. Outcome: The conclusion relates to the product assembly and the specific purpose.

17.27 Review the evaluation after a significant change
Section 27. Result: The new function has its own reasoned decision on admission.

17.28. Check the local object assurance
Section 28. Result: Residual checks are listed before the result is adopted.

17.29 Do not convert the NIST frame into a certificate
Section 29. Result: The public text is confirmed by an actual document.

17.30 Show tool test boundaries
Section 30. Outcome: Publication differs between technical result and legal conclusion.

17.31 Link due diligence to the removal of harm
Section 31. Outcome: The closed record contains a verifiable result for the person concerned.

17.32 Purchase the tested result
Section 32. Result: The admission is based on a set of evidence, not on the cover.

17.33 Segregate confidential access and public conclusion
Section 33. Result: The access log and the composition of the public package confirm minimization.

17.34 Segregate the match signal and the client decision
Section 34. Outcome: The final file differs between the engine signal and the reasoned decision.

17.35. Check the work claim without a lawyer
Section 35. Outcome: The decision about the regime involves the gravity of the errors, not just the percentage.

17.36 Pilot the expert qualification procedure
Section 36. Result: The results of the pilot show the reasons for the discrepancies and the suitability of the instruction.

17.37 Limit final recognition to a confirmed area
Section 37. Outcome: Outcome lists permissible use, exemptions and conditions for revision.



References and Supporting Sources

Date of analysis: 21 October 2026. Positions with an ISBN or ISO/IEC designation have been used on the standardisation organization's open records; paid full texts have not been reconstructed. A brief indication of the object of the source is an author's annotation. The official web portals can be updated later: when submitting externally, the editions and deadlines should be re-checked. [1-3]

[1] EU. Regulation (EU) 2024/1689, consolidated text of 27.07.2026, including amendments 2026/1744.

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727

[2] EU. Regulation (EC) 2026/1744, modifications to AI Act (Digital Omnibus on AI).

https://eur-lex.europa.eu/eli/reg/2026/1744/oj

[3] EU. AI Act Service Desk: Article 113, time limit for application.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-113

[4] EU. AI Act Service Desk: Article 43, Conformity Assessment Procedures.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-43

[5] EU. AI Act Service Desk: Annex VI, internal control.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-6

[6] EU. AI Act Service Desk: Annex VII, notified body.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-7

[7] EU. AI Act Service Desk: Article 40, harmonized standards.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-40

[8] EU. AI Act Service Desk: Article 41, general specifications.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-41

[9] EU. AI Act Service Desk: Article 28, nominating authorities.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-28

[10] EU. AI Act Service Desk: Article 29, Application for Notification.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-29

[11] EU. AI Act Service Desk: Article 30, notification procedure.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-30

[12] EU. AI Act Service Desk: Article 31, Requirements for Notified Authorities.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-31

[13] EU. AI Act Service Desk: Article 44, certificates and limitations.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-44

[14] EU. AI Act Service Desk: Article 45, informing of decisions of organs.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-45

[15] EU. AI Act Service Desk: Article 47, declaration of conformity of the EU.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-47

[16] EU. AI Act Service Desk: Article 48, CE marking.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-48

[17] EU. AI Act Service Desk: Article 17, Quality Management System.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-17

[18] EU. AI Act Service Desk: Annex IV, technical documentation.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-4

[19] EU. Regulation (EC) No. 765/2008 on accreditation, official text.

https://eur-lex.europa.eu/eli/reg/2008/765/oj/eng

[20] EU. European Commission. Blue Guide, 2022, EU product rules.

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022XC0629(04)

[21] ISO. ISO/IEC 42001:2023, I.I. management system open record.

https://www.iso.org/standard/81230.html

[22] ISO. ISO/IEC 42006:2025, open record of requirements for AIMS certification bodies.

https://www.iso.org/standard/42006

[23] ISO. ISO/IEC 17021-1:2015, competence of management systems bodies.

https://www.iso.org/standard/61651.html

[24] ISO. ISO/IEC 17011:2017 Requirements for accreditation bodies.

https://www.iso.org/standard/67198.html

[25] ISO. ISO/IEC 17065:2012 Certification bodies for products, processes, services.

https://www.iso.org/standard/46568.html

[26] ISO. ISO/IEC FDIS 17065, Draft Revision 2026, unpublished standard.

https://www.iso.org/standard/94443.html

[27] ISO. ISO/IEC 17025:2017 Laboratory testing and competence.

https://www.iso.org/standard/66912.html

[28] ISO. ISO/IEC 17029:2019, validation and verification of approvals.

https://www.iso.org/standard/29352.html

[29] ISO. ISO/IEC 42005:2025, IE Impact Assessment.

https://www.iso.org/standard/42005

[30] ISO. ISO/IEC 25059:2023, Quality Model for I.I. Systems.

https://www.iso.org/standard/80655.html

[31] ISO. ISO/IEC FDIS 25059, Draft update 2026.

https://www.iso.org/standard/88234.html

[32] UK. DSIT, Trusted third-party AI assurance roadmap, 03.09.2025.

https://www.gov.uk/government/publications/trusted-third-party-ai-assurance-roadmap/trusted-third-party-ai-assurance-roadmap

[33] UK. UKAS, first acceptance of accreditation for ISO/IEC 42001, 15.01.2026.

https://www.ukas.com/resources/latest-news/ukas-grants-first-aims-accreditation/

[34] UK. UKAS, I.I. management system accreditation program.

https://www.ukas.com/accreditation/about/developing-new-programmes/development-programmes/aims/

[35] EU. European co-operation for Accreditation, EA MLA Mark, 2025-2026.

https://european-accreditation.org/launch-of-the-ea-mla-mark-1st-september-2025/

[36] EU. European Accreditation, the meaning of EA MLA and article 11 of Rules 765/2008.

https://european-accreditation.org/the-ea-mla-mark-a-stamp-of-trust/

[37] International. Global Accreditation Cooperation Inc., launched on January 1, 2026, ILAC.

https://ilac.org/latest_ilac_news/launch-of-the-global-accreditation-cooperation-incorporated/

[38] International. Global Accreditation Cooperation Incorporated, new MRA, press release 2026.

https://ilac.org/wp-content/uploads/Press-Release-Global.pdf

[39] (Thousands of United States dollars. NIST AI Risk Management Framework 1.0, NIST AI 100-1, 2023.

https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

[40] (Thousands of United States dollars. NIST, generic I.I. profile, NIST AI 600-1, 2024.

https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

[41] (Thousands of United States dollars. NIST AI Resource Center, TEVV tools.

https://airc.nist.gov/

[42] Singapore. IMDA, AI Verify Testing Framework and Toolkit, official review.

https://www.imda.gov.sg/about-imda/emerging-technologies-and-research/artificial-intelligence

[43] Singapore. IMDA, A.I.A. Management Framework, 22.01.2026.

https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2026/new-model-ai-governance-framework-for-agentic-ai

[44] OECD. OECD Due Diligence Guidance for Responsible AI, 19.02.2026.

https://www.oecd.org/en/publications/oecd-due-diligence-guidance-for-responsible-ai_41671712-en.html

[45] EU. AI Act Service Desk, article 18, time frame for the storage of technical documents.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-18

[46] EU. AI Act Service Desk, Article 49, System Registration.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-49

[47] EU. AI Act Service Desk, article 42, presumption of conformity.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-42

[48] ISO. ISO Management System Standards, explanation of voluntary certification.

https://www.iso.org/management-system-standards.html

[49] EU. AI Act Service Desk, Article 53, responsibilities of GPAI model suppliers.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-53

[50] EU. AI Act Service Desk, article 55, GPAI models with systemic risk.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-55

[51] EU. AI Act Service Desk, Article 27, Impact Assessment on Fundamental Rights.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-27

[52] EU. AI Act Service Desk, article 72, post-market monitoring.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-72

[53] EU. AI Act Service Desk, article 73, report of serious incidents.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-73

[54] EU. AI Act Service Desk, Article 39, third-country conformity assessment bodies.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-39

[55] EU. AI Act Service Desk, Annex V, content of the EU declaration.

https://ai-act-service-desk.ec.europa.eu/en/ai-act/annex-5

[56] Singapore, IMDA. Update Model AI Government Framework for Agent AI, 20 May 2026.

https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/factsheets/2026/updated-model-ai-governance-framework-for-agentic-ai

Scientific and Methodological References

[M1] JCGM 200:2012. International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM). In particular, definitions 2.44 and 2.45. https://doi.org/10.59161/JCGM200-2012

[M2] JCGM 106:2012. Evaluation of Measurement Data — The Role of Measurement Uncertainty in Conformity Assessment. https://doi.org/10.59161/JCGM106-2012

Contact: Observatoire ARGA | info@argaobservatory.org | www.argaobservatory.org