Observatoire ARGA

INTERNATIONAL ANALYTICAL REPORT

Study #25

Artificial intelligence and the protection of fundamental rights

Impact assessment, discrimination and complaints mechanisms

Organization: Observatoire ARGA

English version

21 rue de l'Aviation, 64600 Anglet, France

info@argaobservatory.org | www.argaobservatory.org

October 23, 2026

Contents

Research Summary....4
1 Subject of the study and method of analysis....4
2 Fundamental rights and limits of permissible interference....5
3 The AI Act calendar and the independence of other legal regimes....6
4 Qualification of a high-risk system....7
5 Impact assessment under Article 27....8
6 FRIA and DPIA in a single management system....9
7 Automated decisions and meaningful human control....10
8 Explanation of the decision and access to evidence....10
9 International guidelines and the limits of their legal force....11
10 France and access to the grounds of the administrative decision....12
11 Comparative benchmarks of the UK and the US....13
12 Map of affected people and participation of interested groups....13
13 Discrimination and quality of problem formulation....14
14 Equality Test Data and Special Categories....15
15 Test sample and quality of the reference result....15
16 Error Metrics and the Impossibility of a Single Universal Fairness Assessment....16
17 Causal Hypotheses and Limits of Statistical Inference....17
18 Severity of harm and decision on the acceptability of residual risk....18
19 Independence of the audit and contractual terms....18
20 Police data and international cooperation....19
21 Migration and asylum....20
22 Banking services and sanctions control....20
23 Employment and availability of algorithmic selection....21
24 Social Services and Digital Exclusion....22
25 Appeal and selection of a competent addressee....22
26 Preservation of evidence and reproducibility....23
27 Urgent protection and correction of consequences....24
28 System procurement and supplier relations....25
29 Post-implementation monitoring....25
30 Public reporting and the right to object....26
31 Proposed model of independent control of the Observatoire ARGA....27
32 Practical implementation program....27
33 Proposals for International Consultations....28
34 Conclusion....29
Appendix 1 Passport of the application being evaluated....31
Appendix 2 Legal Calendar Card....31
Appendix 3 Map of Rights and Harm Scenarios....32
Appendix 4 Participation Plan for Affected Groups....32
Appendix 5 Sample and standard passport....33
Appendix 6 Calculation of group differences on training data....33
Appendix 7 Human Control Verification Protocol....34
Appendix 8 Impact Assessment Worksheet....35
Appendix 9 Register of Evidence....35
Appendix 10 Structure of the Request for Explanation....36
Appendix 11 Complaint Review Card....36
Appendix 12 Conflict of Interest Map....37
Appendix 13 Matrix of contractual requirements for the supplier....37
Appendix 14: Study Case on False Financial Matching....38

Appendix 15: Study case on translation error in migration procedure....39
Appendix 16: Educational file on unavailable personnel test....39
Appendix 17: Study case on digital denial of social services....40
Appendix 18 Register of Corrective Measures....41
Appendix 19 Structure of the independent opinion....41
Appendix 20 Checklist before external distribution....42
Sources and regulatory materials....44

Research Summary

Assessing the impact of artificial intelligence on fundamental rights makes practical sense if it allows for a change in decision regarding the use of a system, preventing harm, and restoring the rights of individuals. A list of risks that is not linked to the power to stop the system, access to evidence, and an appeals procedure does not ensure this outcome. This study proposes a unified evidentiary procedure for organizations using AI, independent auditors, and representatives of affected individuals.

The focus of this study is the use of AI systems in decision-making regarding individuals: service provision, candidate selection, lending, customer screening, immigration review, and law enforcement analysis. The report covers legal grounds, discrimination detection, human oversight, independent verification, and decision-correction mechanisms. It does not claim that any specific organization uses a particular model or violates applicants' rights. The practical situations and calculations below are for educational purposes only.

The main conclusion is the need to link four objects: a specific right, a mechanism for potential violation, verifiable evidence, and a protective measure. The technical accuracy of a model does not answer the question of the permissibility of intervention. A management system certificate does not confirm the legality of each individual decision. The presence of an employee in the process does not prove the meaningfulness of human oversight. All these circumstances require independent assessment.

The legal deadline is set for 23 October 2026. The version used for the AI Act reflects the amendments to Regulation (EU) 2026/1744. The general deadlines for the application of the requirements of Sections 1–3 of Chapter III have been postponed: for systems under Article 6(2) and Annex III – to 2 December 2027, for product systems under Article 6(1) and Annex I – to 2 August 2028. This is significant for Article 27 on fundamental rights impact assessment. However, the independent requirements on data protection, equality, administrative procedure and judicial protection continue to apply; the special conditions for the use of individual biometric systems cannot be postponed by analogy with the general calendar. [1–3]

The proposed ARGA Observatoire model includes preliminary legal qualification, a map of affected groups, data quality analysis, testing on a relevant sample, verification of actual employee behavior, consideration of individual complaints, and periodic review of the assessment. The independent review should have a defined mandate and transparent limitations. It does not grant the Observatoire state oversight powers and does not replace the decision of the competent authority.

1 Subject of the study and method of analysis

The subject of the evaluation is not the model as an abstract software object, but the decision made using it. The same technology can sort internal correspondence, select applications for review, or effectively determine a person's access to a bank account. In the first case, an error is sometimes limited to a delay in work; in the latter, it can interfere with wages and housing payments. Hence the need to describe the intended use before selecting metrics and legal qualifications.

The proposed unit of study is the sequence "input data - processing - system output - organizational action - human consequence." For each link, the decision maker, the data available to them, and the

possibility of error correction are identified. This approach allows us to distinguish between a source defect, a matching error, a questionable threshold, a misreading of the result, and an independent decision by the employee. Without this, the organization risks correcting the model when the problem is in the instructions, or blaming the operator for not having the necessary information.

The method involves analyzing official regulatory texts and court decisions, comparing institutional approaches, and developing independent verification procedures. The application forms are proposals from Observatoire ARGA and are not intended to be mandatory government templates. The numerical examples serve to verify the logic and do not describe the empirical results of existing projects.

The study distinguishes three levels of assertions. The legal level establishes the content of the norm and the conditions of its application. The factual level records the observable behavior of the system and its participants. The evaluative level explains why the evidence supports a certain conclusion and what alternatives remain. In the final conclusion, an unknown fact cannot be replaced by a normative assumption: the obligation to maintain a log does not prove that the log exists and contains the required entry.

The audit is limited to the accessible object. Analysis of the demo version does not confirm the state of the system as of the date of the contested decision. Extracting data from one department does not allow for an assessment of the entire organization. Lack of access to the source data does not necessarily render the study useless, but it does limit the permissible scope of the conclusion. These limitations are included in the substantive conclusion, not hidden in a general disclaimer on the last page.

Object of verification	The main question	Minimal evidence
Solution	What has changed for man?	Notice and actual consequence
System	Which version influenced the result?	Version ID and execution log
Data	What was known at the time of the decision	Snapshot of input data with provenance
Employee	Could he have changed the outcome?	Authority, access and recording of reasons
Protection	How to fix the error	Review procedure and confirmation of execution

2 Fundamental rights and limits of permissible interference

The impact of AI on rights is not limited to privacy. It can also affect dignity, equality, privacy, personal data protection, freedom of expression, access to employment, education, social support, and effective legal remedies. In immigration and criminal proceedings, liberty, security, fairness of proceedings, and the prohibition of treatment incompatible with absolute guarantees are also important. The legal map is drawn according to jurisdiction and facts. [4–6]

For each measure, the legal basis and the competence of the actor are first established. Next, the legitimate objective, the suitability of the chosen means, the necessity of the intervention, and its proportionality are examined where the relevant law allows for limitations. The technical feasibility of increasing the risk detection rate does not constitute an independent legal basis. Saving time may be important when choosing an organizational model, but it does not make prohibited treatment permissible.

The evaluation of alternatives must be specific. Available methods for achieving the same goal are compared: process modification without AI, limited automation, use of less sensitive features, additional document verification, or shorter retention periods. The statement "manual work is too expensive" does not answer the question of whether it is possible to eliminate automatic rejection and leave the system with only preliminary sorting. At the same time, the expert should not assume that a manual procedure is inherently fair: its errors are also measurable.

Particular attention is required to the cumulative effect. Minor delays in multiple systems can deprive a person of access to an account, housing, and medical services simultaneously. A single organization may only see its own stage, while the applicant faces a chain of denials. Therefore, when collecting information, subsequent transfers of the risk assessment and the use of the initial conclusion for new purposes are recorded. Verifying the legitimacy of each transfer does not negate the analysis of the cumulative harm.

The EU Charter applies within the limits of its Article 51; it cannot be automatically declared the basis for any private dispute outside Union law. The European Convention on Human Rights, national constitutional guarantees, and sectoral legislation have their own terms of application. In a cross-border case, it is necessary to specify which obligation applies to the state, which to the organization, and the procedural method by which it can be implemented.

3 The AI Act calendar and the independence of other legal regimes

The version of the AI Act as of 27 July 2026 amends the original 2024 calendar. Article 113 retains the general date of application of 2 August 2026, but contains specific exceptions. For the requirements of Sections 1–3 of Chapter III, with the exception of Article 6(5), the dates are 2 December 2027 for high-risk Annex III systems and 2 August 2028 for Annex I systems. Article 27 is located in Section 3; therefore, it would be incorrect to represent its general obligation as already applicable to all relevant operators in October 2026.[1–3]

This does not imply a moratorium on the protection of rights. The lawfulness of personal data processing, limitations on automated decisions, prohibitions on discrimination, and requirements for administrative justification are independent considerations. An organization may be required to conduct a DPIA under Article 35 of the GDPR even if the general deadline for applying an FRIA under Article 27 of the AI Act has not yet begun. A voluntary early FRIA can also serve as a reasonable preparation measure if its status is clearly defined.

Provisions regarding prohibited practices and narrow exceptions to prohibitions are being reviewed separately. Specifically, the requirements for the permissible use of real-time remote biometric identification cannot be assessed solely under the deferred Chapter III: the special construction of Article 5 contains its own conditions, including an impact assessment in the specified cases. The

new provisions for 2026 also have separate dates. A section-by-section review is required for a specific system, rather than a simple "AI Act not yet in force" check.

It is necessary to distinguish between the effective date of a regulation, the date of application of a specific obligation, the transitional rule for a previously introduced system, and the date of a material change to the product. The mere fact of purchasing before the general effective date does not create a permanent exemption from any obligation. Likewise, updating a model does not always automatically constitute a legally material change; an assessment of the content and terms of the relevant regulation is required.

For practical purposes, it is recommended to maintain a schedule of deadlines, in which each conclusion is linked to a specific article, revision, and object characteristic. When the schedule changes, not only the report summary is updated, but also contract forms, supplier requirements, user notifications, and the audit program. A legislative proposal, a political agreement, and a published, effective act should have different statuses.

Question for October 23, 2026	Working conclusion	What to check separately
General regime of Chapter III for Appendix III	Application of the relevant sections from December 2, 2027	Special rules and transitional provisions
High-Risk Product Systems Appendix I	The corresponding date is August 2, 2028	Applicable industry legislation
GDPR and national guarantees	Maintain independent action	Basis, subject, territory and exceptions
Early voluntary FRIA	Can be used as an internal procedure	Don't pass it off as government approval.
Biometric exceptions to the prohibitions	Article 5 requires separate verification.	Permitted purpose, powers and guarantees

4 Qualification of a high-risk system

The AI Act provides for different grounds for recognizing a system as high-risk. The product route is linked to the legislation in Annex I and the third-party compliance assessment it provides. The other route relates to the areas of application listed in Annex III. The product name, the presence of a neural network, or the advertising term "compliance" do not in themselves determine the status. The purpose and actual role of the system in the decision-making process are examined. [1, 7]

For applications listed in Article 6(2), a limited exemption from recognizing a system as high-risk is provided if the conditions of Article 6(3) are met. The term "preparatory task" cannot be used as a blanket exemption: the significant impact on the outcome, the nature of the risk, and the fulfillment of the stipulated conditions are assessed. A system that formally only issues a recommendation may, in practice, determine the decision if staff do not have the time and resources to review it. Profiling individuals within the relevant systems of Appendix III is of particular importance for qualification.

The independent auditor's job is to examine the documentation and the actual process, not to assign a general legal classification to the entire software suite. A single suite may contain a standard document search, a prioritization algorithm, and a trustworthiness assessment module. The risks and regulatory bases for these functions differ. Too broad a description conceals a dangerous module; too narrow a description artificially disconnects the model from the consequences of its use.

In the area of AML and sanctions control, it is especially important not to automatically label any system as high-risk under Appendix III. Creditworthiness, insurance assessment, and certain law enforcement tasks are listed separately; internal financial monitoring requires its own qualifications. However, the absence of a specific AI Act category does not negate the GDPR, financial regulations, or obligations to prevent unlawful discrimination.

The qualification must result in a reasoned assessment. It describes the intended use, affected parties, the role of the supplier and operator, the applicable clause, any possible exceptions, and the consequences of changing functions. If the evidence is insufficient, the uncertainty and a list of required documents are recorded. A "low risk" declaration without access to the essential characteristics does not constitute a complete assessment.

5 Impact assessment under Article 27

Article 27 of the AI Act does not provide for an FRIA for any owner of any algorithm. Its design covers certain operators of Article 6(2) systems: public authorities, private entities providing public services, and operators of the creditworthiness and risk assessment and pricing systems for life and health insurance specified in Annex III. An exception is provided for the scope of paragraph 2 of Annex III. The qualifications of the addressee are as important as the qualifications of the technology. [1, 8]

The regulatory core of the assessment includes a description of the application process, the duration and frequency of use, the categories of affected individuals and groups, specific risks of harm, human oversight, and response measures should the risk materialize. The assessment is related to initial use and must be updated whenever material circumstances change. The article requires notification of the supervisory authority of the results within the established framework. The presence of an assessment and notification does not constitute court approval or a universal license for the system.

The practical format proposed in the report expands this core with evidentiary fields. For each risk, the source of information, the proven mechanism, the certainty of the conclusion, the responsible person, and the conditions for review are specified. Thus, the risk of denial of a social service is described through specific data categories, classification error, a human verification step, and the possibility of urgent correction. The phrase "possible discrimination" without further clarification is ineffective in making a decision.

An FRIA should be conducted early enough to allow for a realistic opportunity to abandon or modify the project. If the assessment is conducted after procurement, deployment, and personnel training, the resulting costs may create pressure on those who draw the conclusion. This doesn't render a late assessment pointless, but it does require identifying organizational conflicts and determining who has the right to suspend use contrary to commercial expectations.

Reusing an assessment is permissible only after verifying the comparability of contexts. The same model in different municipalities may process different documents and serve groups with different access to digital channels. Transferring findings without analyzing the environment turns an impact assessment into a product description. The new report should clearly indicate which sections were borrowed, which were re-examined, and why the remaining findings are applicable.

6 FRIA and DPIA in a single management system

FRIAs and DPIAs have overlapping but distinct subject matters. A DPIA under Article 35 of the GDPR applies to personal data processing operations that may pose a high risk to rights and freedoms. A FRIA examines the impact of certain uses of AI on fundamental rights. The absence of personal data in a technical module does not preclude the impact on rights; the presence of a DPIA does not mean that service availability, discrimination, procedural safeguards, and the distribution of powers have been examined. [5, 8]

It makes practical sense to use a common set of facts and different legal opinions. A single data structure, purpose description, and list of participants can support both assessments. However, the final form must allow for verification of the relevant grounds, mandatory elements, the decision to consult with the data protection authority, the role of the DPO, and the terms of notification under the AI Act. A single, undifferentiated document makes it difficult to determine which obligation has actually been fulfilled.

An example is the pre-triage system for requests. The DPIA examines the legality of text use, sensitive information, access, storage, and security. The FRIA additionally checks whether requests from people with limited language proficiency are chronically "non-prioritized," whether another channel is available, and who handles the urgent situation. Data protection and protection from de facto exclusion reinforce each other here, but they do not replace one another.

Within the organization, it's necessary to identify the process owner, rather than delegate everything to the DPO. The data protection officer provides legally required advice and exercises appropriate oversight; the decision to launch is delegated to the competent manager. The technical team confirms the system's features and limitations, the industry specialist verifies the correctness of the procedure, and the independent reviewer verifies the sufficiency of the evidence within their scope of work.

A change matrix is useful when conducting joint assessments. A new data source may require a review of the legality of processing and assessing discrimination. Changing the interface without retraining the model can eliminate human oversight. A new country of use changes the legal regime and accessibility to appeal. Therefore, updating should not depend solely on the model version number.

Component	DPIA	FRIA
Central object	Operations with personal data	Use of AI and its impact on rights
Typical evidence	Data schema, foundations, security	Decision context, groups, consequences and measures
General part	Objectives, participants, sources, access	Uses the same proven facts

Component	DPIA	FRIA
Independent conclusion	Data protection compliance	Acceptability of impact and effectiveness of safeguards
Revision	Change in processing risk	Change of use or impact

7 Automated decisions and meaningful human control

Article 22 of the GDPR applies to decisions based solely on automated processing and producing legal or similarly significant effects. All elements are required for the rule to apply, and exceptions have their own conditions and guarantees. It cannot be argued that all use of AI is prohibited, or, conversely, that the presence of a confirmation button is sufficient to opt out of Article 22. The nature of genuine human involvement is essential. [5, 9]

In SCHUFA, C-634/21, the Court of Justice of the EU considered a situation where an automatically generated score plays a decisive role in a third party's decision. The significance of this approach for independent verification lies in the need to examine the entire chain, including the recipient of the score. The division of labor between the score provider and the decision-making body should not automatically eliminate the protection afforded to the affected party. [10]

The effectiveness of controls is verified through employee access to primary data, understanding of limitations, time for analysis, and authority to change the results. If an employee only sees a red indicator, doesn't know its basis, and must confirm hundreds of cases per shift, a formal instruction on independence provides little proof. At the same time, a low cancellation rate doesn't in itself indicate a lack of control: the verified results may indeed be accurate. A random check of the rationale and alternatives is needed.

It is proposed to examine complex cases separately. The sample includes borderline results, incomplete data, language differences, and contradictory documents. The employee is presented with situations where the automated recommendation clearly requires additional verification. Not only the final responses are assessed, but also the time spent, requests for additional information, and the ability to detect errors. Mock audits must be organized legally and not worsen the situation of real applicants.

For citizens, the outcome of the review is important: the opportunity to raise objections, have a competent person review the review, and obtain a correction. A review, in which the same algorithm reprocesses unchanged data, doesn't become humane just because the response was signed by an employee. The file records new information, the reasons for the decision, and the actions taken to address the consequences of the initial error.

8 Explanation of the decision and access to evidence

An explanation must answer the question of why a specific person obtained a specific result. A general presentation of machine learning principles does not replace information about the data used, the role of the system, and the underlying rationale behind the decision. However, explainability does not always require public disclosure of the entire source code. The required

scope is determined by applicable law and the ability to effectively verify and challenge the result. [5, 11]

In Dun & Bradstreet Austria, C-203/22, the ECJ clarified the importance of substantive information about the logic of processing. The mere provision of a complex formula does not make the process understandable. Invoking a trade secret should also not automatically terminate scrutiny: a mechanism for assessing the relevant interests by a competent authority or court is required. In practice, this means preparing two coherent layers of explanation—one that is understandable to the affected party and one that can be technically verified by an expert. [11]

Article 86 of the AI Act has its own conditions: the operator's decision must be based on the output of a specific high-risk system in Annex III and must create a significant adverse effect as provided for in the regulation. Point 2 of Annex III is excluded from the envisaged scope; restrictions under Union and national law, as well as the relationship with existing rights of explanation, are taken into account. The general deferral of sections of Chapter III cannot be mechanically transferred to all provisions of other Chapters. [12]

It is suggested that the evidence package be composed of the solution identifier, timestamp, system version, actual input data, threshold, employee role, and the reasons for the resulting action. If exact replication is impossible, the reasons are described: data deletion, non-determinism, external service change, or configuration loss. The subsequent explanation using the language model should not be presented as a reliable record of how the original system actually arrived at its result.

It's useful to test a personal explanation. After receiving it, the recipient should understand what information is incorrect, what circumstances were overlooked, where to submit an objection, and what will happen before it's considered. If the document consists solely of phrases like "the decision was made in accordance with internal policy," the right to appeal remains formal. The quality of the explanation is assessed based on its content and accessibility, including language, disabilities, and lack of digital skills.

9 International guidelines and the limits of their legal force

The Council of Europe's Framework Convention on Artificial Intelligence links the use of AI to human rights, democracy, and the rule of law. Its provisions on safeguards, procedural guarantees, and risk assessment are important for building an institutional model. However, the treaty's commitment must be correlated with the participation of a specific state, the treaty's entry into force for that state, and national implementing measures. Signature should not be presented as complete ratification or an immediate universal obligation for every company. [13]

HUDERIA provides a stand-alone, non-binding impact assessment methodology. The 2026 update includes COBRA contextual analysis resources. Their practical value lies in the consistent consideration of the conditions of application, development, and implementation. Adherence to the methodology does not constitute a certificate of compliance with the Convention and does not transform a private assessor into a Council of Europe body. This report utilizes its general principle: the properties of a technology are assessed together with the institutional environment. [14–15]

The UN Guiding Principles on Business and Human Rights distinguish between the State's duty to protect rights, the responsibility of businesses to respect them, and access to legal remedies. For AI, this means considering the impact not only of its own operations but also of its relationships with

suppliers and customers. Contractually allocating due diligence costs does not eliminate the human impact. However, the Guiding Principles cannot automatically infer specific civil liability without an applicable national basis. [16]

The NIST AI RMF offers a voluntary organizational risk management framework. It is useful for assigning responsibilities and conducting regular assessments, but does not replace legal qualifications. The Observatoire Work Map can match management tasks with rights and evidence, maintaining the distinction between voluntary practice and legal obligation. An overall assessment of an organization's maturity should not replace an analysis of an individual decision. [17]

It's useful to compare international instruments based on practical issues: who evaluates, when, based on what data, with whom, who can stop implementation, and who provides the remedy. Such a table is more useful than ranking countries by the number of published principles. For a cross-border project, it also shows which elements can be reused and which must be prepared separately for each jurisdiction.

10 France and access to the grounds of the administrative decision

In France, Article L311-3-1 of the Code of Public and Administrative Relations provides for the information of the person concerned about an individual decision taken on the basis of algorithmic processing and the provision, upon request, of the rules and main characteristics of its application, taking into account the envisaged limitations. Subordinate provisions specify the content of the information, including the input of the algorithm, the processed data, the sources, and the parameters applied. This is an independent legal route that should not be confused with the future general application of the FRIA. [18]

When preparing a request, the existence of an administrative decision and the role of processing in its adoption must first be established. The request must be specific to the decision, date, agency, and stage of interest. Requesting "all the administration's artificial intelligence" does not help obtain relevant information. It is practical to request the applicant's input data, its origin, the role of the automated result, information about human review, and how the error was corrected.

Access to administrative documents, the rights of the data subject, and appealing the decision itself constitute distinct procedures. A response to one does not necessarily authorize the other. Timeframes and the competence of the addressees must be carefully considered: a request for documents should not automatically be considered a complaint about a denial of service, nor should an appeal to the CNIL be considered a guarantee of suspension of the administrative decision. For each issue, a legally prescribed mechanism is selected, and the procedural feasibility of urgent protection is verified.

When working with an independent organization, it's important not to promise the applicant access to any closed algorithms. Legal restrictions, the need to protect third parties, and special rules for certain areas are possible. However, the restriction must be justified and sufficiently specific. If the agency only provides general instructions, it should be verified that they allow for an understanding of how the processing applies specifically to the applicant and formulate any missing questions.

Documenting the response has value in its own right. The text of the request, confirmation of receipt, response, list of attachments, and reasons for refusal are all preserved. The expert evaluates not only what was disclosed but also what material circumstances remained unverified. However,

the absence of a response does not automatically prove discriminatory intent; it may confirm a procedural problem and a limitation of available remedies.

11 Comparative benchmarks of the UK and the US

The UK's automated decision-making regime should be assessed in light of the Data (Use and Access) Act 2025 and the latest ICO guidance. The reform expands the use of general processing grounds for significant automated decisions while maintaining safeguards and enhanced protection for special categories of data. Therefore, simply copying the original text of Article 22 of the EU GDPR into the UK dossier creates a risk of error. For each specific decision, the current version, date, and nature of the data are checked. [19–20]

What's of comparative interest is not the formal leniency or strictness of the system, but rather the accessibility of information, the right to present a position, challenge the result, and request human intervention. These guarantees require organizational implementation. Expanding the permissible grounds for automation does not mean that applicants need only declare a legitimate interest and decline to explain the material outcome. An independent review should examine the content and speed of objection resolution.

In the United States, the rights of affected individuals are often determined by industry-specific and territorial regulations. EEOC documents confirm the applicability of existing anti-discrimination protections to the use of AI in employment; Department of Justice documents specifically address the risks for individuals with disabilities. When evaluating hiring, the accessibility of the test method, reasonable accommodations, and the connection of the assessed characteristic to the job function are important. Purchasing a tool from an external provider does not make the impact on the candidate solely someone else's problem. [21–22]

The report does not claim the existence of a single American equivalent of the FRIA for all industries and states. An external opinion requires a map of applicable legislation, including the terms of the relevant act and permissible remedies. A voluntary audit or public technical report should not be passed off as a government permit. Comparisons with the EU should preserve the differences in the competence of agencies, evidentiary rules, and the role of private lawsuits.

The general practical conclusion is that a single factual dossier can be used across different legal pathways. The model version, input data, revision history, and human consequences are useful in any jurisdiction. However, legal requirements, deadlines, and request formulations cannot be automatically transferred. International cooperation should facilitate fact-checking while preserving the accuracy of national legal qualifications.

12 Map of affected people and participation of interested groups

Impact assessments begin with the people who will be affected by the process, including those who are not registered users of the product. An applicant may be unaware of the system's existence, a family member may experience the consequences of a benefit denial, a representative may expend resources to correct an error, and an employee may be under pressure from an automated performance metric. The map should cover direct and predictable indirect consequences without transforming any remote effect into demonstrable harm.

For each group, the conditions for accessing the service, the ability to opt out of automation, language and physical limitations, dependency on the solution, and access to legal assistance are described. Vulnerability is considered contextually. A person may be an experienced user of technology but lack the ability to challenge a decision in a foreign language or be subject to time constraints. The juxtaposition of the "regular user" and the unchanging group of "vulnerable" often obscures the real mechanisms of exclusion.

Stakeholder participation should influence the research questions. Post-project consultation is not a substitute for early risk discussion. It is recommended to determine in advance which decisions can still be changed, how the technology is explained to participants, and how responses to comments are recorded. Group disagreement with the project does not necessarily constitute a legal veto, but it requires a reasoned review rather than a mere mere meeting note.

Participation should not become a new source of pressure. Access to social services should not be tied to mandatory consent to disclose experiences of discrimination. Sensitive data should be minimized, secure reporting methods should be provided, and the research should be separated from the applicant's personal file. If feedback is anonymous, the limits of fact-checking should be clearly described, without automatically dismissing it.

FRA materials demonstrate the importance of context in areas such as asylum, education, and employment. For the Observatoire's own program, this justifies incorporating real-world process observation into its evaluation, rather than limiting it to interviews with the developer. At the same time, the findings of another study should not be presented as the results of one's own survey or generalized to a specific organization without data. [23]

13 Discrimination and quality of problem formulation

Discriminatory effects can occur before the model is trained. An organization chooses a goal, historical success rates, acceptable features, and the costs of various errors. If the target reflects not human needs but the ease of processing their documents, the model may effectively replicate undesirable practices. Therefore, the question of "what the system predicts" precedes the question of its accuracy.

A distinction must be made between technical bias, statistical inequality, and legally established discrimination. Differences in results between groups require investigation, but in themselves do not complete the legal qualification. This depends on the protected characteristic, the applicable norm, the nature of the difference, the permissible justifications, and the distribution of the burden of proof. Equal average results also do not prove the absence of a violation of a specific right.

Direct use of a protected attribute and the use of a correlating indicator pose different evidentiary challenges. Postal code, employment history, speech patterns, or document type may indirectly reflect social status or origin. Removing an explicit field does not guarantee neutrality. However, an examiner should not declare any correlating indicator illegal: the mechanism, purpose of use, and applicable restriction must be established.

Testing an alternative problem statement is useful. A system trained to predict the "probability of an application being rejected" can reproduce previous administrative decisions, including their errors. Sometimes, a more appropriate task is to identify a specific missing document and provide the

opportunity to supplement it. Such a change can mitigate harm without creating a universal model of human moral evaluation.

The study of each indicator identifies its origin, connection to the objective, the availability of correction, and the risk of replacing a legally significant circumstance with a convenient proxy. For example, continuous employment may be linked to the requirements of a specific position, but should not be converted without analysis into a general indicator of reliability across all services. The more serious the consequence, the more important is the justification of the causal and normative connection.

14 Equality Test Data and Special Categories

Assessing differences between groups sometimes requires data that itself enjoys enhanced protection. This problem cannot be resolved by either one of two extremes: collecting any sensitive data "for the sake of fairness" or declaring equality assessments impossible without analyzing available methods. A specific purpose, applicable justification, verification of necessity, and access restrictions are necessary.

In the current version of the AI Act, Article 4a regulates the exceptional processing of special categories of data for the purpose of detecting and correcting bias. It sets forth a set of conditions, including strict necessity, the impossibility of effectively achieving the purpose with other data, protection against reuse, access restrictions, and erasure at a specified time. The provision, in its construction, also applies to certain other providers and processors. It does not create a general obligation to conduct any such collection and does not constitute permission to freely transfer data to an external auditor. [24]

This is especially important for independent verification: a researcher cannot consider their own interests sufficient grounds for obtaining data on people's health or ancestry. Controlled access within the data holder's infrastructure, the transfer of aggregated indicators, the use of anonymized samples, and other approaches are possible, provided there are legal grounds. The choice of solution depends on whether it allows for the verification of the required hypothesis, not on the ease of file sharing.

Self-identification of a group, administrative records, and probabilistic algorithmic inference are not interchangeable. Inferring ethnicity from a person's name introduces new errors and inherent risks. If indirect estimates are used, the method, uncertainty, and limitations of the inference must be disclosed. A confident claim about discrimination against a specific group cannot be published when the identity of individuals within that group is unreliably established.

The verification program includes an access log, identifier separation, a minimum set of fields, and a retention period. Small groups require protection from re-identification. Aggregation can reduce the risk of disclosure but conceal an important issue; in this case, a private expert view is chosen over a public, detailed table. The legal regime and evidentiary needs must be agreed upon before data extraction.

15 Test sample and quality of the reference result

A metric cannot be more reliable than the sample on which it is calculated. Validation requires a specific population, time period, inclusion criteria, and selection mechanism. Extracting only

appealed decisions allows for the examination of complaints, but does not allow for an assessment of the error rate across all decisions. Victims' voluntary reports have significant qualitative value, but they cannot be considered a random, representative sample.

An independent assessment must also examine the benchmark against which the system is compared. If the "true" result is considered to be the previous decisions of the same organization, the study can only measure the degree to which its practices are replicated. Disputed cases require a separate review, sometimes by multiple experts, with rules for resolving disputes. Documents obtained after the initial decision are noted separately: they may prove the result to be erroneous, but not necessarily the original action to be unscrupulous.

It's important to prevent information leakage between training and testing. Repeated statements from the same person, duplicate documents, and decisions on related cases can artificially inflate quality. Separating data into individual rows is insufficient when the unit of independence is an individual, household, or organization. The evaluator records the chosen unit and justifies it.

For rare, severe errors, a standard random sample may be insufficient. A separate, targeted sample of complex cases is permitted, but its results are published separately from the overall frequency estimate. Mixing specially selected failures with the main sample distorts the estimate; excluding such cases from the analysis makes the test overly optimistic. Both samples are useful if their purpose is clear.

The minimum data sheet for a set contains the origin, period, criteria, number of units, omissions, duplicates, markup version, and applicability limitations. A verification of the absence of errors in a small sample size does not prove their impossibility. For independent testing and rare events, the "three divided by n" benchmark can be used as an approximate upper 95% frequency limit for zero observations; its assumptions and approximation must be specified.

16 Error Metrics and the Impossibility of a Single Universal Fairness Assessment

A binary solution requires defining what constitutes a positive event. In document verification, a positive result might be a detected violation, while in candidate selection, it might be admission to the next stage. Without a clear definition, identical metric names create conflicting interpretations. The error matrix should be accompanied by a description of the consequences of each type of error for the individual and the organization.

Sensitivity measures the proportion of detected real cases among all such cases in the reference sample. Specificity measures the proportion of correctly missed negative cases. Positive predictive value measures the proportion of positives that are actually confirmed. The latter metric depends on the prevalence of the event: even impressive overall accuracy does not guarantee the reliability of an individual signal in a rare coincidence problem.

In a training sample of 10,000 clients, 100 have true matches. With a sensitivity of 90%, this results in 90 correct signals and 10 missed signals. If the false alarm rate among the remaining 9,900 is 1%, this results in 99 false matches. Of the 189 signals, 90 are confirmed, or approximately 47.6%. This doesn't render the system useless as a preliminary screening tool, but it prevents every detection from being considered sufficient grounds for a final negative decision.

Comparing groups requires several metrics. Equal shares of positive decisions, equal frequencies of certain errors, and identical calibration answer different questions and may conflict with different base rates. The choice of priority is linked to the system's purpose and legal implications. It cannot be concealed behind the mere word "fairness." An explanation of the tradeoffs and a test of less harmful alternatives are needed.

It's practical to publish the numerator, denominator, absolute number of errors, and uncertainty. The difference between 1 error in 10 cases and 100 errors in 1,000 cases has the same point fraction but different statistical reliability. For small groups, one should avoid drawing overly confident conclusions and, at the same time, not exclude them from analysis simply because of measurement inconvenience.

Element of the case study	Number	Significance for the solution
Real coincidences	100	The proportion of the event in the sample is 1%
True signals	90	Requires confirmation of procedure
False signals	99	Risk of unlawful restriction of clients
Missed matches	10	Separate risk to control effectiveness
All signals	189	Positive predictive value 47.6%

17 Causal Hypotheses and Limits of Statistical Inference

The difference discovered between groups does not automatically explain its origin. Differences in initial conditions, document quality, service availability, labeling errors, employee behavior, and the model itself are possible. The expert's task is to consistently test competing explanations. This does not mean seeking justification for every difference; on the contrary, it allows one to identify effective corrective measures and avoid superficial conclusions.

It's impossible to blindly "control" for all available variables. Some indicators may be the result of previous discrimination or an intermediate link in the mechanism being studied. If their influence is automatically excluded, the real problem will disappear from the statistical report, but persist in reality. The selection of variables must be accompanied by a meaningful framework of causal hypotheses and a sensitivity analysis.

A counterfactual test, in which one feature is changed while the others remain the same, is useful for studying system behavior. However, it does not always reproduce a plausible real-life situation. Changing a name, language, or address can affect several related features. Therefore, the result of such a test is considered evidence of a certain sensitivity of the model, rather than definitive proof of any discrimination.

If data are available, it's advisable to compare results before and after a procedure change, use comparable groups, and record external events. However, a simple improvement following the implementation of a new version does not establish its causal effect: the applicant pool, operating

rules, or document quality may have changed. The external assessment should indicate which alternative explanations have been tested and which remain.

A legal conclusion requires separate argumentation. The expert describes the statistics, mechanism, and limitations; the lawyer relates them to the applicable prohibition, the possibility of substantiation, and the evidentiary rules. A conclusion of "a difference requiring verification has been discovered" can be a powerful and useful result if formulated precisely. Artificial confidence increases the likelihood that the entire analysis will be rejected, along with bona fide observations.

18 Severity of harm and decision on the acceptability of residual risk

The risk to rights cannot be assessed solely by the product of probability and monetary damage. Severity, duration, reversibility, the number of affected parties, the ability to defend themselves, and the nature of the right are all relevant. A low probability of an irreversible violation may require a more severe measure than a frequent but quickly remediable technical delay. A numerical scale is useful only as a means of discussion, not as a license to deviate from a mandatory rule.

For each scenario, it's necessary to describe the initiating event, the path to the outcome, and any existing barriers. For example, a false name match in itself doesn't necessarily lead to a failure: an intermediate date-of-birth check can halt the chain. If such a check isn't actually performed, there's no organizational barrier. The registry should evaluate verified measures, not instructions that exist only on paper.

Residual risk is determined after testing corrective measures. Its score cannot be reduced simply because a responsible person has been appointed or training has been scheduled. It is necessary to verify whether the system and employee behavior has changed. Sometimes the correct outcome is to limit the scope of use, prohibit automatic failure, require a second opinion, or refrain from implementation until evidence is obtained.

It is proposed to distinguish four management decisions: permit a specific use; permit with verifiable restrictions; suspend until specific deficiencies are corrected; and refuse a given use. Each decision contains the grounds, limits, owner, and review date. The "agreed with comments" category, without conditions or enforcement, makes it difficult to understand what exactly is permitted.

The independent auditor is not obligated to assume residual risk on behalf of the organization. Their job is to evaluate the evidence and clearly communicate any significant deficiencies. The competent manager makes decisions within the law and within their own authority; the government agency exercises the required oversight. This consistency preserves accountability and does not transform the external report into a vehicle for transferring it to a consultant.

19 Independence of the audit and contractual terms

Independence is defined by the ability to obtain relevant data, choose a methodology, report unfavorable findings, and preserve their meaning after editorial review. The absence of an employment contract with the client does not guarantee independence. Financial dependence, involvement in the development of the system being reviewed, and the promise of a positive outcome can create conflict, even if the research is formally commissioned from an external organization.

The mandate must describe the object, period, issues, access, restrictions, and addressees of the conclusion. The agreement must stipulate the procedure to be followed in the event of a material breach, denial of access, or an attempt to remove the conclusion. Confidentiality may protect legitimate interests, but it should not be transformed into an obligation to sign a knowingly misleading public document. The terms of disclosure to a government agency are determined by applicable law, not by a universal formula.

Payment should not be contingent on the system being deemed safe or on the absence of any comments. Separating the consulting and review teams reduces conflict, but does not eliminate it. If the organization participated in the development of the measures it is subsequently evaluating, this fact is disclosed, and the significant findings are shared with a separate reviewer. In some situations, the assignment should be declined if independence cannot be ensured.

The client is given the opportunity to correct factual errors and submit objections. This does not imply the right to rewrite the expert opinion. The review log contains the initial comment, response, evidence, and the author's final decision. Disagreement with the interpretation can be expressed directly. This procedure protects both the organization being reviewed from inaccuracies and the expert from informal pressure.

The final report should state what was verified and what was not covered. The wording "the system complies with human rights" is too broad for a limited audit. It is preferable to specify specific requirements, verified operations, significant deficiencies, and assurance limits. The public summary should adhere to these limitations, otherwise good technical work could be used as misleading advertising.

20 Police data and international cooperation

In law enforcement, an algorithmic error can impact a person's freedom, the intensity of review, and the treatment of evidence. It's important to distinguish between an information signal, an operational hypothesis, procedural evidence, and the basis for a coercive measure. Sharing a result between agencies does not, in itself, increase its credibility. Each new use must preserve its origin, limitations, and the possibility of correction.

Competent authorities processing data for the purposes of preventing, investigating, detecting, or prosecuting criminal offences or enforcing penalties must review the provisions of Directive (EU) 2016/680 and national legislation. It is not possible to automatically describe all police processing solely in terms of the GDPR. Article 11 of the Directive contains separate provisions on automated individual decisions and guarantees; the conditions for processing special categories of data are also structured separately. [6]

In an international search, an algorithmic match of a name or image does not establish an individual with the degree of certainty required for any subsequent action. Date of birth, citizenship, transliteration variants, the original source of the image, and the context of the recording are all verified. An erroneous association of two individuals can generate further "confirmations" if new databases replicate the initial error. Therefore, it is necessary to restore the independence of sources, rather than simply counting the number of matching records.

The report does not claim that INTERPOL or its Commission for the Control of Files (CCF) use a specific AI model. Issues of access, correction, and deletion of data in INTERPOL's information

system are addressed within the CCF's own rules and purview. Technical analysis may support an appeal, but does not replace its subject matter and procedural requirements. National actions and remedies exist separately. [25]

For the Observatoire, the practical task is to prepare a verifiable dossier: what data pertains to the individual, where the error occurred, how it spread, and why the existing conclusion cannot be used without further verification. The recommendation must be specific. A statement such as "the algorithm is wrong" is weaker than a proven identifier discrepancy, confirmed by independent documentation and linked to a specific record.

21 Migration and asylum

In migration procedures, AI can be used for document sorting, translation, finding inconsistencies, or assessing information. These functions differ significantly. A tool that helps an employee find pages is not the same as a system that determines the credibility of an applicant's story. Legal classification depends on the actual role of the result and the designated category of Appendix III, not on the general term "digitalization of migration."

A particular problem arises when meaning changes during translation and transcription. An error in name, negation, time of an event, or relationship between individuals can appear as a contradiction in testimony. If a second algorithm then analyzes the now erroneous translation, a chain of technically consistent but factually incorrect conclusions emerges. For significant circumstances, preservation of the original, the possibility of comparison, and verification by a person competent in the language and context are essential.

Insufficient documentation is not always a sign of dishonesty. Document availability depends on the circumstances of travel, conflict, the functioning of government agencies, and personal safety. A model trained on favorable administrative data may systematically downgrade the scores of individuals in the most challenging situations. Therefore, the analysis includes missing data as a separate phenomenon rather than automatically replacing it with an unfavorable value.

The remedy must take into account urgency. Correcting the record after a few months may be insufficient if the appeal period expires before then, or if a decision with irreversible consequences is enforced. The file must clearly outline the procedural deadlines, the possibility of seeking interim measures, and the availability of a competent representative. A complaint against the algorithm is not expected to automatically suspend the migration decision.

The proposed verification uses training and legally prepared examples of various languages, document types, and recording conditions. It must not reproduce sensitive information from real applicants in a publicly available collection. The published result describes the method and limitations, and individual materials are transmitted only in an authorized manner. Accessibility of explanation and secure communication with the individual are essential elements of the procedure's quality.

22 Banking services and sanctions control

In the financial sector, it's important to distinguish between creditworthiness, anti-money laundering, sanctions screening, and commercial decisions about working with a client. Their objectives, legal grounds, and permissible disclosure restrictions differ. A single, general

assessment of a "high-risk client" should not obscure the specific outcome that led to a refusal, transaction delay, or request for additional documentation.

When checking for matches, common names, transliterations, incomplete dates of birth, and outdated information are of particular importance. The algorithmic signal must be compared with the original source and additional identifiers. Citizenship should not be used without analysis as evidence of a specific individual's sanction status. At the same time, the assessor cannot declare any differentiation by country illegal: specific mandatory restrictions and acceptable risk assessments must be taken into account.

Rights to information and review may coexist with legitimate restrictions on disclosure, including the confidentiality of certain financial monitoring measures. A proper procedure seeks an acceptable scope of explanation and a competent verification route, without promising the client to disclose a protected communication. Restricting one category of information should not automatically preclude the ability to correct an incorrect name, document, or erroneous association of persons.

For independent assessment, indicators such as the time to resolve false matches, the number of repeat requests for the same document, the proportion of unjustified repeated restrictions, and the execution of review decisions are useful. The overall number of detected risks may encourage excessive triggering. Quality is assessed based on the balance between confirmed detections and the harm caused by erroneous actions, taking into account mandatory industry requirements.

In this scenario, the client submits a document excluding a match with a person on the list. Changing a record in just one service system is insufficient if the previous flag remains in the repository from which other products are updated. The correction must be distributed to dependent processes within acceptable limits. The file records the source of the error, the correction, and confirmation of the cessation of consequences, not just a formal "request closed" note.

23 Employment and availability of algorithmic selection

During hiring, the system may evaluate a candidate's resume, task performance, video recording, or behavior. First and foremost, it's necessary to establish a connection between the measured characteristics and the actual job function. High predictive accuracy relative to the recruiter's previous decisions does not confirm professional suitability or eliminate the possibility of past biases. The screening process must distinguish between a convenient signal and a valid criterion.

The accessibility of the assessment method is important in its own right. Speech, motor, visual, or hearing difficulties may impact interaction with the instrument without determining the ability to perform the task. An accessible alternative method and procedure for providing the necessary accommodations within the framework of applicable law are required. A candidate's refusal of a single technical format cannot be considered evidence of insufficient motivation. [21–22]

Assessing emotions and other sensitive characteristics requires separate verification of the AI Act's prohibitions and exceptions. A product's promise to determine honesty, engagement, or stress from a facial image does not constitute scientific evidence for the admissibility of such an assessment. An expert must verify the construct's validity: whether the claimed property is actually measured and whether it is relevant to the legitimate task. The lack of such a connection cannot be compensated for by a beautiful interface.

To test the recruiting process, it is proposed to analyze stages separately. A resume may be rejected before the test, an invitation may not be delivered, and an interview assessment may only influence a portion of candidates. The final acceptance rate conceals where the differences originate. The funnel should contain denominators by stage and documented reasons for attrition, including voluntary withdrawal.

Human review requires access to the original candidate and the ability to consider information outside the algorithmic profile. A request to "take the same test again" may be helpful in the event of a technical glitch, but does not correct the discriminatory design of the test. The final report separately describes recommendations for the model, interface, instructions, and the contract with the vendor.

24 Social Services and Digital Exclusion

Automated pre-assessment of eligibility for a service can speed up the process, but it can also exclude people whose circumstances are poorly represented in the standard form. Lack of a digital address, unstable housing, irregular income, or incomplete records should not automatically be converted into evidence of ineligibility. The legal provision of the service and the technical data entry scheme must be compared before implementation.

Access to an alternative channel is assessed based on actual availability. A phone number with no staff available or only available during hours inaccessible to most applicants does not provide a viable choice. The wait time, cost of the call, availability of translation, and the availability of representation should be checked. The lack of a digital interface should not, in itself, lead to a negative assessment of credibility.

In the educational sphere, special attention is required for minors and the sustainability of long-term outcomes. A low score at one stage can perpetuate a limited educational trajectory, even if the initial circumstance was temporary. The assessment should consider the possibility of change, explanation to the student and representative, adjustment of information, and periodic review. A single-digit profile does not capture all of a child's interests.

When distributing a limited resource, fairness isn't limited to equal shares among recipients. Legally established criteria and differences in needs must be taken into account. However, the model cannot arbitrarily create new grounds for refusal simply because they improve the serviceability forecast. If a technical rule effectively changes public policy, this must be identified and brought to the appropriate decision-making level.

An independent review is examining the fate of erroneously rejected applications. Correcting the algorithm for future cases does not address the issue of those who have already lost service. A search for affected decisions, to the extent permitted, notification, review, and verification of the possibility of reinstating the missed provision are required. Such a mechanism is planned in advance, while the data and authority are still available.

25 Appeal and selection of a competent addressee

Effective protection begins with identifying the contested action: data processing, denial of service, discriminatory practice, non-compliance with the AI Act, or a procedural decision of an authority. These subjects may overlap, but the addressees and powers differ. The internal service, the data

protection authority, the equality authority, the market surveillance agency, and the court are not interchangeable recipients of a single, universal letter.

Article 85 of the AI Act provides for the right of a natural or legal person to lodge a complaint with the relevant market surveillance authority if they believe that the provisions of the Regulation have been infringed. Such a complaint does not in itself constitute a general claim for compensation and does not automatically entail the annulment of any affected decision. The subject of the claim and the available powers must be formulated separately. [26]

The GDPR provides specific rights to complaint, legal protection, and compensation under established conditions. In discrimination cases, the relevant substantive and evidentiary rules apply. The timeframe and procedure for challenging an administrative decision are critical. The representative must assess which procedures can be used in parallel and how to avoid losing the urgent route while waiting for a response to the information request.

It is proposed to create a "defense card": event, date of knowledge, consequence, norm, addressee, purpose, time limit, and evidence. A request for an explanation, a request to preserve a log, a request to correct data, and a request for a suspension may have different legal status. They should not be combined in such a way that the addressee does not understand the required action. Furthermore, duplication of sensitive materials to multiple agencies must be justified by necessity.

The organization handling complaints must ensure that the review process is reasonably independent of the original decision. Authority is required to request materials from the supplier, evaluate new evidence, and remedy the consequences. An automatic response repeating the original result does not constitute a meaningful dispute resolution. For quality control purposes, the reasons, deadlines, and actual implementation of the complaint decision are recorded.

Subject of the appeal	Possible route	Main result
Inaccuracy of personal data	Controller and data protection authority	Correction and verification of consequences
Individual administrative decision	The prescribed administrative or judicial procedure	Revision, cancellation or urgent protection
Violation of the AI Act	Competent market supervision	Compliance review and enforcement measures
Discrimination	Equality Authority or Court as applicable	Establishment of a violation and the remedy provided
INTERPOL data	CCF procedure within its competence	Consideration of the appeal regarding data

26 Preservation of evidence and reproducibility

Disputes over algorithmic decisions often arise after a system update. Therefore, the control program must determine in advance which records will allow the restoration of material circumstances. The version identifier, configuration, input data (to the extent permitted), output,

time, employee action, and subsequent changes must be preserved. The retention period is determined by legal obligation and necessity, not by the principle of "everything forever."

Reproduction can be exact or limited. For a deterministic system, it is sometimes possible to repeat the calculation. For an external generative model, previous weights, hidden settings, and the service version may be unavailable. In this case, the actual response, query parameters, and information about the available conditions are stored. A distinction should be made between reproducing the original result and testing a similar system again; their evidential value is different.

The chain of custody records receipt, copying, access, transformation, and transmission. A checksum confirms the immutability of specific bytes after computation, but not the veracity of the statement they contain. A signature verifies a specific relationship with the signatory when verification conditions are met, not the truth of all facts. These constraints are especially important when technical features are presented to a non-expert as final proof.

Materials must be stored with restricted access. Originals are separated from working copies; changes and masking are recorded. Personal information of third parties is removed from the public version unless there is an independent basis for disclosure. Full client files cannot be uploaded to a third-party tool simply for ease of analysis without checking the legal framework, terms of use, and security.

If evidence is lost, the expert describes which issues remain resolvable. A notice, independent documents, and a history of actions may allow one to establish data errors even without the source code. However, a lack of access should not be translated into arbitrary assumptions about a hidden mechanism. The conclusion separates the identified problem, the probable cause, and the untested hypothesis.

27 Urgent protection and correction of consequences

The complaint review must take into account that harm may continue during the investigation. Account restrictions, risk of loss of housing, termination of a necessary service, or enforcement of a migration decision require a separate assessment of urgency. The organization must have a mechanism for interim measures within its authority. If judicial protection is necessary, the applicant is advised of the intended route, and an internal review is not presented as a substitute for judicial protection.

It is proposed to separate the establishment of the final cause from the prevention of additional harm. Temporarily referring a case for an expanded human review is sometimes possible until the technical assessment is completed. This does not automatically mean admitting a violation or permitting an operation despite a mandatory prohibition. The decision on a temporary measure must take into account the legal obligation, the nature of the risk, and available alternatives.

Correction involves not only recording in the database but also the dependent decisions. It is necessary to determine to whom the erroneous information was transmitted, which processes are reusing it, and how to prevent the old flag from being restored. Where the law requires notification of recipients about the correction, compliance is verified separately. In other cases, the necessity and admissibility of further communication is assessed.

After a system error is confirmed, an analysis of the affected community is conducted. It's not enough to limit the review to those who were able to file a complaint if the general mechanism of the violation is known and there is a legal possibility of finding similar cases. A mass review must have specific criteria and avoid collecting redundant data. Both the completeness of the correction and the preservation of the rights of others are important.

Recovery quality indicators include time to temporary protection, time to final correction, repeated restrictions, enforcement of the decision, and human comprehensibility of the message. The number of closed tickets is not a standalone indicator of rights protection. Closing a case must confirm a specific result or a reasoned absence of grounds for the requested action, with an accessible appeal process.

28 System procurement and supplier relations

Affected party rights are difficult to ensure if the necessary capabilities are not included in the procurement. The contract must specify access to documentation, notification of changes, retention of relevant records, assistance with complaints, and the possibility of independent verification. Technical integration that prevents the identification of the version of a response creates a future evidentiary gap that cannot be remedied by a general promise of transparency.

The supplier and the operator have different knowledge. The former knows the architecture and development results, the latter the conditions of use and the actual consequences. Their cooperation is essential, but responsibility should not be lost in correspondence. For each question, a person, deadline, response format, and escalation procedure are defined. If the supplier refuses to disclose commercially sensitive materials to the operator, secure access for an independent reviewer may be discussed, provided there are legal grounds.

An upgrade should be treated as a risk management event. Changes to the purpose, data, thresholds, interface behavior, and dependent services are verified. The organization needs a way to defer implementation, perform verification, and restore the previous valid configuration where technically and legally feasible. Automatic upgrades without notification can invalidate previous assessments.

Public procurement should consider accessibility, representation of the interests of affected groups, and the ability to operate without a single supplier. Reliance on a proprietary service makes it difficult to correct and exit the project. However, open source code alone does not guarantee legitimacy or independence: the data, process, and organization's ability to utilize available information must be assessed.

Acceptance criteria should include rights and remedies, along with accuracy and cost. Examples include the ability to generate a customized explanation, register an objection, and correct a dependent record. These functions are tested using scenarios, not accepted based on presentation. If a requirement is not met, the contract should allow for a correction or restriction of the relevant use.

29 Post-implementation monitoring

Initial assessments don't cover all future usage conditions. Populations, documents, sources, employee behavior, and the regulatory environment change. Therefore, monitoring should track not only technical drift but also changing impacts. A stable average accuracy may be accompanied by increasing errors for a small group or an increase in the time it takes to review.

The indicators are divided into technical, procedural, and human rights-based. The former describe errors and the reliability of processing. The latter focus on the actual work of employees, complaints, and the implementation of corrections. The third focus on the severity and duration of the consequences for people. Comparisons allow us to identify situations where a model has been formally improved, but the organization has begun to use its output more harshly, thereby increasing the harm.

Each indicator must have a denominator and source. The percentage of cancellations due to complaints does not reflect all decisions, as not everyone complains. A decrease in the number of requests may indicate an improvement in the process or a decrease in channel availability. An increase in the number of registered incidents sometimes indicates better detection. Interpretation requires context, not automatic assignment of a green color.

It is proposed to establish mandatory review events: change of purpose, significant update, new data source, confirmed serious error, unexplained group difference, supplier denial of required information, and change of entitlement. Regular review schedules are determined based on risk. A universal annual review may be insufficient for a rapidly changing process and excessive for a narrow, stable function.

Monitoring concludes with action. If a threshold is violated, the person, response time, time limits, and conditions for reinstatement must be determined. The threshold without consequences is merely a decorative indicator. The external report must explain how this procedure actually worked during the reviewed period and distinguish between the measures taken and those planned.

30 Public reporting and the right to object

Publicity supports trust by allowing understanding of the subject of the assessment and the limitations of the conclusion. Full publication of source materials is not always legal or necessary. It is recommended to publish a public summary describing the purpose, methodology, significant risks, and measures, while sensitive evidence is retained in a secure appendix. The degree of disclosure should ensure verifiability to the extent consistent with third-party rights and applicable restrictions.

The summary must distinguish between the absence of detected violations, insufficient data, and confirmed compliance with specific requirements. These statuses have different meanings. The wording "no violations detected" is misleading if the auditor did not receive the logs and was unable to examine the underlying mechanism. An honest conclusion about insufficient evidence may be grounds for limiting the system's use until further verification.

The audited organization must have the opportunity to respond to factual allegations. The timing and scope of such a response are determined by the procedure and urgency. Received objections are considered on their merits and not published as a formal ritual. If a comment changes a conclusion, the correction must be visible and linked to the previous version. Hidden file replacement complicates historical reconstruction and may encourage the dissemination of outdated conclusions.

Data on clients, complainants, and employees should not be used to enhance the credibility of a publication without independent justification. Case studies are designated as such. Aggregate figures are not presented as industry-wide statistics if they are based on a limited, voluntary sample.

The Observatoire should publish the method for selecting reports, the verification rules, and the proportion of cases for which a conclusion remains open.

The use of the report in advertising is subject to separate controls. The client must not quote a favorable line, removing a significant limitation on its scope. It is advisable to include an agreed-upon summary and a mechanism for correcting misattribution in the terms of use of the report. This protects not only the appraiser's reputation but also those who might mistake the advertising statement for a guarantee of their rights.

31 Proposed model of independent control of the Observatoire ARGA

For Observatoire ARGA, a limited evidentiary model, linked to its existing expertise in cross-border procedures and rights protection, is most appropriate. Its core is the analysis of a specific decision, the origin of the data, the applicable legal regime, and the available remedy. Creating a broad "ethics" rating of all systems would require significantly different resources and should not be declared an existing function.

The first module is research. It includes a register of regulatory changes, methodological materials, and a general risk analysis. The second is for reviewing specific situations, given the mandate, legal access, and necessary expertise. The third is for preparing proposals for consultations and institutional discussions. These modules should have different statuses: an analytical report is not a forensic examination, and participation in a consultation is an acknowledgment of supervisory authority.

Incoming materials are reviewed for their subject matter, representative authority, relevance, and urgency. If information is insufficient, specific questions are formulated. If competence is lacking, the case is referred to the appropriate professional route without promise of a result. Conflicts of interest are identified before access to the full file is granted to prevent a situation where a refusal occurs after receiving sensitive information from a competing party.

To ensure a substantive conclusion, two independent review functions are proposed: legal and technical. These do not necessarily require two permanent units, but must be performed by competent individuals. Disagreements are documented. If technical review does not support the legal conclusion, the formulation is revised; if there is no legal basis, the technical feasibility of processing does not justify the collection.

International collaboration should be built around verifiable materials: a standardized description of the problem, method, constraints, and proposal. External submissions and publications are subject to organizational approval. This report provides tools for such a process but does not confirm the actual submission of documents, the existence of agreements with government agencies, or the conduct of field audits.

32 Practical implementation program

The proposed program begins with a limited pilot using legally usable materials. The goal of the first stage is to test the functionality of the system, the rights map, the evidentiary register, and the peer-review procedure. It is not advisable to begin by accumulating large amounts of sensitive data. A few controlled training situations and then a small number of real-world assignments, subject to appropriate authority and competence, are sufficient.

During the first period, roles, access rules, and assignment rejection criteria are approved. Forms are checked for consistency: identical identifiers must link the decision, evidence, risk, and corrective action. Next, a trial run of several different scenarios is conducted. Any identified inconveniences are recorded as methodological shortcomings rather than being compensated for through informal agreements between participants.

The second period is dedicated to testing independence and reproducibility. Another specialist receives the authorized set of materials and attempts to reconstruct the conclusion. If this is impossible, it is determined whether any facts, definitions, or evaluation rules have been lost. Simultaneously, the response to the reviewed party's objection and the modification of the conclusion in the face of new evidence are tested. The ability to correct the conclusion is a necessary quality characteristic, not a sign of project weakness.

The third period concludes with a decision on the scope of further implementation. Possible outcomes include continuing within a narrow scope, engaging specialists for specific functions, modifying the procedure, or eliminating certain types of tasks. The decision specifies available resources and unresolved limitations. The announcement of a new direction should not outpace the actual commitment to ensuring confidentiality and verifiability.

Stage of the proposed program	Result	Transition condition
Preparation	Mandate, roles, forms and data mode	Competence and legality of access verified
Training pilot	Several reproducible analyses	Significant logical errors have been corrected.
Limited practice	Testing on approved real materials	Independent review and safety monitoring completed
Grade	Decision on the permissible scope of work	Constraints are disclosed and resourced
Periodic review	Updated methods and law map	There is an actual verification of corrective measures

33 Proposals for International Consultations

The first proposal concerns the minimum evidentiary content of impact assessments. Public forms should be linked to identifiable uses, groups, sources, and specific mitigation measures. Filling out generic fields like "low, medium, high risk" without justification is insufficient. At the same time, small organizations need clear tools that allow them to identify truly significant issues without creating unnecessary documentation.

The second proposal is to ensure the portability of the explanation of an individual decision between participants in the chain. The supplier and operator should have an agreed-upon set of information about the version, input data, and the role of the output. Universal public disclosure of closed models is not required, but a legal mechanism for independent and regulatory verification

should exist. Lack of access should not be removed from the final report under the guise of commercial confidentiality.

The third suggestion concerns the assessment of actual human oversight. Formal job descriptions should be supplemented with verification of competence, time, access, and authority. Organizations benefit from examples of tests that demonstrate whether an employee can recognize an error and change the outcome. However, mandatory metrics for reversing decisions can create a negative bias; the quality of the review should be assessed, not whether the quota of disagreements with the algorithm is met.

The fourth proposal is to develop safe methods for verifying group differences. Practical clarifications are needed regarding access to special categories, verifier independence, and reporting results for small groups. Equality testing should not become an excuse for excessive data collection. At the same time, confidentiality requirements should not be used as a blanket prohibition on identifying discriminatory effects.

The fifth proposal concerns the correction of decisions already made. Monitoring recommendations should include the link between technical corrections and the restoration of rights of affected parties. The corrected model does not automatically correct past service losses. It is proposed to develop verifiable procedures for identifying similar cases, notifying, and implementing reconsiderations, taking into account applicable law.

These proposals are formulated as a position for further internal coordination and adaptation to the subject of a specific consultation. They are not designated as requirements already adopted by the regulator or submitted requests. For each external submission, the wording must be compared with the published questionnaire and the organization's current mandate.

34 Conclusion

The protection of fundamental rights when using AI depends on whether humans can understand the substantive reasons for a decision, provide relevant information, and obtain an effective remedy. Technical tests are necessary, but without legal analysis and an organizational mechanism, they do not answer these questions. A legal declaration without access to the facts is equally insufficient.

The proposed methodology combines legal qualification, contextual research, data verification, error analysis, and remedies. Each conclusion must have a specific scope and limit. Where facts are not established, uncertainty must be maintained; where a serious deficiency is confirmed, it must not be dissolved into an average quality score or a favorable summary.

The current AI Act calendar requires careful attention to revisions and transitional provisions. Postponing a portion of the high-risk regime does not cancel existing rights and does not allow for the entire work to be postponed until the final date. Early preparation, based on independent responsibilities and voluntary impact assessments, with a clear indication of its status, is practically justified.

Observatoire ARGA prioritizes reproducible dossiers and independent peer review within its mandate. This model enables the preparation of substantive research, support for advocacy in specific proceedings, and the formulation of proposals for international discussion. Its credibility is

determined by the quality of the evidence, the clarity of its limitations, and the ability to correct its own errors.

Appendix 1 Passport of the application being evaluated

The data sheet is completed for a specific use, not for an entire product line. It must allow another specialist to identify the object without verbal explanation. Unknown information is marked as unknown, with the person responsible for clarification indicated; a blank field is not considered a negative answer.

Field	What needs to be recorded
Identifier	Unique application number and passport version
Organization and process	Who makes the decision and what function is automated?
Purpose	What result is created and how is it used?
Affected persons	Users, applicants and foreseeably affected third parties
Technological object	Vendor, model, version, interface, and dependent services
Input information	Categories, sources, period and order of correction
Exit and action	Score, recommendation or text; employee follow-up action
Territory and time	Countries of use and dates of tested solutions
Legal qualification	Applicable rules, exceptions and rationale
Restrictions	Unavailable Materials and Limits of Acceptable Output

The passport is signed by the process owner; technical and legal specialists confirm the information within their competence. A change in the purpose or essential condition creates a new version, preserving the previous one. The model name does not replace the description of the configuration and the person's role.

Appendix 2 Legal Calendar Card

The form prevents confusion between current law, transitional provisions, and reform proposals. Each requirement requires its own line. The verification date is not the same as the effective date of the provision.

Requirement	Legal basis	Date and condition	Practical action
FRIA General Regime Annex III	Articles 27 and 113 of the AI Act	December 2, 2027, subject to special conditions	Prepare the process and check the operator's coverage
Product requirements of the relevant sections of Chapter III	Article 113(c)(ii)	August 2, 2028	Compare with the existing industry regime
DPIA	Article 35 of the GDPR	In the presence of a foreseen high risk of processing	Carry out prior to appropriate processing

Requirement	Legal basis	Date and condition	Practical action
Special biometric conditions	Article 5 of the AI Act and applicable national law	Determined by a separate rule	Check the purpose, authority and necessary guarantees
Explanation of the individual solution	GDPR, Article 86 AI Act or national law	Each basis has its own subject and conditions	Form a separate reasoned conclusion

When updating a card, the link to the official text, the revision, and a brief description of the change are retained. The draft and political message are entered into a separate monitoring section. They must not change the mandatory date without confirmation of the adopted act.

Appendix 3 Map of Rights and Harm Scenarios

The map links an abstract right to a specific mechanism. A single system may require multiple rows for a single right if the reasons and protections differ. The overall score does not replace the individual description.

Right or interest	Scenario	Proof	Possible measure
Equality of access	Error in recognizing documents of one group leads to refusal	A verified sample and analysis of the reasons	Correcting the module and an alternative method
Privacy	Redundant source aggregation	Data schema and actual logs	Restricting sources and access
Data protection	The obsolete entry determines the solution	Source, date and confirmation of change	Fixing and monitoring dependent systems
Effective appeal	The person receives only a general refusal.	Notifications and test procedures	Individual explanation and competent review
Service availability	The only interface is unavailable.	Testing with allowed scenarios	Equivalent alternative channel

In the working copy, severity, reversibility, probability with justification, and responsible party are added. If the connection to the consequence is still hypothetical, this is stated explicitly. The map should not transform any technical inconvenience into an established violation of law.

Appendix 4 Participation Plan for Affected Groups

Before the consultation, the organizer explains its purpose, potential impact on the project, and how the information will be used. Participation should not require sharing a complete personal profile. Group representatives can report systemic barriers without disclosing their individual history.

Stage	Question	The result being tested
Selection of participants	Who can be excluded from the usual channels?	A well-founded map of groups and communication methods
Informing	Are the task and consequences clear?	Brief accessible description

Stage	Question	The result being tested
Discussion	What risks does the developer not see?	Recording comments without redundant data
Solution	What has changed and why	Response to each significant comment
Feedback	Did the participants know the result?	Accessible message without disclosure to third parties

The organizer specifically sets forth limitations on representativeness. The opinions of a few active participants are not declared to represent the position of all affected individuals. Reimbursement of reasonable expenses should not be contingent on support for the project or refraining from criticism.

Appendix 5 Sample and standard passport

The form is completed before calculating the key indicators. Changes to criteria after reviewing the results must be visible and justified. Otherwise, there is a risk of selecting only those cases that support the desired conclusion.

Field	Content
Population	All decisions to which the conclusion is supposed to apply
Period	Beginning and end, reasons for choice
Unit	A person, statement, transaction, or other object
Selection	Random, continuous, stratified or targeted
Exceptions	Reason and number of excluded units
Duplicates	Rule for detecting and processing related observations
Reference	The source of the correct result and the procedure for disagreement
Independence	No leakage between training and testing
Uncertainty	Small groups, gaps and limits of generalization
Reproducibility	Versions of the set, code and calculation rules

A separate sample of severe cases is retained as a diagnostic sample. It is not combined with the representative sample when calculating the overall incidence. All published denominators must be retrievable from the case log and exception log.

Appendix 6 Calculation of group differences on training data

A fictitious selection process is considered, where the positive outcome is admission. In Group A, 300 out of 500 people are admitted, and in Group B, 180 out of 400 are admitted. The percentages are 60% and 45%. The difference is 15 percentage points; the ratio of the percentage of B to the percentage of A is 0.75. These numbers describe the selection outcome and do not in themselves establish legal discrimination.

Indicator	Group A	Group B
Total observations	500	400
Positive decisions	300	180
Negative decisions	200	220
The share of positive decisions	60%	45%

Assuming independent observations and the applicability of the simple binomial model, the standard error of the difference is approximately 0.0331. The normal approximation yields an approximate 95% interval of the difference from 8.5 to 21.5 percentage points. This calculation requires modification for household clustering, repeated applications, or complex designs. The statistical interval does not correct for selection bias or incorrect reference.

The next step is to examine the stage at which the distinction arose, the quality of the source data, the relevance of the criteria to the objective, and the availability of an alternative process. Applicable anti-discrimination law is examined separately. The 0.8 ratio threshold is not used here as a universal legal threshold: known practical rules from individual regimes cannot be generalized to all jurisdictions and decision types.

To demonstrate uncertainty, let's compare 30 errors out of 1,000 cases and 3 errors out of 100. In both cases, the proportion is 3%, but the approximate Wilson 95% intervals differ: 2.11–4.25% and 1.03–8.45%. Therefore, identical point estimates do not imply equal evidential strength. All data in this appendix are provisional.

Appendix 7 Human Control Verification Protocol

The test should demonstrate the employee's ability to identify a significant error, not their knowledge of the wording of the instructions. Legally prepared scenarios are used; real people should not be given an unfavorable decision for the sake of the experiment.

Scenario	What to watch for	Sign of sufficient control
Matching name with a different date of birth	Checking additional identifiers	The employee does not accept the signal as an identified person.
High score with a controversial document	Working with the original source	The contradiction is investigated and motivated
Incomplete translation of an essential phrase	Request for original and competent verification	The decision is not based on an unverified transcript
Borderline score	Understanding the threshold and alternatives	There is a reasonable individual assessment
New evidence after complaint	Possibility to change the output	The revision takes into account the information and corrects the consequences

The protocol specifies the time, available materials, the employee's explanation, and the result. Failures are analyzed for reasons such as lack of training, lack of authority, unavailable data, or workload. A cancellation quota is not established as an independent criterion for fairness.

Appendix 8 Impact Assessment Worksheet

This form is a proposal from the ARGA Observatory, not an official AI Office questionnaire. Its use requires verification of the applicability of Article 27 and coordination with other obligations. Each reference to evidence must refer to a specific registry object.

Chapter	Contents of the conclusion
Purpose and process	How and why the system is used
Addressee of the duty	Why an organization is or is not included in the relevant scope
Time and frequency	Period, intensity, revision events
People and groups	Who is affected and why certain groups require attention
Risks	Law, mechanism, severity, reversibility and basis for assessment
Human supervision	Competence, data, authority and verification of actual behavior
Measures	Prevention, detection, emergency response and recovery
Appeal	Channels, competence, availability and implementation of the solution
Contacting the DPIA	General facts and independent conclusions
Result	Permitted Use, Restrictions, Responsibility and Review

A log of disagreements and a list of missing materials are attached to the form. A significant gap should not be eliminated when shortening the management document. If the conclusion is dependent on a future test, this condition is included in the launch decision itself.

Appendix 9 Register of Evidence

A single object can support multiple derivations, but its origin is described only once. Derivative files are linked to the original. Data on real individuals is stored in a secure area, while the public version uses anonymized identifiers.

Field	Filling rule
Number	Unique object identifier
Description	What exactly was received and what event does it relate to?
Source	From whom and on what basis
Time	Receipt, creation and attributable period differ
Format and integrity	Original, copy, checksum if necessary
Transformations	Translation, extraction, masking and responsibility

Field	Filling rule
Access	Who has the right to see the full object?
Related findings	Numbers of statements being verified
Restrictions	Incompleteness, inability to confirm authenticity or context
Storage	The basis for the term and the procedure for revising the necessity

The registry should not contain more sensitive information than is necessary for evidence management. The mere fact of registering an object does not confirm the veracity of its contents. Verification of authenticity, relevance, and sufficiency is conducted separately.

Appendix 10 Structure of the Request for Explanation

This example is intended to be adapted to the applicable law and the addressee. It does not imply unconditional access to any information or automatic suspension of the decision.

"I request information regarding the role of automated processing related to the decision dated [date, number]: what information about me was used, from what sources it was obtained, how the system's results influenced the decision, and what key circumstances determined the unfavorable outcome. I also request information on the procedure for correcting inaccuracies and reviewing objections by a competent employee."

"I consider the following circumstances to be material: [specific facts and documents]. I request that they be verified before reusing the disputed information. If certain information cannot be provided, I request that the legal basis for the restriction and the accessible procedure for challenging it be indicated. I specifically request that the need for permissible interim measures be considered, taking into account the following consequences: [specific description]."

Only the necessary documents are attached to the request. The legal basis, deadline, and identification requirements are determined separately. Do not include a promise that the request will inevitably lead to disclosure of the source code, reversal of the decision, or compensation.

Appendix 11 Complaint Review Card

The map links the organization's actions to the applicant's actual situation. The "closed" status is granted only after a reasoned decision and verification of compliance with the prescribed measures.

Stage	Examination	Entry in the case
Receipt	Subject, authority, urgency	Date, channel, number and person responsible
Preservation	What records may be lost	Legal basis and limit of preservation
Temporary measure	Is there any ongoing serious harm?	Decision and limits of authority
Study	Data, version, motivation and new information	Evidence and controversy

Stage	Examination	Entry in the case
Solution	Confirm, change or cancel within the limits of competence	Clear reasons and further route
Execution	Have dependent records and actions changed?	Confirmation of the actual result
Systems analysis	Are there any similar cases raised?	Decision on additional verification

The time limits on the map are set based on applicable law and the urgency of the specific case. The internal benchmark should not be defined as a single legal deadline for all states. Access to the map is limited to essential participants.

Appendix 12 Conflict of Interest Map

The review is performed before accepting the assignment and is updated as circumstances change. Conflict can be manageable or non-participatory; the decision must be motivated, not just marked "not identified."

Circumstance	Risk	Possible action
Participation in the development of the system	Evaluation of one's own decision	A separate independent team or refusal
Payment depends on a favorable conclusion	Pressure for results	Change payment terms
Financial dependence on one client	Restriction of freedom of conclusion	Additional review and disclosure
Representation of the opposing party	Confidentiality and loyalty	Legal assessment and, if necessary, refusal
The customer controls the publication	Removal of significant restrictions	To establish precise rules for using the report

A control measure is considered fulfilled upon its actual application. The intention to engage a reviewer is not the same as receiving a review. The final report discloses the material conflict to the extent consistent with legitimate confidentiality.

Appendix 13 Matrix of contractual requirements for the supplier

Requirements are tailored to the parties' roles, the product, and the legal regime. The agreement may not grant access to personal data contrary to applicable restrictions.

Condition	Verifiable execution
Version identification	Each significant result is associated with a version and configuration
Documentation	The operator receives a description of the purpose, limitations and necessary conditions

Condition	Verifiable execution
Updates	Changes are communicated in a timeframe that allows for risk assessment.
Complaints	The channel, the person responsible and the required assistance have been identified.
Proof	Relevant logs are retained in the legal volume
Independent verification	Allowed access and privacy protection are set
Errors	There is a procedure for notification, correction and verification of the result
Exit	Termination of use and acceptable data transfer are provided

Acceptance is conducted using scenarios. The supplier demonstrates that the organization is capable of explaining the decision, obtaining information on the complaint, and determining which version was in effect at the time. Simply referring to a general transparency policy is not sufficient.

Appendix 14: Study Case on False Financial Matching

Factual facts: a client receives a transaction restriction after their name automatically matches an entry in an external list. The date of birth in the entry is different, but the field was missing from the initial request. The client submits their passport and reports the discrepancy. A standard message is received in response, with no further verification details. This situation is fictitious and does not describe the practices of a specific bank.

The first step is to determine the subject of the action. It's important to distinguish between a sanctions ban, an internal suspension for verification, and an independent commercial decision. The client's documents may rule out an identification match, but they don't automatically allow for a determination of all other grounds. The representative formulates a request to verify a specific inaccuracy and determine an accessible route, without assuming the right to disclose protected financial monitoring information.

Hypothesis	Necessary verification	Possible conclusion
Only the common name matched	Source and additional identifiers	The signal is insufficient to establish identity.
Date of birth was not submitted to the system.	Data schema and query log	Lack of integration, not necessarily model
The employee did not see the document	Access history and interface	Procedural defect in consideration
The fix has not been distributed.	Dependent records and rechecking	The continuing consequence of the initial mistake

An independent analysis should not begin with a claim of proven discrimination. The data error and its role must first be established. If systematic inequality is identified for certain groups of names, a separate sample and legal qualification are required. At the same time, permissible interim measures are possible if the ongoing restriction causes significant harm.

The completion criterion is an actual verification of the identity and consequences, not the sending of another standard response. The report records the original entry, the correction, the responsible department, and the presence of other independent grounds that were established. An unknown ground is not replaced by an assumption favoring either the client or the organization.

Appendix 15: Study case on translation error in migration procedure

Factual facts: the system transcribes the interview and skips the denial. The secondary analytical summary reveals inconsistencies between the interview and the written statement. The employee uses the summary to prepare an unfavorable decision. The original recording is preserved; the complainant specifies a specific time period. The case is fictitious.

The review begins with capturing the original and comparing the disputed fragment with a competent translator. The error's path is then examined: the original transcript, the translation, the analytical summary, and the final justification. It cannot be assumed that correcting one file will automatically correct the already prepared solution. Each derivative version is linked to the original source.

Link	Question	Measure
Audio recording	Has the controversial fragment survived?	Legal preservation and control of integrity
Transcript	Is the negation or tense changed?	Competent verification with indication of the fragment
Summary	Did the change affect the output?	Re-analysis on corrected material
Solution	Was the contradiction used as a substantial basis?	Due process review
Urgency	Is there a near term or irreversible consequence?	Separate assessment of available urgent protection

The technical report identifies the specific distortion and the limits of its impact. It should not confirm the veracity of the entire story simply because one error has been confirmed. The legal representative will use the result in the appropriate procedure, verifying the deadlines and the authority of the recipient. A request to the software provider does not replace a timely appeal of the decision.

As a preventative measure, similar language constructs are checked against the approved set. If the error persists, the organization will review the automatic summary and implement a check for significant fragments. Public publication must not reveal the voice, identity, or circumstances of the actual applicant without proper justification.

Appendix 16: Educational file on unavailable personnel test

Facts: The candidate must complete a timed video assessment. Due to speech impediments, the system assigns a low score, even though the position does not require a speed assessment. There is no alternative format, and the reviewer suggests repeating the same test. This case study is intended for methodological analysis and is not specific to a specific employer.

The first question is what exactly the indicator measures and how it relates to the job function. A description of the criterion, validation tests, and the role of the outcome should be obtained. Then, the accessibility of the procedure and the possibility of accommodations provided by applicable law are checked. The presence of a uniform interface for everyone does not prove equal actual opportunity to demonstrate competence.

The element being tested	Relevant evidence
Labor function	Description of duties and required skills
Test design	What is intended to be measured and why
Availability	Possibility of passage with various restrictions
The role of the score	Threshold, selection stage and recruiter's powers
Revision	Considering alternative skill validation

An independent assessment may establish that a criterion is inappropriate for the task or that an effective adjustment procedure is lacking. Legal qualifications depend on the country, the employer's status, the applicable ban, and the facts. Simply writing "the algorithm is biased" without describing the mechanism is insufficient. Correction sometimes requires changing the evaluation format itself, rather than retraining the model on a larger number of similar videos.

The outcome criterion is the provision of a legitimate and meaningful opportunity for skills assessment, not guaranteed employment. If a similar obstacle has affected other candidates, a permissible review of the relevant decisions is considered. Disability information is not included in the open technical recruitment process without a legal basis and the necessary guarantees.

Appendix 17: Study case on digital denial of social services

Factual facts: the system rejects the application due to the lack of standard proof of address. The applicant does not have a permanent residence, but the law allows for alternative proofs. The employee only sees the result "incomplete," and the paper channel is essentially non-functional. This case is fictitious.

The problem may lie in the interface rules and process organization, even if the model accurately recognizes the uploaded documents. The assessment must compare the legal conditions for receiving the service with technically acceptable options. If the system does not allow the submission of legally recognized evidence, its high accuracy in processing the remaining applications does not eliminate the defect.

Stage	Examination
Norm	What evidence is admissible?
Interface	Is it possible to imagine every possible option?
Employee	Can an exceptional case be registered?
Alternative channel	Is it actually available?

Stage	Examination
Consequence	Has the right been lost or the deadline missed?
Correction	Has the possibility of consideration been restored and have similar cases been verified?

A temporary measure may consist of manually accepting and recording the date of the request within the organization's authority. A final measure is to change the process and inform employees. However, the independent expert does not make a decision on payment eligibility in lieu of the competent authority. Instead, he or she creates an obstacle to the legally mandated review.

The post-correction quality indicator includes the number of successfully registered alternative confirmations and the processing time, not just the reduction in technical errors. If people previously abandoned their application, their absence from the database should not be considered the absence of a problem.

Appendix 18 Register of Corrective Measures

The measure is linked to the identified cause and verified after implementation. Phrases such as "strengthen controls" and "train employees" are insufficient without concrete results.

Field	Content
Flaw	A proven fact and its associated consequence
Cause	Established or clearly stated probable cause
Measure	A specific change to data, system, or procedure
Owner	A person with the necessary powers
Term	Legal or internal with status indication
Examination	Method of confirming performance and effectiveness
Residual risk	What remains after the measure and why
Solution	Resume, restrict, continue suspension
Past cases	The necessity and limits of retrospective correction

Closing a measure requires proof of its effectiveness. If a new version reduces one type of error and increases another, the tradeoff is reconsidered. Significant residual risks are communicated to a competent person rather than hidden in a technical log.

Appendix 19 Structure of the independent opinion

The conclusion is intended for readers who were not involved in the review. It should allow them to establish the object, basis, and limits of the conclusion without recourse to the author's verbal explanations.

Chapter	Mandatory content of the proposed form
Mandate	Customer, questions, period and addressees
Object	Application, versions, process and affected persons
Grounds	Norms, criteria and their status
Materials	Obtained objects and significant gaps
Methods	Sampling, testing and rules of interpretation
Results	Facts separate from legal assessment
Restrictions	What conclusions cannot be drawn
Objections	Substantive comments and reasoned response
Measures	Priority, basis and method of verification
Result	Narrowly formulated conclusion and conditions for revision

An example of an acceptable final formula: "The sample analyzed confirmed that the absence of an additional identifier resulted in false matches. The data does not allow us to estimate their frequency across the entire system. Until the integration is corrected, it is recommended to exclude a final adverse finding based on one such signal, taking into account applicable mandatory requirements."

It is unacceptable to turn this result into a formula such as "the system is illegal in all cases" or "the audit guarantees non-discrimination." The terms of use of the public summary must preserve the scope and limitations of the output.

Appendix 20 Checklist before external distribution

The review concerns the content and presentation. It does not confirm the conduct of field research if the report is a normative and methodological analysis.

Examination	Completion Criteria
Regulatory edition	The date, official source and transitional provisions are indicated.
Competence	The organization is not assigned any state powers.
Facts	Verified data is separated from the training examples.
Statistics	Numerators, denominators and calculations are reproduced
Sources	The links correspond to the statements and status of the document.
Terms	FRIA, DPIA, audit and certification do not mix
Confidentiality	There is no unnecessary information about real persons
Editing	No missing chapters, duplicates or empty service fields
Word and PDF	The table of contents, tables, symbols and numbering are legible

Examination	Completion Criteria
Spreading	The established internal coordination has been completed

All proposals for organizing oversight require adaptation to the actual mandate and resources. The publication of a report does not in itself signify the adoption of a program, the commencement of audit activities, or the distribution of materials to external recipients.

Sources and regulatory materials

The regulatory snapshot and date of access to web sources is October 23, 2026. Numbers in square brackets link the analysis to the list below. Explanations from agencies, voluntary methodologies, and proprietary proposals are separated from mandatory standards. The statistical calculations and four case studies are for educational purposes only; a field audit of a specific organization is not supported by this study.

[1] Regulation (EU) 2024/1689 on artificial intelligence. Consolidated version of 27 July 2026. Regulatory framework for the study.Official source

[2] Regulation (EU) 2026/1744 of 8 July 2026. Digital Omnibus on AI. Officially published amending act.Official source

[3] European Commission. AI Act Service Desk, Article 113. Current application calendar; explanatory resource, text based on the consolidation of 27 July 2026.Official source

[4] Charter of Fundamental Rights of the European Union, 2012/C 326/02. In particular Articles 1, 7, 8, 21, 47, 51 and 52.Official source

[5] Regulation (EU) 2016/679, GDPR. In particular Articles 5, 6, 9, 12–16, 19, 22, 35–36 and 77–82.Official source

[6] Directive (EU) 2016/680. Processing of data by competent authorities for criminal law purposes, in particular Article 11. Implemented through relevant national law.Official source

[7] European Commission. AI Act Service Desk, Article 6. Qualification of high-risk systems and conditions for exemptions.Official source

[8] European Commission. AI Act Service Desk, Article 27; Navigating the AI Act, updated 7 August 2026. Subject matter of the FRIA and relationship with the DPIA.Official source

[9] EDPB. Automated decision-making and profiling. Endorsement of WP29 guidelines, WP251rev.01, 25 May 2018. Explanatory document, not legislative act.Official source

[10] Court of Justice of the EU. Judgment of 7 December 2023 in Case C-634/21, SCHUFA Holding (Scoring), ECLI:EU:C:2023:957.Official source

[11] Court of Justice of the EU. Judgment of 27 February 2025 in Case C-203/22, Dun & Bradstreet Austria, ECLI:EU:C:2025:117. Official text of the judgment.Official source

[12] European Commission. AI Act Service Desk, Article 86. Conditions for the right to an explanation of an individual decision.Official source

[13] Council of Europe. Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225. Official review and access to treaty materials.Official source

[14] Council of Europe. HUDERIA Methodology, adopted by CAI on 28 November 2024. A stand-alone non-binding methodology for risk and impact assessment.Official source

[15] Council of Europe. HUDERIA Model: COBRA Resources Approved, 25 February 2026. Updated context analysis resources.Official source

[16] UN, OHCHR. Guiding Principles on Business and Human Rights, 2011. Principles on the protection of, and respect for, rights and access to legal remedies.Official source

[17] NIST. AI Risk Management Framework, AI RMF 1.0, 2023. Voluntary risk management framework.Official source

[18] France. Code of Public and Administration Relations, Articles L311-3-1 and R311-3-1-1–R311-3-1-2. Access to information on an algorithmic administrative decision.Official source

[19] UK. Data (Use and Access) Act 2025, Chapter 18. In particular, provisions on automated decision-making.Official source

[20] ICO. The Data Use and Access Act 2025 – what does it mean for organisations. Official explanations of the reform, current web version checked on 23 October 2026.Official source

[21] EEOC. Employment Discrimination and AI for Workers, April 29, 2024. Effect of Anti-Discrimination Protections on Employer Use of AI.Official source

[22] US Department of Justice. Algorithms, Artificial Intelligence, and Disability Discrimination in Hiring. Official guidance on accessibility and discrimination in hiring.Official source

[23] FRA. Assessing High-risk Artificial Intelligence: Fundamental Rights Risks, December 4, 2025. An empirical study of the application of AI in selected significant areas.Official source

[24] European Commission. AI Act Service Desk, Article 4a. Conditions for the treatment of special categories for the detection and correction of biases.Official source

[25] INTERPOL. Commission for the Control of INTERPOL's Files. Official description of the Commission's functions and application routes.Official source

[26] European Commission. AI Act Service Desk, Article 85. Right of appeal to the competent market surveillance authority.Official source