

What can I help
you with?

Positive AI Label: Guide for a Responsible Artificial Intelligence

Contents

1.	Positive AI and Label presentation	5
1.1.	Positive AI Vision	5
1.2.	Positive AI Ambition	6
1.3.	History and methodology of the Positive AI Label	6
1.4.	Document reading guide	7
2.	Label governance	109
2.1.	Positive AI Association	109
2.1.1.	Presentation	109
2.1.2.	Composition	109
2.1.3.	Role and missions	109
2.2.	Positive AI Members	116
2.2.1.	Presentation	116
2.2.2.	Role and missions	116
2.3.	Labelling Committee	129
2.3.1.	Presentation	124
2.3.2.	Composition	124
2.3.3.	Role and missions	124
2.3.4.	Functioning	124
2.4.	Positive AI Referenced Auditors	136
2.4.1.	Presentation	136
2.4.2.	Role and missions	136
3.	Labelling Process	157
3.1.	Labelling cycle	157
3.1.1.	Membership	157
3.1.2.	Eligibility Verification	164
3.1.3.	Submitting of the application and contracting with a Positive AI Referenced Auditor	164
3.1.4.	Execution of the Labelling Audit	164
3.1.5.	Labelling Decision	176
3.1.6.	Surveillance Audits	184
3.1.7.	Renewal of the Label	184
3.1.8.	Suspension, withdrawal, termination of the Positive AI Label	184
3.2.	Definition of the Labelling Audit scope	198
3.2.1.	About the Labelling scope	198
3.2.2.	Criteria to define the Labelling scope	198
3.2.3.	Setting the Labelling scope	212
3.2.4.	Dealing with changes in scope	212
3.3.	Conformance evaluation criteria	228
3.3.1.	Defining conformance rating	232
3.3.2.	Documenting non-conformances	242

4.	Positive AI Label framework	2624
4.1.	Key Responsible AI Principles	2624
4.2.	Structure of the Positive AI Label framework	2726
4.2.1.	Global structure	2726
4.2.2.	Control objectives	2825
4.3.	Pillar 1 of the framework: Governance of the AI design and usage	2826
4.3.1.	Section 1: Governance	2826
4.3.2.	Section 2: Policies	2826
4.3.3.	Section 3: Responsible AI Operationalization Strategy	2927
4.3.4.	Section 4: Use case inventory	2927
4.3.5.	Section 5: AI lifecycle	3128
4.3.6.	Section 6: Production Control Framework	3128
4.3.7.	Section 7: Data Management and Privacy Policies	3128
4.3.8.	Section 8: Overall IT infrastructure (IT stack)	3229
4.3.9.	Section 9: Responsible AI at the level of the organization and incorporation in supporting functions (Human Resources, Internal Audit, etc.)	3229
4.4.	Pillar 2 of the framework: Validation of algorithmic models	3229
4.4.1.	Section 1: Model documentation	3430
4.4.2.	Section 2: Model design Assessment	3430
4.4.3.	Section 3: Ongoing Performance Monitoring	3534
4.4.4.	Section 4: Functional Process Controls and Implementation Tests	3534
4.4.5.	Section 5: IT & Architecture Process Controls and Implementation Tests	3534
4.4.6.	Section 6: Identification, assessment and management of bias and risks to fairness	3732
4.4.7.	Section 7: Explainability & Interpretability of the model and of its results	3732
4.4.8.	Section 8: Sensitivity analysis of the model to its features	3732
4.4.9.	Section 9: Hypothesis Testing when designing the algorithm	3833
4.4.10.	Section 10: Performance Testing	3833
4.4.11.	Section 11: Impact Assessment of known weaknesses and open improvement points	3934
4.4.12.	Section 12: Compliance review to the regulatory and legal requirements	3934
5.	Commitments of the labelled Entities	4136
5.1.	General commitments	4136
5.2.	Use of the trademark	4136
6.	Appeals and complaints handling	4338
6.1.	Positive AI complaints mechanism	4338
6.2.	Appeal procedure	4338
6.3.	Claims	4338
7.	Appendices	4540
7.1.	List of acronyms and glossary	4540
7.2.	List of Positive AI Referenced Auditors	4641
7.3.	Sources	4641



1. Positive AI and Label presentation

1.1. Positive AI Vision

A label for all organizations committed to a Responsible AI approach

The Positive AI Label aims to support all organizations that manage and/or develop Artificial Intelligence (AI) systems and that wish to voluntarily commit to a responsible AI approach, whatever their size, level of maturity or sector of activity. The label is not only aimed at the creators of the systems but also at the users. Indeed, even if an algorithm is designed in a responsible way, its management must also respect the principles of a trusted AI.

The Positive AI Label aims to be a provider of solutions (as an operational framework to assess the level of trust of companies in the implementation and the monitoring of their AI systems) for organizations that wish to make rapid progress on the issue of Responsible AI, as well as a trust indicator for the public who wishes to see responsible issues more respected in the development of AI.

Definition of Responsible AI

Positive AI defined the Responsible AI as an AI whose stakeholders (developers, business teams, employees, consumers, etc.) are aware of the impact of their AI systems on the society and the environment, and act to reduce this impact as much as possible through risk analysis and dedicated controls.

The European Commission defines a trusted AI in 3 points:

- It must be lawful and ensure compliance with applicable laws and regulations.
- It must be ethical and reflect respect for and adherence to ethical principles and values.
- It must be technically and socially robust, to ensure that, even with good intentions, artificial intelligence systems do not cause unintended harm. Trusted AI validates the trustworthiness of the AI system itself, but also the trustworthiness of all the processes and actors that are part of the system's life cycle.

Responsible AI must also meet the 7 ethical principles defined by the European Commission:

- Human action, human control, and fundamental rights.
- Technical robustness and security.
- Privacy and data governance.
- Transparency of data, systems, and models.
- Diversity, non-discrimination, and equity.
- Societal and environmental well-being.
- Accountability.

1.2. Positive AI Ambition

The idea of the Positive AI Label was born out of the observations shared by the four founding companies that:

- As the primary space for the creation and deployment of AI, companies have a role to play in thinking about its practical application within organizations.
- To face these major challenges, companies need solutions to operationalize the recommendations on Responsible AI.
- AI questions all economic processes and the organization of societies. Its omnipresence raises questions for citizens, consumers and employees who see in it the opportunity to improve their daily lives but also the risk of losing their business or control over some of their freedoms.

Sharing this observation, the Positive AI founding members, leaders of their sector (Telecom, Cosmetics and Beauty, Health and Consulting), launched Positive AI to make Responsible AI concrete and accessible for all organizations.

The Positive AI's ambition is double:

- To provide organizations that wish to implement Responsible AI with a place for exchange, work, and share practical tools and good practices. This will be possible first thanks to the Positive AI Label framework, assessing the maturity of an organization in its application of Responsible AI, through the analysis of its governance, management of its AI systems and the algorithms that pose the greatest ethical risks and their design methods.
- To contribute to the public debate on the regulation and the future of AI, by working with other public and private players, both French and foreign, to pool a common progress and create a common approach at European level, in particular within the context of the IA Act, the EU's approach to AI centers on excellence and trust, aiming to boost research and industrial capacity while ensuring safety and fundamental rights.

1.3. History and methodology of the Positive AI Label

The Positive AI decision to create a Label is explained by several factors:

1. At the heart of current and future major innovations, Artificial Intelligence, however promising it may be, raises questions for citizens and public authorities and creates a demand for transparency and ethics. To respond to and make progress on this subject, companies must assume their responsibilities and commit themselves to Responsible AI for their customers, their employees and society.
2. Labelling is a recommendation of most reports on AI (European Commission White Paper, Montaigne Institute, Villani Report).
3. The Positive AI Label aims to make Responsible AI concrete and accessible to all organizations. This approach is in line with the proposals of the public authorities and aims to make them operational.

The label was designed by the data scientists of the founding companies to be as pragmatic and concrete as possible to enable any organization to improve the level of confidence in its AI systems. This framework was then submitted to an external committee of independent experts for their opinion and comments.

The Positive AI Label will be awarded based on a reference framework developed around the ethical principles of Responsible AI defined by the European Commission. Three ethical principles are prioritized by Positive AI:

- Fairness and Equity.
- Transparency and Explainability.
- Human agency and Oversight.

To encourage all actors in the implementation of a Responsible AI, including those less mature on the subject, the label has been designed in a commitment and progress approach according to 3 gradual levels, that reflect the increasing commitment, the progress, and the continuous improvement of the labelled organizations in the implementation of Responsible AI.

Organizations can follow a step-by-step path from a first level demonstrating willingness and first steps towards the implementation of Responsible AI, to a third level demonstrating a strong maturity or even excellence through a high level of compliance with the control objectives of the Positive AI Label framework. The organizations are given a certain amount of time to complete each stage and are encouraged to move from one level to the next through their Positive AI membership.

The Positive AI Label is an asset for an organization for 3 reasons:

- First, it ensures that the organization respects its own objectives and commitments on Responsible AI.
- Secondly, it allows an organization to demonstrate its ability to build Responsible AI systems, which is a demand of most developers working on these topics.
- Finally, it offers a guarantee to the public, who are increasingly aware of the challenges of responsible development of AI and sends a positive employer brand message to current and future employees.

The Label is delivered in its initial version 1 from 2023. Positive AI plans to manage versioning of the Label to ensure traceability of the Positive AI Label framework evolutions and to make it possible to link the version of the Label and the Label obtained by the organizations. Any evolution of the Positive AI Label framework will be formally recorded and communicated to the organizations.

Labelling must be distinguished from any form of certification. The Positive AI Label responds to a specific framework (refer to section §4.) that is meant to assess the conformance of eligible organizations to the Positive AI requirements within the labelling scope, while certification is generally meant to assess the conformance to a specific norm issued by a public administration or regional/international organizations for standardization (CEN, ISO etc.). In addition, in the context of the AI Act, especially its implementation and enforcement, it is important to underline that the issue of a Labelling Certificate to the organization, even though it demonstrates strong responsible values with regards to audited AI, does not directly induce or imply the organization's conformance to the IA Act or any other certifications existing on the market. This allows the organization to build a foundation that will facilitate its conformance without replacing regulators, their requirements and assessment of the measures taken by the organizations.

1.4. Document reading guide

The present document “Positive AI Label: Guide for a Responsible Artificial Intelligence” sets out how Positive AI members should conform with Positive AI standards and Label framework (refer to section §4.). It applies to all Positive AI members who wish to apply for the Label.

This is a living document and Positive AI reserves the right to revise it based on implementation experience and emerging good practices. The official language of this document is English, and the most recent version is available on Positive AI website: refer to <https://positive.ai/fr>.

This is the version 1, published in March 2023, and effective from March 1st, 2023, for all labelling audits conducted against the Positive AI Label framework version 1.



2

Label
governance

2. Label governance

2.1. Positive AI Association

2.1.1. Presentation

The Positive AI Association is a non-profit membership association founded in 2022 by four companies:

- BCG Gamma.
- L'Oréal Group.
- Malakoff Humanis.
- Orange France.

2.1.2. Composition

The founding companies, as all future members of Positive AI, will participate in the animation of the Association to promote exchanges, the sharing of good practices, and the creation of tools to support all organizations in the implementation of Responsible AI.

Regarding the label, the founding companies and the future members of the Association will decide the changes to be made over time. Some companies will also be present on the Labelling Committee which, based on the reports of independent third-party auditor, will decide on the award of the label to the applicant organization (refer to section §2.3.).

Moreover, the Positive AI founders have set up an external committee of independent experts to provide advice ensuring that Positive AI meets the expectations of civil society and continues to enrich its Label framework. The committee of experts is composed of three researchers recognised for their research work on Artificial Intelligence domain, who regularly share their opinions to the Association. It is an independent advisory body, responsible for instructing Positive AI's advances in support of the advanced knowledge on Responsible AI and the diversity of skills of each of its members.

2.1.3. Role and missions

Positive AI Association is expected to:

- Regularly review and update its Positive AI Label framework to ensure it remains relevant and fits for purpose with regards to the evolution of the legislation in force.
- Support and oversee the quality, integrity, and credibility of audit processes.
- Offer members' support.
- Keep relevant online information (including labelling information) up to date.
- Administer the Positive AI Complaints Mechanism and carry out disciplinary proceedings where required.
- Publicly report the impact of Positive AI programs.
- Engage industry and other stakeholders to build awareness of the value of Positive AI Label.

Note that Positive AI will not act as auditor or undertake audits of members in its own right.

2.2. Positive AI Members

2.2.1. Presentation

Any organization that submits to Positive AI a completed membership form with the attached supporting documents, undertakes the payment of the annual Positive AI commercial membership (depending on the membership type) and for which the application is accepted and validated by Positive AI, becomes a Member of the Positive AI Association.

Three types of Positive AI Members have been defined by the Association:

- Founding members: BCG Gamma, L'Oréal Group, Malakoff Humanis and Orange France.
- Active members.
- Associate members.

Positive AI Members are part of the general assembly of the Association.

Any organization that wishes to apply for a Labelling Audit by a Positive AI Referenced Auditor must be a Positive AI Member.

Note that it is not mandatory for a Positive AI Member to apply for a Labelling Audit. It is up to the organization to indicate its readiness and willingness to start a labelling process.

2.2.2. Role and missions

Positive AI Members are expected to:

- Operate their business in a way that meets the requirements of Positive AI Label.
- Dedicate resources to ensure ongoing compliance with these requirements over time.
- Communicate and train personnel about the Positive AI Label and how to meet the control objectives.
- Promote responsible business practices within their organization an AI uses.
- Participate in discussions on the Positive AI Label framework and on good practices in development of Responsible AI.

Positive AI Members wishing to apply to the Positive AI Label are, in addition, expected to:

- Engage an auditor to carry out their Labelling Audits.
- Give auditors access to facilities, personnel and any information and records they need to assess conformance with the Positive AI Label.
- Implement corrective action or improvement plans, as appropriate.

Thematic committees or consultations (for example on the evolution of the Positive AI Label framework, supported by feedbacks from Labelling Audits, developments in the field of Responsible IA and research documentation) can also be organised by Positive AI on specific projects, particularly with Positive AI Members.

2.3. Labelling Committee

2.3.1. Presentation

The Labelling Committee is a body created by the Positive AI Association with the objective to issue the Positive AI Label.

2.3.2. Composition

The Labelling Committee is composed of:

- Representatives of the four founding companies: BCG Gamma, L'Oréal Group, Malakoff Humanis and Orange France.
- An independent expert recognized in AI research and its implication within the society.
- Representatives of some Positive AI members.

The stakeholders are appointed by Positive AI for a determined period in a nominative capacity. Applications are open to all Positive AI Members. The selection is made by the Labelling Committee (at its launch by the founding members and the independent expert, then by the Committee Labelling) based on several criteria, namely, the candidate's expertise, motivation, involvement in Positive AI and the representativeness of their organization in relation to Positive AI Members (activity, size, etc.).

Each candidate may be accompanied by one substitute within the same organization and after validation by the Labelling Committee. The substitute may replace him/her during the Labelling Committees if necessary.

At the end of the period, each organization is free to present a new candidacy or not to renew his mandate. In case of withdrawal during the term of office, the organization shall inform the Labelling Committee of their decision by email.

2.3.3. Role and missions

Labelling committee is expected to make labelling decisions and issue label agreement to the applicant organizations that comply to the Positive AI Label framework, following audits carried out by external auditors based on Positive AI Label framework.

2.3.4. Functioning

The Labelling Committee meets periodically regarding the volume of labelling processes to review and assess audit reports to take labelling decisions based on the recommendation of the Positive AI Referenced Auditors.

The Labelling Committee shall take its decisions by seeking consensus among its members. In the event of disagreement, decisions shall be taken by a simple majority vote. Each member of the Labelling Committee shall have one vote.

The Labelling Committee can only take a decision when a quorum of two-thirds of the members is present.

Functioning principles and conflict of interest management

The Labelling Committee issues labelling decisions based on the following principles:

- A founding member shall not be represented in the discussions and labelling decision concerning its application.
- A member of the Labelling Committee should not be represented in the discussions and labelling decision concerning the application of an organization from the same sector.

Confidentiality

An abridged version of the audit report is produced by the Positive AI Referenced Auditor for the Labelling Committee. This version does not contain any sensitive information about the Entity's AI activity to ensure confidentiality and protection of its intellectual propriety.

Any information submitted to Positive AI by the third-party auditor will be kept strictly confidential by Positive AI and Its labelling committee. Information will be used solely for the purposes of Labelling decision.

2.4. Positive AI Referenced Auditors

2.4.1. Presentation

The Positive AI Label is based on an audit carried out by Positive AI Referenced Auditors, independent and competent in the Responsible AI domain. Auditors must be Referenced by Positive AI to carry out Labelling Audits.

The third-party auditor will perform the Labelling Audit with respect to the main following principles: integrity, fairness, due professional care (to apply a professional judgement), confidentiality of information, independence, evidence-based approach (to follow a rational method), and risk-based approach (to consider risks and benefits).

2.4.2. Role and missions

A Positive AI Referenced Auditor is expected to:

- Carry out initial, surveillance and renewal audits regarding the Positive AI Label framework.
- Identify and categorize non-conformances to the label and suggest improvements to the organization (when applicable).
- Recognize when audit objectives are unattainable and tell the organizations and Positive AI why this is so.
- Prepare audit reports, including a conformance evaluation, for the organization and the Positive AI Labelling Committee.
- Review progress in implementing any corrective actions by the organization.

Note that Auditors have a legal relationship with the Positive AI Member applying for a Labelling Audit (named the Entity) that has hired them. The company doing a third-party audit cannot advise or assist an Entity in any project related to artificial intelligence or help the Entity to be compliant to the Positive AI Label framework, that could lead to a conflict of interest. Nevertheless, the Auditor can perform a preliminary diagnosis of conformance if requested by the Entity.



3

Labelling
process

3. Labelling Process

The Positive AI Label framework is designed for use by Positive AI Referenced Auditors to ensure an Entity's compliance for the purposes of granting Positive AI Label. The Entity's labelling scope is defined by the Entity seeking to be labelled (refer to section §3.2.).

In the rest of the document, an Entity corresponds to an organization, member of Positive AI, that decides to enter the Labelling process and to comply with the Positive AI Label framework.

The steps for Positive AI Labelling process are as follows:

1. **Membership:** The Entity becomes a Member of the Positive AI Association (refer to section §3.1.1.).
2. **Eligibility verification:** The Entity checks its eligibility to the Positive AI Label following different criteria (refer to section §3.1.2.).
3. **Submitting of the application and contracting with a Positive AI Referenced Auditor:** The Entity can submit its application to Positive AI Referenced Auditors chosen among auditors selected and referenced by Positive AI (refer to section §3.1.3.), then contract with one Auditor.
4. **Execution of the Labelling Audit:** During the Labelling Audit, the Auditor verifies the Entity has policies and systems in place that conform to the Positive AI Label framework. Minor and major non-conformances will be noted, and the Entity will be directed to address them within an improvement action plan (refer to section §3.1.4.).
5. **Labelling decision:** Based on the Auditor's Recommendation Report and the improvement action plan proposed by the Entity, Positive AI makes labelling decision, and can issue the label agreement to the Entity (refer to section §3.1.5.).
6. **Surveillance Audits:** Depending on the level of labelling obtained by the Entity, the Referenced Auditor conducts one or two Surveillance Audits of the Labelled Entity to verify that policies and systems in place continue to work effectively and non-conformances found during the Labelling Audit have been addressed (refer to section §3.1.6.).
7. **Renewal of the Label:** After the validity period of two years or three years following the level of labelling obtained by the Entity, a Labelling Audit would be required to renew the Entity's Labelling (refer to section §3.1.7.).
8. **Suspension, withdrawal, termination of the Positive AI Label:** The Labelling of the Entity may be suspended, withdrawn, or terminated for different reasons. The Entity can also make the choice to withdraw from the Positive AI Label during its Labelling validity period (refer to section §3.1.8.).

3.1. Labelling cycle

3.1.1. Membership

Any organization that submits to Positive AI a completed membership form with the attached supporting documents, undertakes the payment of the annual Positive AI commercial membership and for which the application is accepted and validated by Positive AI, becomes a Member of the Positive AI Association (refer to section §2.2.).

Any organization that wishes to apply for a Labelling Audit by a Referenced Auditor against the Positive AI Label framework must be a Positive AI member. It is not mandatory for a Positive AI Member to apply for a Labelling Audit.

3.1.2. Eligibility Verification

Positive AI established the following set of eligibility criteria:

- The organization must be a Positive AI Member, whatever its membership type.
- The Positive AI Label is defined as open: It is aimed at any type of organization, in the private or public sector, whatever their size, their configuration, their domain of activity and their place in the AI system value chain in accordance with the EU AI Act (user, distributor, importer, provider), based on a voluntary approach of the organization. The Responsible Label is neutral for a point of sectoral and technological view.
- The organization must be established within the European Union (holding or subsidiaries) and placing on the market or putting into service AI systems or using AI systems in the Union.
- The organization must have Artificial Intelligence Systems in production (at least one AI system in production).
- The organization is engaged in a process to implement Responsible AI principles through formalized commitments (policies, governance, processes, controls, etc.).

After its own eligibility verification regarding these criteria, the organization can submit its application to Positive AI Referenced Auditors (refer to section §3.1.3.). If needed, the organization can request for clarification and information to Positive AI to confirm or invalidate its eligibility.

3.1.3. Submitting of the application and contracting with a Positive AI Referenced Auditor

In case of eligibility, the Entity shall address its application to one (or several if desired) of the Positive AI Referenced Auditor and provide it with information on the nature and composition of the AI systems to analyse its eligibility and to define its Labelling scope (refer to section §3.2.).

The Auditors referenced by Positive AI are listed in Appendix 2.

The Auditor will then review the application form and supporting documents. The Entity may be required to respond to requests for clarification, information, etc.

Based on this review, the Auditor will inform the Entity of the result of the check that will include the following outcomes:

- The Entity application is validated.
- The Entity application is invalidated in cases where it's not possible for the Entity to comply with the Positive AI Label requirements.

If the Entity application is validated, the Entity shall:

- Contract with the Auditor to prepare and plan the execution of the Labelling Audit (refer to section §3.1.4.).
- Accept all terms and conditions which apply to the issue of the Labelling Decision (refer to section §3.1.5.).

3.1.4. Execution of the Labelling Audit

The initial Labelling Audit is performed through different inquiries and interviews with the main Entity stakeholders involved in the AI domain, and through the inspection of the requested documentation (refer to sections §4.3. and §4.4. regarding examples of supporting evidence to be provided by the Entity). As the Labelling Audit takes place over several days, the Auditor can organize intermediate meetings and/or emails exchanges to clarify any ambiguities.

The Labelling Audit is carried out remotely, or on site if requested by the Entity.

The Entity shall ensure the availability of:

- The key interlocutors with a deep knowledge of the audited subjects (the list of profiles and skills required will be communicated by the Auditor during the preparation of the audit with regards to the work to be carried out).
- The documentation required by the Positive AI Label framework and demonstrating that policies and processes comply with the requirements and control objectives of the framework.

During a closing meeting, the Labelling Audit conclusions will be shared by the Auditor to the Entity.

The Labelling Audit is based on a sampling of information and piece of evidence. The absence of non-conformity constitutes a presumption of conformance and not a proof of full conformance with the control objectives of the Positive AI Label framework.

The Entity can be issued with findings such as a minor non-conformance and/or major non-conformance during the Labelling Audit, for which an improvement action plan must be established and shared to the Auditor within one month after the Labelling Audit closing meeting (refer to section §3.3.2.). This improvement action plan is a formal document that will outline the improvement areas and remediation action to respond to each issued non-conformance, with associated responsibilities and timeframes.

The Auditor considers the Entity's exchanges and comments during the closing meeting when establishing the final report.

3.1.5. Labelling Decision

Positive AI Labelling Committee decides upon granting the Positive AI Label to the Entity following the below conditions that must be met:

- The review of the Auditor's Recommendation Report based on the Labelling audit he/she has conducted and including the conformance evaluation (refer to section §3.3.).
- The assessment of the improvement action plan provided by the Entity to address minor and/or major non-conformances and agreed by the Auditor.

The Labelling decision shall be communicated to the Entity, including justification should the decision be not to grant the Positive AI Label (refer to section §6.2.). If the conformance evaluation can be communicated internally to encourage teams to improve and become more involved in the process, it is not possible to communicate externally on this note (refer to section §5.2.).

Once the Labelling decision is ratified and providing the Entity's acceptance of the terms of the Labelling Agreement, Positive AI will issue a Labelling Certificate to the Entity that will be authorized to communicate on the obtained Label and to use the associated logo corresponding to its Labelling Level, with respect to Positive AI conditions.

The Positive AI Label is granted for two years for the Labelling Level 1 and for three years for the Labelling Levels 2 and 3. Positive AI reserves the right to suspend or withdraw labelling during this period if the Entity does not carry their duties appropriately (refer to section §3.1.8.).

3.1.6. Surveillance Audits

Depending on the Labelling Level obtained by the Entity, the Referenced Auditor conducts one or two Surveillance Audits of the Labelled Entity to verify that policies and systems in place continue to work effectively and non-conformances found during the Labelling Audit have been addressed:

Labelling Level	Validity period	Audit types	When it happens	Why it's needed
Level 1	2 years	Initial audit	Once requested by an eligible Entity	To obtain the Positive AI Label
		Surveillance audit	12 months after the initial audit	To review progression against the improvement action plan
		Renewal audit	2 years after the initial audit	To continue demonstrating the Positive AI Label conformance and using the Label logos
Level 2 & Level 3	3 years	Initial audit	Once requested by an eligible Entity	To obtain the Positive AI Label
		Surveillance audits	12 months and 24 months after the initial audit	To review progression against the improvement action plan
		Renewal audit	3 years after the initial audit	To continue demonstrating the Positive AI Label conformance and using the Label logos

The Surveillance Audit, shorter than the initial Labelling Audit, is carried out in a remote documentary way based on the key elements of the Responsible AI governance and systems and on the evidence of corrective actions and improvement plan from the previous audit. The Auditor focuses on the main points of the initial Labelling Audit.

A follow-up report is issued, complementary to the initial Labelling Audit report. The Surveillance Audit does not result in a new conformance evaluation.

3.1.7. Renewal of the Label

Six months before the end of the validity period of the Labelling Certificate, the Entity shall request for a Renewal Audit. To avoid any break in Labelling, it is mandatory to carry out the Renewal Audit before the expiration of the Labelling Certificate, in a delay of three months before the expiration.

The Renewal Audit is equivalent to the initial Labelling Audit and performed following the same rules. However, it will be performed on different AI systems put into service (preferably systems developed since the last Labelling Audit, if any) that those included in the scope of the initial Labelling Audit (refer to section §3.2.).

3.1.8. Suspension, withdrawal, termination of the Positive AI Label

The Labelling of the Entity may be suspended or withdrawn for the following reasons:

- Non-compliance with the contractual conditions of the Positive AI Labelling Agreement.
- Refusal by the Entity to carry out the Surveillance Audit or the Renewal Audit within the deadline notified by the Auditor.
- Refusal to implement corrective actions within the deadline notified by Positive AI (refer to section §3.3.).
- Important changes made in the organization's AI systems, not declared to Positive AI and thus not taken into account.

To withdraw from the Positive AI Label, the Entity must inform Positive AI by email and suspend its contract with its Auditor with respect of the contractual conditions.

In case of suspension, withdrawal or termination of the Positive AI Label, any use by the Entity of the Labelling Certificate and of the Positive AI Label logo will be prohibited.

3.2. Definition of the Labelling Audit scope

3.2.1. About the Labelling scope

The Positive AI Labelling scope refers to the extent of an Entity to which the Positive AI label framework will apply. The Positive AI label framework requires a clearly and accurately documented Labelling scope that is established by the Referenced Auditor to ensure appropriate auditing activities.

Each Entity's Labelling scope will be different, depending on its organization extent, and on the number and risk level of audited use cases. Once established, the Labelling scope will define exactly which Positive AI Label framework's control objectives apply. In all cases, an Entity's labelling scope is contractually validated between the Entity and the auditors before starting the Labelling Audit.

Positive AI encourages the Labelled Entities to contact them if they believe anything has been left out or misrepresented in an Entity's Labelling scope.

3.2.2. Criteria to define the Labelling scope

The following definitions and considerations apply:

AI system definition

For the Positive AI Label, the definition of AI system mentioned in the IA Act applies.

As of 25 November 2022, the following definition in the Article 3 of the AI Act applies: "Artificial Intelligence system (AI system) means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts."

AI systems risk level assessment

The applicant organization must have a methodology for rating the risks of its AI systems to classify them as: Low / Medium / High risk AI system. This risk rating methodology will be audited during Pillar 1 audit and should be based on the impact of AI systems on the humans with whom it interacts: customers, third parties and employees refer to section §4.).

The risk assessment methodology should answer the following questions:

- To what extent does the company's use of the AI system pose an ethical risk to the customer and/or to third parties?
- Does the use of the AI system impact the work of the employees? Does this impact entail ethical risks for the employees?

In addition, the Entity should consider the classifications rules for forbidden and High-Risk AI systems mentioned in the Article 6 of the AI Act.

Finally, it could integrate reflections on the following questions as the volume of recommendations provided, the existence of human interactions, the level of decision supported.

Based on these definitions and considerations, the Entity can define the perimeter of AI systems to enter in the Labelling scope consideration.

Positive AI Label framework for definition of the Labelling scope

The Positive AI Label framework is divided into 2 pillars: one related to Responsible AI governance defined by the organization and the second related to the responsible AI consideration in the development and the use of the different AI systems developed by the organization (refer to section §4.).

Pillar 1 of the Positive AI label framework: Governance of the AI design and usage

The Pillar 1 applies to the overall organization of the Entity that is part of the Labelling Audit. All the company's activities integrating AI are concerned, as well as the activities supporting them (IT, security, purchasing, recruitment, etc.). The Entity must demonstrate its conformance with the standard for the entire brand it wishes to be labelled (holding group, subsidiaries, etc.).

Pillar 2 of the Positive AI label framework: Validation of algorithmic models

For the Pillar 2, the Labelling scope varies according two different criteria:

- The Labelling Level the Entity applies that determines the number of audited use cases.
- The risk level of the Entity's use cases. For the Labelling Levels 2 and 3, a minimum number of High-Risk use cases must be audited (when existing). Entities without High-Risk use cases are eligible to the Labelling Levels 2 and 3 if they comply with the associated requirements.

The below table presents the Labelling scope principles for the Pillar 2:

Labelling level	Level 1	Level 2	Level 3
Number of pre-identified use cases	≥4	≥8 including 4 High-Risk use cases if existing	No pre-identified use cases. Random selection on the overall population of existing use cases.
Number of audited use cases	1	2	From 1 to 5 (Refer to below details)
Of which, audited high risk use cases	≥0	≥1 (when existing)	≥ 50% of audited use cases (when existing)

The following rules apply:

- For the Labelling Levels 1 and 2, the organization pre-identifies a set of use cases that it wishes to submit for the Labelling Audit. The Auditor randomly selects use cases to audit: 1 for the Level 1 (whatever the risk level of the use case) and 2 for the Level 2 (from which 1 High-Risk use case if existing).
- For the Labelling Level 3, the organization does not pre-identify a set of use cases and submits for the Labelling Audit all its existing and inventoried use cases. The Auditor randomly selects use cases to audit (from 1 to 5) following its number of existing use cases (the sampling for the selection of use cases is performed statistically with a centred normal distribution):

Number of existing use cases	Number of audited use cases
From 1 to 3 use cases	100% examination

From 4 to 8 use cases	3
From 9 to 33 use cases	4
Greater than or equal to 34 use cases	5

Note that the criteria for selecting use cases and classifying their risk level will be assessed by the Positive AI Referenced Auditor with the Entity during the Labelling Audit.

3.2.3. Setting the Labelling scope

Each Entity is responsible for setting its own labelling scope. The information below sets out the general rules to follow. The Entity can contact Positive AI for any query about the Labelling Scope.

Pillar 1 of the Positive AI Label Framework: Governance of the AI design and usage

The applicant Entity decides the scope on which it wishes to be labelled (holding group, subsidiaries, facilities, etc.). Then, the Pillar 1 applies to the entire organization and all the activities of the Entity on the chosen scope that is part of the Labelling Audit.

Pillar 2 of the Positive AI Label Framework: Validation of algorithmic models

The applicant Entity assesses its use cases risks level according to its own methodology with regards to the Positive AI Label requirements, then chooses a Labelling Level to apply for, regarding its maturity in Responsible AI. Considering the chosen Labelling Level, the Entity prepares the list of use cases to be submitted for the Labelling Audit (refer to section §3.2.2.).

3.2.4. Dealing with changes in scope

Positive AI encourages the Labelled Entities to contact them if they believe anything has been left out or misrepresented in an Entity's Labelling scope.

Moreover, the Labelled Entity must inform Positive AI of any changes to its organization and / or AI systems that may affect the Labelling scope.

The Labelling Scope may change if some changes occur at the level of the Labelled organization such as (not exhaustive list):

- An organizational restructuring.
- Divestments, acquisitions, or changes to the equity share.
- New put into service IA systems, especially High-Risk use cases.
- Etc.

Positive AI will assess how the Entity's changes impact or not the validity of the issued Label, and the conduct of the different audits (initial, surveillance, renewal).

Positive AI encourages the Labelled Entities to contact them if they believe anything has been left out or misrepresented in an Entity's Labelling scope.

Also, if a Labelled Entity detects or is informed of a major or critical incident on one (or several) of its AI systems, the Entity must inform Positive AI.

3.3. Conformance evaluation criteria

3.3.1. Defining conformance rating

For each control objective (refer to section §4.2.), the auditor assesses the maturity of the audited Entity as follows:

- **A:** The control objective is met with a high level of monitoring.
- **B:** The control objective is achieved with existing good practices.
- **C:** The control objective is partially achieved, with little or no formalization.
- **D:** The control objective is not achieved; nothing is put in place.

A correspondence between the degree of maturity / performance and the rating is achieved such as:

Evaluation	Points
A	20
B	15
C	5
D	0

The **conformance rate** is calculated for the Pillar 1, and for each use case audited within the Pillar 2, as follows: Number of points obtained / Maximum number of points of the Pillar. To obtain the Labelling Level, each conformance rates must achieve the minimum conformance rate defined as threshold.

The **global conformance rate** is the average of the conformance rates obtained on the Pillar 1 and on each use case audited within the Pillar 2.

The maximum number of points is defined as follows:

Pillar	Number of control objectives	Maximum number of points
Pillar 1	15 + 4 as optional ^(*)	300
Pillar 2 - Low-Risk use case	5	100
Pillar 2 - Medium-Risk use case	11	220
Pillar 2 - High-Risk use case	17	340

^(*) Some control objectives are, for now, considered as optional requiring a high maturity in the Responsible AI approach: they will allow Entities to obtain bonus points where these practices are in place, without penalizing Entities that do not address these issues.

To encourage all actors to engage in the Responsible AI approach, regardless of their level of maturity, the Positive AI Label will be composed of 3 Labelling Levels, corresponding to gradual requirements regarding the compliance with the Positive AI Label framework:

- **Level 1:** For organizations achieving a minimum global conformance rate of 40%. It demonstrates the commitment of the organization to the Responsible AI approach and a minimum maturity.
- **Level 2:** For organizations achieving a minimum global conformance rate of 60%. It demonstrates a more advanced maturity of the organization in the Responsible AI approach.
- **Level 3:** For organizations achieving a minimum global conformance rate of 80%. This corresponds to the final Labelling Level, which rewards a high level of maturity of the organization in Responsible AI.

3.3.2. Documenting non-conformances

At the end of the Labelling Audit, the Entity can be issued with findings such as a minor non-conformance and/or major non-conformance during the Labelling Audit, for which an improvement action plan must be established and shared to the Auditor within one month after the Labelling Audit closing meeting. The following definitions, timescale and implications apply:

Finding	Correspondence	Definition	Timescale for closure	Implications if no demonstration of progress
Minor non-conformance	<ul style="list-style-type: none"> • Correspondence with a control objective's evaluation "C: The control objective is partially achieved, with little or no formalization". 	<ul style="list-style-type: none"> • The Entity's IA practices perform in a way that does not wholly conform to the relevant Positive IA Label framework, due to an isolated lapse of a control. 	<ul style="list-style-type: none"> • Corrective actions proposed within 1 month after the Labelling Audit closing meeting. • Progress of the implementation of corrective actions assessed during the Surveillance Audit. 	<ul style="list-style-type: none"> • Upgraded to a major non-conformance.
Major non-conformance	<ul style="list-style-type: none"> • Correspondence with a control objective's evaluation "D: The control objective is not achieved; nothing is put in place". 	<ul style="list-style-type: none"> • The Entity's IA practices perform in a way that does not conform to the relevant Positive IA Label framework, due to the total absence of implementation or evidence of a control. 	<ul style="list-style-type: none"> • Corrective actions proposed within 1 month after the Labelling Audit closing meeting. • Progress of the implementation of corrective actions assessed during the Surveillance Audit. 	<ul style="list-style-type: none"> • Suspension of labelling.

A photograph of a woman with long dark hair, wearing a white lab coat, looking directly at the camera. She is surrounded by a shower of colorful confetti (blue, green, red, yellow) falling around her. The background is dark with some blurred lights. On the left side of the image, there are three overlapping pink geometric shapes: a rectangle at the top, a triangle pointing right in the middle, and another rectangle at the bottom.

4 ■

Positive AI
Label
framework

4. Positive AI Label framework

About the Label framework

The Positive AI Label framework sets out how Positive AI Referenced Auditors should assess an organization's conformance with Positive AI requirements.

The Audit Framework is a living document and Positive AI reserves the right to revise it based on implementation experience and emerging good practices. The official language of the document is English. The version of the Positive AI Label framework associated with a labelling decision will be outlined in the labelling agreement.

Disclaimer

No guarantee, warranty or representation is made as to the accuracy or completeness of the Positive AI Label framework and other documents, or information sources referenced in it. The framework is not intended to, nor does it, replace, contravene, or otherwise alter the requirements of any applicable national, state, or local governmental statute, law, regulation, ordinances, or other requirements.

The Positive AI Label framework gives general guidance only and should not be regarded as a complete and authoritative statement on any of the topics covered by it.

Feedback

Positive AI welcomes comments and feedback from all stakeholders on this audit framework. These can be submitted using the following link: <https://positive.ai/fr/contact>.

4.1. Key Responsible AI Principles

The Positive AI Label is awarded according to a Label framework mainly developed around 3 key principles of Responsible AI defined by the European Commission (refer to [Ethics guidelines for trustworthy AI | Shaping Europe's digital future \(europa.eu\)](#)) and prioritized by Positive AI in its first version of its Label:

1. **Fairness and Equity (Diversity, Non-discrimination, and fairness):** This is to prevent discrimination biases that may exist between data and design, and to ensure regular evaluation and auditing of algorithms.
2. **Transparency and Explainability:** This is to ensure that the data and algorithms at the heart of AI systems are accessible and understandable, but also that the results obtained by an algorithm are both explainable and reproducible each time an identical case is submitted to the algorithm. Finally, the use of the results should also be transparent.
3. **Human agency and Oversight:** This is to ensure that human feedback mechanisms are in place in organizations, but also the ability to review automated decisions.

These principles are evaluated according to the impact on customers, third parties and employees:

- **Customers and third parties:** To what extent does the company's use of the AI systems pose an ethical risk to the customer and/or to third parties?
- **Employees:** Does the AI systems' use have an impact on the work of employees? Does this impact entail ethical risks for them?

The evolution of the Positive AI Label and associated framework may consider other key principles. In particular, the principal related to societal and environmental well-being in one of the long-term priorities of the Label.

4.2. Structure of the Positive AI Label framework

4.2.1. Global structure

The Positive AI Label framework is divided into 2 pillars: the first one related to Responsible AI Governance defined by the organization and the second related to the Responsible AI consideration in the development and/or the use of the different AI systems of an organization.

Each pillar is divided into different sections:

Pillar 1: Governance of the AI design and usage

1. Overall Governance Operating Model with respect to data science use cases
2. Policies
3. Responsible AI Operationalization Strategy
4. Use Case Inventory
5. AI Lifecycle
6. Production control framework
7. Data Management and Privacy Policies
8. Overall IT infrastructure (IT stack)
9. Responsible AI at the level of the organization and incorporation in support functions (Human Resources, Internal Audit, etc.)

Pillar 2: Validation of algorithmic models

1. Model Documentation
2. Model Design Assessment
3. Ongoing Performance Monitoring
4. Functional Process Controls and Implementation tests
5. IT & architecture Process Controls and Implementation Tests
6. Identification, assessment and management of bias and risks to fairness
7. Explainability & Interpretability of the model and of its results
8. Sensitivity Analysis of the model to its features
9. Hypothesis Testing when designing the algorithm
10. Performance Testing
11. Impact Assessment of known weaknesses and open improvements points
12. Compliance Review to the regulatory and legal requirements

4.2.2. Control objectives

For each of the two main dimensions analysed by the label (the overall organization and governance of AI, as well as algorithms and design methods), the Positive AI Label framework seeks to measure how the organization is compliant regarding a set of control objectives (19 and 17 respectively), to assess its maturity and compliance level.

Each control objective is associated with one or more of the ethical principles prioritized by Positive AI (Fairness and Equity, Transparency and Explainability, Human agency and Oversight) and corresponds to an internal activity and/or internal control for which an Entity must ensure conformance. The control objectives are described in an Audit framework containing examples of supporting evidence that could be requested by the Auditor (non-exhaustive list) (refer to section §4.3.).

4.3. Pillar 1 of the framework: Governance of the AI design and usage

4.3.1. Section 1: Governance

This is the key section of the Positive AI Label framework. The Referenced Auditor assesses the vision and strategy of the organization in relation to ethical and responsible artificial intelligence.

An AI model is a complex computer object with important ethical issues (sensitivity of processed data, impact of predictions). To mitigate the associated risks as well as possible, it seems essential to frame AI practices with adapted policies and governance.

The following control objectives are audited:

Control objectives	Example of supporting evidence
The organization defines the roles and responsibilities allowing an effective implementation of AI governance policies.	<ul style="list-style-type: none"> - Corporate Governance Framework - AI Organigram - AI Governance Policy - Board and Executive Committee Charters - Educational Qualifications and Suitably attested education/training certificates of the AI Personnel
The organization established committees responsible for formulating and implementing Responsible AI principles. For example, operating committees are in place for signing off on the compliance of individual use cases with the Responsible AI principles.	<ul style="list-style-type: none"> - Charters and sub-unit structures for all AI related operations. - Educational Qualifications and Suitably attested education/training certificates of the AI Personnel
The organization establishes internal reporting standards for AI systems such that accurate, actionable, and timely information regarding AI systems (across model life cycle) is communicated to all relevant stakeholders (which reports, to whom, what frequency).	<ul style="list-style-type: none"> - Internal reporting Templates - AI - External Reporting Templates - AI - AI Model Policy - AI Model Framework
The organization carries out policies and procedures pertaining to transparency (towards individuals exposed to or subject to AI systems) and disclosure (general representations for the use by third parties).	<ul style="list-style-type: none"> - AI Transparency or Disclosure Rules - AI Stewardship Requirements - AI Governance Rules - AI Collection Rules
The EU AI Act (draft 2022) points out that all parties involved in the putting in the market of an AI system have a responsibility. The organization assesses its place in the AI system value chain (for each use case) in accordance with the EU AI act (User, distributor, importer, provider).	<ul style="list-style-type: none"> - Corporate Governance Documents - Business Strategy Guidance

4.3.2. Section 2: Policies

This section aims to analyse the policies that the organization puts in place to reflect its ethical and responsible vision.

Implementation of ethical and responsible AI cannot take place without a clear definition (and then internal communication) of the ethical and responsible principles with which the organization wishes to comply.

The following control objectives are audited:

Control objectives	Example of supporting evidence
The organization incorporates the principles of Responsible AI principles and details supervision of IA risks into the group policy or into existing policies where relevant.	<ul style="list-style-type: none"> - AI Model Risk Policy - AI Model Governance Policy - AI Model Validation Framework - AI Model Monitoring Framework - AI Model Implementation Framework

4.3.3. Section 3: Responsible AI Operationalization Strategy

This section aims to analyse more precisely the organization’s operationalisation strategy, a critical phase between the design and use of AI algorithms, with the entry of ethical risks associated with the ways in which the algorithm is used by the business teams.

Operation phase of AI is a key phase in its life cycle during which the governance of the AI system has shifted from data science to the business and a new articulation is needed to maintain the same level of control over the data and its impact as well as over the possible drifts of the algorithm.

The following control objective is audited:

Control objectives	Example of supporting evidence
The organization defines a clear strategy for the operationalisation of Responsible AI.	<ul style="list-style-type: none"> - Corporate Governance Framework - Overall Enterprise AI policy - Senior Management notes and vision on responsible AI - Meeting Minutes from AI committee meetings

4.3.4. Section 4: Use case inventory

This section aims to analyse the management of the organization regarding the use case inventory, which is at the heart of the overall AI management and risk analysis to ensure that the entirety of the use cases is visible.

The increasing use of AI technologies to support and inform the business generally results in many use cases within an organization. Despite this volume, it is necessary to maintain visibility and control over all the use cases, so rigorous inventory and documentation work is key to monitoring.

The following control objectives are audited:

Control objectives	Example of supporting evidence
The organization clearly defines the information regarding the purpose and risk rating of the use cases in accordance with the Responsible AI guidelines, that must be captured in the use case inventory.	<ul style="list-style-type: none"> - Inventory Tool Documentation - Documentation of Upstream and Downstream functionalities to the tool - Tool Controls checklist & Escalation Matrix
The organization uses a centralised tool capable of recording and tracking accurate and complete information about the various AI use cases.	<ul style="list-style-type: none"> - Inventory Tool Documentation - Documentation of Upstream and Downstream functionalities to the tool - Tool Controls checklist & Escalation Matrix
The organization operates appropriate processes to capture accurate and complete information about AI use cases.	<ul style="list-style-type: none"> - Inventory Tool Documentation - Documentation of Upstream and Downstream functionalities to the tool - Tool Controls checklist & Escalation Matrix

4.3.5. Section 5: AI lifecycle

This section aims to analyse the management of the AI life cycle by the organization and the assurance of risk management throughout this process.

An AI model is a complex computer object that can evolve as it is learned. Tracing the stages of its development and evolution makes it possible to constitute an end-to-end documentation of its life cycle, a prerequisite for reproducing or auditing a model.

The following control objectives are audited:

Control objectives	Example of supporting evidence
The organization defines and follows AI lifecycle management processes with appropriate controls.	<ul style="list-style-type: none"> - AI Model / Use case Family Standards - AI Model / Use case Technical Document - AI Model / Use case Development Documentations - AI Model / Use case Monitoring and Validation Reports - Controls Checklist for development, monitoring and validation of the AI Model / Use case
The organization documents all use cases in an inventory in a complete manner, for all defined project stages, and in a timely basis.	<ul style="list-style-type: none"> - AI Model / Use case Family Standards - AI Model / Use case Technical Document - AI Model / Use case Development Documentations - AI Model / Use case Monitoring and Validation Reports - Controls Checklist for development, monitoring and validation of the AI Model / Use case

4.3.6. Section 6: Production Control Framework

This section aims to analyse the organization's management of the production environment. This environment is key in limiting the ethical risks that may impact AI.

Production of an algorithm is a key phase involving a set of risks of fairness, explicability, and human intervention (data used, chosen design method, modifications made, etc.).

The following control objective is audited:

Control objectives	Example of supporting evidence
The organization ensures sufficiently robust controls framework for the AI production environment.	<ul style="list-style-type: none"> - AI Model / Use case Family Standards - AI Model / Use case Technical Document - AI Model / Use case Implementation Documents

4.3.7. Section 7: Data Management and Privacy Policies

This section aims to analyse data management and compliance with data protection and personal data protection regulations.

Use of personal or confidential data carries the risk of exposure, which can have very detrimental consequences for the producers, managers, or subjects of such data. In particular, in data science projects, they must therefore be protected and the risks of them being leaked or exposed must be minimised.

The following control objective is audited:

Control objectives	Example of supporting evidence
--------------------	--------------------------------

The organization establishes and follows data management and data privacy policies, applicable and taken into account in Responsible IA governance.

- Data Dictionary
- Data Governance Policy and Documentation
- GDPR policy implementation (if applicable)
- Controls checklist for data management

Please note that issues relating to the technical and security safeguards associated with data retention are addressed in the control objective below.

4.3.8. Section 8: Overall IT infrastructure (IT stack)

This section aims to analyse the management of the IT infrastructure in which the AI use case will be inserted and which, in the absence of adapted safeguards, constitutes a real integrity risk for the algorithm.

As the AI algorithm fits into a more complex IT environment (integrating other systems, with a defined architecture and security, etc.), it is necessary that this environment is adapted to the risks represented by AI and complies with ethical and accountability requirements (data security, traceability of human interventions, third-party algorithms complying with the issues of equity and justice, etc.).

The following control objectives are audited:

Control objectives	Example of supporting evidence
The organization deploys sufficient cyber security measures for each of the AI use cases.	<ul style="list-style-type: none"> - IT & Information Security policy - Change Management & cyber resilience policy - Controls checklist for IT Governance, change management and cyber resilience
The organization follows clearly defined third- party AI risk management policies and practices within its overall AI governance.	<ul style="list-style-type: none"> - IT /IS Policy - Third Party engagement & due diligence policies

4.3.9. Section 9: Responsible AI at the level of the organization and incorporation in supporting functions (Human Resources, Internal Audit, etc.)

This section aims to analyse other key dimensions of organizational functioning, which may impact on the handling of ethical and responsible AI.

Beyond the core business, the challenges of AI affect the whole company, their mitigation must be considered more widely.

The following control objectives are audited:

Control objectives	Example of supporting evidence
Existing firm policies (E.g., Firm Conduct, Ethics Principles, ESG Policies, etc.) incorporate Responsible AI guidelines.	<ul style="list-style-type: none"> - Corporate Governance Policy - Firm wide ESG policy - Ethics, Conducts and Risk Appetite Frameworks
The organization defines and maintains appropriate policies and procedures to ensure that well qualified employees are hired for developing and putting in production AI systems, and that all employees involved in AI governance are well trained, up to date and aware of the latest Responsible AI guidelines.	<ul style="list-style-type: none"> - Risk Appetite Framework - ERM policy - Corporate Governance
The organization establishes an appropriate Internal Audit approach for Responsible AI.	<ul style="list-style-type: none"> - Internal Audit Reports & Framework

4.4. Pillar 2 of the framework: Validation of algorithmic models

The Pillar 2 of the Positive AI Label framework aims at auditing the organization functioning regarding IA conception and usage, at the scale of the algorithms. The Referenced Auditor checks if best practices are indeed put in place when designing, monitoring, and using IA algorithms. The use cases' risk level determines the control objectifs to be audited.

4.4.1. Section 1: Model documentation

This section is at the basis of the second pillar of the Positive AI Label framework. The Referenced Auditor assesses use case documentation to get a general overview of the use case.

The following control objectives are audited:

Control objectives	Use case risk levels	Example of supporting evidence
The organization defines a set of documentation of the model, including the business purpose, the modelling choice, the performance measures, and the Operational Design Domain. Moreover, there should be a version of this model overview accessible to layman business users.	Low, Medium, High	Model Document (and associated help files for model modules) <ul style="list-style-type: none"> ▪ Scope ▪ Purpose & usage ▪ Model Design & methodology ▪ Input & Output data ▪ Risk category & Risk analysis ▪ Implementation environment ▪ Model performance ▪ Policies and standard of care documents
The organization establishes the model documentation with the objective of compliance to the documentation standards of the entity and good modelling practices.	Low, Medium, High	<ul style="list-style-type: none"> - Risk Appetite Framework - ERM policy - Corporate Governance

4.4.2. Section 2: Model design Assessment

This section aims to analyse the model design choices regarding data used and models selected by data scientists and to ensure they are in line with the modelling good practices.

The following control objectives are audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization assesses its Model Design and deploys a model/algorithm approach to ensure its overall appropriateness for addressing the stated business purpose, and in line with good modelling practices.	Low, Medium, High	1. Model Document (and associated help files for model modules) <ul style="list-style-type: none"> ▪ Literature review ▪ Model limitations & assumptions ▪ Transparency and clarity of information on the modelling approach
The organization assesses its Model Design and deploys a data selection approach to ensure the data used are overall appropriate for addressing the stated business purpose, and in line with good modelling practices.	Low, Medium, High	<ul style="list-style-type: none"> ▪ Functional information and source of the model/model components (inhouse, open source, vendor) ▪ Data Review <ul style="list-style-type: none"> o Data source, nature of data, data transformation/pre-processing o Training data set construction and governance decisions
The organization assesses its Model Design and deploys a model/algorithm approach to ensure its overall appropriateness for addressing the stated business purpose, and in line with good modelling practices.	Low, Medium, High	2. Sign-off mail from compliance (GDPR sense), data management policies <ul style="list-style-type: none"> ▪ Data protection and privacy considerations

4.4.3. Section 3: Ongoing Performance Monitoring

This section aims to analyse more precisely the monitoring strategy put in place and the consideration of the performance reports. The algorithm/use case having been designed to operate within certain boundaries (e.g., absence of bias with regards to certain segments, similar accuracy across ranges of application, ...), the organization must define the monitoring strategy implemented to assure that such boundaries are respected on an ongoing basis.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
<p>The organization defines an ongoing performance monitoring strategy and a framework of algorithmic performance tests in line with the business purpose of the IA system.</p> <p>Next to the design of performance metrics and their alert levels, the organization also considers the outcome of actual algorithm performance reports.</p>	Low, Medium, High	<ol style="list-style-type: none"> Model monitoring plan/Documentation/Utility tools <ul style="list-style-type: none"> Strategy for model performance monitoring <ul style="list-style-type: none"> Process details, methodology frequency Input data Alert threshold Roles and responsibilities Model Monitoring historical reports/Mail communication <ul style="list-style-type: none"> Model overlays/overrides specification and justification

4.4.4. Section 4: Functional Process Controls and Implementation Tests

This section aims to analyse the overall management of the “running” phase of the project, especially to mitigate risks linked to data used or created through human interactions.

The following control objectives are audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization operates appropriate process controls regarding the operational running and business use of the model, including the data management and the input/output data flows.	Medium, High	<ol style="list-style-type: none"> Model Document/User manual Communication mails Model Monitoring Report Desktop view of AI user interface
The organization operates appropriate process controls regarding the interactions with expert users and data subjects, and the model version change management.	Medium, High	

4.4.5. Section 5: IT & Architecture Process Controls and Implementation Tests

This chapter aims to analyse the management of the implementation phase by the organization, to ensure that appropriate controls and infrastructure are in place to secure this important phase in IA projects.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization performs and documents standard implementation test procedures of the production model to assess whether the implemented model corresponds to the design and to the model documentation. It also considers the operational robustness of the implementation, and the creation of audit logs.	Medium, High	<ol style="list-style-type: none"> Model Document <ul style="list-style-type: none"> Implementation testing plan Functional components test references Reference to test cases and implementation testing of feeder models Sign-off documents/mails for UAT/PROD testing/System migration testing <ul style="list-style-type: none"> Evidence of sign-off

4.4.6. Section 6: Identification, assessment and management of bias and risks to fairness

This section aims to analyse the broad management of bias and risks to fairness with a special analysis of the concept definition, measurement and mitigation and the motoring of bias risks in data and design used for the algorithm.

The following control objectives are audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization defines the risks to fairness and model bias with regard to business objectives and defines appropriate quantitative metrics that are being monitored and whose objectives are set by business owners. The organization implements risk mitigation approaches.	Medium, High	1. Model document <ul style="list-style-type: none"> ▪ Identification of bias ▪ Monitoring bias ▪ Root cause analysis ▪ Mitigation of bias 2. Discussion with the stakeholders/supporting document memorandum with used case examples for bias
For each algorithm with increased risk for unfair model outcome and bias, the organization explores and identifies the presence of bias in the training data (data bias).	Medium, High	
For each algorithm with increased risk for unfair model outcome and bias, the organization explores and identifies the presence of bias in the algorithm (design bias).	Medium, High	

4.4.7. Section 7: Explainability & Interpretability of the model and of its results

This section focuses on the actions and documentation in place to ensure a good understanding of the use case functioning, at different level of the organization's hierarchy.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization clearly articulates the model explainability in function of the business use. In order to manage bias/fairness challenges, it is important to understand how the model reached its outcome. This needs to be both at an aggregate level and possibly at the level of individual outcomes. In function of the business application, such analysis may need to be available in real time for the data subject.	High	1. Model Document/Mail Records <ul style="list-style-type: none"> ▪ Need for explainability based on ▪ Requirement for disclosure of information ▪ Approach for explainability and robustness of the same 2. Feedback from stakeholders

4.4.8. Section 8: Sensitivity analysis of the model to its features

This section assesses methods and tools used to tests the sensitivity of the model when designing it. This analysis is needed to assure the right configuration and evaluation of the applicability's perimeter of the algorithm.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
<p>The organization performs a sensitivity analysis of the model.</p> <p>Statistical models often have a multitude of parameters, that can be fixed with varying degree of precision (e.g., based on external data, on the basis of expert judgement, ...). This includes so-called hyperparameters. The purpose of sensitivity analysis is to have a good insight in which parameters have the most influence on model outcome and model performance and understand how their values have been set. The sensitivity analysis should also support the definition of the domain of applicability of the algorithm, i.e. when does the model work, when not</p>	High	<ol style="list-style-type: none"> Model Document (supporting files) <ul style="list-style-type: none"> Model parameters inventory Sensitivity analysis results Considerations for hyperparameters and non-linear behaviour of the model Model Monitoring report <ul style="list-style-type: none"> Controls to counter instability of the model

4.4.9. Section 9: Hypothesis Testing when designing the algorithm

This section aims to analyse more deeply the design phase of the algorithm and to question design choices made by data scientists (model, features, algorithm, approach, etc.).

The following control objectives are audited:

Control objectives	Use case risk level	Example of supporting evidence
<p>During the model development process, different modelling choices are explored. The organization justifies and documents the options explored and the final approach retained.</p>	High	<ol style="list-style-type: none"> Model document (supporting files) <ul style="list-style-type: none"> Relevance of hypothesis testing Coverage of different hypothetical scenarios Details of the hypothesis testing (completeness, adequacy etc.)
<p>In case of a non-conformance identified on the control objective N° 14, the auditors could perform independant testing with a challenger.</p> <p>The implementation of a challenger aims to explore several aspects of the model development process: choice of features, choices of algorithm, choice of implementation approach and packages.</p>	High	<ol style="list-style-type: none"> Model Document <ul style="list-style-type: none"> Challenger model (development, performance, approach etc.)

4.4.10. Section 10: Performance Testing

This section aims to analyse actions put in place to assess the model performance under the “real” market conditions.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
<p>The organization assesses the performance of the model under different usage conditions and regarding all the components of the model.</p>	High	<ol style="list-style-type: none"> Model document/supporting files <ul style="list-style-type: none"> Assessment of model performance under different market scenarios Various components of the model and their performance Completeness of the model based on performance criteria

4.4.11. Section 11: Impact Assessment of known weaknesses and open improvement points

This section aims to analyse measures in place to maintain an overview of weaknesses and improvement points.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization maintains for each algorithm an overview of known weaknesses and open improvement points.	High	<ol style="list-style-type: none"> 1. Model document <ul style="list-style-type: none"> ▪ Overview of model limitations ▪ Impact of model limitations on the model results ▪ Mitigating actions for limitations 2. Mails/Minutes of meetings <ul style="list-style-type: none"> ▪ Issues identified during the audit process

4.4.12. Section 12: Compliance review to the regulatory and legal requirements

This final section focuses on maintaining and ensuring compliance of the model to the regulatory and legal requirements.

The following control objective is audited:

Control objectives	Use case risk level	Example of supporting evidence
The organization ensures the compliance of the model to the regulatory and legal requirements.	High	<ol style="list-style-type: none"> 1. Model document/Sign-off mails from compliance team <ul style="list-style-type: none"> ▪ Overview of the relevant regulatory/ legal requirements ▪ Compliance of the AI system

A woman with dark hair and glasses is seen from the chest up, holding a red marker and drawing a red line graph on a glass surface. She is wearing a green top and a yellow and white striped cardigan. The background is a blurred office setting with windows. On the left side of the image, there are several overlapping pink geometric shapes: a square at the top, a triangle below it, and another triangle below that. A large white number '5' is centered on the page, with a small yellow square to its right.

5 ■

**Commitments
of the labelled
Entities**

5. Commitments of the labelled Entities

5.1. General commitments

5.2. Use of the trademark



6

Appeals and complaints handling

6. Appeals and complaints handling

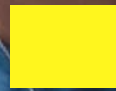
6.1. Positive AI complaints mechanism

6.2. Appeal procedure

6.3. Claims



7



Appendices



7. Appendices

7.1. List of acronyms and glossary

Label Level	Audit types
Artificial Intelligence (AI) System	For the Positive AI Label, the definition of AI system mentioned in the IA Act applies. As of 25 November 2022, the following definition in the Article 3 of the AI Act applies: “Artificial Intelligence system (AI system) means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts”.
Conformance	The Entity’s policies, procedures and processes related to AI governance and systems perform in a manner that is conformant with the Positive AI Label framework.
Entity	An organization for implementation of the Positive AI Label framework.
Improvement action plan	Formal document that will outline the improvement areas and remediation action to respond to each issued non-conformance during the Labelling Audit, with associated responsibilities and timeframes. This document must be shared to the Auditor.
Labelling Agreement	Agreement that defines rights and undertakings of each party in respect of the use of the Positive AI Label
Labelling Audit	The Labelling Audit provides the independent, third-party verification that Responsible IA policies and systems have been implemented to meet the Positive AI framework’s control objectives.
Labelled Entity	An Entity that is Member of Positive AI and conforms with the Positive AI Label framework, as verified by a Positive AI Referenced Auditor.
Labelling Scope	The Labelling Scope is defined by the Member/Entity seeking the Labelling agreement considering the rules established in the section §3.2.
Member	<p>Any organization that fills out a membership form, undertakes the payment of the annual Positive AI commercial membership and for which the application is accepted and validated by Positive AI, becomes a Positive AI Member.</p> <p>Any organization that wishes to apply for a Labelling Audit by a Referenced Auditor against the Positive AI Label framework must be a Positive AI member. It is not mandatory for a Positive AI Member to apply for a Labelling Audit.</p>
Non-Conformance	<p>There are two types of non-conformances that could be noted by a Referenced Auditor:</p> <ul style="list-style-type: none"> Minor non-conformance: The Entity’s IA practices perform in a way that does not wholly conform to the relevant Positive IA Label framework, due to an isolated lapse of a control. Major non-conformance: The Entity’s IA practices perform in a way that does not conform to the relevant Positive IA Label framework, due to the total absence of implementation or evidence of a control.
Positive AI Label framework	The Positive AI Label framework is designed for use by Positive AI Referenced Auditors to verify an Entity’s conformance for the purposes of granting Positive AI Label.
Referenced Auditor	An independent, third-party external organization meeting the Positive AI’s objective selection criteria and referenced by Positive AI to carry out the Labelling and Surveillance Audits for the Positive AI Label.
Surveillance Audit	Following the level of labelling obtained by the Entity, the Referenced Auditor conducts one or two Surveillance Audits of the Labelled Entity to verify that policies and systems in place continue to work effectively and non-conformances found during the Labelling Audit have been addressed.

7.2. List of Positive AI Referenced Auditors

The Labelling Audit is carried out by one Positive AI Referenced Auditor as of today:

- EY France: [Page de contact | EY - France](#)

7.3. Sources

The Positive AI Label has drawn inspiration and insights from the following key documents and sources:

- Ethics guidelines for trustworthy AI
<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- AI Act: Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (AI Act) and amending certain Union legislative acts, on 29/11/2021
<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Livre blanc - Intelligence artificielle Une approche européenne axée sur l'excellence et la confiance
https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_fr.pdf
- AI Risk Management Framework: Initial Draft March 17, 2022, NIST
<https://www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf>

[Page de fin]