JUNE 2024

AI POLICY TEMPLATE

Build Your Foundational Organizational AI Policy

Responsible Artificial Intelligence Institute

AI Policy Template Disclaimer

*This Template [**AI Policy**] includes various provisions throughout that must be reviewed and, potentially, revised based on the specifics of your business and your use of artificial intelligence technologies. You are advised to confirm that all pre-populated information is accurate and appropriate for your business.*

This Policy Template was developed to reflect sound practices for responsible AI management at the time it was created. The Policy is not intended to, and does not purport to, satisfy particular legal requirements in every jurisdiction. In all cases, you are advised to consult an attorney for guidance on the laws and regulations applicable to your business and to determine whether this Policy is adequate to address such legal requirements.

Artificial intelligence is a rapidly evolving field, and the best practices for responsible AI management are likewise evolving. We recommend revisiting your policy regularly (at least annually) and updating it as needed to reflect current standards or legal requirements.

You use this Policy Template at its own risk. This Policy Template does not constitute legal advice, and by using all or any part of this Policy Template, you agree to this disclaimer. You are advised to (i) consult independent legal advice before adopting or publishing your policy; (ii) read this Policy with care and modify, delete or add all and any areas as necessary; and (iii) not rely on this Policy for any purpose without seeking legal advice from a licensed attorney in your jurisdiction. This Policy Template is provided only for informational purposes and may or may not reflect the most current legal developments; accordingly, it is not promised or guaranteed to be correct or complete.

Table of Contents

Introduction to the AI Policy Template 4

I. Purpose and Scope 5

II. AI Principles 5

III. AI Objectives and Strategy 7

IV. Governance 9

V. Data Management 13

VI. Risk Management 17

VII. Project Management 27

VIII. Stakeholder Management and Engagement 33

IX. Workforce Management 35

X. Regulatory Compliance 37

XI. AI Procurement 38

XII. Documentation Management 41

XIII. Review and Enforcement of the AI Policy 43

XIV. Conclusion/Acknowledgement 44

Appendix A 45

Introduction to the AI Policy Template

As artificial intelligence (AI) unlocks opportunities for businesses to increase efficiency and generate novel insights and offerings, organizations seek to build responsible AI practices that align with ethical principles, mitigate potential risks, and foster trust with stakeholders. For many, an organizational-level policy can be a foundational document to establish guiding principles, objectives, and management direction for all AI-related activities according to business requirements. A standalone policy for AI can also centralize and highlight an organization's responsible AI strategy to accelerate internal adoption and external awareness. Developing an AI Policy is a requirement of an AI management system under leading standard ISO/IEC 42001, and it is also a main policy recommendation under the Responsible AI Institute's framework for organizational maturity.

This document serves as a template for an AI Policy, which can be used to establish a comprehensive framework for an organization's development, procurement, supply, and use of AI technologies. The template provides one interpretation of how global and regional guidance, including the NIST AI Risk Management Framework (RMF), and ISO/IEC 42001, can be operationalized through corporate policy, informed by RAI Institute's deep expertise in accelerating member organizations' maturity for AI. Organizations, depending on their size, existing policies, and industry requirements, may find that some schemes detailed in the template may need to be adapted or substituted to better fit their context.

Organizations are encouraged to customize this Policy Template to address the specific needs and risks of their AI use cases. In alignment with ISO/IEC 42001, organizations are recommended to use a suite of factors to inform their AI Policy, including but not limited to business strategy; organizational values and culture; organizational risk environment and tolerance; statutory, regulatory and contractual requirements; and possible AI risks and impacts of its use cases.[1]

The following template is RAI Institute's first draft version. The Institute will be accepting feedback into July 2024 for an updated version, to be released in the following months.

[1] Adapted from ISO/IEC 42001 B.2.2.

I. Purpose and Scope

A. At *[organization name]* _____, we recognize the transformative potential of artificial intelligence (AI) to enhance our operations, products, and services. This Policy outlines our commitment to responsible AI *[e.g., development, deployment, supply, use]* _____ to ensure ethical considerations are upheld, AI risks are managed, and compliance to emerging regulation is achieved.

B. The following *[policy document name e.g. "AI Policy"]* _____, provides a framework to guide all activities at *[organization name]* _____ related to AI. This Policy applies to all *[e.g., employees, contractors, and third-party suppliers]*_____ involved in *[all relevant AI activities e.g. "the development, procurement, and use of AI"]*_____ within *[organization name]* _____.[2]

C. We **define AI as** *[e.g., "an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy"[3]]* _____.

1. We **define an AI model as** *[e.g., "a component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs"[4]]* _____.
2. We **define an AI system as** *[e.g., "any data system, software, hardware, application, tool, or utility that operates in whole or in part using AI"[5]]* _____.
3. We have developed these definitions in alignment with leading industry and ecosystems definitions, including *[e.g., the OECD, EU AI Act, U.S. EO]* _____, to define the bounds of AI life cycle management and inventorying and of compliance requirements, as appropriate for our context and use cases.

II. AI Principles

A. In alignment with the established enterprise values, including *[e.g., sustainability, diversity, equity, and inclusion, corporate social responsibility]* _____, we are committed to upholding ethical principles in the *[AI activities e.g., procurement, development, deployment, supply, and/or use]* _____ of AI technologies.[6]

---

[2] Aligned with ISO/IEC 42001 4.3.
[3] Definition of AI provided by the National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF). Adapted from: OECD Recommendation on AI:2019; ISO/IEC 22989:2022.
[4] Definition of AI model provided by United States Executive Order No. 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.
[5] Definition of AI system provided by United States Executive Order No. 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.
[6] Aligned with ISO/IEC 42001 A.2.2.

B. Our AI principles encompass the following[7]:

1. **System Trustworthiness:** Every AI system that is *[bought, built, used, or sold]* _____ by *[organization name]*_____ shall strive to achieve appropriate levels of all trustworthy characteristics, as defined by the National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF).[8] The achievement of each of these characteristics will depend on the use case and may require tradeoffs between characteristics, which will be justified and documented.

    a) Validity and Reliability: All AI systems shall consistently provide accurate outputs or otherwise behave within a defined range of acceptability when subject to expected conditions of use.

    b) Safety: No AI system shall endanger human life, health, property, or the environment.

    c) Security and Resiliency: All AI systems shall withstand unexpected adverse events or unexpected changes in their environment or use, maintaining confidentiality, integrity, and availability in the event of adversarial or unauthorized actions.

    d) Accountability and Transparency: Meaningful and timely information about every AI system shall be provided to all relevant stakeholders, tailored to the expected knowledge and accessibility needs of each audience. An accountability structure governs each decision made related to an AI system.

    e) Explainability and Interpretability: All AI systems shall be designed and documented to answer how and why a decision was made by the system, to the fullest extent possible.

    f) Privacy-Enhanced: All AI systems shall safeguard human autonomy, identity, and dignity with respect to privacy to the fullest possible extent.

    g) Fairness with Harmful Bias Managed: All AI systems shall meet a defined metric of fairness appropriate for its context and shall manage all forms of harmful bias, including system bias, computational and statistical bias, and human-cognitive bias.[9]

2. **Human Oversight and Accountability:** We commit to ensuring that all processes and materials related to AI in our organization are subject to proper oversight mechanisms to enable responsible development and use of AI. We also embrace all sources of external accountability, including seeking independent audits and

[7] Aligned with ISO/IEC 42001 B.6.1.2.
[8] The following statements by characteristic are aligned with the definitions of each characteristic provided in the NIST AI RMF. Additionally aligned with GOVERN 1.2.
[9] NIST has identified these three as major categories of AI bias to be considered and managed. Each of these can occur in the absence of prejudice, partiality, or discriminatory intent.

certifications, monitoring by governmental organizations, and meaningful transparency with interested public parties.

3. **Beneficence, Equity, and Ethics:** We shall align our AI strategy with the broader interest of our organization to preserve and promote societal well-being. This includes our commitments to ethical frameworks related to *[e.g., sustainability, equity, human rights]* _____.
4. **Continual Learning:** As the technological and regulatory environment of AI rapidly develops, we are committed to a culture of open-mindedness, flexibility, and dialogue. We shall engage with partners, peers, stakeholders, and the public to invest in shared knowledge and a shared vision for responsible AI.

III. AI Objectives and Strategy

A. As a *[entity in a context e.g. "leading company in the financial services industry in the United States"]* _____, AI has the power to *[benefits]* _____, adding significant enterprise value and bolstering our competitive advantage. We will *[build, buy, and/or sell]* _____ AI *[components, systems, and/or applications]* _____, and aim to apply AI in *[list of functions or use cases]* _____.[10]

B. However, we also foresee challenges to responsible *[procurement, development, deployment, supply, and/or use]* _____ of AI, given the context of our organization and of our use cases. This includes *[examples of regulatory requirements, areas of deficient organizational capacity or governance, industry-specific data challenges]* _____.[11]

C. With these potential benefits and challenges in mind, a long-term AI business strategy shall be developed, documented, and implemented. The **key objectives** that shall guide this strategy are as follows[12]:

1. *[Key objective with an explanation of how progress along it will be measured]*
2. *[Key objective with an explanation of how progress along it will be measured]*
3. *[Key objective with an explanation of how progress along it will be measured]*
4. This strategy shall align with *[organization name]* _____'s broader objectives and commitments, including *[e.g., achieving net-zero by 2030, investing in leadership by historically marginalized groups within and beyond the organization]* _____.

[10] Aligned with NIST AI RMF MAP 1.3 and ISO/IEC 42001.
[11] Aligned with ISO/IEC 42001 4.1.
[12] Aligned with NIST AI RMF MAP 1.3 and ISO/IEC 42001 6.2 and B.6.1.2.

D. Recognizing the need for bespoke, flexible, and unobtrusive organizational adaptation for AI, the contents of this AI Policy shall be used as an augmenting layer on top of *[organization name]* _____'s existing governance, policies, and processes.[13]

E. While *[organization name]* _____'s AI strategy will evolve into a complete practice over time, the urgency of AI risks requires that *[organization name]* _____ identify areas of highest priority as a *[developer, procurer, and/or supplier]* _____ of AI:

1. **For bought systems[14]:**
   a) Responsibly buying AI will require significant investment in the following capabilities:
      (1) A rigorous and principles-driven procurement process that sufficiently weighs marketplace options and comprehensively assesses the risks attached to potential suppliers and their product or service;
      (2) A legal means to clarify liability and other requirements with suppliers, to whatever extent is not articulated by regulation;
      (3) Role- and system-specific training and educational materials for employees such that procured AI can be used safely and responsibly; and
      (4) *[additional areas of strategic focus]* _____.

2. **For built systems:**
   a) Responsibly building AI will require significant investment in the following capabilities:
      (1) A product management program that balances the forward inertia of innovation with the necessary governance gates and other processes, such as AI Impact Assessments, to sufficiently manage risk;
      (2) Processes and tools to enable systematic and comprehensive documentation of AI systems throughout their life cycle, in accordance with regulatory requirements;
      (3) Responsible AI training for all technical and nontechnical roles involved in an AI system's life cycle across design, development, deployment, and operation; and
      (4) *[additional areas of strategic focus]* _____.

3. **For sold systems[15]:**
   a) Responsibly selling AI will require significant investment in the following capabilities:
      (1) Development of audience-specific guidance and documentation for each AI system, such as for potential buyers, users, or the general public;
      (2) A legal means to clarify liability and other requirements with buyers, to whatever extent is not articulated by regulation;

[13] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[14] Aligned with ISO/IEC 42001 A.10.3.
[15] Aligned with ISO/IEC 42001 A.10.4.

(3) Regular assessment of downstream impacts of AI products, enabled by transparency and shared learning agreements with buyers while also protecting users' privacy and other rights; and
(4) *[additional areas of strategic focus]* _____.

F. *[organization name]* _____ maintains a commitment to continually improve the suitability, adequacy, and effectiveness of this AI Policy and of its AI management system.[16]

IV. Governance

A. The following executive or senior management positions are designated as the *[owners, champions, and/or sponsors]* _____ of *[organization name]* _____'s responsible AI approach. Executive *[owners, champions, and/or sponsors]* _____ bear responsibility for ensuring that *[organization name]* _____'s responsible AI strategy is developed and executed effectively and are ultimately accountable for its success.[17]

1. The *[position title e.g., CIO, CTO, COO]* _____ is accountable for *[responsibilities and key results e.g., leading the Steering Committee, tracking, reporting, and owning overall progress]* _____.
2. The *[position title e.g., VP of Legal, VP of Compliance]* _____ is accountable for *[responsibilities and key results e.g. ensuring that AI systems do not violate legal or regulatory requirements]* _____.
3. *[Additional positions and fields of accountability]* _____.

B. *[organization name]* _____'s responsible AI strategy shall be led by two major **governance bodies**[18]:

1. A High-Level Board/Steering Committee shall provide executive leadership and oversight, including direction, mandates, and resourcing for responsible AI efforts, in a timely manner.[19] The Steering Committee has representation across senior-level management, including *[e.g., CIO, CDO, Board members, Chief AI Officer]* _____.
2. An Operational Committee shall direct the implementation of Steering Committee objectives, oversee the life cycle progression and impact of AI systems, and act as a

[16] Aligned with ISO/IEC 42001 10.1. An organization's AI management system, in accordance with the scope of ISO/IEC 42001, refers to the collective body of management structures and processes established for an organization to responsibly perform their role with respect to AI systems (e.g. to use, develop, monitor or provide products or services that utilize AI). This AI Policy has been designed as a top-level framework for an organization's AI management system.
[17] Aligned with NIST AI RMF GOVERN 2.3.
[18] ISO/IEC 42001 3.22 notes that "not all organizations, particularly small organizations, will have a governing body separate from top management." The creation of separate steering and operational groups may not be necessary.
[19] ISO/IEC 42001 5.1 provides a list of activities for top management to demonstrate leadership and commitment in this capacity.

convening power for responsible AI efforts. The Operational Committee has representation across functions and/or divisions of *[organization name]* _____, including *[e.g., Head of Information Security, Head of Procurement, Legal, Head of HR]* _____.

C. The Operational Committee[20] shall convene every *[e.g. two weeks]* _____. Its responsibilities include:
    1. Approving progression of an AI system's development to the next stage;
    2. Directing the development and updating of both cross-functional and function-specific AI guidance and tools alongside function or division leaders;
    3. Developing and managing AI-related inventories, such as of AI systems and AI incidents;
    4. Developing responsible AI training;
    5. Determining and implementing changes to the AI management system in a planned manner[21]; and
    6. *[additional responsibilities]* _____.
D. The Steering Committee shall convene every *[e.g. two months]* _____. Its responsibilities include[22]:
    1. Developing an organizational-level AI strategy with a clear timeline and measures of success (KPIs);
    2. Articulating how its AI strategy amplifies or creates trade-offs with other organizational objectives or commitments (highest-level SWOT or ROI analyses);
    3. Determining organizational AI risk tolerances;
    4. Determining and allocating the resources needed for the establishment, implementation, maintenance, and continual improvement of the AI management system[23];
    5. Aligning workforce planning with responsible AI human capital needs; and
    6. *[additional responsibilities]* _____.
E. Additional individual roles shall be created to direct and support *[organization name]* _____'s AI strategy, including *[e.g., Chief AI Officer, Responsible AI Chair]* _____ which shall be responsible for *[description of responsibilities]* _____.

[20] The Operational Committee is often named differently in practice, such as simply "Responsible AI (RAI) Team."
[21] Aligned with ISO/IEC 42001 6.3.
[22] See ISO/IEC 42001 5.1 for further guidance on responsibilities of top management.
[23] Aligned with ISO/IEC 42001 7.1.

F. More specific areas related to AI, such as *[e.g., cybersecurity, supplier relationships, and data management]* _____, can be managed by specific *[e.g. functions]* _____, in collaboration with or with broad oversight by the Operational Committee.[24]

   1. Leaders of *[organization name]* _____'s *[functions, segments, and/or divisions]* _____ shall direct the development of AI-specific processes and tools tailored for their operations, informed by Operational Committee directives, similar efforts by peer *[functions, segments, and/or divisions]* _____, and consultations with domain experts and internal or external stakeholders.

G. *[organization name]* _____ shall additionally determine **roles, responsibilities, and accountability** of internal AI actors with respect to AI systems and the systems' external AI actors.[25] The position of each role in a chain of accountability shall be made clear.

   1. **Internal AI actors** include all employees who contribute to AI design, development, deployment, operation and monitoring; TEVV (test, evaluation, verification, and validation) tasks; risk management and impact assessment tasks[26]; procurement tasks; governance and oversight tasks; and *[additional task areas]* _____.

      a) This can include developers; data scientists; data engineers; product managers; system integrators; system operators; domain experts; socio-cultural analysts and experts (e.g. DEI, accessibility, governance); human factors experts (e.g., UX/UI design); procurers; AI governance and oversight professionals; legal and privacy officers; and *[additional internal AI actor profiles]* _____.

   2. Internal AI actors shall be assigned responsibilities to appropriately manage relationships with **external AI actors**, which include all individuals, groups, and society members that are involved in AI systems' life cycle or may be impacted by the system.

      a) This can include suppliers and partners (of data or AI platforms, products, or services); third-party assessors or evaluators (of data, algorithms, models, and/or systems); clients; data subjects; end users; members of impacted communities; and *[additional external AI actor profiles]* _____.

H. **Communication and feedback channels** shall be augmented to enable proper information sharing and issue management between different AI actors.[27] These channels shall facilitate various processes percolating upwards or downwards *[organization name]* _____'s structure, including[28]:

   1. A mechanism for employees to report their concerns about any of *[organization name]* _____'s AI activities or with respect to an AI system throughout its life

[24] ISO/IEC 42001 B.3.2 provides examples of areas where roles and responsibilities can be defined.
[25] Adapted from NIST AI RMF Appendix A.
[26] Aligned with NIST AI RMF GOVERN 2.1.
[27] Aligned with ISO/IEC 42001 7.4.
[28] Aligned with NIST AI RMF GOVERN 2.1.

cycle.[29] Such a mechanism shall *[e.g., protect individuals from identification and reprisals; be accessible to all workforce members; have appropriate personnel and capabilities (including investigation, resolution, escalation, and reporting powers); respond and act in a timely manner[30]]* _____;

2. Readily accessible means for employees to contact a human representative within the organization for support and guidance on AI-related inquiries, including *[e.g., on where to find responsible AI upskilling material, to identify executive RAI champions]* _____;
3. Formalized and regular performance and progress reporting between organization *[functions, segments, or divisions]* _____ to AI governance bodies to top management;[31] and
4. A formalized and regular process to proactively solicit feedback from internal AI actors on the suitability, adequacy, and effectiveness of *[organization name]* _____'s AI management system.

I. Governance gates shall be established to facilitate the standardized approval of AI systems' life cycle progression and the collection of documentation at each life cycle stage. Governance gates shall accomplish the following[32]:

1. Be performed and managed by *[structure e.g. "the Operational Committee with high-level oversight by the Steering Committee"]* _____ through *[process e.g. "asynchronous and digital voting, bimonthly meetings to confirm approval decisions, and triggers and mechanisms for escalation of decisions"]* _____;
2. Pause or terminate an AI proof of concept (PoC) or project once a risk has been identified that measures beyond *[organization name]* _____'s risk tolerance (see Risk Management Section G);
    a) Approval requirements are tied to measures of AI risk criteria and enable exercise of AI risk triage processes in alignment with risk priorities and tolerance.[33]
    b) Approval requirements are tied to continued alignment with organizational RAI principles and objectives, determined via established measures.
        (1) Approval is also dependent on demonstrated achievement or projected achievement of the system's intended purpose and stated objectives.[34]
        (2) Approval is also dependent on the completion of all required responsible AI tasks at the current life cycle stage, including the documentation of all relevant decisions made and AI actors involved.

[29] Aligned with ISO/IEC 42001 B.3.3.
[30] Aligned with ISO/IEC 42001 B.3.3.
[31] Aligned with ISO/IEC 42001 5.3.
[32] Aligned with ISO/IEC 42001 B.6.1.3.
[33] Aligned with ISO/IEC 42001 6.1.1.
[34] Aligned with NIST AI RMF MANAGE 1.1.

c) Depending on the catalyst for pause/termination, projects may proceed at a later date once concerns have been remediated, or they may need to be resubmitted at the beginning of the life cycle pipeline as a new PoC.

3. Reprioritize an AI project in relation to its counterparts depending on the risk level and the human and technological resource requirements identified, in the context of overall AI objectives, as necessary;
4. Utilize AI Impact Assessments (see Risk Management Section I) as a tool to inform approval decisions and set conditions for progression, such as *[e.g. specific design or use requirements]* _____,[35] performed by *[details on roles e.g. "mainly developer and procurement teams with support of and consultation with compliance, external stakeholders, etc."]* _____[36] at the appropriate level of comprehensiveness and at the appropriate life cycle stages; and
5. Be responsive to risk triggers, both established and unexpected, and outline a review or recourse process to manage the cause of the trigger. Risk triggers include *[e.g., feedback or AI performance requiring immediate attention, switching vendors, changes to applicable regulation, reports of a serious incident from a similar use case]* _____.

V. Data Management

A. Existing data management at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing data management framework wherever absent.[37]

B. Additionally, *[organization name]* _____ identifies and aligns with existing regulations and guidelines for data management and reporting, including *[e.g., appropriate use and disclosure of IP-protected content, proper and industry-specific consent mechanisms for data subjects]* _____.[38]

C. *[organization name]* _____ shall be committed to meaningful **data transparency**, characterized by the ongoing delivery and testing of concise and audience-specific data information and mechanisms for recourse, informed by engagement with all interested parties, including *[e.g., users, impacted groups, third-party auditors, researchers]* _____.

D. During the early stages of a project, teams shall identify, characterize, and justify the type and quantity of data needed for an AI system.

[35] Aligned with ISO/IEC 42001 B.5.2.
[36] Aligned with ISO/IEC 42001 B.5.2.
[37] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[38] Aligned with NIST AI RMF 1.2.2 and GOVERN 1.1.

E. Rigorous exploratory data analysis shall be conducted for each potential data set to gain insights into underlying structure and statistical properties and to assess alignment with data fit-for-purpose and quality standards.

F. General information about every selected and utilized data set shall be logged in an **enterprise-wide data inventory system**, including *[e.g., in which projects it is used, internal cross-project usage permissions, data source and contact information, data type and features]* _____.[39]

G. Systems and standards shall be developed by *[e.g. functions such as Information Security]* _____ to manage the secure and organized storage of data sets enterprise-wide.

H. For all data sets from third-party sources, project teams shall follow a standard protocol set forth by the procurement team to communicate procurement needs during project scoping and to allow the procurement team to manage the procurement process.

I. For each data set used in a project, the project team shall document as part of project-level documentation[40]:

    1. Data source[41]: internal or external, including how data is acquired and/or accessed from the source;

        a) If internal, how the data is created and contact information of the data owner/manager;

        b) If external, from which supplier and their contact information, and how the data is created, if known;

    2. Data types and features[42]: whether the data is directly observable, collected from subjects and their demographics/characteristics, and/or indirectly inferred/derived from other data; includes sensitive personal identifiable information (PII); associated metadata; structured or unstructured; continuous or discrete; or multimodal, time series, etc.;

    3. Data consent process and results: any legal requirements for consent in the use case; whether data subjects provided free, prior, and informed consent; how data can be excluded/removed in response to retracted consent;

    4. Data provenance process and results[43]: information about the creation, update, transcription, abstraction, validation and transferring of the control of data within; data sharing and data transformation;

[39] Aligned with ISO/IEC 42001 A.4.3. Documenting resources at the organizational level in addition to the project level enables more efficient tracking and handling, such as for estimating data storage needs and reusing data sets, when appropriate.
[40] Aligned with ISO/IEC 42001 B.4.3.
[41] Aligned with ISO/IEC 42001 A.7.3.
[42] Aligned with ISO/IEC 42001 B.7.3.
[43] Aligned with ISO/IEC 42001 A.7.5 and B.7.3.

5. Internal data access and usage: permissions and restrictions on the access (permission to view, edit, or share) and use of the data set, including whether it can be used in other projects based on collection conditions (with supplier, with respect to consent);
6. External data access and cybersecurity process and results[44]: terms of data sharing with external parties, such as model suppliers or researchers, and measures to prevent unauthorized access to or control over the data;
7. Data privacy process and results[45]: the required level of privacy based on the use case and privacy-enhancing technologies and techniques applied; sensitive data or PII anonymization, etc.;
8. Data proxy risks and fairness: identifying potential proxies for protected class characteristics, and mitigating bias risks to align with fairness standards;
9. Data collection and preparation process and results[46]: the collection, preparation, and data transformation methods used and a justification for each based on intended use(s) and system context, and including categories of data for machine learning (e.g. training, validation, test and production data);
10. Data quality process and results[47]: assessment of data set quality along defined metrics, such as fit-for-purpose, representativeness, completeness, consistency, and accuracy, and any methods used to increase data quality and resulting measured improvements;
11. Data retention and disposal: protocols for retaining or disposing of data used for training, operating, and maintaining the AI system in a timely manner, if not prohibited by regulations and other transparency requirements;
12. Data drift and versioning: a process to ensure data do not become outdated as the culture and context around the system change, tracked by a data set version control system.

J. For bought systems[48]:

1. In partnership with the procurement team, project teams shall request information on all data used to develop the system from the supplier. Based on the data documentation requirements listed in Section I, suppliers of systems shall be required to disclose *[e.g., data types and features, sources, provenance, consent policy, privacy policy, collection and preparation process, quality assessment, disposal policy]* _____, to advance in the procurement process.

[44] Aligned with ISO/IEC 42001 B.7.2.
[45] Aligned with ISO/IEC 42001 B.7.2.
[46] Aligned with ISO/IEC 42001 B.7.3 and B.7.6, which lists common methods.
[47] Aligned with ISO/IEC 42001 A.7.4 and B.7.2 and NIST AI RMF MAP 2.3.
[48] Aligned with ISO/IEC 42001 B.10.3.

2. Suppliers shall be required to disclose additional details related to data used to train, validate, and test the system, including *[e.g., IP compliance, consent and privacy processes specific to use in training]* _____, to advance in the procurement process.
3. In partnership with the procurement team, project teams shall clarify and establish data control and sharing requirements with the supplier, including *[e.g., what data received or created by the system will be shared with the supplier, what enterprise (buyer) data will be visible to the supplier, what data related or resulting from use of the system will be retained by the supplier and/or used to train the system, what data related or resulting from use of the system will be controlled by the enterprise (buyer)]* _____.
4. Project teams shall document all additional data used to support the operation of the system, including *[e.g., enterprise reference databases, data sets for fine-tuning or prompt engineering, customer data]* _____, and shall document details on how such data is collected, utilized, stored, and/or disposed.
5. Project teams shall document post-deployment and/or post-decommission data protocols to manage data generated through the operation of the system, including *[e.g., input and output logs, event and incident logs, feedback from users]* _____, and shall document details on how such data is collected, utilized, stored, and/or disposed.

K. For built systems:

1. If any components (e.g., data sets, AI models, platforms) of the built system are procured, project teams shall request information on all data used to develop them, in partnership with the procurement team. Based on the data documentation requirements listed in Section I, suppliers shall be required to disclose *[e.g., data types and features, sources, provenance, consent policy, privacy policy, collection and preparation process, quality assessment, disposal policy]* _____, to advance in the procurement process.
2. Project teams shall document how data sets, regardless of origin, are used in the development of the system. Teams shall disclose additional details related to the data used to train, validate, and test the system, including *[e.g., IP compliance, consent and privacy processes specific to use in training]* _____.
    a) If an AI model is procured, suppliers shall be required to disclose additional details related to data used to train, validate, and test the model, including *e.g., IP compliance, consent and privacy processes specific to use in training]* _____, to advance in the procurement process.
3. Project teams shall document how data sets are used to support the operation of the system, including *[e.g., enterprise reference databases, data sets for iterative*

retraining, fine-tuning, or prompt engineering, customer data] _____, and shall document details on how such data is collected, utilized, stored, and/or disposed.

a) If an AI model is procured, project teams shall, in partnership with the procurement team, clarify and establish data control and sharing requirements with the supplier, including *[e.g., what data received or created by a model will be shared with the supplier, what enterprise (buyer) data will be visible to the supplier, what data related or resulting from use of the model will be retained by the supplier and/or used to train the system, what data related or resulting from use of the model will be controlled by the enterprise (buyer)]* _____.

4. Project teams shall document post-deployment and/or post-decommission data protocols to manage data generated through the operation of the system, including *[e.g., input and output logs, event and incident logs, feedback from users]* _____, and shall document details on how such data is collected, utilized, stored, and/or disposed.

L. For sold systems[49]:

1. In partnership with compliance, legal, and sales teams, project teams shall determine and compile data documentation that is relevant, required, or requested by buyers of systems.

2. In partnership with compliance, legal, and sales teams, project teams shall clarify and establish data control and sharing requirements with the buyer, including *[e.g., what data received or created by the system will be shared with the enterprise (supplier), what buyer data will be visible to the enterprise (supplier), what data related or resulting from use of the model will be retained by the enterprise (supplier) and/or used to train the system, what data related or resulting from use of the model will be controlled by the buyer]* _____.

VI. Risk Management

A. Existing risk management at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing risk management framework wherever absent.[50]

B. *[Organization name]* _____ shall determine to what extent discipline-specific risk management processes sufficiently integrate AI considerations for those specific aspects, such as for *[e.g., information security, safety, or privacy]* _____.[51] AI-specific

[49] Aligned with ISO/IEC 42001 B.10.4.
[50] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[51] Aligned with ISO/IEC 42001 B.5.2.

risk management processes and tools shall be implemented once existing processes are exhausted.

C. Additionally, *[organization name]* _____ identifies and aligns with existing regulations and guidelines for risk management, including established documentation, reporting, and disclosure requirements and industry or use case-specific requirements, such as *[e.g. "reporting results of safety tests of high-risk models, as required by the U.S. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"]* _____.[52]

D. *[Organization name]* _____ **defines an AI impact as** *[e.g. "a negative effect experienced by the organization, individuals, groups of individuals, or societies as a result of the organization's development, use, or otherwise interaction with AI technologies"]* _____. This definition is developed to be operationalized through our AI Impact Assessment, and therefore *[e.g. "considers negative impacts exclusively but additionally considers impacts to the organization"]* _____.

   1. This definition was developed in alignment with leading industry and ecosystem definitions, including *[e.g., industry definitions of harm, the scope of AI impacts as detailed by regulators requiring an AI Impact Assessment]* _____.
   2. Using this definition, *[organization name]* _____ has also developed an **impact taxonomy** that enables the identification and categorization of specific impacts and assigns a severity score to each impact, based on a justified methodology of measuring magnitudes of harm.

E. *[Organization name]* _____ **defines an AI incident as** *[e.g. "an event precipitated by the organization's mismanagement of an AI system at any point of its life cycle which has the potential to lead to an AI impact"]* _____. This definition is developed to align with our definition and taxonomy for AI impacts, and therefore *[e.g. "does not make a distinction between 'incidents', 'accidents', 'hazards', and 'near-misses' nor limits 'incidents' to only unexpected events, unlike in other disciplines"]* _____.

F. *[Organization name]* _____ **defines an AI risk as** *[e.g. "the composite measure of an AI event's probability of occurring (likelihood) and the magnitude of the consequences of the corresponding AI impact"[53]]* _____. This definition is developed in alignment with leading industry and ecosystem definitions, including *[e.g. the NIST AI RMF]* _____.

   1. AI systems and individual AI risks identified for a system can be each assigned one of the following risk levels: *[e.g. "minimal (1), moderate (2), high (3), and very high (4)"]* _____. A system's risk level is calculated as the maximum risk level of all known AI risks for that system; individual AI risks are calculated using a risk matrix

[52] Aligned with NIST AI RMF 1.2.2 and GOVERN 1.1.
[53] Aligned with NIST AI RMF 1.1.

constructed from the two independently measured dimensions of likelihood and severity.[54]

G. At *[organization name]* _____, we recognize that tolerating risk is necessary to pursue innovation and remain competitive. As we balance the opportunities of AI with its risks, we define the following boundaries of *[organization name]* _____'s **risk tolerance**[55] with respect to AI:

   1. Use cases that carry unacceptable risks are prohibited without condition.[56] Unacceptable risks are aligned with leading and emerging requirements, such as *[e.g. the EU AI Act]* _____, and include *[e.g., the use of emotion recognition systems, manipulative behavior-change techniques, and social scoring algorithms]* _____, among others.

      a) Based on our own values and risk appetite, we additionally prohibit systems that introduce the following unacceptable risks: *[e.g., unfair and unexplainable promotion or termination decisions for employees based on an automated decision, use of a specific Large Language Model (LLM)]* _____.

   2. The following **risk thresholds** apply to all systems immediately prior to deployment:

      a) Systems that remain very high-risk or high-risk post-treatment shall not be allowed to deploy.[57]

      b) Systems deemed moderate-risk post-treatment shall be subject to more rigorous and frequent testing and monitoring requirements, including impact assessments, than those deemed minimal-risk.

H. Components of *[organization name]* _____'s risk management process for AI systems are as follows:

   1. Impact Identification with AI Impact Assessments
   2. Risk Measurement
   3. Risk Prioritization
   4. Risk Treatment
   5. Residual Risk
   6. Impact and Risk Tracking, Inventorying, and Transparency
   7. Impact Contingency Planning and Recourse
   8. Impact Reassessment

[54] An organization's approach to characterizing and calculating risk of an AI system can vary in structure and quantitative method. The provided approach is only one example. NIST AI RMF MEASURE 1.1 recommends that the most significant AI risks are determined, measured, and addressed first.
[55] Aligned with NIST AI RMF GOVERN 1.3 and MAP 1.5.
[56] Aligned with ISO/IEC 42001 6.1.1.
[57] Aligned with NIST AI RMF MEASURE 2.6.

I. Impact Identification with AI Impact Assessments: *[organization name]* _____'s **AI Impact Assessment (AIIA)** identifies potential impacts of an AI system through a snapshot evaluation of current and anticipated system development, deployment, and operation.[58]

1. Meant to guide teams to expand their impact considerations, improve their responsible AI practices throughout the system life cycle, and determine next steps for risk treatment, the AIIA can be used at different points in a system's life cycle at varying levels of specificity based on system maturity.[59]

    a) A Low-Touch AIIA is completed after the *[e.g. Plan and Design]* _____ stage with the appropriate stakeholders, including *[e.g., AI developers and representatives from the compliance team, procurement team, data engineering team, system user and/or operator groups, domain experts, socio-cultural analysts, potentially impacted individuals and groups of individuals, and responsible AI (RAI) Operational Committee]* _____.

    b) A Medium-Touch AIIA is completed during the *[e.g. Verify and Validate]* _____ stage once the system has reached reliable performance, with the appropriate stakeholders, including *[e.g., AI developers and representatives from the data science team, TEVV (Test, Evaluation, Verification, and Validation; e.g. red-teaming) team and experts, system user group, domain experts, socio-cultural analysts, potentially impacted individuals and groups of individuals, and responsible AI (RAI) Operational Committee]* _____. This AIIA can be skipped for systems designated minimal-risk.

    c) A High-Touch AIIA is completed immediately before deployment with the appropriate stakeholders, including *[e.g., AI developers and representatives from the compliance team, procurement team, MLOps team, TEVV (Test, Evaluation, Verification, and Validation; e.g. red-teaming) team and experts, system user group, domain experts, socio-cultural analysts, potentially impacted individuals and groups of individuals, and responsible AI (RAI) Operational Committee]* _____.

2. As a tool for impact identification, the AIIA identifies potential AI incidents and their sources and outcomes (i.e. impacts). The AIIA considers the impacts on[60]:

    a) Individuals and their *[e.g., legal position, life opportunities, physical or psychological well-being]* _____.

        (1) Relevant individuals include those *[e.g. "using the AI system (users) or whose PII are processed by the AI system (data subjects)"]* _____;

[58] Aligned with NIST AI RMF MEASURE 4.1.
[59] Aligned with ISO/IEC 42001 B.6.1.3. ISO/IEC 42001 B.5.2 lists potential circumstances under which an AI system impact assessment should be performed, should an organization desire an AI impact assessment process independent of life cycle stage.
[60] Aligned with ISO/IEC 42001 B.5.2 and B.5.4.

b) Groups of individuals and their *[e.g., legal position, life opportunities, physical or psychological well-being]* _____.

(1) Relevant groups of individuals include *[e.g., children, impaired persons, elderly persons, marginalized ethnic groups, and workers]*_____;

c) Societies, which can include impacts to[61]:

(1) Environment sustainability, including *[e.g., the impacts on natural resources and greenhouse gas emissions from compute, data storage, etc.]* _____[62];

(2) Economic opportunity and protections, including *[e.g., access to financial services, employment opportunities, taxes, trade and commerce]* _____;

(3) Government, including *[e.g., legislative processes, misinformation for political gain, national security, and criminal justice systems]* _____;

(4) Health and safety, including *[e.g., access to healthcare, medical diagnosis and treatment, and potential physical and psychological harms]* _____;

(5) Norms, traditions, culture and values, including *[e.g., human rights, accessibility rights, potential misinformation that leads to biases or harms to individuals]* _____; and

d) The financial and reputational health of *[organization name]* _____.[63]

3. The AI Impact Assessment shall:

a) Produce consistent, valid and comparable results across different systems and levels of specificity[64];

b) Assess the system according to trustworthiness characteristics, including validity and reliability; safety; security and resiliency; accountability and transparency; explainability and interpretability; privacy-enhanced; fairness with harmful bias managed[65];

c) Calibrate its assessment criteria based on various aspects of the system and its specific technical and societal context, including *[e.g., expected scope of use, the data used for the development of the AI system, the AI technologies used, the third-party components or services used,[66] the functionality of the overall system, and use case context, including deployment and operating environments[67]]* _____;

[61] Aligned with ISO/IEC 42001 B.5.5.
[62] Aligned with NIST AI RMF MEASURE 2.12.
[63] Aligned with ISO/IEC 42001 6.1.2.
[64] Aligned with ISO/IEC 42001 6.1.2.
[65] Aligned with ISO/IEC 42001 A.5.4 and NIST AI RMF 3 and MEASURE 2.5-2.11.
[66] Aligned with NIST AI RMF GOVERN 6.1 and MAP 4.1.
[67] Aligned with ISO/IEC 42001 6.1.4 and B.5.2. Also aligned with NIST AI RMF MEASURE 4.1.

d) Incorporate consultation insights and feedback on potential individual and societal impacts from external AI actors[68];

e) Continually improve its ability to identify potential AI impacts by tracking and applying historical and emerging data from *[e.g., past uses of AI systems in similar contexts, public incident reports, external stakeholder feedback, results from other impact assessments[69]]* _____; and

f) Be documented, have its results retained for a defined period, and be made available to all relevant and impacted audiences.[70]

4. The AI Impact Assessment shall be compatible with, but distinct from, other existing impact or risk assessments conducted at *[organization name]* _____, including *[e.g., financial impact assessments, security risk assessments, privacy risk assessments, business impact assessments]* _____.

a) Where domain-specific assessments do not exist, the AIIA shall be comprehensive and domain representatives shall be involved in the design and completion of the AIIA.

b) Otherwise, to reduce redundancy, the AIIA can focus on AI-driven risk sources[71] and AI-specific risks.[72] Means shall be established to ensure that the assessments collectively cover the entire known AI risk landscape and that all AI-related risks across assessments can be efficiently aggregated for each AI system.

J. Risk Measurement: Once potential impacts are identified through the AI Impact Assessment and assigned a severity score, AI risk is calculated through a likelihood analysis.

1. Likelihood and magnitude of each identified impact are determined in a standardized manner through an organizational risk and impact taxonomy delineated with common characteristics of AI systems and their use case context.[73]

a) Various factors can be used to determine the likelihood and magnitude of impacts,[74] including *[e.g., whether the AI system is trained on large data sets composed of sensitive or protected data such as personally identifiable information; whether it uses any third-party resources or components[75]; whether it is designed or deployed to directly interact with humans; whether its outputs have*

[68] Aligned with NIST AI RMF GOVERN 5.1, MEASURE 1.3, and MEASURE 4.1-4.2.
[69] Aligned with NIST AI RMF MAP 5.1.
[70] Aligned with ISO/IEC 42001 B.5.3.
[71] Examples can be found in ISO/IEC 42001 Annex C and ISO/IEC 42001 B.5.2.
[72] Examples can be found in NIST AI RMF Appendix B.
[73] Aligned with NIST AI RMF MAP 5.1.
[74] Aligned with NIST AI RMF 1.2.3.
[75] Aligned with NIST AI RMF GOVERN 6.1 and MAP 4.1.

direct impacts on humans; and how the organization itself measures risk based on its own values and context] _____.

b) *[organization name]* _____'s risk and impact taxonomy shall be reasonably aligned with applicable regulatory regimes, including *[e.g. the EU AI Act]* _____, to ease compliance efforts.

2. The risk and impact taxonomy is informed and continually improved through the tracking of historical and emerging data in an **AI Incident, Impact, and Risk (IIR) database**, including from sources like *[e.g., past uses of AI systems in similar contexts, public incident reports, external stakeholder feedback, results from other impact assessments[76]]* _____.

3. Risks or system trustworthiness characteristics that cannot be measured[77] and the specific reasons why[78] shall be properly documented.

K. Risk Prioritization: Recognizing that risk management of AI systems must also be cost-effective to bolster a competitive AI strategy, *[organization name]* _____ also establishes the following risk priorities[79] and risk triage process[80] to enable the efficient use of resources to treat risks[81]:

1. The order of risk treatment is decided first by a cost-benefit analysis at the system-level; systems with the highest upside potential and the lowest total cost of treatments that reduce the system risk level most effectively are placed at the top of the queue.[82]

a) Risk triage shall be performed at the level at which developer resources are self-contained, for example, at the level of a function, segment, or division of the organization.

b) Risk treatment of a system must be comprehensive; while a system's risk-level is determined by its highest individual risk, all known risks of a system must be appropriately treated in one round of effort.

c) The order of prioritization for AI system risk treatment shall be mainly automated using metrics as follows[83]:

- Highest net opportunity-to-treatment cost ratio
- Estimated risk residual level is within deployment threshold (minimal or moderate risk)

[76] Aligned with NIST AI RMF MAP 5.1.
[77] Aligned with NIST AI RMF MEASURE 1.1.
[78] NIST AI RMF 1.2.1 lists potential challenges to accurate measurement of risks.
[79] Aligned with NIST AI RMF GOVERN 1.4.
[80] Aligned with NIST AI RMF GOVERN 1.4.
[81] Aligned with NIST AI RMF 1.2.3.
[82] Aligned with NIST AI RMF MANAGE 1.2.
[83] Aligned with NIST AI RMF MANAGE 1.2.

- Highest net opportunity/benefit-to-risk residual and monitoring cost ratio
- Absolute highest opportunity/benefit measure
- Absolute lowest risk residual level

L. Risk Treatment: Risk treatment of a system shall consist of an appropriate and justified response to each identified risk and a system-level estimate of the resources needed to complete all risk responses collectively.[84]

1. Responses to a specific identified risk shall first consider possible options for avoidance of or safeguarding from the risk. Avoidance techniques intervene to reduce the likelihood that an AI incident occurs in the first place, while safeguarding techniques attempt to insulate individuals or society from harm after an AI incident has already occurred.[85]
    a) Other risk response options can include *[e.g. "transferring the risk to another entity, such as a downstream enterprise, or accepting the risk if no other treatment action is possible"]* _____.
2. Risk treatment shall aim to address weaknesses in a system's trustworthiness characteristics in a holistic, complete manner; a system is only as trustworthy as its weakest characteristic.[86]
    a) Progress toward any one characteristic may be dependent on or even in conflict with progress toward others. A risk treatment strategy shall balance the tradeoffs among trustworthiness characteristics for a system.
3. Teams shall estimate (and report after the fact) the resources needed for the risk treatment of a system, including *[e.g. tooling, time, and human resources]* _____ as part of the broader risk prioritization and resource allocation strategy.[87]
    a) Cognizant of existing workflows and how to efficiently allocate resources, teams shall also be afforded the agency to prioritize risk treatment efforts across their own projects, as long as this is aligned with top-down risk triage directives.

M. Residual Risk: Residual risk of a system, or the unmitigated risk remaining after risk treatment, shall be documented and made transparent, and resources needed to manage residual risk shall be documented.[88]

1. For each system, a residual risk level shall be assigned (which shall determine whether the system can be deployed) and individual residual risks shall be documented in a manner compatible for impact reassessment after potential deployment.

[84] Aligned with ISO/IEC 42001 6.1.3.
[85] Aligned with NIST AI RMF MANAGE 1.3.
[86] Aligned with NIST AI RMF GOVERN 1.2.
[87] Aligned with NIST AI RMF MANAGE 2.1.
[88] Aligned with NIST AI RMF 1.2.3 and MANAGE 1.4.

a) Residual risks shall measure the full scope of a system's potential impact, including to downstream acquirers of the system and to end users.[89]

2. Residual risk of a deployed system shall be publicly reported to inform end users and society about potential negative impacts of the system.[90]
3. Internal risk controls to manage residual risk of a system and of each of its components (including third-party technologies) for the duration of its operation shall be developed and documented.[91]
    a) The resources needed to manage residual risk of a system for the duration of its operation shall be estimated and later measured, in the same manner as for the system's initial risk treatment.

N. Impact and Risk Tracking, Inventorying, and Transparency: Materialized impacts from deployed systems shall be documented in an internal inventory and shall be broadly and publicly reported.[92]

1. Deployed systems shall be subject to continuous impact monitoring and regular impact measurement at a specificity-level and cadence commensurate with residual risk level and types. The process and tools (including AI actors involved, metrics and technology used) established for both continuous monitoring and discrete measurement of each system shall be documented and justified. Artifacts and insights shall be properly stored and made accessible.
2. AI incidents and the impacts that may have resulted from the incident shall be entered into *[organization name]* _____'s AI Incident, Impact, and Risk (IIR) database.
    a) The database shall track existing, unanticipated, and emergent AI risks based on factors such as intended and actual performance in deployed contexts.[93]
        (1) Additional risk tracking approaches are considered for settings where AI risks are difficult to assess using currently available measurement techniques or where metrics are not yet available.[94]
    b) This database shall have a standardized entry card that aligns with *[organization name]* _____'s risk and impact taxonomy and that allows for quick reference by teams.

O. Impact Contingency Planning and Recourse: *[organization name]* _____ shall have timely and clear incident management policies, including contingency plans and

[89] Aligned with NIST AI RMF MANAGE 1.4.
[90] Aligned with NIST AI RMF 1.2.3.
[91] Aligned with NIST AI RMF GOVERN 4.2 and MANAGE 3.1. Also aligned with ISO/IEC 42001 8.1.
[92] Aligned with NIST AI RMF GOVERN 4.2 and GOVERN 4.3.
[93] Aligned with NIST AI RMF MEASURE 3.1.
[94] Aligned with NIST AI RMF MEASURE 3.2.

recourse protocols, to respond to and recover from an AI impact or incident when it is identified or has occurred.[95]

1. Contingency plans shall include *[e.g., assigned and understood internal responsibilities, standardized protocols to quickly determine the proper response (automated when appropriate), processes to address issues related to third-party providers or buyers]* _____.[96] Plans are inclusive of processes to handle failures or incidents originating from third-party data or AI systems.[97]
2. AI systems that display performance or outcomes inconsistent with intended use, as demonstrated through performance and safety metrics or impact reassessment results, shall be subject a proper recourse protocol, including to *[e.g., supersede, disengage, or deactivate]* _____ the system.[98]
    a) In cases when an AI system presents unacceptable negative risk levels, such as when *[e.g., significant negative impacts are imminent, harms are measurably occurring, or catastrophic risks are identified]* _____, the system shall be immediately taken offline in a safe manner. The system shall only be redeployed according to a formal protocol that determines whether the impact or risk has been appropriately managed and how operation and monitoring of the system can be adjusted henceforth.[99]
3. In the aftermath of an AI incident, *[organization name]* _____ shall engage with affected users, operators, data subjects, and third parties according to the following requirements[100]:
    a) Conspicuous notifications with information relevant to the audience are delivered quickly;
    b) Individuals have clear and simple means to report any adverse experiences and expect a response within a reasonable time frame[101]; and
    c) Updates are made to system characteristics or operations, and individuals are provided timely updates of any relevant changes.

P. Impact Reassessment: In addition to continuous impact monitoring and regular impact measurement, an AI system shall undergo impact reassessment using a level of AI Impact Assessment (AIIA) commensurate to its risk-level in specific circumstances, including when *[e.g., an incident has precipitated significant changes to system design or operations, the business or context scope of a system is significantly updated or*

[95] Aligned with NIST AI RMF MANAGE 2.3.
[96] Aligned with NIST AI RMF MANAGE 2.4.
[97] Aligned with NIST AI RMF GOVERN 6.2.
[98] Aligned with NIST AI RMF MANAGE 2.4.
[99] Aligned with NIST AI RMF 1.2.3.
[100] Aligned with ISO/IEC 42001 A.8.
[101] Aligned with ISO/IEC 42001 A.8.3.

broadened, technical components are switched out in a manner that changes third-party relationships with vendors] _____.

1. The results of impact reassessments shall be used to update system-level documentation and the AI Incident, Impact, and Risk (IIR) database.
2. If the reassessment is triggered by an AI incident, its results shall be incorporated into project retrospective reports and used to update organization-wide best practices.

VII. Project Management

A. Existing project and/or product management at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing project/product management framework wherever absent.[102]

B. Additionally, *[organization name]* _____ identifies and aligns with existing regulations and guidelines impacting or guiding project and product management, including *[e.g. requirements to register the development of foundation models or other models with national or global impact]* _____.[103]

C. Project management of an AI system encompasses the processes of all or a subset of the following AI system life cycle stages, depending on the built or bought status of the system or of its components: *[e.g., Plan and Design, Collect and Process Data, Build and Use Model, Verify and Validate, Deploy and Use, Operate and Monitor[104]]* _____.[105]

1. Project management leaders and *[e.g. the RAI Operational Committee]* _____ shall determine and build the tools and processes (including tech platforms or applications) necessary to operationalize cross-functional collaboration and to standardize documentation creation and collection across AI systems' life cycles.
2. Project leads/managers, with the support of relevant personnel, shall ensure that all requirements detailed in other sections of this Policy, including *[e.g., Governance, Data Management, Risk Management, Stakeholder Management, Regulatory Compliance, AI Procurement]* _____, are properly executed throughout each AI system's life cycle.
3. **For bought systems:**
    a) Project leads and system integrator teams shall work closely with procurement teams to collect sufficient information to assess potential suppliers' responsible

[102] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[103] Aligned with NIST AI RMF 1.2.2.
[104] Aligned with NIST AI RMF 2.
[105] For more detailed guidance on management of a system across all lifecycle stages, refer to Responsible AI Institute's system-level resources.

AI development practices and to evaluate and verify potential systems' performance across trustworthiness characteristics, in alignment with the processes outlined in AI Procurement.

b) Project leads shall ensure that a selected (validated and deployed) system is used according to its intended uses and in alignment with documented objectives and processes for the responsible use of AI systems.[106]

4. **For built systems:**

a) Project leads shall ensure that a system is developed in accordance with documented objectives and processes for its responsible design and development.[107]

b) Developer teams shall work closely with procurement teams to collect sufficient information to assess potential suppliers' responsible development practices and to evaluate and verify potential components' performance in the system across trustworthiness characteristics, in alignment with the processes outlined in AI Procurement.

c) Project leads shall ensure that the system is used according to its intended uses and in alignment with documented objectives and processes for the responsible use of AI systems.[108]

5. **For sold systems:**

a) Project leads shall integrate buyer expectations and needs into how a system is developed and can be used.[109]

D. Human-AI Interaction and Configurations: Project teams shall clearly define and differentiate the various human roles and responsibilities when using, interacting with, or managing AI systems.[110]

1. Human oversight protocols shall be established and tested.[111]

a) Oversight roles can depend on the system's degree of autonomy. More autonomous or low-risk systems may be largely maintained by system operators (responsible for maintaining infrastructure, monitoring automated performance logging, and executing contingency plans) while less autonomous or higher risk systems may require human intervention, interpretation, or manipulation as part of regular operation (i.e. a human-in-the-loop).

[106] Aligned with ISO/IEC 42001 A.9.
[107] Aligned with ISO/IEC 42001 A.6.1.
[108] Aligned with ISO/IEC 42001 A.9.
[109] Aligned with ISO/IEC 42001 A.10.4.
[110] From NIST AI RMF Appendix C.
[111] Aligned with ISO/IEC 42001 B.9.3 and NIST AI RMF GOVERN 3.2, MAP 2.2, and MAP 3.5.

2. Evaluations of whether users (and/or end users), defined as *[e.g. "relevant interested parties who make decisions or are subject to decisions based on the AI system outputs"]* _____, can adequately interpret the AI system outputs shall be regularly conducted to inform the design of system output interfaces, communication methods, and user documentation.[112]
3. Design of sites of human-AI system interaction (such as between the system and operator or between the system and users) shall be informed by research and consultation on human biases and individual preferences, traits, and skills that may influence interpretation of system outputs and generate or amplify harms.[113]

E. DEI and Stakeholder Engagement: Activities across a system life cycle, including *[e.g., AI system design and development[114], evaluation[115], performance monitoring[116], and impact assessments[117]]* _____, should all be **informed by a representatively diverse immediate team** (e.g., diversity of demographics, disciplines, experience, expertise, and backgrounds)[118] and by **consultation with and feedback from internal and external AI actors**.[119] Refer to Workforce Management for more on DEI guidance and to Stakeholder Management and Engagement on consultation and feedback mechanisms.

F. Managing Adaptation and Drift: Project owners shall implement mechanisms to sustain the value of deployed AI systems against unintended scope drift.[120]

1. Systems and their components (e.g., data for fine-tuning, pre-trained models[121]) shall be continually monitored for drift with respect to characteristics, quality, suitability, and behavior. Undesired drift with respect to performance criteria and other documented systems requirements shall be corrected and measures shall be applied to prevent future drift. All other relevant types of drift, including increased capabilities from adaptive learning, shall be documented once identified.
2. Project owners, with support of *[e.g. the RAI Operational Committee]* _____, shall proactively determine the current system's boundaries for business use case scope and technical scope. Change or expansion beyond these boundaries requires *[e.g., an impact assessment reassessment, a new project proposal]* _____.

G. System-level Documentation: As artifacts of a system's progression across life cycle stages, system-level documentation shall be created and maintained, including:

---

[112] Aligned with ISO/IEC 42001 B.6.2.4.
[113] From NIST AI RMF Appendix C.
[114] Aligned with NIST AI RMF GOVERN 5.2.
[115] Aligned with NIST AI RMF MEASURE 3.3.
[116] Aligned with NIST AI RMF MEASURE 4.3.
[117] Aligned with NIST AI RMF MAP 5.2 and MEASURE 1.3.
[118] Aligned with NIST AI RMF GOVERN 3.1.
[119] Aligned with NIST AI RMF GOVERN 5.1.
[120] Aligned with NIST AI RMF MANAGE 2.2.
[121] Aligned with NIST AI RMF MANAGE 3.2.

1. Entries into enterprise-wide inventories, including the **AI system inventory**[122] and *[e.g., Data inventory, Project inventory, AI Incident, Impact, and Risk (IIR) database*[123]*]* _____;

    a) Entries into inventories for resource documentation and management can include for *[e.g., AI system components, data resources*[124]*, tooling resources*[125]*, system and computing resources*[126]*, human resources*[127]*]* _____.[128] Resource documentation shall be incrementally updated across each life cycle stage and be used to inform future risk triage, project prioritization, and resource allocation efforts.

2. **Business use case documentation** that can describe, but is not limited to:

    a) Intended purpose, business value, and context(s) of business use[129];

        (1) Potential benefits of intended AI system uses, capabilities, and performance.[130]

        (2) Laws, norms and expectations specific to the prospective settings in which the AI system will be deployed.[131]

    b) Targeted application scope, based on the system's capability, established context, and AI system categorization[132];

        (1) Viable non-AI alternative systems, approaches, or methods based on the system's intended purpose and scope.[133]

    c) *[organization name]* _____'s role and activities with respect to the system's intended purpose, in alignment with regulatory definitions such as *[e.g., AI Provider, Deployer, Importer, Distributor]* _____ and *[e.g., making available on the market, putting into service]* _____, respectively.[134]

---

[122] Aligned with NIST AI RMF GOVERN 1.6.
[123] See Appendix A for more examples.
[124] See ISO/IEC 42001 B.4.3 for a list of potential documentation on data resources utilized for the AI system.
[125] See ISO/IEC 42001 B.4.4 for a list of potential documentation on tooling resources utilized for the AI system.
[126] See ISO/IEC 42001 B.4.5 for a list of potential documentation on system and computing resources utilized for the AI system.
[127] See ISO/IEC 42001 B.4.5 for a list of potential documentation on human resources utilized for the AI system.
[128] Aligned with ISO/IEC 42001 B.4.2.
[129] From NIST AI RMF MAP 1.4.
[130] From NIST AI RMF MAP 3.1.
[131] From NIST AI RMF MAP 1.1.
[132] From NIST AI RMF MAP 3.3.
[133] From NIST AI RMF MANAGE 2.1.
[134] Aligned with ISO/IEC 42001 4.1.

d) System requirements, including for new AI systems or material enhancements to existing systems[135];

e) Assumptions and related limitations about AI system's purpose or knowledge and how system output may be utilized and overseen by humans[136];

f) Potential risks across the development or product AI life cycle[137]; and

(1) Potential costs, including non-monetary costs, which result from expected or realized AI errors or system functionality and trustworthiness[138]; and

(2) Potential positive and negative impacts of system uses to individuals, communities, organizations, society, and the planet.[139]

g) Specific set or types of users and their expectations[140];

3. **System technical documentation** that is made appropriate for each relevant or interested party, including users, partners, and supervisory authorities.[141] The documentation can describe, but is not limited to[142]:

a) Technical assumptions and limitations (e.g. related to system interoperability, run-time environment, data quality, AI explainability)[143];

b) Human-AI configurations, including operator, user, or human-in-the-loop capabilities and instructions[144], and configuration evaluations that meet applicable requirements (including human subject protection) and are representative of the relevant population[145];

c) The system's intended capabilities and the implementation methods and architecture (e.g., classifiers, generative models, recommenders)[146];

d) Test, evaluation, validation, and verification (TEVV) and system metrics, test sets, and tools calibrated to specific scientific integrity considerations, including *[e.g. those related to experimental design, data collection and selection (e.g., availability, representativeness, suitability), and construct validation.[147]]* _____[148];

[135] Aligned with NIST AI RMF MAP 1.6 and ISO/IEC 42001 B.6.2.2.
[136] From NIST AI RMF MAP 1.1 and MAP 2.2.
[137] From NIST AI RMF MAP 1.1.
[138] From NIST AI RMF MAP 3.2.
[139] From NIST AI RMF MAP 1.1.
[140] From NIST AI RMF MAP 1.1.
[141] See ISO/IEC 42001 B.6.2.7 for more details on potential technical documentation elements.
[142] Aligned with ISO/IEC 42001 B.6.2.4 and B.6.2.6.
[143] Aligned with ISO/IEC 42001 B.6.2.7.
[144] Aligned with ISO/IEC 42001 B.6.2.7.
[145] From NIST AI RMF MEASURE 2.2.
[146] From NIST AI RMF MAP 2.1.
[147] From NIST AI RMF MAP 2.3.
[148] Aligned with NIST AI RMF GOVERN 4.3, MAP 1.1, and MEASURE 2.1.

e) Results from an internal audit or technical assurance process to enable impartial system evaluation by those outside of the developer team, if deemed necessary[149];

f) Release criteria requirements, including acceptable ranges for operational factors and performance errors, and any acceptable factors that affect a system's ability to reach minimum release criteria[150];

g) Qualitative or quantitative performance or assurance criteria, calibrated to deployment setting(s)[151];

h) Roles and processes for the responsible operation of the AI system, including for monitoring and improvements[152] (including of pre-trained models)[153], appeal and override, decommissioning[154], incident management[155] (and enabling systems to fail safely)[156], recovery, and change management[157];

4. Complete, up-to-date, and accurate **user documentation**, including technical specifications but also general notification and information about their interaction with an AI system, including *[e.g., necessary information to properly interpret system outputs[158], limitations of system's accuracy and performance, disclosure of recent system updates, incidents, or impacts, contact information, means to report feedback or harms, links for additional informational materials]* _____, presented in an accessible and understandable manner.[159]

5. **Documentation** of processes, decisions, and results **across each life cycle stage**, with justifications of each, including of:

   a) **Design choices** made during *[e.g. "the Plan and Design, Collect and Process Data, and Build and Use Model"]* _____ stage(s), including *[e.g., the choice between models, machine learning architecture and methods, data sets, infrastructure options for compute and interoperability, response to security threats, user or output interface design, human-AI configuration]* _____[160];

   b) Measures for **verification and validation** of the system during *[e.g. "the Verify and Validate"]* _____ stage(s), including *[e.g. the process and results of evaluating the system with respect to the environmental impact and sustainability of model*

[149] Aligned with ISO/IEC 42001 9.2.2.
[150] Aligned with ISO/IEC 42001 B.6.2.4.
[151] From NIST AI RMF MEASURE 2.3.
[152] Aligned with NIST AI RMF MANAGE 4.2.
[153] Aligned with NIST AI RMF MANAGE 3.2.
[154] Aligned with NIST AI RMF GOVERN 1.7.
[155] Aligned with NIST AI RMF MANAGE 4.3.
[156] Aligned with NIST AI RMF MEASURE 2.6.
[157] From NIST AI RMF MANAGE 4.1.
[158] Aligned with NIST AI RMF MAP 2.2.
[159] Aligned with ISO/IEC 42001 B.8.2.
[160] See ISO/IEC 42001 B.6.2.3 for a list of design choices to be documented.

training and management activities[161] *and across each trustworthiness characteristic*[162]*, in support of but also beyond the AI Impact Assessment process]* _____[163];

c) Design and implementation of a **deployment plan** for *[e.g. "the Deploy and Use"]* _____ stage(s) that *[e.g. details roles and required approvals, a timeline, and a strategy for testing and feedback (e.g., phased roll-outs, pilots, beta or full release)]* _____ tailored to deployment contexts and requirements and to the risk profile of the system[164]; and

d) Measures for post-deployment system **operation and monitoring** during *[e.g. "the Operate and Monitor"]* _____ stage(s), including *[e.g., event logs*[165]*, monitoring logs of system behavior and user input*[166]*, incident reports, impact reassessment results]* _____.[167]

H. Retrospective Learning and Integration: Project owners, with support of *[e.g. the RAI Operational Committee]* _____, shall direct efforts to integrate lessons learned and best practices from AI projects (attempted or deployed) into relevant enterprise-level databases (e.g., on unexpected risks, into a bank of red-teaming prompts) and educational resources, to update and augment the AI and responsible AI capabilities of *[organization name]* _____'s workforce for future AI projects.

VIII. Stakeholder Management and Engagement

A. Existing stakeholder management and engagement at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing stakeholder management strategy wherever absent.[168]

B. AI actors, both internal and external to *[organization name]* _____, shall be engaged through regular and formalized processes to inform and improve AI systems throughout their life cycles, the knowledge of and responses to risks and impacts from systems, and *[organization name]* _____'s AI management system as a whole.

C. Managing and engaging stakeholders (internal and external AI actors) consist of the following activities:

[161] Aligned with NIST AI RMF MEASURE 2.13.
[162] See AI Principles. Aligned with NIST AI RMF MEASURE 2.5-2.12 and NIST AI RMF 3.
[163] Aligned with ISO/IEC 42001 9.1 and A.6.2.4.
[164] Aligned with ISO/IEC 42001 B.6.2.5.
[165] Aligned with ISO/IEC 42001 B.6.2.8.
[166] Aligned with NIST AI RMF MEASURE 2.4 and MANAGE 4.1.
[167] Aligned with ISO/IEC 42001 9.1 and B.6.2.6.
[168] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.

1. Consultation and Feedback[169]: All relevant and interested parties, including *[e.g., representatives of internal functions like Legal or Procurement, the intended user and human-in-the-loop groups, potentially impacted groups; and domain and socio-cultural experts]* _____, shall be regularly consulted and solicited for feedback throughout a system's life cycle.

    a) Consultation and feedback processes shall be formal and compulsory but tailored to the context and risk level of the system. Those invited to participate in consultations or to provide feedback shall be identified through a balanced and accessible recruitment strategy that prioritizes demographic diversity and broad domain and user experience expertise.[170] Participants shall also be compensated for their efforts except in proper extenuating circumstances.

    b) The needs, expectations, concerns, and experienced impacts of those consulted or providing feedback shall be used to inform system requirements and to guide development and improvement of the system, including related to *[e.g., system design and implementation[171]; performance[172]; system trustworthiness[173]; assessments[174]; potential risks and impacts[175]; continual improvements[176]]* _____.[177]

    c) Voluntary feedback mechanisms shall also be made accessible for all interested parties who wish to provide feedback at any time. Such mechanisms shall be directly available from a system's output interface and provide options for different types of feedback (e.g., rating of output quality, corrective actions for RLHF, options to send a message, escalation to a report of concerns or impacts (see Reporting and Response)).

2. Notification and Communication[178]: *[organization name]* _____ shall provide necessary, accessible, and appropriate information to all relevant and interested parties, including *[e.g., representatives of internal functions like Legal or Procurement, of the intended user and human-in-the-loop groups, of potentially impacted groups; domain and socio-cultural experts; policymakers and other civil society actors; and third-party assessors]* _____.

    a) Plans for notifications and other communication of information about systems, including *[e.g., that one is interacting with or will be subject to a decision made (in*

[169] Aligned with NIST AI RMF GOVERN 5.1 and MAP 1.2.
[170] Aligned with NIST AI RMF MAP 1.2.
[171] Aligned with NIST AI RMF GOVERN 5.2.
[172] Aligned with NIST AI RMF MEASURE 4.3.
[173] Aligned with NIST AI RMF MEASURE 4.2.
[174] Aligned with NIST AI RMF MEASURE 1.3.
[175] Aligned with ISO/IEC 42001 B.5.4 and NIST AI RMF MAP 5.2 and MEASURE 4.1.
[176] Aligned with NIST AI RMF MANAGE 4.2.
[177] Aligned with ISO/IEC 42001 4.2.
[178] Aligned with ISO/IEC 42001 7.4 and A.8 and NIST AI RMF GOVERN 4.3.

part) by a system; user documentation[179]; incidents, impacts, and risks[180]] _____ shall be proactive, timely, and meet legal, regulatory, and stakeholder-driven transparency obligations.

3. Reporting and Response[181]: All relevant and interested parties, including *[e.g., representatives of internal functions like Legal or Procurement, of the intended user and human-in-the-loop groups, of potentially impacted groups; domain and socio-cultural experts; policymakers and other civil society actors; and third-party assessors]* _____, shall be be provided mechanisms to report concerns and impacts related to a system and receive corrective action from *[organization name]* _____.
    a) Such a mechanism shall *[e.g., protect individuals from identification and reprisals; be accessible to all workforce members; have appropriate personnel and capabilities (including investigation, resolution, escalation, and reporting powers); respond and act in a timely manner[182]]* _____;
    b) Reports shall be used to inform system requirements and to improve the system and *[organization name]* _____'s overall AI management system. Any individuals who find the response or remedy provided for their report insufficient have means to note their public dissatisfaction and escalate their issue.
4. Public and Ecosystem Engagement: In the spirit of continual learning and beneficence, *[organization name]* _____ shall contribute to collaborative, shared, and open-source initiatives with the public and with the broader AI ecosystem to promote the responsible development, use, procurement, and supply of trustworthy AI.
    a) Activities can include *[e.g., creating thought leadership; leveraging network and marketplace synergies to create stronger technological or informational resources (e.g., developing an industry-specific bias evaluation toolkit, sharing common lessons from the AI Incident, Impact, and Risk (IIR) database; providing insights to shared learning spaces like working groups]* _____.

IX. Workforce Management

A. Existing workforce management at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing workforce management framework wherever absent.[183]

[179] Aligned with ISO/IEC 42001 A.8.2.
[180] Aligned with ISO/IEC 42001 B.8.5 and NIST AI RMF MANAGE 4.3.
[181] Aligned with ISO/IEC 42001 B.8.4 and NIST AI RMF MEASURE 1.2 and MEASURE 3.3.
[182] Aligned with ISO/IEC 42001 B.3.3.
[183] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.

1. Diversity, Equity, and Inclusion (DEI) shall be integrated into every component of *[organization name]* _____'s AI strategy.[184]

    a) DEI personnel at *[organization name]* _____, if present, shall work with *[e.g. the RAI Operational Committee]* _____ to integrate DEI objectives and requirements into responsible AI efforts. DEI experts shall also be represented in cross-functional stakeholder processes for AI, including *[e.g., AI impact assessments, system approval for progression]* _____.

    b) Internal and external AI actors across all AI-related processes shall reflect a diversity of backgrounds, skills, and capacities. Gaps in diverse representation shall be identified and remediated through *[e.g., consultation with external AI actors, hiring]* _____.

    c) All interactions between AI actors shall be conducted with mutual respect and foster an environment of belonging that values unique and dissenting perspectives about *[organization name]* _____'s AI activities and AI systems.

    d) Existing DEI policies, initiatives, and resources, including *[e.g., Employee Resource Groups (ERGs), development programs, events for underrepresented employees]* _____ shall be updated or augmented as necessary to support employees in AI-specific upskilling or in AI-specific equity concerns.

    e) The use of AI at *[organization name]* _____ shall not violate existing DEI commitments nor regulatory and legal requirements (e.g., discrimination, harassment, and equal employment opportunity), for example, in the context of *[e.g., employee performance evaluations, productivity tracking using computer vision]* _____.

B. Employee Awareness and Competence[185]: Employees at *[organization name]* _____ shall develop the necessary awareness of and competence for their AI role and responsibilities through training, educational resources, and other upskilling initiatives implemented by *[e.g. the RAI Operational Committee]* _____.

    1. Employees and relevant partners of *[organization name]* _____ shall receive responsible AI training to enable them to understand and perform their duties with respect to the AI management system outlined in this Policy and to other related policies and processes.[186]

        a) Employees shall demonstrate adherence to responsible use policies and processes, including only using approved systems (to avoid "shadow AI" risks) and only using systems as intended by system documentation.[187]

[184] Aligned with NIST AI RMF GOVERN 3.1 and MAP 1.2.
[185] Aligned with ISO/IEC 42001 7.2 and 7.3.
[186] Aligned with NIST AI RMF GOVERN 2.2.
[187] Aligned with ISO/IEC 42001 B.9.2 and B.9.4.

2. Educational and RAI outreach efforts within *[organization name]* _____ shall help to cultivate a RAI culture, including a critical thinking and safety-first mindset in the design, development, deployment, and use of AI systems,[188] and can include *[e.g., Communities of Practice for function- or role-specific RAI collaboration, a Center of Excellence to centralize RAI guidance and resources]* _____.
3. Employees shall be provided with role- and/or system-specific training and educational tools to enable and evaluate their understanding and fulfillment of RAI requirements for a system.[189]
    a) Employees interacting with a system (e.g., as a user, operator, human-in-the-loop) shall demonstrate proficiency with respect to knowledge of relevant system information and to the ability to responsibly execute and complete tasks.[190]

C. Hiring for Responsible AI: *[organization name]* _____ shall identify deficiencies in AI and RAI capacity and skills, identify which gaps can be addressed by training and which are relevant to hiring, and seek new talent as necessary. RAI workforce planning shall be informed by human resource inventorying activities and documentation.[191]

X. Regulatory Compliance

A. Compliance with existing data, analytics, and technology regulation at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance for compliance for AI systems shall be incorporated into existing compliance activities wherever absent.[192]

B. The compliance team shall actively monitor and understand all existing and emerging legal and regulatory requirements relevant to all AI activities.[193]

C. The compliance team shall map the compliance requirements of each AI system, or otherwise develop tools and processes, such as checklists or compliance meetings, to enable technical teams to determine requirements at each stage of a system's life cycle.

D. The compliance team shall be ultimately accountable for the perceived or realized compliance of every system at each stage of the life cycle, from initial project scoping to post-deployment or post-sale to potential decommission, and shall be provided with timely and sufficient transparency to enable this responsibility.

---

[188] From NIST AI RMF GOVERN 4.1.
[189] Aligned with NIST AI RMF MAP 1.6.
[190] Aligned with NIST AI RMF MAP 3.4.
[191] Aligned with ISO/IEC 42001 B.4.6. Refer to Project Management (System-level Documentation) for more on resource documentation.
[192] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[193] Aligned with NIST AI RMF GOVERN 1.1.

E. Compliance team members shall receive appropriate foundational legal and AI training, developed and distributed by *[e.g. the RAI Operational Committee]* _____, to enable successful implementation of compliance organization-wide.[194]

F. The compliance team shall develop role- or system-specific compliance training and other guidance materials for the workforce, with the support of function leadership or system owners, respectively.

XI. AI Procurement

A. Existing procurement at *[organization name]* _____ is implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following guidance shall be incorporated into the existing procurement strategy wherever absent.[195]

B. The procurement team shall engage with product teams and intended users of procured products/services in the earliest stage of use case scoping, and shall develop standardized processes to initiate and manage such engagements, such as *[e.g., ticket submissions for procurement requests]* _____.

C. The procurement team shall identify and document all external parties involved during the life cycle of an AI system, including all partners, suppliers, and buyers.[196]

D. Procurement team members shall receive appropriate foundational compliance and AI training and maintain continual and informed use of risk management tools, including *[e.g., all levels of the AI Impact Assessment, the AI Incident, Impact, and Risk (IIR) database, the AI resources inventory]* _____.[197]

E. Procurement teams shall support the creation of analyses for third-party resources in the **cost-risk/benefit repository**, and shall use the repository as a resource to inform future procurement efforts.[198]

F. In the case of products or services that do not employ AI directly but may be delivered in part by AI used internally by the supplier, the procurement team shall subject every potential supplier to a **Responsible Supplier Assessment** to assess the responsible AI maturity of the supplier organization. Results from the Responsible Supplier Assessment shall inform the decision to accept the supplier.

    1. If an existing procured product or service that does not employ AI directly but may be newly delivered in part by AI used internally by the suppliers, the procurement team shall initiate a Responsible Supplier Assessment with the supplier. Results from the

[194] Aligned with NIST AI RMF GOVERN 2.2.
[195] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[196] Aligned with ISO/IEC 42001 A.10.2.
[197] Aligned with NIST AI RMF GOVERN 2.2.
[198] Aligned with NIST AI RMF MANAGE 3.1.

Responsible Supplier Assessment shall inform the decision to maintain a relationship with the supplier.

G. For built systems[199]:

1. The procurement team shall clarify with the compliance team, product team, and intended users (if internal) the purpose, technical needs, and legal requirements of an AI component (e.g., data sets, AI models, platforms) in its use case to develop a must-have list for potential suppliers.
2. The procurement team shall consider multiple possible suppliers during a process, while understanding that it is possible that none of the suppliers will fulfill all technical and responsible AI needs, and therefore, none may be accepted.
3. The procurement team shall subject every potential supplier to a Responsible Supplier Assessment (for AI products like models) or to a standard supplier assessment (for other components), depending on the development or use of AI to deliver the product.
4. For every potential component, the procurement team shall request sufficient information from the potential supplier to conduct a Low-Touch AI Impact Assessment of the proposed system, including the component. If the information is refused, the procurement team shall not move forward with the supplier.
5. The procurement team shall determine what documentation the potential supplier must provide for the responsible procurement and integration of the component. This includes *[e.g., data documentation, training process, change policies]* _____.
6. The procurement team shall establish absolute thresholds for the level of responsible AI maturity required, level of AI risk allowed, and type of documentation required for all suppliers, which shall determine whether an individual prospective supplier is accepted.
7. If the potential supplier is accepted, the procurement team shall negotiate and document terms of use with the supplier, including *[e.g., delineation of liability and responsibility for corrective action in downstream impacts, data sharing, deletion, and privacy agreements, force majeure clause, termination clause]* _____.
8. Upon acceptance of the terms and conditions, the procurement team shall clarify and document ongoing communication and transparency processes with the supplier, including *[e.g., timely notification from suppliers when the component is changed or an unmitigated risk has emerged, high-level performance monitoring reports to the supplier in the case of AI models, a feedback or dispute mechanism with the supplier]* _____.
9. The procurement team shall share information about the component and supplier with all relevant functions and audiences, including *[e.g., compliance teams, product*

[199] Aligned with ISO/IEC 42001 B.10.3.

teams, the RAI Operational Committee, the user group] _____ to ensure responsible use and integration of the component into built systems.

H. For bought systems[200]:

1. The procurement team shall clarify with the compliance team, product team, and intended users (if internal) the purpose, technical needs, and legal requirements of the use case to develop a must-have list for potential suppliers.
2. The procurement team shall consider multiple possible suppliers during a process, while understanding that it is possible that none of the suppliers will fulfill all technical and responsible AI needs, and therefore, none may be accepted.
3. The procurement team shall subject every potential supplier of an AI system to a Responsible Supplier Assessment to assess the responsible AI maturity of the supplier organization.
    a) If an existing procured tool or application has rolled out new AI capabilities, the procurement team shall initiate a Responsible Supplier Assessment with the supplier.
4. The procurement team shall request sufficient information from the potential supplier to conduct a Medium-Touch AI Impact of the system in the context of its intended use, and if refused, shall not move forward with the supplier.
    a) If an existing procured tool or application has rolled out new AI capabilities, the procurement team shall conduct its own High-Touch AI Impact Assessment, requesting information from the supplier as needed.
5. The procurement team shall determine what system documentation the potential supplier must provide for the responsible procurement and use of the system. This includes *[e.g., the supplier's AI risk or impact assessment, data documentation, summary of incidents, change policies]* _____.
    a) If an existing procured tool or application has rolled out new AI capabilities, the procurement team shall request additional system documentation, including *[e.g., the supplier's AI risk or impact assessment, data documentation, training process, summary of incidents, change policies]* _____.
6. The procurement team shall establish absolute thresholds for the level of responsible AI maturity required, level of AI risk allowed, and type of documentation required for all suppliers, which shall determine whether an individual prospective supplier is accepted.
    a) If an existing procured tool or application has rolled out new AI capabilities, the procurement team shall establish absolute thresholds for the level of responsible AI maturity required, level of AI risk allowed, and type of documentation required

[200] Aligned with ISO/IEC 42001 B.10.3.

to retain the product. All relationships with suppliers who do not meet the required thresholds shall be terminated, within reason.

7. If the potential supplier is accepted, the procurement team shall negotiate and document terms of use with the supplier, including *[e.g., delineation of liability and responsibility for corrective action in downstream impacts; data sharing, deletion, and privacy agreements; PII processor or controller role[201]; force majeure clause; termination clause]* _____.
    a) If an existing procured tool or application has rolled out new AI capabilities and the procurement team has determined that the product will be retained, the procurement team shall review existing terms of use with the supplier and negotiate and document any updates needed for responsible procurement and use.
8. If the terms and conditions are agreed upon, the procurement team shall clarify and document ongoing communication and transparency processes with the supplier, including *[e.g., prompt notification from suppliers when the system is changed or an unmitigated risk has emerged, high-level performance monitoring reports to the supplier, a feedback or dispute mechanism with the supplier]* _____.
    a) If an existing procured tool or application has rolled out new AI capabilities and the procurement team has determined that the product will be retained, the procurement team shall review ongoing communication and transparency processes with the supplier and negotiate and document any updates needed for responsible procurement and use.
9. The procurement team shall share information about the system and supplier with all relevant functions and audiences, including *[e.g., compliance teams, product teams, the RAI Operational Committee, the user group]* _____ to ensure responsible use and integration of the system and enable the development of system-specific guidance.

I. For sold systems[202]:

1. If the system is intended to be sold, the procurement team shall additionally determine what information shall be requested from suppliers to enable downstream transparency with buyers.

XII. Documentation Management

A. Existing documentation processes at *[organization name]* _____ are implemented by *[governance and policies]* _____ and *[tools and processes]* _____. The following

[201] Aligned with ISO/IEC 42001 A.10.2.
[202] Aligned with ISO/IEC 42001 B.10.3 and B.10.4.

guidance shall be incorporated into existing documentation practices wherever absent.[203]

B. Additionally, *[organization name]* _____ identifies and aligns with existing regulations and guidelines for documentation and reporting for AI systems, including *[e.g. "reporting results of safety tests of high-risk models, as required by the U.S. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"]* _____.[204]

C. *[Organization name]* _____ identifies the following roles as responsible for proper maintenance and control of documentation relevant to its management of AI systems and its broader AI strategy[205]:

   1. The *[e.g. RAI Operational Committee]* _____ shall maintain *[e.g. "all enterprise-level AI-specific policy documents and artifacts"]* _____. This includes *[e.g., charter documents for new AI bodies such as the Steering/Operational Committees and working groups, AI component inventories, and databases for AI impacts]* _____, among others;
   2. The *[e.g. "business lead, with support of the technical lead"]* _____ shall maintain *[e.g. "documentation for each AI project and/or system they oversee"]* _____. This includes *[e.g., business use case documentation, system technical documentation, AI impact assessment results, and approval gate documentation]* _____, among others;
   3. The *[e.g., function leaders, owners of specific risk areas]* _____ shall maintain *[e.g. "all related documentation, including policies and system-specific artifacts"]* _____. This includes *[e.g., data management policies and standards, governance for AI procurement, and privacy risk assessment results]* _____, among others; and
   4. The *[role or group]* _____ shall maintain *[documentation category]* _____. This includes *[examples of documents]* _____, among others.

D. Maintenance and control of documentation encompasses the following responsibilities[206]:

   1. Enablement and review of the creation or update (i.e. version control) of documentation by the proper AI actors;
   2. Storage and preservation of documents' legibility and suitability for use, including through *[features e.g. "identification and description, format, and media"[207]]* _____;

[203] Aligned with ISO/IEC 42001 A.2.3 and NIST AI RMF GOVERN 1.2.
[204] Aligned with NIST AI RMF 1.2.2.
[205] See Appendix A for a list of documents under each identified documentation category. In accordance with ISO/IEC 42001 7.5.3, organizations shall also identify and control necessary external documentation.
[206] Aligned with ISO/IEC 42001 7.5.3.
[207] See ISO/IEC 42001 7.5.2 for examples of each listed feature.

3. Enforcement of access controls for permissions to view, change, retrieve, use, or distribute documentation, to enable traceability and meet reporting or audit trail requirements; and
4. Retention and disposition, in accordance with regulatory requirements, for as long as other organizational policies, and with adequate protection of documentation in line with confidentiality and usage requirements.

XIII. Review and Enforcement of the AI Policy

A. The *[role or governance body e.g. RAI Operational Committee]* _____ shall maintain this Policy and is responsible for **monitoring** and **reviewing** the effectiveness of the Policy and of the AI management system described therein *[e.g. every 3 months]* _____, with ongoing input from cross-functional stakeholders and periodic review by executive *[owners, champions, and/or sponsors]* _____.[208]

   a) **Corrective updates** and other changes to the Policy shall be planned, and their effects are reviewed with any adverse effects mitigated.[209]

   b) Each component of the AI Policy shall undergo periodic review and ongoing monitoring, with the *[e.g. RAI Operational Committee]* _____ assigning sub-roles and responsibilities.[210]

B. The *[role or governance body e.g. RAI Operational Committee]* _____ shall be responsible for procedures to **detect** and **respond to deviations and violations** of the Policy.[211]

[208] Aligned with ISO/IEC 42001 5.2, 9.2.1, 9.3.1 and A.2.4. ISO/IEC 42001 9.3.2-9.3.3 provides guidance on the desired inputs and results of a review of the AI Policy and the AI management system described therein.
[209] Aligned with ISO/IEC 42001 6.3, 8.1, and 10.1-10.2. ISO/IEC 42001 10.2 provides guidance on actions to take when a nonconformity occurs.
[210] Aligned with NIST AI RMF GOVERN 1.5.
[211] Aligned with ISO/IEC 42001 5.3. ISO/IEC 42001 10.2 provides guidance on actions to take when a nonconformity occurs.

XIV. Conclusion/Acknowledgement

A. This Policy was developed using the Responsible AI Institute's AI Policy Template.
B. Signatories:

_____	_____
[Name, Role]	[Date]
_____	_____
[Name, Role]	[Date]
_____	_____
[Name, Role]	[Date]

Appendix A

Appendix A provides a potential list of documents and artifacts that can be linked to the AI Policy to provide a complete ecosystem view of the organization's AI strategy.

Enterprise-level Policy Documents:

- Governance structure policies, including charter documents for new AI bodies or groups
- Existing ethics, DEI, ESG, social responsibility, corporate human rights policies
- Stated and achieved voluntary commitments related to AI (e.g., Frontier AI Safety Commitments, 2023 White House AI commitments)
- Data management policies and standards
- Risk management policies and standards
- Product development or project management policies and standards
- Stakeholder engagement policies
- Procurement framework and policies
- Infosecurity and/or cybersecurity policies and standards
- Workforce policies, including usage policies, planning and hiring protocols
- Any other existing policies that have been augmented with AI-specific guidance

Enterprise-level Artifacts:

- Risk/impact taxonomy
- Complete and detailed list of prohibited use cases
- AI system inventory with model cards
- AI regulatory tracker
- AI Incident, Impact, and Risk (IIR) database
- Resource utilization tracker
- Data inventory
- Project inventory
- Best practices for AI database
- Workforce RAI development, education, and training resources

System-level Artifacts (per AI system):

- AI technical documentation, including role-specific documentation (e.g. for users, operators)
- AI business use case documentation
- Contracts, with suppliers or buyers
- AI impact assessment results
- Documentation across life cycle stages (e.g., data set quality test results, model evaluation results, performance and event logs, governance gate approval forms)
- Relevant external documented information, such as from partners or suppliers

Supporting You on Your RAI Journey

The Responsible AI Institute is here to support organizations on their AI journeys. Becoming a member enables essential support and direction for making significant progress toward future-proofing RAI governance and implementation.

Click below to learn more about how we help our members achieve their Responsible AI goals:

About Responsible AI Institute (RAI Institute)

Founded in 2016, the Responsible AI Institute (RAI Institute) is a global and member-driven non-profit dedicated to enabling successful responsible AI efforts in organizations. We accelerate and simplify responsible AI adoption by providing our members with AI assessments, benchmarks and certifications that are closely aligned with global standards and emerging regulations.

Where to connect with us: