AI Policy Framework

The appropriate and ethical use of artificial intelligence is a top priority for Anthology and education institutions around the world. That is why, in 2022, we established a cross-functional and diverse working group to implement a dedicated **Anthology Trustworthy AI program**. This program was formally implemented in 2023 and is led by our Trustworthy AI Officer.

As part of the Trustworthy AI program, we are committing ourselves to implementing the following seven (7) principles, all of which are based on and aligned to the principles of the **NIST AI Risk Management Framework**, the **EU AI Act,** and the **OECD Principles**:

Trustworthy AI: 7 Principles

- Fairness
- Accountability
- Reliability
- Value Alignment
- Humans in Control
- Transparency & Explainability
- Privacy, Security, Safety

Given that these principles are based on several international standards, they can be a good starting point for higher education institutions who are interested in developing and adopting specific policies and programs on the ethical use of AI within their institution.

Going through the implementation of our Trustworthy AI program, we have also learned important lessons on how to set up a governance model that involves the necessary stakeholders, how to draft related policies and documents, and how to implement such a program.

Through this suggested policy framework, Anthology wants to provide higher education institutions with some guiding questions for them to consider when starting to think about their ethical AI policies and procedures and share insights from our implementation.

Please remember that this suggested policy framework is not a one size fits all document, but rather a guiding document that institutions can use according to their needs and specific circumstances.

Anthology®

Stakeholder Identification and Meetings

One of the most important aspects of a successful implementation of ethical AI policies is to have diverse and cross-functional voices represented.

Broad representation and input from across the institution and from multiple levels is critical. When identifying individuals to be consulted and informed, institutions should make sure they include all stakeholders in the policy formation process; this ideally includes students or student representative bodies. Not all stakeholders will have equal input on the policy formation, and some may only need to be kept informed of the process.

The following questions should help institutions easily identify who the stakeholders are:

- Who are the stakeholders likely to use, benefit from, or be impacted by (generative) AI at your institution?
- Who are the stakeholders who have a role to play in managing the risks of using (generative) AI?
- Are all appropriate constituencies represented or able to contribute to the discussion?

Once the stakeholders have been identified, the following questions may be helpful to guide discussions and meetings:

- Which function/team should take on the role of coordinating and overseeing the efforts?
- Should there be a senior executive sponsor of the initiative who can support the efforts within senior management?
- Is there a core group of stakeholders that could drive the formulation and implementation of ethical AI policies (while ensuring that the wider group of stakeholders are informed and consulted with as needed)?
- Are there any boards or committees that should be informed and consulted?
- How often will the stakeholders meet?
- What is the timeline for enacting a program and the respective policies?
- What is the final approval process and ongoing governance process for the policy?
- Who will be responsible for the program and implementation of the policies across the institution?
  - What is the process?
  - Who is responsible for the communication plan?

Identifying the stakeholders, assigning people responsibilities, and scheduling a meeting cadence will assist in the policy adoption process and inspire a continuous working group environment.

Defining Institutional Position on Generative AI

The first thing an institution might want to consider is what their stance on (generative) AI is. There might be a culture of exploration and innovation, or conversely, the culture may be one of risk-reduction and control.

The following are key questions the institution should ask when defining its overall policy position:

- Does the institution, or specific areas of the institution, already have a stance on the use of AI, and in particular generative AI?
- From an institutional culture perspective, what is the general attitude toward the use of generative AI tools?
- How are different parts of the institution currently using generative AI and how are they planning to use it in the near future? (This will also help with developing an inventory of generative AI use)
- Does the institution have sufficient knowledge and expertise regarding the technology and the risks involved (or should it aim to upskill key stakeholders and/or look for external expertise)?
- What are the risks and harms of using generative AI in a manner that is illegal or unethical for the institution? (See for instance the NIST AI Risk Management Framework for a description of typical risks and harms and how they can be managed)
- Are there similar efforts (e.g., privacy, security, compliance, risk management programs, procurement/vendor risk management) that can be leveraged for aspects of the development and implementation of the generative AI policies?

The answers to these questions should be in line with institutional values, mission statements, and codes of conduct. This alignment will facilitate the adoption and implementation of the AI policy.

Understanding Existing Policy Environment

In parallel with identifying what their position on generative AI is (see above), the institution should identify how this position overlaps or not with existing institutional policies. A review of the existing policy environment is suggested to identify where existing policies are aligned and where they conflict with the goals of the new AI policy.

The following are key questions the institution should ask when reviewing their existing policy environment and processes:

- What existing policies, if any, specifically address the use of generative AI? (Dishonesty policies? Ethics codes? Privacy policies? Security policies?)
  - What does the existing policy cover?
    - What specific types of AI or machine learning are addressed in the policy?

        - Which domain or discipline is/are covered?
        - Is the policy aimed at faculty, students, or staff?
    - How were the policies developed?
    - Who developed or most recently revised the existing policies?
- Can existing academic dishonesty policies cover the use of generative AI by students?
- Which policies will need to be revised or rewritten to align with a new master policy (see below)?
- What processes may need to be altered in the face of new policy guidelines?

The answers to these questions will help the institution minimize conflicts between existing and new policies and avoid duplicating efforts.

Program and Policy Development

At the same time as defining its position on (generative) AI and understanding its policy environment, the institution should focus on developing a consistent program that helps them regulate their approach on this subject. It is important to keep in mind that the program should reflect the institutional position on (generative) AI and that its implementation should be aligned with existing policies. It should also be an invitation to other departments or specific areas within the institution to be a part of the policy and program and avoid duplicating efforts and having different approaches to the same concern. It may be better to start small and then incrementally build out the program to ensure that specific policies can be implemented and applied quickly. Detailed and complex policies will require more time to be agreed upon, approved, and implemented.

We recommend that each institution include all matters relevant to them.

The following are some matters that we believe should be included in each institution's AI program, and the way each of these matters is addressed will depend on each institution's needs and priorities. It might be one single policy or different policies. Also, please note that each institution is unique and as such, these matters may or may not be relevant. We recommend that each institution include all matters relevant to them.

- Governance (adoption of AI principles)
    - Identification and mitigation of bias
    - Transparency, accountability, privacy, and security
    - Equal access to AI tools for all users
- Teaching and Learning
    - Use of alternative assessments
    - Instructor agency, including non-use of AI in courses
    - Academic integrity in the use of AI
    - Attribution and intellectual property requirements of AI-generated work
- Operational and Administrative
    - Training and "AI literacy" needs

  - Appropriate use of AI in engaging directly with students via chatbots or other automated services
  - Potential changes in staffing that may be required
- Copyright and Intellectual Property
  - Is the approach to generative AI creating copyright and intellectual property risks?
    - If so, how can these risks be mitigated?
  - How is the institution protecting themselves against copyright and intellectual property infringements?
- Research
  - Allowable uses of AI assistance in research functions
  - Attribution and intellectual property requirements of AI-generated work
- Academic Dishonesty
  - How can students reference AI-generated work?
  - Are generative AI tools permitted as a class tool?
  - Will/should the approach on generative AI tools change evaluation/ assessment criteria?
- Policy continuous update/revision
  - How frequently should this policy be updated/revised?
  - Who should be part of the update/revision committee?
  - How are the updates/revisions going to be communicated?
- What are the consequences of non-compliance with the policy?
  - Who will be monitoring and enforcing the policy?

Getting to a final version of a policy is an iterative process, and keeping the stakeholders engaged and providing feedback remains important at this stage. The final policy should strive to balance addressing potential risks with enabling innovation and experimentation without being overly prescriptive by using a risk-based approach.

Implementing the Program

Once the policy is finalized, institutions should move to the implementation phase. At this point two things are critical: first, a written policy is not enough. It requires governance, training, and other processes to ensure the policy is properly operationalized. Second, the responsibility for the implementation and overall program needs to be clearly assigned at this point, with roles and responsibilities of stakeholders defined. When implementing, we encourage institutions to follow their defined process for policy implementation. Nonetheless, the following are points to keep in mind when implementing the new institution-wide AI policy:

- Implementation time frame
  - What is the official effective date of the policy?

    - Is there a phase-in or is it effective immediately (for instance, having a vendor review process may require extra time)?
- Communications
    - Establish a communication plan
        - Who are the audiences?
        - What channels will be used (email, print, social, websites, etc.)?
        - Are the messages clear and consistent?
        - What is the messaging frequency?
        - Are communications multidirectional?
            - Are there opportunities for questions and feedback?
- Training
    - Establish a training plan
        - Identify the different audiences (e.g., do IT staff need more detailed training, do instructors need different training?)
        - Are there existing training tools and processes that can be leveraged? (e.g. legal/compliance or security training)
        - Is there third party-provided content that the institution can use, or is it better to internally develop the training?
- Monitoring
    - Who will monitor the adoption of the policy? (This should be the function assigned with responsibility of the program)
    - How will non-conformity situations be addressed?

We hope this framework assists you and your institution in the adoption and development of AI policies and procedures. Remember that each institution has distinct needs, different structures, and unique goals. This document is intended to help each institution through their policy adoption by providing them with different matters which might be useful for them to consider.

We appreciate any feedback you might have. Please feel free to reach out to trustworthy-ai@anthology.com

You can access the **AI Policy Framework Blog and Resources** now.