

Appendix 2: Goals and Guidance for the AI Ethics Board

Before establishing an AI ethics committee to create a set of AI ethics principles, decide what the board will be expected to accomplish, and the guidance to give them. Preparing this in advance will help the board of directors recruit members for the ethics board, and select members who can achieve these goals. The following is a list of goals and guidance for the board of directors to consider when establishing the committee.

Before beginning

Before establishing an ethics board, the board of directors should discuss the following:

- **AI ethics concerns:** What issues does the board of directors wish the AI ethics board to address? What needs – can be articulated to the AI ethics board members?
- **Readiness:** How ready is the organization to address ethical issues as they pertain to technology? Who is aware of AI ethics issues among the management team and employees? Have people in the organization begun to experience AI's ethical quandaries? How well does the company understand and address the ethics issues it has faced up to this point?
- **Existing boards:** Does the organization currently have an ethics board? If yes, how should the AI ethics board work with the existing board as it draws up a set of AI ethics principles?
- **Permanence:** Are AI ethics best served by a permanent body devoted to those issues, a permanent ethics body with responsibilities for other ethics issues, or a temporary task force? How will compliance with the AI ethics code be effectively monitored and enforced in the future?
- **Independence:** What steps should the board of directors and management take to ensure the AI ethics board maintains its independence? In addition to its selection of the board (see Appendix 3), how should the board of directors ensure that the AI ethics board members will stay independent?
- **Resources:** Will the AI ethics board require a budget and its own staff to carry out its work? How large a budget should be provided, and from what funds?

Goals

- **Recommend whether to update existing codes, develop a new code specifically for AI, or a broader code covering technology and data.** This includes identifying whether codes now exist, evaluating their effectiveness, and whether AI deployment requires a code for the entire company or just a small number of individuals.
- **Develop an actionable set of principles.** While broad and simple principles are memorable, they should be specific enough to provide meaningful guidance for people making real and often difficult decisions. This includes translating broad principles to specific guidelines for different kinds of employees, managements and suppliers.

- **Recommend benchmarks and performance indicators.** This includes recommending which ethical AI design, development and research standards the organization should follow, key performance indicators of ethical behaviour, and which ethics benchmarks or rankings to measure the organization against.
- **Suggest how to help people learn and comply with the code.** Recommend ways to make the code easier and more convenient to understand and apply, such as ethics hotlines and advisers, ethics training and online ethics chatbots.
- **Recommend how frequently the code should be updated.** Suggest a minimum period that should pass before revisiting the code, and triggers that would require a code update before that period ends. Triggers could include new laws and regulations, the emergence of new AI technologies and capabilities, the development or purchase of new AI systems, new labour agreements and acquisitions of companies. The ethics board should also ask whether the organization is equipped to update standards, and what needs to be done to improve readiness.
- **Recommend whether to appoint an ethics officer specifically for AI.** Consider whether the company should establish such a role rather than appoint a chief AI ethics officer or add AI to the current ethics officer's portfolio. If so, recommend what their responsibilities would be and how to judge their success.
- **Recommend how to resolve potential conflicts between principles.** The ethics board could identify situations where principles may conflict with laws or other principles, who would be affected, and recommend a course of action for reconciling or choosing between the conflicting elements. These may include conflicts between employee and customer protections or between human rights standards and local laws.
- **Identify any parts of the code that should not be made public.** Consider whether there are parts of the code that would violate agreements with suppliers or customers if they were made public.
- **Recommend how the code should be implemented.** This includes who should review the code besides the board and management team, whether the code should be piloted before being finalized, how to educate people about the code, and creating a timeline for roll-out and implementation.
- **Suggest whistleblowing mechanisms:** This includes avenues by which employees, contractors, customers and providers can safely report ethics violations and problems.

Guidance

- **Support the organization's principles and goals.** The recommendations of the AI ethics board should be consistent with the organization's principles and support the goals of the board of directors in terms of: organizational survival; achievement of the organization's mission; increasing the value created by the company; compliance with laws and regulations; meeting mandated obligations.
- **Support the greater good.** Consistent with the World Economic Forum's commitment to improving the state of the world, ethics boards should also consider including principles within the AI ethics code that go beyond the organization's self-interest and benefit humanity and the planet as well.

- **Respect human rights:** Consistent with the UN Guiding Principles on Business and Human Rights, the AI ethics code:
 - should help companies avoid causing or contributing to adverse human rights impacts through their own activities
 - should seek to prevent or mitigate such impacts when the enterprise is directly linked to them via a business relationship
 - should engage in ongoing due-diligence processes to identify, prevent and mitigate salient human rights risks
 - should provide remediation for human rights impacts through legitimate mechanisms when adverse human rights impacts do occur
- **Be consistent:** Ethics codes should be internally consistent. The principles of the code should follow the same logic, support rather than undermine one another and be applicable to all. The provisions of the ethics code should apply definitions equally, and guiding principles should support one another. The code should be applicable to all who are affiliated with the organization: no leaders, staff or AI users should be exempt. Ask whether different groups of employees or the public would find them to be inconsistent.
- **Meet external standards:** At a minimum, the codes should meet the ethical standards set forth in law, and of the most important customers. The code should also meet or exceed the standards of the company’s industry associations, its vendors and partners, and the professional associations of its employees – or have defensible reasons if it does not.
- **Be defensible:** Codes should be able to withstand public scrutiny. They should be written so that they are as transparent as possible to the public. Even sections that are judged as inappropriate for release (for instance, for financial reasons) ought to be defensible if they were made known.
- **Be clear and easy to use:** The code should be clearly written without requiring technical knowledge. It should also be organized so that any reader can easily find sections pertinent to their situation. The committee should also consider how to make the code more convenient to implement.
- **Protect the vulnerable:** Children can succumb to addictive behaviour when algorithms learn how to gain and retain their attention; unlike adults, they are not able to learn why and thus change their behaviour. Minorities can lose out on economic opportunities or be treated unfairly due to biases in algorithms and datasets. Ethics codes should protect these and other vulnerable populations.
- **Seek outside voices:** The ethics committee should seek out the views of people who are or will be affected by the organization’s use of AI. This may require inviting outsiders to address meetings, holding open hearings where the public can address the ethics committee and soliciting written statements.
- **Consider local conditions:** Codes must be operable wherever the organization conducts business. While difficult to achieve, ethics guidelines for AI should be flexible enough to accommodate local cultures and laws without compromising core principles. Ethics boards should be prepared to suggest how to find the best balance for their organization.
- **Consider industry requirements:** Similarly, codes must be operable within the company’s business. They should be written so they are plainly applicable to work processes and the specific kinds of decisions made in the company. Ethics codes must be especially careful to distinguish discriminatory and non-discriminatory activity – for example, AI decisions that legitimately deny services to a member of a group from decisions that discriminate due to deliberate or inadvertent bias against that group.”