The Code of Data Ethics

Authors: Katia Bonacina, Enrico Panai
Last updated: 05/21/2024

This Body of Knowledge (BoK) provides an overview of the fundamental principles and characteristics of a Code of Data Ethics within professional and organizational contexts

Introduction to ForHumanity

ForHumanity's purpose is to mitigate downside risks posed by AI, algorithmic and autonomous systems to humans. ForHumanity is a 501 (c) 3 tax-exempt public charity. ForHumanity endeavours to be a beacon, examining the impact of AI & Automation on jobs, society, our rights, and our freedoms. We focus on mitigating risk in the areas of ethics, bias, privacy, trust, and cybersecurity (EBPTC) at the corporate and public policy levels. Always on behalf of humanity.

1. **Definition**

A **Code of Data Ethics (CoDE)** is a set of principles, definitions, procedures, methods, and practices that helps develop ethical reasoning on issues involving data, information, algorithms, and digital infrastructure. The **Code of Data Ethics** provides the enterprise with the tools to align the values and ethical principles expressed in its **Code of Ethics** with practice.

The **CoDE** is built upon the principles of **information ethics**, emphasizing the integrity and responsible management of data, and it is especially addressed to those who deal with data governance (data stewards and data custodians) and those who directly manipulate and wrangle data (data workers, data scientists, developers).

0. **Entities in charge**

The document is drafted and kept up to date by the **Chief Data Ethics Officer (CDEO)** and the **Data Ethics Committee (DEC)**. In organizations with advanced AI capabilities, the CDEO may collaborate with a **Chief AI Ethics Officer (CAIEO)**—an AI ethicist with strong statistical expertise—to ensure ethical oversight of AI systems.

The chief ethics officer and/or the ethics committee are accountable for the company's code of ethics and assist the CEO and the board of directors in formalising the organisation's ethical principles and values. While the chief ethics officer works on a higher level of ethics addressing both external societal responsibilities and internal employee relations, other roles focus on specific aspects of data governance. The **Chief Data Officer (CDO) or Data Protection Officer (DPO)** ensures compliance with legal regulations, such as the **GDPR (EU), CCPA (California), and the EU AI Act (2024/1689)**, while **data workers** adopt a technical and performance-driven approach to data and algorithms.

To align ethical and legal responsibilities, the **ethical framework for data** must comply with relevant legal standards. The **CEO and board of directors** should be aware of its implications and liabilities, while the **CDEO** ensures its consistency with the broader code of ethics. This framework requires **annual re-endorsement** or updates when necessary.

From an operational point of view, the code of data ethics applies to all data workers, both employees and external collaborators, who collect, store, process, use, sell, exploit or disseminate data.

The **Chief Data Ethics Officer** plays the critical role of creating and fostering this ethical environment by:

- setting and maintaining the code of data ethics,

- ensuring its internal coherence,

- promoting its dissemination to data workers and its adoption at the top-level,

- explicitly defining the decision-making perimeter of data ethics.

Additionally, the **CDEO** should work alongside an **Algorithmic Risk Committee (ARC)** to assess the potential risks of AI algorithms and autonomous systems before deployment. The **ARC** is an ad hoc committee that decides on the production and reviews the performance of the organisation's AI systems.

0. **Goals of the CoDE**

The **CoDE** uses a **Level of Informational Abstraction (LoA)** in order to achieve the following goals:
i) to shape **governance**,
ii) perform ethical analyses of systems (Axiological Analysis),
iii) predict possible **ethical risks** (Ethical Analysis Forecasts), and
iv) allocate responsibility (distributed morality) for actions generated by complex agents (**Multi Agent System**) in the space to which **Data Ethics** applies.

Using LoA ensures that all parties have aligned their perspectives, allowing for clear communication and management. This avoids endless discussions where objectives, agencies and often even languages are adapted to different frames of reference.

This approach ensures that moral responsibilities are properly allocated across human, artificial, and hybrid agents (**Multi-Agent Systems, MAS**). The framework consists of four core areas:

. **Ethics of Practices** – Establishes governance structures, defines roles, and ensures ethical decision-making in data handling.
. **Ethics of Data** – Addresses the ethical collection, processing, and management of data, including issues related to bias, privacy, and transparency.
. **Ethics of Algorithms** – Evaluates the ethical implications of AI and algorithmic decision-making, ensuring fairness, accountability, and explainability.

A **Code of Data Ethics** serves to drive the **Digital Governance** of an organisation by:
- helping to shape the processes and methods used by data stewards and data custodians;
- defining strategies for allocating moral responsibility for the consequences of handling data and information;
- applying the principles expressed in the code of ethics (by developing satisfactory axiology to solve problems related to data, information, algorithms, infrastructures, and practices);
- highlighting the limits of its implementation;
- using ethics as an innovation trigger and analysis tool to avoid opportunity costs by enhancing ethical reasoning.

0. **The CoE and the CoDE**

In order to function as intended, a reasonable code must, at the very least, draw on the principles of the organisation in order to support axiological analyses and moral assessments of digital governance and compliance with digital regulation. Therefore, it is necessary that the **Code of Data Ethics** is preceded or accompanied by a sufficient or mature **Code of Ethics.**

In a nutshell, the **Code of Ethics** identifies business values and principles, while the **Code of Data Ethics** provides the theoretical framework and operational tools to be able to govern digital processes, from data to autonomous artificial intelligence systems. It is essential for an appropriate use of **Digital Ethics** in an organisation. One cannot imagine applying it without an adequate and explicit business ethics framework already defined in the organisation's **Code of Ethics**.

In the absence of a **Code of Ethics** or a set of principles that are not clearly shared or defined, this code may be inspired by the **Unified Framework of Ethical Principles for AI** (Floridi et al., 2018):

- Beneficence: understood as "promoting well-being, preserving dignity, and sustaining the planet";
- Nonmaleficence: understood as "operating 'within secure constraints'", preventing "of infringements on personal privacy", enhancing "cautions against the various negative consequences of overusing or misusing AI technologies";
- Autonomy: understood as "the power to decide to decide", in its three forms (decide, to decide-to-delegate, decide to decide again);
- Justice: understood as "promoting prosperity, preserving solidarity, avoiding unfairness";
- Explicability: understood as "incorporating both the epistemological sense of intelligibility—as an answer to the question 'how does it work?'—and in the ethical sense of accountability—as an answer to the question 'who is responsible for the way it works?'".

The **Code of Data Ethics (CoDE)** consists of two parts, a Public Part on general principles and methodology and a Private Part, where the values and approaches used at each level of the organisation and each stage of the pipeline are explained. The Private Part concerns the operational approach to data ethics. Although it can be published as an explanatory supplement to the Public Part, the Private Part should remain an internal operational document and thus distributable within the organisation that is to use it. The Private Part is necessary to guarantee the ethical operability of the processes.

5. Criteria for a good CoDE:

N.B. 'Must' statements = sufficient CoDE, 'Should' statements = mature CoDE

1. Must be compatible with, if not explicitly affirm, those values which are fundamental to the **Independent Audit of AI Systems (IAAIS)**. These include transparency, accountability, and trust.
2. Must be divided into two parts: the Public Part, containing the principles, and the Private Part, the more operational document.
3. The Public Part must be publicly disclosed.
4. The Private Part must be distributed within the organisation that is to use it.
5. Must affirm that the entity also has a **Code of Ethics**.
6. Must outline the relation between the **Code of Ethics** and the **Code of Data Ethics** as well as their respective scope and purpose.
7. Must clearly define the shared moral framework behind the CoDE in the Public Part.
8. In the Private Part the Organization should define all the procedures, definitions, and policies it is using to manage data.
9. Must be compatible with the values of **transparency**, **accountability**, and **trust.**
10. Must be **transparent** i.e., sufficiently detailed and precise in its formulation of a shared moral framework to allow for stakeholders to reasonably judge whether or not the entity and/or its employees is/are abiding by said moral framework.
11. Must be drafted, kept up to date, and overseen by the entity's **Chief Data Ethics Officer (CDEO)** and the **Data Ethics Committee (DEC).**
12. Those responsible for the **Data Ethics** must have a well-established competence in digital ethics.
13. Must take into account/address **Relevant Legal Frameworks** (show that the entity takes relevant risks into account):

    13.1 Law as it applies to Data Subjects (specific to the jurisdiction of Data Subject)

13.2 Consideration of Human Rights, equalities, anti-discrimination law, access to goods and services, children's law, laws regarding AI/data processing platforms or sector

13.3 Other relevant risks (personal gain, sustainability, etc.)

0. Must consider the principles of **Data Security:** an environment suitable for ethical reasoning cannot be developed if data security is not guaranteed first.
0. Must be sufficiently detailed to give context to **Ethical Choice** (explicitly = preferred because instances must be disclosed).
0. Must allow for **accountability** through trust as a consequence of **transparency** (as outlined above).
0. Must be **internally coherent** (must not contradict itself).
0. Must consider the entity's **Diversity Policy**, it being the constructive foundation to avoid reinforcing biases due to uniformity of view, thus reducing the risk of AI, algorithmic or autonomous systems producing negative consequences (refer back to the Code of Ethics for diversity policy details).
0. Must be contained in the **Employee Handbook.**
0. Should divide the Private Part to cover three main areas:

20.1 **Ethics of Practices**

20.2 **Ethics of Data**

20.3 **Ethics of Algorithms**

0. Should state the rules and the policies of the organization in the Ethics of Practices to understand how the roles should be distributed inside the organization.
0. Should set out the principles and core values in data management, indicate the chain of responsibility, and regulate the use of data and algorithms as much as possible.
0. Should contain a commitment to making employees aware of the **Code of Ethics** and its presence in the **Employee Handbook**.
0. Should explicitly affirm the values of **transparency**, **accountability**, and **trust.**
0. Should contain a framework or set of procedures to guide ethics officers in determining the source of reasonably diverse inputs and multi-stakeholder feedback as well as how to judge whether or not these inputs are reasonable in and of themselves i.e., the extent to which they should be heeded or incorporated.

FAQs

What does [axiological analysis](#) mean?

Axiology itself is the study of values—things like fairness, honesty, privacy, and human dignity—that shape our moral judgments. Axiological ethics is that portion of ethics that is concerned specifically with these values. It centres on questions of what is worth pursuing or promoting and what should be avoided, along with issues of what such questions mean.

What is Ethics of Practices?

Ethics of Practices including **Professional Ethics** and **Deontology** address the pressing questions concerning the responsibilities and liabilities of people and organizations in charge of data processes, strategies and policies, including data scientists, with the goal to define an ethical framework to shape a professional code about responsible innovation, development, and usage which may ensure ethical practices fostering both the progress of data science and the protection of the rights of individuals and groups.

What is Ethics of Data?

The **Ethics of Data** focuses on ethical problems posed by the collection and analysis of large datasets and on issues ranging from the use of big data in biomedical research and social sciences, to profiling, advertising and data philanthropy as well as open data. In this context, key issues concern possible re-identification of individuals through data-mining, -linking, -merging and re-using of large datasets, as well as risks for so-called 'group privacy', when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination (e.g., ageism, ethnicism, sexism) to group-targeted forms of violence.

Data ethics is a latent space in companies that operate in informationally mature societies. ethical choices are already being made about data, but that the ethical decision-making process is not explicit: it is not supported by any axiological analysis.

What is Ethics of Algorithms?

The **Ethics of Algorithms** addresses issues posed by the increasing complexity and autonomy of algorithms broadly understood [...], especially in the case of machine learning applications. In this case, some crucial challenges include moral responsibility and accountability of both designers and data.

[What is the difference between a Code of Ethics and a Code of Data Ethics?](#)

The Code of Ethics is a document with company-wide scope that lays out the general values of a business, providing guiding principles for ethical choices that all employees throughout a business must abide by. The Code of Data Ethics, on the other hand, provides detailed guidance on handling data and algorithms in an ethical manner, and may be longer and contain more detail regarding the individual responsibilities of different positions within an entity than the Code of Ethics.

What is the Level of Abstraction method?

The method of levels of abstraction (LoA) makes it possible to choose a good question answer. Already used in mathematics (where it is closely related to model theory), in computer science (mainly in object-oriented programming or OOP), LoA has been used by information philosophy to shift the focus from humans to informational entities.
The method of levels of abstraction allows a point of observation to be chosen. As the level of observation changes, so do the "observables". In practice, specific observables refer to a specific level of abstraction (LoA).
Information ethics changes the level of abstraction compared to classical macroethics and considers all elements as informational entities, emphasising the importance of the moral situation as an envelope within which agents (human, artificial or hybrid) perform actions that impact on recipients or patients (human, artificial, hybrid or passive informational entities)

FH DEFINITIONS

Algorithm Ethics: A sub field of Ethics focused on instances of Ethical Choice emerging from AI, algorithmic and autonomous systems. Training and expertise include areas such as Necessity, Proportionality, Benchmark setting, Validity, Reliability, Concept Drift, and thresholds for Bias mitigation.

Ethical Choice: Awareness of a set of options to be made in the context of automated and intelligent systems, using a set of principles and rules concerning moral obligations and regards for the rights of humans and for nature, which may be specified by a given profession or group. The result, outcome or judgment is made using a shared moral framework or set of moral principles based upon the entity's Code of Ethics.

Relevant Legal Framework: A Relevant Legal Framework can contain a broad range of applicable law such as the laws that govern an entity or organisation, that govern the rights and privileges of a Data Subject, that restrict the activities and behaviours of a Data Controller or Data Processor or put positive obligations upon an entity. These include consideration for human rights, equalities and anti-discrimination law, access

to goods and services (having due regard to who is included/excluded from such goods and services), children's law, and laws with regard to the platform and/or laws with regard to the sector in and through which the AI (and data processing) is being provided.

LINKED FOR HUMANITY UNIVERSITY COURSE

[ALGETH](#)

Algorithm Ethics teaches students the necessary skills and knowledge to adjudicate instances of Ethical Choice associated with AI, Algorithmic and Autonomous Systems. A graduate will earn an Expert accreditation and will be well qualified to sit on standing and empowered Ethics Committees. They will be an expert in understanding, executing and satisfying Independent Audit of AAA Systems compliance requirements.

LINKED BODY OF KNOWLEDGE

Code of Ethics https://forhumanity.center/bok/code-of-ethics/

Bibliography

ISO/IEC/IEEE 24748-7000:2022 - Systems and software engineering — Life cycle management Part 7000: Standard model process for addressing ethical concerns during system design https://www.iso.org/fr/standard/84893.html

Floridi, L. (2018), *Semantic Capital: Its Nature, Value, and Curation*. Philos. Technol. 31, 481–497. https://doi.org/10.1007/ s13347-018-0335-1

Floridi, L., (2018), *Soft Ethics: Its Application to the General Data Protection Regulation and Its Dual Advantage*. Philos. Technol. 31, 163–167. https://doi.org/10.1007/s13347-018-0315-5

Panai, E. (2023). *The latent space of data ethics*. Springer-Verlag London Ltd.