Ethics and Human Rights Guidelines for *Big Data for Development* Research #3

Summary of Review of Existing Codes of Ethics for Big Data and Artificial Intelligence

March 2020

Author: **Amber Sinha and Pranav Bhaskar Tiwari**

Produced as part of the **Big Data for Development** network supported by **International Development Research Centre**, Canada

The Centre for Internet and Society
https://cis-india.org

Note

This is the third part of a four-part review of guideline documents for ethics and human rights in big data for development research. Please read the entire document here: [link]

Contents

1. Introduction

In this third part of our research of guiding ethical principles for big data and AI, we consider the following sets of Code of Ethics articulated in the context of big data, artificial intelligence and computational and statistical sciences: Asilomar AI Principles,[1] IEEE Ethically Aligned Design,[2] ACM Code of Ethics,[3] Computer Ethics Institute,[4] Data for Democracy Principles,[5] Data Science Code of Professional Conduct,[6] IEEE Code of Ethics (COE),[7] Electronic Privacy Information Center Code of Ethics,[8] Information Systems Security Association Code of Ethics,[9] Information Security Consortium Code of Ethics,[10] Internet Architecture Board COE, IFLA Professional Code of Ethics for Librarians,[11] INFORMS Ethical Guidelines,[12] ASA

[1] The Future of Life Institute. Asilomar AI Principles. (2017). Retrieved from https://futureoflife.org/ai-principles/ (Hereinafter Asilomar AI Principles).
[2] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2. IEEE, 2017. Retrieved from http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html (Hereinafter IEEE Ethically Aligned Design).
[3] ACM Ethics. ACM Code of Ethics and Professional Conduct. (2018). Retrieved from https://ethics.acm.org/ (Hereinafter ACM Code of Ethics).
[4] Computer Ethics Institute. Ten Commandments of Computer Ethics. Retrieved from http://computerethicsinstitute.org/publications/tencommandments.html (Hereinafter Ten Commandments of Computer Ethics).
[5] Data for Democracy Global Data Ethics Project. Global Data Ethics Principles. (2018).
[6] Data Science Association. Data Science Code of Professional Conduct. Retrieved from http://www.datascienceassn.org/code-of-conduct.html (Hereinafter Data Science Code of Professional Conduct).
[7] IEEE. IEEE Policies, Section 7 - Professional Activities (Part A - IEEE Policies)-7.8 IEEE Code of Ethics. Retrieved from https://www.ieee.org/about/corporate/governance/p7-8.html (Hereinafter IEEE Policies).
[8] Electronic Privacy Information Center. The Code of Fair Information Practices. Retrieved from https://epic.org/privacy/consumer/code_fair_info.html (Hereinafter EPIC Code of Fair Information Practices).
[9] Information Systems Security Association. ISSA Code of Ethics. Retrieved from https://www.issa.org/page/CodeofEthics (Hereinafter ISSA Code of Ethics).
[10] Information Security Consortium. (ISC)2 Code Of Ethics Retrieved from https://www.isc2.org/Ethics (Hereinafter (ISC)2 Code Of Ethics).
[11] International Federation of Library Associations and Institutions. IFLA Code of Ethics for Librarians and other Information Workers. (2016). Retrieved from https://www.ifla.org/publications/node/11092 (Hereinafter IFLA Code of Ethics for Librarians).
[12] The Institute for Operations Research and the Management Sciences. INFORMS Ethics Guidelines. Retrieved from https://www.informs.org/About-INFORMS/Governance/INFORMS-Ethics-Guidelines (Hereinafter INFORMS Ethics Guidelines).

Ethical Guidelines for Statistical Practice,[13] League of Professional Systems Administrators,[14] AI at Google Principles,[15] and Microsoft AI principles.[16] This was based on a global review of such attempts. While this list is not exhaustive, we have attempted to represent different kinds of stakeholders engaging in such exercises: Private sector, educational and academic institutions, global bodies representing professionals, civil society and states. Similarly, we have tried to represent attempts from different related disciplines such as Data Science, Computing, Electronics & Electrical Engineering, Knowledge Theory, Statistics, Operations, Systems Administration, Information Theory and Artificial Intelligence.

Each of these guidelines were analysed for their adherence to well recognised principles from the fields of normative ethics, applied ethics and law. Based on this exercise, we have both high level conclusions about representations of principles across guidelines, as well as analysis of reasons for preferences made. While there have been several attempts at arriving at guidelines to govern the use of data and algorithms, and some critical scholarship analysing them, the attempt behind this exercise is to provide a frame to recognise and evaluate them.

[13] American Statistical Association. Ethical Guidelines for Statistical Practice. (2018). Retrieved from https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx (Hereinafter ASA Ethical Guidelines for Statistical Practice).
[14] The League of Professional Systems Administrators. Code of Ethics. Retrieved from https://www.lopsa.org/CodeOfEthics (Hereinafter LOPSA Code of Ethics).
[15] Pichai, S. (2018, June 7). AI at Google: our principles. Google. Retrieved from https://www.blog.google/technology/ai/ai-principles/ (Hereinafter Google AI principles).
[16] Microsoft. Microsoft AI principles. Our Approach to AI. Retrieved from https://www.microsoft.com/en-us/ai/our-approach-to-ai (Hereinafter Microsoft AI principles).

2. Broad Conclusions

Out of the eighteen Ethical Codes of Conducts examined, seven of them[17] had a broadly Utilitarian approach,[18] six of them[19] subscribe to a Consequentialist approach,[20] five[21] draw from a Deontological approach[22] while four of them[23] can be seen to have Kantian leanings.[24] Seven out of the Eighteen Codes adhered to more than one normative ethical theory, whereas five[25] did not clearly align with any of the four ethical theories mentioned above.

Out of the three Ethical Codes on AI, two aligned with Utilitarianism,[26] two with Consequentialism,[27] two with Deontology[28] and 1 with Kantianism.[29] Microsoft aligned with three normative ethical theories with the exception of Utilitarianism.

---

[17] Asilomar AI Principles, Principle 14; Google AI principles, Objective 1 Be socially beneficial; ACM Code of Ethics Section 1.1 and 1.2; Global Data Ethics Principles, FORTS Framework, Social Benefit; EU General Data Protection Regulation permits collection of sensitive data for public good; IAB Ethics and the Internet Policy Statement, Point (c) and; ASA Ethical Guidelines for Statistical Practice, Purpose of the Guidelines.

[18] Utilitarianism entails that an action is ethical if it increases utility by maximizing happiness or minimizing harm for maximum people.

[19] Google AI principles, AI applications we will not pursue; Microsoft AI principles, Responsible bots guidelines; ACM Code of Ethics, Section 1.2; Ten Commandments of Computer Ethics, Commandment 9 and 10; Global Data Ethics Principles, Principle 10 and; ASA Ethical Guidelines for Statistical Practice, Preamble.

[20] Consequentialism entails that an action is ethical if its consequence is morally upright. Actions whose consequences merely increase utility are not encompassed in consequentialism for the purposes of this research, and are considered a part of utilitarianism.

[21] Asilomar AI Principles, Principle 10 and 11; Microsoft AI principles, requiring that AI systems should treat all people fairly; Data Science Code of Professional Conduct, Part 2 [Data Scientist - Client Relationship] emphasizes on the duties of the data scientists while pointing out limited circumstances where derogation is permissible and; ISSA Code of Ethics, Point 5.

[22] Deontology entails that an action is ethical if it is in consonance with ones' duties.

[23] Microsoft AI principles, requiring that AI systems should treat all people fairly; IEEE Ethically Aligned Design, Principle 1; Ten Commandments of Computer Ethics, Commandment 1 and 4 and; ASA Ethical Guidelines for Statistical Practice, Responsibilities to Research Subjects.

[24] Kant, E. (2012) Groundwork of the Metaphysics of Morals (1785). Revised Edition. Ed. M. Gregor and J. Timmermann. Cambridge: Cambridge University Press. Kantianism entails that every action in violation of a Categorical imperative, i.e., a universal moral principle, is unethical.

[25] (ISC)2 Code Of Ethics; IFLA Professional Code of Ethics for Librarians; INFORMS Ethical Guidelines and; IAB Ethics and the Internet Policy Statement.

[26] Asilomar AI Principles, Principle 14 and; Google AI principles, Objective 1 Be socially beneficial.

[27] Google AI principles, AI applications we will not pursue; Microsoft AI principles, Responsible bots guidelines.

[28] Asilomar AI Principles, Principle 10 and 11; Microsoft AI principles, requiring that AI systems should treat all people fairly.

[29] Microsoft AI principles, requiring that AI systems should treat all people fairly.

Out of the three Ethical Codes on Data Science, two aligned with Utilitarianism,[30] 1 with Consequentialism,[31] 1 with Deontology[32] and none with Kantianism.

3. Reflections on Ethical Principles

Asilomar AI Principles purport a deontological approach[33] but the ambiguous nature of the duties entrusted may render their existence futile considering that they might not always lead to certain results while resolving ethical dilemmas. Microsoft AI principles cast a fundamental duty on its developers that AI systems should treat all people fairly. This duty not only renders the Code deontological in nature but this duty also has a universal moral character and accordingly rises to the level of a Categorical Imperative. But what does treating fairly imply? While framing the Code the drafters did not consider elaborating on the duty entrusted.

The above malady of mitigating ambiguity in understanding a duty is well tackled in the IEEE Ethically Aligned Design. Herein, the Code not only mandates the observance of a Categorical Imperative (respecting Human Rights)[34] but also provides Guidelines on the practical application of the principles. For instance the code refers to the Ruggie principles[35] which provide methods to pragmatically implement human rights ideals within business or corporate contexts that could be adapted for engineers and technologists. The IEEE Ethically Aligned Design is well referenced with similar guidelines for resolving practical scenarios. Most of the codes reviewed are open ended and do not certainly/unambiguously define what is expected, this is where the role of additional guidelines to resolve practical issues kicks in. But the issue with additional guidelines is that it renders the codes too heavy to be absorbed. For instance the IEEE Ethically Aligned Design, though a very instructive resource, runs into 266 odd pages and is referenced with additional

[30] Global Data Ethics Principles, FORTS Framework, Social Benefit.
[31] Global Data Ethics Principles, Principle 10.
[32] Data Science Code of Professional Conduct, Part 2 [Data Scientist - Client Relationship] emphasizes on the duties of the data scientists while pointing out limited circumstances where derogation is permissible.
[33] Asilomar AI Principles, Principle 10 and 11.
[34] IEEE Ethically Aligned Design, Principle 1.
[35] Id . at 25.

documents for guidance. In such circumstances rigorous training sessions would help the employees appreciate the code and understand their ethical obligations.

The ACM Code of Ethics also provides additional attachments which explain the practical application of codes for troubleshooting in different circumstances like for dealing with Malware Disruption, Medical Implant Risk Analysis, Abusive Workplace Behavior, Autonomous Active Response Weapons, Dark UX Patterns, Malicious Inputs to Content Filters etc.

The ASA Ethical Guidelines for Statistical Practice is a good example of a balanced code. Its Preamble provides for avoiding harm (utilitarian) and describes using statistics in pursuit of unethical ends as inherently unethical (Consequentialist). It further delineates the statistician's Responsibilities towards Research Subjects (towards humans) which is of a Kantian nature.

Only one Code confers rights to individuals.[36] Two of the Codes provide for sanctions for violation of the Code of Conduct,[37] whereas none of them provides for a legal remedy and or notes on compensation.

Sixteen of the Codes are only substantive in nature, whereas Information Security Consortium COE is both substantive and procedural.

Both Asilomar AI Principles[38] and Google AI Principles[39] categorically prohibit indulgence in an arms race in lethal autonomous weapons. The IEEE Ethically Aligned Design throws light on the various issues in relation with the use of AI in autonomous weapons and recommends how they may be resolved.[40]

Virtue and applied ethics feature prominently across the different codes of ethics. Social Benefit consideration[41] was evident in fourteen of the Codes.[42]

[36] IEEE Ethically Aligned Design, Individual Rights in Section 2 — 'Governmental Use of A/IS: Transparency and; Individual Rights'.
[37] ACM Code of Ethics Section 4.2; and (ISC)2 Code Of Ethics, disciplinary action.
[38] Asilomar AI Principles, Principle 18.
[39] Google AI principles, AI applications we will not pursue.
[40] IEEE Ethically Aligned Design, Future Technology Concerns (Reframing Autonomous Weapons).
[41] The principle of Social Benefit encompasses a utilitarian approach envisages actions to be ethical if they benefit the society.

The principle of benevolence[43] was a part of only the IFLA Professional Code of Ethics for Librarians on whom there is a responsibility to make provisions so that the less privileged also get access to information.[44]

Eleven Codes of ethics[45] align with the Principle of Harm[46] which illuminates the increasing inclination towards utilitarianism. Nine of the Codes[47] align with the Principle of Honesty.[48]

Seven of the Codes[49] assert on the actions of the practitioners being legal, while twelve Codes require that the practitioners ensure that their work is Reliable and Safe. Fourteen Codes posit that the Privacy of the subjects should and clients must be ensured.

There are 11 Codes which have deliberated on the Security, which is a paramount concern for use of AI and Data Scientists and five out of six codes on AI and Data

[42] Asilomar AI Principles, Principle 14 and 1; Google AI principles, Objective 1 Be socially beneficial; ACM Code of Ethics Section 3.1 and 3; Global Data Ethics Principles, FORTS Framework, Social Benefit; EU GDPR, Article 1- data is collected to maintain law and order (control and regulate crime and criminals) and for preventing threats to public security; IAB Ethics and the Internet Policy Statement, Introduction; IEEE Ethically Aligned Design, Objectives and Well-being Metrics; Ten Commandments of Computer Ethics, Commandment 9; IEEE Policies, Rule 1; (ISC)2 Code Of Ethics, Canons; INFORMS Ethics Guidelines, Principle 1; IFLA Code of Ethics for Librarians, Section 1; IAB Ethics and the Internet Policy Statement, Introduction; ASA Ethical Guidelines for Statistical Practice, Purpose of the Guidelines and; LOPSA Code of Ethics, Education & Social Responsibility.
[43] The principle of Benevolence envisages helping those in need as opposed to helping everyone.
[44] IFLA Code of Ethics for Librarians, Section 1. Access to information- Responsibility to make provisions so that the less privileged also get access to information.
[45] Google AI principles, AI applications we will not pursue; ACM Code of Ethics Section 1.2; Ten Commandments of Computer Ethics, Commandment 1; Global Data Ethics Principles, FORTS Framework- Social Benefit; Data Science Code of Professional Conduct, Rule3(b) & 8(g); EU GDPR, the declaration is formulated to protect the society thus the corollary encapsulates the principle of harm; IEEE Policies, Rule 9and 1; ISSA Code of Ethics, Rule 6; IAB Statement of Policy, Point (c) & (d); INFORMS Ethics Guidelines, Article 2. Our Organisation and; ASA Ethical Guidelines for Statistical Practice, Purpose of the Guidelines.
[46] The principle of Harm considers an action as unethical if it causes harm.
[47] ACM Code of Ethics, Section 1.3; Ten Commandments of Computer Ethics, Commandment 5; Data Science Code of Professional Conduct, Rule 8(f); IEEE Policies, Rule 3; ISSA Code of Ethics, Rule, 5; (ISC)2 Code Of Ethics; Canon 2; INFORMS Ethics Guidelines, Principle 1; ASA Ethical Guidelines for Statistical Practice, Professional Integrity and Accountability; and LOPSA Code of Ethics, Personal Integrity.
[48] The principle of Honesty requires one to be honest and truthful.
[49] Google AI principles, AI applications we will not pursue 4; IEEE Ethically Aligned Design, Goals & Objectives; ACM Code of Ethics, Section 2.3; Data Science Code of Professional Conduct, Rule 8 (e) & 9(c); Article 4(1)(a) and Article 8; ISSA Code of Ethics, Rule 1; (ISC)2 Code Of Ethics, Canon 2; and LOPSACode of Ethics, Laws and Policies.

Science have provisions for ensuring security. Seven Codes make sure that they are Inclusive and do not discriminate.[50]

In our analysis, nine Codes are seen to have deliberated on Transparency which is another important concern for use of AI and Data Scientists, whereas eight of them assert on the need of accountability. 4 out of 6 codes on AI and Data Science adhere with the principle of accountability.

[50] Inclusivity entails assimilation of all and non-discrimination on parameters like caste, creed, sexual-orientation, gender, race etc.