Ethics and Human Rights Guidelines for *Big Data for Development* Research #2

Review of Principles of Ethics in Computer Science

March 2020

Author: **Rajashri Seal and Amber Sinha**
Editor: **Sumandro Chattapadhyay**

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

Contents

1. Introduction

The rapid expansion in the volume, velocity, and variety of data available, together with the development of innovative forms of statistical analytics, is generally referred to as "big data";[1] though there is no single agreed upon definition of the term. Big data promises to provide new insights and solutions across a wide range of sectors. Despite enormous optimism about the scope and variety of big data's potential applications, many remain concerned about its widespread adoption, with some scholars suggesting it could generate as many harms as benefits.[2] The predecessor disciplines of data science such as computer sciences, applied mathematics, and statistics have traditionally managed to stay out of the scope of ethical frameworks, based on the assumption that they do not involve humans as subject of their research. While critical study into big data is still in its infancy, there is a growing belief that there are significant discontinuities between the rapid growth in big data and the ethical framework that exists to govern its use. In this document we look at them in detail.

The role of ethical frameworks in the discipline of computer science has been debated extensively. Due to the assumption that this discipline does not have clear impacts on human subjects of research, it has generally avoided establishing institutional ethical frameworks. However, in a few decades old history of these disciplines has been peppered with attempts to articulate ethical frameworks that could govern their practice and research. The following sections in this document document and analyse key frameworks that have been put forth in the domain of computer ethics. Unlike the biomedical sciences, the computational sciences have only seen nascent evolution of ethical frameworks, with limited literature available on conflicts, practical application and critiques of the frameworks.

[1] Viktor Mayer Schoenberger and Kenneth Cukier, Big Data: A Revolution that will transform how we live, work and think" John Murray, London, 2013.
[2] Raghupathi, W., & Raghupathi, V. Big data analytics in healthcare: promise and potential. Health Information Science and Systems, (2014).

2. Gotterbarn's professional ethics

The evolution of professional ethics are often driven by in part, fears[3] about the impact of that discipline, and in other part, by the desire to have minimal standards of practice.[4] At times, it is also guided by pressures from society which seeks to ensure that its interests are safeguarded. The impact of computational sciences on the global society has grown significantly over the last century, and exponentially in the recent years. However, certain attributes of computer sciences, such as the facts that it can be joined without any professional course or that it does not have a representative licensing body, make the process of developing a code of professional ethics a difficult task.

The shift from the traditional conceptualisation of a regulatory mode of ethics to the normative mode in the fields of computer science and scientific computing needs to be taken into account. These norms were evolved as merely ethical rules of thumb which are to be followed. This shift can be witnessed from the code revision of institutions such as IEEE, ACM and BCS.[5] Taking from this, Gotterbarn tries to draw a new understanding of the working of ethics in a profession. This is known as the fiduciary model of professionalism.[6] According to this model, the relationship between the professional and the client is not guided by the traditional social contract understanding, but it establishes a standard of trust as in a fiduciary relationship. In this model, even though the client retains autonomy and authority in the process of decision making, as the professional has an advantageous position, special obligations towards the client will be imposed on such a professional.[7]

[3] Melford, R.J. "If We Don't License Ourselves, Who Will? proceeding of the PASE conference Ross, W.D. *The Right and the Good* (Oxford: Clarendon, 1930).
[4] Gotterbarn, D. (1992), "Software Engineering Ethics," *Encyclopedia of Software Engineering*, Ed. John J. Marciniak, New York :John Wiley & Sons, Inc.
[5] *Id*.
[6] *Id*.
[7] Bayles, M. (1981), *Professional Ethics*, Belmont,CA.: Wadsworth

3. Floridi's informational ethics (FIE)

According to the informational ethics theory of Floridi, the purview of computer ethics should be broadened to include much more than simply human beings, their actions, intentions and characters.[8] FIE is conceptualised as macroethics which is intended to be applicable to all situations. The name "information ethics" is appropriate to Floridi's theory, because it treats everything that exists as "informational" objects or processes.[9]

As FIE considers everything as informational, it allows one to go beyond the human centric approach of looking at computer ethics and includes every existing entity – humans, other animals, plants, organizations, even non-living artifacts, electronic objects in cyberspace, pieces of intellectual property. Everything can be considered as an entity which has the potential of affecting other entities and be affected by other such entities. Floridi envisages an "infosphere" which is the totality of all that exists. Any damage or destruction to the information is termed as "entropy", which is parallel to an "impoverishment of the infosphere".[10] Floridi offers four "fundamental principles" :

1. entropy ought not to be caused in the infosphere;
2. entropy ought to be prevented in the infosphere;
3. entropy ought to be removed from the infosphere; and
4. the flourishing of informational entities as well as the whole infosphere ought to be promoted by preserving, cultivating and enriching their properties.

FIE is based upon the idea that everything in the infosphere has at least a minimum worth that should be ethically respected, even if that worth can be overridden by

[8] (2004), "Computer Ethics," in L. Floridi (ed.), *The Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell, 65–75.
[9] Id.
[10] (2010b), "Philosophy in the Information Age," in P. Allo (ed.), *Luciano Floridi and the Philosophy of Information*, Cambridge, UK: Cambridge University Press, 420–442.

other considerations.[11] FIE suggests that there is something even more elemental than life, namely being – that is, the existence and flourishing of all entities and their global environment – and something more fundamental than suffering, namely entropy.[12]

4. Górniak's hypothesis of universality

According to the 1995 ETHICOMP paper by Górniak, the system of computer ethics will evolve into global ethics and will no longer be considered as a separate branch of ethics. Górniak predicted that this will be applicable to every culture on earth.[13]

Górniak argues that as the political philosophy and ethical theories of Bentham and Kant were developed in response to the printing press revolution, similarly, the ethical theory in relation to the computer ethics will evolve as a response to the computer revolution. She believes that such a development of ethics will be non-spatial in nature due to the very nature of computers. As computer networks are not restrained within a border, therefore, the development of computer ethics cannot be restricted inside a particular border. Due to this, the development of computer ethics will encompass the entire globe.

As the computer ethics will fail to be functional unless it is recognised by the majority of the world's population, Górniak argues that the new principles of computer ethics will be respected globally and will be universal in nature. Here, she is presenting an argument which rethinks the foundation of human ethics and the nature of human life itself.[14]

---

[11] (2006a), "Evaluating the Social and Cultural Implications of the Internet," *Computers and Society*, 36(3): 41–44.
[12] (2006a), "Evaluating the Social and Cultural Implications of the Internet," *Computers and Society*, 36(3): 41–44.
[13] Krystyna Górniak-Kocikowska (1996), "The Computer Revolution and the Problem of Global Ethics" in Terrell Ward Bynum and Simon Rogerson, eds., *Global Information Ethics*, Opragen Publications, 1996, pp. 177 – 190, (the April 1996 issue of Science and Engineering Ethics).
[14] *Id*.

According to the Górniak hypothesis, "local" ethical theories like Europe's Benthamite and Kantian systems and the ethical systems of other cultures in Asia, Africa, the Pacific Islands, etc., will eventually be superseded by global ethics evolving from today's computer ethics. "Computer" ethics, then, will become the "ordinary" ethics of the information age.[15]

5. The Johnson hypothesis

In the 1999,[16] Deborah Johnson presented a view on computer ethics which at first sight might seem similar to the Górniak hypothesis, however, it is the opposite of that view. Johnson paints a picture where computer ethics has become ordinary ethics and ordinary ethics has become computer ethics. This idea of Johnson is different from Górniak's hypothesis as it does not discard the presence of the philosophies of Bentham and Kant. She is arguing that these philosophies will continue to exist and will be the base on which the computer ethics are to be developed.

Therefore, the idea that Johnson tries to put forward is that the distinction between computer ethics and ordinary ethics will disappear as the same old ethical principles will be adopted in the field of computer science. In such a situation, as the foundation of all the ethics is the same, the distinction between ordinary or computer ethics cannot be sustained.

[15] Terrel Ward Bynum, *A Very Short History of Computer Ethics*, American Philosophical Association's Newsletter on Philosophy and Computing. *Available at* http://www.cs.utexas.edu/~ear/cs349/Bynum_Short_History.html
[16] Deborah G. Johnson (1985), *Computer Ethics*, Prentice-Hall. (Second Edition 1994).

6. The Menlo Report

The Menlo Report sets up ethical principles and guidelines for computer and information security research. It is based on the principles set forth in the 1979 Belmont Report.[17] This Report adopts the three core principles from the Belmont Report: Respect for Persons, Beneficence and Justice. However, as Information and Communication Technology (ICT) raises new challenges due to the interaction between humans and communication technologies, it adds the fourth principle of Respect for Law and Public Interest. These principles need to be understood in the context of Information and Communication Technology Research (ICTR).

Under the Belmont Report, the principle of Respect for Persons means that individuals should be treated as autonomous agents who have the right to be protected. In this context, it only applies to research subjects with sufficient awareness, either by taking their consent or by obtaining informed consent from legal representatives. However, in the context of ICTR, it also includes computer systems and data that directly impact persons who are not research subjects themselves. According to the Belmont Report, the principle of Justice is applied through fairness in the selection of research subjects, and equitable distribution of the burdens and benefits of research according to individual need, effort, societal contribution, and merit. Fairness should guide the initial selection of the subjects, as well as the apportionment of burdens to those who will most likely benefit from the research. In the ICTR context, this principle implies that research should not arbitrarily target persons or groups based on attributes including (but not limited to): religion, political affiliation, sexual orientation, health, age, technical competency, national origin, race, or socioeconomic status. Neither should ICTR target specific populations for the sake of convenience or expediency.

[17] D. Dittrich and E. Kenneally, "The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research", Tech. rep., U.S. Department of Homeland Security, Aug 2012.

Under the principle of Beneficence, an appropriate balance between harm and likelihood of enhanced welfare from the research needs to be present. Broadening this principle under ICTR, it demands a framework for systematic identification of risks and benefits for a range of stakeholders, diligent analysis of how harms are minimized and benefits are maximized, pre-emptive planning to mitigate any realized harms, and implementing these evaluations into the research methodology. Finally, the Menlo Report seeks to introduce a new principle of Respect for Law and Public Interest. This principle tries to increase the compliance and transparency of methodologies and results and the accountability for actions. Therefore, Menlo Report takes the principles from the Belmont Report and broadens it in the context of ICTR.

7. Concluding remarks

It is evident that the ethical principles evolved in biomedical research have been most influential in discussion on ethics in other disciplines. In the context of ethics in computational sciences also, the focus has been on extending existing formulation of ethics to newer contexts. Our research identifies three key reasons for this discontinuity. First, the dynamic norms and research methods in big data render traditional systems of research ethics untenable.[18] For instance, the focus of big data methods is to collect as much data as possible, often without having clear expectations of the exact purpose of the data with the anticipation that it would be useful. This run counter to the ethical principles of informed consent of data subjects. Second, research ethics have evolved primarily in the domain of biomedical sciences, and the tensions that have existed between these set of ethics

[18] Metcalf, Jacob and Crawford, Kate, Where are Human Subjects in Big Data Research? The Emerging Ethics Divide (May 14, 2016). *Big Data and Society*, Spring 2016. Available at: https://ssrn.com/abstract=2779647

and social sciences are aggravated by big data research.[19] This is so because unlike medical research, newer methods in disciplines such as (big) data sciences do not go through a rigorous peer review examination and are often applied directly in the field. Finally, often laws exempt research using publicly available datasets, or anonymised datasets, as they expect little or no harm to data subjects from such research. However, this assumption is severely tested when data is used for secondary purposes, or when it is used in combination with other datasets.[20] We look at each of these three issues in some detail below.

The primary focus of ethical principles for research has been on the protection of human subjects. While big data research still involves the traditional idea of ethical principles drawing from the need to protect the human subject, the manner in which they manifest themselves may be very different. When data can be repurposed and connected with other datasets, it renders questionable traditional ethical frameworks which rely primarily on the idea of "research data being temporally and contextually constrained and restricted by technical infrastructures and financial cost." Research data is also no longer simply connected to the direct data subjects it relates to, but also to larger distributions of groups that the data subjects belong to, as the data could be used in ways that may impact larger groups. While this has to some extent always been the case, with the use of big data allows researchers to derive insights in ways that impact other members of the group that the data subject belongs to more directly.

Historically speaking, research ethics have evolved largely in the context of biomedical sciences, and have gradually been applied across other disciplines as well. However, this adoption has not been without its tensions. Tom Beauchamp, one of the authors of the Belmont Report (which is discussed above in detail) felt that the ethical principles in the biomedical sciences use the delineation between

[19] Zwitter, Andrej, Big Data Ethics (November 20, 2014). *Big Data & Society* 2014 1: DOI: 10.1177/2053951714559253; University of Groningen Faculty of Law Research Paper 2015/17. Available at: https://ssrn.com/abstract=2553758

[20] National Commission for the Protection of Human Subjects, of Biomedical and Behavioral Research and The National Commission for the Protection of Human Subjects (1979) *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Available at: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html

practice and research to determine application of the principles: "The general rule is that if there is any element of research in an activity, that activity should undergo review for the protection of human subjects."[21] While this approach was convenient in biomedical sciences as the unique fiduciary nature of physician-patient relationship lends itself well to this framework, the nature of practice in other domains, particularly an evolving domain such as data science, do not have clear demarcations between research and practice. For instance, when big data is used in practical applications such as Google Flu Trends, or to identify loan applicants by a fintech company, the demarcations between practice and research is not clear, leading to the whole exercise being seen as practices devoid from a need for internal review boards. Therefore, an ethical framework which was built around the idea of this distinction may not have clear triggers for application in the context of big data for development research.

The extension of the Belmont model has been heavily critiqued by researchers who view the indiscriminate application of these principles, particularly in the case of "fast developing technologies, it is often difficult to define the actual object of Internet research precisely."[22] Kate O'Riordan and Elizabeth Bassett argue that the 'Internet as a space' metaphor leads to incorrect classification the Internet as a whole, thus not respecting the heterogeneous nature of activities being conducted on it, not all of them of social nature. Therefore, while it is increasingly clear that computational sciences such as those involving big data need to evolve ethical frameworks to address and limit the direct and indirect impact they have on human subjects, there is also a need to critically examine the nature of such ethical frameworks and the suitability of their application for governing computational analysis and decision-making, especially when deployed to monitor, plan, and implement global sustainable development initiatives.

The final issue deals with the legitimacy of the use of publicly available materials by big data practitioners. So far, the use of data available in the public domain has often been considered as legitimate without questioning the way in which such data

[21] Id.
[22] Bassett EH and O'Riordan K (2002) Ethics of Internet research: Contesting the human subjects research model. *Ethics and Information Technology* 4(3).

may be used in research. In fact, given the rigid barriers to accessing big data of significant quality and quantity, publicly available data, often a result of publicly funded collection and research, has served as the great equalizer in the research community.[23] However, personal data is available in the public domain, or may be inferred through data processing practices by analysing disparate points of data. On their own, these disparate points of data may be perfectly innocuous (or not) but in combination with other data may reveal intimate and sensitive personal details about an individual, which may be used as broad parameters for decision-making not just with respect to the individual, but also to a large aggregated group the individual is seen as a part of. Anonymised data sets made available publicly have often been re-identified, and compromise the privacy of data subjects.[24]

Therefore, there is a clear need to articulate ethical guiding principles that must inform big data research and practice. The experience of ethical principles in the biomedical sciences and attempts to extend them to computational sciences would be extremely instructive for any process of this nature. We want to be especially cognizant of the tensions that have already existed between biomedical ethics and social sciences ethics. The research-practice dichotomy has been central to how we understand research ethics, and needs to be revisited when we look at evolving technologies such as big data and artificial intelligence. Additionally, the new tensions that the big data paradigm introduces renders some conventional formulations of ethical principles hard to implement. In the next document of this guidelines series, we will review ongoing initiatives to articulate ethical principles for big data and related disciplines, including emerging literature on Fairness, Accountability, and Transparency in Machine Learning (FAT ML) and articulate our approach to evolving principles.

[23] Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A. & Christakis, N. (2008) 'Tastes, ties, and time: A new social network dataset using Facebook.com', *Social Networks*, vol. 30, no. 4.
[24] Zimmer, M. (2008) 'More on the "Anonymity" of the Facebook dataset – it's Harvard College', MichaelZimmer.org Blog, [Online] Available at: http://www.michaelzimmer.org/2008/01/03/more-on-the-anonymity-of-the-facebook-dataset-its-harvard-college/