Ethics and Human Rights Guidelines for *Big Data for Development* Research #1

Review of Principles of Ethics in Biomedical Science

March 2020

Author: **Manjri Singh 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 first 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

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.

In this document, first of a multi-part guideline document for ethics and human rights in big data for development research, we trace the history of ethical principles in biomedical research. While embarking on an exercise of evolving ethical guiding principles in big data for development and their application, we feel it important to begin with a review of the body of literature of ethical principles in other domains.

Despite the mass of regulatory codes released by multiple organizations laying down ethics for human subject research, very few are globally accepted or referred to as general guidelines. We divide this document into three sections. The first section looks at the history of regulatory guidelines as an institutional response to much publicized discoveries of grossly unethical studies; the second section views the

[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).

development as a response, by different interested parties, to the Nuremberg Code; and the final section attempts a brief critique and comparison of selected codes.

2. Evolution of ethics as a response to highly publicised unethical experiments

This section looks at the Nuremberg Code,[3] the Declaration of Helsinki,[4] the Belmont Report,[5] and the International Conference on Harmonization on Good Clinical Practice[6] in terms of the historical context in which they came about. The first formulation of a code came at the end of World War II, as a result of the Doctor's Trial, for the prosecution of German physicians involved in medical experiments of Jewish (and other) prisoners.[7] The Code countered many of the arguments that had been placed as a justification by the defendants. The most important of these arguments were first, the assertion that there was no universal standard of research ethics, second, that the state determined the necessity of experimentation; third, that people doomed to die, as a consequence of being a public health problem, had been used (implying that the Hippocratic Oath had not been betrayed); and finally, that consent was tacit in the absence of documents to the contrary.[8] At the end of trial, fifteen German physicians were convicted and the Code had been developed to prevent further human subject abuse.

The Nuremberg Code however was ignored by many researchers.[9] In 1956, the Willowbrook Study was initiated by deliberately infecting children, from an

[3] Available at: https://history.nih.gov/research/downloads/nuremberg.pdf
[4] Available at: http://www.who.int/bulletin/archives/79%284%29373.pdf
[5] Available at: https://videocast.nih.gov/pdf/ohrp_appendix_belmont_report_vol_2.pdf
[6] Available at: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf
[7] Ralph Slovenko, The Evolution of Standards for Experimental Treatment or Research, 33 J. Psychiatry & L. 129 (2005).
[8] *Id*.
[9] Das, N. K., & Sil, A. (2017). Evolution of Ethics in Clinical Research and Ethics Committee. *Indian Journal of Dermatology*, 62(4), 373–379. http://doi.org/10.4103/ijd.IJD_271_17

institution for the mentally retarded, with mild hepatitis.[10] Further, in 1963, live cancer cells were injected into senile patients incapable of giving consent (Jewish Chronic Study).[11] Both these studies led to extensive public debate and in light of the emerging situation, the General Assembly of the World Medical Association promulgated the Declaration of Helsinki that has since, been revised and clarified. The declaration binds all physicians with a single principle highlighting the utmost supremacy of the health of the patient over any other value.[12]

The infamous Tuskegee Study surfaced in 1972 in the United States, and highlighted the callous medical rationing (or discrimination) based on race, class and sex. The U.S. Government embarked on a systematic formal regulation and in 1974, the National Research Act was passed that established the National Commission. The Belmont Report was the work of the Commission and established the three basic unifying principles: autonomy, beneficence, and justice that are widely used.[13]

With the increasing multiplicity of regulations and transnational studies, having a uniform code becomes an emerging need. This development led to the International Conference on Harmonisation in 1996 by the European Union, Japan and the United States (constituting the majority of the drug market) with Australia, Canada, the Nordic countries and the World Health Organization as observers. The goal was to accomplish centralized or uniform regulation in response to the consumer demands for improved access to drugs and greater efficiency of drug approvals.[14]

---

[10] Institute of Medicine (US) Committee on Clinical Research Involving Children; Field MJ, Behrman RE, editors. *Ethical Conduct of Clinical Research Involving Children*. Washington (DC): National Academies Press (US); 2004. 1, Introduction. Available from: https://www.ncbi.nlm.nih.gov/books/NBK25549/

[11] Goodwin, Michele and Whelan, Allison, Law, Bioethics, and Biotechnology (May 8, 2015). *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed., Vol. 13, James D. Wright, Editor, 2015, Forthcoming; UC Irvine School of Law Research Paper No. 2015-52. Available at SSRN: https://ssrn.com/abstract=2604187

[12] Available at: http://www.who.int/bulletin/archives/79%284%29373.pdf

[13] Goodwin, Michele and Whelan, Allison, Law, Bioethics, and Biotechnology (May 8, 2015). International Encyclopedia of the Social & Behavioral Sciences, 2nd ed., Vol. 13, James D. Wright, Editor, 2015, Forthcoming; UC Irvine School of Law Research Paper No. 2015-52. Available at SSRN: https://ssrn.com/abstract=2604187

[14] David V. Eakin, *The International Conference on Harmonization of Pharmaceutical Regulations: Progress or Stagnation*, 6 Tulsa J. Comp. & Int'l L. 221 (1999).

3. Evolution of ethics as a response to the Nuremberg Code

Another method of viewing the evolution in ethics is to look at the motivations that resulted in the promulgation of new regulations. Subsequent research guidelines have been shaped by the criticisms and to the convenience of groups such as drug companies, politicians, and researchers.[15] This section uses the Nuremberg Code as a baseline, and then traces the release of the Declaration of Helsinki, the International Ethical Guidelines for Health-related Research Involving Humans,[16] the Belmont Report and the International Conference on Harmonisation Good Clinical Practice regulations with reference to the responsible parties and their interests.

The Nuremberg Code came as a consequence of human experiments during World War II. At this time, human experimentation was viewed as a suspect activity and the judges in the Nuremberg trial sought to regulate it as much as possible. However, many medical institutions and organizations were not keen on accepting the principles in the Nuremberg Code because it laid down an absolute requirement for consent. The medical scientists and researchers found the code inapplicable to their activities and an alternative was desired. Therefore, the World Medical Association issued the Declaration of Helsinki that allowed for proxy consent, and experimentation on the young, mentally incapacitated and other individuals who lacked the ability to give legally valid consent. The Helsinki Declaration did not refer to the Nuremberg Code as it was deemed to not be pertinent to the subject matter of the Declaration. The inhumane Nazi experiments were condemned, but the Nuremberg Code seemed to attach itself with the idea that it was only suitable in extreme circumstances of war and brutality, and not relevant to civilized doctors

[15] George J. Annas, The Changing Landscape of Human Experimentation: Nuremberg, Helsinki, and Beyond, *2 Health Matrix* 119 (1992).
[16] Available at: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

during peacetime. Unsurprisingly, the Declaration of Helsinki was more popular, amongst the medical community, than the Nuremberg Code.

Scientific organizations were even more unreceptive to the Nuremberg Code than the medical community.[17] In 1978, CIOMS and the World Health Organization (WHO) issued new guidelines that focused on the ethical justification and scientific validity of experimentation on human beings.[18] Another set of guidelines (International Ethical Guidelines for Health-related Research Involving Humans) were issued in 1982 underscoring the need for independent institutional review.[19] This placed many of the ethical concerns in the hands of other scientists, and obliviated the need for a judgment on the touchstone of human rights. The rights of the subject were subordinated to the researcher's idea of welfare. It placed more importance on the rationale of the researcher.

Unlike the Declaration of Helsinki and International Ethical Guidelines for Health-related Research Involving Humans, which were issued by medical and scientific organizations for essentially private actors, the Belmont report arose from the need to regulate publicly funded, or government department led research. The Nuremberg Code, as discussed above, was not very influential across the globe. As a consequence, multiple federal regulations followed, the emphasis of which remained on Public Health Service funded research and governmental research, resulted in a uniform policy for approvals of grants and created a 'Common Rule'.[20] The various regulatory measures are very indicative of the uniqueness of the American regulatory framework in light of its needs. For example, Institutional Review Boards in the US, must have a minimum of five members of diverse backgrounds to properly safeguard interests of the subjects by taking into account cultural factors, and must contain a non-member when reviewing a proposal.[21] That

[17] George J. Annas, The Changing Landscape of Human Experimentation: Nuremberg, Helsinki, and Beyond, *2 Health Matrix* 119 (1992).
[18] Alfano, S. L. (2013). Conducting Research with Human Subjects in International Settings: Ethical Considerations. *The Yale Journal of Biology and Medicine*, 86(3), 315–321.
[19] George J. Annas, The Changing Landscape of Human Experimentation: Nuremberg, Helsinki, and Beyond, *2 Health Matrix* 119 (1992).
[20] Federal Policy for the Protection of Human Subjects ('Common Rule'). Available at https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html.
[21] *Id*.

the International Conference on Harmonisation issued Good Clinical Practice (ICH - GCP) regulations emerged from the need to connect the markets, expedite drug approval, and lower costs has been detailed sufficiently in the previous section. The ICH - GCP catered to the demands of the drug consumer and supplier market.[22]

4. Critique and comparison of selected codes

This final section is an attempt to contrast the significant codes and regulations in order to bring out an evolution of values as they developed in the field of experimental research. The features and nature of the codes alo to afford an understanding of the progression of ethical guidelines. The documents selected for this section are the Nuremberg Code, the Declaration of Helsinki and the Belmont Report.

The judges at the Nuremberg Trial sought to elaborate on the already existing Hippocratic ethics by articulating ten principles that focused on the research subject rather than the physician through the grant of individual rights. The principles deviated from the Hippocratic Oath by bringing the provision of informed consent and right to withdraw from participation of an experiment.[23] Along with consent, the Code's other provisions required welfare of the subjects to be considered. These provisions cannot be waived off by the subject. The Nuremberg Code also allows for the subject to actively claim protections as well.

The Nuremberg Code is based on a natural law. The Code combined the Hippocratic ethics with human rights. As the authors of the Code had a Hippocratic view on medical research, they failed to envision the entirety of risks faced by subjects. This is because Hippocratic Ethics, merge the research subject's autonomy with the

[22] Available at https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf.
[23] Evelyne Shuster, Ph.D., Fifty Years Later: The Significance of the Nuremberg Code. *N Engl J Med* 1997; 337:1436-1440, DOI: 10.1056/NEJM199711133372006. Available from: https://www.nejm.org/doi/full/10.1056/NEJM199711133372006

physician's judgment of what is best for the subject. The Nuremberg Code, however, is still an authoritative legal document in the field of ethics.[24]

The Declaration of Helsinki, on the other hand, privileged the facilitation of advancement of science rather than the emphasis on protection of the subject to the extent that the first official version of the Declaration released in 1964, did not have a requirement for consent in the Basic Principles laid. This allowed for an expansion of potential subjects as now minors or mentally incapacitated individuals could also be involved in clinical trials.[25] It was therefore increasingly accepted across the globe though it lacked legal character. The Declaration replaced the human rights based principles of the Nuremberg Code with a more lenient paternalistic model of guidelines. The Declaration distinguishes between therapeutic and scientific research and follows the principles of risk/benefit and informed consent. The standards vary for each case. The Helsinki Declaration has also been revised a number of times. In 1975, the requirement of informed consent was introduced through a revision. The same amendment also provided for formal peer review that diluted the informed consent requirement. In this aspect, and many others, the Nuremberg Code and the Declaration of Helsinki provide conflicting directions and are dissimilar. While the Nuremberg code sets out the rights on the subjects, the Helsinki Declaration focuses on the obligations of the researcher instead. The Code is a legalistic document insisting on consent whereas the Declaration is more of a non-legal ethical document that scraps the requirement of consent in certain cases. Despite its lack of legal character, the Declaration has been increasingly preferred over the Code.

The Belmont Report incorporated many of the elements of the Nuremberg Code and the Helsinki Declaration. It requires consent (though providing for legal representatives to give consent on behalf of the subject) and places immense importance on peer review of research protocol and design through Institutional Review Boards. There are numerous other regulations in place in the United States. However, the regulatory system also suffers from deficiencies such as being

[24] Id.
[25] Henning Rosenau, Legal Prerequisites for Clinical Trials under the Revised Declaration of Helsinki and the European Convention on Human Rights and Biomedicine, 7 *Eur. J. Health L.* 105 (2000).

expensive, slow and inadequately covers serious injuries sustainable through experimentation.[26] It does not sufficiently deal with threats to the subjects. While the regulatory system deficiencies is a shortcoming within the implementation phase of the Report, the substance of the Report itself has also been much criticized. The Report lays down three principles (Respect for Persons, Beneficence and Justice) to be respected in carrying out human subject research. The primary issue with the Belmont Report lies in the excessive vagueness of these principles and various interpretations possible of the given ethics.[27] Furthermore, the relation of the rules to each other is unclear. This becomes problematic in situations where the principles are in conflict with each other, as there is no hierarchy or harmonization of the principles. There is also not a single underlying rule to interpret the principles by. A final issue is that the nature and scope of the Report itself is subject to controversy. Two alternative approaches lie in the application of the principles, one that they serve as higher rules that must be considered when drafting ethics/guidelines/regulations/ codes or approving proposals and the second is to treat the principles themselves as substantial rules.

5. Concluding remarks

Ethical principles concerning research involving human subjects have evolved primarily in the domain of biomedical science, and the tensions that have existed between these set of ethics and social sciences are aggravated by big data research.[28] 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

[26] Cohen, I. Glenn and Lynch, Holly Fernandez, Introduction to: 'Human Subjects Research Regulation: Perspectives on the Future' (July 28, 2014). *Human Subjects Research Regulation: Perspectives on the Future* (I. Glenn Cohen & Holly Fernandez Lynch eds., MIT Press 2014). Available at: https://ssrn.com/abstract=2473261

[27] Miller, Richard B. (2003) "How the Belmont Report Fails,"Essays in Philosophy: Vol. 4: Iss. 2, Article 6. Available from: https://commons.pacificu.edu/cgi/viewcontent.cgi?article=1089&context=eip

[28] 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

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.[29] 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 science, and have gradually been applied across other disciplines. 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 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."[30] While this approach was convenient in biomedical sciences as the unique fiduciary nature of physician-patient relationship

[29] 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
[30] Id.

lends itself well to this framework, the nature of practice in other domain, 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."[31] 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 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

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

community.[32] 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.[33]

Therefore, there is a clear need to articulate ethical guiding principles that must inform big data research and practice. 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.

[32] 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.
[33] 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/