

Is it time to rethink institutional review boards for the era of big data?

To the Editor — We read with interest the recent Editorial ‘Shared, but not up for grabs’¹. In particular, the editor’s reflections about consent and privacy in connection to the use of publicly shared images of humans in the computer science community led us to think about the role of an institutional review board (IRB): how would an IRB examine a waiver of consent in such situations and review whether research “will not adversely affect the rights and welfare of the subjects”²?

An IRB is the stronghold that curbs unethical research practices. Historically, it has been routine practice in biomedical, behavioural and social sciences, with few exceptions. Most research entities in the aforementioned fields nowadays enforce mandatory IRB training for personnel upon hire, even if not immediately working on human subjects’ research. Investigators often complain about the leviathan bureaucracy accompanying the process of protocol set-up and approval by an IRB³. On the other hand, many IRBs face overwhelming challenges in streamlining the review process while maintaining rigour and quality⁴. These challenges are even more tortuous in cases where research bridges different fields and makes use of big ‘organic’ data that have been acquired from various online sources, even if the official questions remain the same⁵: ‘Is it research? Does the research

involve human subjects?’ Of course, data science practices may not be framed as research or datasets may not be considered human subjects, managing to entirely bypass ethics regulations⁵.

As the editor pointed out, even if data have a Creative Commons licence, this is not a sufficient guarantee that an IRB is not needed. Besides scraping online photos, this issue is also relevant for social media data mining. A systematic review of Twitter use for health research found that only 32% of the articles mentioned ethical approval and only 12% mentioned participant consent⁶.

The definition of minimal risk to subjects reads² “the probability and magnitude of harm or discomfort anticipated in the proposed research are not greater in and of themselves than those ordinarily encountered in daily lives of the general population or during the performance of routine physical or psychological examinations or tests.” Even if still subject of debate⁷, such a definition is workable in clinical settings, but hardly translatable to big data research. Would online photos or social media data scraping pose minimal risk to the subjects? Which research designs and objectives can be sought to minimize risks? We are long overdue for modernization of the procedural burden, evaluation and handling of new problematic grey areas in human

subject studies, and regulation of research integrity with big data^{8,9}.

We should honestly consider how much we can currently rely on IRB training for our big data research endeavours. *The Belmont Report*¹⁰, one of the basic reads for research ethics, is still inspiring us after 40 years. The report’s core philosophical principles of ‘respect for persons, beneficence and justice’ should guide us and make sure that IRBs will not be a merely irrelevant administrative exercise. □

Mattia Prosperì* and Jiang Bian

University of Florida, Gainesville, FL, USA.

*e-mail: m.prosperi@ufl.edu

Published online: 24 May 2019

<https://doi.org/10.1038/s42256-019-0059-7>

References

1. *Nat. Mach. Intell.* **1**, 163 (2019).
2. *Code of Federal Regulations* (US Department of Health and Human Services, 2009).
3. Silberman, G. & Kahn, K. L. *Milbank Q.* **89**, 599–627 (2011).
4. Shoenbill, K., Song, Y., Cobb, N. L., Drezner, M. K. & Mendonca, E. A. *J. Clin. Transl. Sci.* **1**, 176–183 (2017).
5. Metcalf, J. & Crawford, K. *Big Data Soc.* <https://doi.org/10.1177/2053951716650211> (2017).
6. Sinnenberg, L. et al. *Am. J. Public Health* <https://doi.org/10.2105/AJPH.2016.303512> (2017).
7. Resnik, D. B. *J. Med. Ethics* **31**, 35–38 (2005).
8. Ienca, M. et al. *PLOS ONE* **13**, e0204937 (2018).
9. Raymond, N. *Nature* **568**, 277 (2019).
10. *The Belmont Report* (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979).