

[Open Peer Review on Qeios](#)

# Big Data

Marcello Ienca, Agata Ferretti, Samia Hurst, Milo Puhan, Christian Lovis, Effy Vayena

## Source

Marcello Ienca, Agata Ferretti, Samia Hurst, Milo Puhan, Christian Lovis, Effy Vayena. (2018). *Considerations for ethics review of big data health research: A scoping review*. *PLoS ONE*, vol. 13 (10), e0204937.

In general terms, big data involves large sets of data with diverse levels of analysable structuration, coming from heterogeneous sources (online data, social media profiles, financial records, self-tracked parameters, etc.), produced with high frequency and which can be further processed and analysed using computational techniques. While the term big data has become nearly ubiquitous, there is controversy over what data volumes are sufficiently large to obtain the big data label. Dumbill, for example, suggested that data should be considered big when they cross the threshold of the conventional databases systems' capacity in processing information<sup>1</sup>.

<sup>1</sup> Dumbill E. Making sense of big data. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2013.