

Introduction



Cite this article: Floridi L, Taddeo M. 2016
What is data ethics? *Phil. Trans. R. Soc. A* **374**:
20160360.
<http://dx.doi.org/10.1098/rsta.2016.0360>

Accepted: 3 October 2016

One contribution of 15 to a theme issue
'The ethical impact of data science'.

Subject Areas:

e-science, human–computer interaction

Keywords:

data ethics, data science, ethics of data, ethics
of algorithms, ethics of practices, levels of
abstraction

Author for correspondence:

Luciano Floridi
e-mail: luciano.floridi@oii.ox.ac.uk

What is data ethics?

Luciano Floridi^{1,2} and Mariarosaria Taddeo^{1,2}

¹Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford
OX1 3JS, UK

²Alan Turing Institute, 96 Euston Road, London NW1 2DB, UK

 LF, 0000-0002-5444-2280; MT, 0000-0002-1181-649X

This theme issue has the founding ambition of landscaping *data ethics* as a new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values). Data ethics builds on the foundation provided by computer and information ethics but, at the same time, it refines the approach endorsed so far in this research field, by shifting the level of abstraction of ethical enquiries, from being information-centric to being data-centric. This shift brings into focus the different moral dimensions of all kinds of data, even data that never translate directly into information but can be used to support actions or generate behaviours, for example. It highlights the need for ethical analyses to concentrate on the content and nature of computational operations—the interactions among hardware, software and data—rather than on the variety of digital technologies that enable them. And it emphasizes the complexity of the ethical challenges posed by data science. Because of such complexity, data ethics should be developed from the start as a macroethics, that is, as an overall framework that avoids narrow, ad hoc approaches and addresses the ethical impact and implications of data science and its applications within a consistent, holistic and inclusive framework. Only as a macroethics will data ethics provide solutions that can maximize the value of data science for our societies, for all of us and for our environments.

This article is part of the themed issue 'The ethical impact of data science'.

1. Introduction

Data science provides huge opportunities to improve private and public life, as well as our environment (consider the development of smart cities or the problems caused by carbon emissions). Unfortunately, such opportunities are also coupled to significant ethical challenges. The extensive use of increasingly more data—often personal, if not sensitive (big data)—and the growing reliance on algorithms to analyse them in order to shape choices and to make decisions (including machine learning, artificial intelligence and robotics), as well as the gradual reduction of human involvement or even oversight over many automatic processes, pose pressing issues of fairness, responsibility and respect of human rights, among others.

These ethical challenges can be addressed successfully. Fostering the development and applications of data science while ensuring the respect of human rights and of the values shaping open, pluralistic and tolerant information societies is a great opportunity of which we can and must take advantage. Striking such a robust balance will not be an easy or simple task. But the alternative, failing to advance both the ethics and the science of data, would have regrettable consequences. On the one hand, overlooking ethical issues may prompt negative impact and social rejection, as was the case, for example, of the NHS care.data programme.¹ *Social acceptability* or, even better, *social preferability* must be the guiding principles for any data science project with even a remote impact on human life, to ensure that opportunities will not be missed. On the other hand, overemphasizing the protection of individual rights in the wrong contexts may lead to regulations that are too rigid, and this in turn can cripple the chances to harness the social value of data science. The LIBE amendments initially proposed to the European Data Protection Regulation offer a concrete example of the case in point.²

Navigating between the Scylla of social rejection and the Charybdis of legal prohibition in order to reach solutions that maximize the ethical value of data science to benefit our societies, all of us and our environments is the demanding task of *data ethics*. In achieving this task, data ethics can build on the foundation provided by computer and information ethics, which has focused for the past 30 years on the main challenges posed by digital technologies [1–3]. This rich legacy is most valuable. It also fruitfully grafts data ethics onto the great tradition of ethics more generally. At the same time, data ethics refines the approach endorsed so far in computer and information ethics, as it changes the level of abstraction (LoA) of ethical enquiries from an information-centric (LoA_I) to a data-centric one (LoA_D).³

Ethical analyses are developed at a variety of LoAs. The shift from LoA_I to LoA_D is the latest in a series of changes that have characterized the evolution of computer and information ethics. Research in this field first endorsed a human-centric LoA [6], which addressed the ethical problems posed by the dissemination of computers in terms of professional responsibilities of both their designers and users. The LoA then shifted to a computer-centric one (LoA_C) in the mid-1980s [7], and it changed again at the beginning of the second millennium to LoA_I [8].

These changes responded to rapid, widespread and profound technological transformations. And they had important conceptual implications. For example, LoA_C highlighted the nature of computers as universal and malleable tools. It made it easier to understand the impact that computers could have on shaping social dynamics as well as on the design of the environment surrounding us [7]. LoA_I then shifted the focus from the technological means to the content (information) that can be created, recorded, processed and shared through such means. In doing so, LoA_I emphasized the different moral dimensions of information—i.e. information as the

¹See <https://www.england.nhs.uk/ourwork/tsd/care-data/>.

²Amendments 27, 327, 328 and 334–337 proposed in Albrecht's Draft Report, http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/922/922387/922387en.pdf.

³The method of abstraction is a common methodology in computer science [4] and in philosophy and ethics of information [5]. It specifies the different LoAs at which a system can be analysed, by focusing on different aspects, called observables. The choice of the observables depends on the purpose of the analysis and determines the choice of LoA. Any given system can be analysed at different LoAs. For example, an engineer interested in maximizing the aerodynamics of a car may focus upon the shape of its parts, their weight and the materials. A customer interested in the aesthetics of the same car may focus on its colour and on the overall look and may disregard the shape, weights and material of the car's components.

source, the result or the target of moral actions—and led to the design of a macroethical approach able to address the whole cycle of information creation, sharing, storage, protection, usage and possible destruction [8].

Data science, as the latest phase of the information revolution, is now prompting a further change in the LoA at which our ethical analysis can be developed most fruitfully. In a few decades, we have come to understand that it is not a specific technology (computers, tablets, mobile phones, online platforms, cloud computing and so forth), but what any digital technology manipulates that represents the correct focus of our ethical strategies. The shift from *information* ethics to *data* ethics is probably more semantic than conceptual, but it does highlight the need to concentrate on what is being handled as the true invariant of our concerns. This is why labels such as ‘robo-ethics’ or ‘machine ethics’ miss the point, anachronistically stepping back to a time when ‘computer ethics’ seemed to provide the right perspective. It is not the hardware that causes ethical problems, it is what the hardware does with the software and the data that represents the source of our new difficulties. LoA_D brings into focus the different moral dimensions of data. In doing so, it highlights the fact that, before concerning information, ethical problems such as privacy, anonymity, transparency, trust and responsibility concern data collection, curation, analysis and use, and hence they are better understood at that level.

In the light of this change of LoA, data ethics can be defined as the branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values). This means that the ethical challenges posed by data science can be mapped within the conceptual space delineated by three axes of research: the ethics of data, the ethics of algorithms and the ethics of practices.

The ethics of data focuses on ethical problems posed by the collection and analysis of large datasets and on issues ranging from the use of big data in biomedical research and social sciences [9], to profiling, advertising [10] and data philanthropy [11,12] as well as open data [13]. In this context, key issues concern possible re-identification of individuals through data-mining, -linking, -merging and re-using of large datasets, as well as risks for so-called ‘group privacy’, when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination (e.g. ageism, ethnicism, sexism) to group-targeted forms of violence [14,15]. Trust [16,17] and transparency [18] are also crucial topics in the ethics of data, in connection with an acknowledged lack of public awareness of the benefits, opportunities, risks and challenges associated with data science [19]. For example, transparency is often advocated as one of the measures that may foster trust. However, it is unclear what information should be made transparent and to whom information should be disclosed.

The ethics of algorithms addresses issues posed by the increasing complexity and autonomy of algorithms broadly understood (e.g. including artificial intelligence and artificial agents such as Internet bots), especially in the case of machine learning applications. In this case, some crucial challenges include moral responsibility and accountability of both designers and data scientists with respect to unforeseen and undesired consequences as well as missed opportunities [20,21]. Unsurprisingly, the ethical design and auditing [22] of algorithms’ requirements and the assessment of potential, undesirable outcomes (e.g. discrimination or the promotion of antisocial content) is attracting increasing research.

Finally, the ethics of practices (including professional ethics and deontology) addresses the pressing questions concerning the responsibilities and liabilities of people and organizations in charge of data processes, strategies and policies, including data scientists, with the goal to define an ethical framework to shape professional codes about responsible innovation, development and usage, which may ensure ethical practices fostering both the progress of data science and the protection of the rights of individuals and groups [23]. Three issues are central in this line of analysis: consent, user privacy and secondary use.

While they are distinct lines of research, the ethics of data, algorithms and practices are obviously intertwined, and this is why it may be preferable to speak in terms of three axes defining a conceptual space within which ethical problems are like points identified by three values. Most of them do not lie on a single axis. For example, analyses focusing on data privacy will also address issues concerning consent and professional responsibilities. Likewise, ethical auditing of algorithms often implies analyses of the responsibilities of their designers, developers, users and adopters. Data ethics must address the whole conceptual space and hence all three axes of research together, even if with different priorities and focus. And for this reason, data ethics needs to be developed from the start as a macroethics, that is, as an overall ‘geometry’ of the ethical space that avoids narrow, ad hoc approaches but rather addresses the diverse set of ethical implications of data science within a consistent, holistic and inclusive framework.

This theme issue represents a significant step in such a constructive direction. It collects 14 other contributions, each analysing a specific topic belonging to one of the three axes of research outlined above, while considering its implications for the other two. The articles included in this issue were initially presented at a workshop on ‘The Ethics of Data Science, The Landscape for the Alan Turing Institute’ hosted at the University of Oxford in December 2015. The issue shares with the workshop the founding ambition of landscaping data ethics as a new area of ethical enquiries and identifying the most pressing problems to solve and the most relevant lines of research to develop.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

Acknowledgements. Before leaving the reader to the articles, we express our gratitude to the authors and the reviewers for their contributions, as well as to the Alan Turing Institute for funding the landscaping workshop as part of its research strategy. The workshop and this theme issue would not have been possible without the strong support and continuous encouragement of many colleagues, but in particular of Prof. Helen Margetts, Director of the Oxford Internet Institute, and of Prof. Andrew Blake, Director of the Alan Turing Institute. We are also grateful to Bailey Fallon, the journal’s Commissioning Editor, and to the editorial office of *Philosophical Transactions A* for their great help during the process leading to the publication of this theme issue.

References

1. Bynum T. 2015 Computer and information ethics. In *The Stanford encyclopedia of philosophy* (ed. EN Zalta), Winter 2015. See <http://plato.stanford.edu/archives/win2015/entries/ethics-computer/>.
2. Floridi L. 2013 *The ethics of information*. Oxford, UK: Oxford University Press.
3. Miller K, Taddeo M (eds) 2017 *The ethics of information technologies. Library of essays on the ethics of emerging technologies*. New York, NY: Routledge.
4. Floridi L. 2008 The method of levels of abstraction. *Minds Mach.* **18**, 303–329. (doi:10.1007/s11023-008-9113-7)
5. Hoare CAR. 1972 Notes on data structuring. In *Structured programming* (eds OJ Dahl, EW Dijkstra and CAR Hoare), pp. 83–174. London, UK: Academic Press.
6. Parker DB. 1968 Rules of ethics in information processing. *Commun. ACM* **11**, 198–201. (doi:10.1145/362929.362987)
7. Moor JH. 1985 What is computer ethics? *Metaphilosophy* **16**, 266–275. (doi:10.1111/j.1467-9973.1985.tb00173.x)
8. Floridi L. 2006 Information ethics, its nature and scope. *SIGCAS Comput. Soc.* **36**, 21–36. (doi:10.1145/1195716.1195719)
9. Mittelstadt BD, Floridi L. 2015 The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci. Eng. Ethics* **22**, 303–341. (doi:10.1007/s11948-015-9652-2)
10. Hildebrandt M. 2008 Profiling and the identity of the European citizen. In *Profiling the European citizen* (eds M Hildebrandt, S Gutwirth), pp. 303–343. Dordrecht, The Netherlands: Springer.
11. Kirkpatrick R. 2013 A new type of philanthropy: donating data. *Harvard Business Review*, March 21. See <https://hbr.org/2013/03/a-new-type-of-philanthropy-don>.

12. Taddeo M. 2016 Data philanthropy and the design of the infraethics for information societies. *Phil. Trans. R. Soc. A* **374**, 20160113. (doi:10.1098/rsta.2016.0113)
13. Kitchin R. 2014 *The data revolution: big data, open data, data infrastructures and their consequences*. Beverley Hills, CA: Sage.
14. Floridi L. 2014 Open data, data protection, and group privacy. *Philos. Technol.* **27**, 1–3. (doi:10.1007/s13347-014-0157-8)
15. Taylor L, Floridi L, van der Sloot B (eds). In press. *Group privacy: new challenges of data technologies*. Philosophical Studies, Book Series. Hildenberg, The Netherlands: Springer.
16. Taddeo M. 2010 Trust in technology: a distinctive and a problematic relation. *Knowl. Technol. Policy* **23**, 283–286. (doi:10.1007/s12130-010-9113-9)
17. Taddeo M, Floridi L. 2011 The case for e-trust. *Ethics Inf. Technol.* **13**, 1–3. (doi:10.1007/s10676-010-9263-1)
18. Turilli M, Floridi L. 2009 The ethics of information transparency. *Ethics Inf. Technol.* **11**, 105–112. (doi:10.1007/s10676-009-9187-9)
19. Drew C. 2016 Data science ethics in government. *Phil. Trans. R. Soc. A* **374**, 20160119. (doi:10.1098/rsta.2016.0119)
20. Floridi L. 2016 Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions. *Phil. Trans. R. Soc. A* **374**, 20160112. (doi:10.1098/rsta.2016.0112)
21. Floridi L. 2012 Distributed morality in an information society. *Sci. Eng. Ethics* **19**, 727–743. (doi:10.1007/s11948-012-9413-4)
22. Goodman B, Flaxman S. 2016 European Union regulations on algorithmic decision-making and a ‘right to explanation’. See <http://arxiv.org/abs/1606.08813>.
23. Leonelli S. 2016 Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems. *Phil. Trans. R. Soc. A* **374**, 20160122. (doi:10.1098/rsta.2016.0122)