



The Philosophy and Ethics of Information

Academic Year	2017-18, Hilary Term
Day and Time	Thursdays, Weeks 1-9, 15:00-17:00
Location	Meeting Room, Oxford Internet Institute, 41 St Giles, Oxford, OX1 3LW
Course Providers	Professor Luciano Floridi, Oxford Internet Institute, luciano.floridi@oii.ox.ac.uk Dr Jenny Krutzinna, jenny.krutzinna@oii.ox.ac.uk Dr Brent Mittelstadt, brent.mittelstadt@oii.ox.ac.uk Dr Mariarosaria Taddeo, Mariarosaria.taddeo@oii.ox.ac.uk Dr Sandra Wachter, sandra.wachter@oii.ox.ac.uk Mr James Williams, james.williams@balliol.ox.ac.uk
Prerequisites	None

Background

In this course, philosophy will be interpreted as the conceptual design of cogent and relevant answers to open questions of a foundational nature. In the case of the philosophy and ethics of information (PEI), the questions concern the conceptual nature and basic principles of information, including its dynamics, utilisation, and sciences, and the elaboration and application of an information-theoretical approach to classic and new philosophical problems. The course has three goals. It explains what PEI is, its problems, approaches, and methods. It introduces some key concepts and phenomena related to information. And it seeks to answer some crucial theoretical questions of great philosophical significance prompted by the development of the information society. No prior knowledge of philosophy, ethics or logic is presupposed, but some preparatory readings are recommended (see below).

Key Themes

- The Information Revolution.
- Open questions in the philosophy and ethics of information.
- The method of abstraction.
- The nature and logic of information.
- Digital ethics as a new form of environmentalism.
- The ethics of data
- The ethics of algorithms
- Cyberconflicts

Learning Objectives

At the end of this course students will

- have obtained a critical understanding of the basic problems, concepts and methodology of the philosophy and ethics of information;
- know how to analyse the conceptual nature of fundamental information-theoretic phenomena;
- be able to formulate research questions that are amenable to philosophical analysis and use relevant conceptual tools to design cogent answers to them;
- understand how new ethical challenges arise from the development of digital environments

and unprecedented forms of agency and what approaches may be fruitful in order to deal with them;

- be familiar with important work on the philosophy and ethics of information authored by distinguished researchers in the field.

Teaching Arrangements

There will be eight weekly two-hour classes. The classes will meet in weeks 1-4 and 6-9 of Hilary term.

Note

Students should note that over the course of the year, small changes may be made to the content, dates or teaching arrangements set out in this reading list, at the course provider's discretion. These changes will be communicated to students directly and will be noted on the internal course information website.

Summative Assessment

Assessment is by a single essay not to exceed 5000 words in length, details of which will be released to students at noon on Monday Week 1 Hilary Term (15 January) via the Assignment Submission Weblearn site. This essay must be submitted via WebLearn by 12 noon on Monday, Week 1 of Trinity Term (23 April).

Formative Assessment

Students will be required to write one short essay. The advised length is 2000-2500 words. This essay will provide a means for students to obtain feedback on the progress they have achieved. Additional details will be given the first week of class.

Submission of Assignments

The summative assignment for this course is due on Monday of Trinity Term Week 1 (23 April) by 12.00pm and should be submitted electronically via the [Assignment Submission WebLearn Site](#). The assignment should also be submitted electronically by 5:00 pm on the same day to teaching@oii.ox.ac.uk. If anything goes wrong with your submission, email teaching@oii.ox.ac.uk immediately. In cases where a technical fault that is later determined to be a fault of the Weblearn system (and not a fault of your computer) prevents your submitting the assessment on time, having a time stamped email message will help the Proctors determine if your assessment will be accepted. Please note that you should not wait until the last minute to submit materials since Weblearn can run slowly at peak submission times and this is not considered a technical fault.

Full instructions on using WebLearn for electronic submissions can be found on Plato under General Information. There is also an FAQ page on the Assignment Submission WebLearn Site.

Please note that work submitted after the deadline will be processed in the standard manner and, in addition, the late submission will be reported to the Proctors' Office. If a student is concerned that they will not meet the deadline they must contact their college office or examinations school for advice. For details on the regulations for late and non-submissions please refer to the Proctors website at <https://www.admin.ox.ac.uk/proctors/examinations/candidates/>.

Any student failing this assessment will need to follow the rules set out in the OII *Examining Conventions* regarding re-submitting failed work.

General Readings

Information and communication technologies (ICTs), their related sciences, the issues they generate, and the society they are shaping are all changing very quickly and deeply, so texts about such topics soon become outdated. The following reading list seeks to provide a balance between (a) recent texts that are up to date, (b) classic texts that have withstood the test of time, and (c) texts that are less likely to have been encountered in other courses for this degree. Texts in **bold** are set references and the only compulsory readings for this course, whereas (*) denotes a more advanced text.

Philosophy and Ethics of Information: the lectures will be based on the following four books, but only two, in bold, will serve as set references: **(Floridi 2010c)**, (Floridi 2011)*, (Floridi 2013a)*, **(Floridi 2014a)**. (Floridi 2003) is an introduction to the philosophy of information and computing that is now getting old. (Floridi forthcoming) may become available in time for the course. The Society for the Philosophy of Information publishes the following open access textbook: (Illari 2012).

Philosophy: (Russell 2001) is a classic, short introduction to some philosophical problems by one of the greatest philosophers of the twentieth century. It is useful in order to become acquainted both with some crucial issues in the field and with an abstract style of thinking. Of the many textbooks that introduce philosophical methods, (Baggini and Fosl 2010) provides an accessible and fairly comprehensive overview. Philosophical dictionaries and encyclopaedias can serve as useful references to check concepts, theories, authors, terminology etc. that appear throughout the course. Almost any will do, but (Flew and Priest 2002) is recommended. The Stanford Encyclopedia of Philosophy is excellent, online and free, (Craig 1998) is an exhaustive encyclopaedia; the version online is regularly updated.

Ethics: (Benn 1998) is a bit old but still a good introduction; (Deigh 2010) is more recent but does not cover applied ethics; on this a good reference is (Ryberg, Petersen, and Wolf 2007), see also (Driver 2007). (Floridi 2010a) is a recent handbook on information and computer ethics.

Classics: the following texts are must-read for anyone interested in the philosophy and ethics of information, (Turing 1936) and (Turing 1950), see (Turing 2004); (Shannon and Weaver 1949 rep. 1998); and (Wiener 1954), (Wiener 1961), (Wiener 1964), and (Simon 1996).

Artificial Intelligence: a standard reference book is (Russell and Norvig 2010), to be consulted. Two philosophical books about AI are: (Copeland 1993), a bit old now, but still very insightful, and (Walmsley 2012). The book by the father of ELIZA is also worth reading (Weizenbaum 1976).

Computation: this is another field with a huge number and variety of introductions. For this course, (St. Amant 2013) is a good choice. (Sipser 2012)* is a very authoritative text for anyone interested in a more mathematical approach to computation; (Boolos, Burgess, and Jeffrey 2007)* is the text for anyone interested in a more logico-philosophical approach. Both require knowledge of discrete mathematics and first order logic.

Logic: there is a huge number and variety of introductory textbooks about mathematical logic. The following is recommended for the course: (Nolt et al. 2012), an abridged version of (Nolt, Rohatyn, and Varzi 2011). For many years (Hodges 2001) was synonymous with logic for any student in Oxford taking a first course in the subject; (Jeffrey and Burgess 2006)* covers the same material but more technically and systematically; so does (Smullyan 2014), which is almost a classic. (Priest 2000) and (Papineau 2012)* introduce philosophical issues in logic; (Huth and Ryan 2004)* and (Loveland, Hodel, and Sterrett 2014)* cover logic in computer science.

Mathematics: there is no royal road to geometry, as Euclid once remarked, nor to mathematics in general, but the following texts can help to make a few steps in the right direction: (Haggarty 2001) and (Makinson 2008) provide a solid foundation useful for this course, (Devlin 1981) is shorter and still very good.

Information theory: few introductory texts to information theory do not require some expertise in mathematics, computation, or logic. (Pierce 1980) is old but still one of the most accessible books on the topic. (Floridi 2010c) is a much shorter and more recent alternative; see also (Bremer and Cohnitz 2004) Popular books on the nature of information and the history of information technologies include (Brown and Duguid 2002), (Baeyer 2003), and more recently (Gleick 2011). (Goldstine 1972) remains a classic with regard to the history of computers.

Week 1: The Fourth Revolution

Instructor: Prof. Luciano Floridi

ICTs are not just tools merely modifying how we deal with the world, like the wheel or the engine. They are above all formatting systems, which increasingly affect how we conceptualise reality, how we relate to it, how we see ourselves, and how we interact with each other. In this lecture, we shall discuss the information revolution as a fourth revolution (Turing's, after Copernicus', Darwin's and Freud's), and analyse some of its long-term implications.

Essential Reading: (Floridi 2014a), chapters 1-5.

Optional Reading: (Floridi 2010b) chapter 1 is an overview of the same topics. *The Onlife Manifesto*, in (Floridi 2014b), please note that this is also freely available online as open access; (Bolter 1984); (Bynum and Moor 1998), esp. chapter 1.

Week 2: Digital Ethics as E-nvironmental Ethics

Instructor: Prof. Luciano Floridi

Moral life is a highly information-intensive game, so any technology that radically modifies the "life of information" is bound to have profound moral implications for any moral player. ICTs, by radically transforming the context in which moral issues arise, not only add interesting new dimensions to old problems, but lead us to rethink, methodologically, the very grounds on which our ethical positions are based. This lecture introduces Digital Ethics as a general approach to moral issues by comparing it to environmental ethics.

Essential Reading: Floridi, Luciano and Taddeo, Mariarosaria, What is Data Ethics? *Phil. Trans. R. Soc. A*, Volume 374, Issue 2083, December 2016. Available at https://www.academia.edu/30234860/What_is_Data_Ethics; (Floridi 2010a), chapters 1-5; for a quick overview, covering the same topic, see (Floridi 2010c), chapter 8 and the Epilogue; (Floridi 2014a), chapter 10.

Optional Reading: (Castells 2000); (Floridi 2013a)*, chapters 1-2, 4-6.

Week 3: Agency and Power

Instructor: Prof. Luciano Floridi

The post-Westphalian Nation State developed by becoming more and more an Information Society. However, in so doing, it progressively made itself less and less the main information agent, because what made the Nation State possible and then predominant, as a historical driving force in human politics, namely ICTs, is also what is now making it less central, in the social, political and economic life of humanity across the world. ICTs fluidify the topology of politics. They do not merely enable but actually promote (through management and empowerment) the agile, temporary and timely aggregation, disaggregation and re-aggregation of distributed groups around shared interests across old, rigid boundaries represented by social classes, political parties, ethnicity, language barriers, physical barriers, and so forth. This is generating a new tension between the Nation State, still understood as a major organisational institution, yet no longer monolithic but increasingly morphing into a multiagent system itself, and a variety of equally powerful, indeed sometimes even more politically influential and powerful, non-Statal organisations. This lecture discusses such a tension and how an information society could be designed to take full advantage of the socio-political progress made so far, while being able to deal successfully with the new global challenges (from the environment to the financial markets) that are undermining the legacy of that very progress.

Essential Reading: (Floridi 2014a), (Floridi 2014f).

Optional Reading: (Steger 2013); (Webster 2014). The following classics are available in many editions, some of which are free online, in brackets are the editions that are recommended for this course because of their scholarly apparatus: Locke, *A Letter concerning Toleration* (Locke 1991); Mill,

On Liberty (Mill 1991); Kant, *Toward Perpetual Peace* (Kant 2006); (Keynes 1933) is freely available online on many websites; (Rawls 1985) is a short overview of his influential theory of justice.

Week 4: Conflicts in Information Societies, Normative Challenges

Instructor: Dr Mariarosaria Taddeo

A relation of mutual influence exists between the way conflicts are waged and the societies waging them. Much like other social activities, conflicts mirror the values of societies while relying on their technological and scientific developments. In turn, the principles endorsed to regulate conflicts play a crucial role in shaping societies. The same mutual relation exists between information societies and cyber conflicts, making the regulation of the latter a crucial aspect contributing to shape current and future information societies. This lecture will focus on the normative problems that cyber conflicts pose, analyse the limits of existing approaches to address these problems, and describe new normative frameworks to regulate cyber conflicts.

Essential Reading: **(Taddeo 2014), (Taddeo 2016), (Schmitt 2013), (Ryan 2017)**

Optional Reading: (Hoisington 2009), (Lucas 2012), (Dipert 2013) (Cornish 2016)

Cornish, Paul. 2016. "Deterrence as the Basis for Ethical Constraints on Conflict in the Cyber Domain." In *Ethics and Policies for Cyber Warfare*. Philosophical Studies. Berlin Heidelberg: Springer.

Dipert, Randall. 2013. "The Essential Features of an Ontology for Cyberwarfare." In *Conflict and Cooperation in Cyberspace*, edited by Panayotis Yannakogeorgos and Adam Lowther, 35–48. Taylor & Francis.

<http://www.crcnetbase.com/doi/abs/10.1201/b15253-7>.

Lucas, G. R. 2012. "Just War and Cyber Conflict 'Can There Be an "Ethical" Cyber War?'" presented at the Naval Academy Class 2014.

Hoisington, Matthew. 2009. "Cyberwarfare and the Use of Force Giving Rise to the Right of Self-Defense." SSRN Scholarly Paper ID 1542223. Rochester, NY: Social Science Research Network.

<http://papers.ssrn.com/abstract=1542223>.

Week 5: Break

Week 6: The Ethics of Biomedical Big Data

Instructor: Dr Jenny Krutzinna

Data is increasingly being collected in all areas of our lives. When it comes to biomedical data, this is especially problematic, because it concerns a particularly sensitive aspect: our health. This has wide-ranging implications for all of us: In no other context is the conflict between individual and public interests so stark, because on the one hand, there are important privacy rights that warrant protection, and on the other hand, there is a moral imperative to improve the health of all. In the Western world, demographic changes and the resulting aging population put a strain on public health systems, leading to political pressures to manage health care costs more efficiently. Developing strategies for more effective prevention, diagnosis and treatment of ill-health have become a priority for local governments, with much hope being placed in big data analytics to provide insight into disease determinants. In the developing world, biomedical big data is hoped to help in the prevention and management of infectious diseases, and in improving health outcomes for a wider population.

New technologies and data-driven services promise to deliver the pathway to more effective disease-management. The greatest expectations appear to be placed in the growing fields of epigenetics and epigenomics. The Chief Medical Officer recently called for the NHS to deliver "the genomic dream" in the treatment of cancer patients. At the same time, private corporations are marketing aggressively to individuals to have their genome sequenced and to track their biological functions and lifestyle. Calls are being made to widely share such data in order to advance medicine and improve public health, leading to increasing concerns about privacy, confidentiality and individual autonomy with regard to one's medical data. This lecture and seminar will explore the conflicting interests in the context of

health, and consider ethical approaches to balancing such interests in an increasingly interconnected and data-driven health system.

Essential Reading: **(Mittelstadt and Floridi 2016)**, **(Vayena et al. 2015)**, (Nuffield Council on Bioethics 2015), (**“Chief Medical Officer Annual Report 2016: Generation Genome - GOV.UK”** 2017) Chapter 1

Optional Reading: (Aicardi et al. 2016)

Emerging ethical issues regarding digital health data. On the World Medical Association Draft Declaration on Ethical Considerations Regarding Health Databases and Biobanks

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4856186/>

Week 7: Ethics, Law and AI – Opening the Black Box

Instructor: Dr Sandra Wachter

We are increasingly being evaluated and assessed by machines. This alone is not a problem. It could help to make fairer and more accurate decisions; it may prove to be the case, for example, that algorithms can make less biased decisions than humans if programmed well or fed neutral data. It only becomes a problem when we do not understand how the algorithm reached its conclusion and the actions that are based on these conclusions adversely affect us, such as being fired, denied parole, or declined insurance. We have a tendency to trust these technologies more than human assessment, simply because they consider lots of data points and we see them as ‘smart’ or ‘intelligent’, even though they are often opaque ‘black boxes’ that are difficult to scrutinise. Healthcare, job market, and the criminal justice sector are areas that affect us massively. When a decision is taken about us in these areas, it is natural to want to make sure the decision is fair—that we are not being discriminated against or unfairly being treated different than others. This lecture and seminar will explore the regulatory efforts to ensure fair, transparent, and accountable algorithms and shed light on the current legislative difficulties.

Essential Reading: On ethical challenges in the use of algorithms – **(Mittelstadt et al., 2016)**; **(Wachter et al., 2017)**; **Wachter et al., 2016)**; **Barocas and Selbst 2016)**.

Optional Reading: (Friedman and Nissenbaum 1996); (Burrell 2016).Burrell, Jenna. 2016. How the machine thinks: understanding opacity in machine learning systems. *Big Data and Society* Jan-June 2016: 1-12. <http://journals.sagepub.com/doi/pdf/10.1177/2053951715622512>; Friedman, Batya, and Helen Nissenbaum. 1996. Bias in Computer Systems. *ACM Transactions and Information Systems* Vol 14:3 July 1996 pp 330-346.

<https://pdfs.semanticscholar.org/2e85/ebac82315863cd1b0426de96b0d641839e1a.pdf>

Week 8: Ethics of Algorithms

Instructor: Dr Brent Mittelstadt

In information societies, operations, decisions and choices previously left to humans are increasingly delegated to algorithms, which may advise, if not decide, about how data should be interpreted and what actions should be taken as a result. More and more often, algorithms mediate social processes, business transactions, governmental decisions, and how we perceive, understand, and interact among ourselves and with the environment. But despite their importance and potential benefits, decision-making algorithms also raise many ethical concerns. Often, the decisions made by algorithms are based on evidence that is inconclusive, produced with inscrutable methods, and potentially misguided. Problems can arise from biases and flaws in both input data and decision-

making models. Even when a sound evidence base exists, algorithmic decision-making can produce unfair or discriminatory outcomes, and have transformative effects on how individuals and society. Defining and enforcing standards of algorithmic fairness are fundamentally normative tasks, requiring reconciliation between potentially incompatible notions of fairness, privacy, equality, and other core ethical concepts. Determine which individuals, institutions, and artificial agents should be held accountable for the effects of algorithmic decision-making presents an overarching challenge. This lecture will provide an overview of ethical issues arising from decision-making algorithms in modern information societies. Specific attention will be given to algorithmic accountability pursued via pre-deployment ethical certification mechanisms, and post-deployment ethical auditing. Potential gaps between current work on ethical certification and auditing of algorithmic systems, and existing work in philosophy, ethics, and social theory will be discussed.

Essential Reading: (Mittelstadt et al. 2016); (Sandvig et al. 2014); (Romei and Ruggieri 2014); (Friedler et al. 2016)

Optional Reading: (Mittelstadt 2016); (Adler et al. 2016); (Floridi 2017); (Mittelstadt 2017)
Mittelstadt, Brent. 2016. 'Auditing for Transparency in Content Personalization Systems'. *International Journal of Communication* 10 (0): 12.; Adler, Philip, Casey Falk, Sorelle A. Friedler, Gabriel Rybeck, Carlos Scheidegger, Brandon Smith, and Suresh Venkatasubramanian. 2016; 'Auditing Black-Box Models by Obscuring Features'. *ArXiv:1602.07043 [Cs, Stat]*, February. <http://arxiv.org/abs/1602.07043>.; Floridi, Luciano. 2017. 'Group Privacy: A Defence and an Interpretation'. In *Group Privacy*, 83–100. Springer;Mittelstadt, Brent. 2017. 'From Individual to Group Privacy in Big Data Analytics'. *Philosophy & Technology*, February. doi:10.1007/s13347-017-0253-7.

Week 9: The Ethics of Attention and Persuasion

Instructor: Mr James Williams

When information becomes abundant, attention becomes the scarce resource. Today, the capture and exploitation of scarce human attention is the dominant business model, and therefore the central design goal, of most of the digital products and services we use. In the so-called 'attention economy,' the most 'persuasive' designs are the ones that ultimately win—and, given the intense, global nature of the competition for our attention, they must increasingly exploit our psychological vulnerabilities to have any chance of succeeding. To date, ethical analysis of digital technology has largely focused on questions related to the management of our information; very little, by comparison, has addressed the challenges posed by this industrial-scale management of our attention. As a result, the harms of persuasive design have been erroneously minimized as minor 'distractions' or 'annoyances.' While in the short term these influences can indeed 'distract' us from doing what we want to do, in the longer term they can distract us from living the lives we want to live, or, even worse, undermine our capacities for reflection and self-regulation, making it harder, in the words of Harry Frankfurt, to 'want what we want to want.' As a result, the digital attention economy poses major ethical challenges for user freedom and self-determination. This lecture and seminar will provide an overview of the current state of research in this emerging area, as well as a discussion about potential 'infraethical' interventions that could serve to realign the attention economy with true human needs and interests.

Essential Reading: (Simon, H. 1971); (Broadbent, S., & Lobet-Maris, C. 2015); (Berdichevsky, D., & Neuenschwander, E. 1999); (Verbeek, P. P. 2009);(Williams, J. 2016).

Optional Readings: (Crawford M. 2015); (Wu, Tim 2016); (Alter, Adam 2017); Lanham, R. A. (2006); Schüll, N.D. (2012).Crawford M. (2015). *The World Beyond Your Head: How to Flourish in an Age of Distraction*. Penguin; Wu, Tim (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads* (Part V). Knopf; Alter, Adam (2017). *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked*; Lanham, R. A. (2006). *The economics of attention: Style and*

substance in the age of information. University of Chicago Press; Schüll, N.D. (2012). *Addiction by Design: Machine Gambling in Las Vegas*. Princeton University Press.

List of Readings

- Baeyer, Hans Christian von. 2003. *Information: the new language of science*. London: Weidenfeld & Nicolson.
- Baggini, Julian, and Peter S. Fosl. 2010. *The philosopher's toolkit: a compendium of philosophical concepts and methods*. 2nd ed. Oxford: Wiley-Blackwell.
- Barocas, Solon and Andrew D Selbst. 2016. Big Data's Disparate Impact. *California Law Review* Vol 104:671.
<https://pdfs.semanticscholar.org/1d17/4f0e3c391368d0f3384a144a6c7487f2a143.pdf>
- Benn, Piers. 1998. *Ethics*. London: UCL Press.
- Berdichevsky, D., & Neuenschwander, E. 1999. *Toward an ethics of persuasive technology*. *Communications of the ACM*, 42(5), 51-58.
- Bolter, J. David. 1984. *Turing's man: Western culture in the computer age*. London: Duckworth.
- Boolos, George, John P. Burgess, and Richard C. Jeffrey. 2007. *Computability and logic*. 5th ed. Cambridge: Cambridge University Press.
- Bremer, Manuel Eugen, and Daniel Cohnitz. 2004. *Information and information flow: an introduction*. Frankfurt, Lancaster: Ontos Verlag; Gazelle Books.
- Broadbent, S., & Lobet-Maris, C. 2015. Towards a Grey Ecology. In *The Onlife Manifesto* (pp. 111-124). Springer International Publishing.
- Brown, John Seely, and Paul Duguid. 2002. *The Social Life of Information*. Boston: Harvard Business School Press.
- Bynum, Terrell Ward, and James Moor. 1998. *The digital phoenix: how computers are changing philosophy*. Oxford: Blackwell.
- Chief Medical Officer annual report 2016: *Generation Genome*, Chapter 1 Chief Medical Officer's summary (pp.1-23)
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/624628/CMO_annual_report_generation_genome.pdf
- Copeland, B. Jack. 1993. *Artificial intelligence: a philosophical introduction*. Oxford: Blackwell.
- Craig, Edward. 1998. *Routledge encyclopedia of philosophy*. 10 vols. London: Routledge.
- Deigh, John. 2010. *An introduction to ethics*. Cambridge: Cambridge University Press.
- Driver, Julia. 2007. *Ethics: the fundamentals*. Oxford: Blackwell.
- Flew, Antony, and Stephen Priest. 2002. *A dictionary of philosophy*. New ed. London: Pan.
- Floridi, Luciano, ed. 2003. *The Blackwell Guide to the Philosophy of Computing and Information*. Oxford, New York: Blackwell.
- Floridi, Luciano, ed. 2010a. *The Cambridge Handbook of Information and Computer Ethics*. Cambridge: Cambridge University Press.
- Floridi, Luciano. 2010b. *Information : a very short introduction*. Oxford: Oxford University Press.
- Floridi, Luciano. 2011. *The Philosophy of Information*. Oxford: Oxford University Press.
- Floridi, Luciano. 2013. *The Ethics of Information*. Oxford: Oxford University Press.
- Floridi, Luciano. 2014a. *The Fourth Revolution - How the infosphere is reshaping human reality*. Oxford: Oxford University Press.
- Floridi, Luciano, ed. 2014b. *The Onlife Manifesto - Being Human in a Hyperconnected Era*. New York: Springer.
- Floridi, Luciano, ed. 2016. *The Routledge Handbook of Philosophy of Information*. London: Routledge.
- Friedler, Sorelle A., Carlos Scheidegger, and Suresh Venkatasubramanian. 2016. 'On the (Im) Possibility of Fairness'. *ArXiv Preprint ArXiv:1609.07236*. <https://arxiv.org/abs/1609.07236>.
- Gleick, James. 2011. *The information: a history, a theory, a flood*. London: Fourth Estate.
- Goldstine, Herman H. 1972. *The computer from Pascal to von Neumann*. Princeton, N.J.: Princeton University Press.
- Hodges, Wilfrid. 2001. *Logic*. 2nd ed. London: Penguin.
- Huth, Michael, and Mark Ryan. 2004. *Logic in computer science: modelling and reasoning about systems*. 2nd ed. Cambridge: Cambridge University Press.
- Jeffrey, Richard C., and John P. Burgess. 2006. *Formal logic: its scope and limits*. 4th ed. Indianapolis: Hackett Pub. Co.

- Loveland, Donald W., Richard E. Hodel, and Susan G. Sterrett. 2014. *Three views of logic : mathematics, philosophy, and computer science*. Princeton: Princeton University Press.
- Nolt, John, Dennis A. Rohatyn, and Achille C. Varzi. 2011. *Outline of Logic*. 2nd ed. New York: Schaum; London: McGraw-Hill.
- Nolt, John, Dennis A. Rohatyn, Achille C. Varzi, and Alex M. McAllister. 2012. *Easy Outline of Logic*. New York; London: McGraw-Hill.
- Nuffield Council on Bioethics, Guide to the Report: "The collection, linking and use of data in biomedical research and health care." <http://nuffieldbioethics.org/wp-content/uploads/Biodata-a-guide-to-the-report-PDF.pdf>
- Mittelstadt, B.D. and Floridi, L. 2016. "The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts.", *Science and engineering ethics*. 22 (2) 303-341.
- Mittelstadt, Brent, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. 'The Ethics of Algorithms: Mapping the Debate'. *Big Data & Society* 3 (2). doi:10.1177/2053951716679679.
- Papineau, David. 2012. *Philosophical devices: proofs, probabilities, possibilities, and sets*. Oxford: Oxford University Press.
- Pierce, John Robinson. 1980. *An introduction to information theory: symbols, signals & noise*. 2nd ed. New York: Dover Publications.
- Priest, Graham. 2000. *Logic: a very short introduction*. Oxford: Oxford University Press.
- Romei, Andrea, and Salvatore Ruggieri. 2014. 'A Multidisciplinary Survey on Discrimination Analysis'. *The Knowledge Engineering Review* 29 (05): 582–638. doi:10.1017/S0269888913000039.
- Ryan, N. J. 2017. "Five Kinds of Cyber Deterrence." *Philosophy & Technology*, January. doi:10.1007/s13347-016-0251-1.
- Russell, Bertrand. 2001. *The problems of philosophy*. 2nd ed. Oxford: Oxford University Press.
- Russell, Stuart J., and Peter Norvig. 2010. *Artificial intelligence: a modern approach*. 3rd ed. Boston, Mass.: Pearson.
- Ryberg, Jesper, Thomas S. Petersen, and Clark Wolf. 2007. *New waves in applied ethics*. Basingstoke: Palgrave Macmillan.
- Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms'. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*. <http://social.cs.uiuc.edu/papers/pdfs/ICA2014-Sandvig.pdf>.
- Schmitt, M. 2013. "Cyberspace and International Law: The Penumbra Mist of Uncertainty." *Harvard* 126 (176): 176–80.
- Shannon, Claude Elwood, and Warren Weaver. 1949 rep. 1998. *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Simon, H. (1971). Designing Organizations for an Information-Rich World. *Computers, Communications, and the Public Interest*, pp. 37-72
- Simon, Herbert A. 1996. *The sciences of the artificial*. 3rd ed. Cambridge, Mass.: MIT Press.
- Sipser, Michael. 2012. *Introduction to the theory of computation*. 3rd ed. Boston, MA: Cengage Learning.
- Smullyan, Raymond M. 2014. *A beginner's guide to mathematical logic*. Mineola: Dover Publications.
- St. Amant, Robert. 2013. *Computing for ordinary mortals*. New York, NY: Oxford University Press.
- Taddeo, Mariarosaria. 2014. "Just Information Warfare." *Topoi*, April, 1–12. doi:10.1007/s11245-014-9245-8.
- . 2016. "On the Risks of Relying on Analogies to Understand Cyber Conflicts." *Minds and Machines* 26 (4): 317–21. doi:10.1007/s11023-016-9408-z.
- Turing, Alan Mathison. 1936. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematics Society, 2nd series* 42: 230-65.
- Turing, Alan Mathison. 1950. "Computing machinery and intelligence." *Mind* 59 (236): 433-460.
- Turing, Alan Mathison. 2004. *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence and artificial life, plus the secrets of Enigma*. Oxford: Clarendon Press.
- Verbeek, P. P. (2009). Ambient intelligence and persuasive technology: the blurring boundaries between human and technology. *Nanoethics*, 3(3), 231-242.
- Vayena, Effy and Urs Gasser. Between Openness and Privacy in Genomics. *PLoS Med* 13(1): e1001937
<http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001937>

- Wachter, Mittelstadt, B. and Floridi, L. (2017) "[Transparent, explainable, and accountable AI for robotics](http://digitalethicslab.oii.ox.ac.uk/sandra-wachter/#tab-7456a56b194183e7f18)", *Science Robotics*. 2 (6).
<http://digitalethicslab.oii.ox.ac.uk/sandra-wachter/#tab-7456a56b194183e7f18>
- Wachter, S., Mittelstadt, B. and Floridi, L. (2016) Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation.
<https://academic.oup.com/idpl/article/doi/10.1093/idpl/ix005/3860948/Why-a-Right-to-Explanation-of-Automated-Decision>
- Walmsley, Joel. 2012. *Mind and machine*. Basingstoke: Palgrave Macmillan.
- Weizenbaum, Joseph. 1976. *Computer power and human reason: from judgment to calculation*. San Francisco: W.H. Freeman.
- Wiener, Norbert. 1954. *The human use of human beings: cybernetics and society*. 2d ed. Garden City, New York: Doubleday.
- Wiener, Norbert. 1961. *Cybernetics or Control and Communication in the Animal and the Machine*. 2nd ed. Cambridge, Mass.: MIT Press.
- Wiener, Norbert. 1964. *God and Golem, inc.; a comment on certain points where cybernetics impinges on religion*. Cambridge, Mass.: M.I.T. Press.
- Williams, J. (2016) Why It's OK to Block Ads. In *Philosophers Take on the World* (pp. 223-226). Oxford University Press.