[8] Bovens, M. 2010. Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. *West European Politics* 33(5): 946–67. doi.org/10.1080/01402382.2010.486119.

[9] Bovens, M.; Schillemans, T.; and Goodin, R. E. 2014. Public Accountability. *The Oxford Handbook of Public Accountability*, edited by M. Bovens; and R. E. Goodin 1(1): 1–22. doi: 10.1093/oxfordhb/9780199641253.001.0001.

[10] Brundage, M. 2016. Artificial intelligence and responsible innovation. *Fundamental Issues of Artificial Intelligence*, edited by V. E. Müller, 543-555. Switzerland: Springer International Publishing.

[11] Brundage, M.; Avin, S.; Wang, J.; Belfield, H.; Krueger, G.; Hadfield, G.; Khlaaf, H. et al. 2020. Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. arXiv preprint. ArXiv:2004.07213 [Cs.CY]. Available at: arxiv.org/abs/2004.07213.

[12] Bryson, J. 2018a. Patiency Is Not a Virtue: The Design of Intelligent Systems and Systems of Ethics. *Ethics and Information Technology* 20(1): 15–26. doi.org/10.1007/s10676-018-9448-6.

[13] Bryson, J. 2018b. AI & Global Governance: No One Should Trust AI. *United Nations University Centre for Policy Research* (blog). 2018. Available at: cpr.unu.edu/ai-global-governance-no-one-should-trust-ai.html.

[14] Bryson, J.; and Winfield, A. 2017. Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. *Computer* 50(5): 116–19. doi.org/10.1109/MC.2017.154.

[15] Buolamwini, J.; and Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In Proceedings of Machine Learning Research, edited by A. F. Sorelle; and C. Wilson, 81: 77–91. New York, NY, USA: PMLR. proceedings.mlr.press/v81/buolamwini18a.html.

[16] Burrell, J. 2016. How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* 3(1): 1-12. doi.org/10.1177/2053951715622512.

[17] Büthe, T.; and Mattli, W. 2011. *The New Global Rulers: The Privatization of Regulation in the World Economy*. Oxford; Princeton, N.J.: Princeton University Press.

[18] Cabitza, F.; Ciucci, D.; and Rasoini, R. 2019. A Giant with Feet of Clay: On the Validity of the Data That Feed Machine Learning in Medicine. *Organizing for the Digital World*, edited by F. Cabitza; C. Batini; and M. Magni, 121–36. Lecture Notes in Information Systems and Organisation. Cham: Springer International Publishing. doi.org/10.1007/978-3-319-90503-7_10.

[19] Cabitza, F.; Rasoini, R.; and Gensini, G. F. 2017. Unintended Consequences of Machine Learning in Medicine. *JAMA* 318(6): 517–18. doi.org/10.1001/jama.2017.7797.

[20] Canca, C. 2020. Operationalizing AI Ethics Principles. *Communications of the ACM* 63(12): 18–21. doi.org/10.1145/3430368.

[21] Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; and Elhadad, N. 2015. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '15: 1721–30. Sydney, NSW, Australia: Association for Computing Machinery. doi.org/10.1145/2783258.2788613.

[22] CEN-CENELEC. 2020. CEN-CENELEC Focus Group Report: Road Map on Artificial Intelligence (AI). CEN-CENELEC Road Map Report on AI, version 2020-09. Available at: ftp.cencenelec.eu/EN/EuropeanStandardization/Sectors/AI/CEN-CLC_FGR_RoadMapAI.pdf.

[23] Chatila, R.; and Havens, J. C. 2019. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Robotics and Well-Being*, edited by M. I. A. Ferreira; J. S. Sequeira; G. S. Virk; M. O. Tokhi; and E. E. Kadar, 95: 11–16. Cham: Springer International Publishing. doi.org/10.1007/978-3-030-12524-0_2.

[24] Cihon, P.; Kleinaltenkamp, M. J.; Schuett, J.; and Baum, S. D. In review. AI Certification: Advancing Practice by Reducing Information Asymmetries. IEEE Transactions on Technology and Society.

[25] Davies, C. A. 2008. Reflexive Ethnography – A Guide to Researching Selves and Others. London & New York: Routledge.

[26] Delimatsis, P. 2015. *The Law, Economics and Politics of International Standardisation*. Cambridge International Trade and Economic Law. Cambridge, United Kingdom: Cambridge University Press.

[27] Denzin, N. K. 2013. Writing and/as Analysis or Performing the World. *The SAGE Handbook of Qualitative Data Analysis*, edited by U. Flick, 569-584. London, UK: Sage Publications.

[28] Deo, R. C. 2015. Machine Learning in Medicine. *Circulation* 132(20): 1920–30. doi.org/10.1161/CIRCULATIONAHA.115.001593.

[29] Diakopoulos, N. 2015. Algorithmic Accountability. *Digital Journalism* 3(3): 398–415. doi.org/10.1080/21670811.2014.976411.

[30] Diakopoulos, N.; Friedler, S.; Arenas, M.; Barocas, S.; Hay, M.; Howe, B.; Jagadish, H. V. et al. 2017. Principles for Accountable Algorithms and a Social Impact Statement for Algorithms. *FAT/ML*. Available at: fatml.org/resources/principles-for-accountable-algorithms.

[31] Dignum, V. 2019. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Cham, Switzerland: Springer. doi.org/10.1007/978-3-030-30371-6.

[32] Doshi-Velez, F., Kortz, M.; Budish, R.; Bavitz, C.; Gershman, S.; O'Brien, D.; Scott, K. et al. 2019. Accountability of AI Under the Law: The Role of Explanation. arXiv preprint. ArXiv:1711.01134 [Cs.AI]. Available at: arxiv.org/abs/1711.01134.

[33] Ensign, D.; Friedler, S. A.; Neville, S.; Scheidegger, C.; and Venkatasubramanian, S. 2018. Runaway Feedback Loops in Predictive Policing. In Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81: 160–71. proceedings.mlr.press/v81/ensign18a.html.

[34] Esteva, A.; Kuprel, B.; Novoa, R. A.; Ko, J.; Swetter, S. M.; Blau, H. M.; and Thrun, S. 2017. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature* 542(7639): 115–18. doi.org/10.1038/nature21056.

[35] Eubanks, V. 2018. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York, NY, USA: St. Martin's Press.

[36] Floridi, L. 2017. Infraethics–on the Conditions of Possibility of Morality. *Philosophy & Technology* 30: 391–94. doi.org/10.1007/s13347-017-0291-1.

[37] Floridi, L. 2018. Soft Ethics, the Governance of the Digital and the General Data Protection Regulation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376(2133). doi.org/10.1098/rsta.2018.0081.

[38] Floridi, L. 2019. Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology* 32(2): 185–93. doi.org/10.1007/s13347-019-00354-x.

[39] Floridi, L.; Cowls, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C. et al. 2018. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines* 28(4): 689–707. doi.org/10.1007/s11023-018-9482-5.

[40] Forsythe, D.; and Hess, D. J. 2001. *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence*. Writing Science. Stanford, CA: Stanford University Press.

[41] Geertz, C. 1973. *The Interpretation of Cultures*. New York, NY, USA: Basic Books.

[42] Ghorbani, A: Abid, A.; and Zou, J. 2019. Interpretation of Neural Networks Is Fragile. In Proceedings of the AAAI Conference on Artificial Intelligence 33(01): 3681–88. doi.org/10.1609/aaai.v33i01.33013681

[43] Goodman, B.; and Flaxman, S. 2017. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". *AI Magazine* 38(3): 50–57. doi.org/10.1609/aimag.v38i3.2741.

[44] Guidotti, R.; Monreale, A.; Ruggieri, S.; Turini, F.; Giannotti, F.; and Pedreschi, D. 2018. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys* 51(5) Article 93: 1-42. doi.org/10.1145/3236009.

[45] Gulshan, V.; Peng, P.; Coram, M.; Stumpe, M. C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S. et al. 2016. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 316(22): 2402–10. doi.org/10.1001/jama.2016.17216.

[46] DARPA. 2017. Broad Agency Announcement: Explainable Artificial Intelligence (XAI), DARPA-BAA-16-53. Available at: darpa.mil/attachments/DARPA-BAA-16-53.pdf.

[47] Hagendorff, T. 2020. The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines* 30(1): 99–120. doi.org/10.1007/s11023-020-09517-8.

[48] Halkier, B. 2011. Methodological Practicalities in Analytical Generalization. *Qualitative Inquiry* 17(9): 787–97. doi.org/10.1177/1077800411423194.

[49] ISO. 2013. ISO/IEC 27001:2013. International Organization for Standardization. Available at: iso.org/isoiec-27001-information-security.html.

[50] ISO. 2016. ISO 13485:2016 Medical Devices — Quality Management Systems — Requirements for Regulatory Purposes. International Organization for Standardization. Available at: iso.org/standard/59752.html.

[51] ISO. 2021a. Consumer Standards: Partnership for a Better World. International Organization for Standardization. Available at: iso.org/sites/ConsumersStandards/index.html#top

[52] ISO. 2021b. Popular Standards. International Organization for Standardization. Available at: iso.org/popular-standards.html.

[53] Jaton, F. 2017. We get the algorithms of our ground truths: Designing referential databases in digital image processing. *Social Studies of Science* 47(6): 811-840. doi.org/10.1177/0306312717730428.

[54] Jobin, A; Ienca, M.; and Vayena, E. 2019. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence* 1(9): 389–99. doi.org/10.1038/s42256-019-0088-2.

[55] Kannel, W. B.; Doyle, J. T.; McNamara, P. M.; Quickenton, P.; and Gordon, T. 1975. Precursors of Sudden Coronary Death. Factors Related to the Incidence of Sudden Death. *Circulation* 51 (4): 606–13. doi.org/10.1161/01.CIR.51.4.606.

[56] Kemper, J.; and Kolkman, D. 2019. Transparent to Whom? No Algorithmic Accountability without a Critical Audience. *Information, Communication & Society* 22 (14): 2081–96. doi.org/10.1080/1369118X.2018.1477967.

[57] Kooi, T.; Litjens, G.; van Ginneken, B.; Gubern-Mérida, A.; Sánchez, C. I.; Mann, R.; den Heeten, A.; and Karssemeijer, N. 2017. Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions. *Medical Image Analysis* 35 (January): 303–12. doi.org/10.1016/j.media.2016.07.007.

[58] Kroll, J. A. 2018. The Fallacy of Inscrutability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2133): 20180084. doi.org/10.1098/rsta.2018.0084.

[59] Kroll, J. A. 2020. Accountability in Computer Systems. *The Oxford Handbook of Ethics of AI*, 181.

[60] Kroll, J. A.; Huey, J.; Barocas, S.; Felten, E. W.; Reidenberg, J. R.; Robinson, D. G.; and Yu, H.. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165: 74.

[61] Kvale, S. (2008). *Doing interviews*. London, UK: Sage Publications.

[62] Lakkaraju, H.; and Bastani, O. 2020. 'How Do I Fool You?': Manipulating User Trust via Misleading Black Box Explanations. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 79–85. AIES '20. New York, NY, USA: Association for Computing Machinery. doi.org/10.1145/3375627.3375833.

[63] Larsson, S. 2020. On the Governance of Artificial Intelligence through Ethics Guidelines. *Asian Journal of Law and Society*, 1–15. doi.org/10.1017/als.2020.19.

[64] Latour, B. 1987. Science in action: How to follow scientists and engineers through society. Harvard University Press.

[65] Lepri, B.; Oliver, N.; Letouzé, E.; Pentland, A.; and Vinck, P. 2018. Fair, Transparent, and Accountable Algorithmic Decision-Making Processes. *Philosophy & Technology* 31(4): 611–27. doi.org/10.1007/s13347-017-0279-x

[66] Lipton, Z. C. 2018. The Mythos of Model Interpretability. *Queue* 16(3): 31-57. doi.org/10.1145/3236386.3241340

[67] Malterud, K. 2012. Systematic text condensation: a strategy for qualitative analysis. *Scandinavian Journal of Public Health* 40(8): 795–805. doi.org/10.1177/1403494812465030

[68] Martin, K. 2019. Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics* 160(4): 835–50. doi.org/10.1007/s10551-018-3921-3

[69] Miller, T. 2019. Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence* 267: 1–38. doi.org/10.1016/j.artint.2018.07.007.

[70] Mittelstadt, B. 2019. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence* 1 (11): 501–7. doi.org/10.1038/s42256-019-0114-4.

[71] Mittelstadt, B.; Russell, C.; and Wachter, S. 2019. Explaining Explanations in AI. In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*, 279–88. Atlanta, GA, USA: ACM Press. doi.org/10.1145/3287560.3287574.

[72] Montavon, G.; Lapuschkin, S.; Binder, A.; Samek, W.; and Müller, K.-R. 2017. Explaining Nonlinear Classification Decisions with Deep Taylor Decomposition. *Pattern Recognition* 65: 211–22. doi.org/10.1016/j.patcog.2016.11.008.

[73] Morley, J.; and Floridi, L. 2019. Enabling Digital Health Companionship Is Better than Empowerment. *The Lancet Digital Health* 1 (4): e155–56. doi.org/10.1016/S2589-7500(19)30079-2.

[74] Nissenbaum, H. 1994. Computing and Accountability. *Communications of the ACM* 37 (1): 72–81. doi.org/10.1145/175222.175228

[75] OECD. 2020. OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. Organisation for Economic Cooperation and Development.

[76] O'Neil, C. 2017. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. London: Penguin Books.

[77] Power, M. 1997. *The Audit Society: Rituals of Verification*. Oxford, [England]; New York: Oxford University Press.

[78] Raji, I. D.; Smart, A.; White, R. N.; Mitchell, M.; Gebru, T.; Hutchinson, B.; Smith-Loud, J.; Theron, D.; and Barnes, P.. 2020. Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. ArXiv:2001.00973 [Cs], January. Available at: arxiv.org/abs/2001.00973.

[79] Rességuier, A.; and Rodrigues, R. 2020. *AI ethics should not remain toothless!* A call to bring back the teeth of ethics. *Big Data & Society*, 1-5. doi.org/10.1177/2053951720942541

[80] Ribeiro, M. T.; Singh, S.; and Guestrin, C. 2016. 'Why Should I Trust You?': Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–44. KDD '16. New York, NY, USA: Association for Computing Machinery. doi.org/10.1145/2939672.2939778.

[81] Rudin, C. 2019. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence* 1 (5): 206–15. doi.org/10.1038/s42256-019-0048-x.

[82] Russell, S. J.; and Norvig, P. 2016. *Artificial Intelligence: A Modern Approach*, 3rd Edition. Pearson Education, 2009.

[83] Sculley, D.; Holt, G.; Golovin, D.; Davydov, E.; Phillips, T.; Ebner, D.; Chaudhary, V.; and Young, M. 2014. Machine Learning: The High Interest Credit Card of Technical Debt. In *SE4ML: Software Engineering for Machine Learning (NIPS 2014 Workshop)*.

[84] Shickel, B.; Tighe, P. J.; Bihorac, A.; and Rashidi, P. 2018. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics* 22 (5): 1589–1604. doi.org/10.1109/JBHI.2017.2767063.

[85] Shwartz-Ziv, R.; and Tishby, N. 2017. Opening the Black Box of Deep Neural Networks via Information. *ArXiv:1703.00810 [Cs]*. Available at: arxiv.org/abs/1703.00810.

[86] Sinclair, A. 1995. The Chameleon of Accountability: Forms and Discourses. *Accounting, Organizations and Society* 20 (2–3): 219–37. doi.org/10.1016/0361-3682(93)E0003-Y.

[87] Smith, H. 2020. Clinical AI: Opacity, Accountability, Responsibility and Liability. *AI & SOCIETY*, July. doi.org/10.1007/s00146-020-01019-6.

[88] Sossin, L.; and Smith, C. W. 2003. Hard Choices and Soft Law: Ethical Codes, Policy Guidelines and the Role of the Courts in Regulating Government. *Alberta Law Review*: 867–89. doi.org/10.29173/alr1344.

[89] Spradley, J. P. 1980. *Participant Observation*. New York: Holt, Rinehart and Winston.

[90] St. Pierre, E. A. 2011. Post qualitative research: The critique and the coming after. In *The SAGE Handbook of Qualitative Research*, edited by A. K. Denzin and Y. S. Lincoln, 4th Edition, 11-25. London, UK: Sage Publication.

[91] Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.

[92] Sweeney, L. 2013. Discrimination in Online Ad Delivery. *Queue* 11 (3): 10–29. doi.org/10.1145/2460276.2460278

[93] the European Commission. 2020. Commission Implementing Decision (EU) 2020/437. Available at: eur-lex.europa.eu/eli/dec_impl/2020/437/oj.

[94] Theodorou, A.; and Dignum, V. 2020. Towards Ethical and Socio-Legal Governance in AI. *Nature Machine Intelligence* 2(1): 10–12. doi.org/10.1038/s42256-019-0136-y.

[95] Tupasela, A.; Snell, K.; and Tarkkala, H. 2020. The Nordic Data Imaginary. *Big Data & Society* 7 (1): 2053951720907107. https://doi.org/10.1177/2053951720907107.

[96] Tutt, A. 2016. An FDA for Algorithms. *69 Admin. L. Rev. 83*. Rochester, NY: Social Science Research Network. Available at: papers.ssrn.com/abstract=2747994.

[97] Wachter, S.; Mittelstadt, B.; and Floridi, L. 2017. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law* (7): 76–99. Oxford University Press. doi.org/10.1609/aimag.v38i3.2741.

[98] Weller, A. 2019. Transparency: Motivations and Challenges. *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, edited by Samek, W.; Montavon, G.; Vedaldi, A.; Hansen, L.; Müller, K. R. 23–40. Lecture Notes in Computer Science. Cham: Springer International Publishing. doi.org/10.1007/978-3-030-28954-6_2.

[99] Whittaker, M.; Crawford, K.; Dobbe, R.; Fried, G.; Kaziunas, E.; Mathur, V.; West, S. M.; Richardson, R.; Schultz, J.; and Schwartz, O. 2018. *AI Now Report 2018*. AI Now Institute at New York University, New York.

[100] Whittlestone, J.; Nyrup, R.; Alexandrova, A.; and Cave, S. 2019. The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions, In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (*AIES '19*): 195–200. Association for Computing Machinery, New York, NY, USA. doi.org/10.1145/3306618.3314289

[101] Wieringa, M. 2020. What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 1–18. FAT* '20. Barcelona, Spain: Association for Computing Machinery. doi.org/10.1145/3351095.3372833.

[102] Winfield, A. F. T.; and Jirotka, M. 2018. Ethical Governance Is Essential to Building Trust in Robotics and Artificial Intelligence Systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2133): 20180085. doi.org/10.1098/rsta.2018.0085.

[103] Yin, R. K. 2009. *Case Study Research: Design and Methods*, 4th Edition. Los Angeles, LA, USA: Sage Publications.

[104] Zerilli, F. M. 2010. The Rule of Soft Law. *Focaal—Journal of Global and Historical Anthropology* (56): 3–18. doi.org/10.3167/fcl.2010.560101.

[105] Zhang, J. M.; Harman, M.; Ma, L.; and Liu, Y. 2019. Machine Learning Testing: Survey, Landscapes and Horizons. *IEEE Transactions on Software Engineering*. doi: 10.1109/TSE.2019.2962027.