Mois Navon
Adjunct Professor
Netanel Wiederblank
Adjunct Professor

Yeshiva University
**THE ZAHAVA AND MOSHAEL STRAUS CENTER
FOR TORAH AND WESTERN THOUGHT**

ETHICS IN
BIG DATA & ARTIFICIAL INTELLIGENCE
– FROM A JEWISH PERSPECTIVE

Ethics in Artificial Intel - JST 2665H 331 – HONORS
CRN: 85042
Class Schedule: Tues and Thurs 3:00 pm - 4:15 pm
In-person Class Location: Glueck Center G518
Semester and Year: Fall 2022

Professor: Mois Navon	Office: NA
Phone: +972 (50)-854-4612	Online Office Hours: by appointment
Email: mois.navon@divreinavon.com	Use this office hours appointment link
Professor: Netanel Wiederblank	Office: Glueck 508
Phone: 646-592-4463	Online Office Hours: by appointment
Email: wiederblank@hotmail.com	I can usually be found in the 2nd floor Glueck BM

COURSE DESCRIPTION

This course offers an introduction to the ethical issues inherent in the technologies spawned by big data and artificial intelligence. In particular it seeks to understand the Torah's perspective on these timely and controversial questions. As these fields enter every area of human existence, the ethics surrounding them is becoming ever critical. The course – seeking to touch upon all the prominent technologies driven by big data and artificial intelligence – will divided into five parts. The first part will focus on ethics, what is it and why it is important. This will include an overview of various approaches to ethics in general and the ethics of technology in specific. Next, we will dive into the big issues in big data. The third part will then focus on the ethical dilemmas involved in the autonomous systems driven by these technologies. Following this we will delve into the ethical issues of robotics and then conclude with a look into the future to discuss the ethical issues surrounding "the singularity."

COURSE MEETING SCHEDULE

Mode of Instruction: This course is an in-person class with Professor Navon participating remotely.

COURSE LEARNING OBJECTIVES

The goals of this course are to imbue the student with an appreciation for the value and necessity of ethics in general and as it applies to engineering in specific. The student will acquire the tools to engage in the ethical questions at home and in the workplace and to apply himself accordingly. Moreover, the course seeks to demonstrate how core Torah values address modern ethical questions.

PRE-REQUISITES

None.

COURSE REQUIREMENTS AND EXPECTATIONS:

Attendance: Regular and timely attendance is required. Material will be covered in class that is not covered in the readings.

ASSIGNMENTS

All students must make an oral presentation and submit a ten-page paper rigorously addressing an ethical question associated with AI based technology. When choosing a dilemma to address, it should have two reasonable sides and some level of tension; or result in different conclusions depending on moral approaches applied. You may suggest a resolution or simply present arguments in favor of multiple solutions. In either case, all arguments must be well supported. The paper should include a brief literature review and may include practical recommendations. It should incorporate both secular and Torah sources. You also must make an oral presentation to the class. This will be comprised of two parts: (1) A six-minute video presentation (it may be NO MORE than six minutes) and should focus on the key points of your findings, and (2) your ability to answer questions from your classmates and professors.

•   Oct 25th - Proposal. Submit a proposal for a topic. You may do this prior to this date insofar as we are going to try to avoid overlap with class and other presentations on a first come first served basis.

•   Nov 8th: Presentation of initial findings. You must schedule a meeting to discuss direction and initial findings with one of the instructors.

•   Dec 1st: You have an option of submitting a rough draft and getting feedback from one of the instructors in advance of their presentation (recommended but not mandatory).

•   December 20 and 22: Presentations.

•   Dec 28th: Paper due. Your paper should address issues raised by fellow students and faculty during the oral presentations.

The paper will be graded based on **six criteria**: (i) approach (creativity, etc.); (ii) informativeness; (iii) relevance and interest; (iv) cogency of argumentation; (v) quality of writing; and (vi) incorporation of critiques raised in classroom presentation.

GRADING COMPONENTS

Item	Percentage
Participation	5
Proposal	5
Initial findings	10
Oral presentation	30
Written paper	50

Mois Navon
Adjunct Professor
Netanel Wiederblank
Adjunct Professor

COURSE OUTLINE

#	Inst.	Date	Subject	Detail	Required Reading	Optional Reading
1	MN	Aug 25	Course Introduction	Overview, Motivations, Appetizers	(Soloveitchik 2012, pp. 6-11, 59-60)	
2	NW	Aug 30	Introduction to Ethics	Defining Ethics and the Need for an Ethical System	(Chazon Ish, Emuna U'Bitachon, Ch. 3); (R. Twersky 2003)	(R. Joseph Soloveitchik, The Halakhic Mind); (Kolbrener 1996)
3	NW	Sept 1	Introduction to Ethics	Sources of Morality	(R. Lichtenstein 2004, "Does Judaism Recognize an Ethic Independent of Halakhah?"); (Amital 2005, Ch. 2, "Natural Morality")	(R. Avraham Grodzinski *Torat Avraham* p. 128-132); (R. Yosef Yehuda Leib Bloch, Introduction to *Shiurei Da'at*); (Wiederblank 2014); (Plato, Euthyphro);
4	NW	Sept 6	Introduction to Ethics	Deontology, Utilitarianism, Virtues Ethics, and more	(Harris 2008)	On Canvas
5	MN	Sept 8	Introduction to Ethics	Difficulties with Moral Approaches (1) & (2)	(Nietzsche 1974, p181-2); (Berkovits 2004, ch 11); (Warburton 2019, sec. "existentialism", sec. "humanism"); (Blau 2002); (Plato 1997, Euthyphro); (Rachels and Rachels 2015a, 4.2)	(Rachels and Rachels 2015b); (Wurzburger 1994); (Descartes [1637] 2017); (Lamm 2007); (Shatz 2012); (Wurzburger 2008, 75-81); (Spero 2016); (Shatz, 1996)
6	NW	Sept 13	Ethics and Technology	Our Societal Responsibility	On Canvas	On Canvas
7	NW	Sept 15	Ethics and Technology	Our Relationship	(Soloveitchik 2012, pp. 6-11, 59-60)	On Canvas

Yeshiva University
THE ZAHAVA AND MOSHAEL STRAUS CENTER
FOR TORAH AND WESTERN THOUGHT

#	Inst.	Date	Subject	Detail	Required Reading	Optional Reading
				with Technology		
8	MN	Sept 20	Data Matters (1)	Bias In Models	(Hardt 2014);	(Crawford 2018); (Stewart 2020); (Solon Barocas, Moritz Hardt, and Arvind Narayanan 2019); (Johnson, 2020); (Hao 2019)
9	MN	Sept 22	Data Matters (2)	Fairness, Accountability, Transparency	(Corbett-Davies et al. 2016); (Lepri et al. 2017, p619-622)	(Lum and Isaac 2016); (Barocas and Selbst 2016); (Nesterak 2021); (Helbing 2017); (Xiang 2021)
10	MN	Oct 6-- Remote	Ethical Dilemmas in Big Data and AI	Surveillance, Privacy, and Public Welfare	(Pimple 2013: Ch. 4); (Bamberger and Mayse 2020, pp3-9)	(Roessler 2017); (Munch 2021); (ACM Staff 2021); (Kozlowska 2018); (Xu et al. 2016); (Pimple 2013: Ch. 3-6);
11	MN	Oct 20	Ethical Dilemmas in Big Data and AI	Manipulations, Nudges and Dark Patterns	(Susser, Roessler, and Nissenbaum 2019, p2-4); (Helbing 2017: see Big Nudging); (Hirsch 2010, Vay. 19:11);	(Noggle 2020); (Zimmerman 2016a); (Zimmerman 2016b)
12	MN	Oct 25- Proposal due	Ethical Dilemmas in Big Data and AI	GANs, Deepfakes and Fake News	(Rini 2020, p7-8 "Backstop Crises"); (Levinas 1997, p6-9); (Ha-Levi [1523] 1978, #74)	(Rini 2017); (Citron and Chesney 2019); (Gino 2017)
13	MN	Oct 27	Ethical Dilemmas in Big Data and AI	Virtual Reality (VR) - Real Implications	(Lanier 2017, ch.4); (Plato 1997, Republic, 514-	(Chalmers 2017); (Madary and Metzinger 2016); (Navon 2011)

#	Inst.	Date	Subject	Detail	Required Reading	Optional Reading
				for Unreal Activity	516e); (Nozick 1974, p1-2);	
14	MN	Nov 1	Ethical Dilemmas in Autonomous Systems	Autonomous Moral Agents (1)	(Muehlhauser and Helm 2012, ch. 4); (Moor 2011); (Purves, Jenkins, and Strawser 2015, ch. 3); (Penrose 1991, p423-4);	(Whitby 2011); (Moor 1995); (Spaulding 2020); (Wallach and Allen 2010, p.7); (Johnson and Miller 2008b)
15	MN	Nov 3	Ethical Dilemmas in Autonomous Systems	Autonomous Moral Agents (2)	(Sparrow 2021, sec. 3-6); (Lichtenstein 2002, p3-8); (Anderson and Anderson 2009);	(Dehghani et al. 2011); (Anderson 2011); (Brand 2020); (Beavers 2012); (Bringsjord et al. 2011); (Turkle 2011b); (Jiang et al. 2021); (Navon 2008)
16	MN	Nov 8- Present initial findings	Ethical Dilemmas in Autonomous Systems	Autonomous Vehicles	(Thomson 1985); (Thomson 2008)	(Lin 2016); (Navon 2022); (Harris 2008)
17	MN	Nov 10	Ethical Dilemmas in Autonomous Systems	Autonomous Weapon Systems (AWS)	(Arkin 2008a); (Sparrow 2016, 7-9, 14-15); (Asaro 2020, 227-230); (Shatz 2007, xiv-xvii)	(Sparrow 2007); (Purves, Jenkins, and Strawser 2015); (Müller 2016); (Tegmark 2018: Weapons); (Arkin 2010); (Arkin 2008b); (Leveringhaus 2016); (Orend 2006); (Broyde 2007); (Klapper, Ish-Shalom, and Broyde 2006);

#	Inst.	Date	Subject	Detail	Required Reading	Optional Reading
						(Bleich 2013, Ch. XI)
18	NW	Nov 15	Introduction to Roboethics	7. What is Human? Criteria	On Canvas	On Canvas
19	MN	Nov 17	Introduction to Roboethics (1)	Ontology, Consciousness and the Hard Problem	(Coeckelbergh 2013, 62-64)	(Nagel 1974); (Searle 1980); (Chalmers 1995);
20	MN	Nov 22	Introduction to Roboethics (2)	Machine Consciousness and Other Minds	(Penrose 1991, pp447-449); (Turing 1950); (Sebo 2018, p60-63)	(Reining 2020); (Hales 2014, ch 14); (Kastrup 2018); (Kastrup 2017); (Jiménez-Rodríguez 2017); (Warwick 2012); (Elamrani and Yampolskly 2019); (Torrance 2011); (Nagel 1974); (Bringsjord 2010)
21	MN	Nov 29	Ethical Dilemmas in Robotics	Mindless Robots	(Cappuccio, Peeters, and McDonald 2019, p25-28)	(Navon 2021); (Shmalo 2012); (Bryson 2010); (Heider and Simmel 1944); (Turkle 2011a)
22	MN	Dec 1—optional rough draft due	Ethical Dilemmas in Robotics	Consciousness Robots	(Petersen 2017, p3-11); (Maimonides 1168); (Turing 1950, p443); (Bleich 1998, p57-60)	(Signorelli 2018); (Grau 2011); (Gunkel 2018); (Idel 2019); (Scholem 1969); (Kogan 1989, p132-137); (Bryson 2010);

Mois Navon
Adjunct Professor
Netanel Wiederblank
Adjunct Professor

#	Inst.	Date	Subject	Detail	Required Reading	Optional Reading
						(Loike and Tendler 2003);
23	MN	Dec 6	Ethical Dilemmas in Robotics	Social Robots (1)	(Nyholm 2020, p109-115); (Aristotle 2004, Intro p.xxix-xxxii)	(Richardson 2015); (Mayor 2018); (LaGrandeur 2013, ch 1); (Nyholm 2020, ch. 5); (Sparrow and Sparrow 2006); (Turkle 2011a);
24	MN	Dec 8	Ethical Dilemmas in Robotics	Social Robots (2)	(Sparrow and Sparrow 2006, p153-156); (Berkovits 2002, pp.115-120); (Bertolini and Arian 2020, p44-49)	(Sparrow and Sparrow 2006); (Berkovits 2002, ch.4); (Darling 2017); (Turkle 2011a); (Soloveitchik 1989, ch. V); (Bertolini 2018); (Leibtag 2003, PART TWO); (Lichtenstein 2015)
25	NW	Dec 13	End of Days Ethics	Labor Issues	On Canvas	On Canvas
26	NW	Dec 15	End of Days Ethics	Eschatology	On Canvas	On Canvas
27		Dec 20	Reports 1-8			
28		Dec 22	Reports 8-15			

READING LIST

Bibliography

ACM, ed. 2021. "Differential Privacy: The Pursuit of Protections by Default." *Communications of the ACM* 64 (2): 36–43. https://doi.org/10.1145/3434228.

Anderson, Susan Leigh. 2011. "Philosophical Concerns with Machine Ethics." In *Machine Ethics*, edited

by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Anderson, Susan Leigh, and Michael Anderson. 2009. "How Machines Can Advance Ethics." *Philosophy Now*, 2009. https://philosophynow.org/issues/72/How_Machines_Can_Advance_Ethics.

Aristotle. 2004. *Nicomachean Ethics*. Translated by Roger Crisp. Cambridge: Cambridge University Press.

Arkin, Ronald C. 2008a. "Governing Lethal Behavior." *Proceedings of the 3rd International Conference on Human Robot Interaction - HRI '08*. https://doi.org/10.1145/1349822.1349839.

———. 2008b. "Governing Lethal Behavior - PART 2: Formalization for Ethical Control." In *Artificial General Intelligence 2008 Proceedings of the First AGI Conference*, edited by Pei Wang, Ben Goertzel, and Stan Franklin. Amsterdam: IOS Press.

———. 2010. "The Case for Ethical Autonomy in Unmanned Systems." *Journal of Military Ethics* 9 (4): 332–41. https://doi.org/10.1080/15027570.2010.536402.

Asaro, Peter. 2020. "Autonomous Weapons and the Ethics of Artificial Intelligence." In *Ethics of Artificial Intelligence*, edited by S. Matthew Liao, 212–36. Oxford: Oxford Univerity Press. https://doi.org/10.1093/oso/9780190905033.001.0001.

Bamberger, Kenneth A., and Ariel Mayse. 2020. "Privacy in Society: Jewish Law Insights for the Age of Big Data." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3731770.

Barocas, Solon, and Andrew D Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104 (3): 671. https://lawcat.berkeley.edu/record/1127463.

Beavers, Anthony F. 2012. "Moral Machines and the Threat of Ethical Nihilism." In *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by Patrick Lin, Keith Abney, and George Bekey. MIT Press.

Berkovits, Eliezer. 2002. *Essential Essays on Judaism*. Edited by David Hazony. Jerusalem: Shalem Press.

———. 2004. *God, Man and History*. Edited by David Hazony. Jeruselem: Shalem Press.

Bertolini, Andrea. 2018. "Human-Robot Interaction and Deception." *Osservatorio Del Diritto Civile E Commerciale, Rivista Semestrale* 2 (December): 645–59. https://doi.org/10.4478/91898.

Bertolini, Andrea, and Shabahang Arian. 2020. "3 Do Robots Care?" In *Aging between Participation and Simulation : Ethical Dimensions of Social Assistive Technologies*, edited by Joschka Haltaufderheide, Johanna Hovemann, and Jochen Vollmann, 35–52. Berlin: De Gruyter. https://doi.org/10.1515/9783110677485-003.

Blau, Yitzchak. 2002. "Ivan Karamazov Revisited: The Moral Argument for Religious Belief." *The Torah U-Madda Journal* 11: 50–60. http://www.jstor.org/stable/40914687.

Bleich, J. David. 1998. "Survey of Recent Halakhic Periodical Literature: Cloning: Homologous

Reproduction and Jewish Law." *Tradition* 32 (3). https://www.jstor.org/stable/23261122.

———. 2013. *Contemporary Halakhic Problems*. Vol. 3. NY: KTAV.

Brand, Lukas. 2020. "Why Machines That Talk Still Do Not Think, and Why They Might Nevertheless Be Able to Solve Moral Problems." In *Artificial Intelligence: Reflections in Philosophy, Theology, and the Social Sciences*, edited by BenediktPaul Göcke and Astrid Rosenthal-Von der Putten, 203–17. Boston: Brill.

Brey, Philip. 2010. "Philosophy of Technology after the Empirical Turn." *Techné: Research in Philosophy and Technology* 14 (1): 36–48. https://doi.org/10.5840/techne20101416.

Bringsjord, Selmer. 2010. "Meeting Floridi's Challenge to Artificial Intelligence from the Knowledge-Game Test for Self-Consciousness." *Metaphilosophy* 41 (3): 292–312. http://www.jstor.org/stable/24439827.

Bringsjord, Selmer, Joshua Taylor, Bram van Heuveln, Micah Clark, Ralph Wojtow, and Konstantine Arkoudas. 2011. "Piagetian Roboethics via Category Theory." In *Machine Ethics*, edited by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Broyde, Michael J. 2007. "Just Wars, Just Battles and Just Conduct in Jewish Law: Jewish Law Is Not a Suicide Pact!*." In *War and Peace in the Jewish Tradition*, edited by Lawrence H Schiffman and Joel B Wolowelsky. New York: Michael Scharf Publication Trust of the Yeshiva University Press.

Bryson, Joanna. 2010. "Robots Should Be Slaves." In *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues (Natural Language Processing; Vol. 8)*, edited by Y. Wilk, 63–74. John Benjamins Publishing Company. https://researchportal.bath.ac.uk/en/publications/robots-should-be-slaves.

Bunge, Mario. 1977. "Towards a Technoethics." *The Monist* 60 (1): 96–107. https://www.jstor.org/stable/27902461.

Cappuccio, Massimiliano L., Anco Peeters, and William McDonald. 2019. "Sympathy for Dolores: Moral Consideration for Robots Based on Virtue and Recognition." *Philosophy & Technology* 33 (1): 9–31. https://doi.org/10.1007/s13347-019-0341-y.

Chalmers, David. 1995. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (3): 200–219.

———. 2017. "The Virtual and the Real." *Disputatio* 9 (46): 309–52. https://doi.org/10.1515/disp-2017-0009.

Cholbi, Michael. 2017. "Suicide (Stanford Encyclopedia of Philosophy)." Stanford.edu. 2017. https://plato.stanford.edu/entries/suicide/.

Citron, Danielle, and Robert Chesney. 2019. "Deep Fakes: A Looming Challenge for Privacy,

Democracy, and National Security." *California Law Review* 107 (6): 1753.
https://scholarship.law.bu.edu/faculty_scholarship/640.

Coeckelbergh, Mark. 2013. "The Moral Standing of Machines: Towards a Relational and Non-Cartesian
Moral Hermeneutics." *Philosophy & Technology* 27 (1): 61–77. https://doi.org/10.1007/s13347-
013-0133-8.

Corbett-Davies, Sam, Emma Pierson, Avi Feller, and Sharad Goel. 2016. "A Computer Program Used for
Bail and Sentencing Decisions Was Labeled Biased against Blacks. It's Actually Not That Clear."
The Washington Post, October 17, 2016. https://www.washingtonpost.com/news/monkey-
cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/.

Crawford, Kate. 2018. "The Hidden Biases in Big Data." *Harvard Business Review*, January.
https://hbr.org/2013/04/the-hidden-biases-in-big-data.

Danaher, John. 2019. *Automation and Utopia : Human Flourishing in a World without Work.* Harvard
University Press.

Darling, Kate. 2017. "'Who's Johnny?' Anthropomorphic Framing in Human-Robot Interaction,
Integration, and Policy." In *Robot Ethics 2.0*, edited by Patrick Lin, Ryan Jenkins, and Keith
Abney, 173–88. NY: Oxford University Press.

Dehghani, Morteza, Ken Forbus, Emmett Tomai, and Matthew Klenk. 2011. "An Integrated Reasoning
Approach to Moral Decision Making." In *Machine Ethics*, edited by Michael Anderson and
Susan Leigh Anderson. NY: Cambridge University Press.

Descartes, Rene. (1637) 2017. *Discourse on the Method of Rightly Conducting One's Reason and Seeking
Truth in the Sciences*. Translated by Jonathan Bennett. *Annals of Neurosciences*. EARLY
MODERN PHILOSOPHY. https://doi.org/10.5214/ans.0972.7531.2009.160207.

Dietrich, Eric. 2011. "Homo Sapiens 2.0: Building the Better Robots of Our Nature." In *Machine Ethics*,
edited by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Elamrani, A., and R. Yampolskly. 2019. "Reviewing Tests for Machine Consciousness." *Journal of
Consciousness Studies* 26 (5-6). https://philpapers.org/rec/ELARTF.

Gaon, Saadia. (933AD) 1989. *The Book of Beliefs and Opinions*. Translated by Samuel Rosenblatt. New
Haven: Yale University Press.

Geraci, R. M. 2008. "Apocalyptic AI: Religion and the Promise of Artificial Intelligence." *Journal of the
American Academy of Religion* 76 (1): 138–66. https://doi.org/10.1093/jaarel/lfm101.

Gino, Francesca. 2017. "Dazzled by Creativity: Why We Excuse Dishonest Acts." Scientific American.
March 14, 2017. https://www.scientificamerican.com/article/dazzled-by-creativity-why-we-
excuse-dishonest-acts/.

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. 2018. "Viewpoint: When

Will AI Exceed Human Performance? Evidence from AI Experts." *Journal of Artificial Intelligence Research* 62 (July): 729–54. https://doi.org/10.1613/jair.1.11222.

Grau, Chistopher. 2011. "There Is No 'I' in 'Robot.'" In *Machine Ethics*, edited by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Gunkel, David J. 2018. *Robot Rights*. Cambridge, Mass: MIT Press.

Ha-Levi, Aaron. (1523) 1978. *Sefer HaHinnuch: The Book of [Mizvah] Education*. Translated by Charles Wengrov. New York: Feldheim.

Hales, Colin G. 2014. *The Revolutions of Scientific Structure*. Singapore World Scientific Pub. Co.

Hao, Karen. 2019. "This Is How AI Bias Really Happens—and Why It's so Hard to Fix." MIT Technology Review. February 4, 2019. https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.

Hardt, Moritz. 2014. "How Big Data Is Unfair." Medium. September 26, 2014. https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de.

Harris, Michael J. 2008. "Consequentialism, Deontologism, and the Case of Sheva Ben Bikhri." *The Torah U-Madda Journal* 15: 68–94. https://www.jstor.org/stable/40914728.

Heidegger, Martin. 2013. *The Question Concerning Technology and Other Essays*. New York: Harper.

Heider, Fritz, and Marianne Simmel. 1944. "An Experimental Study of Apparent Behavior." *The American Journal of Psychology* 57 (2): 243–59. https://doi.org/10.2307/1416950.

Helbing, Dirk. 2017. "Will Democracy Survive Big Data and Artificial Intelligence?" *Scientific American*, February 25, 2017. https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/.

Hirsch, Samson Raphael. 2010. *The Hirsch Chumash : The Five Books of the Torah*. Edited by Daniel Haberman and Gertrude Hirschler. Jerusalem: Feldheim.

Idel, Moshe. 2019. *Golem : Jewish Magical and Mystical Traditions on the Artificial Anthropoid*. Brooklyn, N.Y.: KTAV.

James, Aaron. 2020. "Planning for Mass Unemployment." In *Ethics of Artificial Intelligence*, edited by S. Matthew Liao, 183–211. Oxford: Oxford Univerity Press. https://doi.org/10.1093/oso/9780190905033.003.0007.

Jiang, Liwei, Jena D Hwang, Chandrasekhar Bhagavatula, Ronan Le Bras, Maxwell Forbes, Jon Borchardt, Jenny Liang, Oren Etzioni, Maarten Sap, and Yejin Choi. 2021. "Delphi: Towards Machine Ethics and Norms." *ArXiv* abs/2110.07574.

Jiménez-Rodríguez, Luis O. 2017. "Human Uniqueness or Anthropocentrism? Semantic, Anthropological and Theological Clarifications in Dialogue with Damasio's Neuroscience." In *Issues in Science*

ETHICS IN BD & AI - FALL 2022
Mois Navon
Adjunct Professor
Netanel Wiederblank
Adjunct Professor

Yeshiva University
THE ZAHAVA AND MOSHAEL STRAUS CENTER
FOR TORAH AND WESTERN THOUGHT

and Theology: Are We Special? Human Uniqueness in Science and Theology, edited by Michael Fuller, Dirk Evers, Anne Runehov, and Knut-Willy Sæther, 191–207. Springer International Publishing. https://doi.org/10.1007/978-3-319-62124-1₁4.

Johnson, Deborah G, and Keith Miller. 2008a. *Computer Ethics : Analyzing Information Technology*. Upper Saddle River, N.J.: Prentice Hall.

Johnson, Deborah G., and Keith W. Miller. 2008b. "Un-Making Artificial Moral Agents." *Ethics and Information Technology* 10 (2-3): 123–33. https://doi.org/10.1007/s10676-008-9174-6.

Johnson, Gabbrielle M. 2020. "Algorithmic Bias: On the Implicit Biases of Social Technology." *Synthese* 198 (June): 9941–61. https://doi.org/10.1007/s11229-020-02696-y.

Jonas, Hans. 2000. *The Imperative of Responsibility : In Search of an Ethics for the Technological Age*. Chicago (Ill.): The University Of Chicago Press.

Kastrup, Bernardo. 2017. "An Ontological Solution to the Mind-Body Problem." *Philosophies* 2 (4): 10. https://doi.org/10.3390/philosophies2020010.

———. 2018. "Sentient Robots, Conscious Spoons and Other Cheerful Follies." Scientific American. 2018. http://blogs.scientificamerican.com/observations/sentient-robots-conscious-spoons-and-other-cheerful-follies.

Keynes, John Maynard. (1930) 2009. "Economic Possibilities for Our Grandchildren." In *Essays in Persuasion*. New York: Classic House Books. http://www.econ.yale.edu/smith/econ116a/keynes1.pdf.

Klapper, Aryeh, Benjamin Ish-Shalom, and Michael J. Broyde. 2006. *CONVERSATION: Halakhah and Morality in Modern Warfare*. 6th ed. Vol. 1. Meorot. http://www.edah.org/backend/journalarticle/conversation%20-%20final.pdf.

Kogan, Brain S. 1989. "'What Can We Know and When Can We Know It'? Maimonides on the Active Intelligence and Human Cognition." In *Moses Maimonides and His Time*, edited by Eric L Ormsby, 121–37. Washington, D.C.: Catholic University of America Press.

Kozlowska, Iga. 2018. "Facebook and Data Privacy in the Age of Cambridge Analytica." The Henry M. Jackson School of International Studies. April 30, 2018. https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/.

Kroes, Peter, and Peter-Paul Verbeek, eds. 2014. *The Moral Status of Technical Artefacts*. Springer.

LaGrandeur, Kevin. 2013. *Androids and Intelligent Networks in Early Modern Literature and Culture Artificial Slaves*. NY: Routledge.

Lamm, Norman. 2007. "Amalek and the Seven Nations: A Case of Law vs. Morality." In *War and Peace in the Jewish Tradition*, edited by Lawrence H Schiffman and Joel B Wolowelsky. New York: Michael Scharf Publication Trust of the Yeshiva University Press.

Lanier, Jaron. 2017. *Dawn of the New Everything : Encounters with Reality and Virtual Reality*. New
York: Henry Holt And Company.

Leibtag, Menachem. 2003. "PARSHAT VA'YERA - the AKEYDA (and More)." Tanach.org. 2003.
https://tanach.org/breishit/vayera2.txt.

Lepri, Bruno, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck. 2017. "Fair,
Transparent, and Accountable Algorithmic Decision-Making Processes." *Philosophy &
Technology* 31 (4): 611–27. https://doi.org/10.1007/s13347-017-0279-x.

Leveringhaus, Alex. 2016. *Ethics and Autonomous Weapons*. London: Palgrave Macmillan.
https://doi.org/10.1057/978-1-137-52361-7.

Levinas, Emmanuel. 1997. *Difficult Freedom : Essays on Judaism*. Baltimore: Johns Hopkins University
Press.

Lichtenstein, Aharon. 2002. "The Human and Social Factor in Halakha." *Tradition* 36 (1).
https://traditiononline.org/the-human-and-social-factor-in-halakha/.

———. 2015. "On Raising Children." Torah Musings. May 22, 2015.
https://www.torahmusings.com/2015/05/on-raising-children/.

Lin, Patrick. 2016. "Why Ethics Matters for Autonomous Cars." In *Autonomous Driving*, edited by
Markus Maurer, Gerdes J Christian, Barbara Lenz, and Hermann Winner, 69–85. Springer Berlin
Heidelberg. https://doi.org/10.1007/978-3-662-48847-8$_4$.

Loike, John D., and Moshe D. Tendler. 2003. "Ma Adam Va-Teda-Ehu: Halakhic Criteria for Defining
Human Beings." *Tradition* 37 (2). https://traditiononline.org/ma-adam-va-teda-ehu-halakhic-
criteria-for-defining-human-beings/.

———. 2014. "Tampering with the Genetic Code of Life: Comparing Secular and Halakhic Ethical
Concerns." *Hakira* 18. https://hakirah.org/Vol18LoikeTendler.pdf.

Lum, Kristian, and William Isaac. 2016. "To Predict and Serve?" *Significance* 13 (5): 14–19.
https://doi.org/10.1111/j.1740-9713.2016.00960.x.

Madary, Michael, and Thomas K. Metzinger. 2016. "Recommendations for Good Scientific Practice and
the Consumers of VR-Technology." *Frontiers in Robotics and AI* 3 (February).
https://doi.org/10.3389/frobt.2016.00003.

Maimonides, Moses. 1168. "Introduction to Mishnah (EXCERPT)." Israel: Mois Navon.

———. 1956. *The Guide for the Perplexed*. Translated by Michael Friedlander. New York: Dover.

Mayor, Adrienne. 2018. *Gods and Robots : The Ancient Quest for Artificial Life*. Princeton Princeton
Univ Press.

Mill, John Stuart. 1863a. "Utilitarianism." *Https://Www.earlymoderntexts.com/*. EARLY MODERN
PHILOSOPHY. https://www.earlymoderntexts.com/assets/pdfs/mill1863.pdf.

———. 1863b. "Utilitarianism." *Https://Www.earlymoderntexts.com/*. EARLY MODERN

PHILOSOPHY. https://www.earlymoderntexts.com/assets/pdfs/mill1863.pdf.

Miller, Keith W. 2011. "Moral Responsibility for Computing Artifacts: 'the Rules.'" *IT Professional* 13

(3): 57–59. https://doi.org/10.1109/mitp.2011.46.

Moor, James H. 1995. "Is Ethics Computable?" *Metaphilosophy* 26 (1/2): 1–21.

https://www.jstor.org/stable/24439044.

———. 2011. "The Nature, Importance, and Difficulty of Machine Ethics." In *Machine Ethics*, edited by

Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Muehlhauser, Luke, and Louie Helm. 2012. "The Singularity and Machine Ethics." *The Frontiers

Collection*, 101–26. https://doi.org/10.1007/978-3-642-32560-1_6.

Müller, Vincent C. 2016. "Autonomous Killer Robots Are Probably Good News." In *Drones and

Responsibility*, edited by Ezio Di Nucci and Filippo Santonio de Si, 67–81. London: Ashgate.

https://doi.org/10.4324/9781315578187-4.

Munch, Lauritz Aastrup. 2021. "Privacy Rights and 'Naked' Statistical Evidence." *Philosophical Studies*

178 (11): 3777–95. https://doi.org/10.1007/s11098-021-01625-0.

Nagel, Thomas. 1974. "What Is It like to Be a Bat?" *The Philosophical Review* 83 (4): 435–50.

https://doi.org/10.2307/2183914.

Navon, Mois. 2007. "One Thousand Years of Shabbat." Jerusalem: Chidushei Torah Journal.

http://www.divreinavon.com/pdf/one_thousand_years_of_shabbat.pdf.

———. 2008. "Halacha, Ethics and Aesthetics." *The Everett Journal of Jewish Ethics*.

http://www.divreinavon.com/pdf/HalachaEthicsAesthetics.pdf.

———. 2011. *The Matrix – a Mystical Modern Midrash*. *Pop Media*. Divreinavon.

www.divreinavon.com.

———. 2021. "The Virtuous Servant Owner—a Paradigm Whose Time Has Come (Again)." *Frontiers in

Robotics and AI* 8 (September). https://doi.org/10.3389/frobt.2021.715849.

———. 2022. "The Trolley Problem Just Got Digital." *B.D.D*.

http://www.divreinavon.com/pdf/EthicalDilemmasinProgrammingAutonomousVehicles_MoisNa

von.pdf.

Nesterak, Evan. 2021. "A Conversation with Daniel Kahneman about 'Noise' - by Evan Nesterak -

Behavioral Scientist." Behavioral Scientist. May 24, 2021. https://behavioralscientist.org/a-

conversation-with-daniel-kahneman-about-noise/.

Nietzsche, Friedrich. 1974. *The Gay Science*. Translated by Walter Kaufmann. New York: Random.

Noggle, Robert. 2020. "The Ethics of Manipulation." Edited by Edward N. Zalta. Stanford Encyclopedia

of Philosophy. 2020. https://plato.stanford.edu/entries/ethics-manipulation/.

Nozick, Robert. 1974. "The Experience Machine (Excerpted)."

https://rintintin.colorado.edu/~vancecd/phil3160/Nozick1.pdf.

Nyholm, Sven. 2020. *Humans and Robots : Ethics, Agency, and Anthropomorphism*. NY: Rowman &
Littlefield Publishing Group.

Orend, Brian. 2006. *The Morality of War*. Peterborough, Ont.: Broadview Press.

Penrose, Roger. 1991. *The Emperor's New Mind : Concerning Computers, Minds and the Laws of
Physics*. NY: Penguin.

Petersen, Steve. 2017. "Is It Good for Them Too? Ethical Concern for the Sexbots." In *Robot Sex: Social
Implications and Ethical*, edited by John Danaher and Neil McArthur, 155–71. Cambridge, USA:
MIT Press.

Pimple, Kenneth, ed. 2013. *Ethics and Pervasive Augmented Reality: Some Challenges and Approaches*.
Emerging Pervasive Information and Communication Technologies (PICT). Springer.
https://doi.org/10.1007/978-94-007-6833-8_8.

Plato. 1997. *Plato : Complete Works*. Edited by John M Cooper. Indiana: Hackett.
https://www.holybooks.com/wp-content/uploads/plato-complete-works.pdf.

Purves, Duncan, Ryan Jenkins, and Bradley J. Strawser. 2015. "Autonomous Machines, Moral Judgment,
and Acting for the Right Reasons." *Ethical Theory and Moral Practice* 18 (4): 851–72.
https://doi.org/10.1007/s10677-015-9563-y.

Rachels, Stuart, and James Rachels. 2015a. *The Elements of Moral Philosophy*. New York: Mcgraw-Hill
Education.

———, eds. 2015b. *The Right Thing to Do : Basic Readings in Moral Philosophy*. New York: Mcgraw-
Hill Education.

Reining, Stefan. 2020. "Revisiting the Dancing-Qualia Argument for Computationalism." In *Artificial
Intelligence: Reflections in Philosophy, Theology, and the Social Sciences*, edited by
GöckeBenedikt Paul and Astrid Rosenthal-Von der Pütten. Brill.

Richardson, Kathleen. 2015. *An Anthropology of Robots and AI Annihilation Anxiety and Machines*. NY:
Routledge.

Rini, Regina. 2017. "Fake News and Partisan Epistemology." *Kennedy Institute of Ethics Journal* 27 (2):
43–64. https://doi.org/10.1353/ken.2017.0025.

———. 2020. "Deepfakes and the Epistemic Backstop." *Philosophers' Imprint* 20 (24).
https://philpapers.org/rec/RINDAT.

Roessler, Beate. 2017. "Privacy as a Human Right." *Proceedings of the Aristotelian Society* 117 (2): 187–
206. https://doi.org/10.1093/arisoc/aox008.

Sandler, Ronald L., ed. 2014. *Ethics and Emerging Technologies*. London: Palgrave Macmillan UK.

https://doi.org/10.1057/9781137349088.

Schneider, Susan. 2019. "Should You Add a Microchip to Your Brain?" *The New York Times*, June 10, 2019, sec. Opinion. https://www.nytimes.com/2019/06/10/opinion/future-artificial-intelligence-transhumanism.html.

Scholem, Gershom Gerhard. 1969. *On the Kabbalah and Its Symbolism*. New York: Schocken Books.

Searle, John R. 1980. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (03): 417–57. https://doi.org/10.1017/s0140525x00005756.

Sebo, Jeff. 2018. "The Moral Problem of Other Minds." *The Harvard Review of Philosophy* 25: 51–70. https://doi.org/10.5840/harvardreview20185913.

Shatz, David. 1996. "Beyond Obedience: Walter Wurzburger's Ethics of Responsibility." *Tradition* 30 (2): 74–95.

———. 2007. "Introduction." In *War and Peace in the Jewish Tradition*, edited by Lawrence H Schiffman and Joel B Wolowelsky. New York: Michael Scharf Publication Trust of the Yeshiva University Press.

———. 2012. *Ethical Theories in the Orthodox Movement*. *Oxford Handbooks Online*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199736065.013.0015.

Shmalo, Gamliel. 2012. "Orthodox Approaches to Biblical Slavery." *The Torah U-Madda Journal* 16. https://www.jstor.org/stable/23596054.

Signorelli, Camilo Miguel. 2018. "Can Computers Become Conscious and Overcome Humans?" *Frontiers in Robotics and AI* 5 (October). https://doi.org/10.3389/frobt.2018.00121.

Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2019. *Fairness and Machine Learning*. fairmlbook.org. https://fairmlbook.org/.

Soloveitchik, Joseph. 1989. *Reflections of the Rav : Lessons in Jewish Thought - Vol. 2*. Edited by Abraham R Besdin. Hoboken, NJ: KTAV.

———. (1965) 2012. *The Lonely Man of Faith*. Jerusalem: Maggid. https://traditiononline.org/wp-content/uploads/2019/09/LMOF.pdf.

Sotala, Kaj. 2012. "Advantages of Artificial Intelligences, Uploads, and Digital Minds." *International Journal of Machine Consciousness* 04 (01): 275–91. https://doi.org/10.1142/s1793843012400161.

Sparrow, Robert. 2007. "Killer Robots." *Journal of Applied Philosophy* 24 (1): 62–77. https://doi.org/10.1111/j.1468-5930.2007.00346.x.

———. 2016. "Robots and Respect: Assessing the Case against Autonomous Weapon Systems." *Ethics & International Affairs* 30 (1): 93–116. https://doi.org/10.1017/s0892679415000647.

———. 2021. "Why Machines Cannot Be Moral." *AI & Society* 36 (3): 685–93.

https://doi.org/10.1007/s00146-020-01132-6.

Sparrow, Robert, and Linda Sparrow. 2006. "In the Hands of Machines? The Future of Aged Care."
Minds and Machines 16 (2): 141–61. https://doi.org/10.1007/s11023-006-9030-6.

Spaulding, Norman W. 2020. "Is Human Judgment Necessary? Artificial Intelligence, Algorithmic
Governance, and the Law." In *The Oxford Handbook of Ethics of AI*, edited by Markus Dirk
Dubber, Frank Pasquale, and Sunit Das. New York: Oxford University Press.

Spero, Shubert. 2016. "The Good, the Right, and the Morality of Judaism." *The Torah U-Madda Journal*
17: 202–17. https://www.jstor.org/stable/26203067.

Stewart, Matthew. 2020. "Programming Fairness in Algorithms - towards Data Science." Medium.
Towards Data Science. July 2, 2020. https://towardsdatascience.com/programming-fairness-in-
algorithms-4943a13dd9f8.

Susser, Daniel, Beate Roessler, and Helen F. Nissenbaum. 2019. "Online Manipulation: Hidden
Influences in a Digital World." *4 Georgetown Law Technology Review 1*.
https://doi.org/10.2139/ssrn.3306006.

Tegmark, Max. 2018. *Life 3.0*. NY: Vintage.

Thomson, Judith Jarvis. 1985. "The Trolley Problem." *The Yale Law Journal* 94 (6): 1395–1415.
https://doi.org/10.2307/796133.

———. 2008. "Turning the Trolley." *Philosophy & Public Affairs* 36 (4): 359–74.
https://www.jstor.org/stable/40212830.

Tirosh-Samuelson, Hava. 2012. "Transhumanism as a Secularist Faith." *Zygon* 47 (4).
https://doi.org/https://doi.org/10.1111/j.1467-9744.2012.01288.x.

Torrance, Steve. 2011. "Machine Ethics and the Idea of a More-Than-Human Moral World." In *Machine
Ethics*, edited by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University
Press.

Turing, A. M. 1950. "Computing Machinery and Intelligence." *Mind* LIX (236): 433–60.
https://doi.org/10.1093/mind/lix.236.433.

Turkle, Sherry. 2011a. *Alone Together : Why We Expect More Form Technology and Less from Each
Other*. New York: Basic Books.

———. 2011b. "Authenticity in the Age of Digital Companions." In *Machine Ethics*, edited by Michael
Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Vallor, Shannon. 2018. *Technology and the Virtues a Philosophical Guide to a Future Worth Wanting*.
NY: Oxford University Press.

Vaughn, Lewis. 2015. *Beginning Ethics : An Introduction to Moral Philosophy*. New York: W.W. Norton
& Company.

Verbeek, Peter-Paul. 2014. "Some Misunderstandings about the Moral Significance of Technology." In *The Moral Status of Technical Artefacts*, edited by Peter Kroes and Peter-Paul Verbeek, 75–88. Springer. https://doi.org/10.1007/978-94-007-7914-3_5.

———. 2015. "Beyond Interaction: A Short Introduction to Mediation Theory." *Interactions* 22 (3): 26–31. https://doi.org/10.1145/2751314.

Wallach, Wendell, and Colin Allen. 2010. *Moral Machines : Teaching Robots Right from Wrong*. Oxford ; New York: Oxford University Press.

Warburton, Nigel. 2019. "Philosophy Now." Philosophynow.org. 2019. https://philosophynow.org/issues/15/A_students_guide_to_Jean-Paul_Sartres_Existentialism_and_Humanism.

Warwick, Kevin. 2012. "Robots with Biological Brains." In *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by Patrick Lin, Keith Abney, and George Bekey. MIT Press.

Whitby, Blay. 2011. "On Computable Morality: An Examination of Machines as Moral Advisors." In *Machine Ethics*, edited by Michael Anderson and Susan Leigh Anderson. NY: Cambridge University Press.

Wurzburger, Walter S. 1994. *Ethics of Responsibility : Pluralistic Approaches to Covenantal Ethics*. Philadelphia; Jerusalem: Jewish Publication Society.

Wurzburger, Walter S. 2008. *Covenantal Imperatives : Essays by Walter S. Wurzburger on Jewish Law, Thought and Community*. Edited by Eliezer L Jacobs and Shalom Carmy. Jerusalem: Urim Publications.

Xiang, Alice. 2021. "Reconciling Legal and Technical Approaches to Algorithmic Bias." *Tennessee Law Review* 88 (3). https://ssrn.com/abstract=3650635.

Xu, Lei, Chunxiao Jiang, Yan Chen, Jian Wang, and Yong Ren. 2016. "A Framework for Categorizing and Applying Privacy-Preservation Techniques in Big Data Mining." *Computer* 49 (2): 54–62. https://doi.org/10.1109/mc.2016.43.

Zimmerman, Binyamin. 2016a. "Geneivat Daat - Misleading, Misrepresentation and Deception." Torat Har Etzion. Yeshivat Har Etzion. 2016. https://etzion.org.il/en/philosophy/issues-jewish-thought/issues-mussar-and-faith/geneivat-daat-misleading-misrepresentation.

———. 2016b. "True to the Heart and Tokho Ke-Varo." Torat Har Etzion. Yeshivat Har Etzion. 2016. https://etzion.org.il/en/philosophy/issues-jewish-thought/issues-mussar-and-faith/true-heart-and-tokho-ke-varo.