

Debate: What is Personhood in the Age of AI?

David J. GUNKEL^{a,*} and Jordan Joseph WALES^b

^a*Department of Communication, Northern Illinois University, De Kalb, IL, USA*

^b*Department of Philosophy and Religion, Hillsdale College, Hillsdale, MI, USA*

AI and Law—Ethical, Legal, and Socio-political Implications

Special Issue of *AI & Society*

Ed. John-Stewart Gordon

2021

Full text available at:

<https://rdcu.be/cc00T>

PDF available by subscription or for purchase at:

<https://doi.org/10.1007/s00146-020-01129-1>

In a friendly interdisciplinary debate, we interrogate from several vantage points the question of “personhood” in light of contemporary and near-future forms of social AI. David J. Gunkel approaches the matter from a philosophical and legal standpoint, while Jordan Wales offers reflections theological and psychological. Attending to metaphysical, moral, social, and legal understandings of personhood, we ask about the position of apparently personal artificial intelligences in our society and individual lives. Re-examining the “person” and questioning prominent construals of that category, we hope to open new views upon urgent and much-discussed questions that, quite soon, may confront us in our daily lives.

References

- Anatolios K (2018) *Retrieving Nicaea: The Development and Meaning of Trinitarian Doctrine*. Baker Academic, Grand Rapids, Mich.
- Athanasius of Alexandria (1980a) *Orations Against the Arians, Book III [Selections]* [ca. 339–343]. In: Norris RA (ed) *The Christological Controversy, Re-Typeset Ed.* Fortress Press, Philadelphia, pp 65–78
- Athanasius of Alexandria (1980b) *Orations Against the Arians, Book I* [ca. 339–343]. In: Rusch WC (ed) *The Trinitarian Controversy, Re-Typeset Ed.* Fortress Press, Philadelphia, pp 55–104
- Augustine of Hippo (2012a) *The Confessions* [397–401], 2nd edn. New City Press, Hyde Park, N.Y.
- Augustine of Hippo (2012b) *The Trinity* [399–419], 2nd edn. New City Press, Hyde Park, N.Y.
- Augustine of Hippo (1996) *Teaching Christianity [De doctrina Christiana]* [396–426], 1st edn. New City Press, Hyde Park, N.Y.
- Augustine of Hippo (1887) *The City of God, Against the Pagans* [413–427]. In: *St. Augustine’s City of God and Christian Doctrine*. Christian Literature Publishing Co., Buffalo, N.Y.
- Augustine of Hippo (2004) *The Literal Meaning of Genesis [De Genesi ad litteram]* [401–415]. In: *On Genesis*. New City Press, Hyde Park, N.Y., pp 168–506

* Corresponding author. Department of Communication, Northern Illinois University, De Kalb IL, USA 60115; e-mail: dgunkel@niu.edu.

References

- Bauckham R (2008) *Jesus and the God of Israel: God Crucified and Other Studies on the New Testament's Christology of Divine Identity*. Eerdmans, Grand Rapids, Mich.
- Bendel O (2019) *The Morality Menu*. https://maschinenethik.net/wp-content/uploads/2019/12/Bendel_MOME_2019.pdf. Accessed 21 Aug 2020
- Benford G, Malartre E (2007) *Beyond Human: Living with Robots and Cyborgs*, 1st edition. Forge Books, New York
- Bentley B, Branicky R, Barnes CL, et al (2016) The Multilayer Connectome of *Caenorhabditis elegans*. *PLOS Comput Biol* 12:e1005283. <https://doi.org/10.1371/journal.pcbi.1005283>
- Block N (1978) Troubles with Functionalism. *Minn Stud Philos Sci* 9:261–325
- Bloom P (2016) *Against Empathy: The Case for Rational Compassion*. Ecco, New York
- Boethius (1973) *Contra Eutychen* [ca. 513]. In: *Theological Tractates. The Consolation of Philosophy*. Harvard University Press, Cambridge, Mass
- Bryson JJ (2015) Artificial Intelligence and Pro-Social Behaviour. In: Misselhorn C (ed) *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation*. Springer International, Cham, pp 281–306
- Bryson JJ (2010) Robots Should Be Slaves. In: Wilks Y (ed) *Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues*. John Benjamins Publishing Company, Philadelphia, Pa., pp 63–74
- Bryson JJ (2018) Patience is not a virtue: the design of intelligent systems and systems of ethics. *Ethics Inf Technol* 20:15–26. <https://doi.org/10.1007/s10676-018-9448-6>
- Bryson JJ, Diamantis ME, Grant TD (2017) Of, for, and by the people: The legal lacuna of synthetic persons. *Artif Intell Law* 25:273–291
- Buford TO (2019) Personalism. *Internet Encycl. Philos.*
- Burgess A, Plunkett D (2013) Conceptual Ethics I. *Philos Compass* 8:1091–1101. <https://doi.org/10.1111/phc3.12086>
- Carpenter J (2016) *Culture and Human-Robot Interaction in Militarized Spaces*. Ashgate, Burlington, Vt.
- Chalmers DJ (2010) *The Character of Consciousness*, 1st edn. Oxford University Press, New York
- Chalmers DJ (2011) A Computational Foundation for the Study of Cognition. *J Cogn Sci* 12:323–357
- Chalmers DJ (1997) *The Conscious Mind: In Search of a Fundamental Theory*, Revised ed. Oxford University Press, New York
- Chopra S, White LF (2011) *A Legal Theory for Autonomous Artificial Agents*. University of Michigan Press, Ann Arbor, Mich.
- Coeckelbergh M (2018) Why Care About Robots? Empathy, Moral Standing, and the Language of Suffering. *Kairos J Philos Sci* 20:141–158. <https://doi.org/10.2478/kjps-2018-0007>
- Committee on Legal Affairs (2016) *Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics*. European Parliament
- Danaher J (2020a) Welcoming Robots Into the Moral Circle: A Defence of Ethical Behaviourism. *Sci Eng Ethics* 26:2023–2049. <https://doi.org/10.1007/s11948-019-00119-x>
- Danaher J (2019) The Philosophical Case for Robot Friendship. *J Posthuman Stud* 3:5–24. <https://doi.org/10.5325/jpoststud.3.1.0005>
- Danaher J (2020b) Robot Betrayal: a guide to the ethics of robotic deception. *Ethics Inf Technol* 22:117–128. <https://doi.org/10.1007/s10676-019-09520-3>
- Darling K, Nandy P, Breazeal C (2015) Empathic concern and the effect of stories in human-robot interaction. In: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp 770–775
- Dennett DC (1998) *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press, Cambridge, Mass.
- Derrida J (2005) *Paper Machine*, 1st edition. Stanford University Press, Stanford, Calif
- Douglass F (2016) *Narrative of the Life of Frederick Douglass, an American Slave: Written by Himself* [1845]. Critical Edition. Yale University Press, New Haven, Conn.
- Dyschkant A (2015) Legal personhood: How we are getting it wrong. *Univ Ill Law Rev* 2015:2075–2110

References

- Gelin R (2016) The Domestic Robot: Ethical and Technical Concerns. In: Ferreira MIA, Sequeira JS, Tokhi MO, et al. (eds) *A world with robots (International Conference on Robot Ethics: ICRE 2015)*. Springer, New York
- Gill C (1996) *Personality in Greek Epic, Tragedy, and Philosophy: The Self in Dialogue*. Oxford University Press, Oxford
- Gordon J-S (2020a) Artificial moral and legal personhood. *AI Soc*. <https://doi.org/10.1007/s00146-020-01063-2>
- Gordon J-S (2020b) What Do We Owe to Intelligent Robots? *AI Soc* 35:209–223. <https://doi.org/10.1007/s00146-018-0844-6>
- Gregory I (1997) *Moralia in Iob; Commento Morale a Giobbe 3 (XIX-XXVII) [586–590]*. Città Nuova, Rome
- Gregory I (1992) *Moralia in Iob; Commento Morale a Giobbe 1 (I-VIII) [586–590]*. Città Nuova, Rome
- Gross T (2018) How American Corporations Had A “Hidden” Civil Rights Movement. *Fresh Air*
- Gunkel DJ (2018) *Robot Rights*. The MIT Press, Cambridge, Massachusetts
- Hamilton AF de C (2009) Goals, intentions and mental states: challenges for theories of autism. *J Child Psychol Psychiatry* 50:881–892. <https://doi.org/10.1111/j.1469-7610.2009.02098.x>
- Harnad S (1994) Computation is just interpretable symbol manipulation; cognition isn't. *Minds Mach* 4:379–390. <https://doi.org/10.1007/BF00974165>
- Harvey C (2015) Sex Robots and Solipsism: Towards a Culture of Empty Contact. *Philos Contemp World* 22:80–93. <https://doi.org/10.5840/pew201522216>
- Heider F, Simmel M (1944) An Experimental Study of Apparent Behavior. *Am J Psychol* 57:243–259. <https://doi.org/10.2307/1416950>
- Hill E (1991) Introduction. In: *The Trinity*, 1st edn. New City Press, Hyde Park, N.Y.
- Hume D (1980) *A Treatise of Human Nature [1738-40]*. Oxford University Press, New York
- Hurtado LW (2005) *Lord Jesus Christ: Devotion to Jesus in Earliest Christianity*, Paperback. Eerdmans, Grand Rapids, Mich.
- Hurtado LW (2018) *Honoring the Son: Jesus in Earliest Christian Devotional Practice*. Lexham Press, Bellingham
- Jabr F (2012) The Connectome Debate: Is Mapping the Mind of a Worm Worth It? *Sci. Am.*
- Kim J (2010) *Philosophy of Mind*, 3rd edn. Routledge, Boulder, Colo.
- Kittel G (1964) δόξα. In: Friedrich G, Kittel G (eds) *Theological Dictionary of the New Testament*. Eerdmans, Grand Rapids, Mich., pp 233–255
- Kurki VA (2019) *A Theory of Legal Personhood*. Oxford University Press, Oxford, United Kingdom
- Kurki VAJ, Pietrzykowski T (eds) (2017) *Legal Personhood: Animals, Artificial Intelligence and the Unborn*. Springer International Publishing
- Lanzoni S (2015) *A Short History of Empathy*. The Atlantic
- Leite I, Castellano G, Pereira A, et al (2014) Empathic Robots for Long-term Interaction. *Int J Soc Robot* 6:329–341. <https://doi.org/10.1007/s12369-014-0227-1>
- Leite I, Pereira A, Mascarenhas S, et al (2013) The influence of empathy in human–robot relations. *Int J Hum-Comput Stud* 71:250–260. <https://doi.org/10.1016/j.ijhcs.2012.09.005>
- Leong B, Selinger E (2019) Robot Eyes Wide Shut: Understanding Dishonest Anthropomorphism. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, pp 299–308
- Levin J (2018) Functionalism. In: Zalta EN (ed) *The Stanford Encyclopedia of Philosophy*, Fall 2018. Metaphysics Research Lab, Stanford University
- Levy D (2008) *Love and Sex with Robots: The Evolution of Human-Robot Relationships*. Harper Perennial, New York
- Locke J (1998) *An Essay Concerning Human Understanding [1689], Revised Edition*. Penguin Classics, London
- Low P (2012) The Cambridge Declaration on Consciousness. In: Panskepp J, Reiss D, Edelman D, et al. (eds). *Churchill College, University of Cambridge*
- Markoff J (2015) *Machines of Loving Grace: The Quest for Common Ground between Humans and Robots*. Ecco
- Mauss M (1985) A Category of the Human Mind: the Notion of Person; the Notion of Self. In: Carrithers M, Collins S, Lukes S, Mauss M (eds) *The category of the person: anthropology, philosophy, history*. Cambridge University Press, Cambridge

References

- Mbiti JS (1990) *African Religions & Philosophy*, 2nd edn. Heinemann, Portsmouth, N.H.
- Menkiti IA (1984) *Person and Community in African Traditional Thought*. In: Wright RA (ed) *African Philosophy*, 3rd edn. University Press of America, Lanham, pp 171–182
- Misselhorn C (2010) Empathy and Dyspathy with Androids: Philosophical, Fictional, and (Neuro)Psychological Perspectives. *Konturen* 2:101–123. <https://doi.org/10.5399/uo/konturen.2.1.1341>
- Nagel T (1974) What Is It Like to Be a Bat? *Philos Rev* 83:435. <https://doi.org/10.2307/2183914>
- Nagel T (1978) Panpsychism. In: *Mortal Questions*. Cambridge University Press, Cambridge, UK, pp 181–195
- Najork M (2016) Using Machine Learning to Improve the Email Experience. In: *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*. p 891
- Newell A, Simon HA (1976) *Computer Science as Empirical Inquiry: Symbols and Search*. *Commun ACM* 19:113–126. <https://doi.org/10.1145/360018.360022>
- Ogude J (2018) *Ubuntu and Personhood*. Africa World Press, Trenton
- Pearl J (2019) The Limitations of Opaque Learning Machines. In: Brockman J (ed) *Possible Minds: Twenty-Five Ways of Looking at AI*, 1st edn. Penguin Press, New York, pp 13–19
- Piccinini G (2010) The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism. *Philos Phenomenol Res* 81:269–311. <https://doi.org/10.1111/j.1933-1592.2010.00356.x>
- Preuss HD (1995) *Old Testament Theology*. Westminster John Knox Press, Louisville, Ky.
- Reeves B, Nass C (2003) *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, New Edition. CSLI, Stanford, Calif.
- Richardson K (2016a) Are Sex Robots as Bad as Killing Robots? In: Seibt J, Nørskov M, Andersen SS (eds) *What social robots can and should do: Proceedings of Robophilosophy 2016 / TRANSOR 2016*. IOS Press, Amsterdam, pp 27–31
- Richardson K (2016b) Sex Robot Matters: Slavery, the Prostituted, and the Rights of Machines. *IEEE Technol Soc Mag* 35:46–53. <https://doi.org/10.1109/MTS.2016.2554421>
- Richardson K (2015) The Asymmetrical ‘Relationship’: Parallels Between Prostitution and the Development of Sex Robots. *SIGCAS Comput Soc* 45:290–293. <https://doi.org/10.1145/2874239.2874281>
- Rist JM (2020) *What is a Person?: Realities, Constructs, Illusions*, 1st edn. Cambridge University Press, New York
- Russell S, Norvig P (2009) *Artificial Intelligence: A Modern Approach*, 3rd edn. Pearson, Upper Saddle River
- Salmond JW (1902) *Jurisprudence, Or the Theory of the Law*. Stevens and Haynes, London
- Sandry E (2015) *Robots and Communication*. Palgrave Pivot, New York
- Schafer WR (2018) The Worm Connectome: Back to the Future. *Trends Neurosci* 41:763–765. <https://doi.org/10.1016/j.tins.2018.09.002>
- Schweizer P (2019a) Triviality Arguments Reconsidered. *Minds Mach* 29:287–308. <https://doi.org/10.1007/s11023-019-09501-x>
- Schweizer P (2019b) Computation in Physical Systems: A Normative Mapping Account. In: Berkich D, d’Alfonso MV (eds) *On the Cognitive, Ethical, and Scientific Dimensions of Artificial Intelligence: Themes from IACAP 2016*. Springer International Publishing, Cham, pp 27–47
- Searle JR (1980) Minds, Brains, and Programs. *Behav Brain Sci* 3:417–457
- Searle JR (1992) *The Rediscovery of the Mind*. MIT Press, Cambridge, Mass.
- Siedentop L (2014) *Inventing the Individual: The Origins of Western Liberalism*, 1st edn. Belknap Press, Cambridge, Mass.
- Singer P (2009) Speciesism and Moral Status. *Metaphilosophy* 40:567–581
- Solum LB (1992) Legal Personhood for Artificial Intelligences. *N C Law Rev* 70:1231–1287
- Spaemann R (2006) *Persons: The Difference between “Someone” and “Something.”* Oxford University Press, New York
- Strawson PF (1959) *Individuals: An Essay in Descriptive Metaphysics*. Methuen & Co., London
- Stueber K (2019) Empathy. *Stanf. Encycl. Philos.*
- Taylor C (1985) The Person. In: Carrithers M, Collins S, Lukes S (eds) *The Category of the Person*, First Paperback Edition. Cambridge University Press, Cambridge, pp 257–281

References

- Tegmark M (2017) *Life 3.0: Being Human in the Age of Artificial Intelligence*. Knopf, New York
- Tertullian of Carthage (2011) *Against Praxeas* [ca. 213]: The Text Edited, with an Introduction, Translation, and Commentary, Bilingual, Reprint Edition. Wipf & Stock Pub, Eugene, Ore.
- Turing AM (1950) Computing Machinery and Intelligence. *Mind New Ser* 59:433–460
- Turner J (2018) *Robot Rules: Regulating Artificial Intelligence*, 1st edn. Palgrave Macmillan
- United States Supreme Court (1964) *Jacobellis v. Ohio*, 378 U.S. 184 (1964)
- Van Nieuwenhove R (2017) Contemplation, Intellectus, and Simplex Intuitus in Aquinas: Recovering a Neoplatonic Theme. *Am Cathol Philos Q* 91:199–225. <https://doi.org/10.5840/acpq2017227108>
- Wales JJ (2018) Contemplative Compassion: Gregory the Great’s Development of Augustine’s Views on Love of Neighbor and Likeness to God. *Augustin Stud* 49:199–219. <https://doi.org/10.5840/augstudies201861144>
- Wiener N (1988) *The Human Use Of Human Beings: Cybernetics And Society*, New Edition. Da Capo Press, New York
- Williams TD, Bengtsson JO (2018) Personalism. *Stanf. Encycl. Philos.*
- Žižek S (2006) Philosophy, the “unknown knowns,” and the public use of reason. *Topoi* 25:137–142. <https://doi.org/10.1007/s11245-006-0021-2>
- (2016) Open letter to the European Commission (<http://www.robotics-openletter.eu/>)