Smart Machines and Wise Guys: Can Intelligent Machines be Wise?

Edward Spence, University of Sydney, Australia

Extended Abstract:

This paper will address an important topic in the Philosophy and Ethics of Technology, which is that of *artificial mediated agency and autonomy* and its moral significance and impact on human wellbeing. Furthermore, it will address the key question of this paper whether artificial agency and autonomy can be *extended to the notion of wisdom*: if agency is mediated and distributed between humans and intelligent machines, can there also be mediated wisdom? Although several well-known philosophers of technology, such as Luciano Floridi (2014 and 2013), Peter-Paul Verbeek (2014, 2011), Richard Heersmink, (2017), to mention but a few, have researched and written on the issue of *artificial moral mediation*, very little research has been published on the topic of wisdom with regard to AI technology and its impact on wellbeing. This paper sets out to address and fill that gap.

According to Verbeek, "only by recognizing this interweaving of human and technologies can we take responsibility for the ways in which *technologies have an impact on society and on human existence* [my emphasis]– in practices of technology design, implementation, and use" (2014: Ch 5, p.76). In a clear summary Heersmink summarizes Verbeek's distributed moral agency (2017; 473) thus:

> By drawing on actor-network theory, Verbeek (2011) develops the notion of "distributed moral agency". He specifically focuses on the moral aspects of obstetric ultrasound imaging, an imaging technology used to visualize a foetus in the mother's womb, which is usually done a week twelve and twenty of pregnancy. This technology does not just provide a neutral and transparent window to the womb, but has several morally important aspects…so obstetric ultrasound imaging mediates and transforms the relationship parents have with their foetus, providing them with information on the health of their foetus that is important for moral decision -

making regarding pregnancy and abortion. For these reasons, moral agency should be seen as distributed across humans and technology. In Verbeek's words: Ultrasound imaging actively contributes to the coming about of moral actions and the considerations behind moral actions. The example therefore shows that moral agency should not be seen as an exclusive human property; it is distributed among human beings and nonhuman entities (2011, p. 38). In a later chapter he points out: Moral mediation always involves an intricate relation between humans and nonhumans, and the "moral agency" that results from this relation therefore always has a hybrid rather than a "purely human" character….moral agency is distributed among humans and nonhumans; moral actions and decisions are the product of human-technology associations (2011, p.53).

According to Heersmink, "Verbeek…argues that moral agency is distributed across humans and technology, they cannot be understood in isolation from each other because they are integrally connected. So, one way to understand the moral relevance of technological artefacts is to say that artifacts are co-constitutive of moral agency" (2017, 438).

However, Verbeek, Floridi, Heersmink and the other aforementioned philosophers, do not refer to the important role that wisdom should play in recognizing and assessing the mediated impact of technology on wellbeing. It is precisely that missing inclusion of wisdom and its gap that this paper intends to address. Namely, to examine and evaluate the impact of mediated AI technologies on our wellbeing, in terms of our agency, autonomy, morality, and most importantly as we shall examine in this paper our human dignity, and do so through the concept of wisdom. One key question this paper will examine is whether there can be mediated wisdom along the lines that Floridi, Verbeek and Heersmink apply to mediated moral agency or is some other way that retains the important role of human autonomy and dignity, which is of central importance to human agency. book.

In an important innovative paper "Artificial wisdom: a philosophical framework" Cheng-hung Tsai examines the notion of artificial wisdom (2020). In his paper he explains "why artificial wisdom matters and how artificial wisdom is possible (in principle and in practice)….the result is a conceptual framework that guides future research on creating artificial wisdom" (2020, p. 937). Contrary to that view, Jason Swartwood and Valerie Tiberius (chapter 2, 2019:26- 27)) argue that:

A wise person's understanding is *uncodifiable:* it is possible to derive good guidance in any particular situation (Hursthouse 1999: 39-40). Morality is too complex , Virtue Ethicists argue, to capture with such a set of principles. For instance, someone might try to defend the following moral principle: An action is right when, of all the actions possible in a situation, it could be expected to produce the most total long-term happiness when the effects on everyone are added up. If this principle accurately described what matters in all situations, then a non-virtuous person could use it to derive (after suitable inquiry into the likely consequences of different actions) reliable guidance about what to do. Virtue Ethicists argue that principles like this do not succeed in codifying wise understanding. Julia Annas provides additional arguments against codifiability (Annas 2004). If Virtue Ethicists are correct that wise understanding is uncodifiable, then that would imply that a reliable understanding of how to conduct oneself must be acquired through some sort of experience and reflection….Aristotle thought we need reflection that focuses on both the universal and the particulars of situations (NE 1141b10-15, 1142a14).

Swartwood and Tiberius raise an important characteristic of wisdom that resists codifiability, that is, wisdom is a key characteristic of Virtue Ethics and by extension Stoic philosophy and Hellenistic philosophy generally, which is that wisdom is a way of life for the attainment of eudaimonia and thus specific to individual lives. This paper will examine in detail whether even if uncodifiable, wisdom, however, could be mediated in some way to be defined, through AI technology as suggested with regard to morality by Verbeek and Floridi, whilst at the same time, preserving the moral agency of wise human beings that remains under their full control.

References

Brey, Brey., Briggle, A. and Spence, E. (2012) (Eds.) *The Good Life in a Technological Age .* New York, London: Routledge, Taylor & Francis Group.

Floridi, L. (2014) *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford: Oxford University Press.

Heersmink, R. (2017) Distributed selves: personal identity and extended memory systems. *Synthese* (2017) 194:3135–3151 DOI 10.1007/s11229-016-1102-4.

Spence, E. (2022). Promoting the Human Good: The Dual Obligation Wisdom Theory and the Duties of Ethics, in Stephen Ward (ed,), *The Handbook of Global Media Ethics*

Spence, E. (2022). The Privacy Paradox, in Michael Boylan and Wanda Teays (eds.) *Ethics In The AI, Technology, And Information Age*. Lanham: Rowman & Littlefield.

Spence, E. (2021) *Stoic Philosophy and Technology*: *Caught in the Web*. Rowman and Littlefield.

Spence E. (2020) The sixth estate; tech media corruption in the age of information, *Journal of Information, Communication and Ethics in Society,* 2020, Emerald Insight at: https://www.emerald.com/insight/1477-996X.htm

Spence, E. (2018) Us and Them: The Ethics of AI, Proceedings for the Annual OZSW2018 Conference November 9 and 10, University of Twente, Netherlands.

Hui, J. and **Spence, E**. (2016). Internet Addiction and Well-Being: Daoist and Stoic Reflections. *Dao: A Journal of Contemporary Philosophy.* Springer.

Spence, E. (2014). Wisdom and Well-being in a Technological Age, *Proceedings **of** Well-Being, Flourishing, and ICTs*: *Annual Conference of Computer Ethics and Philosophical Enquiry (CEPE) 2014,* Paris, France, June 22-25, 2014.

Spence, E. (2013) Wisdom and Well-being in a Technological Age, *Synesis: A Journal of Science, Technology, Ethics, and Policy* 2013.

Spence, E. (2011) Is Technology Good for Us? A Eudaimonic Meta-Model for Evaluating the Contributive Competence of Technologies for a Good Life. *Nanoethics,* 5:335–343.

Spence, E. (2011a) Information, Knowledge and Wisdom: Groundwork for the Normative Evaluation of Digital Information and Its Relation to the Good Life, *Ethics and Information Technology*, Volume 13, Number, 3, Pages 261-275.

Tsai, Chen-hung (2020). Artificial wisdom: a philosophical framework, *AI & Society* volume 35, pages 937-944. Springer.

Swartwood, J. and Tiberius, V. In Robert Sternberg & Judith Gluek (eds.), *The Cambridge Handbook of Wisdom (*2019). Cambridge, UK: Cambridge University Press. pp. 10-39.

Verbeek, P.P. (2014*) The Moral Status of Technical Artefacts**. (Eds. Peter-Kroes and Peter-Paul Verbeek, Springer.*

Verbeek, P.P. (2011*). Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago: The University of Chicago Press.