„Ethos" 36(2023) nr 4(144) 223-230

Łukasz MŚCISŁAWSKI

WHEN WISDOM AND BIGTECH-AI COLLIDE

> Where is the wisdom we have lost in knowledge?[1]
> Thomas Stearns Eliot

> In the age of information, it seems that we would be better off with more wisdom and a little less information. Switching over our smartphone and tablet to Plato and Epictetus, sometimes, may be a good start.[2]
> Edward Spence

The subject of artificial intelligence (AI) provokes a wide variety of reactions—from extreme enthusiasm for the possibilities of this technology in the daily press to Cassandra—like visions almost straight out of the film *Terminator*. The fact that AI is becoming more and more present in new areas of human functioning provokes discussions on both the potential benefits and possible dangers.

In this context, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web* by Edward H. Spence is a position worthy of attention.

It must be stressed that the title is somewhat misleading. Spence addresses a much broader range of issues than dealing with a philosophical analysis of narrow, technological issues related to AI, while ultimately focusing indeed on issues related to this technology. Place them, however, in the much broader perspective of a specific view of human beings and how they function, in contact with technology in general and AI technology in particular.

Spence considers two most pressing issues affecting human existence: (1) climate change and (2) the rapid technological advances in the field of autonomous AI agents for information processing (see p. 1).[3] The author attempts to unravel the latter issue in a rather surprising way by asking what technology is good for and trying to explore the impact of technology on human well-being, especially when it comes to information processing technologies (including AI). Thus, it can be said that the starting point is not technology, but the human being affected by that technology. Human beings—and therefore, as one might suspect, thinking beings. It is precisely the defence of human thinking

[1] Thomas Stearns E l i o t, *The Rock* (New York: Harcourt, Brace, and Company, 1934), 7.

[2] Edward S p e n c e, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web* (Lanham: Rowman & Littlefield, 2021), 93.

[3] The page numbers in parentheses refer to the discussed work by Spence.

and functioning, and the defence of human dignity, that the author is concerned with. All of this takes place against the backdrop of the question of who has power over information and the technologies of producing and disseminating information, above all AI, used by the giants of the IT market, referred to as BigTech.[4] When it comes to the immersion of modern man in the infosphere, the problem, as the author notes, is not so much the use of various types of information processing technologies themselves, but how, by whom and for what purpose they are used. By virtue of the amount of information and the speed and range of its propagation, the author seems to conclude that the problem is not one of an individual or local community, but a global one. Taking also into account that it is not a purely technical problem, he suggests that this issue has also a moral aspect and should be solved collectively by the entire global community[5] (see p. 59f.). The question that arises here is whether there are not already some means in the history of human thought that would make it easier, or, in fact, possible, to analyse such issues from an ethical perspective? After all, human nature has not fundamentally changed in the last hundreds or thousands of years. Does it automatically follow from the fact that AI is a contemporary technology that such means have yet to be developed? And if they are already available, could they be modified and applied to the issues mentioned? Does the rapid development of technology mean that it is only possible to analyse the situation ex post, which, due to the dynamics of this development, results in perpetual lagging behind? Or, on the other hand, is it possible to develop tools of reflection that anticipate and modify the way advanced research on modern technologies is designed and, at the same time, make it possible to react in advance to the undesirable effects of the use of these technologies, especially in the area of information processing in the broadest sense? Edward H. Spence's proposal seems to be an interesting attempt to find answers to such questions.

A brief reconstruction of Spence's proposal is as follows. Because of the increasing accumulation of information and the increasing impact of technology on human beings and their well-being some kind of theoretical tool is needed to assess the impact of technology on human quality of life. Such a theory should be able to determine whether and how technology can assist humans in enhancing their quality of life (see p. 1). It is important to emphasise that Spence's posing the question in this way opens the way to the question of whether any technology is able to contribute to the well-being, ethical quality of a person's life or their quest for fulfilment. It is a question that, it seems, is too rarely asked in the context of contemporary technological developments.

The problem, of course, is to answer the question of what a good life is. Spence goes in favour of the eudaimonic direction here, stating that a good life is one that at least minimally allows a person to achieve self-fulfilment, well-being, or simply happiness (see p. 15). A theory that could be used for the aforementioned evaluation must meet the relevant formal and material conditions (see p. 15f).

The pursuit of the good life—or, ultimately, happiness—when confronted with an ocean of information is not so simple. Spence argues that there is no simple transition between taking in information and li-

[4] This usually includes technology giants such as Amazon, Apple, Facebook, Google, and Microsoft.

[5] Here is one of the weaker points of Spence's work. For it is difficult to consider that there is any homogeneous global community. Some of the inhabitants of our planet may not be interested in information processing problems at all.

ving well. Hence, a tool is needed to assess whether a given piece of information (or technology) will contribute to the good life. Such a tool, for him, is wisdom (see p. 18). He sees its role in multiple ways. Wisdom is thus a certain meta-knowledge, a principle of evaluation (ibidem), but it can also be seen as a meta-technology of the self (see p. 20).

In the context of the search for an appropriate criterion for evaluating technology, including communication and information processing together with AI, Spence presents the views of the Stoics and Neo-Stoics. In doing so, he emphasises that, according to the Stoics, a virtuous life and wisdom are necessary and sufficient conditions for happiness. He also juxtaposes the views of the Stoics and neo-stoics with those of Gewirth as a representative of the rationalist movement in the context of the key points: (a) the ultimate goal; (b) a good life in harmony with nature, (c) happiness, (d) the relationship between virtue and eudaimonia. In doing so, he notes that the latter's views coincide with those of the Neo-Stoicists. According to Spence, particularly significant here is the similarity between the concept of self-fulfilment and the Stoic concept of eudaimonia. In the case of (d), the juxtaposition of Gewirth and the Stoics is very interesting, as both are concerned to show the role of the conformity of life to the requirements of reason. Here, however, Spence emphasises the role of the principle of general consistency in Gewirth, which is to guarantee the achievement of the best possible version of oneself in conformity with human rights and leads to a view of man as someone whose nature is to act rationally, purposefully (see p. 39f.). Another important point for Spence is that, in order to avoid frustration or disappointment, the Stoics propose doing two things: controlling those things that are within our control, and remaining unmoved by the things we cannot control (see p. 31). This explains, at least in part, why Spence is interested in the views of the Stoics, the Neo-Stoics and the views of Gewirth, which he believes to be strongly aligned with these two currents. This is because the fundamental issue at the heart of the author's interest is the problem of control, especially in the context of modern technology and its impact on humans (see p. 60). Here, Spence highlights the fact that by treating philosophy as a kind of therapy of the mind, they have noticed that there are things we can control and things we cannot control. In order to be able to achieve happiness, he proposes, in relation to technology, the application of the basic Stoic principles of the good life: (a) dealing only with things over which we have control; (b) considering the ultimate goal of life—that is, achieving happiness; (c) living a virtuous life and living wisely; (d) cosmopolitanism; (e) living in harmony with nature; (f) growing in virtue and internalising it, (g) treating philosophy as a way of life (see p. 62–69). As Spence points out, in the digital age, Stoic philosophy provides a systemically unified practical proposal for living fruitfully and achieving happiness. This is made possible by epistemic, ethical and eudaimonic principles centred on the concept of wisdom. It also provides the means, both theoretical and practical, needed to address the problem of control in the context of contemporary technology, especially in relation to information processing and AI (see p. 70). Spence devotes the first four chapters of his book ("Introduction: Who Is in Control?", "What Is Technology Good For", "Stoic and Neo-Stoic Philosophy", and "Application of Stoic Philosophy to Technology") to outlining the problems and preparing the relevant tools, based mainly on the two philosophical currents mentioned above.

In chapter five ("Wisdom and Well-Being: The Dual Obligation Information-Wisdom Theory"), Spence offers the reader

a particularly useful tool—the Dual Obligation Information-Wisdom Theory (DOIT-Wisdom Model). It allows issues related to technology to be framed and analysed in two categories: epistemic and ethical. The author devotes the fifth chapter to this. Using previously developed tools provided by the Stoics and Neo-Stoics, he shows wisdom as a kind of meta-information and meta-knowledge, which include epistemic, ethical, and eudaimonic dimensions, and also provides the possibility to capture the link between information, knowledge, good life, and well-being. In general, it allows to combine into a conceptual scheme to find a direct link between information and communication technologies and AI technologies and well-being and good life. The proposed concept of freedom also provides normative criteria for assessing the dissemination and impact of information (see p. 74f.).

Central to this is the DOIT-Wisdom Model already mentioned (see p. 90f.). Within it, Spence proposes to link the concepts of information and wisdom with values into a coherent whole. The context of treating information as a message allows him to link issues of information processing to the values of truthfulness, reliability, and robustness. The context of treating the communication of information as a kind of action, he links to the preservation of the right to freedom and well-being. Treating wisdom as a kind of meta-informational action, he links it to epistemic, ethical and eudaimonic values. In doing so, it is noteworthy that he proposes a shift from epistemic to eudaimonic values (see p. 21) when dealing with issues related to the impact of technology on humans and the way they function. The whole is motivated by Spence's reference to three aspects. The metaphysical (or ontological) aspect is related to the fact that nowadays man lives not only in the biosphere, but also in the infosphere, and that information and its processing increasingly permeate his action (see p. 91). He refers to the second aspect as technological: the digitalisation of information allows for the instantaneous creation and dissemination of information. This can have a positive impact, but also a negative one, e.g., due to the unreflective use of this possibility. The latter possibility instantly results in an impact on a person's existential situation, which Spence illustrates with numerous examples. This suggests the necessity for a culture to learn how to handle information, hence also his encouragement to use tools that are already present in human culture—e.g., his proposed use of Stoic and Neo-Stoic philosophy (see p. 93). A third aspect that Spence proposes to consider is that prudence and foresight are needed in dealing with information and the means of processing it (see p. 94). The entire chapter is interspersed with insights into the role of major corporations in information processing and uptake. Spence points out that, in fact, information or content that reaches us is very precisely selected with the tools used by the digital giants. The problem is that these tools operate in a non-transparent way, and it is not clear who could be held accountable for their actions. In places, his observations may seem too intrusive to the reader, but it is nevertheless difficult to refuse to be right (see p. 95).

Chapter Six ("Tech Media Corruption in the Age of Information") is devoted to discussing the irregularities of the information processing giants. The starting point here is to present the implicit part of their business model, to compare them with the operation of traditional media, and to point out the places where the interests of users of information processing platforms and their owners are at odds with each other and the radical, albeit implicit, violation of the social order. These include, in particular, violations of the rights to: informed consent, privacy and autono-

my (see p. 111). The basic tool that the author uses is the DOIT-Wisdom model proposed earlier, supported by the identification of the basic characteristics of corruption: the possession of power, the disposition to exercise that power, the opportunity to use it, a kind of invisibility of actions or their concealment, consideration only of one's own gain, violation of the duty of the fiduciary—in this case, the data, provided by the users (see p. 112f.). Spence's analysis of the case of Facebook's involvement in the Cambridge Analytica affair leads the author to conclude that the business model of surveillance used has resulted in a corruption of communication and information, a corruption that violates democratic principles and strikes at society (see p. 117). In this context, he asks a very important question: what kind of society are such models of so-called Big-Tech functioning creating? He also stresses that the real problem is precisely the methods and practices of non-intrusive, non-transparent, unjust, and unaccountable (or accountable) collection and use of user data (see p. 120). An additional difficulty here is the invisibility of these practices (such as the transfer of user data to advertising companies or the preparation of information) to the user. Hence, the question is whether humans can be happy in a world ruled by Technological Oligarchs? Moreover, this raises the question of who actually controls our autonomous intentional action, which is a necessary condition for well-being (see p. 121). The problem is acute because factual knowledge of the world is part of the information on which we make decisions. If this information is appropriately and invisibly crafted, then we find ourselves in a situation where we are not even in control of our decision-making process. On the basis of Spence's reflections, another question can be posed: will we then have power over our own interiority, our own intellectual sphere, and to what extent? However, Spence suggests that a collective multidisciplinary approach, involving the community and the relationships within it, could be a solution of the problem.

The next two chapters ("Normative Impact of ICT Technologies on Well-Being" and "The Normative Impact of AI Technologies on Well-Being") deal with this issue. In chapter seven, the author places particular emphasis on autonomy, integrity, and dignity, as seen in the context of achieving self-fulfillment or the pursuit of happiness. In practical terms, this means the well-understood self-control characteristic of the Stoics (see p. 127). This self-control is in some danger due to BigTech's actions, but it can be sustained at least at the level of our information generation. Therefore, one of the measures that can prevent an undesirable situation is paying attention to the dignity of the person or—alternatively—according to Spence—self-respect (see p. 129–33). The combination with Gewirth's principle of general consistency leads to very interesting conclusions, namely that the basic category to which attention now needs to be drawn is the category of human dignity, closely related to the dignity of a specific person. It is this category, in relation to issues related to the control of information, that Spence proposes to take special seriousness (see p. 137–46). Moreover, the previously introduced principles of Stoic philosophy and looking at the human situation from this very perspective give us the basis to see in the light of reason this inviolable dignity of people who are not alone—it is united by something that Gewirth defines as a community of rights (see p. 136).

The previously prepared tools allow the author in chapter eight to face the problem of controlling or influencing an individual or society, also on a global scale. This issue is discussed in the context of the impact of information processing technology and AI,

in particular the use of two techniques: machine learning and neural networks. With regard to the latter, Spence draws attention in particular to the black box problem—a situation in which it is not known exactly on what basis a given technology obtains its output data.[6] Such a situation means that the lack of transparency of the functioning of such technologies results in the impossibility of assigning responsibility for the decisions made. This, in turn, leads to a situation where it is impossible to monitor AI-based systems in order to eliminate the results of their actions potentially harmful to society (see p. 149).

In this context, Spence emphasizes (see p. 150) that the monopolistic practices of the IT industry giants undermine the individual power of individuals over their own will. People's sovereignty is threatened. Referring to the words of Luciano Floridi, the author emphasizes the need to rethink the very concept of sovereignty (see p. 151). In a very clear way, this ultimately leads to the issue of protecting our individual sovereignty, especially in the age of global information processing and dissemination. This is an urgent problem as individual sovereignty is one of our natural rights, essential to our freedom, well-being and dignity as persons.

A major problem with the lack of transparency with regard to machine learning is that: (1) there is some degree of loss of control over the activities of technology companies by their human workers; (2) through invasive persuasive algorithms, operating in an invisible way, BigTech companies have an impact on decision-making by users (see p. 153). Spence lists a whole list of dangers of such a situation, partly resulting from the very nature of the operation of this type of technology: inconclusive evidence, inscrutable evidence, misguided evidence, unfair outcomes,[7] transformative effects (see p. 153–62) traceability.

A separate problem of key importance that appears in the context of functioning in cyberspace is a kind of disintegration of identity, and more specifically of our identity data. Information processing systems, based on the data held, create a certain representation of the user, and a group digital identity over which these systems have control. Usually, it is not complete, and therefore it is in contradiction with our natural identity, which, unfortunately, also becomes dangerous under the influence of persuasive BigTech techniques. This causes a fundamental paradox: the autonomy of our natural identity is threatened by persuasive technologies that are callously used to shape the digital group identity (see p. 161f.).

In this context, Spence emphasizes the importance of the presented model of neostoic philosophy, based on the rationalist ethics of laws. This is because it enables us to identify and assess the violation of our autonomy (see p. 164). He also emphasizes that with the help of modern technologies, BigTech companies systematically undermine and violate our ethical and cognitive rights, as well as the common good and well-being—both for individuals and for society (see p. 165). Another very important issue arises here, introduced by, so to speak, immoderate enthusiasm for modern information processing technologies. This is the trend of trusting the results of programs more than individuals or the community. This leads to an increasing dehumanization of both individuals and

[6] This is a known problem in which AI algorithms can report the outcome of some decision-making process, but neither the systems nor their designers can justify this and not the other outcome.

[7] Outcomes of work of given neural network is strongly dependent on way in which this network is trained. Here appears another problem: who and how is training neural networks.

communities (see p. 165f.). It should also be emphasized here, which, unfortunately, Spence does not do, that unlimited trust in the results presented by IT systems is at best imprudence and getting rid of the habit of critical thinking.[8] In the whole of this puzzle, Spence emphasizes the fact that—based on the social commitment of the Stoic philosophy—greater social control of the functioning of BigTech companies is needed[9] (see p. 166–73).

Spence's very valuable contribution is drawing attention to the fact that in the era of mass digital information processing, the use of the concept of the dignity of a person in the context of functioning in the infosphere should be restored. He emphasizes that AI algorithms are based on statistical inference. Meanwhile, statistical correlations are not yet causal relationships. Hence, among other things, there is a conflict between our digital representation created by AI algorithms and our real identity. In this context, Spence suggests referring to the concept of the dignity of the person in order to provide us with certain digital rights that could not be appropriated by BigTech companies,[10] and the right to identity. The last postulate can be realized if the algorithms of artificial intelligence, especially machine learning, are trained taking into account the norm of preserving the dignity of a person (see p. 173–85). It seems that another difficulty arises here: clarifying how the dignity of a person should be understood and how to implement their norms in neural network training procedures.

In chapter nine ("Smart Machines and Wise Guys") Spence is outlining a certain future perspective, a kind of modus procedendi for the future. It begins with a short summary of the considerations so far (see p. 189–92). Next, Spence points to the predatory policy of BigTech, which tries to change our preferences and the basis of autonomous choices (see p. 192f.). This policy is based on the use of AI technology. So the question is, what will happen if super-AI technologies are developed? Interestingly, Spence shows a certain amount of optimism here and seems to have a somewhat similar approach to the matter here as Wooldridge.[11] He notes that we do not need to be particularly afraid of this, as long as we act prudently. More precisely, instead of being afraid of what AI will do, we need to teach AI what our goals are, learn to include and keep our goals in the solutions it searches for (see p. 193–97). However, the problem may arise at the meta-level: how to take into account the dignity of the person already when designing AI systems and their implementation, so that two classes of people are not formed: an elite, operated by people, and a lower class, managed by machines, what is the real threat (see p. 197f.)? From a technological point of view, a value-sensitive design (VSD) approach can be considered here (see p. 217). However, on the human side, perhaps our expectations and goals need to be changed. While the influence

[8] Suffice it to say that the program itself may work fine, but the basic ideas behind creating an algorithm, assumptions about what it should do and how, and the interpretation of its results, may be wrong. Interestingly, Spence points out to this, using his experience as an accountant and auditor, that it is important not only that the results match, but also that the methodology used by the company is correct (see p. 168–69).

[9] It can be seen as a kind of irony that similar statements can be found in the statements of BigTech representatives, such as Mark Zuckerberg, cited by Spence (see p. 166f.).

[10] H e n c e, some international regulations, such as GDPR-EU.

[11] Michael W o o l d r i d g e, *The Road to Conscious Machines: The Story of AI* (Dublin: Pelican Books, 2021).

of machines is essentially external to us, as Spence points out, we do have power over our inner self. Therefore, we need, in line with the indications of Stoic philosophy, to learn internal control, including the already mentioned modification of our expectations and goals (see p. 203). This in turn leads to the suggestion of a symbiotic relationship between us and machines (see p. 214) and thinking in terms of designing machines that accomplish our goals and defer to man (see p. 215).[12] Such an approach to our attitude, however, requires wisdom, able to comprehend our goals, expectations, laws and knowledge about the reality around us into one coherent whole. Due to the global scale of the impact of AI-based technologies used by BigTech, it would require some kind of cultural movement, also of a global nature, promoting the ethos of life based on wisdom—perhaps also based on the foundations of stoic thought (see p. 208).

Unexpectedly, Spence emphasizes the role of the concept of love in this kind of ethos and the relationship between people and machines,[13] by analogy to the literal understanding of the term philosophy, but also by emphasizing its importance in various religious traditions, including the Christian vision of the relationship between God and man (see p. 224).

Edward Spence's book is undoubtedly unique. The author skillfully tries to apply the tools available in philosophical culture since antiquity to solve one of the most pressing problems of contemporary times, related to the development of AI technology and its almost ubiquitous presence. The solidity of building the arguments and the clarity of the whole should be emphasized. Undoubtedly, the enrichment of the whole is also the reference to abundant literature. However, the question arises whether, with all the solidity of the argument, it can be persuaded to change the way of functioning and thinking of a person who is accustomed to the current situation and who is simply comfortable in this situation? It seems that this sense of comfort, or simply the modern pace of life, can very effectively prevent both reflection on the problems related to the increasing interpenetration of human life and modern technology, as well as changing this situation to one more compatible with human nature.

It seems that for anyone who is familiar with the ethical side of the development of modern technologies and some sort of ordering and taming this extremely complex issue, Spence's book should be one of the basic readings.

Contact: Faculty of Management, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: lukasz.mscislawski@pwr.edu.pl

[12] It is worth noting that similar conclusions were presented by Paweł Polak and Roman Krzanowski (see Paweł P o l a k and Roman K r z a n o w s k i, "How to Tame Artificial Intelligence (AI)? A Symbiotic AI Model for Beneficial AI," Ethos 36, no. 3 (143) (2023): 92–106).

[13] It also raises here the intriguing question, can AI, or super-AI, be taught some kind of love for humans?