distribuzione 747 - anno CLIV - fascicolo 3 2022 - I.S.S.N. 0391 5646

Estratto

ARCHIVIO GIURIDICO

Filippo Serafini

dal 1868

STEM Mucchi Editore

I.S.S.N. 0391 5646

ARCHIVIO GIURIDICO

Filippo Serafini

dal 1868

già diretto da
GIUSEPPE DALLA TORRE

Direzione
GERALDINA BONI
Ord. Università di Bologna

Anno CLIV - Fascicolo 3 2022

STEM Mucchi editore

Archivio giuridico Filippo Serafini - ISSN 0391 5646

Amministrazione: Stem Mucchi editore S.r.l.
Direzione: Via Zamboni, 27/29 - 40126 Bologna
Redazione: Via Zamboni, 27/29 - 40126 Bologna; Via della Traspontina, 21 - 00193 Roma
Autorizzazione: del Tribunale di Modena, n. 328 dell'11-05-1957
Direttore responsabile: Marco Mucchi

Periodico trimestrale, prezzi abbonamento

Formato cartaceo Italia	€ 114,00
Formato cartaceo estero	164,00
Formato digitale (con login)	98,00
Formato digitale (con ip)	107,00
Formato cartaceo Italia + digitale (con login)	136,00
Formato cartaceo estero + digitale (con login)	185,00
Formato cartaceo Italia + digitale (con ip)	145,00
Formato cartaceo estero + digitale (con ip)	194,00
Fascicolo singolo cartaceo*	30,00
Fascicolo singolo digitale	25.00

Tutti i prezzi si intendono iva e costi di spedizione inclusi. *Escluse spese di spedizione.

L'abbonamento decorre dal 1° gennaio di ogni anno e dà diritto a tutti i numeri dell'annata, compresi quelli già pubblicati. Al fine di assicurare la continuità nell'invio dei fascicoli gli abbonamenti si intendono rinnovati per l'annata successiva se non annullati (tramite comunicazione scritta a info@mucchieditore.it) entro il 31 dicembre del corrente anno. I fascicoli non pervenuti all'abbonato devono essere reclamati entro 10 giorni dal ricevimento del fascicolo successivo. Decorso tale termine si spediscono, se disponibili, contro rimessa dell'importo (più spese di spedizione). Per ogni effetto l'abbonato elegge domicilio presso l'amministrazione della Rivista. Le annate arretrate sono in vendita al prezzo della quota di abbonamento dell'anno in corso. Si accordano speciali agevolazioni per l'acquisto di più annate arretrate, anche non consecutive, della Rivista.

Il cliente ha la facoltà di revocare gli ordini unicamente mediante l'invio di una lettera raccomandata con ricevuta di ritorno alla sede della Casa editrice, o scrivendo a info@pec.mucchieditore.it entro le successive 48 ore (identificazione del cliente e dell'ordine revocato). Nel caso in cui la merce sia già stata spedita il reso è a carico del cliente e il rimborso avverrà solo a merce ricevuta Per gli abbonamenti eventuale revoca deve essere comunicata entro e non oltre il 7° giorno successivo alla data di sottoscrizione.

Via Jugoslavia, 14 - 41122 Modena - Tel. 059.37.40.94
e-mail: info@mucchieditore.it - info@pec.mucchieditore.it
indirizzi web: www.mucchieditore.it - www.archiviogiuridiconline.it
facebook - twitter - instagram

Tipografia, impaginazione, web: Stem Mucchi Editore (MO). Stampa: Legodigit (TN).

Finito di stampare nel mese di ottobre del 2022.

Archivio giuridico, anno CLIV, fasc. 3 2022, pp. 559-575

Laura Palazzani

TECHNOLOGICAL HUMANISM: A PHILOSOPHICAL FRAMEWORK FOR EDUCATION*

Sommario: 1. What remains of being human in the age of AI and robotics. – 2. The artificial 'beyond' the natural: the techno-centric perspective. – 3. Humans in the age of AI: the human-centric perspective. – 4. For technological humanism: ethical requirements for using AI and robotics. – 5. Educating to acquire critical awareness (AI-robotic literacy).

1. *What remains of being human in the age of AI and robotics*

If in the past bioethics focused its attention, in the context of scientific-technological progress, on the invasiveness of technology on the body and mind (discussing proportion and disproportion in patient care), the relationship between humans-animal and humans-environment, today occupying centre stage in the philosophical discussion in the field of emerging technologies is the 'new technological wave', the human-machine relationship and human enhancement.

The rapid evolution of robotics and artificial intelligence (AI) technologies in recent years has been characterised by 'disruptive' progress, due to their complexity, breadth of applications, exponential acceleration, blurring lines between therapy/enhancement. Think of the progress of AI, mainly due to the increase in computing power, the availability of enormous amounts of data (which constitute 'examples' for the machine, big data, forms the basis on which machines 'learn', machine learning) and the development of algorithms to identify hidden relationships between data (correlations and forecasts/predictions). AI includes all machines that imitate/simulate

* Contributo sottoposto a valutazione.

doi 10.53148/AG202203001

certain aspects of human intelligence, even bringing together interaction with neuroscientific discoveries, using information technologies, being able both to interact and learn from the external environment, and to reason and make decisions with increasing degrees of 'independence' with respect to specific instructions for that particular task. Robotics is moving towards the embodiment of AI, providing AI with an android mechanical body (which mimics the human form), creating an artificial body for the artificial mind. With the ambitious goal of designing machines that *imitate* humans (in body and mind) eventually to *replace* them.

The question arises: can machines act and think like humans? The technical question is if, by 'possible' we mean 'technically feasible'. But philosophy is not concerned with this. The philosophical question is anthropological: is a human being reducible to a machine? Granted that machines can become technologically similar to humans, is it good for 'machines to become human' or for 'humans to become machines'[1]?

2. *The artificial 'beyond' the natural: the techno-centric perspective*

Technophiles in the horizon of a techno-scientistic vision, in favour of any intervention to technologise the body and mind, outline horizons in which the artificial becomes increasingly similar to the natural and tends to 'merge', to intentionally cancel the difference between humans and machine, in a symbiosis between human being and technology, between organic and inorganic life. Technophiles push in the direction of the development of converging technologies and robotics/AI that replaces and surpasses the human being, the only way to overcome the biological limits of the body and the neurological-cognitive limits of the mind towards a trans-human, post-human or even 'super-human' perfection.

[1] M. Coeckelbergh, *AI Ethics*, Cambridge (Mass.), 2020; M.A. Boden, *AI. Its Nature and Future*, Oxford, 2016.

The elaboration of the concept of 'technological/electronic/artificial person', in the dual direction of a human who becomes artificial and the machine that becomes human anthropomorphically[2], philosophically presupposes a reductionist, functionalist and dualist conception.

A reductionist conception stems from the materialistic vision of human beings, whose body is reduced to an object or the sum of parts that can also be mechanically modified and replaced and the mind reduced to neural components that can be modified and replaced by computers.

The reductionist view is strictly related to the conception of functionalism which believes that the exercise or performance of functions is what counts to identify the person, regardless of nature, that is considered irrelevant. The nature of the human living organism is ontologically emptied. What matters are the functions expressed or exhibited: it is unimportant whether these belong to a human being or a machine.

There is a return to the anthropological body/mind dualism derived from Platonic-Cartesian thinking: according to this vision the person is split into body and mind, where being 'embodied' in a biological body and in a neurological mind, is considered a 'burden', which can and must be overcome by the machine, that is capable of exercising the kinetic, rational and cognitive function without bios, albeit artificially (robot as an artificial body, AI as an artificial mind). Technological alteration becomes an 'evolutionary duty'[3] that allows the 'technological shortcut' for the improvement the human being and the species as well as a 'technological imperative', whose omission can be considered fault and negligence towards technological evolutionism. A future horizon would open up techno-prophetically that would lead to a radical alteration of the nature of human beings, up to the total artificialisation of humans. The convergence of technologies would become hu-

[2] P. Benanti, *The cyborg. Corpo e corporeità nell'epoca del postumano*, Assisi, 2012.

[3] J. Harris, *Enhancing Evolution. The Ethical Case of Making Better People*, *Enhancing Evolution. The Ethical Case for Making Better People*, Princeton, 2007.

man/artificial convergence. The use of the conditional here is a must: these are futuristic and purely speculative scenarios.

Transhumanism promotes the abandonment of the organic and the transition to the virtual/artificial/digital in order to expand human capabilities, to have better lives and better minds. The motives of transhumanism[4] lie in the desirability of supra-human and hyper-human perfection, which is expressed in a moderate way[5] in the quest for increased beauty, physical endurance and life expectancy (aesthetic, physical and biological enhancement); in a radical way in the cancellation of the human condition itself, perceived and experienced as a limit. The goal is expressed in the Central Meme of Transhumanism: «it is ethical and desirable to employ technoscientific means to overcome the human condition (data)». In this perspective, the enhancement of human beings is achieved in the 'techno-human condition'[6] of enhanced post-human[7]. It is the theoretical horizon that will tend to empty flesh and blood human bodies, reducing them to mechanical and electronic biotechnological components, to mutant information flows capable of assisting and replacing the vital processes of the organism, with the promise of unlimited perfection. What counts is that the machine 'works' as perfectly as possible.

Perhaps these are «foolish fantasies» or «plausible prognoses», belated «eschatological needs» or unprecedented science fiction models[8], as J. Habermas argues. However, these

[4] World Transhumanist Association, *Transhumanism Declaration*, in *http//humanityplus.org/learn/philosophy/transhumanist-declaration*; N. Bostrom, *Superintelligence*, Oxford 2014.

[5] M.J. McNamee, S.D. Edwards, *Transhumanism. Medical Technology and Slippery Slope*, in *Journal of Medical Ethics*, 32, 2006, pp. 513-514.

[6] Cfr. FM-2030, *Are you a Transhuman?*, London, 1989; N. Bostrom, *Welcome to a World of Exponential Change*, in *Better Humans? The Politics of Human Enhancement and Life Extension*, ed. by P. Miller, J. Wilsdon, London, 2006, pp. 40-50; J. Savulescu, H. Maslen, *Moral Enhancement and Artificial Intelligence: Moral AI?*, in *Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*, ed. by J. Romporti, E. Zackova, J. Kelemen, Switzerland, 2015, pp. 79-96.

[7] R. Pepperell, *The Posthuman Condition Consciousness. Beyond the Brain*, Bristol-Portland (OR), 2003.

[8] J. Habermas, *The Future of Human Nature,* London, 2003.

are scenarios that are being prefigured at least speculatively (think of superintelligence and the singularity theory), with respect to which we cannot limit ourselves to simply taking note, a critical reflection that perceives these issues is indispensable. It is not a question of adhering to techno-phobic techno-catastrophism, but it is essential to develop a critical philosophical reflection on human life, on its meaning and value, which identifies the possible limits of technologies. It is not a question of exalting technology by despising humans or exalting humans by despising technology: the goal is to allow interventions on humans without distorting their identity, avoiding irreparable transformations.

Starting from the consideration that: technology is not a 'destiny', but we are the ones who build the technology. In this sense, we must not limit ourselves to taking note of what 'remains human' in technologies, but we must go further and consider also and above all what 'must remain human' in technologies.

3. *Humans in the age of AI: the human-centric perspective*

No matter how many advances have been made in the search for a machine that imitates and reproduces complex kinetic processes and cerebral organisation and all the characteristics of the human body and intelligence, it seems out of reach of realistically predictable technical developments. Certainly machines (software and computer programs) are now able to perform considerably complex functions and operations comparable – or in some contexts, even outperforming – those of humans: take for instance the collection, selection and cataloguing/archiving of information faster than is humanly possible, or the ability of calculation and performinag all languages[9]. More complex to accomplish and more costly,

[9] Italian Committee for Bioethics, Italian Committee for Biosafety, Biotechnology and Life Sciences, *Artificial Intelligence and Medicine: ethical aspects*, 2020.

both financially and in terms of energy consumption is the actual mobility of the robot[10] itself.

Techno-optimists consider it to be only a matter of time, but the project is clear: in the short term to replace every function to overcome physical limit, those of a fragile and mortal body, with robotic constructions; the limits of a mind that could be extended quantitatively and qualitatively in an indefinite and indefinable way to operate as and beyond the human in every context, sphere and situation.

But are there qualitative or essential differences between artificial intelligence and human intelligence? Will machines be able to fully replace humans? And, assuming it is technologically possible, would it be desirable? The question concerns the possibility and desirability of reproducing machines that can replace humans in *any sphere* of typically human thought and action, the possibility or desirability of *fully* reproducing human intelligence and of even surpassing it. That is, to reproduce the functions/operation and the same structure or constitution of human intelligence as a substrate.

Are there reasons to believe that it is 'better' on an anthropological, ontological and moral level to defend *the non-reproducibility, non-substitutability and uniqueness* of human intelligence? There are some aspects that make it difficult, in fact, to think that one day machines will be able to completely replace human intelligence.

There is a human dimension that does not pertain to machines*: self-awareness*. Thinking machines are not self-aware, they do not self-identify, being unable to detail the sum of the properties/functions and the series of acts/operations they perform to a unit that constitutes their synthesis. AI does not have an identity over time that persists with the modification of characters.

In the context of Western thought, philosophy has also recognised the relevance of *affection, emotions, feelings* in the field of human cognition. Even today from the experimen-

[10] Italian Committee for Bioethics, Italian Committee for Biosafety, Biotechnology and Life Sciences, *Development of Robotics and Roboethics*, 2017.

tal point of view it emerges that emotions influence the decision-making and rational processes of choices. Human choice does not derive exclusively from an entirely logical process that collects information, processes and calculates stored information in view of the decision that maximises benefits and minimises risks. As far as we know today, AI is incapable of reproducing emotional and affective aspects: and this already places a limit on the technological dream of strong artificial intelligence. Those designing 'thinking machines' prefer the dimension of 'calculation' moving in the direction of convenience in line with utilitarian thinking. However, an intelligent machine wanting to imitate and fully reproduce human intelligence should also include the general influence of emotive-emotional states on cognition and decisions. The recognition of the close connection between the cognitive and emotional dimension, within individual and inter-individual contexts, has highlighted the so-called «Descartes' error»[11], shedding light on the complexity of subjectivity.

An emerging problematic aspect of (strong) artificial intelligence concerns the distinction – that is clear in the philosophy of contemporary language – between syntax and *semantics*. Computers and software, which support AI, can operate on the syntactic links between symbols, whereas they do not consider the semantic dimension, that is, the interpretation of meanings. This would pose a further obstacle to the possibility of comprehensively creating an AI that can imitate *in toto* human intelligence. Semantics is not reducible to syntax and computing power no matter how advanced it may be in future, it will not be able to deal with complex semantic aspects on an interpretive level, possible for human intelligence. Human reasoning is based on concepts and meanings. If syntax is constituted by a series of rules of general composition applicable to large classes of linguistic elements without regard to meaning, semantics is not only the network of relationships between terms, but it is also *real experience*. Therefore,

[11] A.R. Damasio, *Descartes Error: Emotion, Reason and the Human Brain,* London, 2014.

on the basis of this, computing power and the syntactic computation of symbols would not be sufficient to create thinking machines, what would be needed is the capacity to interpret and perhaps also try out, to gain experience/feel, like a human being would (to suffer and rejoice, desire and fear, see, hear, touch, smell and taste). Therefore, at this point, AI should also be designed with an 'artificial life', thus, not only 'thinking machines' but also 'living machines'.

There is also a *motivational* dimension which, at this moment in time, is only human. The motivation to seek truth, to know reality, to invent, to imagine, to create. A thinking machine wanting to reproduce human intelligence should not only be able to perform intelligent operations to solve the tasks assigned by human beings, but it should also be able to self-assign tasks, to aspire towards the growth of knowledge. Today, sophisticated neural networks are able to perform intelligent operations (in cognitive and computational terms) outperforming human beings, with programs capable of self-modifying to improve their performance. But without motivation, which, to date, always comes from human input.

What the machine lacks and is uniquely human is the *possibility of perceiving the self as 'I' and being recognized by others as 'you'*. A human being is able to perceive relationality as 'being' (i.e., as an ontological and anthropological condition) and as 'duty of being' (i.e., as an ethical and normative condition). In this sense, the exceptionality of humans compared to other living beings and machines is justified: only human is recognised as having an intrinsic dignity as he/she is the only being that is able to recognise his relational duty, therefore is able to act morally (with respect to other beings, that are similar and dissimilar to her). In this sense, a human being, and not the machine, is properly autonomous, as he/she is able in principle (in the ontological sense) to give herself a moral norm, while the machine can at most be unpredictable, using algorithms.

The degree to which we identify human specificity, both on a phenomenological and ontological level, and the *irreplaceable dimensions* of humans, relates to how we will impose *lim-*

its on the technological possibility of building android robots or thinking machines: limiting technological dominance and the aspiration to '*remake*' nature, insofar as technology can also 'inhumanise' and 'dehumanise'[12]. The human person counts and must be preserved, even in the age of machines. There is *dignity intrinsecally in human nature*, independent of the functions a human being exercises and which has a value, that must be protected. It is not the functions, present or amplified by technological enhancement, that make it 'more worthy': improvement is also possible in the natural environment, by implementing intrinsic potential or possibilities, without technological shortcuts, but on the basis of virtue, personal, regular and active commitment in the direction of 'achievement' or of the maximum expression and 'flourishing' possible of one's own natural abilities.

4. *For technological humanism: ethical requirements for using AI and robotics*

The refounding of technological humanism means avoiding on one side 'anthropomorphising' the machine (humanising technology) or artificialising humans into a 'biological machine' (technologising humans). The 'machine' can be the ally of human beings and integration between natural and artificial can become *complementary*, always giving priority to respect for humans, safeguarding what 'counts' for the human person, within the framework of fundamental human rights.

Moreover, this is the evident direction of many ethical reflections on AI and robotics. AI is designed and produced by a human being (data is collected and selected by humans, algorithms are built by humans): we are the ones who have to decide how to produce robotics and AI. We know today what is human and what we want to preserve of what is human in

[12] *Readings in the Philosophy of Technology*, ed. by D.M. Kaplan, Lanham (MD), 2004.

a human-centred perspective (so-called human-centric or human-centered) in the construction and design of technology.

Many documents on the subject of the human-machine relationship prepared by commissions of experts at international, regional and national level, discussed in an interdisciplinary and pluralistic way, have drawn up minimum ethical requirements for regulating technoscience in the horizon of fundamental human rights (UNESCO, WHO, the European Group of Ethics in Science and New Technologies at the European Commission, the DH-BIO Committee for Bioethics of the Council of Europe, or ad-hoc groups on AI, National bioethics committees in Italy, Sweden, UK). In this perspective, ethics plays the role of critical reflection in understanding and evaluating AI and robotics, which justifies the regulatory requirements seeking on the one hand to open innovative technological opportunities 'for' humans and humanity and on the other to avoid or at least control and manage risks.

Even as humans build AI, they must maintain control and supervision over what they design, program, apply; machines should continue to be a 'support' for human decision making, to cognitively 'assist' human decisions, but not replace them.

Machines should not 'compete', but 'complete' human actions. In this sense, even language must be kept anthropocentric, considering AI machines as 'automatic' rather than 'autonomous' in learning. The term 'autonomy' cannot be applied to artifacts, even if these are complex or very advanced cognitive systems. The terminology of 'autonomous' systems is, however, widely used in scientific literature and public debate in reference to the highest degree of automation and the highest degree of independence from humans in terms of operations and decisions[13].

[13] The High-Level Expert Group on Artificial Intelligence, in *Ethics Guidelines for Trustworthy AI* (April 2019) emphasises the «"*human-centric*" dimension of the new technologies». European Group on Ethics in Science and New Technologies states «the importance that humans – and not computers and their algorithms – should ultimately remain in control, and thus be morally responsible»; «The principle of human dignity, understood as the recognition of the inherent human state of being worthy of respect, must

The need to maintain human control also remains essential to avoid the possible problem of 'technological delegation'. An expert system that is optimal in suggesting 'decisions' to humans poses the danger of reducing human attention with the possible consequence of reducing human skills (the so-called phenomenon of de-professionalisation), reducing responsibility by moving towards the artificialisation of choices that can impoverish and even cancel interpersonal relationships. In this sense, it is important to regulate productive synergy as complementarity between humans and machine, seeking ways of intelligent 'support' that allow a human being to have 'meaningful human control' in terms of attention, contribution, control and responsibility. The debate on reliability, explainability, traceability, transparency and algor-ethics are an important example of this[14].

A relational conception of human dignity that is characterised by our social relations, requires us to be aware of when and if we are interacting with a machine or another human being, and for us to reserve the right to assign certain tasks to humans or machines. In this ethical framework, AI ethics is the ethics of human beings: the machine cannot obscure action, which is human. Humans conceive, design, use AI, and humans should be kept at the centre. Given the technical possibility of creating artificial systems that can be confused with human beings, there should be a right to know the human or artificial nature of the interlocutor[15]. Ignorance or unclear

---

not be violated by 'autonomous' technologies» (p. 16). The European Commission, *White Paper On Artificial Intelligence - A European Approach to Excellence and Trust*, 19 February 2020, accepts that «the specific characteristics of many AI technologies, including […] partially autonomous behaviour, may make it hard to verify compliance with, and may hamper the effective enforcement of rules of, existing EU law meant to protect fundamental rights» (p. 12). Recommendation of the Council on OECD, *Legal Instruments Artificial Intelligence*, 2020.

[14] L. Palazzani, *AI and Health: Ethical Aspects for Regulation*, in *Teoria e Critica della Regolazione Sociale*, 1, 2021, pp. 1-16.

[15] Report of UNESCO, COMEST, *Robotics Ethics*, 2018: «Dignity is inherent to human beings, not to machines or robots. Therefore, robots and humans are not to be confused even if an android robot has the seductive ap-

understanding of the nature of the interlocutor could lead to misunderstandings, betray the expectation of empathic understanding and could affect human dignity (the problem of deception).

With any machine or technology, the design must be safe; safety is an ethical requirement for every machine/technology as it is for pharmaceuticals, food, transport, etc. This must also apply to AI and robotics. All the 'products' of AI and robotics should be compared, through studies conducted according to the rules of controlled clinical trials, with decisions made independently of the technology, together with computer experts and engineers. Controlled clinical trials remain the 'gold standard' for demonstrating the safety and efficacy of technologies. A new methodology is needed to control software applied to medicine, including the problem of the mechanism changing over time and validation that requires monitoring and further checks. Therefore, it will be necessary to demonstrate that AI is safe, to ensure that even unintentional harm can be minimised and prevented and to ensure technical robustness based on control beginning with the database (quality, accuracy, the interoperability of clinical data, both collected and compared), the algorithms that are applied, the advantage in terms of benefits and risks of their application. Only by doing so will it be possible to demonstrate the reliability of these systems through certification/validation/monitoring that guarantee their usability in clinical practice.

An additional ethical requirement is explainability. AI systems are also defined as 'opaque'[16]. 'Opacity' refers to the unexplainability or the limitation of explainability of the algorithms that interpret the data. In some circumstances it is impossible even for software programmers and computer scientists to explain how the system has achieved certain results

pearance of a human, or if a powerful cognitive robot has learning capacity that exceeds individual human cognition».

[16] L. Floridi, J. Cowls, *A Unified Framework of Five Principles for AI in Society*, in *Harvard Data Science*, (1) 1, 2019; S. Wachter, B. Mittelstadt, *A Right to Reasonable Inferences. Re-thinking Data Protection Law in the Age of Big Data and AI*, in *Columbia Business Law Review*, 494, 2019.

(the 'black box problem', in which only the inputs and outputs are known, but not what lies in between). It is virtually impossible for a human being (even an expert) to analyse the huge number of calculations carried out by the algorithm and determine exactly how the machine managed to decide. Automation can lead to 'opacity' or absence/lack of transparency on the pathways followed by the machine: the machine does not provide complete information on the correlations of the data and/or on the logic adopted to reach a conclusion or propose a decision. Non-transparent systems or non-intelligible systems make it more difficult to identify errors and therefore can also undermine the trustworthiness of AI. This involves the need to develop technology capable of explaining each step of the decision or at least to inform users of the risk of opacity, in order to acquire critical awareness. Human beings must be aware that they are interacting with an AI system as well as the capabilities and limitations of the system itself (informed consent)[17].

To have trustworthy and secure AI, it is necessary to have 'data quality' (in addition to quantity with 'big data'), specifically, accuracy in data collection and 'quality algorithms' that must be inclusive (with respect to age, gender, ethnicity) and not discriminate. The protection of privacy and confidentiality is often highlighted as an obstacle to the development of AI, which is based on big data. In an era of mass and massive collection of data through digital communication technologies, the right to the protection of personal information and the right to respect for privacy are put to the test. In this sense, technologies are becoming increasingly 'opaque' and the users increasingly 'transparent'. In the era of AI and the need to use data for medical research, the possibility/opportunity of 'sharing' data arises, as a 'social/common good' for the advancement of scientific knowledge. However, this calls for a specific

[17] C. Rudin, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,* in *Nature Machine Intelligence*, 1, 2019, p. 206.

regulation, with a view to the protection of scientific progress and at the same time protecting patients.

5. *Educating to acquire critical awareness (AI-robotic literacy)*

It is extremely important to introduce specific education on AI and robotics, but also on ethics in this sector of emerging technologies in order to avoid the so-called 'polarisation of skill' of employees, that is 'retraining' workers, in the face of the developments in emerging technologies. This raises concerns addressed by the European Group on Ethics in Science and New Technologies (EGE), in its opinion *Future of Work, Future of Society* (2018), which emphasises the 'polarisation of skills' that can hide new forms of discrimination, excluding those who fail to secure the new 'skills' required. The problem of new professions, even in the medical field, therefore, remains that high-level skills will be required. It is necessary to redesign medical education programs, dedicating a significant part of the training of future doctors to the problems deriving from the virtualisation, digitisation and artificialisation of medicine which are the basis of AI technologies. Interdisciplinary and transdisciplinary courses are needed in training health professionals to constantly adapt to technological change and the possible 'convergence' of traditionally separate disciplinary sectors (e.g., medicine and computer science or physics or data science).

The 'new education' should ensures adequate skills in the era of new technologies, but the danger is for those that get left 'behind'. However, an adequate educational project that is truly inclusive must not only aim at improving the skills of those with already high skill levels, but it should also accompany people during the process of the technological transition of work, in a process of progressive and gradual adaptation of personal skills, realising how to 'govern' the changes in the human-machine relationship. The educational problem, on the ethical level, should not only be to achieve, in the fastest and most efficient way, acquisition of the technological skills

required in the era of digitisation and automation, but it also involves educating on the best way to create the conditions to inform and train on how to develop 'personal skills', that are integrated and complementary to technological skills (such as creativity, interpersonal relationships, social participation).

Therefore, education needs to be rethought to strengthen the training of those who work, in a continuous way, constantly adapting to technological transformations, and the training of those who enter employment, paying specific attention to guaranteeing conditions for 'decent' jobs which can reconcile digital skills with the improvement of personal skills. It is also important to educate towards the risks of over-reliance on technologies which can even lead to a weakening of skills or de-skilling; the transparent use of new technologies, without opacity and discrimination; having the goal of inclusiveness and justice, reliability and responsibility.

Education should also convey a capacity for 'critical thinking' beyond digital skills, such as the ability to interact with technologies. Many points regarding vulnerable groups are highlighted in the European Parliament document *Digital Skills in the EU Labour Market* (2017). The question of how society and citizens, in particular vulnerable groups such as those with disabilities or the unemployed, can be trained in new technologies is a fundamental ethical question. The technologies themselves must be used to facilitate education and the motivation to learn, even for people with disabilities, or the elderly, improving the possibility of innovative and personalised teaching, that is flexible and adaptable to the needs of individuals.

Training also needs to be renewed, introducing ethics and bioethics courses for engineers, computer scientists, IT and data scientists, with particular reference to ethics in technologies (ethics by design/in design/for designers) and to the design, methodology and application of technologies. This is the only way to ensure ethical awareness and understanding from the very beginning of technology design. The AI designer and programmer, particularly in 'machine' and 'deep learning', could benefit from interdisciplinary training on the ethical,

social and legal aspects of their activity. Designing AI is not just a technical activity, but it involves, intentionally or unintentionally, ethical and legal concepts.

It is also desirable to promote public debate on the developments and limitations of AI in medicine, so that all individuals – present and future patients – can acquire the basics of 'AI literacy', promoting active participation in social discussion. These are the prerequisites for a possible overcoming of the 'digital divide' in medicine, avoiding the marginalisation, stigmatisation and exclusion of people lacking the technologies and skills and the motivation to use them, within the framework of inclusiveness, in order not to leave anyone behind. The right to free AI education should be provided as part of compulsory education, both for young people (from primary school upwards) and for adults (in universities and vocational education). There is a need for a common understanding of AI, its pros and cons, including the acquisition/development/critical awareness of its ethical issues, in order to bridge the AI gap and ensure equal access to opportunities and inclusive growth.

Laura Palazzani, Technological humanism: a philosophical framework for education

The article focuses attention on the disruptive transformations of the new technological wave, with specific reference to AI and robotics, as an embodiment of AI. The central question is not technical, but philosophical by seeking, within the horizon of reason, to define the boundaries of the person and human nature with respect to the machine. The response from philosophy of the person justifies a human-centred horizon vs. the techno-centric horizon, recognizing the centrality of the dignity of the human person irreducible to a machine or function. The article identifies some specific implication for education.

Key words: new technological wave, AI, robotics, ethics, trans-humanism.

Laura Palazzani, Umanesimo tecnologico: premesse filosofiche per l'educazione

L'articolo focalizza l'attenzione sulle trasformazioni dirompenti della nuova ondata tecnologica, con specifico riferimento all'IA e alla robotica, come incarnazione dell'IA. La questione centrale non è tecnica, ma filosofica, cercando, nell'orizzonte della ragione, di definire i confini della persona e della natura umana rispetto alla macchina. La risposta della filosofia della persona giustifica un orizzonte che pone al centro l'uomo (umanocentrico) contro l'orizzonte tecnocentrico, riconoscendo la centralità della dignità della persona umana irriducibile a una macchina o a una funzione. L'articolo individua alcune implicazioni specifiche per l'educazione.

Parole chiave: nuova ondata tecnologica, intelligenza artificiale, robotica, etica, transumanesimo

NOTE SUI COLLABORATORI DEL FASCICOLO 3 2022

Giuseppe de Vergottini, Professore emerito di Diritto costituzionale, *Alma Mater Studiorum* – Università di Bologna

Paolo Biavati, Professore ordinario di Diritto processuale civile, *Alma Mater Studiorum* – Università di Bologna

Jorge Otaduy Guerín, Profesor ordinario de Derecho eclesiástico del Estado, Universidad de Navarra

Laura Palazzani, Professoressa ordinaria di Filosofia del diritto, Libera Università Maria Ss. Assunta (Lumsa) di Roma e componente del Comitato internazionale di Bioetica dell'Unesco

Michele Sesta, Professore ordinario di Diritto civile, *Alma Mater Studiorum* – Università di Bologna

Valerio Gigliotti, Professore associato di Storia del diritto medievale e moderno, Università di Torino

Elena López Barba, Profesora titular de Derecho civil, Universidad de Huelva

Stefano Barbati, Ricercatore di Diritto romano e diritti dell'antichità, Università di Torino

Edmund J. Mazza, Full Professor of History, Azusa Pacific University

Claudio Gentile, Dottore in Diritto canonico, Pontificia Università Urbaniana

Giovanni Parise, Dottore in Diritto canonico, Pontificia Università della Santa Croce

Ilaria Samorè, Dottoranda di ricerca in Diritto ecclesiastico e Diritto canonico, *Alma Mater Studiorum* – Università di Bologna

Archivio giuridico Filippo Serafini

INDICE DEL FASCICOLO 3 2022

Miscellanea

***Giuseppe Dalla Torre, Scritti su* Avvenire. *La laicità serena di un cattolico gentile.* Presentazione del volume a cura di Geraldina Boni, Edizioni Studium, Roma, 2021, pp. 1-568 (Bologna, 29 marzo 2022)**

ARCHIVIO GIURIDICO *Filippo Serafini*

Periodico Fondato nel 1868
Pubblicazione trimestrale

Caratteristica dell'*Archivio giuridico* è stata, sin dall'inizio, quella di essere visto in Italia e all'estero, come un autorevole e qualificato punto di riferimento sui progressi della dottrina giuridica italiana in una visione che, pur non rifuggendo dalla specializzazione in sé, ne evita peraltro ogni eccesso.

I Collaboratori sono pregati di inviare i loro contributi via e-mail (scritti in formato.doc). Ogni lavoro dovrà essere corredato di: Nome, Cognome, Qualifica accademica, Indirizzo postale, Indirizzo e-mail, Numero di telefono (è gradito anche un numero di cellulare). Ogni articolo dovrà essere corredato di un titolo in lingua inglese e un riassunto in lingua italiana e inglese di non più di 200 parole specificando: scopo, metodologia, risultati e conclusioni; e di almeno tre parole chiave in lingua italiana e inglese. Gli articoli, salvo casi eccezionali non potranno superare le 32 pagine (intendendosi già impaginate nel formato della rivista, ovvero circa 16 cartelle in formato A4 corrispondenti a 88.000 battute spazi e note inclusi). Le opinioni esposte negli articoli impegnano solo i rispettivi Autori.

La Rivista adotta la procedura di revisione *double-blind peer review*.
I contributi pubblicati sono indicizzati nelle seguenti banche dati nazionali ed internazionali: Articoli italiani di periodici accademici (AIDA); Catalogo italiano dei Periodici (ACNP); DoGi Dottrina Giuridica; ESSPER Associazione periodici itailani di economia, scienze social e storia; Google Scholar; IBZ online International bibliography of periodical literature in the humanities and social sciences; SCOPUS.

La casa editrice fornirà, ai rispettivi Autori, estratto degli articoli in formato pdf. Possono altresì essere forniti fascicoli cartacei degli 'estratti', a pagamento. Chi fosse interessato è pregato di richiedere preventivo di spesa a: **info@mucchieditore.it.**

Recensioni e segnalazioni bibliografiche: gli Autori ed Editori di pubblicazioni giuridiche sono pregati di mandare un esemplare di ogni volume alla Redazione dell'*Archivio giuridico Filippo Serafini*. Sarà gradito un foglio di accompagnamento con i dati bibliografici, classificazione, sommario, etc. La Direzione della Rivista si riserva di recensire le opere che, a suo insindacabile giudizio, risulteranno di maggior interesse.