Dirk Messner

Redefining and Renewing Humanism in the Digital Age

Preserving Cornerstones of the Enlightenment
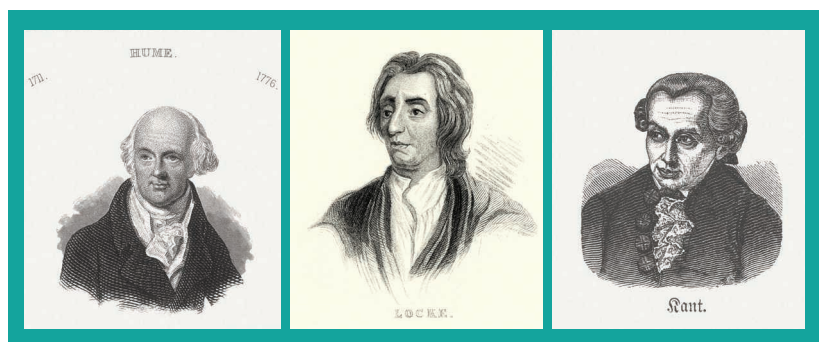

We are at a crossroads. The convergence of Artificial Intelligence (AI), big data analytics, and machine learning has rapidly produced some remarkable technologies: voice, face, and odor recognition technologies for precise personal identification; statistical algorithms predicting human behavior and preferences; camera and sensor networks tracking, monitoring, and analyzing human activities; and bots of sufficient sophistication to masquerade as human actors in 280 characters.

However, when all these digital technologies meet authoritarian governments or powerful private actors (creating business models based on personalized data, not embedded in effective public governance structures for data protection), privacy, human rights, democracy, freedom, and dignity are at stake. Repressive regimes no longer need to rely on the blunt tools of repression, oppression, and suppression [1], [2]. With unrestricted access to and control of information technology and communication infrastructures it is possible to use these digital instruments to manipulate and shape data collection, information flows,

public debate, and public opinion. Now regimes can build surveillance societies [3], manipulate electoral districts and democratic elections, undermine the independent press and academic freedom, and control and destroy democratic movements and the right to protest.

Bearing in mind that in 2020 more than 50% of the global population is living in non-democratic states, and keeping in mind the dis-

turbing trend to authoritarianism of populist leaders in supposedly democratic countries, it is easy to develop dystopian scenarios about the destructive potentials of digitalization and AI for the future of freedom, privacy, and human rights. But here it will be argued: this is only half the story, the dark side. The light side is that AI and digital innovations could also be enablers of a Renewed Humanism in the Digital Age.

Digitalization Challenging Democracy

Let us first zoom into some potential threats. The combination of authoritarianism and the technological reservoir of digital tools is frightening. Armin Grunwald (4, pp. 168–175) shows how the characteristics of these technologies could challenge the fundamentals of democracy. He detects five main risks.

The first is that an exponentially growing pool of data, which is a new source of power, is currently controlled by a small group of globally engaged private firms and remains largely unregulated. One important characteristic of the power of these data driven firms is that their tax payments are well below the average. How could and should states and citizens control these new digital superpowers?

Second, algorithms and deep learning systems are very good at discovering patterns based on big data analysis; with growing complexity of data and deep learning systems it becomes more and more difficult for humans (experts included) to understand how algorithms actually produce their sometimes astonishing results. "Believing" (a fundamental characteristic of the Middle Ages) technical systems would substitute for "knowing" or "understanding" causal relationships (fundamental characteristics of the enlightenment). If we were to transfer more and more decisions from policymakers, judiciaries, administrations, and parliaments to deep learning systems, without solving the problem of *lack of transparency*, technological progress would result in societal regression.

Third, digitalization (and autonomous technical systems) promises to drive the acceleration of decision-making processes. While this may be beneficial for the optimization of production processes, it could undermine democratic and deliberative processes and ultimately be destructive for societies. Democracy is about the participation of many people, finding compromises between different interests and values, and creating public dialogues — democracy needs time to balance the complex tensions in diverse communities and societies. Digitalization is about automation, acceleration, and optimization. Aligning societies and democracies to the acceleration dynamics of digitalization could undermine their power and responsiveness.

Fourthly, democracy is cumbersome. Algorithms seem to be neutral, quick, objective. But could we imagine election campaigns between (social democratic, conservative, liberal, green) algorithms, representing the different values systems within societies; and *Algorithms* mainly driving or dominating decision making processes in administrations? This would be a technocratic, expert-based utopia (or dystopia), substituting or undermining the complex decision-making processes and deliberations of democracies and open societies.

Fifth, the collection of multiple personalized data makes it possible to detect behavioral patterns, lifestyles, politics, or sexual attitudes of individuals. One example is that of the different types of social scoring systems (of private firms or public institutions) that are emerging around the globe. Large scale societal experiments based on social scoring technologies are already being used by authoritarian regimes, but also by so called non-liberal democracies or even liberal democracies. From psychology we know that people, being "observed" (now by social scoring systems) feel controlled and start behaving differently (5). These trends could undermine individual freedom, privacy, and public engagement.

Grundwald shows that digitalization does not automatically translate into strengthening democratic institutions. Governance is needed to align digitalization with democracy.

Four Layers to Think about a Renewed Humanism in the Digital Age

Against this background of dystopian scenarios or risks we will develop a four-step approach to describe some cornerstones for a forward looking, optimistic future of humanism in the Digital Age:

- Defending the achievements of the Enlightenment in the Digital Age;
- Fixing the weaknesses of Enlightenment with forward looking sustainability concepts;
- Developing normative guardrails to handle the challenges of "technologization of humans" and the "humanization of machines"; and
- Mobilizing the potentials of digital innovations for a Renewed Humanism in the Digital Age.

Defending the Achievements of the Era of Enlightenment in the Digital Age

Our starting point is to defend the essential achievements of the Enlightenment against the described threats of digital disruptions to democracy, privacy, freedom, and dignity. How can we align the normative standards of Enlightenment with digital innovations and dynamics? What would an institutional setting look like, that embeds technological change into open and democratic societies. Let us remember, what we should defend (6, p. 294): The project of Enlightenment is well summarized in Immanuel Kant's "emergence of man from his self-imposed immaturity" (7). Absolutism and God's grace were

replaced step by step by human reasoning as the new authority of judgement. The Enlightenment consisted of currents of rationalism (Descartes), empiricism (Hume, Locke), encyclopaedism (Diderot), and universalism (Kant), and created basic foundations for the idea of human rights, science, rule of law, and democracy. The self-image of humans changed in the transition towards the Era of Enlightenment, as they were then understood to be rational (Kant), autonomous (Descartes, Locke), and communicative (Habermas) subjects with natural rights (Kant, Rousseau). These essential ingredients of the Enlightenment should be preserved in the epoch of digital change to avoid the re-emergence of a "self-imposed immaturity," now not driven by Absolutism and religious beliefs, but by technological systems with the potential to undermine the basic structures of privacy, freedom, democracy, and autonomy. We must consider the following questions (6):

- How can human "maturity" and democratic voting processes be preserved in the context of automated decision making-processes?
- How can processes of deliberation in societies be protected in view of massive increases in knowledge, opinions, fake news?
- How can freedom, equality, and privacy be protected for all people in digital and virtual spaces?
- How can the sovereignty of people and rule of law be protected against digital surveillance?
- How can the basic values of the Enlightenment be built into algorithms and ML systems?
- How could AI experts and producers of algorithms be trained to consider the normative pillars of democratic societies? Could a functional equivalent to the Hip-

pocratic Oath for medical professionals make sense for the AI community?

These questions are even more important, as democracy is currently not only under pressure from digital disruptions, challenging some fundamentals of democracy, and from some digital protagonists, questioning the eloquence of democratic procedures vis a vis AI driven technical systems (8). In parallel democracy is also undermined by right wing, authoritarian, populist movements, and governments — often using digital tools for their purposes. Putting democracy, privacy, and human dignity first, and thinking about digital innovations as instruments, supporting the basic normative values of open, democratic societies, is therefore of high importance in these turbulent times.

The Blind Spots of Enlightenment – Being Cured by Sustainability Thinking

A further developed humanism for the 21st century would give us the chance to examine the blind spots of the Enlightenment Era. Four factors are of relevance in this regard (6).

Firstly, the protagonists of the Enlightenment focused with good arguments on "reasoning," instead of "irrationality" as a basic foundation of societies. In the 21st century we should combine the power of reasoning and cognitive capabilities with the power of social intelligence of humans (9)–(11).

Secondly, the protagonists of the Enlightenment focused with good reasons on the rights of individuals against the background of absolutism. In the 21st Century, we cannot ignore that a culture of individualism needs to be embedded into a culture of the common good (12), (13), recognizing that the freedom and creativity of individuals depends on

societal structures and social networking. Humans are social beings — we become individuals, because we interact with others (11), (14). Societies are more than only the sum of many individuals (15)–(17).

Thirdly, protagonists of the Enlightenment argued, with good reasons, that humans needed to emancipate themselves from nature, in an era in which the dynamics of nature (droughts, extreme weather events, fire) threatened (the survival of) human communities. In the 21st century, the era of the Anthropocene (18), humans need to understand that human existence is coupled with the complex systems of the biosphere. Meeting the needs of 10 billion people in 2050, while protecting the globally interconnected ecosystems within the planetary boundaries (19), in order to avoid dangerous tipping points in the Earth System (20), is one of the most urgent challenges for human societies.

Finally, protagonists of the Enlightenment invented the idea of universalism and equal rights (human rights) for all. This was a fundamental "moral revolution" (21) in an era still being shaped by colonialism and slavery. In the 21st century, the epoch of global interconnectivity, we are learning that universalism needs to be embedded in vital local cultures and institutions. A locally rooted universalism or cosmopolitanism (22) might be a concept for a globalized world, being composed of highly interconnected local spaces — increasingly intertwined with multiple virtual spaces.

It is interesting to observe that these blind spots from the Enlightenment Era are key elements of the sustainability discourses of the last 40 years: balancing rationality and the power of social intelligence, combining individual rights and the idea of commons, strengthening

universalism and local identities alike, developing economies, compatible with planetary boundaries. Sustainability concepts and narratives (like the 2030 Agenda) build on the achievements of Enlightenment — and develop those further. Therefore the basic achievements of Enlightenment and of the epoch of Sustainability should be defended in the Digital Age. The big question is: how can digitalization be aligned with sustainability transformations [6]?

At a Crossroads – Redefine the Very Nature of Humans in the Digital Age

The Enlightenment era focused on the nature of humans (rational, communicative, autonomous), as subjects with natural and equal rights, and distinct rights vis a vis states. Digital innovations, and the revolutionary entanglement between humans and technical systems, combined with synthetic biology, could enable a transformation of our understanding of human nature, and a transformation of human nature itself. Two main discourses and trends (e.g., in research) are of importance in this regard. Firstly, there is a debate concerning the "technologization of human beings" through digital technologies (combined with synthetic biology) to overcome biological limitations. "Human enhancement" [23]–[25] moves far beyond fighting diseases, aiming at "developing further," transforming humans, and even creating new, hybrid species through artificial evolution [24]. There are also research strands and debates about the "humanization of machines," the creation of human-like, sentient, and or even independent artificial species [26], [27].

Both trends would alter societies enormously and mark a crossroad in human civilization. After humans became the major drivers of Earth system changes during the 20th century (the debate of the Anthropocene), humans are now becoming capable of fundamentally transforming themselves as a species. They are creating machines and technical systems with cognitive (and potentially even emotional) capabilities similar to their own. There is an urgent need to define and describe our understanding of humans and human self-image in the Digital Age. How does the "Eigenart" of humans look in these new contexts? What normative guardrails for human transformations would we like to see emerge? This ethical discussion should not only be decided by scientists, experts, elites. and firms, it must be driven by broad-based societal processes.

The Utopian Potentials of Digitalization, Enabling a Renewed Humanism

So far, digital disruptions have mainly been described as threats, challenges and risks to the future of humanisms and democracy – that need to be "controlled," "governed," "embedded in institutional and normative settings," and "aligned with open societies". But the digital transformations themselves could also be seen as radical enablers towards a next step of human civilization, moving beyond the horizon of the Era of Enlightenment. How could we leverage digital dynamics of change to "invent" a Renewed Humanism for the 21st century [6]? This question is explored here in four main points.

Firstly, digitalization and AI are creating an explosion of knowledge in research, in knowledge driven firms, and in society. The main question is, whether and how educational systems are being prepared to empower people to translate this explosion of knowledge into individ-ual and societal progress. The invention of the printing press in the 15th century and the following emergence of modern science and educational systems during the last 200 years translated into fundamental transformations of human civilization. None of the economic, social, and political successes of the last two centuries would have happened without this first great knowledge transformation in human history. AI and deep learning machines, multiplied by quantum computing, will result in a second large scale knowledge transformation in human civilization. This knowledge revolution will create a much larger reservoir of solutions and opportunities for societies (and the risks described above) than ever before in human history. Empowering humans, by investing in educational and other knowledge systems will be key. Important to also consider is that the expanding knowledge pools as such does not automatically result in greater human well-being and societal progress. They can be misused to exploit people, damage the planet, and result in unequal development, making the few wealthy and powerful. It is relevant to consider how and for what purpose knowledge is being used. Equal access to excellent educational systems, combined with linking knowledge explosions to societal normative frameworks (like a Renewed Humanism or the 2030 Agenda) is therefore key [28].

Second, digitally-based, transnational communication infrastructures are creating the technical preconditions for a global culture of cooperation [29]–[31]. Next steps in building virtual spaces for people across borders and beyond geography to communicate, meet, connect, and work are opening up the chance for real time interactions between 10 billion people on

Earth in 2050. Digitally enabled communication infrastructures, virtual spaces, and the opportunities of hologram techniques drive and multiply — at a transnational scale — interactions between people, joint learning processes, joint routines, and connectedness. They trigger and shape common, cross-border mindshifts, heuristics, frameworks, and narratives. All of this could enable the emergence of a culture of global cooperation — rooted in a web of deep transnational interconnectedness between humans, which has never before existed in human history. During the last few centuries, horses, trains, cars, airplanes, as well as telephone infrastructures, shrunk distances and redefined the importance of geography for human interaction. These innovations made the emergence and stabilization of larger scale human communities (such as nations, the European Union, global value chains) possible. During future years and decades, the Internet, multiple global and cheap communication structures, and the rising importance of virtual realities and spaces (which are adding a fourth dimension to our three-dimensional world that humans learned to live with during the last 250 000 years) will exponentially drive dense transnational network building and deep connectivity across national borders. Immanuel Kant, living and working in Königsberg all his life (1724–1804), "invented" the idea of a "world population" and "world citizenship." The digital drivers of transnational human networks could "create" an interconnected, interactive world community, potentially being able to interact densely in multiple ways — regardless of physical borders.

Third, this new generation of technical systems and infrastructure will shape our world views and

transform our human self-image profoundly. It began with the first photos of the fragile, blue planet, made by Apollo Missions in the 1960s, which changed humans' perspectives of the Earth system and of humans as a species. Digital innovations and tools now allow humans to monitor, track, visualize, analyze the planet and the universe, local and global ecosystems, societies and individuals, the metabolism of local, national, and global economic activities, local and transnational infrastructures, and local and global knowledge flows, much more comprehensively than ever before. Satellites can create pictures of the planet every day in real time, observing and analyzing Arctic ice sheets, refugee flows, the Amazon forest, transnational infrastructure, and shrinking and growing human settlements. Humans can now "look" at the whole planet, observe it, zoom into local spaces, and discuss all of this with others in virtual spaces, regardless their physical location. The new opportunities of looking at (macro structures) of the planet as a whole and zoom into its pieces, is being complemented by digital instruments which open up the doors to the micro and nano world. Digital cameras and robots can visit our human bodies and make visible atomic structures. In addition, individualized and personalized AI systems might support us in understanding ourselves better, in day to day decision making, and in life-long learning [24]. Telephones, trains, and cars created the technical infrastructures to build modern nation states, related to the emergence of "national consciousness" and the idea of an international system. "World consciousness," "Earth system consciousness," "humans as one species consciousness," "humans, interwoven with the planetary bio-sphere conscious-

ness" [6] might now emerge — and help humans and societies to cope with the challenges of the Anthropocene in the digital age.

Fourth, it is uncertain whether and when something like "universal AI" will succeed [23], [25]. Nevertheless, it is already clear that AI will surpass the cognitive capabilities of humans in an increasing number of areas. This might open up opportunities for renewing the human self-image. The emerging "competition" between humans and technical systems in the field of cognitive and analytical skills could help us to focus (more than in the past) on the human characteristics (human "Eigenart") which clearly distinguish us from machines. Social intelligence and emotional capabilities — which drive attitudes like empathy, altruism, solidarity, love — (neglected in the Era of Enlightenment with its focus on reasoning, rationality, and cognition; see above) have likely been as important for the process of human civilization as the cognitive eloquence of our species [32], [34]. Or as the WBGU [6, p. 6] puts it: "AI would possibly grant us a certain amount of emancipation from the latter (cognitive tasks) and allow us to focus more on skills such as empathy, care and solidarity. In contrast to the "hard clichés" of the superhuman with the computer brain in a world of steel, this would delineate a "soft" vision of societal progress."

Renewed Humanism

We are at a crossroads. On one hand, AI and digital technologies, ungoverned, threaten to undermine privacy, democracy, freedom, human rights, and dignity. On the other, digital transformations offer the potential for enhanced self-knowledge, revolutionary machine-human interactions, and global interconnectedness. As we move forward, the need to establish the basis for a Renewed

Humanism in the digital age is becoming more pressing. To accomplish this we must first preserve the cornerstones of humanism that emerged from the Enlightenment Era.Second, we must consider and address its blind spots along with modern elements of current sustainability approaches (i.e., the 2030 Agenda). Third, we must establish normative guardrails for a newly emerging digital society in which the "technologization of humans" (human enhancement) and the "humanization of machines" become possible — (re)defining the human self-image and building consensus about "the future we want."

Finally, we must mobilize the digital potentials for a civilizational shift towards a Renewed Humanism. This entails channeling the ever-expanding knowledge explosion through open and inclusive education systems, and directing transnational communication and virtual systems to support an emerging global culture of cooperation. It requires supporting an emerging "world and Earth system consciousness" driven by multiple digital tools and systems that allow humans to manage the Earth system collectively, and directing human learning processes — confronted with "intelligent" machines — towards a fundamental valorization of social intelligence and emotional capabilities as key drivers of the next phase of human civilization.

Acknowledgment

Author Information

Dirk Messner is President of the German Environment Agency, Prof. for Political Science at University Duisburg-Essen/Germany, and past (2019) Co-Chair of the German Advisory Council on Global Change (WBGU). Email: Dirk.Messner@uba.de.

References

[1] A. Wolin, *Democracy Incorporated: Managed Democracy and the Specter of Inverted Totalitarianism.* Princeton, NJ: Princeton Univ. Press, 2008.

[2] S. Feldstein, "How artificial intelligence is reshaping repression," *J. Democracy*, vol. 30, no. 1, pp. 40-52, 2019.

[3] S. Zuboff, *The Age of Surveillance Capitalism*, New York, NY: Public Affairs, 2018.

[4] A. Grunwald, *Der unterlegene Mensch. Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern.* München, Germany: Riva-Verlag, 2019.

[5] M. Manner and J. Gowdy, "The evolution of social and moral behavior," *Ecological Economics*, vol. 69, no. 4, p. 753–761, 2010.

[6] WBGU (German Advisory Council on Global Change), *Towards our Common Digital Future.* Berlin, German: WBGU, 2019.

[7] I. Kant, *An Answer to the Question: 'What is Enlightenment?'.* Berlinische Monatsschrift, 1784.

[8] C. Botero Marino, M.J. Camilleri, and C. Cortes, "Democracy in the digital age," *Rule of Law Rep.*, Nov. 2017.

[9] J. Jacquet, C. Hauert, A. Traulsen, and M. Milinski, "Shame and honour drive cooperation," *Biology Letts.*, vol. 7, no. 6, pp. 899–901, 2011.

[10] E. Fehr and S. Gächter, "Altruistic punishment in humans," *Nature*, vol. 415, pp. 137-140, 2002.

[11] A. Sen, "Rational fools: A critique of the behavioural foundations of economic theory," *Philosophy and Public Affairs*, vol. 6, no. 4, pp. 317–344, 1977.

[12] E. Ostrom, *Governing the Commons. The Evolution of Institutions for Collective Action.* Cambridge, New York: Cambridge Univ. Press,1990.

[13] D. Messner and I. Scholz, "Globale Gemeinwohlorientierung als Fluchtpunkt internationaler Kooperation für nachhaltige Entwicklung," *Ein Perspektivwechsel Zeitschrift für Außen- und Sicherheitspolitik*, vol. 4, pp. 561–572, 2019.

[14] M. Tomasello, *Origins of Human Communication*, Cambridge, MA: M.I.T. Press, 2010.

[15] J. Habermas, *Auch eine Geschichte der Philosophie*, Frankfurt, Germany: Suhrkamp, 2019.

[16] M. Tomasello and H. Rakoczy, "What makes human cognition unique? From individual to shared to collective intentionality," *Mind and Language*, vol. 18, no. 2, pp. 121–147, 2003.

[17] D. Messner, *The Network Society.* London, U.K.: Frank Cass, 1997.

[18] W. Steffen, J. Grinevald, P. Crutzen, and J. McNeill, "The Anthropocene. Conceptual and historical perspectives," *Philosophical Trans. Royal Society*, vol. 369, pp. 842–867, 2011, 1938.

[19] J. Rockstroem *et al.*, "A safe space for humanity," *Nature*, vol. 461, pp. 472–475, 2009.

[20] T. Lenton *et al.*, "Tipping elements in the Earth's climate system," *PNAS*, vol. 105, no. 6, pp. 1786–1793, 2009.

[21] K.A. Appiah, *The Honor Code: How Moral Revolutions Happen*, New York, NY: Norton, 2011.

[22] J. Brasset and D. Bulley, "Ethics in world politics. Cosmopolitanism and beyond," *Int. Politics*, vol. 44, pp. 1–18, 2007.

[23] N. Bostrom, *Superintelligence. Paths, Dangers, Strategies.* Oxford, U.K.: Oxford Univ. Press, 2014.

[24] P. Domingos, *The Master Algorithm.* New York, NY: Basic, 2015.

[25] M. Tegmark, *Life 3.0 – Being Human in the Age of Artificial Intelligence.* New York, NY: Knopf, 2017.

[26] Y.N. Harari, *Homo Deus*. London, U.K.: Harvill Secker, 2015.

[27] Y.N. Harari, "Reboot for the AI revolution," *Nature*, vol. 550, pp. 324–327, 2017.

[28] J. Renn, *The Evolution of Knowledge, Rethinking Science in the Anthropocene.* Princeton, NJ: Princeton Univ. Press, 2020.

[29] D. Messner and S. Weinlich, "The evolution of human cooperation. Lessons learnt for the future of global governance," in *Global Cooperation and the Human Factor in International Relations,* D. Messner and S. Weinlich, Eds. London, New York: Routledge, 2016, pp. 3–64.

[30] M. Tomasello, *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard Univ. Press, 1999.

[31] R. Stichweh, *Die Weltgesellschaft: Soziologische Analysen.* Frankfurt, Germany: Suhrkamp, 2000.

[32] M. Tomasello, *Why We Cooperate.* Cambridge. MA: M.I.T. Press, 2009.

[33] M. Nowak, "Five rules for the evolution of cooperation," Science, vol. 314, no. 5805, pp. 1560–1563, 2006.

[34] D. Messner, A. Guarìn, and D. Haun, "The behavioral dimension of international cooperation," in *Global Cooperation and the Human Factor in International Relations*, D. Messner and S. Weinlich, Eds. London, New York: Routledge, 2016, pp. 47–65.

TS