Digital Humanism: Navigating the Tensions Ahead

Helga Nowotny

Abstract The assumption of digital humanism that a human-centered approach is possible in the design, use, and further development of AI entails an alignment with human values. If the more ambitious goal of building a good digital society along the co-evolutionary path between humans and the digital machines invented by them is to be reached, inherent tensions need to be confronted. Some of them are the result of already existing inequalities and divergent economic, social, and political interests, exacerbated by the impact of digital technologies. Others arise from the question what makes us human and how our interaction with digital machines changes our identity and relations to each other. If digital humanism is to succeed, a widely shared set of practices and attitudes is needed that sensitize us to the diversity of social contexts in which digital technologies are deployed and how to deal with complex, non-linear systems.

The availability of masses of data, efficient algorithms, and unprecedented computational power has pushed humans on a co-evolutionary path with the digital machines we have created. Seen from an evolutionary perspective, this might look like another of the many evolutionary trials and errors whose outcome leads either toward a dead end or toward new forms of life. Although this is impossible to predict, we should remind ourselves that cultural evolution, spearheaded by science and technology, has overtaken biological evolution. It has equipped the human species with cognitive capabilities that have enabled it to generate the digital entities, devices, and infrastructures with which humans interact in ever-more intricate and intimate ways. We should know them better than they know us—yet, we are repeatedly plagued by the anxiety that in the end they might dominate us.

Thus, we oscillate between trust in the digital technologies that have become our daily companions while being aware that there are many reasons for distrust and caution. Concerns about privacy and fear of surveillance co-exist with the collusion

H. Nowotny (✉)
ETH Zürich, Zurich, Switzerland
e-mail: helga.nowotny@wwtf.at

H. Werthner et al. (eds.), *Perspectives on Digital Humanism*,
https://doi.org/10.1007/978-3-030-86144-5_43

of voluntarily handing over our data to the large corporations (Zuboff 2018). The possibilities of abuse and malfunctioning of vulnerabilities to hacking and other forms of cyber-insecurity persist, while optimistic scenarios of new opportunities continue to be acclaimed. We rightly insist that in critical situations, humans ought to be the ones whose judgement trumps automated response and decisions and that accountability must be built into the process in case things turn bad (Christian 2020; Russell 2019). On this co-evolutionary journey and despite the uncertainty of its outcome, we feel encouraged by what may turn out to be an illusion: that we have been dealt the slightly better cards in the co-evolutionary game and human ingenuity will prevail. This is one of the premises on which digital humanism rests, the belief that human values can be instilled in the digital technologies and that a human-centered approach will guide their design, use, and future development (Werthner et al. 2019).

For digital humanisms, such aspirations serve as the necessary preconditions for gaining momentum, but ought not obscure the difficulties that lie ahead. In the long history of technological inventions and innovations, humans always attempted to be in control. What began as deploying tools thousands of years ago to carve out a precarious living from the natural environment turned into massive intervention and large-scale change of the natural environment during industrialization, with devastating consequences for the latter on which we depend. The peak of the belief that humans were in complete control of technology and mastering their future came during modernity (Scott 1999). A turning point was reached in the mid-twentieth century, when it became clear that we were no longer in control over the radioactive waste left behind from the production of the atomic bomb. After the end of the war, the world's population began to grow dramatically, and so did GDP and living standards. At the same time, the impact of human intervention in the earth system and its services began to be noticeably felt. Called "The Great Acceleration," the convergence of these two large-scale developments has not abated since (Steffen et al. 2015; McNeill and Engelke 2015). Today, we are faced with a major sustainability crisis, while digitalization is rapidly gathering momentum with profound and far-reaching implications for what it means to be human and what a good digital society could be. We have arrived in the Anthropocene, and it will be a digital Anthropocene.

Digital humanism thus emerges at a crucial moment, at the intersection of the sustainability crisis and the opportunities offered by digitalization. In order to gauge the challenges it faces, we ought to remind ourselves of the continuities and ruptures it entails. It aspires to build upon some of the great cultural transformations that are part of the European heritage, exploring human nature and adopting a human-centered approach under rapidly changing global circumstances. But digital humanism also harbors a rupture that is less obvious. It marks the transition from the linearity in thinking and understanding the world which was one of the hallmarks of modernity toward coping with the non-linear processes of complex adaptive systems. Just as it is no longer possible to rely on the linearity of technological progress that will inevitably lead to a future being better than the past and the present, digital

humanism must guide us in thinking in non-linear terms when we increasingly face uncertainty and complexity (Nowotny 2015).

Digital humanism therefore must navigate the different strands of our existence that emerge from the inherent tension between humans and machines. In philosophical terms, we speak about life and non-life, about organic and inorganic matter, about different rates of energy conversion needed to keep us and machines going, and, ultimately, about consciousness and their absence in machines (Lee 2020). But as there is little agreement on the definition of these terms and their meaning, the entangled interaction between humans and the digital machines continues in practice as a blurred and messy process. Digital humanism, if it is to be enacted, must be prepared to navigate the manifest and hidden tensions that come to the fore in expected and unexpected ways and in different constellations.

Digitalization exacerbates already existing and familiar tensions between divergent economic, political, and social interests, as amply demonstrated during the COVID-19 pandemic when societal inequalities and fissures were laid bare. Fake news and conspiracy theories continue to circulate freely in the social media, turning science into mere opinion and risking the further destabilization of already fragile liberal democracies. Many unresolved conflicts are linked to rising inequalities. As the digital divide deepens, the fear persists that digitalization will replace jobs faster than new ones will be generated (Susskind 2020). These manifest tensions can ignite major conflicts and further tear apart the social fabric already under considerable stress. Digital humanism cannot abstain from entering this contested arena. It cannot retreat to pursue the ideal of a humanistic and digitally sophisticated individual without considering the digital society that shapes how we live together. Digital humanism will have to come up with designs for new modes of digital governance that can meet the challenges of what a good digital society fit for the twenty-first century could be.

Other tensions are less visible; some are latent or emerging. They hover above the question that constitutes the core of digital humanism: what makes us human and how does the interaction with digital machines change us? Some of these tensions fuel identity anxieties that are directly or indirectly related to social media or the feeling that an algorithm knows us better than we know ourselves. If the experience of acceleration dominated modernity, the prevailing experience in the digital age is informational overload and emotional overextension. As the past reaches into the present and the future has already arrived, at least in the visible form of the latest digital devices, the present becomes densified and further compressed. Digital humanism is challenged to create new spaces in this overheated and hyper-reactive atmosphere in which physical presence needs to be reconciled with virtual space in ways that have yet to be invented. The virus has taught us much about the needs of our bodies in a digital world. Whatever lessons we draw, digital humanism will have to open new pathways for implementing them.

The uncanny efficiency of predictive algorithms and their practical take-over in decision-making pervading our individual and collective life mark another tension-ridden domain for digital humanism to navigate. Whether it is the entire health sector or individual lifestyles, our consumption behavior, or the functioning of our

institutions, predictive algorithms extrapolate from the past to let us see further ahead into the future. Yet, in doing so, they lure us into transferring agency unto them. Once we start to believe that an algorithm can predict what will happen in the future and supportive digital decision-making systems are widely adopted, the point may be reached when human judgement seems superfluous and algorithmic predictions turn into self-fulfilling prophecies (Nowotny 2021).

Thus, the stakes for a digital humanism are high. In order to navigate these tensions, it will have to come up with concrete propositions that include the deeper humanistic layers, going beyond technological solutions. Important as appeals to ethical principles are, they will not suffice unless they can draw in very practical terms on a widely shared set of attitudes and practices that are inspired and guided by a humanistic ideal as a way of living together. This involves devising new ways of tackling problems that go beyond technological fixes and to acknowledge that "wicked problems" exist for which no solutions are in sight, yet they too must be confronted. Digital humanism draws its strength from the conviction that a better digital society is possible, mustering the courage to experiment with what is needed for shaping it.

In practice, this means to cultivate a humanistic sensitivity for the diversity of social contexts in which digital technologies are deployed and efficacious. Currently, neither predictive algorithms nor the data on which they train are sufficiently context-sensitive. Digital humanism can let us discover hitherto unknown features of who we are without determining what we will be. It can teach us the irreplaceable value of critical human judgement when we face the illusionary promise of predictive algorithms that they know the future, which is not determined by any technology but remains uncertain and open. The major benefits of digital processes do not only consist in being "smart," but other potential benefits wait to be explored with an open and curious mind. Digital humanism can sensitize us how to deal with complexity which is closer to our intuitive understanding of what is means to be human than a linear, cause-effect way of thinking. It can attune us to emergent properties and to what remains unpredictable—the ultimate sign of life that keeps evolving.

References

Christian, B. (2020) The Alignment Problem. Machine Learning and Human Values. New York: Norton & Company.

Lee, E. A. (2020) The Coevolution: The Entwined Futures of Humans and Machines. Cambridge, MA: MIT Press.

McNeill, J. R. and Engelke, P. (2015) The Great Acceleration: An Environmental History of the Anthropocene since 1945. Cambridge, MA: Harvard University Press.

Nowotny, H. (2015) The Cunning of Uncertainty. Cambridge, UK: Polity Press.

Nowotny, H. (2021) In AI We Trust. Power, Illusion and Control of Predictive Algorithms. Cambridge, UK: Polity Press.

Russell, S. (2019) Human Compatible: AI and the Problem of Control. London: Allen Lane.

Scott, J 1999, Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed, Yale University Press, New Haven.

Steffen, W. et al. (2015) 'The Trajectory of the Anthropocene: The Great Acceleration', The Anthropocene Review 2:1, pp. 81–98.
Susskind, D. (2020) A World Without Work: Technology, Automation, and How We Should Respond. London: Allen Lane.
Werthner, H et al. (2019) Vienna Manifesto on Digital Humanism, viewed 20 June 2019, <www.informatik.tuwien.ac.at/dighum/>.
Zuboff, S. (2018) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, New York: Hachette Book Group.