IIPP0008 Written submission

Option 4: Carlota Perez's scholarship on techno-economic paradigms and the innovation process has been central to this module. Consider the recent and ongoing development of the ICT-based paradigm (including the prospect of further technological development), and evaluate the prospect that it will lead to more equitable social outcomes, or instead pose the threat of increased inequality and hardship. You should draw upon Perez's historical understanding of how paradigms tend to evolve to discuss what kind of institutions (social, political, legal, economic, etc.) will be necessary to bring about improved outcomes for society.

December 17, 2022

Contents

Technological Revolutions and Techno-economic Paradigms **2**

The promises of AI and the ICT-based paradigm **2**

The perils of AI and the ICT-based paradigm **3**

The institutions of AI and the ICT-based 'Golden Age' **5**

Bibliography **7**

Technological Revolutions and Techno-economic Paradigms

The last 250 years have seen waves of new technologies that have transformed society often in ways that were unimaginable to previous generations. The seminal work of Carlota Perez has demonstrated how *surges of development* propagate new technologies across society. These Technological Revolutions (TRs) change the way people live, work, travel, communicate, and even what we value. This shift in techno-economic paradigm (TEP) eventually becomes the new 'common sense', society having been changed by, and having influenced, technological development (Perez 2002).

In Perez's explanation, each surge has two periods, *Installation* and *Deployment*, separated by a moment (which may be some years!) of transition, the *Turning Point*. With the Installation Period, two phases are apparent, *Irruption* and *Frenzy*, when the new technology first challenges the existing paradigm and then generates a frenzy of speculation. Similarly, the Deployment Period comprises phases of *Synergy* and *Maturity*, when the new technology and society are aligning but then its limits are reached. The current (5^{th}) surge, forming the *Age of Information and Telecommunications* (TR5), began in 1971 with the creation of the microprocessor. This irrupted into the previous surge that formed the *Age of Oil, the Automobile and Mass Production* (TR4), which is considered to have started around 1908 with the Model-T automobile (Perez 2002).

The current techno-economic paradigm (TEP5), arising from TR5, values access to information much as the paradigm of TR4 (TEP4) valued convenience goods. In many ways the mass production of TEP4 is paralleled in TEP5 in massive processing of data to produce insights. However, unlike the uniformity of mass production, the rapid collection, storage, serving and processing of data facilitates diversity and adaptability in goods and services in TEP5. Being virtual and global, the communication infrastructures are shifting the paradigm to decentralised organisation of many aspects of society, disrupting top-down co-ordination regimes and blurring the line between producer and consumer.

According to Perez's analysis, we find ourselves at the Turning Point of TR5. The old paradigm is still in place for much of society, regrettably driving one of humanity's most perilous creations, Anthropogenic Global Warming (AGW). The new paradigm is emerging, offering solutions to societal problems and yet creating a whole new set of hazards. My own field of work, Artificial Intelligence (AI), embodies many of these promises and perils. In the rest of this essay, I will outline the mixed blessings of AI for society, and describe broadly some of the institutions that will be required to bring about the most equitable outcomes.

The promises of AI and the ICT-based paradigm

The new ICT-based paradigm (TEP5), and especially recent advances in AI, have the potential to revolutionise lifestyles in ways that promote human wellbeing. For instance by the personalisation of health and social care (Dias and Torkamani 2019; Topol 2019) and education (Bhutoria 2022), and the reduction in barriers to cooperation such as language (Wang *et al.* 2021) or culture (IVOW AI Developers 2021). Work in the field of AI can be deeply creative and stimulating. Jobs are accessible to people with a wide range of physical abilities and neurodiversity. There is currently a deficiency of data scientists and research scientists, programmers, and AI practitioners, meaning that in some sectors these

roles command a good wage (Nolan 2022). Further, as during previous technological revolutions, the development of TEP5 will generate a new system of service roles such as engineers, designers and trades (Perez 2002, p34; Perez and Murray-Leach 2018).

AI can take on dull repetitive work and thus boost productivity whilst leaving the more interesting roles to humans (Lopez 2020). It offers the potential for knowledge-discovery from existing data, speeding up advances in science and engineering (e.g. Tshitoyan *et al.* 2019) and optimisation of the design of materials, components and machines (Montes de Oca Zapiain *et al.* 2021; Aziz *et al.* 2021). Even the uncertainties of government decision-making may be reduced by applying AI to determine the likely outcomes of policies and possible intervention points under a range of conditions (e.g. Castilla-Rho *et al.* 2017).

Perhaps most critical is that TR5 may bring solutions to the inequity created by TR4, such as the the damage caused by AGW to lives and livelihoods of communities in the Global South and future generations. In their thorough survey of the application of machine learning to AGW, Rolnick *et al.* (2022) identify 33 broad applications that span mapping, measuring and monitoring the physical and biological environment, researching and managing systems, energy and materials, and identifying and precipitating the potential technological and social transitions. Further, non-learning AI such as agent-based systems and knowledge modelling could offer alternative tools for discovering and testing technological and social interventions for mitigation of, and adaptation to, AGW (Castro *et al.* 2020).

The perils of AI and the ICT-based paradigm

In her work on technological revolutions, Perez (2002, p39) identifies many dimensions of the economy that experience tension as a consequence of the irruption of the new technology. These tensions largely arise from the divergent fortunes between the old and the new paradigms with a loss of jobs and displacement of activities to new regions and industries Perez (2002, p26-7). We have been witnessing this over the last decades as the wealth and comfort of those established in some aspects of the ICT-based industries and financial services have been able to command well-paid jobs and benefits compared to mass-production-based industries.

The high wages commanded by workers in AI and related roles are generally in the private sector, particularly in Big Tech (Mazzucato *et al.* 2022). This also causes a drain of some of the brightest out of the public sector into roles that benefit society much less, such as creating algorithms that provoke excesses within the financial markets (Banerjee and Duflo 2019, p245). Further, AI also requires an invisible workforce to complete what Gray and Suri (2019) refer to as "technology's last mile", such as labelling data and content moderation. Gray and Suri refer to this as "Ghost Work", likening it to the piecework of the Industrial Revolution, because it is undertaken in private homes without the protections enjoyed by those in mainstream employment. Consequently, these jobs are precarious, poorly paid (when they are paid) and can present physical and mental hazards to workers (Gray and Suri 2019).

Within Perez's understanding of technological revolutions the Frenzy phase is a particularly destructive time of high inequality due to speculation on financial markets. This speculation can also occur within the technology itself leading to "Science Bubbles"

whereby scientists and technologists are incentivised to inflate the merits of their work and to work only in the most fashionable fields (Pedersen and Hendricks 2014; Arenas and Gil-Lafuente 2021). Within AI, the failure to meet expectations has previously resulted in "AI Winters" when investment and thus progress stalled (Mitchell 2021) and recent focus of machine learning on competition-style problems reduces empirical rigour, ultimately risking further downturns (Sculley *et al.* 2018). Further, the competitive pressures in the field cause most AI practitioners to neglect ethical considerations (Strümke *et al.* 2021) much as environmental considerations were neglected in the drive to accelerate mass production.

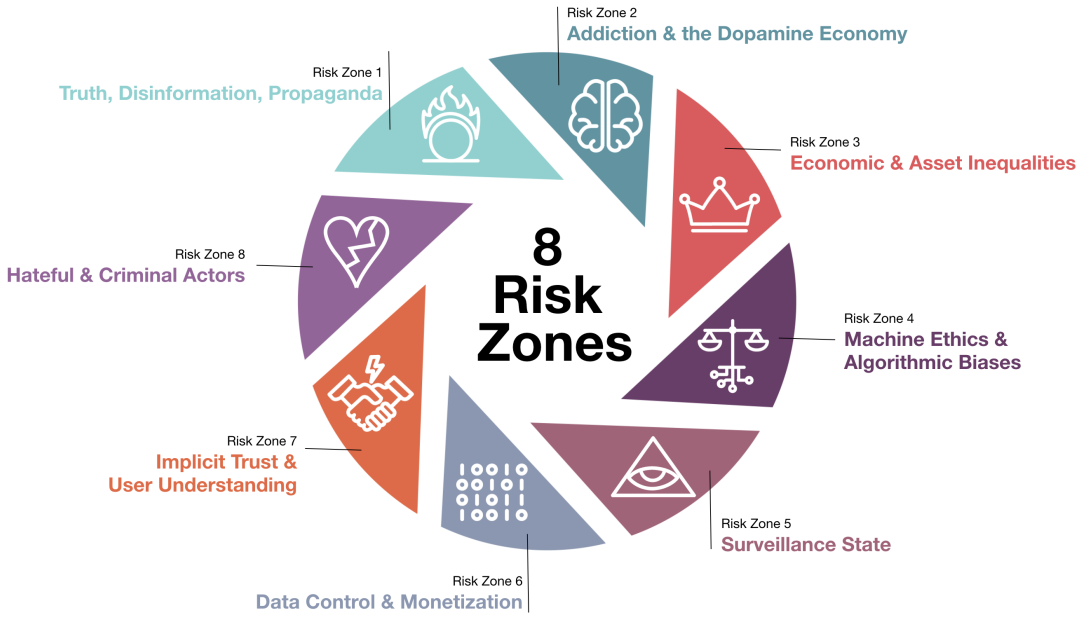

Figure 1: The "Ethical OS" work from Omidyar Network and Institute for the Future identified 8 risk zones (with the capacity for users to define further risk zones) to generate discussion about the risks posed tomorrow by what is being built today (Omidyar Network and Institute for the Future 2018)

AI, having the power to rapidly derive insights from vast volumes of data brings with it a broad range of threats to individuals, communities and society. The work of the Omidyar Network (Omidyar Network and Institute for the Future 2018) identifies 8 'Technical Risk Zones'[1] that can arise from a wide range of technologies, especially AI (see Figure 1). These are particularly pernicious considering the ability of AI to extract data from every interaction we have with devices and lack of awareness of most users of this, as described by Zuboff (2015) and allowing Big Tech and third parties to exploit our vulnerabilities. Because data are never unbiased (Hooker 2021), AI can amplify inequality by deliberately or accidentally profiling populations or simply excluding them from the decision-making process (e.g. Brown *et al.* 2020; Banerjee, Bhimireddy, *et al.* 2021; Amuasi 2022) with particular impact on those already disadvantaged. Efforts to manage these ethical risks have the potential to deepen harm when they are being undertaken by Big Tech itself (Abdalla and Abdalla 2021).

[1]This is one of many frameworks identifying possible harms from AI e.g. Palm and Ove Hansson 2006; The IEEE Global Initiative 2019; Bird *et al.* 2020; Yang 2021

The institutions of AI and the ICT-based 'Golden Age'

A range of new institutions is required to realise the promises and avert the perils of TEP5 (Perez 2002, p52-53; Perez and Murray-Leach 2018). A regime of regulation and legislation will be needed to avert the harms of Surveillance Capitalism, Ghost Work, investment bubbles, and other ethical hazards arising from AI and related technologies. These would play a similar role to the regulation and enforcement organisations such as UK Environment Agency, which is tasked with monitoring and prosecuting the environmental damage that tends to arise from TEP4. There are moves towards this, such as regulation of AI in the European Union (2021), the development of an AI Council in the UK, which is forming an AI strategy with ethical foundations (The AI Council 2021), and standards for design of robotic and autonomous systems (British Standards Institute 2016). Such institutions will need to have authority and independence such that even state use of AI is fully scrutinised.

AI will pervade every sector of the economy and society and needs to be understood to at least a foundational level by everyone to prevent misinterpretation of its products and services. Therefore education programmes in schools, universities and throughout industry will be necessary. Further, training must be made available to support those in industries that are being lost to the new paradigm, potentially creating a large workforce of specialists in AI and related roles that fill the gaps in worker supply that we are currently experiencing.

The new paradigm requires new organisational structures in both the private and public sector to coordinate and create efficiencies in particular by easing the flow of information between different spheres of activity (Simon 1991; Perez 2010). These must centre ethical safety in the same way that health and safety needed centring in the factories of mass production (Ayling and Chapman 2020; Kelley 2022), employing mechanisms to prevent AI amplifying inequality and inaccuracy, exposing personal data, and enabling malicious activities (Leslie 2019; Open Data Institute 2019; Davis 2022). It may also be pertinent to bring AI activities closer to social welfare and environmental protection bodies to accelerate the development of solutions to our most pressing challenges.

The promises of TEP5 are unlikely to be fully realised unless the TR emerges out of the Turning Point and enters the prosperous Synergy phase (Perez 2002). The advent of this 'Golden Age' requires state institutions that are willing to tame market speculation, and set clear direction for business and society towards the new paradigm. This seems a distant prospect under current political conditions. Thus, new institutions of civil organisation may be essential to precipitate change both in terms of AI and action on AGW (Zeppini *et al.* 2014; Lenton *et al.* 2022). If this is achieved, new decision-making institutions will be vital to fully engage citizens in creating a more egalitarian society. Citizens who are already immersed in TEP5 are accustomed to the ease of cooperating and sharing knowledge. As the paradigm embeds, citizens will expect to be involved in decisions about civil society such as specifying the challenges to which the new technologies will be applied and the values to which it must adhere. Citizen engagement will also provide early warning of potential harms. This requires institutional cultures, structures and tools that facilitate co-design of the desired technological direction (Ostrom 1996; Mazzucato 2021, p199-202).

The new techno-economic paradigm promises to overcome the inequalities of the previous

paradigm and yet brings a new array of risks. AI, in particular, is likely to transform every aspect of our lives and remains little understood by many who will be affected by it. Adapting our institutions to this paradigm will require determination and vigilance, and a willingness to involve society in the transition. There are small indications within society and a few organisations of a desire to implement changes to mitigate AI's threats whilst exploiting it to solve our greatest challenges.

Bibliography

Abdalla, Mohamed and Moustafa Abdalla (2021). "The Grey Hoodie Project: Big Tobacco, Big Tech, and the threat on academic integrity". In: *arXiv:2009.13676*. URL: https://arxiv.org/abs/2009.13676.

Amuasi, John (2022). "Representation Learning in the Global South: Societal Considerations-Fairness, Safety and Privacy". In: *International Conference on Learning Representations*. URL: https://iclr.cc/virtual/2022/invited-talk/7235.

Arenas, Laura and Anna María Gil-Lafuente (2021). "Emerging Technologies, Innovation, and Volatility: A MiniReview". In: *International Journal of Sensor Networks and Data Communications*. eprint: https://www.hilarispublisher.com/open-access/emerging-technologies-innovation-and-volatility-a-minireview.pdf.

Ayling, Jacqui and Adriane Chapman (2020). "Putting AI ethics to work: are the tools ft for purpose?" In: *AI and Ethics*. DOI: https://doi.org/10.1007/s43681-021-00084-x. URL: https://link.springer.com/content/pdf/10.1007/s43681-021-00084-x.pdf.

Aziz, Nurhasyimah Abd *et al.* (2021). "Component design optimisation based on artificial intelligence in support of additive manufacturing repair and restoration: Current status and future outlook for remanufacturing". In: *Journal of Cleaner Production* 296, p. 126401. ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2021.126401. URL: https://www.sciencedirect.com/science/article/pii/S0959652621006211.

Banerjee, Abhijit V and Esther Duflo (2019). *Good Economics for Hard Times*.

Banerjee, Imon, Ananth Reddy Bhimireddy, *et al.* (2021). "Reading Race: AI Recognises Patient's Racial Identity In Medical Images". In: arXiv: 2107.10356 [cs.CV]. URL: https://arxiv.org/abs/2107.10356.

Bhutoria, Aditi (2022). "Personalized education and Artificial Intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model". In: *Computers and Education: Artificial Intelligence* 3, p. 100068. ISSN: 2666-920X. DOI: https://doi.org/10.1016/j.caeai.2022.100068. URL: https://www.sciencedirect.com/science/article/pii/S2666920X22000236.

Bird, Eleanor *et al.* (2020). *The ethics of artificial intelligence. Issues and initiatives*. Tech. rep. Panel for the Future of Science and Technology, European Parliamentary Research Service. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf.

British Standards Institute (2016). *BS 8611. Guide to the Ethical Design of Robots and Robotic Systems*. URL: https://shop.bsigroup.com/ProductDetail?pid=000000000030320089.

Brown, Tom B. *et al.* (2020). "Language Models are Few-Shot Learners". In: arXiv: 2005.14165 [cs.CL]. URL: https://arxiv.org/abs/2005.14165.

Castilla-Rho, Juan Carlos *et al.* (2017). "Social tipping points in global groundwater management". In: *Nature Human Behaviour* 1 (9), pp. 640–649. DOI: 10.1038/s41562-017-0181-7. URL: https://www.nature.com/articles/s41562-017-0181-7.

Castro, Juana *et al.* (2020). "A review of agent-based modeling of climate-energy policy". In: *WIREs Climate Change* 11 (4). DOI: 10.1002/wcc.647. URL: https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/wcc.647.

Davis, Jenny (2022). "'Affordances' for Machine Learning". In: *International Conference on Learning Representations*.

Dias, Raquel and Ali Torkamani (2019). "Artificial intelligence in clinical and genomic diagnostics". In: *Genome Medicine*. URL: https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-019-0689-8.

European Union (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence*. URL: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence.

Gray, Mary L. and Siddharth Suri (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*.

Hooker, Sara (2021). "Moving beyond "algorithmic bias is a data problem"". In: *Patterns* 2 (4). DOI: 10.1016/j.patter.2021.100241. URL: https://www.sciencedirect.com/science/article/pii/S2666389921000611.

IVOW AI Developers (2021). *Indigenous Knowledge Graph*. URL: https://www.ivow.ai/ikgstories.html.

Kelley, Stephanie (2022). "Employee Perceptions of the Effective Adoption of AI Principles". In: *Journal of Business Ethics*. URL: https://link.springer.com/article/10.1007/s10551-022-05051-y.

Lenton, Timothy M. *et al.* (2022). "Operationalising positive tipping points towards global sustainability". In: *Global Sustainability*. URL: https://www.cambridge.org/core/journals/global-sustainability/article/operationalising-positive-tipping-points-towards-global-sustainability/8E318C85A8E462AEC26913EC43FE60B1.

Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. Tech. rep. The Alan Turing Institute. DOI: https://doi.org/10.5281/zenodo.3240529. URL: https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf.

Lopez, Elyssa (2020). *Artificial intelligence: friend or foe to Philippine call centre workers?* URL: https://www.scmp.com/week-asia/economics/article/3073999/artificial-intelligence-friend-or-foe-philippine-call-centre (visited on 12/11/2022).

Mazzucato, Mariana (2021). *Mission Economy. A Moonshot Guide to Changing Capitalism*.

Mazzucato, Mariana *et al.* (2022). *Governing artificial intelligence in the public interest*. Tech. rep. UCL Institute forInnovation andPublic Purpose / Stanford Cyber Policy Center. Chap. Working Paper Series (IIPPWP 2022-12). URL: https://www.ucl.ac.uk/bartlett/public-purpose/wp2022-12.

Mitchell, Melanie (2021). "Why AI is Harder Than We Think". In: DOI: 10.48550/ARXIV.2104.12871. URL: https://arxiv.org/abs/2104.12871.

Montes de Oca Zapiain, David, James A. Stewart, and Rémi Dingreville (2021). "Accelerating phase-field-based microstructure evolution predictions via surrogate models trained by machine learning methods". In: *npj Computational Materials* 7 (3). DOI: 10.1038/s41524-020-00471-8. URL: https://www.nature.com/articles/s41524-020-00471-8.

Nolan, Beatrice (2022). *The average data scientist earns almost $100,000 a year - and the barriers to entry for candidates are being broken down*. English. URL: https://www.proquest.com/newspapers/average-data-scientist-earns-almost-100-000-year/docview/2680788252/se-2.

Omidyar Network and Institute for the Future (2018). *Ethical OS Toolkit*. URL: https://ethicalos.org/.

Open Data Institute (2019). *ODI Data Ethics Canvas*. URL: https://theodi.org/article/data-ethics-canvas (visited on 02/04/2021).

Ostrom, Elinor (1996). "Crossing the great divide: Coproduction, synergy, and development". In: *World Development* 24.6, pp. 1073–1087. ISSN: 0305-750X. DOI: https://doi.org/10.1016/0305-750X(96)00023-X. URL: https://www.sciencedirect.com/science/article/pii/0305750X9600023X.

Palm, Elin and Sven Ove Hansson (2006). "The case for ethical technology assessment (eTA)". In: *Technological Forecasting and Social Change*. DOI: https://doi.org/10.1016/j.techfore.2005.06.002. URL: https://www.sciencedirect.com/science/article/pii/S004016250500082X.

Pedersen, David Budtz and Vincent F. Hendricks (2014). "Science Bubbles". In: *Philosophy & Technology* 27, pp. 503–518. URL: https://link.springer.com/article/10.1007/s13347-013-0142-7.

Perez, Carlota (2002). *Technological Revolutions and Financial Capital*. ISBN: 9781781005323.

– (2010). "Technological revolutions and techno-economic paradigms". In: *Cambridge Journal of Economics* 34, pp. 185–202. DOI: 10.1093/cje/bep051.

Perez, Carlota and Tamsin Murray-Leach (2018). "Smart & Green: A New 'European Way of Life' as the Path for Growth, Jobs and Wellbeing". In: *Re:Thinking Europe: Positions on Shaping an Idea, Austrian Council for Research and Technology Development*. Vienna: Verlag Holzhausen, pp. 208–223.

Rolnick, David *et al.* (2022). "Tackling Climate Change with Machine Learning". In: *ACM Computing Surveys* 55.2. ISSN: 0360-0300. DOI: 10.1145/3485128. URL: https://dl.acm.org/doi/full/10.1145/3485128.

Sculley, D. *et al.* (2018). "Winner's Curse? On Pace, Progress, and Empirical Rigor". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=rJWF0Fywf&noteId=rJWF0Fywf.

Simon, Herbert A. (1991). "Organizations and Markets". In: *Journal of Economic Perspectives* 5 (2), pp. 25–44. DOI: 10.1257/jep.5.2.25. URL: https://www.aeaweb.org/articles?id=10.1257/jep.5.2.25.

Strümke, I., M. Slavkovik, and V.I. Madai (2021). "The social dilemma in artificial intelligence development and why we have to solve it". In: *AI Ethics*. DOI: https://doi.org/10.1007/s43681-021-00120-w. URL: https://link.springer.com/article/10.1007/s43681-021-00120-w.

The AI Council (2021). *AI Roadmap*. Office for Artificial Intelligence, Department for Business, Energy & Industrial Strategy, and Department for Digital, Culture, Media & Sport.

The IEEE Global Initiative (2019). *Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. Tech. rep. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. eprint: https://ethicsinaction.ieee.org/wp-content/uploads/ead1e.pdf. URL: https://ethicsinaction.ieee.org/.

Topol, Eric (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books. ISBN: 9781541644649. URL: https://books.google.co.uk/books?id=%5C_EFlDwAAQBAJ.

Tshitoyan, Vahe *et al.* (2019). "Unsupervised word embeddings capture latent knowledge from materials science literature". In: *Nature* 571, pp. 95–98. DOI: 10.1038/s41586-019-1335-8. URL: https://www.nature.com/articles/s41586-019-1335-8.

Wang, Haifeng *et al.* (2021). "Progress in Machine Translation". In: *Engineering*. ISSN: 2095-8099. DOI: https://doi.org/10.1016/j.eng.2021.03.023. URL: https://www.sciencedirect.com/science/article/pii/S2095809921002745.

Yang, Selene (2021). *Feminism, ethics and geospatial data. A brief reflection towards their joint analysis*. Tech. rep. Working paper (n). Montevideo: ILDA. DOI: https://doi.org/10.5281/zenodo.4681033. URL: https://docs.google.com/document/d/1Ee81bssn8CSdRRD2RhlYoNy10r0jhpk1niz8l59VpPo.

Zeppini, Paolo, Koen Frenken, and Roland Kupers (2014). "Thresholds models of technological transitions". In: *Environmental Innovation and Societal Transitions* 11, pp. 54–70. ISSN: 2210-4224. DOI: https://doi.org/10.1016/j.eist.2013.10.002. URL: https://www.sciencedirect.com/science/article/pii/S2210422413000713.

Zuboff, Shoshana (2015). "Big other: Surveillance Capitalism and the Prospects of an Information Civilization". In: *Journal of Information Technology* 30.1, pp. 75–89. DOI: 10.1057/jit.2015.5. eprint: https://journals.sagepub.com/doi/epdf/10.1057/jit.2015.5. URL: https://doi.org/10.1057/jit.2015.5.