

Big Data and A.I. for the SDGs:

Private corporation involvement in SDG data-driven development, policy and decision-making

Ciarán O'Brien (UCD Centre for Sustainable Development Studies)

Abstract

Big Data and A.I. offers vast analytical opportunities for policy makers to assess complex real-time information to track and monitor SDG progress, and to design policies toward SDG advancement. However, Big Data and A.I. based decision-making can reinforce and magnify pre-existing inequalities and bias, whilst distorting or amplifying real world exclusion and discrimination against marginalised groups. This paper discusses private corporation involvement in the provision of Big Data and A.I. services, and their potential negative impact on 2030 Agenda policy and decision-making. Alongside assessment of existing concerns regarding the erosion of data privacy, data colonialism and the enclosure of the global digital commons, through unequal power dynamics between corporations, governments and citizens. The paper's recommendations show how the utilisation of Big Data and A.I. for SDG advancement can be capitalised upon through the construction of dynamic, robust and inclusive public data ecosystems, via UN funded data co-operatives and algorithmic accountability programmes.

Introduction

'Quality, accessible, timely and reliable disaggregated data will be needed to help with the measurement of progress and to ensure that no one is left behind. Such data is key to decision-making'. (United Nations, 2015, p. 13)

Big Data¹ and A.I.² (henceforth, BDAI) presents a vast opportunity for the advancement of the SDGs (AI4SDGs, 2021; Bhutani & Paliwal, 2015; Seele & Lock, 2017;), and offers vast analytical possibilities to track and monitor progress of SDG Indicators (Letouzé, 2015; Wu *et al.*, 2020). BDAI can assist SDG policy-makers (Xu *et al.*, 2020; Milano *et al.*, 2014) with real-time structural information (MacFeely, 2019), regulatory oversight (Yaeger *et al.*, 2019) and governance (Sharma *et al.*, 2020). There exists concerns over BDAI, and its utilisation in sustainable development as BDAI-based decision-making can reinforce and magnify pre-existing inequalities and bias, whilst distorting or amplifying real world exclusion and discrimination against marginalised groups (Karimi *et al.*, 2018), unless diversity and inclusion are factored in (Mitchell *et al.*, 2020). Additionally, there are predictions that BDAI may have negative economic and social impacts on

developing nations in the near future, by reducing employment (Carmody, 2020) and through the reshoring of ICT³ activities back to advanced economies (Baldwin, 2019). Whilst BDAI is a positive enabler of over 79% of SDG Targets, it can negatively affect over 35%, as per Figure 1 below.

Across a variety of UN agencies^{4 5} BDAI is utilised for the SDGs in public-private partnerships (TReNDS, 2021b; UNGP, 2021a), an approach that is reflected by other international development agencies (USAID, 2021). UNESCAP (2017) positions these public-private partnerships (PPP) as a form of data philanthropy, a win-win partnership under the structure of corporate social responsibility. This relationship affords private corporations the opportunity to improve their reputation through the provision of open data services, which allow the UN to undertake complex data analysis

¹ Big Data are data sets whose size, type or complexity are beyond the ability of traditional methodologies or relational databases. For further information please see <https://www.ibm.com/analytics/hadoop/big-data-analytics>

² A.I. - Artificial Intelligence uses machine learning to mimic human problem-solving and decision-making capabilities and in some case try to exceed them. For further information please see <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>

³ Information and communications technology (ICT) - technological and communication tools used to transmit,

store, create or share information and data. For further information please see <http://uis.unesco.org/en/glossary-term/information-and-communication-technologies-ict>

⁴ UN Global Pulse - works towards developing applications of BDAI for sustainable development and humanitarian action on a variety of innovative projects for social good. (UNGP, 2021b)

⁵ SDNSN Thematic Research Network on Data and Statistics (TReNDS) - utilise BDAI towards the SDGs through a variety of data projects. (TReNDS, 2021a)

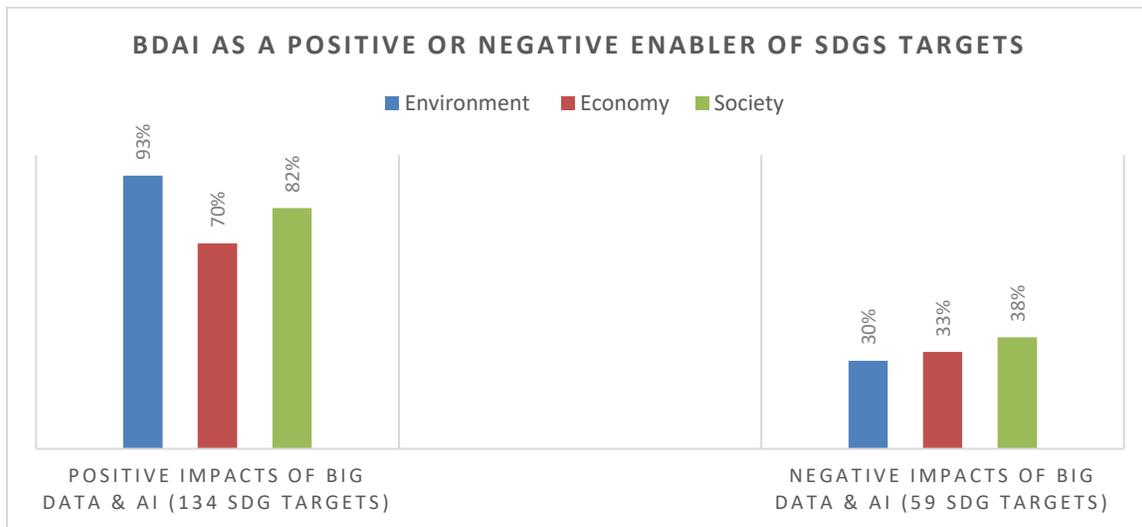
of SDG indicators, subsequently informing policy and decisions, and may allow developing countries to bridge the data gap.

The data gap in question is large, with over four billion people and 93% of households in developing countries without access to the Internet (USAID, 2021). Additionally there exists a vast data disparity between states (Schmidt-Traub *et al.*, 2017), with 72% of countries requiring external assistance (FAO, 2019). As a result, reducing the data gap and digital inequality is key to the SDGs (Hassani *et al.*, 2021), as gaps prevent developing countries adequately assessing their citizen's needs, which impacts upon their ability to design the correct policies to support those (Stuart *et al.*, 2015). Collaborations with the private sector are

promoted as a positive mechanism for SDG outcomes, as they sectors has the economies of scale to both leverage BDAI to foster social and economic good via effective data use, and to surmount the data gap (Viswanathan *et al.*, 2021).

Notwithstanding, a number of questions emerge, such as who should bridge this divide? Who are the end users of these SDGs BDAI ecosystems and algorithms; governments? Are they UN agencies, NGOs or private corporations? If so, what checks and balances are required to ensure accountability and transparency? What mechanisms and frameworks should be in place to ensure that these data ecosystem positively contribute towards the SDGs?

Figure 1. BDAI as a positive or negative enabler of SDGs Targets



Source: The role of artificial intelligence in achieving the Sustainable Development Goals (Vinuesa *et al.*, 2020)

Concerns

The reliance on private corporations within PPPs to collate and control the vast social, economic, environmental and geospatial data required for the 2030 Agenda can lead towards private appropriation of the digital data sphere, which can erode data privacy (Foster & McChesney, 2014). Resulting in 'data colonialism' (Stürmer *et al.*, 2021), as corporations grab vast tracts of the global data market, and enclose the digital commons. Data colonialism also has the possibility of drawing developing states into asymmetrical Global North-South trade and investment scenarios (Howson, 2020).

This leads towards further dependence on the private sector by governments, and creates long-term consequences of data privacy, creating a 'society of control' (Sadowski & Pasquale, 2015), and the

possibility of a big tech global gilded age. This disproportionately would go against the goal of 'leaving no one behind', as these PPPs can reinforce unequal power dynamics between corporations, governments and citizens resulting in unequal distribution of potential benefits of the SDG, (Cheng, 2020), as they may inevitably force governments into reproducing and expanding socioeconomic policies that only benefit private corporations (Wood & Monahan, 2019).

Private driven data collection becomes a form of capital to be extracted and accumulated, resulting in a drive towards data extraction without private consent (Sadowski, 2019) as private corporations fail to respect cultural or political norms and rights, driven by the data imperative (Schildt, 2020) to extract as much data as possible by any means necessary (Fourcade and Healy, 2017). This allows private corporations to control access to the conditions of capital production, in this

case data and the data ecosystems⁶, resulting in 'digital rentierism', granting them the 'ability to derive income (rent) from access to these data assets' (Sadowski, 2020a, p. 565).

Critics point towards the UN Global Pulse's Data for Climate Action as an example of a SDG focused private-public collaboration that allows private corporations to extract and accumulate profit-oriented data by framing 'personal data as necessary evidence for crucial decision making', (Espinoza & Aronczyk, 2021, p. 9).

Private corporations' data-driven systems can reproduce and reflect inequalities as they can be manipulated towards reinforcing societal bias and discrimination (O'Neil, 2016). They can also dehumanise citizens and societal interactions by reducing them into data sets, diminishing their complexity, and allowing elites to exert control over marginalised communities (Sadowski, 2020b). This reduces democratic ability to shape our social, cultural and political spheres (Mayer-Schönberger and Ramge, 2018), decreasing the huge potential of BDAI towards the SDGs in terms of policy and decision making (Ryan *et al.*, 2019).

Recommendations

Utilising BDAI for good is necessary in achieving the SDGs, and requires the construction of dynamic, robust and inclusive data ecosystems for all. BDAI policy and decision-making is highly centralised, resulting in a propensity towards elite control that necessitates a public data ecosystem based on cooperation, self-regulation and participatory empowerment (Helbing & Hausladen, 2022). In the construction of these data ecosystems, UN agencies need to ensure that human and data rights (Soh *et al.*, 2018), alongside civil society perspectives are included in their construction to help identify issues and bias, and to leverage data to ensure effective SDG implementation (The Danish Institute for Human Rights, 2021).

If we wish to leave no one behind we must ensure we give voice and agency to all. Moving governance away from private companies, private-public partnerships or impartial algorithms towards an economy of information that is open and transparent in civic engagement (Couldry & Powell, 2014). This process should be rooted in the three pillar concepts of international data justice (Taylor, 2017), that would integrate transnational data rights and privacy-based

freedoms into publically owned data systems and analytics.

In the avoidance of the potential long-term negative impacts in using data platforms by private companies, UN policies could be directed towards open knowledge and a shared social data economy via digital cooperativism (Scholz, 2016). This can be achieved via open digital commons models and cooperative principles that can foster a positive BDAI ecosystem for the achievement of the SDGs (Morell *et al.*, 2020). Key to this would be 'data co-operatives' that would allow UN agencies, governments, civil society and citizens to avoid the pitfalls of data extractivism and surveillance capitalism (Zuboff, 2019), towards a future of indigenous data sovereignty (Calzada, 2021a).

Data co-operatives act like credit unions for data management storage, with their obligation solely towards benefiting its members (Pentland & Haradjono, 2021a; Pentland & Haradjono, 2021b), creating a citizen-driven and controlled data ecosystem (Calzada, 2021b; Blasimme *et al.*, 2018). These can be combined with strong governance frameworks to ensure democratic platforms that protect citizen's rights (Mathiesen, 2014), and ensure countries can control data flows for data sovereignty, avoiding data colonialism (Mejias, 2019). The EU's General Data Protection Regulation (GDPR) (GDPR, 2021), Strategy for Data (EC, 2020) and the Data Governance Act (EC, 2021), provides a governance framework for national data sovereignty. Whilst the proposed data-cooperative framework would allow citizens to decide how their data is stored and used (Craglia *et al.*, 2021), and is based on Rawl's concept of a 'property-owning democracy' (Loi, *et al.*, 2020).

In summary, this policy brief's recommendations are as follows,

1. UN Agencies should provide the required funding to close technological and data gaps between developing and developed nations.
2. Fund and enable civil societies and government partnerships to construct and scale up data co-operatives to foster community-driven data sovereignty.
3. Fund and/or provide the required training to these co-operatives that allow them to operate independently.

⁶ Data ecosystem is a platform that combines data and information from numerous providers and constructs value through their process and usage. For further information

please see <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/tech-forward/data-ecosystems-made-simple>

4. Establish algorithmic accountability programmes to ensure AI-based systems do not reinforce institutional bias, unequal power structures and inequalities (Diakopoulos, 2015). Within the data co-ops this would also entail consistent systemic examination of pre-, in-, and post-processing methods of data by AI (Ntoutsis *et al.*, 2019).
5. Advance and establish effective good digital governance procedures and regulatory policies (Elmi, 2020), that establishes accountability for global societal data (Shiffman & Shawar, 2020).
6. Ensure data co-operatives and their practitioners are vital stakeholders in designing and implementing policies related to future SDG BDAI decision-making.

'When artificial intelligence and Big Data meet the social reality of human coordination, we can become more sustainable'. (Mayer-Schönberger and Ramge, 2018, p. 211)

References

- AI for Sustainable Development Goals - AI4SDGs (2021) *SDG Projects*. Available at: <https://ai-for-sdgs.academy/topics> (Accessed 09 October 2021)
- Baldwin, Richard (2019). *The Globotics Upheaval: Globalisation, Robotics and the Future of Work*. Oxford and New York: Oxford University Press
- Bhutani, S. and Paliwal, Y. (2015) 'Digitalization: A Step towards Sustainable Development', *OIDA International Journal of Sustainable Development* 08(12)
- Blasimme, A., Vayena, E., and Ernst, H. (2018) 'Democratizing health research through data cooperatives', *Philos. Technol.* 31, pp. 473–479
- Calzada, I. (2021a) 'Data Co-operatives through Data Sovereignty', *Smart Cities*, 4, pp. 1158-1172
- Calzada, I. (2021b) 'Devolving Smart City Citizenship: Smart City-Regions, Data Devolution, and Technological Sovereignty', in Igor, C., (ed.) *Smart City Citizenship*, Elsevier Science Publishing Co Inc.: Cambridge, MA, USA
- Carmody, Pdraig, (2020) 'Meta-trends in Global Value Chains and Development: Interacting Impacts with COVID-19 in Africa', *Transnational Corporations Journal*, 27(2).
- Cheng, H.W.J. (2020) 'Economic properties of data and the monopolistic tendencies of data economy: policies to limit an Orwellian possibility', *UN Department of Economic and Social Affairs Working Paper No. 1'64, ST/ESA/2020/DWP/164*, May 2020.
- Couldry, N. and Powell, A. (2014) 'Big Data from the bottom up', *Big Data & Society*. doi: 10.1177/2053951714539277.
- Craglia, M. *et al.* (2021) *Digitranscope: The Governance of Digitally-Transformed Society*; EUR 30590 EN; Publications Office of the European Union: Luxembourg, 2021; ISBN 978-92-76-30229-2. JRC 123362; Available at: <https://publications.jrc.ec.europa.eu/repository/handle/JRC123362>
- Diakopoulos, N. (2015). 'Algorithmic accountability: Journalistic investigation of computational power structures', *Digital Journalism*, 3(3), pp. 398–415.
- Elmi, N. (2020) 'Is Big Tech Setting Africa Back?', *Foreign Policy*, November 11. Available at: <https://foreignpolicy.com/2020/11/11/is-big-tech-setting-africa-back/> (Accessed 01 November 2021)
- Espinoza M. I. and Aronczyk, M. (2021) 'Big data for climate action or climate action for big data?', *Big Data & Society*, 1-15
- EU GDPR (General Data Protection Regulation) (2021) *General Data Protection Regulation*. Available at: www.gdpr-info.eu (Accessed on 10 October 2021).
- European Commission (2020) *A European Strategy for Data*, European Commission: Brussels, Belgium.
- European Commission (2021) *Proposal Regulation: European Data Governance Act*. Available at: <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-data-governance-data-governance-act> (Accessed on 10 October 2021).
- FAO (2019) *Investing in data for the SDGs: Why good numbers matter*. Available at: <http://www.fao.org/partnerships/resource-partners/news/news-article/en/c/1200471/> (Accessed 21 October 2021).
- Foster J.B. and McChesney, R.W. (2014) Surveillance Capitalism: Monopoly-Finance Capital, the Military-Industrial Complex, and the Digital Age, *Monthly Review*, 66(3)
- Fourcade M and Healy K (2017) 'Seeing like a market', *Socio-Economic Review*, 15(1), pp. 9–29.
- Fuster Morell, M., Espelt, R., and Senabre Hidalgo, E. (2020) 'Data for Sustainable Platform Economy: Connections between Platform Models and Sustainable Development Goals', *Data*, 6(7).
- Hardjono, T. and Pentland, A. (2021b) *Data Cooperatives: Towards a Foundation for Decentralized Personal Data Management*. arXiv. Available at: <https://ui.adsabs.harvard.edu/abs/2019arXiv190508819H> (Accessed on 01 November 2021).

Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, May 2022

- Hassani, H., Huang, X., MacFeely, S. and Entezaruan, M.R. (2021) 'Big Data and the United Nations Sustainable Development Goals (UN SDGs) at a Glance', *Big Data Cogn. Comput.* 5 (28)
- Helbing, D. and Hausladen, C. (2022), 'Socio-Economic Implications of the Digital Revolution', in Chen, P., Elsner, W. and Pyka, A. (eds.), *Handbook of Complexity Economics*, Routledge, London, New York, 2022.
- Howson, P. (2020) 'Crypto-Colonialism: Conjuring Value on the Blockchain Frontiers of the Global South', *Frontiers in Blockchain* 3(22), doi: 10.3389/fbloc.2020.00022
- Karimi, F., Génois, M., Wagner, C., Singer, P., & Strohmaier, M. (2018). 'Homophily influences ranking of minorities in social networks'. *Scientific Reports*, 8.
- Kurt A. Yaeger, Michael Martini, Gal Yaniv, Eric K. Oermann, Anthony B. Costa, (2019) 'United States regulatory approval of medical devices and software applications enhanced by artificial intelligence', *Health Policy and Technology*, 8(2), pp. 192-197, ISSN 2211-8837
- Letouzé, E. (2015) *Reflections on Big Data & the Sustainable Development Goals: Measuring & Achieving Development Progress in the Big Data Era*, February 2015, Data-Pop Alliance, Hard Humanitarian Initiative.
- Loi, M., Dehaye, P.-O., and Hafen, E. (2020) 'Towards Rawlsian 'property-owning-democracy' through personal data platform cooperatives', *Crit. Rev. Int. Soc. Political Philos.* 1–19.
- MacFeely, S. (2019) 'The Big (data) Bang: Opportunities and Challenges for Compiling SDG Indicators', *Global Policy*, 10(1)
- Mathiesen, K. (2014) 'Human Rights for the Digital Age', *J. Mass Media Ethics* 29, pp. 2–18.
- Mayer-Schönberger, V. and Ramge, T. (2018) *Reinventing Capitalism In The Age of Big Data*, Basic Books, New York.
- Mejias, U.A. (2019) 'Why the Global South should nationalise its data' *Aljazeera*, 14 December. Available at: <https://www.aljazeera.com/opinions/2019/12/14/why-the-global-south-should-nationalise-its-data> (Accessed 01 November 2021)
- Milano, M., O'Sullivan, B., and Gavanelli, M. (2014) 'Sustainable Policy Making: A Strategic Challenge for Artificial Intelligence'. *AI Magazine*, 35(3), pp. 22-35.
- Mitchell, M., et al. (2020) *Diversity and Inclusion Metrics in Subset Selection*, AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, February 2020, pp. 117–123
- Ntoutsis, E., et al. (2019) 'Bias in data-driven artificial intelligence systems - an introductory survey', *WIRES: Data Mining and Knowledge Discovery*, 10.
- O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Books, New York.
- Pentland, A.; Hardjono, T. (2021a) 'Data cooperatives', in Pentland, A., Lipton, A., and Hardjono, T. (eds.) *Building the New Economy: Data as Capital*, MIT: Cambridge, MA, USA, 2021.
- Ryan, M., Brooks, L., Antoniou, J., Macnish, K., Stahl, B., and Jiya, T. (2019) *Technofixing the Future: Ethical Side Effects of Using AI and Big Data to meet the SDGs*, 5th IEEE Smart World Congress 2019 - De Montfort University, Leicester, United Kingdom
- Sadowski, J. (2019) 'When data is capital: Datafication, accumulation, and extraction', *Big Data & Society*, 1-12
- Sadowski, J. (2020a) 'The Internet of Landlords: Digital Platforms and New Mechanisms of Rentier Capitalism', *Antipode*, 52(2) pp. 562-580
- Sadowski, J. (2020b) *Too Smart: How digital capitalism is extracting data, controlling our lives, and taking over the world*, The MIT Press, Cambridge.
- Sadowski, J. and Pasquale, F.A. (2015) 'The Spectrum of Control: A Social Theory of the Smart City', University of Maryland Legal Studies Research Paper No. 2015-26, *First Monday*, 20 (7)
- Schildt, H. (2020) *The Data Imperative: how Digitalization is Reshaping Management, Organising, and Work*. Oxford University Press: Oxford.
- Schmidt-Traub, G., Kroll, C., Teksoz, K. and Sachs, J.D. (2017) 'National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards' *Nature Geosci*, 10, pp. 547–555
- Scholz, T. (2016) *Platform Cooperativism. Challenging the Corporate Sharing Economy*, Rosa Luxemburg Stiftung: New York, NY, USA, 2016
- Seele, P. and Lock, I. (2017) 'The game-changing potential of digitalization for sustainability: possibilities, perils, and pathways', *Sustainability Science*, 12, pp. 183-185
- Shiffman, J. and Shawar, Y.R. (2020) 'Strengthening accountability of the global health metrics enterprise', *The Lancet*, 395(10234), pp. 1452-1456
- Soh, C., Connolly, D. and Nam, S. (2018) *Time for a Fourth Generation of Human Rights?*, United Nations Research Institute for Social Development, United Nations, Geneva.
- Stuart, E., Samman, E., Avis, W. and Berliner, T. (2015) *The data revolution: finding the missing millions*, ODI Development Progress, April 2015, Research Report, Overseas Development Institute, UK
- Stürmer, M., Nussbaumer, J. and Stöckli, P. (2021) *Security implications of digitalization: The dangers of data colonialism and the way towards sustainable and sovereign management of environmental data*, Report for the Federal Department of Foreign Affairs FDFA, Universtat Berlin.

Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, May 2022

- Taylor, L. (2017) 'What is data justice? The case for connecting digital rights and freedoms globally', *Big Data & Society*, 1-14.
- The Danish Institute for Human Rights (2021) *Lessons Learned Report: Advancing Inclusive SDG Data Partnerships*, July 2021, International Civil Society Centre, Germany.
- TRENDS (2021a) *About*. TRENDS - UN Sustainable Development Solutions Network, Available at: <https://www.sdsntrends.org/about> (Accessed 09 October 2021)
- TRENDS (2021b) *Big Data and the Sustainable Development Goals: Innovations and Partnerships to Support National Monitoring and Reporting*, July 2021, TRENDS - UN Sustainable Development Solutions Network. Available at: <https://static1.squarespace.com/static/6075cea4d9ce6d4c74e9b489/t/60ecae35fb12492bbb578ccd/1626123836673/Big+Data-report+FINAL.pdf> (Accessed 09 October 2021)
- United Nations (2015) *Transforming our World: The 2030 Agenda for Sustainable Development*, A/RES/70/1. Available at: <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication> (Accessed 09 October 2021)
- United Nations Global Pulse (2021a) *Global Pulse What We Do*. Available at: <https://www.unglobalpulse.org/what-we-do/> (Accessed 09 October 2021)
- United Nations Global Pulse (2021b) *Global Pulse Projects*. Available at: <https://www.unglobalpulse.org/projects/> (Accessed 09 October 2021)
- USAID (2021) *Digital Strategy 2020-2024*, USAID. Available at: <https://www.usaid.gov/usaid-digital-strategy> (Accessed 09 October 2021)
- Vinuesa, R., et al. (2020) 'The role of artificial intelligence in achieving the Sustainable Development Goals', *Nature Communications*, 11(233)
- Viswanathan, S., Karthykeyan, D. and Williams, W. (2021) *Barriers to Data Use in Sustainable Development*, Ahtena Infonomics, Data Values Project, Global Partnership for Sustainable Development Data.
- Wood, D.M. and Monahan, T. (2019) Platform Surveillance, *Surveillance & Society*, 17(1-2)
- Wu, B., Tian, F., Zhang, M., Zeng, H. and Zeng, Y. (2020) 'Cloud services with big data provide a solution for monitoring and tracking sustainable development goals', *Geography and Sustainability*, 1, pp. 25-32
- Xu, Z., et al. (2020) 'Assessing progress towards sustainable development over space and time', *Nature*, 577, pp. 74-78
- Zuboff, S (2019). *The age of surveillance capitalism: the fight for a human future at the new frontier of power*, PublicAffairs, New York