Jeremy Pitt

Being Human in the Days/Daze of Big Data

Irresponsible Design is Responsible for Diminishing Humanity

ISTOCK/GREMLIN

A recent Cambridge University Student Union Debate considered the motion that "This House believes that Big Data Destroys What It Means To Be Human."[1]

While the wording of the motion is, perhaps, intentionally provocative for the sake of the debate, in order to take a position on this proposition, one approach might be to take it at face value and unpack, to some extent, what is actually identified by first, the subject, "Big Data," then the object, "what it means to be human," and finally the action "destroys."

Therefore, beginning with the subject, "Big Data," it is, of course, clear that one cannot take a literal interpretation and the assignment of a "will to action" to Big Data and its ability to pose an existential threat to humanity. The sensing, aggregation, and storage of billions of people's personal data, cannot destroy what it means to be human on its own.[2] Like any technology, there is no intrinsic good or evil, only a purpose to which it is put; and that purpose may be pro-social or anti-social. Technology *per se* is neither good, nor bad. Design, on the other hand — design is never neutral; while pulling the strings backstage can be a more efficient and more effective means of achieving some ends, from executive aggrandizement in the name of democracy but with the intent to subvert democracy (2), or using the mask of convenience and symbiotic benefit to create social and mental entropy (3). Technology for Big Data (BD), and its brother-in-arms Machine Learning (ML), is at the root of, and is the facilitator of, deliberate string-pulling design choices (4), (5). These design choices are made by people, and so

---

[1] Editor's note – Full disclosure: I participated in the debate, and the entire proceedings can be viewed at https://www.youtube.com/watch?v=b3UwIiu6xaA&feature=share. My opening contribution that proposed the motion covered the first half of this article, but I ran out of time to deliver the second half (which is ironic, given the content/context). A balanced perspective on Big Data can be found in (1) (one of whose co-authors spoke against the motion).

[2] That's not to say that the image of a warehouse-sized computer or "The Cloud" becoming sentient, mobile, and hostile and taking out a few capital cities, *Cloverfield*-style, isn't completely unappealing to the wilder margins of one's imagination. But this scenario is *strictly* the preserve science fiction or media hype, is very different from reality, and serves as an artful distraction. In practice, the threat lies elsewhere, is altogether more complex, and is much more insidious – the distraction stops you thinking about it, while the threat is *precisely* that distractive technology is *designed* to stop you thinking about it, or indeed anything else for that matter. But this is getting ahead of the argument.

---

the question actually becomes, do the design choices enabled by Big Data and Machine Learning (BD/ML), have the capacity to alter, diminish and, in the limit perhaps, actually "destroy" what it means to be fundamentally human.

To understand and recognize the scale of action facilitated by BD/ML, it can be observed that there have been, in recent years, a number of accelerating and converging developments in computer hardware, software, and network connectivity, but also in social acceptability. The net result has been firstly, that if the body does it, then someone almost certainly has developed a device or a sensor to measure it;

Technological design choices can affect societal relationships within a community.

and secondly, if the device to measure it is manufacturable, someone has surely minitiaturized that device and shoehorned it into a wearable, portable or implantable device [6].

Thus we can see daily activity and experience trackers, biomarker and bioactivity trackers, bio-enhancement devices, smart clothing, especially for runners, a whole host of devices for sexual health, institutional devices like ankle monitors and employee trackers, smart toilets, and more besides.

These devices generate vast amounts of real-time data, which in itself would not be so much of a problem but for three factors. The first is the willingness of people to carry those devices voluntarily. Secondly, there is the corresponding change in people's habits, lifestyle, and culture which ensures that they are permanently "on grid" (i.e., and with the social acceptance that — indeed, almost expectation that — people will be continually consulting these devices). And thirdly, there is the opportunity for personalization and identification from that data: given "enough" data from a large group, and a few data points of an individual, some relatively simple mathematics (such as logistic regression) can be used to make reliable predictions about behavior, preferences, political affiliation, and so forth.

Moreover, these devices are permanently connected to a global communication network to transport that personalized data. They are linked to a massive expansion of computing power and system architectures (from centralized warehouse-sized server farms to distributed peer-to-peer) to process that data on unprecedented scales. This has effectively made some relatively old algorithms and data structures designed for particular tasks computationally feasible: for example, maintaining consistency in a distributed database (e.g., distributed consensus technology like blockchain), or binary classification (e.g., neural networks, repackaged as "deep learning"). It has also made it possible for some old business models to masquerade as new ones: e.g., brokerage or middleman relabeled as the "platform revolution," or lotteries: taking a small amount of money from a lot of people equals a lot of money (i.e., how some cloud computing services operate).

The consequence is that BD/ML facilitates design choices that have the scope, the scale and the agency to bring about many states of affairs.

So then, if we turn from the subject (Big Data, or rather, the design choices facilitated by Big Data/Machine Learning), to the object, "what does it mean to be human" … well, what does it mean to human?

At which point, one asks a question that has, of course, occupied philosophers, psychologists, and other professions for some thousands of years. Clearly, this article is not going to provide a definitive answer, but instead draws attention to two dimensions of humanity: firstly, the authority and opportunity of a "demos" to make decisions and "get things done," and secondly, the relationship between the "demos" and an empowered elite.

Collectively, one of the great opportunities for what is being called the Digital Transformation is for the creation of a new generation of socio-technical systems. One can see potential applications of such systems everywhere: from the fair distribution of electrical power in decentralized community energy systems based on demand-side management; to systems for maintaining ambient social relations in shared living arrangements or workspaces; to peer production systems with local pooling of tools, machinery, and resources; and to intelligent transportation, including (genuine) sharing economy applications for ride-sharing and journey-as-a-service (and goodness knows we need to get away from the archaic car-ownership model; however watching innovation disappear behind a corporate firewall through the "privatization of invention" is hardly an improvement [7]).

But these systems face three critical problems:

■ firstly, managing sustainability of common-pool resources and

dealing with the existential threat of the climate emergency [8], [9]

- secondly, ensuring distributive justice and reducing inequality in an economy of scarcity; and
- thirdly, choosing a political regime that avoids tyranny and promotes civic participation, protects civic dignity and provides civic education [10].

This is the duality presented by BD/ML design choices. On the one hand, a responsible and ethical service provider can undoubtedly deliver a beneficial and superior quality of service to an end user; and indeed significant opportunities for self-governance and civic problem-solving. But on the other hand, if Nietzsche is right and the *will to power* is the main driving force in humans [11], it offers a previously unsurpassed financial, commercial and even political imperative to aggregate and arrogate power, because of the economic benefits that can be derived from the ownership of trans-national platforms that use processed and aggregated data. It gives asymmetric control, wealth, and influence to the platform owner over the data generators, providing opportunities for: surveillance that the secret police of former dictatorships could only dream about;[3] wealth extraction beyond the most avid dreams of the former colonialist; and the concentration of capital (financial, social, and political) in hands of very few, possibly beyond the reach of the rule of law.

In fact, the essential problem that regulators, citizens, and local communities urgently need to address

is this: The private ownership and control over the means of social coordination, knowledge management, peer production, and digital innovation, with little public oversight, accountability, and transparency, is leading to a global monopoly of just a few platforms dominating each sector of commerce and social life. This monopoly has produced an asymmetric distribution of benefits and the rise of "surveillance capitalism" [13]; the algorithmic reinforcement of confirmation bias and the polarization of public opinion based on false information and preferences rather than evidence and reason; a reduction in opportunities for successful collective action at scale; opportunities for unscrupulous manipulation in pursuit of a hidden agenda (for example, fertility tracker apps being bankrolled by anti-abortion organizations, who also provide links to counsellors who offer misleading or false information [14]); and the concentration of political influence in intermediaries often located beyond national governance, and the growth of unearned income by those same intermediaries (who effectively are neo-colonialists: they are often in the place but not of it, extracting wealth and value from local communities and national economies [15]).

However, if BD/ML design choices are not just side-effecting an entire *community's* capability of resolving public action problems, but it can also affect *social relationships within* that community. One of these relationships is that between an "elite" (those who have been empowered in some way, for example as elected representatives, or experts, or controllers of some social or commercial service)

and a "demos" (i.e., the ordinary citizens). In this relationship, both parties require the capability to appreciate and participate in *the art of rhetoric*: not just as the rhetor (the orator oneself, usually one of the elite) but also as an audience member, the target of rhetoric, i.e., the ordinary citizens [17].

The origins of the scientific study of rhetoric (as part of a larger science of politics) can be attributed to Aristotle [16]. Rhetoric was defined as the counterpoint to dialectic, whereby dialectic was well-suited as a means of philosophical debate and analysis, and rhetoric was well-suited as a means of practical debate. Moreover, it functioned as an important communication channel between the elite and the

> ## Logos is being enfeebled by undermining critical thinking, and the enforcement of rote learning and excessive testing.

"demos" in classical Athens [17]. In Aristotle's exposition, there are three elements of the rhetorical situation: speech, speaker, and audience; *and* three elements of an effective speech (reasoned argument, appeals to character, and appeals to emotion). These elements are identified as *logos*, located in the speech itself, and concerned with reason; *ethos*, the character/reputation of the speaker, the value (or disvalue) attributed to the speaker outside the speech; and *pathos*, located in the audience, being the emotive response to the speaker and the speech. Therefore rhetoric presents three means of persuasion

---

[3]Surveillance or other covert operations by a police agency against political opponents, dissidents, or organized dissent is not restricted to dictatorships and is also a function of quasi-democracies, for example, blacklisting construction workers, or infiltrating undercover police officers into legitimate protest groups (and then having children with the protesters).

that an orator can use to induce the judgement that s/he wants the audience to make: a recognition of reason (logos), a grounding in credibility (ethos), and an appeal to emotion (pathos).

If an orator is using logos, ethos, and pathos to persuade an audience to make a certain judgement, then each member of the audience should be able to "reverse engineer" logos, ethos, and pathos to determine whether or not to make that judgement. Critically, this means being able to distinguish between logical reasoning and fallacies or sophistry, to distinguish between a credible source and a snake-oil salesman or similar charlatan, and to distinguish between emotions being awakened and emotions being manipulated. But in all three cases, the ability of a "demos" (both separately and collectively) to exercise these capabilities with respect to an empowered elite (which as discussed above, is represented by the owners of the social media platforms using BD/ML) is being undermined. Reason (logos) is being devalued; and individual's character (ethos) is being detached from distributed social knowledge about norms and their actual behavior; and emotive responses (pathos) are being manipulated more effectively by using BD/ML to gain knowledge of the (individual or collective) audience and its history of reactions to specific sorts of appeals.

In more detail, logos is being enfeebled not just by the willful replacement of critical thinking in school and university curricula by rote learning and excessive testing and metrication [18], but also by the surreptitious encouragement to outsource critical thinking to digital personal assistants for the sake of convenience [3]. This might be a significant achievement if it led to the genuine augmenta-

tion of human intellect [19] with the DPA working solely on the behalf of the user, but this cannot be guaranteed.[4] Furthermore, there have been several studies that have demonstrated that excessive time online and Internet addiction rewires the brain and shrinks the hippocampus [20], impairing the function of memory. Combined with an acceleration of social life facilitated by technological development, where there is an increased expectation of instant gratification and immediate response, this is accompanied by a loss of time for reflection and refinement,[5] while the sometimes limited means of online expression further diminishes rather than enriches a shared experience.

Similarly, ethos is being undermined by replacing "traditional" markers of quality and value, with ostensibly reasonable but possibly unreliable alternatives: for example, appearing as the first hit or on the first page of an Internet search (which can be paid for or distorted, if the page-rank algorithm can be manipulated). Furthermore, the democratization of the Internet by providing universal access to large bodies of specialist knowledge can be beneficial in a number of ways, but the leveling out of expertise and the coarseness of quantitative ranking (as measured by likes, views, ratings, followers, etc.) can be used by the unscrupulous to lay claim to an authority that is unwarranted, and

to present specious arguments that are based on neither evidence nor merit but purport to be on a par with arguments that have both.[6] This has led to a seemingly acceptable denigration of expertise,[7] a devaluation of educational qualifications and education in general, and being an "online influencer" becoming an aspirational ambition for a career choice.

However, influence is not confined to a peer-peer relationship: while social influencers can often be no more than animated advertising hoardings or shop-window dummies, the centralized elite's use of BD/ML can reduce (peer) pathos to little more than a tropic response. Nowak's Regulatory Theory of Social Influence (RTSI) [22] states that the relationship between a source of social influence and a target of social influence is actually bi-directional: i.e., it is not just that a source focuses on a target to influence, but that a target actively seeks a source by whom to be influenced. Given this psychological predisposition, algorithms for predictive analytics and other types of "persuasive technology" are unsurprisingly effective in the creation of filter bubbles [23], while confirmation bias [24] ensures that in the longer term the relationship is reinforced and unquestioned, and ultimately attention is saturated by online addiction [25]. Coupled with the commonplace confusion between democracy and majoritarian tyranny, the intervention of analytics to customize political advertising can materially affect

---

[4]Robbins [3] recounts the anecdote of a young boy asking a digital personal assistant (DPA) "what's five minus three," even though he knew the arithmetic. If a six-year old becomes accustomed to defaulting to a DPA over questions of trivial arithmetic and being desensitized to the potential self-harm of doing so, there is a risk of growing up to be an eighteen-year old who asks a DPA how to vote — and could we be sure that the answer will be in the best interests and preferences of the individual, or the owners of the DPA?

[5]"I apologize for such a long letter — I didn't have time to write a short one." Attributed to Blaise Pascal and Mark Twain.

[6]As observed in [21]: "The role of the online access to false anti-vaccination information just cannot be understated in examining the rise and spread of the anti-vaccination movement," and this goes hand-in-hand with the diminution of logos and critical thinking.

[7]"I think the people in this country have had enough of experts," Michael Gove, UK politician during the Brexit referendum in 2016, see https://www.youtube.com/watch?v=GGgiGtJk7MA.

the outcome of binary-choice elections and referenda, where opinion is polarized and a relatively small shift can make a substantive difference in the outcome [26]. Critically, though, civic dignity is undermined if the manipulation of pathos effectively tricks the demos into taking decisions that are not for the common good [11].

Finally, then, what of the relationship between the subject, "Big Data," and the object, "what it means to be human." If "what it means to be human" includes the capability at scale to solve public action problems and sustain common-pool resources, to create a civilized society characterized by its tolerance of difference and dissent and support for the weakest, and to answer the political question "how might we live our lives together, better" [11]; and if "what it means to be human" includes the capability of individuals to follow logical reasoning (logos), to establish trustworthy credentials (ethos), and avoid emotional determinism (pathos); and since both of these capabilities are being adversely affected by design choices underpinning the use of BD/ML in the Digital Transformation: then one could reasonably reach the conclusion that unscrupulous use of BD/ML is at least diminishing, if not, in fact destroying, "what it means to be human."

Big Data won't destroy your humanity. Somebody abusing Big Data might just do that. The original proposition might not be true in a literal or definitional sense, but as a process not of individuation but of speciation, it might become true in its spiritual and metaphorical senses. In which case, in conclusion and in the best practice of pathos, the last word can be left to Science Fiction:

"In brief, our collective memories, the richest part of us, have been taken away, and we are poor indeed. In return for castles of the mind, our rulers have given us mud hovels palpable to the touch; a bad exchange for us. …

Generation of cows! Sheep! Pigs! We have not even the spirit of a goat! If Epaminondas was a man, if Achilles was a man, if Socrates was a man, then are we also men?" [27].

Author Information

Jeremy Pitt is Professor of Intelligent and Self-Organizing Systems at Imperial College London, U.K. Email: j.pitt@imperial.ac.uk.

References

[1] V. Mayer-Schönberger and K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think*. John Murray, 2013.
[2] D. Runciman. *How Democracy Ends*. Profile, 2018.
[3] J. Robbins, "If technology is a parasite masquerading as a symbiont, are we the host?," *IEEE Technology & Society Mag.*, this issue, pp. 24-33.
[4] N. Dow Schüll, *Addiction by Design: Machine Gambling in Las Vegas*. Princeton, NJ: Princeton Univ. Press, 2014.
[5] G. Lovink, *Sad by Design: On Platform Nihilism*. Pluto, 2019.
[6] N. Stevens and J. Wernimont, "Seeing 21st century data bleed through the 15th century Wound Man," *IEEE Technology & Society Mag.*, vol. 37, no. 4, pp. 46-54, 2018.
[7] J. Pitt, "Transparent ownership of mobility-as-a-service," *IEEE Technology & Society Mag.*, vol. 38, no. 2, pp. 5-8, 2019.
[8] E. Ostrom, *Governing the Commons*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[9] E. Ostrom, "A polycentric approach for coping with climate change," *Ann. Economics and Finance*, vol. 15, no. 1, pp. 97-134, 2014.
[10] C. Clark, *Wealth and Poverty: On the Social Creation of Scarcity*, J. Economic Issues, vol. 36, no. 2, pp. 415-421, 2002.
[11] J. Ober, *Demopolis: Democracy before Liberalism in Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
[12] F. Nietzsche, *Thus Spoke Zarathustra*, A. del Caro, trans., R. Pippin, Ed. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[13] S. Zuboff, "Big other: surveillance capitalism and the prospects of an information civilization," *J. Information Technology*, vol. 30, pp. 75–89, 2015.
[14] Q. Grundy, K. Chiu, F. Held, A. Continella, L. Bero, and R. Holz, "Data sharing practices of medicines related apps and the mobile ecosystem: Traffic, content, and network analysis," *British Medical J.*, vol. 364, p. 1920, 2019.
[15] M. Lofgren, "The revolt of the rich," *American Conservative*, Aug. 27, 2012; http://www.theamericanconservative.com/articles/revolt-of-the-rich/.
[16] Aristotle, *On Rhetoric: A Theory of Civic Discourse*, 2nd ed., G. Kennedy, trans. Oxford, U.K.: Oxford Univ. Press, 2007.
[17] J. Ober, *Mass and Elite in Democratic Athens: Rhetoric, Ideology, and the Power of the People*. Princeton, NJ: Princeton Univ. Press, 1989.
[18] J. Muller, *The Tyranny of Metrics*. Princeton, NJ: Princeton Univ. Press, 2018.
[19] A. Nowak, P. Lukowicz, and P. Horodecki, "Assessing artificial intelligence for humanity: Will AI be the our biggest ever advance ? Or the biggest threat?," *IEEE Technology and Society Mag.*, vol. 37, no. 4, pp. 26-34, 2018.
[20] D. Mosher, "High wired: Does addictive Internet use restructure the brain?," *Scientific American*, 2011.
[21] A. Hussain, S. Ali, M. Ahmed, and S. Hussain, The anti-vaccination movement: A regression in modern medicine, *Cureus*, vol. 10, no. 7, 2019.
[22] A. Nowak and R. Vallacher, Nonlinear societal change: The perspective of dynamical systems, *British J. Social Psychology*, vol. 58, no. 1 pp. 105-128, 2018.
[23] E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You*. Penguin, 2011.
[24] C. Tavris and E. Aronson, *Mistakes Were Made (But Not by Me): Why We Justify Foolish Beliefs, Bad Decisions, and Hurtful Acts*, rev. ed. Pinter and Martin, 2016.
[25] N. Eyal, *Hooked: How to Build Habit-Forming Products*, Penguin, 2014.
[26] C. Cadwalladr and E. Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach," *Guardian*, Mar. 17, 2018; https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence- us-election.
[27] R. Sheckley, *The Mnemone, in The Same to You Doubled*. Pan, 1972.