Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-29 (2016) 117–122

Data Driven Wave of Certainty- a question of ethical sustainability

Karamjit S Gill

University of Brighton, Brighton, UK (e-mail: kgillbton@yahoo.co.uk).

Abstract: The new wave of artificial super intelligence, exemplified by the evolution in big data and Internet of Things raise an issue of whether this evolution could and should exclude human intuition from human judgment, given its exclusive reliance on data driven certainty of algorithms, analytics, artificial and machine learning. How could we respond to the algorithmic wave of uneven accuracy, deal with uncertainty, ambiguity, conflicts, justifiable judgements, and seek alignment of data and knowledge. We are minded by Weizenbaum's argument (1976) that if human judgement is based on human observation (data) and intuition, then how can science which ultimately rests on a vast array of human value judgements, deny that human value judgements are illusionary**?** The paper reflects on the certainty of data driven singularity and argues for a relational conversation between technology and society**.**

Keywords: artificial intelligence, big data, existential risk, ethics, judgement, dualism, symbiosis

1. RIDING THE DATA DRIVEN WAVE

The new wave of artificial supper intelligence raises a number of serious societal concerns: What are the crises and shocks of the AI machine that will trigger fundamental change and how should we cope with the resulting transformation? What would the implication be if AI machine takes over and transforms the way we live and work? What would technology do to work, employment, economy, governance, state, democracy and professions? What would the social and political implications of employment be if people were replaced by the machine? What if the state disappears, leading to the disappearance of economy, professions, employment, politics as we know them? Can digital economy be regulated, measured, and controlled? Can the AI machine with its embedded machine learning algorithms be monitored and controlled? Would new politics emerge as another digital game, and what would the rules of this game be, and how would these rules change the playing field of the game of politics itself? Would the nasty form of exploitative individualism triumph or would new forms of digital collectives (e.g. consumer collectives) emerge that would be more powerful than corporations? And can humanity live in a simulated state of digital being? (Gill KS 2016)

The questions and issues cited above raise ethical issues and implications of automation (e.g machine learning) of aggregating, processing and manipulation of big data on human judgement and decision making and what are the ethical responsibilities of designers, coders and users of data, especially when data is essentially local, relational and contextual? For example, how can we incorporate the interplay of the ethical dimension of legal rules, social norms, cultural ethos in designing, for example, e-Health systems and how would we measure this dimension from social, technical and policy perspectives. Sha (2016) notes that data are not just pieces of nature lying around for data scientists to pick up like shells on a beach. Data are constructed via very elaborate complexes of theory, politics, judgment calls, plus apparatuses, devices, technologies and procedures that are themselves conditioned by theory, politics, and judgment calls. What is crucial is how data is constructed. The degree of contingency built into the very data as collected before they ever show up in a spreadsheet or database field is already profoundly intertwined with contingent factors like culture, prejudice. Even more important than the volume, velocity of Big Data is what is being done with that data. Moreover, it is not just an amount of data, but how much data is needed to make sense the situation. What matters is the relevant and right kind of data rather than the volume of data.

Pentland (2014) says that though big data gives us unprecedented scope to understand our society, "what works in theory may not translate well in the real world, where complex human interactions cannot always be captured, even by the most sophisticated models." While computational tools such as visualization tools and analytics can dramatically increase the speed of calculation, enhance transparency and information, "they are surprisingly limited when applied to solving society's problems. One reason is that such rich streams of data encourage spurious correlations." It is just not a question of collection, representation and calculation of a massive volume of data, it is also a question of measurement of all possible alternatives, of identifying all sensible and non-sensible results of calculation, and then asking how to computationally deal with "aggregates," of big data beyond the traditional techniques of "averages ", and how do we know that standard statistical tools do not generate nonsensical results. Moreover, although dense, continuous data and modern computation allow us to map many details about society, and to explain how communities might work, such raw mathematical models contain too many variables and complex relationships for most people to understand. He says that "If big data is to be deployed effectively, people must be

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.11.068

able to understand and interpret the relevant statistics." This requires some kind of dialogue between human intuition and the compelling reality of big data, and this in turn requires "a new understanding of human behavior and social dynamics that goes beyond traditional economic and political models. Only by developing the science and language of social physics will we be able to make a world of big data a world in which we want to live." Recognising that big data provides an opportunity to address the big global challenges of the interconnected world of data sharing and information flows, Pentland (ibid.) says that "Big data provides the clearest lens for examining how society functions in fine-grain detail". For example, big data can enable us "to look beyond aggregates (such as markets, classes and parties) and instead examine the fine-grain patterns of society, new opportunities and discoveries emerge." For example, tracking social phenomena down to the individual level and the social and economic connections among individuals. Lohr (2013) notes that the underlying assumption of tilting toward a data-driven society is that data will change attitudes and policy, combating bias and causing policy-making to be more of a science. In a Big Data world, the data-mining for patterns and insights to guide policy will be done automatically — by software algorithms. Of course, algorithms are created by people and they contain inferences and assumptions coded in. Those coded-in values shape the output — computer-generated predictions, recommendations and simulations. That raises questions about the human design and control of the computerized helpers in policy-making, as in other realms of decision-making. Lohr (ibid.) quotes John Henry Clippinger, chief executive of the Institute for Data Driven Design: "At some point, you're in the hands of the algorithm," and further: "You're whistling in the dark if you don't think that day is coming."

Knight (2016) says that since the nature of the progress of artificial general intelligence is unpredictable, there is a need to undertake proactive policy measures and a regulatory framework to mitigate the risks, even if no such breakthroughs currently appear imminent. It is noted that Bostrom's study of "existential risk" (Future of Humanity Institute, 2013, Knight ibid.) argues that artificial intelligence might be the most apocalyptic technology of all. With intellectual powers beyond human comprehension, Bostrom expounds that, self-improving artificial intelligences could effortlessly enslave or destroy Homo sapiens if they so wished. While he expresses skepticism that such machines can be controlled, Bostrom claims that if we program the right "human-friendly" values into them, they will continue to uphold these virtues, no matter how powerful the machines become. Commenting on rhetoric of deep learning, Knight further notes that at the core of deep learning lies the limit of artificial neural networks, in the sense that these artificial neural networks can't compute at the speed or accuracy that our brains do. Thus one of machine learning's most intractable limits is the development of artificial neurons that can function at an accelerated rate.

Commenting on Bostrom's study on existential risk, Geist (2015) says that whilst recognising the limit of the super intelligence machine, AI-enhanced technologies might still be extremely dangerous due to their potential for amplifying human stupidity. So far as the existential risk is concerned, just by enhancing the familiar 20th century technologies, AIs of the future can endanger the future survival of existing societal structures by undermining their precarious strategic balances, for example by making the existing technologies much faster, cheaper and deadlier. If anything, Geist says that machines capable of conceiving and actualizing elaborate plans but lacking self-awareness could be far more dangerous than mechanical analogues of human minds. Baum et. al. (2015) note that for the bulk of the catastrophic threats the literature has thus far focused mainly on philosophical aspects, in particular the moral significance of catastrophic threats and challenges to quantifying their probability, as well as empirical aspects, in particular the nature and size of the various threats. Although there is considerable research into specific threats such as global warming and nuclear war, there is rather a lack of much needed research into existential risk. They further note that catastrophic threats are not merely academic—they actually do threaten humanity, and so for the sake of humanity they should be confronted. For example, there is a need for the "better development of Quantum-safe encryption and its wider deployment to avoid spying on citizens, corporations, and countries, potentially enabling catastrophic totalitarianism and economic chaos."

On the threat from advanced artificial intelligence, Baum et. al. (2015.) note the potential for global government ("singleton"), and for the possibility that humanity exists within a computer simulation. Just as seeking generalised computational solutions to problems of existential risk may be tempting for machine learning ideologues, so is the idea of humanity living in simulations a computational fancy. Baggini (2016) notes the 'fascination' of simulations, of living in virtual world, the thrill in uploading oneself in a virtual world, and living a simulated experience in computer-based virtual reality, a computer game, a thrill of nothingness nonetheless; nothing changes in the real world but fascination remains with creating a perfect simulacrum of a world, even a virtual one. So one wonders what has fundamentally changed from the early days of Eliza (Weizebaum 1976) apart from playing with vast amounts of data, visualisation tools and data analytics. He further says that, "Desire for a virtual life grows from dissatisfaction with the real one we have. This desire can surely only increase, the more disconnected we become from the natural cycle of life and death, and the less able to accept it." So it's an escape from the real world, its complex human problems and challenges of health, hunger, poverty, water, climate, as well as obligations of personal responsibility, let alone social responsibility. We need to be mindful of the differential technological developments in which safe AI technologies are favoured over dangerous ones. Baum et.al. present a practical perspective on the ethics of catastrophic risk. They articulate that the standard ethical argument for confronting catastrophic threats to humanity is based on the far-future benefits of confronting the threats. They posit that those who do not yet share the argument on existential risk, may contribute to long term research by focusing on 'near-future benefits from confronting near future threats, as well "mainstreaming" actions on the threats into existing activities. They survey the threats, finding that