Received 14 June 2019
Revised 16 July 2019
29 July 2019
Accepted 31 July 2019

Bubble to panopticon: dark undercurrents of the big data torrent

Mathew Maavak
Big Data Excellence Centre, Kazimieras Simonavicius University, Vilnius, Lithuania

Abstract

Purpose – Big data are indispensable in scientific endeavours ranging from nuclear research to climate studies. However, there is a growing misperception that congeries of data can be easily reconstructed into competitive business insights. Such notions have been encouraged by a plethora of mainstream techno-utopian forecasts.

Design/methodology/approach – This paper investigated such claims and related big data developments, including its "systems-first" and oligopolistic orientations. Due to the paucity of current scholarship on an admittedly pessimistic topic, the paper studied contrarian developments in the industry by relying on secondary data. The study of experts and scholars; industrial trends; and discrepancies and critical gaps in the mainstream data narrative were sourced to prognosticate the likely trajectory of many data giants.

Findings – A key finding was that the big data industry faces an untenable market bubble worth trillions of dollars. This will have severe consequences for common digital access and social stability worldwide. Evidence presented also suggests that the data industrial complex may undergo a function creep by facilitating a transition from surveillance capitalism to surveillance society.

Research limitations/implications – Primary data for a study of this nature may take years to materialize. This is a "first-pass" study that seeks to illuminate latent dangers facing the big data/AI sector. There is a paucity of scholarly study that even remotely touches on this topic. Therefore, supporting arguments was sourced from contemporary reports and expert study (secondary data).

Practical implications – As control of data may have geostrategic implications, balkanization of the wired ecosystem may be underway with Russia and China leading the way. Future superpowers may be defined by the way they handle data. The concentration of data in fewer hands may also affect citizen innovation.

Social implications – A break-down of the data industrial complex may lead to social mayhem as the monetization of presently free software, blogs and social media platforms may be unfeasible.

Originality/value – This topic has hardly been explored due to the novelty of big data, its applications and the daily hype over its potentials. This paper boldly describes dark countercurrents in the industry.

Keywords Data monopolies, BDA, Artificial intelligence, Google, Search engines, Market bubbles, Panopticon, Foresight, Internet, Information systems, Algorithms, World wide web

Paper type Viewpoint

Introduction

Big data has apparently crossed the planetary limits of its "new oil" tag to become the stellar new sunlight – at least according to Alphabet (Google's parent company) chief financial officer Ruth Porat (Ghosh and Kanter, 2019). As a marketing slogan, it is, indeed, catchy, particularly when articulated at a platform such as the World Economic Forum. Buoyed by such boundless vistas, data giants have effectively delivered an ultimatum to enterprises: embrace big data or perish – in a way reminiscent of the "buy now or be priced out forever" subprime frenzy a decade back.

Kybernetes
Vol. 49 No. 3, 2020
pp. 1046-1060

0368-492X
DOI 10.1108/K-06-2019-0403

At the cursory level, there is enough "big data" to back this optimism. According to the International Data Corporation (IDC), the global data volume will surge from 4.4 zettabytes to 44 zettabytes between 2013 and 2020 (Hajirahimova and Aliyeva, 2017). By 2025, there might be 163 zettabytes of data worldwide (Reinsel *et al.*, 2017). Gartner estimates that the number of connected devices within the emerging internet of things (IoT) will rise from 8.4 billion units in 2017 to 20.4 billion by 2020 (Tung, 2017).

The IDC defines big data and its analytics as "a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery and/or analysis" (Gantz and Reinsel, 2011, p. 6). It represents a novel approach to finding actionable intelligence via the study of varied phenomena and its associations (Cukier and Mayer-Schoenberger, 2013). However, these relatively atheoretical viewpoints are based on the assumption that more data means better inquiry; one where data can provide hypotheses and insights that are superior to traditional forms of discovery such as theory and experimentation (Symons and Alvarado, 2016; Kelling *et al.*, 2009; Kitchin, 2014).

While big data are, indeed, indispensable to scientific endeavours involving prodigious amounts of information, the idea that congeries of data can be digitally reconstructed into competitive insights through specialist software remains a distant promise. As Rosemary Barnett (2017) pointed out in a TechCrunch article:

> Search online for a real-world story of how big data analytics produced a piece of "hidden" or "unexpected" intelligence, based upon which the business took action, with quantifiable commercial results (preferably expressed in one of the major world currencies). You might just detect a conspicuous absence of concrete case studies to validate this "data-insight-action-value" chain as a concept.

Other cautionary observations border on the ominous: "what no one seems to be saying – or recognizing – is that the death of big data is imminent. Big data's demise is only years, maybe even months away, and yes, we will need a big casket" (Phelan, 2012). Although such pessimism has turned out to be rather stillborn, big data certainly faces some formidable countercurrents as outlined in the following sections.

This paper adopted elements from Latour's (2005) oligopticon for its theoretical approach: by focussing on structures, trends or developments that may appear disjointed on their own but which attain cogency when combined as a narrative to produce discovery and knowledge. This contrasts somewhat with the panoramic "big picture" contemporary treatment of the subject.

Oligopolistic "new oil"

Similar to ideologies, beliefs and other conduits of behavioural organization, data tends to shape an oligopolistic worldview (Taylor and Wagman, 2014) that reflects a contemporaneous fractionation of state and corporate power. The control of data translates to an ability to control the destiny of nations. Endowed with such leverage, the common digital domain has become a hotly-contested battlespace for a variety of state actors and special interests. Google plays the kingmaker role here by virtue of its dominance in the search engine optimization field. Former Harvard University academic Benjamin Edelman (2011) noted that while Google claims its search results are "algorithmically-generated", "objective" and "never manipulated"; other search engines rarely or never make such claims.

In fact, there is a growing corpus of reports and scholarly literature on the oligopolistic orientation of Google search. The author himself can be proffered as an Exhibit A in the case

against Google's algorithmic bias. Whenever his name was keyed into Google search over the past few years, it had inexplicably prioritized a ghost page from *The Edge* – a Malaysian-Singaporean publication, which had featured no quote, reference or article from the author; just a hanging byline. How could a phantom page – out of 16,100 search hits at the time of writing – be prioritized over more prestigious publications and institutions that had regularly featured the author's works? Entreaties to remove the ghost link were ignored by both the daily and Google, leaving data scientists consulted on the matter to wonder whether there was a clickbait arrangement between the search giant and a plethora of well-funded media worldwide. A parallel search in Yandex and Baidu incidentally returned results that omitted the phantom page (Maavak, 2019a).

With more than 1.94 billion websites on the internet as of January 2019 (Liedke, 2019), search engines have become the ubiquitous gatekeepers of information in the open-source realm. According to Edelman (2011, p. 18), Google's dominant market share meant "that any bias at Google has a much larger impact than bias at another search engine. In short, Google alone has the power to make or break a site – and Google alone has promised not to abuse that power".

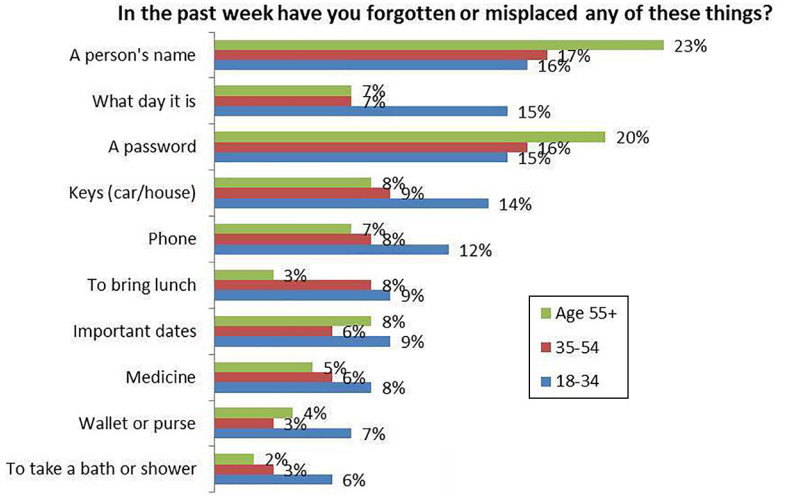
Systems-first panacea for data smog

While Google's avowed neutrality remains an open question, data generated by a bewildering array of devices is outpacing our collective capacity to process them. Enter dig data analytics (BDA) and its alphabet soup of tools, techniques and software as the panacea. BDA, when designed and executed properly, may greatly facilitate information processing in an increasingly complex world. Next-generation weather forecasting models, for example, will crunch exascale-level data into more accurate forecasts. Billions can be saved in the process, in areas ranging from precision farming to open-air sports.

BDA may also alleviate users from information overload and associated malaise such as "data smog" (Shenk, 1997), "infobesity" (Puryear and Root, 2013), "infoxication" (Chamorro-Premuzic, 2014) or information fatigue syndrome (Waddington, 1996). Without reducing data into actionable meaning or utility, individuals may become prone to anxiety, poor decision making and productivity, mnemonic difficulties and strained relationships (Heylighen, 2002; Waddington, 1996; Wurman, 1989). However, these terms and symptoms now seem so *fin de millénaire* amongst generation Google. A recent trending machine poll in the USA revealed how millennials were becoming more forgetful than their seniors (Business Wire, 2013). Scientists increasingly suspect that digital overexposure is gradually atrophying the brain's memory-related temporal lobe in favour of the frontal lobe, the area of synthesis (Mosher, 2011). Figure 1 encapsulates this problem.

Digital-savvy millennials know where and how to retrieve information but lack the impetus for memorization and contemplation, which, in turn, affects their ability to see through data smog. As a hyper-wired IoT emerges, our reading habits are getting poorer; our sentences are getting shorter, and our ideas are parodying the latest mediated fad. We function as real-time individuals in a real-time society. Long-term imperatives are routinely sacrificed to the trivialities of presentism (Rushkoff, 2013).

The *homo digitalis* now only needs an AI overlord to manage his or her "smart life". Technoprophets like Dr. Angel Iscovich are already seeking higher machine forms to organize human life in an increasingly complex world. According to Iscovich, artificial intelligence (AI), which processes big data, may "soon tell us how to live" as "computers can do our thinking for us" (Ashley, 2019). Self-driving vehicle pioneer Anthony Levandowski has even decreed a "way of the future" (WOFT) religion based on an "AI godhead". Despite its stark departure from scientific norms, an AI religion is nonetheless gaining acceptability,



Note: (Graphic: Business Wire, 2013)
Source: The Trending Machine national poll of 800 adults 18+, representative of the US population of adults online

Figure 1.
Age-based breakdown of mundane forgetfulness

at least judging by the amount of print and bytes devoted to the subject. Even a sample verse has been machine-generated, mimicking archaic versions of the King James Bible: "and let thy companies deliver thee; but will with mine own arm save them: even unto this land, from the kingdom of heaven" (Brandon, 2017).

Such empyrean outlooks nonetheless raise some existential questions: Was man made for data or data for man? Would AI, upon attaining super-intelligence, continue to serve humans who had programmed them? While such thinking may invoke apocalyptic scenes from the *terminator* franchise, it is reminiscent of the unfulfilled *idée fixe* of the second industrial revolution: "in the past, the man has been first; in the future, the system must be first" (Taylor, 1911).

Underbelly of data mining

Is big data the new oil as claimed? The global digital ecosystem currently consumes 7 per cent of the world's electricity. This figure will likely rise to 12 per cent by 2020 and is projected to grow annually at an average rate of 7 per cent by 2030. The "average car driven for 1 km produces as many greenhouse gases as a thousand Google searches" while global Bitcoin mining may be "consuming more electricity than 159 countries, including Ireland, Bahrain and the Slovak Republic" (Naughton, 2017).

While data giants consume energy at a voracious pace, is it generating an equivalent value or more in terms of competitive insights, products and services? For the sake of comparison, the crude oil industry generates myriad indispensable products ranging from plastics, specialist chemicals to jet fuel. What does AI extract from the growing zetabytes of irrelevant or toxic information? While the social media may attract billions in terms of users and advertisement revenue, it is also numbing societies to critical long-term issues. Policymakers, for example, are instinctively reacting to senseless Twitters and online trolling instead of tackling the complex task of governance. Few think beyond the next

electoral cycle and policy paralysis is the result. Mass-mediated viral distractions are also impeding the collective focus needed in a creator economy.

Instead of representing the new oil, big data are more comparable to the traditional diamond or gold extraction process. In fact, there are notable parallels between the current data mining fad and the California gold rush of 1848-1855. The latter was marked by frenzied migrations, xenophobia, febrile speculations and the decimation of 100,000 Native Americans in less than 20 years. In the extortionate peak year of 1849 alone, author Steve Boggan (2015) noted that those who "mined the miners" were the ones who had profited the most:

> There were reports of canteens charging a dollar for a slice of bread or two if it was buttered, the equivalent of $56. A dozen eggs might cost you $90 at today's prices; a pick axe would be the equivalent of $1,500; a pound of coffee $1,200 and a pair of boots as much as $3,000 when today you could get a decent pair for around $120.

It is quite telling that the most enduring product from the California gold rush was arguably Levi's jeans; not a native jewellery franchise as one would have logically expected. Will a similar unanticipated winner emerge from the current data rush?

For the time being, data miners in California's Silicon Valley appear better-off than their gold rush predecessors. The Bay Area, which includes San Francisco and Silicon Valley, accounted for nearly 45 per cent of venture capital investments in the USA, totalling more than US$40bn in 2017 alone (Florida, 2018). Yet, this data industrial complex is propped up by an army of offshore programmers who belabour under low wages, poor working conditions and routinely unpaid overtime (Hao, 2019). The low-wage gig economy is one consequence of the current data hype (Gray and Suri, 2019). Recent tragedies involving Boeing 737 Max jets revealed how software components were outsourced to Indian programmers who lacked experience in the aerospace industry, and who were paid as little as US$9 per hour (Robison, 2019).

BDA and AI cannot be detached from actual field experience, evidentiary guidelines and purpose-driven design. The US legal system's use of predictive algorithms, for example, has seen black offenders receiving higher risk scores and concomitant penalties than what demographic evidences suggest (Larson *et al.*, 2016).

Bubbles and phantasms

Much like bubbles created in other industries, the big data industry cannot escape the balloon effect; one where a compressed area ends ups inflating another. Rapid expansions in the industry have similarly aggravated and displaced poverty in California to the doorsteps of tech giants. The streets of San Francisco are now paved with the homeless, human waste and discarded needles from drug abuse. Even a new urban art genre has emerged in the form of poop graffiti (Bendix, 2018). Big data giants, despite market valuations collectively worth trillions of dollars, ironically have little time to crowdsource clean-up campaigns even while they promote human rights, social justice and inclusivity causes abroad.

This lack of empathy is partially attributable to the just-in-time (JIT) algorithmic *modi vivendi* perfected by Big Tech. Ironies abound here. While supply chain gurus regularly unload a logorrhoea of buzzwords over the data-driven marvels of their systems – including how they prevent slave or child labour in the Third World – poorly-paid US poultry workers were resorting to wearing diapers on the job (Singh, 2016). Amazon warehouse and delivery workers fared little better as they sometimes had to urinate into bottles (Liao, 2018) to fulfil their JIT targets. Perhaps, Amazon's book repositories were exhaustively data-mined to

glean Brer Rabbit's parting advice to Brer Fox: "like the saying goes, some folks go up, and some go down! you should make it to the bottom all safe and sound" (Harris, 1881).

The bottom is bubbling with many ominous trends. Decades-long provision of free emails, blogs and social media apps by tech titans have generated a Golconda of personal data. These are marketed as consumer prediction products in what Zuboff (2019) calls "surveillance capitalism". Current user behaviour already activates ubiquitous "smart ads" on screens as well as unsolicited emails. We are led to believe that an AI-led matrix will soon acquire the prescience to anticipate our necessities in advance. However, for the time being, MS Word remains clueless whenever users fumble with less-used forms of text formatting. Surely, Microsoft has two decades of customer big data to meet comprehensive user needs? This begets the question: what exactly is the denouement of the current data ebullience?

In an email correspondence with the author, media theorist Douglas Rushkoff observed that "nearly every Silicon Valley startup and behemoth cites "big data" as its exit strategy. No matter how the purported business is actually doing, it is collecting tons of data and arguing that this data will be valuable enough to justify the company's valuation."

The market valuations of Amazon, Alphabet and Microsoft are projected to surpass US $1tn each in the near future, barring fallouts from a global economic slowdown. Facebook is not far behind in the trillion-dollar race. Digital ad sales are expected to grow from US $251bn in 2018, representing 45 per cent of current global ad revenue, to US$348bn in 2022 (Sweeney, 2018; Levy, 2019). Yet, according to Wikibon, the global big data market will only be worth a paltry US$103bn in 2027 (Columbus, 2018). The trillion-dollar discrepancy between the market valuations of data giants and big data's projected market size is indicative of a bubble build-up:

> "Big data are really just a subset of market research," added Rushkoff. "They use it to make creative decisions about the future. Past data are used to bet on the future, rather than using creativity or intention to create the future. They (eventually) lose any sort of competitive advantage because they are only as good as the data merchant they have hired. If a company's innovation is only as good as the data it is buying, then it is just an extension of the data company."

Renowned economist and forecaster Nouriel Roubini is equally scathing of a parallel development in the big data revolution. Blockchain, according to Roubini (2018), "has been heralded as a potential panacea for everything from poverty and famine to cancer. In fact, it is the most overhyped – and least useful – technology in human history. In practice, blockchain is nothing more than a glorified spreadsheet."

Such dire forecasts have not stalled the chain-ganging of bubble carts along the data analytics train. Betting on that ride may eventually backfire if recent history, such as the 2008 Great Recession, is any indicator. Alarming economic signals were ignored back then, and they continue to be ignored today. A decade later, the global asset bubble has ballooned to US$100tn, amounting to 1.3 times the global GDP. The only question now is when, and how bad the next economic meltdown will be (Xie, 2018). Despite the immense data they possessed, it was not Google, Facebook, Microsoft or Amazon that flagged up the harbingers of the Great Recession. Instead, they persist in invoking techno-utopian digital nostrums ranging from AI-mediated match-making to telepathic communications. One eye-popping poser from Facebook (Ribeiro, 2017) even went like this: So what if you could type directly from your brain?

An intelligently-designed algorithm would have retorted: "That is what hands are for, are they not"? Voice-to-text software also does a decent job. Limbic and oral limitations traditionally forced the mind to be focused and diplomatic, paralleling the cost- and space-constrained information selection process associated with traditional print

(Heylighen, 2002). The brain is a neural melange of impolite intents, jumbled imageries and tacit knowledge.

Tacit knowledge refers to hidden or procedural intelligence (i.e. reflexive skills), ideas and experiences that even experts have difficulty in accessing or codifying. Its extraction, therefore, is considered as the Holy Grail of knowledge conversion. For the time being, it can only be partially drawn out through a gradual process of peer interaction and logical questioning (Polanyi, 1967; Chugh, 2015; Nonaka and Takeuchi, 1995; Child, 1997; Kingston, 2012; Jarche, 2010). AI has not reached a stage where it can extract or discern insights from an individual's tacit realm. It will likely never be able to do so. Behavioural recognition and predictive software only deal with surface signals that are equally discernible to sentient or specially-trained professionals.

Facebook founder Mark Zuckerberg, however, sees the brain as a one-dimensional exit point for profitable data: "our brains produce enough data to stream four HD movies every second. The problem is that the best way we have to get information out into the world – speech – can only transmit about the same amount of data as a 1980s modem. We are working on a (wearable technology) system that will let you type straight from your brain about 5× faster than you can type on your phone today" (Ribeiro, 2017).

One can only imagine the plots and pixels of that high definition movie in a classroom full of hormonally-charged teenagers. Welcome to the age of real-time DIY porn. Then again, the concept may not be a bad idea. Think of the instant pay-per-view slasher flicks that could be generated during the next US presidential debates? To avoid such contretemps, Zuckerberg should conjure up some exogenous mind-control matrix. And here is where big data takes an ominous function creep.

Rise of the surveillance society

What happens when the gravy train of data-driven stock evaluations run off an economic cliff? A Great Depression 2.0 would push the ad-driven mainstream news machinery – a train-wreck in its own right – further into the clutches of Big Tech; thereby consummating the ongoing amalgamation of the techno-media compact. The traditional media is already facing extinction as advertisers and readers gravitate towards digital platforms.

Once the ongoing economic slowdown accelerates, free software services may be rendered inoperable as they can no longer be financially sustained by advertisements from bankrupting industries. Imagine waking up to an economically-depressed world where access to your email account would suddenly be contingent on a periodic fee? However, your credit card has maxed out; and your creditworthiness has been algorithmically junked by the very tech titans who had provided those free services in the first place. Even Sisyphus may have had it easier!

In the aftermath of a market bloodbath, tech giants have a few options but none of them appear good in the long run. Alphabet, for example, may choose to monetize its two billion Android users. This option was in the offing for some time but hardly anyone took notice when then Google CEO Eric Schmidt tinkered with the idea in 2010: "all it would take is $10 per user per year" (Langlois, 2019). However, US$10 today may translate to US$1000 or much more in a hyperinflationary world. Think of how much bread, butter and eggs had effectively cost during the California gold rush or the mountains of German marks carted around to buy these items in the 1920s? Cyber-coolies subsisting on US$9 per hour would be hard-pressed to maintain enterprise status quo. Furthermore, contemporary Android devices include a larger number of apps and components from diverse sources. $10 per device cannot subsidize all of them.

There are many other imponderables on the horizon, reflecting the entwinement of substrate-level global risks (Maavak, 2019b). If present tensions in the Straits of Hormuz –

the global oil and gas chokepoint – morphs into an outright conflict, a barrel of crude may fetch as much as US$300 or more (RT, 2017). This will rapidly impoverish the world and consign hundreds of millions of people, if not billions, to the prospect of being disconnected, disenfranchised and disempowered. Activist-inclined millennials will unlikely cope in such circumstances and if Big Tech is not bailed out, along with government subsidies for its soaring energy bills, there will be social pandemonium. If troubles persist, tech titans have a panoply of social pacification tools that were built on decades-worth of behavioural data and user histories. Is it a coincidence that Google is working with Chinese authorities on the highly-controversial Project Dragonfly? Will this project be expanded into something more global?

The tight entwinement of substrate components in the global data ecosystem had traditionally prevented antitrust motions against US-based tech giants from succeeding. Antitrust measures generally proceed slowly and rarely keep up with fast-paced ephemeralization forces (Fuller, 1938, 1957; Heylighen, 2002) in the industry. In one early instance of antitrust litigation, this disconnect had even reduced judicial mediation to a mere "homiletic contribution" (Sullivan, 2001). The US Justice Department's July 23 2019 antitrust review on "the practices of market-leading online platforms" did not even mention a specific company although proceedings are expected to probe Alphabet, Facebook and Amazon. Investors meanwhile remain unfazed (Kim, 2019).

Unrestrained by legal impediments, the progression from surveillance capitalism to surveillance societies should be smooth, at least by Big Tech's calculus. Thinking of a protest? Zuckerberg's synaptic projector or an analogue may pre-emptively beam them out as holographic montages across city skies. This is the kind of *Minority Report* society being fantasized by Big Tech and Big Government. Amazon Web Services has already secured a US$600mn data storage deal with the CIA for its Secret Region project (Amazon, 2017) while Microsoft is challenging Amazon for the Pentagon's Joint Enterprise Defense Initiative cloud project worth at least US$10bn (Weise, 2019). This is where big money is secured for big data.

As Apple chairman Tim Cook warned at the 40th International Conference of Data Protection and Privacy Commissioners in Brussels: "our own information is being weaponized against us with military efficiency. We should not sugarcoat the consequences. This is surveillance." Cook called for laws to protect citizens from the "data industrial complex" (Gallagher, 2018) – a digital gulag reminiscent of Jeremy Bentham's two-century-old panopticon concept.

Curiously, Amazon owner Jeff Bezos may have been an early victim of the data-industrial complex. He is blaming his recent sexting scandal and ongoing divorce – reportedly the biggest in history – on a "political hit job" linked to US President Donald J. Trump (Smith, 2019).

The spectre of a digital panopticon has become so dire that Tim Berners-Lee, founder of the World wide web, has taken furlough from the Massachusetts Institute of Technology to launch Inrupt, an undertaking aimed at decentralizing the internet once again. As *the American Conservative* noted in a blog (Tepper, 2019):

> "Although the architecture of the internet is still decentralized, the ecosystem of the World Wide Web is not. A few giant companies have near-monopolistic control of traffic, personal data, commerce, and the flow of information. If you had to choose a date for when the internet died, it would be in the year 2014. Before then, traffic to websites came from many sources, and the web was a lively ecosystem. But beginning in 2014, more than half of all traffic began coming from just two sources: Facebook and Google. Today, over 70 per cent of traffic is dominated by those two platforms."

The data-industrial complex can emerge through any number of ways, including sociopolitical turbulence induced by a global economic meltdown or wildcard events such as an internet takedown or a prolonged power grid failure in a major economy.

Balkanized data world

Despite the prevalent hype and bubbles, BDA will remain relevant even if its proper evolution may take years to materialize. Until then, a global data fractionation process is underway in tandem with wealth fractionation, leading to a cartel-like monopoly by five or six global corporations. For the sake of maintaining sovereignty across the PESTLE spectrum, the Russia, India and China (RIC) nations are now pursuing indigenous data paradigms.

China is already working on a repressive social credit system to engineer external behavioural comportment to Taylorian-type ideals laid down by the ruling centre. At some point, a tipping point may emerge when external behavioural diktats are internalized by a "policeman in the mind" (Curtis and British Broadcasting Corporation, 2006). The West's experimentation with Freudian theories of mass pacification has only led to the attritional loss of competitiveness to emerging powers such as the RIC nations, South Korea, Taiwan and Japan. Beijing's primary goal for the moment, however, is to pre-empt French-inspired Yellow Vest-type protests through digital countermeasures.

The Great Firewall of China, as Beijing's net-centric surveillance and censorship system is often called, involves a high degree of complexity. Updated filters are routinely used to deal with the growing sophistication of virtual private networks. Patchworks in such software-intensive systems (SIS) may sooner or later buckle under the forces of ephemeralization and lead to a complexity catastrophe (Heylighen, 2002; Horner and Symons, 2014; Eisenhardt and Piezunka, 2011; Fuller, 1938, 1957).

To avoid an SIS fiasco and concurrently counter Big Tech's digital hegemony, Russia is experimenting with an independent internet system altogether. The Russian internet (Runet) entails an alternative domain name system that enables it to access and route internet traffic by itself (Matsakis, 2019). Ongoing trial runs are probing Runet's internal vulnerabilities as well as evidence of any lingering overreliance on tools, software, and infrastructure outside Russia's borders.

However, as Paul Barford, professor at the University of Wisconsin-Madison recounted to Wired.com (Matsakis, 2019): "because of the complexity across all levels of the protocol stack, there could be catastrophic failures somewhere." Andrew Blum, the author of *Tubes: a journey to the center of the internet* added: "every [web] page is made of 1,000 different things. If you are running a website in Russia, you would have to figure out where everything is coming from."

Several dry runs would be needed before the system attains a degree of robustness. Numerous other conditions must be met. One naturally involves the indigenization of vital cogs in the Runet ecosystem. Another would involve access to vital scientific data and publications anywhere on earth to sustain Russian scholarship, research and innovation. These must be stored on a Russian server (as well as libraries) instead of the external cloud servers of USA data titans. External communications, involving the legitimate transfer of money or internet-enabled joint ventures, may have to be facilitated by a blockchain-secured system. Real-time data such as external stock market movements can be captured or mirrored in real-time for domestic constituents. Mobile connectivity must be facilitated by Russian-made smartphones and operating systems such as the Aurora OS, which has serendipitously gained a potential market from USA sanctioned mobile giant Huawei.

To attribute Russia's internet angst to the current antagonism between itself and the West would be too simplistic. A complex system such as Runet would have taken years of conceptualizing, planning and process modelling. Could it be that Runet's *raison d'être* all along was to ensure a degree of national sovereignty in anticipation of turbulent times ahead? There is absolutely no doubt that Russia intends to remain a major power and control over its data would be a *sine qua non*. The superpowers of tomorrow may be defined more by their adroit use of data rather than military might. Less innovative societies, caught between future data superpowers, may face outright redundancy. The imponderables involved here, and implications regarding a future data-shaped international order, deserve a separate study altogether.

Conclusion: some light for big data

Big data's evolution from a bubble to panoptic phase involves the potentiation of several change dynamics. It is also symptomatic of the omission of foresight in big data undertakings. Change dynamics is a catch-all term that refers to trends, drivers, inflection points, patterns, enablers, wildcards and catalysts that may impact a policy or paradigm (Delaney and Osborne, 2013). In the context of the big data industry, these are roughly sketched out on Figure 2.

In the final analysis, data-driven societies are complex adaptive systems. Big data will entail human supervision of the highest calibre, adroit knowledge modelling and multiple motors of adaptation in sync with planetary developments. This foresight-helmed ecosystem can only be developed through careful planning, checks and balances, trials and errors as well as inputs from a quadruple helix society – one where the citizen is on equal footing with the government, academia and corporations. AI would be the enabler; not the clincher in a future smart society.

Organizational ingenuity and flexibility will determine how much value is harnessed from big data. Through foresight, the "new masters of Big Data" can become the "masters of ingenuity" (Moorthy *et al.*, 2015; Moldoveanu, 2013). It will be a "humans-first" and not systems-first paradigm.

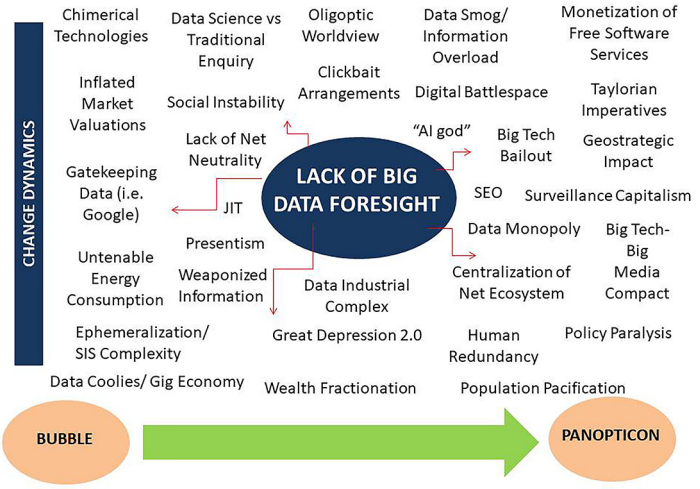


Figure 2.
Big data bubble to panopticon: the change dynamics

There are already nascent developments that presage big data's immense potentials. Researchers from Lawrence Berkeley National Laboratory recently subjected a welter of scientific papers to an AI-driven software called Word2Vec to probe whether there was any "latent knowledge" that scholars had initially overlooked. Using only word associations, the AI made "connections that no scientists could", according to researcher Anubhav Jain. By assessing the language of 3.3 million abstracts from material sciences, the algorithm discovered predictive cues for potential thermoelectric materials that can convert heat into energy (Tshitoyan *et al.*, 2019). When publications from before 2009 were fed into the algorithm, it was able to predict one of the most effective modern-day thermoelectric materials four years before it was actually discovered in 2012 (Zero Hedge, 2019). Likewise, researchers from the Chennai-based Institute of Mathematical Sciences have developed a comprehensive database of endocrine-disrupting chemicals or EDCs. Lessons learnt from this study can be used for developing machine learning-based predictive tools (Sharma, 2019).

This the "sunlight" big data needs. A global database of such successes would greatly inform policy-makers and encourage industry entrants. While such databases remain poorly-interlinked in the digital domain, they will be game-changers once strong linkages are established. This is the immediate challenge and crucial grand response (Popper, 2012) that the big data industry needs.

References

Ashley, M. (2019), "Will artificial intelligence soon tell Us how to live?", Forbes, 27 January, available at: www.forbes.com/sites/cognitiveworld/2019/01/27/will-artificial-intelligence-soon-tell-us-how-to-live/#3b1bb39e2938 (accessed 17 February 2019).

Barnett, R. (2017), "Why the promise of big data hasn't delivered yet", Techcrunch.com, 29 January, available at: https://techcrunch.com/2017/01/29/why-the-promise-of-big-data-hasnt-delivered-yet/ (accessed 27 February).

Bendix, A. (2018), "San Francisco's homelessness crisis is so bad, people appear to be using poop to graffiti the sidewalks", Business Insider US, 20 November, available at: www.businessinsider.my/san-francisco-homelessness-poop-graffiti-2018-11/?r=US&IR=T

Boggan, S. (2015), "Gold rush California was much more expensive than today's Tech-Boom California", Smithsonian.com, 30 September, available at: www.smithsonianmag.com/history/gold-rush-california-was-much-more-expensive-todays-dot-com-boom-california-180956788/#Ryr68162LrUw4iWL.99 (accessed 18 February 2019).

Brandon, J. (2017), "An AI god will emerge by 2042 and write its own bible. Will you worship it? ", Venturebeat.com, 2 October, available at: https://venturebeat.com/2017/10/02/an-ai-god-will-emerge-by-2042-and-write-its-own-bible-will-you-worship-it/ (accessed 17 February 2019).

Business Wire (2013), "Survey shows millennials are more forgetful than seniors", Businesswire.com, 1 August, available at: www.businesswire.com/news/home/20130801006048/en/Survey-Shows-Millennials-Forgetful-Seniors (accessed 13 June 2019).

Chamorro-Premuzic, T. (2014), "How the web distorts reality and impairs our judgement skills. Guardian", 13 May.

Child, J. (1997), "Strategic choice in the analysis of action, structure, organizations and environment: retrospect and prospect", *Organization Studies*, Vol. 18 No. 1, pp. 43-76.

Chugh, R. (2015), "Do Australian universities encourage tacit knowledge transfer?", *Proceedings of the 7th International Joint Conference on Knowledge Discovery*, Knowledge Engineering and Knowledge Management.

Columbus, L. (2018), "10 Charts that will change your perspective of big data's growth", Forbes.com, 23 May, available at: www.forbes.com/sites/louiscolumbus/2018/05/23/10-charts-that-will-change-your-perspective-of-big-datas-growth/#56bcdbbc2926 (accessed 17 February 2019).

Cukier, K.N. and Mayer-Schoenberger, V. (2013), "The rise of big data: how it's changing the way We think about the world", Foreign Affairs, May/June, available at: www.foreignaffairs.com/articles/2013-04-03/rise-big-data (accessed 17 February 2019).

Curtis, A, and British Broadcasting Corporation (2006), *Century of the Self*, BigD Productions [distributor].

Delaney, K. and Osborne, L. (2013), "Public sector horizon scanning – stocktake of the Australasian joint agencies scanning network", *Journal of Futures Studies*, Vol. 17 No. 4, pp. 55-70.

Edelman, B. (2011), "Bias in search results?: diagnosis and response", *The Indian Journal of Law and Technology*, Vol. 7.

Eisenhardt, K.M. and Piezunka, H. (2011), "Complexity theory and corporate strategy", in Allen, P., Maguire, S. and McKelvey, B. (Eds), *The Sage Handbook of Complexity and Management*, Chapter 29, SAGE, LA.

Florida, R. (2018), "The extreme geographic inequality of High-Tech venture capital", Citylab.com, 27 March, available at: www.citylab.com/life/2018/03/the-extreme-geographic-inequality-of-high-tech-venture-capital/552026/ (accessed 18 February 2019).

Fuller, R.B. (1938), *Nine Chains to the Moon*, Anchor Books, N/A.

Fuller, R.B. (1957), "A comprehensive anticipatory design science", Royal Architectural Institute of Canada, 216 via Google Books, 14 September, p. 34.

Gallagher, W. (2018), "Tim cook calls data collection 'surveillance' in blunt, forceful speech on 'data-industrial complex", Appleinsider.com, 24 October, available at: https://appleinsider.com/articles/18/10/24/tim-cook-calls-data-collection-surveillance-in-blunt-forceful-speech-on-privacy (accessed 17 February).

Gantz, J. and Reinsel, D. (2011), "Extracting value from chaos", IDC IView, pp. 1-12.

Ghosh, S. and Kanter, J. (2019), "Google says data is more like sunlight than oil, one day after being fined $57 million over its privacy and consent practices", Business Insider US, 22 January. available at: www.businessinsider.my/google-data-is-more-like-sunlight-than-oil-france-gdpr-fine-57-million-2019-1/?r=US&IR=T (accessed 25 February 2019).

Gray, M.L. and Suri, S. (2019), *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, Houghton Mifflin Harcourt, Boston.

Hajirahimova, M.S. and Aliyeva, A.S. (2017), "About big data measurement methodologies and indicators", *International Journal of Modern Education and Computer Science*, Vol. 9 No. 10, pp. 1-9, available at: www.mecs-press.org/ijmecs/ijmecs-v9-n10/IJMECS-V9-N10-1.pdf

Hao, K. (2019), "The AI gig economy is coming for you", *MIT Technology Review*, 31 May, available at: www.technologyreview.com/s/613606/the-ai-gig-economy-is-coming-for-you/ (accessed 12 June 2019).

Harris, J.C. (1881), *Uncle Remus: His Songs and His Sayings; the Folk-Lore of the Old Plantation*, D. Appleton and Company, New York, NY.

Heylighen, F. (2002), *Complexity and Information Overload in Society: why Increasing Efficiency Leads to Decreasing Control*, CLEA, Free University of Brussels, Pleinlaan 2, B-1050. Brussels.

Horner, J. and Symons, J. (2014), "Reply to Angius and Primiero on software intensive science", *Philosophy and Technology*, Vol. 3 No. 27, pp. 491-494.

Jarche, H. (2010), "Sharing tacit knowledge", available at: www.jarche.com/2o1o/o1/sharing-tacit-knowledge/

Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G. and Hooker, G. (2009), "Data-intensive science: a new paradigm for biodiversity studies", *BioScience*, Vol. 59 No. 7, pp. 613-620.

Kim, T. (2019), "Why investors shouldn't overreact to the DOJ's antitrust review of big tech", Barrons.com, 26 July. Available at: www.barrons.com/articles/why-investors-shouldnt-overreact-to-an-antitrust-review-of-big-tech-51564181107 (accessed 29 July 2019).

Kingston, J.K.C. (2012), "Tacit knowledge: capture, sharing, and unwritten assumptions", *Journal of Knowledge Management Practice*, Vol. 13 No. 3.

Kitchin, R. (2014), "Big data, new epistemologies and paradigm shifts", *Big Data and Society*, Vol. 1 No. 1, pp. 1-12.

Langlois, S. (2019), "This 'hidden asset' is about to deliver the next trillion-dollar company, analyst says", Marketwatch.com, 31 January, available at: www.marketwatch.com/story/this-hidden-asset-is-about-to-deliver-the-next-trillion-dollar-company-analyst-says-2019-01-31 (accessed 19 February 2019).

Larson, J., Mattu, S., Kirchner, L. and Angwin, J. (2016), "How we analyzed the COMPAS recidivism algorithm", Propublica.org, 23 May, available at: www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm (accessed 17 February 2019).

Latour, B. (2005), *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford University Press, Oxford.

Levy, A. (2019), "US Online grocery is just getting started", The Motley Fool, 2 February. available at: www.fool.com/investing/2019/02/02/better-buy-amazoncom-vs-alphabet.aspx (accessed 19 February 2019).

Liao, S. (2018), "Amazon warehouse workers skip bathroom breaks to keep their jobs, says report", Theverge.com, 16 April, available at: www.theverge.com/2018/4/16/17243026/amazon-warehouse-jobs-worker-conditions-bathroom-breaks (accessed 20 February 2019).

Liedke, L. (2019), "100+ Internet statistics and facts for 2019", Updated 26 March, available at: www.websitehostingrating.com/internet-statistics-facts/ (accessed 11 June 2019).

Maavak, M. (2019b), "Developing a net-centric foresight model for the management of emerging risks", Unpublished Doctoral Thesis in Public Policy, Universiti Teknologi Malaysia, February 2019.

Maavak, M. (2019a), "Clickbait arrangement between google and the Deep-Pocketed media? Mathew Maavak's blog", 13 June, available at: https://mathew-maavak.blogspot.com/2019/06/clickbait-arrangement-between-google.html (accessed 14 June 2019).

Matsakis, L. (2019), "What happens if Russia cuts itself off from the internet", Wired.com, 12 February, available at: www.wired.com/story/russia-internet-disconnect-what-happens/ (accessed 15 March 2019).

Moldoveanu, M.C. (2013), "The ingenuity imperative: what big data means for big business", *Rotman Magazine*, pp. 59-63.

Moorthy, J., Lahiri, R., Biswas, N., Sanyal, D., Ranjan, J., Nanath, K. and Ghosh, P. (2015), "Big data: prospects and challenges", *Vikalpa: The Journal for Decision Makers*, Vol. 40 No. 1, pp. 74-96.

Mosher, D. (2011), "High wired: does addictive internet use restructure the brain?", Scientific American, 17 June, available at: www.scientificamerican.com/article/does-addictive-internet-use-restructure-brain/ (accessed 11 June 2019).

Naughton, J. (2017), "The trouble with BITCOIN and big data is the huge energy bill", TheGardian.com, available at: www.theguardian.com/commentisfree/2017/nov/26/trouble-with-bitcoin-big-data-huge-energy-bill (accessed 12 June 2019).

Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-Creating Company*, Oxford University Press. New York, NY.

Phelan, M. (2012), "The death of big data", Forbes.com, 4 October, available at: www.forbes.com/sites/ciocentral/2012/10/04/the-death-of-big-data/#158e1fa779b8 (accessed 12 June 2019).

Polanyi, M. (1967), *The Tacit Dimension*, Anchor Books. New York, NY.

Popper, R. (2012), "Grand challenges and STI foresight", Manchester Institute of Innovation Research, University of Manchester, Instructional Module, 6 April 2012.

Puryear, R.P. and Root, R.J. (2013), "Infobesity: the enemy of good decisions [issue brief]", Retrieved on from Bain and Company Decision Insights website, available at: www.bain.com/publications/articles/infobesity-the-enemy-of-good-decisions.aspx (accessed 17 February 2019).

Reinsel, D. Gantz, J. and Rydning, J. (2017), "Data age 2025: the evolution of data to Life-Critical (PDF)", International Data Corporation, Framingham, MA, available at: www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

Ribeiro, J. (2017), "Facebook wants people to type from their brain and hear with their skin", PCWorld.com, 20 April, available at: www.pcworld.com/article/3191268/facebook-aims-to-let-people-type-from-their-brain-hear-with-skin.html (accessed 19 February 2019).

Robison, P. (2019), "Boeing's 737 max software outsourced to $9-an-hour engineers", Boomberg.com, 29 June, available at: www.bloomberg.com/news/articles/2019-06-28/boeing-s-737-max-software-outsourced-to-9-an-hour-engineers (accessed 16 July 2019).

Roubini, N. (2018), "The big blockchain lie", Project Syndicate, 15 October, available at: www.project-syndicate.org/commentary/blockchain-big-lie-by-nouriel-roubini-2018-10?barrier=accesspaylog (accessed 12 June 2019).

Rushkoff, D. (2013), *Present Shock: When Everything Happens Now*, Current, New York, NY.

Sharma, D.C. (2019), "Indian scientists develop database of everyday chemicals harmful to human health", India Science Wire, 19 July, available at: https://vigyanprasar.gov.in/isw/Indian-scientists-develop-database-of-everyday-chemicals-harmful-to-human-health.html (accessed 26 July 2019).

Shenk, D. (1997), *Data Smog: Surviving the Information Glut*, Harper, San Francisco.

Singh, S. (2016), "Denied breaks, U.S. poultry workers wear diapers on the job", Bloomberg.com, available at: www.bloomberg.com/news/articles/2016-05-11/poultry-workers-in-diapers-as-bathroom-breaks-denied-oxfam-says (accessed 12 May 2016).

Smith, D. (2019), "Trump, 'blackmail' and a pecker: Bezos delivers scandal with something for everyone", TheGuardian.com, 9 February, available at: www.theguardian.com/technology/2019/feb/09/jeff-bezos-trump-national-enquirer-amazon-david-pecker (accessed 10 March 2019).

Sullivan, L.A. (2001), "Is competition policy possible in high tech markets: an inquiry into antitrust, intellectual property, and broadband regulation as applied to the new economy", *Case Western Reserve Law Review*, Vol. 52 No. 1.

Sweeney, E. (2018), "Global ad revenue delivers strongest growth in 8 years with 7.2% increase, per magna", MarketingDive.com, 4 December, available at: www.marketingdive.com/news/global-ad-revenue-delivers-strongest-growth-in-8-years-with-72-increase/543491/ (accessed 17 February 2019).

Symons, J. and Alvarado, R. (2016), "Can we trust big data? Applying philosophy of science to software", Big Data and Society, July–December, pp. 1-17

Taylor, C. and Wagman, L. (2014), "Customer privacy in oligopolistic markets: winners, losers, and welfare", *International Journal of Industrial Organization*, Vol. 34, pp. 80-84.

Taylor, F.W. (1911), *The Principles of Scientific Management*, Harper and Brothers, New York, NY.

Tepper, J. (2019), "The death of the internet", The American Conservative blog, 8 March. Available at: www.theamericanconservative.com/articles/the-death-of-the-internet/ (accessed 20 March 2019).

Tshitoyan, V., Dagdelen, J., Weston, L., Dunn, A., Rong, Z., Kononova, O., Persson, K.A., Ceder, G. and Jain, A. (2019), "Unsupervised word embeddings capture latent knowledge from materials science literature", *Nature*, Vol. 571 No. 7763, pp. 95-98, doi: 10.1038/s41586-019-1335-8.

Tung, L. (2017), "IoT devices will outnumber the world's population this year for the first time", ZDNet.com, 7 February, available at: www.zdnet.com/article/iot-devices-will-outnumber-the-worlds-population-this-year-for-the-first-time/ (accessed 12 June 2019).

Waddington, P. (1996), *Dying for Information? An Investigation into the Effects of Information Overload in the UK and Worldwide*, Reuters Business Information, London.

Weise, K. (2019), "Amazon and microsoft are 2 finalists for $10 billion pentagon contract", *New York Times*, 11 April.

Wurman, R.S. (1989), *Information Anxiety*, Doubleday, New York, NY.

Xie, A. (2018), "The global asset bubble will burst – the only question is when, and how bad it will be", SCMP.com, 6 January, available at: www.scmp.com/comment/insight-opinion/article/2126952/global-asset-bubble-will-burst-only-question-when-and-how (accessed 16 February 2019).

Zero Hedge (2019), "AI pores over old scientific papers, makes discoveries overlooked by humans", Zerohedge.com, 13 July, available at: www.zerohedge.com/news/2019-07-10/ai-pores-over-old-scientific-papers-makes-discoveries-missed-humans (accessed 23 July 2019).

Zuboff, S. (2019), *The Age of Surveillance Capitalism: The Fight for the Human Future at the New Frontier of Power*, Profile Books, London.

Further reading

Harris, M. (2017a), "Inside the first church of artificial intelligence", Wired.com, 15 November. available at: www.wired.com/story/anthony-levandowski-artificial-intelligence-religion/ (accessed 17 February 2019).

Harris, M. (2017b), "Who is Anthony Levandowski, and why is Google suing him?", Theguardian.com, 24 February, available at: www.theguardian.com/technology/2017/feb/23/anthony-levandowski-google-uber-self-driving-cars-lawsuit (accessed 17 February 2019)

Lardinois, F. (2017), "AWS launches a secret region for the US intelligence community", Techcrunch.com, 20 November, available at: https://techcrunch.com/2017/11/20/aws-launches-a-secret-region-for-the-u-s-intelligence-community/ (accessed 19 February 2019).

Vaughan, L. and Thelwall, M. (2004), "Search engine coverage bias: evidence and possible causes", *Information Processing and Management*, Vol. 40 No. 4, pp. 693-707.

Corresponding author

Mathew Maavak can be contacted at: mfmathew@yahoo.com