

Pandemics and big data tyrannies



Computer security threats can make the internet a dangerous place. Computer viruses (such as Stuxnet and CryptoLocker) can quickly spread throughout a network, infect your device, and take over your virtual identity. A virus can send emails on your behalf, make bank withdrawals, post a private picture on your favourite social media platform, or share your latest-greatest dance moves on TikTok. Luckily, however, you don't have to embarrass yourself, because anti-virus software tools are here to the rescue! They can detect the virus and *quarantine* it. Anti-virus companies always provide an updated version to include electronic anti-bodies for all novel viruses. If a virus is programmed by some teenage hacker, anti-virus companies work harder and faster to code the anti-virus and spread it to all users as swiftly as possible.

More dangerously, computer viruses can attack infrastructures, including electric grids, nuclear plants, and any other vital vein of society. For instance, many hospitals use computerised medical equipment, tele-health devices, and in some cases robotics – all of which are hackable and prone to viruses. Imagine losing many patients because devices in the operation room didn't function as they should due to a computer virus? Would doctors/nurses/patients vote to retire all computer devices, close the hospital, *stay home*, and only keep anti-virus software companies open so they can battle the computer virus? I guess you would think to yourself: No.

During outlier events such as the Covid-19 outbreak, most models from the US Centers for Disease Control and Prevention ([CDC](#)) and the World Health Organisation ([WHO](#)) were way off (predictions were very different from actual values). The reason is that those models were built with limited data, without looking into external factors affecting the forecasts, and without considering [multivariate outliers](#). Those models exacerbated the case of *misinformation in the nation*. Albeit most traditional economic models deploy causality analysis to infer policy, regular models don't account much for outliers, and most analysts deal with outliers based on human experience or elimination of such events from predictive models.

As the price of misinformation dissemination during a catastrophe can only be comprehensively measured in retrospect, it is important to not dumb down the metrics of success or failure to the total of positive Covid-19 cases. It is imperative to achieve an *equilibrium* between national productivity, the economy, efficiency of healthcare systems, rural areas access, urban cities' vigour, [well-being of the healthy](#), and the fate of the sick. We cannot manage what we cannot measure, and if our only metric during such exceptional times is one number (positive cases); then we are certainly doomed to be off-balance and fall-off.

Moreover, the current proposed solutions to the virus, including contact tracing, easing of e-health records restrictions, smart-phone tracking, and e-temperature measurements would create big amounts of data but pose many critical questions: Who owns the health data generated by tracing? Who manages the data generated by tracking? Are they selling it to third party vendors? And how are they protecting the data infrastructure from computer viruses and hacker attacks? As global interconnectedness continues to grow, more and more people are finding themselves living in a data republic. A republic that ought to be [democratised](#), but a typical case of big data tyranny occurs when citizens share their data (willingly) so they can achieve safety and good health. The wisdom of Benjamin Franklin states that: "Those who would give up essential liberty to purchase a little temporary safety, deserve neither". Once we give up our data liberties, *digital social distancing* would not be a choice only during exceptional times, rather, an unavoidable reality of electronic isolation and easy targeting.

Data democracy allows personal data ownership, makes data available to citizens, and reduces costs to popularise, disseminate, and crowd-source data. The opposite of that (i.e. a big data tyranny), is a centralisation of all personal *health* datasets in the hands of few companies or entities that have *not* been chosen by data creators. That notion dictates the transformation of the famous saying "*data is the new oil*" (indicating the rising power of data in the global economy) to "*data is the new blood*".

Many studies compared the reactive human behaviour during the [Spanish flu pandemic](#) to that of World War I and World War II, and to that point, a story begs to be told about the city of Pittsburgh. By the 1930s, Pittsburgh was producing half of the nation's demand of steel. Steel proved invaluable during World War II. Interestingly, the day of the Japanese attack upon Pearl Harbor in 1941, Pittsburgh was hosting an "America must stay neutral in this World War" event. Although the city was reluctant to go to war, once war was inevitable, the city mills operated 24 hours a day to produce more than 95 million tons of steel for the war effort. This spike in production created a workforce shortage, which led to a major reduction in unemployment rates. Steel would then be used to build battleships, aircraft carriers, and ammunition. Recovering from Pearl Harbor would have been considerably delayed without Pittsburgh. That would've possibly provided Japan and Germany another chance to launch attacks at the U.S. and potentially tilt the outcome of the war.

During crisis times, healthy citizens need to be much more productive than usual, and not less; and all countries need to figure out ways to be more functional during pandemics. Otherwise, all supply chains are intercepted, and everyone along the way (liberal or conservative) will be harmed. While expecting health and safety in return, you will feel obliged to give up your personal privacy in the shape of a voluntary health data donation.

Contrary to that, by democratising big data and providing the right information to the public, Nietzsche's old adage "he who has a why can live to bear almost any how" can be deployed to call for action and march forward against any visible or invisible enemy.



Notes:

- This blog post expresses the views of its author(s), not the position of the CBI, LSE Business Review or the London School of Economics.
- Featured [image](#) by [geralt](#), under a [Pixabay](#) licence
- When you leave a comment, you're agreeing to our [Comment Policy](#)



Feras A. Batarseh is a [research assistant professor](#) with the College of Science at George Mason University (GMU). His research spans the areas of data science, artificial intelligence, and context-aware software systems. Dr. Batarseh obtained his Ph.D. and M.Sc. in computer engineering from the University of Central Florida (UCF) (2007, 2011), and a graduate certificate in project leadership from Cornell University (2016). He is the director of [Turing Research](#), an AI for policy group at GMU. Additionally, Dr. Batarseh published and edited several book chapters.