

Forecasting the Negative Impacts of the Big Data Era on Political and Social Issues

Dr. Van Hanh NGUYEN

Abstract: Nowadays, big data strongly affects all aspects of social life. It profoundly changes the way humans exist daily and hourly. The article makes some predictions about the negative impacts of the data era on socio-political issues such as the risk of data monopoly and digital dictatorship; liberty loss, mass unemployment, personalized discrimination, and biological inequality. These consequences never appeared in previous societies, so there is a need for distinctive solutions and humanity's unity.

Keywords: big data era; data ownership; data monopoly; impacts of dataism.

Nowadays, the presence of big data has brought about a revolution not only in production but also in the way humans exist. The world is discussing big data; people praise the scientific and technological advances based on this revolution. Indeed, these advances are extremely miraculous, but alongside that, the data era also has negative impacts on present and future society. Humans present and future will face the risks of loss of freedom, democracy, unemployment, and many other consequences. We need to recognize these negative impacts early on to have timely solutions.

1. The big data era

In today's complex scientific world, researchers no longer simplify their study subjects but instead put them in a realistic state, which is constantly evolving and developing. Therefore, it cannot be expressed by linear relationships with a few weakly correlated factors. This requires recording the entire process and a large number of parameters to describe its diversity and complexity. Describing the state of the real world has created a need for big data, as only big data can fully describe the real world comprehensively, vividly, complicatedly, and diversely. The emergence of big data has quickly opened up a new era of data for humanity. Some recent studies suggest that "everything in the world is made up of data, the nature of the world is data, and everything can ultimately be represented by data. Therefore, the world of objects becomes fragmented data, or it can be said that the world is made up of data fragments"⁽¹⁾. The big data era sees the world as a rich, diverse, and distinctive entity; any data has meaning and needs to exist. Therefore,

in the big data era, any data, once generated, must be preserved and stored. In the past, it was believed that it was necessary to rely on the law of cause and effect to explain phenomena and predict the future. But nowadays, data ideology believes that humans can perceive the relationship and laws between objects by using data about those objects instead of necessarily using causal relationships. “This will reduce the level of difficulty in perceiving and grasping objects”⁽²⁾.

In essence, the big data era is not a new era, but rather a development from the industrial society to the post-industrial or information society. The big data era applies advanced techniques such as the Internet, smart sensors, and cloud computing to collect and store information. Along with the emergence of the big data era, data has become an objective and independent type of existence, becoming an information world outside the world of objects and the spiritual world.

Data is a set of facts, such as numbers, words, and images, used to measure, observe, or simply describe objects. In the small data era, collecting and processing data has always been a challenging problem, almost all data is manually collected, such as experimental data, sampling data. However, in the big data era, data collection has been smartened and automated; therefore, the amount of data has increased explosively. The role of data becomes more prominent and becomes the foundation of cognitive science. Many things that could not be quantified, stored, or shared in the past have been digitized. Smart sensors of the big data era transform our location, images, sounds, behaviors, and physical states as well as changes in the physical world into data, which is stored on the Internet cloud. This is consistent with the research findings of Viktor Mayer-Schönberger and Kenneth Cukier in 2013, who stated that “Big data marks a significant step forward for humanity on the path to quantifying and understanding the world. Many things that could not be quantified, stored, analyzed, or shared in the past, have now been digitized. Owning large amounts of data, and mostly unfiltered, has opened up new doors to interpreting our world”⁽³⁾. From there, data has also become an asset that can be owned and traded. Data is a very new object of ownership in today’s era. Media companies, social services have accumulated data to create Big Data, forming big data companies such as social networks: Facebook, Youtuber, Google, Instagram, Twitter, banks, financial companies, healthcare services, education, etc. Data is becoming an important object and a potential for exclusive ownership in the context of the market economy. Owning data is not only an economic issue but also a socio-political issue.

2. Negative impacts of the big data era

2.1. Forecasting the risk of loss of freedom and the emergence of a digital dictatorship

In the past, social effects and directions often followed the trend of the crowd or the guidance of eminent individuals with credibility in society. However, with the new trend of technology today, the bigger the data corporations are, the more they have the ability to direct society and create new monopolies - data monopoly will be a prerequisite for a digital dictatorship. Data today is an invaluable resource and it is being collected and stored, thus data can become an asset to own. In fact, there have been big data companies such as Facebook, Twitter, Youtube, Instagram, etc., but effective regulations on data ownership rights for these companies have not yet been established.

Big data companies such as Google and Facebook are referred to as “attention merchants”. They provide us with information, free entertainment services, and then sell our attention to advertising services. However, their bigger purpose is to collect a massive amount of user information. They store all user information, including everything users post, share, and like. In the long term, by gathering enough data and computational power, these big data companies will “crack the code” of life’s deepest secrets and use this knowledge to control and direct users according to their own purposes. Previously, no tool could measure the inner world of humans, such as thoughts, emotions, trends, aspirations, etc. However, with large amounts of users’ data, algorithms can understand users better than they understand themselves. Companies will collect our data and guide our decisions by sending us suggestions that they want us to choose. As scientists gain a deeper understanding of how humans make decisions, they will gradually digitize this process, and algorithms will increasingly replace human will in decision-making. Today, many people rely on Google Maps for directions rather than their own will. When algorithms make accurate decisions, humans will gradually lose faith in emotions and personal will, relying more and more on algorithms. This process also involves a transfer of decision-making power from feelings and personal will to algorithms based on big data. The danger is not that algorithms can be smarter than humans, but that AI can process data to manipulate the human mind. We have already experienced some of these things in recent elections and referendums around the world when hackers learned how to manipulate individual voters by analyzing their data and exploiting their preconceptions. Then, the freedom of human beings will disappear completely.

In fact, big data companies have partially realized this in advertising products, services, and consumption. However, in the future, unreasonably uncontrolled data ownership could interfere with political affairs and sway voter decisions. The

spread of fake news on Facebook during the 2016 US presidential election, such as “Pope Francis endorses Donald Trump” and “Hillary Clinton sells weapons to the self-proclaimed Islamic State - IS”, was believed to have contributed to the defeat of Hillary Clinton and the Democratic Party, while also creating a surprise victory for Donald Trump and the Republican Party. To this day, there are still suspicions that these fake news were intentionally disseminated. In late 2020, the CEOs of the two leading social networks, Facebook and Twitter, testified before the US Congress on issues related to information control and allegations of political bias. The hearing, titled “Breaking the News: Censorship, Suppression, and the 2020 Election”, was hastily convened after signs that these social networks were “restraining” some election-related information to create bias⁽⁴⁾. Although they rebuffed all accusations, the hearing raised a fierce debate over the ownership of big data and its impact on democracy.

In the future, so as to control diseases or crimes, citizens of nations may be required to wear a monitoring device. If the monitoring system is in the hands of a good government, powerful monitoring algorithms could do the best thing for humanity, which is to ensure health and safety. However, if those big data algorithms fall into the hands of a totalitarian government, citizens could fall into a constant surveillance mechanism. This total surveillance mechanism not only monitors every activity and spoken word, but can also deeply observe human experiences. By using an increasingly deep understanding of the human brain and harnessing the immense power of machines, that system could, for the first time in history, capture what every citizen is thinking every second. In the Covid-19 pandemic, many countries have used technology to identify early signs of illness as well as individuals that those affected ones have contacted with. This is an initial indication that in the future, governments will use technology to manage citizens. Assuming that governments around the world are currently good, humanity still faces the risk of losing freedom because the world in which we live is at the intersection of two major revolutions. Biologists are unlocking the secrets of the human body, specifically the brain and emotions of humans, while computer scientists are providing us with unprecedented data processing capabilities. When the revolution in biotechnology merges with the revolution in information technology, we will create big data algorithms that can track and understand the inner feelings of every human being. From there, it can control our minds. Historian Yuval Noah Harari believes that “the illusion of my free will is likely to be shattered when I daily use smart electronic devices to interact with corporations and government agencies (understanding and controlling what until now has been my inner inviolable world)⁽⁵⁾”.

If users become “enlightened” and try to prevent the flow of data, they will find it increasingly difficult because humans connect to machines like mother connects her foetus in the womb. If we choose to disconnect, banks, insurance companies, hospitals, schools, and other institutions may refuse to serve us, or in other words, the world will disconnect from us.

2.2. Forecasting widespread unemployment risk

Artificial intelligence essentially digitizes human information processing methods, so to have artificial intelligence, a huge amount of data on human information processing methods for a particular issue is needed. However, once artificial intelligence is present, with computing speeds many times faster than the brain, it will excel in certain specialized fields beyond human capability. On February 10, 1996, IBM’s Deep Blue machine defeated world chess champion Garry Kasparov, and later, Google developed the even more advanced AlphaGo software, capable of learning the ancient Chinese game of Go. AlphaGo seals 4-1 victory over the world’s Go champion, Lee Sedol, in Seoul in March 2016⁽⁶⁾. Essentially, the machines have digitized a vast amount of data on chess and process it incredibly quickly, leading human chess players to lose.

Similarly, in the field of healthcare, in 2011, IBM company created the IBM Watson machine, nicknamed the “Super Doctor,” which can browse millions of medical records simultaneously to provide physicians with evidence-based treatment options within seconds, thanks to its ability to synthesize vast amounts of data and powerful processing speed. This “Super Doctor” also allows individuals to access information about their health status. Doctors only need to input patient data for analysis, compare it with available data, and provide accurate treatment recommendations. In a recent laboratory test, an algorithm accurately diagnosed 90% of lung cancer cases, while human doctors only diagnosed 50%. CT and mammography results are regularly checked by specialized algorithms, allowing doctors to gain additional insights, and sometimes the machines can even detect tumors that doctors have missed. Similarly, in 2011, at a San Francisco pharmacy, pharmacists could make mistakes up to 1.7%, but a pharmacy robot dispensed 2 million prescriptions in a year without making any mistakes⁽⁷⁾. This shows that in the future, a significant number of human doctors may be replaced by machine doctors.

Even the creation of artistic works such as music, painting, and poetry, which have traditionally been exclusive to humans, may be replaced by artificial intelligence in the future. David Cope, a music professor at the University of California, Santa Cruz, created a music intelligence experiment called EMI (Experiments in Musical Intelligence), which specializes in imitating the style

of the musical genius Bach. It took him seven years to complete EMI, but after that, EMI only took one year to compose 5,000 hymns in the style of Bach, and then continued to imitate the styles of Beethoven, Stravinsky, and finally, the audience could not distinguish between the music composed by the great composers and that composed by EMI⁽⁸⁾. On November 28, 2021, CNN reported that the University of Oxford (UK) recently organized the first poetry reading by a “robot artist” in the world. Ai-Da (the name of the robot artist) had a public reading of a poem she wrote using an algorithm, which paid tribute to the Italian poet Dante⁽⁹⁾.

In September 2013, two Oxford researchers, Carl Benedikt Frey and Michael A. Osborne, published “The Future of Employment,” in which they announced studies on the likelihood of different professions being taken over by computer algorithms in the next 20 years. These scientists predicted that by 2033, 99% of telemarketing and insurance sales representatives will be replaced by algorithms, 98% of sports referees will lose their jobs, 97% of cashiers and 96% of chefs will face a similar fate, and 94% of waiters will be replaced by robots, along with 94% of paralegals becoming redundant. Of course, some professions will have lower rates of automation, such as bartenders, with 77% expected to lose their jobs, and underwater rescuers, with 67% at risk. The least automatable jobs are those with low profits, but are difficult to replicate, such as archaeologists, who, according to the study, will see only 0.7% of their jobs being taken over by AI⁽¹⁰⁾.

Today, data is being stored and digitized at an ever-increasing rate, and the process is continuing rapidly with no clear end in sight. Essentially, humans have only two forms of labor: manual labor and brainpower. In the centuries before machines replaced manual labor, humans were able to engage in mental labor. However, in today’s world, artificial intelligence replaces human brainpower, so what will humans do? It is true that artificial intelligence, big data, the Internet of Things, and other advancements will create new jobs, but these new professions require very high qualifications that not all workers possess. Furthermore, this group of workers is not the majority. The industrial civilization of the twentieth century relied on the masses to take advantage of cheap labor, raw materials, and markets. However, in the twenty-first century, the post-industrial civilization will depend on Big Data, AI, biotechnology, nanotechnology, and other technologies that will replace human brainpower. Therefore, the biggest political question of the next few decades is “what will humans do?”. If we do not have human-centered data soon, we will face an unprecedented unemployment crisis.

2.3. Forecasting the risk of biological inequality

When wealth is concentrated in just a few people, society is divided into classes, as a result, many people suffer from economic and social inequality. However, in terms of biology, all people are fundamentally equal because no force has yet been able to interfere with this ability. However, in just a few decades, Bigdata algorithms have provided a vast amount of biological data. This has helped to improve healthcare significantly compared to the past. But at the same time, it has also opened new healthcare services such as gene therapy and gene editing. The enormous amount of DNA data, along with advances in biotechnology and genetics, has helped scientists understand the mechanism of DNA and intervene to develop them as desired.

In 2000, Sharon Saarien in West Bloomfield, Michigan (USA) gave birth to a baby girl named Alana. Alana's DNA came from her father Paul and mother Sharon, but through testing, doctors discovered that the couple's mitochondrial gene carried a dangerous mutation. Therefore, the doctors replaced the couple's mitochondrial gene with that of another woman. From a purely technical perspective, Alana had three biological parents. In 2001, the US government banned this method due to ethical concerns⁽¹¹⁾. However, on February 3, 2015, the UK parliament voted for a law called "three-parent embryos" that accepts this treatment method and related research in the UK⁽¹²⁾. Healing is always the defense for any biological intervention, but nowadays there is no clear boundary between healing the sick and upgrading the healthy.

People argue that if we can cure diseases, why not repair defects in unborn children? Suppose a genetic test shows that your unborn child is very intelligent and beautiful but has depression. Would you want to save their future? And while we're at it, why not give them a healthy immune system, a cheerful and optimistic personality, and more? Then people will be willing to spend money to upgrade their future children, and gradually, we will have a generation that is pre-designed. If this service is allowed, then only the rich and super-rich will be able to afford it.

Thus, the combination of information technology and biotechnology could potentially divide humanity into different biological classes: the upper class of wealthy people who can use money to modify and upgrade their genes to enjoy superior abilities, and the lower class of poor, ordinary people who cannot afford genetic intervention. Although this is just a prediction, signs of biological inequality have already emerged in the early 21st century. On May 14, 2013, actress Angelina Jolie decided to undergo a double mastectomy after a DNA test showed she had an 87% chance of breast cancer. The cost of the DNA test

that Angelina Jolie had done was \$3000 USD (not including the cost of surgery, reconstructive procedures, and related treatments). This is happening in a world where one billion people earn less than \$1 USD per day and 1.5 billion people earn \$1-2 USD per day⁽¹³⁾. Plastic surgery was invented during World War I to heal soldiers with facial injuries, but today it has become a beauty service for the wealthy. While over time, all technologies become cheaper, it remains to be seen whether all these achievements can be widely accessible to humanity. This issue may become even worse with the development of AI causing widespread unemployment, pushing millions of people into poverty with no income.

Finally, the convergence of information technology and biotechnology will allow the super-rich to use money to upgrade themselves, buy youth, beauty, and longevity. They will not only own vast wealth but also possess beauty, health, and boundless creativity like “superhumans”, while the rest of the world will be a vast underclass of ordinary people who are “useless” and unable to undergo genetic intervention.

2.4. Forecasting the risk of personalized discrimination

Recently, Google has carried out the “Google Baseline Study” project to build a mammoth database on human health, establishing the “perfect health” profile. Google hopes that this will help people to be aware of various diseases and to anticipate potential illnesses. “Google Baseline Study” is linked to the “Google Fit” product line, which is integrated into wearable accessories such as clothing, wristbands, shoes, glasses, and watches, and will collect an endless series of biometric data⁽¹⁴⁾. Not stopping there, Google also has ambitions to access customers’ DNA. Currently, the DNA testing market is developing rapidly due to its benefits in predicting and diagnosing diseases, and huge databases are the key to such predictions. However, suppose an insurance company gains access to customers’ DNA data and biometric data, and predicts that a particular customer has an 80% risk of developing cancer within a few years. Another customer enjoys risky activities. Will they still be willing to sell insurance to these customers or will they refuse without regret? This raises concerns about the risk of personalized discrimination.

In the future, companies may use algorithms to analyze biometric data, DNA data, or data collected on social networks to select personnel or partners. At that time, the algorithm may find something in the DNA, history, or social media accounts of candidates that the owners of these algorithms do not like. This is personalized discrimination. In the past, discrimination often occurred by group, such as by race, gender, religion, or ethnicity. Therefore, these groups could gather forces to fight against it. However, people who are discriminated

against based on personal data cannot do that because no one experiences discrimination exactly like them. This is personalized discrimination. This is not just an idea but also becoming a reality. Nowadays, more and more corporations and organizations are using algorithms to analyze data and make decisions about their employees, customers, or partners.

The era of big dataism is opening unlimited potential for economic and social development, but it also brings unprecedented challenges for future society. Its negative impacts can undermine the traditional democratic system that humans have built with blood and tears for many centuries, take away the valuable individual freedom of human beings, render the majority of people useless, and replace “nature” to create upgraded human species. These predictions may be too pessimistic, but signs of it are already present in today’s society. Therefore, we need to be aware of this and take action now to prevent it.

(1), (2), (3) Viktor Mayer-Schönberger, Kenneth Cukier (2013), *Big Data Era*, Zhejiang - China, People’s Publishing House, p.215, p.75, p.23.

(4) Nguyen Ha (18.11.2020). *Facebook and Twitter testify before the US Senate on information control*. <https://vtv.vn/cong-nghe/facebook-va-twitter-dieu-tran-truoc-thuong-vien-my-ve-kiem-soat-thong-tin-20201118122624426.htm>, accessed on October 16, 2021.

(5) Yuval Noah Harari (2019), *21 Lessons for the 21st Century*, The Gioi Publishers, Hanoi, p.72.

(6), (7), (8), (13) Yuval Noah Harari (2018), *Homo Deus: A Brief History of Tomorrow*, The Gioi Publishers, Hanoi, p.382, p.377, p.386-387, p.413.

(9) Tuoi Tre Online (28.11.2021), *World’s first “artist” robot writes poetry about Dante*. <https://tuoitre.vn/robot-nghe-si-dau-tien-cua-the-gioi-viet-tho-ve-dante-20211128072331666.htm>, accessed on December 10, 2021.

(10) Carl Benedikt Frey and Michael A. Osborne. (17.9.2013). *The Future of Employment: How Susceptible Are Jobs to Computerization*, <http://www.oxfordmartin.ox.ac.uk>, accessed on October 18, 2021.

(11) Holly Firfer. CNN. (17.6.2004). *How Far Will Couples Go to Conceive?* <http://edition.cnn.com/2004/HEALTH/03/12/infertility.treatment/>, accessed on October 16, 2021.

(12) Khac Nam. (2.4.2015). *The child with “3 parents” will be legitimized*. <https://suckhoedoisong.vn/ua-tre-3-bo-me-se-duoc-hop-thuc-hoa-16994545.htm>, accessed on October 15, 2021.

(14) Alistair Barr. (27.4.2014). *Google’s New Moonshot Project: the Human Body*, Wall Street Journal. <https://www.wsj.com/articles/google-to-collect-data-to-define-healthy-human-1406246214>, accessed on October 16, 2021.