2020-12-23

Data: the Good, the Bad and the Ethical

John D. Kelleher
Technological University Dublin, john.d.kelleher@tudublin.ie

Filipe Cabral Pinto
Altice Labs, filipe-c-pinto@alticelabs.com

Luis M. Cortesao
Altice Labs, luis-m-cortesao@alticelabs.com

Recommended Citation

Kelleher, J., Pinto, F.C. & Cortesao. L. (2020). Data: the good, the bad and the ethical. . *InnovAction*, vol.5, pp. 70-79, 2020. Available at: https://www.alticelabs.com/site/innovaction/

Data: the good, the bad and the ethical

John D. Kelleher
Technological University Dublin
john.d.kelleher@tudublin.ie

Filipe Cabral Pinto
Altice Labs
filipe-c-pinto@alticelabs.com

Luís Cortesão
Altice Labs
luis-m-cortesao@alticelabs.com

Abstract

It is often the case with new technologies that it is very hard to predict their long-term impacts and as a result, although new technology may be beneficial in the short term, it can still cause problems in the longer term. This is what happened with oil by-products in different areas: the use of plastic as a disposable material did not take into account the hundreds of years necessary for its decomposition and its related long-term environmental damage. Data is said to be the new oil. The message to be conveyed is associated with its intrinsic value. But as in the case of real oil, we should take care to ensure that its use does not create harm in the future. We know from recent history that data can be used in harmful ways, but data also has enormous positive potential when applied to the service of communities. In this article, we highlight the opportunities, problems and best practice of using data.

1 Introduction

A[1] piece of data is a measurement of some type. For example, the height of an individual in centimetres, the number of items a customer purchased, the temperature at a location or the stock value at a certain time and date. The related concept of meta-data describes data about data, for example, the timestamp when a measurement was taken. Described in this way, data and meta-data may seem objective, primarily useful for generating reports, and benign. Despite these appearances, however, data is in fact subjective, can be used to do harm, and is a powerful basis

[1]This is a preprint of an article that appeared in InnovAction, please cite as: "Data: the good, the bad and the ethical", John D. Kelleher, Filipe Cabral Pinto, and Luís Cortesão. InnovAction vol. 5 pages 70-79, 2020. Available at: https://www.alticelabs.com/site/innovaction/

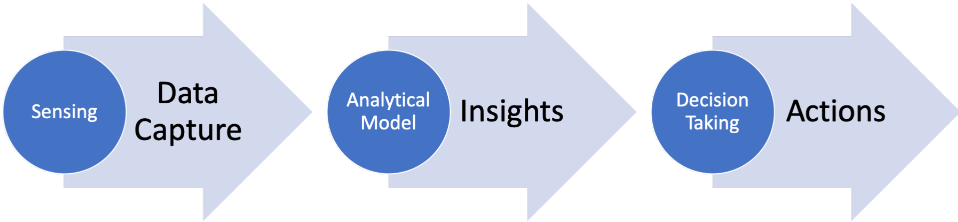

Figure 1: Data-driven decision pipeline

for decision making to drive action and to improve future outcomes. At a high level, the key stages in the data-driven decision pipeline can be understood as: data capture, analysis, insights and decisions (see Figure 1).

The subjectivity of data arises from the variety of decisions that go into its definition, capture, processing and interpretation. For example, there were a series of human decisions involved in the definition of the metric systems. Similarly, a large number of human decisions are involved in the design and deployment of sensors throughout our environment: who decides where video surveillance cameras should be located and why did they choose the locations they did? Furthermore, during the processing of data there are a large number of decisions that affect the outcomes of any data analysis process, concerning how data is cleaned, merged with other data, the questions used to frame the analysis, or the computational models and techniques used to extract insight from the data. And, finally, the contentious nature of interpreting data can be seen in the different views that can be taken on what policies should be put in place in response to the same data, for example, how should society respond to COVID-19 given its most recent health and economic statistics.

The same data sometimes spawns antagonistic interpretations. It is easy to observe the opposite perspectives on the same information when analyzed by members of the government or by the opposition. The same is true in the sports world: for the same fact, fans of a team always have an opinion that is the opposite of that of their opponents. These opposing conclusions are not always purposeful or malicious. Data reveals information that may lead to different conclusions, according to contexts, beliefs or experiences.

Although data is intrinsically historic (the fact that it is a measurement means that it always describes the past) and hence it may seem to be most suitable for reporting what has happened, the emergence of modern data-driven Artificial Intelligence (AI) systems that can make accurate predictions has unlocked the power of data to support decision making and so affect future outcomes. For example, a study of 179 publicly traded firms in 2011 found that firms that adopt data-driven decision-making processes have higher output and productivity than would be expected given their other investments [1]. Somewhat ironically, the power of Big Data and modern AI to drive successful decision making is also one of the key drivers for the growing awareness of the potential harm that can arise from data.

That is evident in the growing discussions on data ethics and new data regulations, such as the General Data Protection Regulation (GDPR). We will discuss these factors in more detail later in the article but, briefly put, the growing prevalence of sensors in modern societies and the tracking of individual behaviour in the online setting raises several important questions about personal privacy, civil liberties, targeted advertising and the rise of targeted misinformation. In the remainder of this article, we will address some of the emerging challenges and opportunities that Big Data and modern AI pose to individuals and societies.

2 The Good

Data can be, and should be, used for good. Data is the basis of a set of innovative businesses enabling creative services and operational optimizations. The evolution of wired and wireless communication systems was key to support digital transformation. The ubiquitous access to data enabled improved operational efficiency and reduced costs across all sectors of activity. The advent of IoT networks was fundamental to the creation of the Smart Cities vision, where data coming from networks of distributed sensors is used to model, in near real-time, and manage the evolving situation of urban space. Improved efficiencies in urban management have been enabled through IoT systems supported via modern telecommunications infrastructure. Similarly, the management of telecommunications networks themselves has also been improved through the use of distributed sensors and data-driven decision-making. For instance, Communication Service Providers (CSP) are using data-driven predictive maintenance to avoid equipment malfunctions and to improve network resource management. At the same time, the explosion of data about humans and society that occurred with the shift to online life and emergence of social media platforms, has enabled the development of innovative commercial consumer engagement services supported on Big Data. Examples of these types of innovation include recommendation systems associated with personalised advertising or cross-selling and even for churn detection and prevention.

Beyond the business opportunities, the access to open datasets about communities and cities, computational resources, and to open-source frameworks for data science create the conditions to use data to support and empower communities. For example, in recent years we've seen the emergence of a vibrant "Data for Good" community. A common theme across these projects is the use of data to positively impact the world and they can often be linked with the objectives of sustainable development, such as reducing poverty, preserving the environment or promoting a healthier life. Naturally, the positive use of data is not limited to these bottom-up grass-roots movements. Governments and international organisations are also keen to leverage the power of data for the public good. Indeed, some of the responses initiated by governments to the COVID-19 pandemic can be understood as using data for good and can be directly linked to the data capture, analysis, insight and decision pattern we described in the introduction. The rapid spread of the virus in the world population has generally forced nations to take exceptional measures to respond to the pandemic. The utilization of personal data in an anonymized way is the basis for new mobile applications aiming at helping to retard the virus propagation. As stated by Grantz et al. [2], mobile phone data can be used in the fight against COVID-19 as a non-pharmaceutical intervention. It can include location-based information, supplied by the CSPs (call detail records) or provided by the mobile GPS system; proximity data through Bluetooth; or even application data explicitly inserted by users. The collected data may be used in different ways, such as following the risk of importing the virus from a region, detecting patterns of mobility or for contact tracing to advise quarantine to potentially infected people.

3 The Bad

However, the promise of using data by governments and large organizations for good can also be a threat to civil liberties. Two arguments are often used to support the adoption of data-driven infrastructures and technologies throughout society. The first argument is that data can be used to improve the efficiency, effectiveness and competitiveness of systems, and the second argument relates to improving security, for example, governments often argue that increased surveillance improves security [3].

With regard to using data to improve efficiency, effectiveness and competitiveness, there is a large body of research that indicates that the more personalised advertising is to an individual, the more effective it is, e.g. [4]. Consequently, companies are encouraged to gather data about their customers in order to target and personalise their offers, thus improving the effectiveness of their advertising. However, although personalisation may appear desirable in many ways at a surface level, personalisation inevitably leads to marginalisation [5]. For example, the targeting of a special offer to one customer necessarily marginalises the customers who do not receive this offer. This form of data-driven discrimination is particularly blatant on websites that use differentiated pricing, where some customers are charged more than others based on their profile [6]. More broadly, data-driven personalisation can be understood as a form of social profiling, where those deemed to be useful are targeted with personalised offers and preferential treatment, and those deemed to be waste are marginalised and ignored.

Beyond marketing and commercial activities, data-driven decision making and AI is often framed as improving the efficiency of governments; for example, smart city technologies are marketed as using data to make public services more efficient and less costly [5, 7]. However, data-driven decision systems work by identifying patterns within data and using the identified patterns to generate output: if the data patterns reflect the prejudices of the society, then these prejudices will be reinforced by these "smart" systems. This systematic reinforcement of prejudice is particularly worrisome when data-driven systems are used for predictive policing or to inform judicial decisions [3].

The emergence of smart city technologies has led to a proliferation of sensors throughout modern societies, for example, video surveillance of roads, offices and shops. These potential sources of real-world surveillance are reinforced by the fact that modern digital technologies make it easier to track people through their mobile phones and credit card usage. Furthermore, as we mentioned above, in online setting individuals are tracked through the search terms they use, websites they visit and items they click on, in order to facilitate targeted marketing. Taken together, these different forms of surveillance mean that it has never been easier to track a person's movement and behaviour. This is obviously worrying from a personal privacy perspective, but the growing awareness by individuals that their behaviour is being tracked can have a self-disciplining effect that curtails personal freedom and has the potential to ultimately undermine democratic processes by diminishing our collective ability to act as political and social agents [8].

To summarise, unless care is taken in the development and deployment of data-driven systems, these systems may lead to marginalisation, the reinforcement of prejudice, and the curtailing of human freedom through the fear of surveillance. However, all of these concerns are related to the improper use of real data. Another set of concerns arise when we consider the growing amount of fake data being generated and distributed. It should be noted that not all fake data is problematic. Fake data can be created for good reasons, for example, fake data is often used for testing systems, or to automatically fill in different forms. Sites such as GenerateData.com provide applications to generate false data. Even Python programming environment provides the Faker library to generate dummy data to applications. Although these data are not true, their use does not imply malevolent usage. However, false data can be created from scratch to distort reality. Its usage can be associated with cheating to make money or with the aim to intentionally harm a person, organisation or even a country.

Disinformation is false information deliberately created and distributed to damage the image of a person or entity. It affects society as it shapes collective minds, even undermining democracy as we know it. The rapid spread via digital platforms, such as Facebook, Google or Twitter, makes them reach everywhere and persist for eternities, creating alternative truths, sometimes injuring the truth of death. Donald Trump used the term fake news to describe news that hindered his candidacy for the presidency of the USA. But fake news has become popular as news that intends to affect the truth to make gains based on lies. The Cambridge Analytica scandal highlighted the weaknesses of democracy. The exploitation of psychological profiles from Facebook data combined with profile-based political ads, apparently exempted from the fact-checking, seems to have influenced the results of Brexit as well as the last presidential election in the USA. In fact, the post-truth seems to be gaining ground in many nations: public opinion is being shaped more by emotional appeals in the form of ads than by objective facts spread by credible sources.

In a world of growing disinformation, the spreading of AI deep learning technologies that can generate deep fakes is a worrying development [9]. Deep fake systems can combine images and sound to create fake videos of people, that are

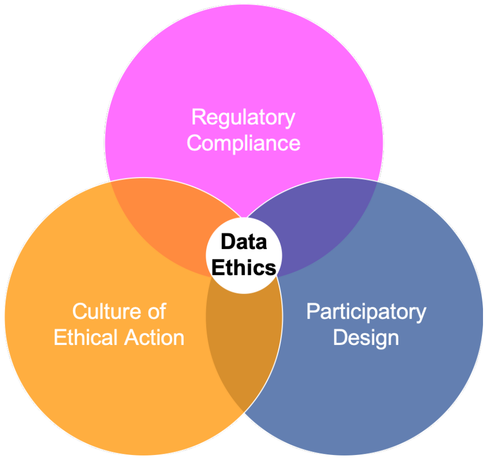

Figure 2: The three key pillars of the ethical use of data

very difficult to distinguish from real videos. The capacity to create a realistic fake video of a person moving and speaking brings the idea of fakes news to a whole new level of danger. It becomes possible to easily undermine the public image of a person or a social group; for example, election results can be distorted by the appearance of a fake video the day before the elections purportedly showing a candidate doing something illegal, such as receiving bribes. The spread of available tools to create false video content makes the digital world a potential battlefield requiring strong policies and ethics in data to avoid chaos.

4 And the Ethical

In this section we address three of the key pillars of the ethical use of data: understanding and compliance with data regulations, the creation of a culture of ethical action within an organisation, and the engagement with stakeholders and the communities potentially affected by data usage and modern AI data-driven technologies throughout the design and development of these technologies (see Figure 2).

Legal frameworks concerning data usage vary across jurisdictions. However, the majority of legal frameworks contain regulations relating to anti-discrimination and also personal data protection. Most anti-discrimination regulations forbid discrimination based on any of the following protected categories: disability, age, sex, race, ethnicity, nationality, sexual orientation and religious or political beliefs. Consequently, apart from special contexts such as medical assessment, data relating to these categories should not be used as the basis for decisions relating to an individual. Complying with this restriction can be more difficult than it might first appear, because often these protected categories can be encoded in data through proxy variables. For example, including an individual's address within a dataset may inadvertently make it possible to predict their race or ethnicity. Furthermore, when datasets are merged, the possibility of this protected types of information becoming identifiable through the combination of features from the merged datasets often becomes feasible (data re-identification). Consequently, care needs to be taken both in the design of datasets, their curation and in the testing of any technologies built using these datasets, to ensure that the resulting decisions driven by the technology are not biased by one or more of these categories. It is also important to highlight that this bias can occur at a group level or individual level. For example, at the group level a system might systematically be biased towards a particular race or ethnicity. However, even if it can be demonstrated that, on average, the decisions made by a system are not biased towards a particular category, it is still problematic if, for a particular individual, the system uses data relating to one of the protected categories to decide for that individual. This is why it is so important to understand how modern AI data-driven systems make decisions, how these decisions are distributed across different communities of people and what data a system accesses to and uses when making a decision about an individual. These are the questions at the core of research fields such as Explainable AI.

With respect to the use of personal data, probably the most significant recent development has been the General Data Protection Regulations (GDPR), from the [10]. The GDPR are legally enforceable across all EU member states; however, perhaps the most broadly accepted personal privacy principles are the Guidelines on the Protection of Privacy and Transborder Flows of Personal Data [11]. Indeed, the GDPR can be traced back to these OECD guidelines. In these guidelines, the concept of personal data is defined as data relating to an identifiable individual, known as the data subject, and the data controller determines the purposes for which and the means by which personal data is processed. There are eight core principles set out in the OECD guidelines:

1. **Collection Limitation Principle**: personal data should only be obtained lawfully and with the consent and knowledge of the individual

2. **Data Quality Principle**: personal data that are collected should be accurate, complete, up to date, and relevant for the purpose for which they are used

3. **Purpose Specification Principle**: at or before the time of collection of data relating to a data subject they should be informed of the purpose for which the data will be used.

4. **Use Limitation Principle**: the use of the data is limited to the purpose that the data subject was informed of, and the data should not be disclosed to third parties with consent from the data subject or by authority of law.

5. **Security Safety Safeguards Principle**: personal data should be protected by security safeguards against theft, deletion, disclosure, modification, or unauthorized use.

6. **Openness Principle**: a data subject should be able to find out with reasonable ease how data relating to them is collected, stored and used.

7. **Individual Participation Principle**: a data subject has the right to access and challenge personal data
8. **Accountability Principle**: a data controller is accountable for complying with the principles.

Although these principles are relatively clear in their meaning and intent, it may be difficult to translate these high-level principles into a culture of ethical action and practice within an organisation [7]. A number of professional bodies have developed guidelines to help their members in translating regulatory principles to practice. For example, the IEEE has produced a call to action for businesses using AI entitled Ethically Aligned Design [12]. This call to action is relevant to the ethical use of data because data is at the heart of modern AI systems. The EAD highlights the value of necessity foregrounding ethical considerations throughout AI technology and data-driven organisations. In particular, developing an ethics-based culture and implementing ethics-based systems and practices within an organisation is the basis for building trust with investors, stakeholders, employees and customers. The EAD suggests a two-stage process for organisations to develop and sustain a culture of ethical practice. In the first phase, people are introduced to ethical concepts relating to AI design and data usage at scale. This first phase includes working with executives to identify the core values and ethical principles of the organisation and launching a communication and training campaign. The second phase involves helping, supporting, and incentivising people to understand and apply these new concepts within the local context of the organisation that they work in. This may include the identification and training of a core-team of strategically positioned employees who can provide local support and evangelise the importance of ethical decision making. It also involves emphasising the consideration of ethical implications as a core function of each person's role and incentivising people to make ethical decisions, the goal here being to move people from awareness of ethics to ethical action.

The widespread use of AI and data-driven systems throughout modern societies means that everyone involved in their design, development and use should be mindful that technology is never neutral and so they should consider and be aware of the ethical implications of any technology. Furthermore, they should not only consider short-term impacts but also long-term impacts and how a technology might affect a future society. Consequently, a technology should be designed and developed so as to align with the values of the society it affects. This means that human needs and the protection of human rights should be at the core of the design and use of AI technology and data. Adopting a human and value-centred approach to the design of a technology requires the ability to understand and empathise with the members of the community that will be affected by that technology. The best way to develop this understanding is to engage with the community and this is why there is a growing need for stakeholders and community engagement throughout the technology development lifecycle. A useful approach for achieving such an engagement is the adoption of participatory design (or co-design) concept. In participatory design, the end-users and those potentially affected by a technology are invited to work with technology designers and developers during the innovation process. Often this engagement happens during the initial problem definition and solution design, and also later during the development to evaluate the solutions as they are implemented and iterated.

5 Conclusions

Data by itself is neither good, nor bad, nor unethical. It is the use that data is put to which is important. Properly used data can improve business, the management of cities and sustainable development, and help to control diseases. On the other hand, fake data can be used to distort the truth and undermine democracy, and large scale data ecosystems may enable surveillance and threaten civil liberties. The use of data, particularly when it involves personal data, raises a set of ethical issues, which affect all the data lifecycle. Responsible utilization of data must prevail regardless of the context in which it is being used. In addition to the processes related to the collection and storage of data, more critical issues arise when data is associated with artificial intelligence mechanisms to extract knowledge and predict an individual's behaviour based on the data collected about them without their awareness. Besides strong regulation, organisations must proactively work to develop a culture of ethical action and also engage with external communities and stakeholders to ensure that ethical practice is at the heart of technical innovation and data usage. Data is like pills: when used according to the rules they can be highly effective, but if used without any control they can be harmful, and can even cause irreversible damage. Data, regardless if it is true or false, can be applied for a meritorious purpose or to misrepresent the truth. In the end, it isn't the data, it is how it is used!

6 Acknowledgements

Some of the findings mentioned in this article result from research conducted at Technological University Dublin with the financial support of the ADAPT Research Centre which is funded by the SFI Research Centres Programme and is co-funded under the European Regional Development Fund (ERDF) through Grant #13/RC/2106.

References

[1] E. Brynjolfsson, L. M. Hitt, and H. H. Kim, "Strength in numbers: How does data driven decision making affect firm performance?" *SSRN Electronic Journal*, 2011.

[2] K. H. Grantz, H. R. Meredith, D. A. Cummings, C. J. E. Metcalf, B. T. Grenfell, J. R. Giles, S. Mehta, S. Solomon, A. Labrique, N. Kishore, C. Buckee, and A. Wesolowsk, "The use of mobile phone data to inform analysis of covid-19 pandemic epidemiology," *Nature Communications*, vol. 11, no. 4961, 2020.

[3] J. D. Kelleher and B. Tierney, *Data Science*. MIT Press, 2018.

[4] S. Hill, F. Provost, and C. Volinsky, "Network-based marketing: Identifying likely adopters via consumer networks," *Statistical Science*, vol. 21, no. 2, pp. 256–276, 2006.

[5] J. D. Kelleher and A. Kerr, "Finding common ground for citizen empowerment in the smart city," *Ethics and Politics*, vol. 22, no. 2, pp. 33–61, 2020. [Online]. Available: http://hdl.handle.net/10077/30978

[6] T. Walker, "How much? the rise of dynamic and personalised pricing," *The Guardian*, 2017. [Online]. Available: https://www.theguardian.com/global/2017/nov/20/dynamic-personalised-pricing

[7] A. Kerr, M. Barry, and J. D. Kelleher, "Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance," *Big Data & Society*, 2020. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/2053951720915939

[8] N. Fitzpatrick and J. D. Kelleher, "On the exactitude of big data: La bêtize and artificial intelligence," *La Delueziana*, pp. 142–155, 2018. [Online]. Available: http://www.ladeleuziana.org/wp-content/uploads/2018/12/Fitzpatrick-Kelleher.pdf

[9] J. D. Kelleher, *Deep Learning*. MIT Press, 2019.

[10] *General Data Protection Regulations of the European Council and Parliament*, European Council and Parliament. Official Journal of the European Union L 119: 1-2016., 2016.

[11] *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. Organisation for Economic Co-operation and Development, 2013.

[12] *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019.