VERITAS, Nº 46 (agosto 2020) 13-35 ISSN 0717-4675

The ethical imperatives of the COVID-19 pandemic: A review from data ethics

GABRIELA ARRIAGADA BRUNEAU*
University of Leeds (United Kingdom)
prgcab@leeds.ac.uk

MARK GILTHORPE**
University of Leeds / Alan Turing Institute (United Kingdom)
m.s.gilthorpe@leeds.ac.uk

VINCENT C. MÜLLER***
Eindhoven University of Technology (TU/e) (Netherlands)
University of Leeds / Alan Turing Institute (United Kingdom)
v.c.muller@leeds.ac.uk

Abstract

In this review, we present some ethical imperatives observed in this pandemic from a data ethics perspective. Our exposition connects recurrent ethical problems in the discipline, such as, privacy, surveillance, transparency, accountability, and trust, to broader societal concerns about equality, discrimination, and justice. We acknowledge data ethic's role as significant to develop technological, inclusive, and pluralist societies.

Key words: Data ethics, data governance, COVID-19, vulnerability, equality, transparency, privacy, digital divide, accountability, fairness, digital epidemiology.

Los imperativos éticos de la pandemia de COVID-19: Un análisis desde la ética de los datos

Resumen

En esta revisión, exponemos algunos de los imperativos éticos observados desde la ética de datos en esta pandemia. Nuestra exposición busca conectar problemas éticos típicos dentro de esta disciplina, a saber, privacidad, vigilancia, transparencia, responsabilidad y confianza, con preocupaciones a nivel social relacionadas con la igualdad, discriminación y justicia. Consideramos que la ética de datos tiene un rol significativo para desarrollar sociedades tecnificadas, inclusivas y pluralistas.

***Palabras clave**: Ética de datos, gobernanza de datos, COVID-19, vulnerabilidad, igualdad, transparencia, privacidad, brecha digital, responsabilidad, justicia, epidemiologia digital.*

---

[*] PhD(c) in Philosophy, University of Leeds, UK. Postgraduate researcher in Artificial Intelligence and Data Ethics at the Inter-disciplinary Applied Ethics Centre (IDEA), University of Leeds. This work is part of my doctoral studies funded by the National Agency for Research and Development (ANID) Scholarship Program, DOCTORADO BECAS CHILE 2018 – Folio 72190086.

[**] Professor of Statistical Epidemiology in the School of Medicine and the Leeds Institute for Data Analytics (LIDA) at the University of Leeds. Latest publication: "Evidence informing the UK's COVID-19 public health response must be transparent" (2020).

[***] Professor of Philosophy at the Technical University of Eindhoven (TU/e). University Fellow, IDEA Centre and LIDA at the University of Leeds. Latest publication: "Ethics of Artificial Intelligence and Robotics" (2020).

Recibido: 10/Junio/2020 - Aceptado: 24/Julio/2020

INTRODUCTION

In this data ethics review, we highlight ethical problems magnified by the circumstances of the pandemic and link them to broader societal worries about fairness and equality.

This crisis demands swift decisions. However, 'the best possible choice' can have multiple ethical, cultural, technical, and even political dimensions. So, what does the best *ethical* choice imply? To clarify the implications of this question, we explore some *ethical imperatives*, i.e. worries that echo essential ethical concerns, aggravated and deepened by these critical circumstances. These imperatives represent two main areas. One area, relates to issues dealing with *trust* and *responsibility*, connected to privacy concerns, data handling[1], and transparency. And the other area considers issues of *justice*, including fair treatment, discrimination, and societal inequality.

Unique elements in this scenario, such as a globalized and unavoidable exposure to the pandemic, can widen our moral perspective when conflictive trade-offs are discussed. This allows us to explore warnings, insights, and lessons on how to achieve the best ethical choices possible and show that data ethics goes beyond the regulation of technicalities; it is a fundamental element in the pursue of *ethical* technified societies.

1. DATA ETHICS: DEALING WITH UNCERTAINTY

In the midst of this pandemic we find ourselves surrounded by uncertainties that escape the ordinary. This creates an interesting scenario that magnifies ethical dilemmas and pre-existing risks. Any disruptions encountered perturb perceptions of safety, normally refuged under the limitations of ordinary lives. This uncertainty affects data-driven endeavours that offer 'solutions' to this crisis as well. However, data and technology can only offer partial redress, to which we *must* add our own adaptations, presenting further challenges. Here we review some of these challenges whilst highlighting related ethical imperatives.

A. Ethical concerns for privacy: the COVID-19 tracing apps

In the middle of the COVID-19 pandemic, tracing apps quickly became a focus of attention because of their privacy risks (Dubov & Shoptaw, 2020). Efforts to take care of pressing factors speeded-up initiatives

---

[1] Data handling here refers to all instances of administration of data, including collection, analysis, interpretation, and distribution.

for increased monitoring. Hence, a trade-off was presented as *necessary*: 'privacy vs data utility' or in other words 'privacy vs common good', i.e. "should we allow widespread and severe risks to public health vs enabling invasive and widespread risks to privacy in the form of corporate or government surveillance?" (Feys, 2020). But this is a false dichotomy, there is *no choice of either/or*, only a misdirection.

The invasiveness of COVID data apps is profiled by some governments as a price worth paying, enforcing them as mandatory (Mozur et al., 2020); cultural and legal differences allow some countries (mostly eastern) to use identifiable and personal data to develop and implement these apps. However, in most western countries, the use of *personal data* is a breach of data privacy regulations. As a response, risks are said to be 'kept under control' by *anonymizing* the data or using *aggregated data* to prevent privacy issues. But these claims are mistakenly reassuring the public about their privacy (Seng Ah Lee, 2020). Furthermore, there are concerning long-term implications. Implementing these apps in critical circumstances can serve as a precedent for data-driven policing, jeopardizing an ethical setback by justifying processes "which risk undermining fundamental values and rights […][that] may be difficult to revert after the pandemic" (Taddeo, 2020: 172).

Accordingly, a fair amount of discussion is dedicated to the risks of re-identification. The main 'ethical slogan' for tracing apps in western countries has been protecting the users' privacy, keeping their information safe by using 'unidentifiable data'. Nonetheless the risk for re-identification is a threat regardless. A study from last year (Rocher et al., 2019) presents a model that successfully re-identifies 99.98% of Americans in any dataset using 15 demographic attributes. The likelihood for individual re-identification can be estimated with high levels of accuracy, "even when the anonymized dataset is heavily incomplete" (Rocher et al., 2019: 2) confirming that: re-identification is a real practical risk; sampling partial datasets does not provide plausible deniability; and even with very low population uniqueness (often considered an argument for labelling a dataset anonymous), many individuals risk to be re-identified.

Rocher et al.'s study questions the effectiveness of using aggregated or anonymized data to protect privacy and it shows that under the GDPR (Protection Regulation (GDPR) Council of European Union. Regulation (EU) 2016/679. *Off. J. Eur. Union* L 119, 1–88) and the California Consumer Privacy Act (CCPA) Cal. Civil Code. Assembly Bill No. 375 §§ 1798.100–1798.198 (2018), basic distinctions for anonymization are insufficient. To meet the standards of anonymization *every* person in the dataset has to be protected. Anonymized data implies that the data subject is *not* *identifiable* by any means. In contrast, pseudonymized data makes re-identification possible, irrespectively of how the process is reversed. Pseudonymized data then *is* personal data.

But these distinctions have been overlooked in practice; an example is the failed Isle of Wight tracing app project in the UK. This project presented a way to *anonymously* monitor users via Bluetooth when they came in 'contact' with each other (NHSX, 2020), with the clear intention to reassure the protection of the public's identities. Addressing potential concerns with privacy, the NHSX released a data protection impact assessment (DPIA)[2] in which they stated that the application: "anonymously records distance over time" and it is "designed to preserve anonymity". Yet, in the same document, it was said that "the data collected by the App is pseudonymised – albeit having never been directly identifiable there is no lookup to users' identities". But according to the GDPR (recital 26)[3], pseudonymisation is not a method of anonymisation. Pseudonymisation is a tool (a recommended security measure) to reduce potential linkability of a dataset with original identifiers of a subject; but it does not guarantee *anonymity*. Therefore, under the rules of the GDPR, what constitutes *personal data* includes profile identifiers. An IP address falls into that category, hence an ID (Installation ID for this app) is also identifiable, i.e. a risk element for re-identification.

The distinctions between pseudonymous and anonymous data must be categorical. In this case, there is an ethical imperative for transparency and clarity when informing these differences. The NHSX presented this app as *anonymous*, but: "the NHSX app does not preserve the anonymity of users, as it primarily processes pseudonymous, not anonymous, personal data" (Veale, 2020: 3). The misuse of the concept 'anonymous' goes beyond a legal technicality, it affects the *trust* of users and the *accountability*, in this case, of the government. Pursuing data-driven technologies requires establishing links of trust with its users, and in a larger scale with society. In this case, "the data in the NHSX app is 'capable' of revealing an individual's identity"(Veale, 2020: 4), and there is a correspondent ethical imperative to inform this risk.

General (non-specific) reassurances for data handling are insufficient to address claims for accountability. The design and development of these

---

[2] Available at: https://faq.covid19.nhs.uk/DPIA%20COVID-19%20App%20PILOT%20LIVE%20RELEASE%20Isle%20of%20Wight%20Version%201.0.pdf

[3] *Personal data which have undergone pseudonymization, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person.* Full recital access https://www.dsgvo-portal.de/gdpr_recital_26.php

apps, particularly in contexts of uncertainty, must acknowledge any potential secondary uses and complement them with evaluations, auditing, and monitoring. Asking a set of straightforward questions about the project, such as: "is it necessary?, is it proportionate?, is it sufficiently effective, timely, popular and accurate?" (Morley et al., 2020: 30).

The failure of some tracing apps[4] has been a common trend when security measures are not met, in a rush to implement the technology as fast as *technically* possible. The hurry to understand on virus-spreading might be well-intended but overlooking ethical criteria can leave these projects ineffectual. This has been the case for various attempts around the world, that failed to inform and/or prevent security risks, making them untrustworthy. These data-driven projects rely on active data provided by users, meaning that they are only *as powerful as their reach* (Goodes, 2020).

Contrasting this lack of success there are also promising projects (it is too soon to talk about success). Germany developed its own tracing app 'Corona-Warn-App'[5] and up to the date of this publication it is fully functional. The app took longer than other European initiatives to be developed, arguing that privacy distinctions were thoroughly considered to align with data regulation and ethical concerns, aiming to get the needed trust from the public. Interestingly, much like the UK app's failed attempt, its German counterpart also uses Bluetooth to alert of nearby contact. The data, however, is stored on each phone (decentralized approach) using an 'Apple-Google system' (APIs) to notify users. The same decentralized features apply for the Italian app 'Immuni'[6] and the Swiss app 'SwissCovid'[7]. The companies involved in the development of these apps opted for an *open source* approach, which not only allowed broader collaboration, it also made their code available for *all* to see.

Data-driven technologies require data sources, often in high volume, to increase their impact and efficiency. Thus, establishing a strong foundational bond of trust with the public can make a project successful. Failure, then, initiates from *ethical* and not *technical* reasons. As a consequence, tracing apps can be functional and accurate in its technical design but become completely useless in practice. These technologies require direct collaboration from the public, and without an ethical frame shaping its design, trust fades. Policymakers and authorities "must have a realistic understanding of what data produced by individuals' mobile phones can and

---

[4] Some failures also include high surveillance issues like the apps developed in Kuwait, Bahrain, and Norway (Amnesty International, 2020).

[5] Available at: https://github.com/corona-warn-app

[6] Available at: https://github.com/immuni-app/immuni-documentation

[7] Available at: https://github.com/DP-3T/dp3t-app-ios-ch

cannot do. […] There is a danger that simplistic understandings of how technology works will lead to investments that do little good, and that invade privacy without producing commensurate benefits" (Stanley & Granick, 2020).

Another concern is the use of *centralised systems* for these apps, instead of decentralized versions (data stored on individual devices) like the DT-3P (Decentralized Privacy-Preserving Proximity Tracing) developed by a group of European universities and with open documents available on GitHub[8]. Using government (centralised) data servers raises risks of function creep[9] and state surveillance, increasing the public's distrust and questioning the ability of centralised based apps to keep their data safe (Zhang et al., 2020). Moreover, it gives room to examine if there is any real benefit in considering centralised approaches or if there are other interests at stake: "[…] there are no advantages to centralized contact-tracing solutions in a democratic society that respects human rights, apart from convenience for the state" says David Muramaki-Wood (Sweeney, 2020: 303).

But there is a risk of framing the ethical debate based on technicalities, which can quickly transform this discussion into a "techno-solutionist" argument where "[…] a lot of really important issues are marginalized by this framing, such as how people who failed to install these apps might be discriminated against, especially those who are already vulnerable or who do not have a lot of agency" argues Lilian Edwards (Sweeney, 2020). Following these worries, the potential limitation of civil liberties (Taddeo, 2015) becomes an unavoidable topic: "there is legitimate concern from civil liberties advocates that the government or corporate entity holding the database of the aggregated information might abuse the power that comes with having such information" (National Security Commission on Artificial Intelligence (NSCAI), 2020: 8). History tells us that extreme social events can change what we deem *reasonable* or *morally acceptable*.

Pushing data-driven technologies in this crisis scenario can be a reasonable response, but only if used with *caution*: "before embracing new forms of surveillance to address the coronavirus, we must ensure that any such responses are proportionate and grounded in evidence" (Diaz, 2020). If we are not careful, the risk of establishing processes that undermine fundamental values and rights (Taddeo, 2020) is eminent. We must keep in mind that: "the quality of contract tracing in a certain population *inherently depends on factors orthogonal to the technological aspects*, which in turn raises

---

[8] Available at: https://github.com/DP-3T/documents

[9] Function creep refers to information being re-purposed from its original use (i.e. its specified purpose).

important questions about fairness" (Chan et al., 2020: 15) (emphasis added).

Privacy and security issues disclose further concerns about fairness. Proper implementation of data-driven technologies implies technical, legal, and ethical distinctions that must ensure trust and accountability. However, privacy and security risks are connected with the *vulnerability* of individuals and social groups, inciting us to use data-driven technologies to *empower* our society, instead of using their power to increase or perpetuate said vulnerabilities.

In the upcoming section we discuss the ethical imperative of fairness by reviewing concerns about discrimination and inequality: the increased data divide, the worrying effect of the pandemic on BAME groups, and the exposure of vulnerability as a moral token.

B. Governing through technology: learning from the pandemic

Fairness, as an ethical imperative of data-driven solutions for social emergencies in the pandemic, is a response to structural social injustices. In this acute situation, we are able to experience —in an imperative fashion— the inherent fragility of our institutions and governments, testing their decision-making abilities and their capacity to adapt and overcome pressing ethical dilemmas. At the same time, issues that would have been otherwise secondary become a priority.

At the end of the previous section reviewing privacy, transparency, and security issues, we identified a risk for asymmetric use of power. Citizens are put in a position of vulnerability when their lack of tech-savviness or technological literacy gives them a disadvantage. Lack of access and knowledge of technologies and digital resources[10] illustrates this inequality. Important parts of the population are actively *ignored* for not being 'on the grid'[11] increasing the *digital divide* (Allmann, 2020).

---

[10] In a World Economic Forum's article, the digital divide is exposed in numbers, showing that almost half of the world population *does not have internet access*; the price for browsing data in Africa can be as costly as 40% of the monthly wage; and in developed countries, important parts of the population have limited access, over 6% in the USA and 13% in Australia. For more see (Broom, 2020).

[11] Smartphone ownership varies among countries. In 'advanced' countries for example nine in ten south Koreans have access to smartphones, whereas this number reduces to six in ten in Greece or Poland. In 'emergent' countries, by contrast, ownership can go from nearly 60% in Brazil to 24% in India. Additionally, these percentages also relate to generational (higher levels amongst younger population) and educational factors (Silver, 2019).

The digital divide is known as a form of "socio-technological inequality that exacerbates, and is exacerbated by, other forms of inequality: socio-economic, geographic, and geopolitical" (Aggarwal et al., 2020). And this pandemic presents additional challenges by imposing remote work, learning, socializing, and the switch of many services to online setups. Technologies have offered prompt solutions to pressing needs, but looking at the bigger picture, this technological switch turned this digital divide into an ethical imperative.

This access disparity means that those alienated from these technologies because of geographic, economic, generational, or educational reasons, are effectively discriminated against, often on top of other systemic injustices. Switching to online education, for example, implies a smooth transition for privileged schools that had hardware, software, and proper training before the pandemic. But this reality is for the few, most schools had precarious facilities in the first place, evidencing educational and territorial inequalities in this pandemic (Cullinane, 2020; LaFave, 2020; Reyes, 2020). And in many cases schools provide more than just educational services; for some children school means having access to a meal (Bauer, 2020). The same has been observed in the forced transition into telehealth, where seemingly innovative programs were not being as successful as intended: "the diminished accessibility to technology based on various societal and social factors […] was being exposed at a critical time in a public health crisis. Frighteningly, there were no measures at the ready to address it" (Ramsetty & Adams, 2020: 1147).

The ethical imperative in this case reinforces the claim for a *fairer* data-driven and technological development: "The digital divide is not merely an unfortunate by-product of the digital revolution: it is one of the major problems undermining a fair information society for all, affecting the very possibility of universal, full citizenship, and some of the necessary conditions for a better life" (Aggarwal et al., 2020). This pandemic is transforming the crisis into a paradigm for data-driven policing that will change the way our society functions, however we cannot expect efficient and fair data-driven technology helping with public safety and further development, if we do not have inclusive (*reliable*) results.

Ethical concerns with fairness and equality are not limited to lack of access to technology. Consider for example the latest statistics showing that COVID-19 affects more people from a BAME (Black, Asian and Minority Ethnic) background. A study by the NHS in collaboration with Cambridge University, shows that in comparison to the overall population, individuals from a BAME background were more likely to be diagnosed, admitted to a hospital, and die from this disease (Alaa et al., 2020).

Data released by the ONS (Office for National Statistics) in the UK provide evidence for a death risk of COVID-19 twice as likely for Chinese and 'other' ethnicities, and more than *four times* more likely for black ethnicity (Elwell-Sutton et al., 2020). And the latest report given by the C.D.C (Centers for Disease Control and Prevention) in the US shows that Latino and African-American residents are *three times* more likely to become infected in contrast to their white neighbours (Oppel Jr. et al., 2020) across states, regions, rural, and urban areas.

These analyses signal (unsurprisingly) to systemic entrenched inequalities. Minority and segregated populations have suffered measurable inequalities[12] for long enough to turn external associated risk factors into a constant risk variable for them, particularly in relation to care (Egede, 2006; Fennell, 2005; Riley, 2012) and SES (Socioeconomic status) or income inequality (Adler & Newman, 2002; Kondo, 2012; Pickett & Wilkinson, 2015; Subramanian & Kawachi, 2004). Moreover, previous pandemics have given us data (Hutchins et al., 2009) showing disproportionate impact on ethnic minorities and socioeconomically disadvantaged populations (Niedzwiedz et al., 2020). So why are we, once again, having a differential impact?

These inequalities, which include the general struggles that come with poverty, i.e. accessibility to healthcare; increased risk of malnutrition or obesity; overcrowded living areas (an important factor to enforce social distancing); and monetary struggles (being forced to work and expose themselves), in combination with pre-existing health issues, do not fully explain this elevated risk (Niedzwiedz et al., 2020). Thus, further research and reforms are imperative to make health systems more responsive to the *highly documented* needs of these minorities: "deliberate care must be taken to enact policy changes both inside healthcare facilities and at all government levels" (Myers, 2020: 7).

These factors limit the possibilities of individuals and groups to have a choice, their autonomy is bounded by the constraints of their social and economic conditions, making them *extremely vulnerable* not only to this pandemic, but also to any structure of power. The treatment given to data is crucial to understand the relations and real impact of these variables at a population level: "Given the complexity of the systems[13] that produce poor outcomes for black and ethnic minority groups, there is a real risk that the imprudent use of statistical adjustment techniques in studies of COVID-19 deaths may obscure the role of some upstream issues" (Elwell-Sutton et al., 2020). A proper analysis can help establish needed moral

---

[12] See (Marmot et al., 2020) for an report of the latest 10 years on Health Equity.

[13] See (Rutter et al., 2017).

tokens. By moral tokens we mean a symbolism — in this case — of inequality, often seen in cases of sexism, genderism, colourism, and racism in association with a state of vulnerability. Establishing these moral tokens can actively prevent overlooking at data and treatment gaps in our society, and accounting for them on decision-making and policy design processes.

Responsible governance then requires (non-exhaustively) elements of privacy, transparency, fairness, equality and a non-naïve trust in data-driven technologies. Linett Taylor from Tilburg University presents some considerations from a *governance* standpoint, to address the uses of data-driven technologies in the efforts to fight the pandemic (*Beyond the Exit Strategy: Ethical Uses of Data-Driven Technology in the Fight against COVID-19*, 2020):

1. *Honesty about the gaps in the data, in an ongoing way*: being clear on what we don't know and having transparency about mechanisms for ensuring systems aren't being built on inaccurate data, and can adapt as our understanding of the data and its accuracy may change.
2. *Guarding against function creep*: we need to ensure our actions are relevant to the current state of affairs in terms of public health and safety during COVID-19, and update to match as this changes over time to prevent unnecessary ongoing measures or security theatre.
3. *Structured protections against irresponsible technology*: that go beyond GDPR and personal data, and see checks and balances via a meaningful role for civil society organisations in protecting privacy and critical civil rights, as well as ensuring representation and diversity, globally, of those involved in the design and development of the technology.
4. *Balance between centralisation and decentralisation*: a key tension is that good privacy technology is decentralised, whilst good government sees centralised information that enables judicious action; we will need skills and organisations to broker these perspectives.
5. *No tech solution should be seen as a political solution*: we will remain in a state of vulnerability and the idea that tech can truly and wholly protect us, is something more dangerous than bad politics or bad government.

This pandemic produces phenomenological experiences that allow us to see these ethical demands as *compelling* and indispensable. One of those experiences is our *shared vulnerability* (exposure to the pandemic), which allows us to contrast the impact of this exposure in different populations, revealing ethical imperatives in an axiomatic manner.

The vulnerability referenced here focuses on the limitations on self-sufficiency and the position of being burdened by factors outside our immediate control. This follows the idea of vulnerability as universal and inherently ontological to the human condition (Fineman, 2016; Shildrick, 2002; Turner, 2006). More specifically, it follows the taxonomy presented by (Mackenzie et al., 2013) including two sources of vulnerability: inherent and situational, which can be either dispositional or occurrent. In simple words, inherent vulnerability considers features of the environment in which we are born (live in), and situational vulnerability is an outcome of genetic, social, and environmental influences. Dispositional refers to a potential state of vulnerability in contrast to an actual state represented by the occurrent one. Consider the following example given by the authors:

> […] all fertile women of childbearing age are dispositionally vulnerable to life-threatening complications in childbirth. But whether or not a pregnant woman is occurrently vulnerable to such complications will depend on a range of factors, both inherent and situational, such as her physical health, medical history, socioeconomic status, geographical location, access to health care, and cultural norms relating to pregnancy and childbirth. The dispositional–occurrent distinction serves to distinguish vulnerabilities that are not yet or not likely to become sources of harm from those that require immediate action to limit harm. (Mackenzie et al., 2013: 8)

The purpose of these distinctions is to show how an accurate recognition and portrayal of these vulnerabilities in data-driven research can help addressing the moral imperatives associated with them. Recognizing the moral tokens at stake allows to conceive and design moral obligations related to these inequalities (Sloane, 2019) into data-driven endeavours, by reviewing the existing exploitation of vulnerabilities based on evidence (e.g. the greater exposure of BAME population, or the privacy settings in data apps), and/or foreseeing potential vulnerabilities observable as trends based on occurring events (future impact). These apply to the demanding circumstances of a pandemic, but they should also be considered an essential element of data ethics more broadly.

C. Moral perspective for trustworthy decision making

To plan and develop *ethical* digital societies further philosophical distinctions motivated by these ethical imperatives are required. A recent example presented by (Floridi, 2020) distinguishes between *validation* (are we building the right system?) and *verification* (are we building the system in the right way?) referring to the risks of COVID-19 tracing apps. The twofold relation between both concepts, as presented by Floridi, tells that a *lack of validation* demands us not to build an app, and a *failed verification* forces us to re-consider validation, i.e. question its necessity. This is not to say that in some scenarios severing validation and verification is not ethically justified, it can be (Floridi, 2020), but this pandemic is not one of those cases. Often, authorities consider that reacting is enough, i.e. 'attempting' to do something in case it is successful, and if it is not at least *something* was done. This *could* (arguably) be justified —as highlighted by Floridi— if costs can be externalized, but in this pandemic, verification should feedback on validation, i.e. if one fails, the whole project needs to be reconsidered to be improved or discarded. Establishing these fundamental philosophical distinctions allows to perform within the bigger picture, achieving levels of transparency and accountability that can build the necessary trust with the public, but most importantly, it means developing data-driven policies that are advantageous and fundamentally *virtuous*.

Ethical imperatives emerging from vulnerability and asymmetry of power between individuals, groups, and institutions, for example, trace how we can cherish technology within digital humanities as an essential element in our societal transformation. Data ethics does not only refer to the limitations and concerns of data collection, manipulation, and/or implementation of data-driven technologies/solutions, it also implies inquiring about *what can be missing* (e.g. data gaps), *what can be overlooked* (e.g. keep it contextualised), and *what are the potential risks* (e.g. verifying/validating) of pursuing any possibilities.

This *carefulness* (i.e. ethical inquiry) is helpful particularly when forced to deal with uncertainty. We might be compelled to provide immediate answers but not having a blind trust in data-driven solutions allows us to recognise that those *expected answers* are not necessarily available. Countries around the world had radically different responses[14], including (but not limited to) testing support, types of lockdowns (schooling, traveling, and working restrictions), access to PPE, economic aids to businesses and individuals (income support), contact tracing, and emergency investment in healthcare (Hale et al., 2020); reflecting the uncertainty around the infectivity of the virus (SARS-CoV-2), the seriousness of the resulting disease (Covid-19), and leading to a spectrum of speculation.

---

[14] Oxford University developed a 'Coronavirus Government Response Tracker' that informs on live data and allows to contrast these responses. Available at: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker

What is tricky is how governments use data for decision-making, for example, *informing and measuring* COVID-19 cases: "in the UK many deaths that occur in care homes are not included in the regularly reported numbers. It is all a balancing act again – trying to understand what the fatality rate is" (Tatem, 2020: 236). This creates a lack of consistency and coherence in how data is handled[15]. How a government informs relevant figures, e.g. the amount of active cases, or how they justify their actions on incomplete data, can create a dangerous state of fragility among citizens, affecting their behaviour and agency, either for good (reassuring a communal approach of collaboration and mutual respect, there is a clear ethical choice) or bad (defiance, lack of encouragement and sympathy for citizens to abide protocols and measures). The legitimacy of any policy is at stake.

Mistrusting data is also an issue within governments. A common trade-off we see being discussed is that of the economy vs. public health. Despite genuine and sustained intentions to *flatten the curve* for the sake of health systems not collapsing, the dichotomy in this case is sacrificing the economy or increasing the risks of infection and death for the most vulnerable.

However, solutions to balance these factors are wrongly understood as categorical. This results in overlooking intermediate solutions. i.e. misclassifying trade-offs as necessary. Thus, data evidencing extremized vulnerability in certain sectors of the population becomes essential. Considering the costs, effects, and benefits of imposing a full, partial, or no lockdown, must include the information and analyses we have on these vulnerabilities. This must also go hand-in-hand with testing, tracing, and isolation (TTI) that allows considering any further relaxations.

If a measure is imposed affecting parts of the population at risk in a radically different way —either directly because of the virus, i.e. pre-existing health related risks, or indirectly because of social inequalities, i.e. orthogonal factors such as historical neglect of basic healthcare access and precarious living environments— *then* we are in front of an ethical imperative that requires us to guarantee the *active* incorporation of any relevant data (i.e. data on vulnerabilities) into the decision-making. These data can translate into complementary measures necessary for a *fair treatment*, informing additional measures that can be taken in present time (e.g. when there is a regional lockdown, mandatory face masks, increase TTI, or even all of the above), and future time (e.g. targeting the vulnerable population

---

[15] It is worth noticing how some governments implemented a successful approach to balance the uncertainty of data with informing the public in a consistent manner, such as New Zealand's Prime Minister Jacinta Ardern, ensuring "robust transparency and accountability" (Ladley, 2020).

to fight inequalities through, for example, public policies). Otherwise, effects of enforced poverty and isolation triggered by strict measures can accumulate over time, making them potentially more dangerous than the exposure to the virus if they are set to rise, notices (Koerner, 2020).

The public being unavoidably afflicted by the virus expect from those in charge the best decision *possible*, based on the availability of knowledge, resources, and data, whilst communicating clearly and transparently all the reasons (and reasoning) behind these decisions. Reinforcing trust and transparency, particularly regarding data gaps, in combination with consistency implementing these measures, is essential to provide a solution that meets the criteria for *a satisfactory ethical response.*

D. Digital epidemiology and ethics

In this final section we consider issues related to digital epidemiology[16] and health data, its limitations, managing expectations of data-driven solutions, and clarifying what an *ethical response* implies.

Data analysis is broadly acknowledged as a *promising* tool. Governments and researchers rush to unlock said potential, confronting limitations such as accuracy, granularity, quality of data sources (Gasser et al., 2020) and precision (even 'big data' has gaps or sparse information on relevant aspects we wish to know). These are known pitfalls, so what are the distinctive imperatives brought by the pandemic?

We face constant change and instability as the pandemic and our understanding of the virus evolve. Access to new and improved information requires *diligence*, contrasting needs, establishing protocols, and ensuring ethical practices. Uncertainty triggers our "human tendency to impose a single interpretation in ambiguous situations […] without entertaining alternatives" (Chater, 2020: 439), making room for our biases to take such interpretation for granted, i.e. confirmation bias. To make decisions and analyses as robust as possible, we need to recognize that we might not find these answers based only on current data or by pushing for promising new solutions. There is a responsibility for public health actors to "uphold rigorous and evidence-based practices" (Gasser et al., 2020: 3) to confront the 'dynamicism' of the pandemic, requiring higher vigilance and moderation (Shao, 2020).

---

[16] By digital epidemiology we follow Salathé's narrow definition: "digital epidemiology is epidemiology that uses digital data that was generated outside the public health system, i.e. […] data that was not generated with the primary purpose of doing epidemiology" (Salathé, 2018: 2).

Various technological solutions lack this solid foundation. Given that technology companies have "greater access to mobile device data, enormous financial resources, and larger teams of data scientists than academic researchers, their data products are being rolled out at a higher volume than high quality studies" (Balsari et al., 2020). This results in lack of depth and experience from expert epidemiologists, turning models into incomplete/insufficient and inconsistent proposals. What is more worrying is that these proposals are advising governments on how to fight the pandemic, by offering projection and prediction models (Balsari et al., 2020). But *prediction* relies on parameters that are unattainable in the midst of a pandemic; and even if improved estimates are obtained farther into the pandemic, the "butterfly-effect" that most parameters experience in the prediction models in which they are deployed still gives rise to huge uncertainties. Even after the fact, many model estimates remain poor and most models are merely conjecture, ensuring that overall totals are consistent with the available data (matching the marginals), but not necessarily guaranteeing any reliable meaning in specific parameter terms.

Fundamental facts are yet to be understood, and they will require further scrutiny since the measurement of these factors (e.g. totality of infected, immunity, or relapse rates) has been irregular across countries and stages of the pandemic. Most likely it remains truly variational due to local factors, resulting in a poorly estimated measurement that is also inconsistent across regions. And without virological testing data, we cannot fit models accurately for any predictive or decision-making purposes (Balsari et al., 2020).

As reviewed in the previous section, there is an ethical imperative for fairness and equality, which in this case relies on data-driven endeavours providing suitable proof to promote counteraction against entrenched health inequalities. In the case of digital epidemiology this implies *capturing* data beyond epidemiological factors, including factors such as SES, which are directly related to disparities in infection rates, thus contextualizing data to avoid further stigmatization (Gasser et al., 2020).

Contextualization is crucial to document and address racial disparities. Insufficient contextualization can end up perpetuating harmful myths of racial biology, undermining the background problem (Chowkwanyun & Reed, 2020). This translates into being aware of *manipulative approaches*, such as geographic disaggregation of data: "Granularity of data allows more fine-grained analyses, including multilevel spatial modelling [...] presented by themselves, such granular data can reinforce what the sociologist Loïc Wacquant has dubbed "territorial stigmatization" (Chowkwanyun & Reed, 2020: 202). This can quickly turn into severe stigmatization of neighbourhoods or areas that are known to have minorities, people of colour, or foreigners at its core.

Thus, issues with analysis and contextualization are imperative to achieve a *fair* process. There is a major cognitive dissonance between data and methods used to derive interpretations that are *meaningless*. Contextualizing gives an opportunity to analyse and *question* data considering possible confounders that could alter our interpretation of initial face values. A couple of months back, a paper based on the 'OpenSAFELY' initiative to hold secure analytics for health records in the NHS, presented factors associated with COVID-19 deaths. However when it was presented on preprint (Collaborative et al., 2020), it received a lot of criticism by epidemiologists, particularly those working with causal inference. The main issue they highlighted had to do with the "Table 2" fallacy (Westreich & Greenland, 2013), which focuses on *misinterpretations* of estimates. In the case of the 'OpenSAFELY' paper on factors associated with COVID-19 deaths, later published in *Nature* (Williamson et al., 2020), criticism pointing out the Table 2 fallacy, has been ignored. The critiques focused on how the study shows smoking having a 'protective' effect *if* this risk factor is misinterpreted as *causal*. In other words, the problem was that the study used the *same* set of adjustment factors for *all* their characteristics. This derives in authors reporting their statistical results with a lack of a concrete frame to understand them; if readers consider a causal interpretation, some factors could be seen as valid causal effects, which they are not.

This lack of attention to feedback not online makes open source initiatives pointless, it can create these *manipulative approaches* highlighted above. Contextualizing means having clarity and transparency not only in methodology, data sources, and ethical protocols; it also means framing the *question* and *develop methods to support that question* accordingly.

On a more positive note, if contextualized and analysed properly, new data sources collected during this pandemic could improve the misrepresentation of certain populations in epidemiological analysis, including those underrepresented by lack of access to healthcare (Mello & Wang, 2020). But these are future benefits.

Aside from this consideration of fairness, digital and traditional epidemiology, face a *demand for ethical and immediate answers*. In practice, acquiring data to answer questions is essential: what is the risk of death by exposing workers to the virus? the risk of morbidity which will lead to time off work? what is the mortality rate of the virus? what are is the symptomatology and incubation period? But expectations require that based on accessible data and/or previous information (retrospectively) other questions are answered too: *Should* we enforce the use of masks in the population? *Should* we stop isolation/distancing measures? Ethical questions that – in abstraction – could be framed under a classical ethical theory (i.e. utilitarianism or deontology) seem to drift away from the rigidness of such theories and seek refuge in the validity of data. If we have enough robustness in a model[17] showing us the risks and benefits of an outcome, or the data is compelling enough, actions at a population level could align consequentialism with deontology, showing the decisive role of data contrasting value judgements (Zohny, 2020).

But we are under uncertain circumstances where data is not always clear (note that this is not always the case even under normal scenarios) or sufficient to give us that reassurance; we must question the integrity of the data and its use. Provenance is vital for understanding *any* potential bias, selection is a powerful form of collider bias, which can severely distort findings (Cole et al., 2009). And this can be done even without realizing it if the provenance of the data is unknown.

Ethical imperatives and ethical reasoning maintain a normative role, but in this case, they focus on *contrasting* the evidence and *filling the blanks*. We take ethical insight to be a *mediator* and not an *enforcer*. Special circumstances like this pandemic put us under a collaborative scheme, where a *common morality* is seen as a priority. The value of ethical thinking is embedded by that common vulnerability (exposure to the pandemic) where an ethical response accounts for the handling, informing, contextualizing, and implementation of data-driven solutions.

CONCLUSIONS

The ethical imperatives presented here magnify pre-existing worries about data-driven ventures, including privacy, surveillance, transparency, accountability, and robustness of the analyses conducted. But they also show its connection with *fairness, equality, and power*. The pandemic has allowed us to contrast existing (and measurable) inequalities in societies that strive for technological progress where material conditions limit this development. Data ethics displays a normative role to question the *necessity* of data-driven solutions, their *appropriateness*, and their ethically aligned use and implementation. Much can be learned from the past and present struggles, to make data-driven technologies contributors in the fight against power asymmetries, discrimination, and disparity, working to achieve

---

[17] A recurrent issue is achieving that robustness, mostly because the same data can lead to a wide variety of interpretations, see (Silberzahn et al., 2018).

technological, pluralistic and inclusive societies. This crisis calls attention to the significant role of data ethics in matters of social justice.

REFERENCES

Adler, N. E., & Newman, K. (2002). Socioeconomic Disparities. In Health: Pathways And Policies. *Health Affairs*, *21*(2), 60–76. https://doi.org/10.1377/hlthaff.21.2.60

Aggarwal, N., Cowls, J., Floridi, L., Mokander, J., Morley, J., Taddeo, M., Tsamados, A., Wang, V., & Watson, D. (2020). *COVID-19 and the Digital Divides*. Oxford Internet Institute. Available at: https://www.oii.ox.ac.uk/blog/covid-19-and-the-digital-divides/

Alaa, A. M., Qian, Z., Rashbass, J., Gomes Pinto, K., Benger, J., & van der Schaar, M. (2020). SAGE (Scientific Advisory Group for Emergencies) meeting 29. *Ethnicity and outcomes of coronavirus (COVID-19) patients in England.* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/890235/s0248-ethnic-composition-outcomes-covid-19-patients-280420-sage29.pdf

Allmann, K. (2020, April 13). *Covid-19 is increasing digital inequality: We need human connectivity to close the digital divide*. Oxford Law Faculty. Available at: https://www.law.ox.ac.uk/research-and-subject-groups/oxfordshire-digital-inclusion-project/blog/2020/04/covid-19-increasing

Amnesty International. (2020, June 16). *Bahrain, Kuwait and Norway contact tracing apps a danger for privacy*. Available at: https://www.amnesty.org/en/latest/news/2020/06/bahrain-kuwait-norway-contact-tracing-apps-danger-for-privacy/

Balsari, S., Buckee, C., & Khanna, T. (2020, May 8). Which Covid-19 Data Can You Trust? *Harvard Business Review*. Available at: https://hbr.org/2020/05/which-covid-19-data-can-you-trust

Bauer, L. (2020, May 6). The COVID-19 crisis has already left too many children hungry in America. *Brookings*. Available at: https://www.brookings.edu/blog/up-front/2020/05/06/the-covid-19-crisis-has-already-left-too-many-children-hungry-in-america/

Beyond the exit strategy: Ethical uses of data-driven technology in the fight against COVID-19. (2020, April 17). [Webinar] Available at: https://www.youtube.com/watch?v=tWL34x1Dt7Y&feature=youtu.be

Broom, D. (2020, April 22). *Coronavirus has exposed the digital divide like never before*. World Economic Forum. Available at: https://www.weforum.org/agenda/2020/04/coronavirus-covid-19-pandemic-digital-divide-internet-data-broadband-mobbile/

Chan, J., Foster, D., Gollakota, S., Horvitz, E., Jaeger, J., Kakade, S., Kohno, T., Langford, J., Larson, J., Sharma, P., Singanamalla, S., Sunshine, J., & Tessaro, S. (2020). PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing. *ArXiv:2004.03544 [Cs]* [preprint] Available at: http://arxiv.org/abs/2004.03544

Chater, N. (2020). Facing up to the uncertainties of COVID-19. *Nature Human Behaviour*, *4*(5), 439–439. https://doi.org/10.1038/s41562-020-0865-2

Chowkwanyun, M., & Reed, A. L. (2020). Racial Health Disparities and Covid-19—Caution and Context. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMp2012910

Cole, S. R., Platt, R. W., Schisterman, E. F., Chu, H., Westreich, D., Richardson, D., & Poole, C. (2009). Illustrating bias due to conditioning on a collider. *International Journal of Epidemiology*, *39*(2), 417–420. https://doi.org/10.1093/ije/dyp334

Collaborative, T. O., Williamson, E., Walker, A. J., Bhaskaran, K. J., Bacon, S., Bates, C., Morton, C. E., Curtis, H. J., Mehrkar, A., Evans, D., Inglesby, P., Cockburn, J., Mcdonald, H. I., MacKenna, B., Tomlinson, L., Douglas, I. J., Rentsch, C. T., Mathur, R., Wong, A., … Goldacre, B. (2020). OpenSAFELY: Factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *MedRxiv*, [preprint] 2020.05.06.20092999. https://doi.org/10.1101/2020.05.06.20092999

Cullinane, C. (2020). COVID-19 and home-schooling: The crisis has exacerbated and highlighted existing educational inequalities. *British Politics and Policy at LSE*. Available at: https://blogs.lse.ac.uk/politicsandpolicy/home-schooling-covid-19/

Diaz, A. (2020). *Coronavirus, Location Tracking, and Civil Liberties*. Brennan Center for Justice. Available at: https://www.brennancenter.org/our-work/analysis-opinion/coronavirus-location-tracking-and-civil-liberties

Dubov, A., & Shoptaw, S. (2020). The Value and Ethics of Using Technology to Contain the COVID-19 Epidemic. *The American Journal of Bioethics*, *0*(0), 1–5. https://doi.org/10.1080/15265161.2020.1764136

Egede, L. E. (2006). Race, Ethnicity, Culture, and Disparities in Health care. *Journal of General Internal Medicine*, *21*(6), 667–669. https://doi.org/10.1111/j.1525-1497.2006.0512.x

Elwell-Sutton, T., Deeny, S., & Stafford, M. (2020). *Emerging findings on the impact of COVID-19 on black and minority ethnic people*. The Health Foundation. Available at: https://www.health.org.uk/news-and-comment/charts-and-infographics/emerging-findings-on-the-impact-of-covid-19-on-black-and-min

Fennell, M. L. (2005). Racial Disparities in Care: Looking Beyond the Clinical Encounter. *Health Services Research*, *40*(6 Pt 1), 1713–1721. https://doi.org/10.1111/j.1475-6773.2005.00489.x

Feys, M. (2020). Data Privacy vs Utility Fallacy: The COVID-19 Tracing App Problem. *PrivSec Report*. Available at: https://gdpr.report/news/2020/05/20/data-privacy-vs-utility-fallacy-the-covid-19-tracing-app-problem/

Fineman, M. (2016). The Vulnerable Subject: Anchoring Equality in the Human Condition. *Yale Journal of Law & Feminism*, *20*(1). Available at: https://digitalcommons.law.yale.edu/yjlf/vol20/iss1/2

Floridi, L. (2020). Mind the App—Considerations on the Ethical Risks of COVID-19 Apps. *Philosophy & Technology*, *33*(2), 167–172. https://doi.org/10.1007/s13347-020-00408-5

Gasser, U., Ienca, M., Scheibner, J., Sleigh, J., & Vayena, E. (2020). Digital tools against COVID-19: Taxonomy, ethical challenges, and navigation aid. *The Lancet Digital Health*, *2*(8), 425-434. https://doi.org/10.1016/S2589-7500(20)30137-0

Goodes, G. (2020). *The Proliferation of COVID-19 Contact Tracing Apps Exposes Significant Security Risks*. Report by Guardsquare (mobile application protection) Available at: https://www.guardsquare.com/en/blog/report-proliferation-covid-19-contact-tracing-apps-exposes-significant-security-risks

Hale, T., Angrist, N., Kira, B., Petherick, A., Phillips, T., & Webster, S. (2020). *Variation in government responses to COVID-19* (Version 6.0 Blavatnik School of Government Working Paper). University of Oxford. Available at: https://www.bsg.ox.ac.uk/research/publications/variation-government-responses-covid-19

Hutchins, S. S., Fiscella, K., Levine, R. S., Ompad, D. C., & McDonald, M. (2009). Protection of racial/ethnic minority populations during an influenza pandemic. *American Journal of Public Health*, *99 Suppl 2*, S261-270. https://doi.org/10.2105/AJPH.2009.161505

Koerner, R. (2020, March 24). *COVID-19 and the Trolley Problem: You're on the Tracks and the Government Is Controlling the Switch*. Foundation for Economic Education. Available at: https://fee.org/articles/covid-19-and-the-trolley-problem-you-re-on-the-tracks-and-the-government-is-controlling-the-switch/

Kondo, N. (2012). Socioeconomic Disparities and Health: Impacts and Pathways. *Journal of Epidemiology*, *22*(1), 2–6. https://doi.org/10.2188/jea.JE20110116

Ladley, A. (2020, April). *New Zealand and COVID-19: Parliamentary accountability in time of emergencies*. ConstitutionNet. Available at: http://constitutionnet.org/news/new-zealand-and-covid-19-parliamentary-accountability-time-emergencies

LaFave, S. (2020, May 4). *How school closures for COVID-19 amplify inequality*. The Hub. Available at: https://hub.jhu.edu/2020/05/04/school-closures-inequality/

Mackenzie, C., Rogers, W., & Dodds, S. (2013). What Is Vulnerability, and Why Does It Matter for Moral Theory? In *Vulnerability: New Essays in Ethics and Feminist Philosophy*. Oxford University Press. Available at: https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199316649.001.0001/acprof-9780199316649-chapter-1

Marmot, M., Allen, J., Boyce, T., Goldblatt, P., & Morrison, J. (2020). *Health Equity in England: The Marmot Review 10 Years On*. [Review]. Institute of Health Equity. Available at: http://www.instituteofhealthequity.org/resources-reports/marmot-review-10-years-on/the-marmot-review-10-years-on-full-report.pdf

Mello, M. M., & Wang, C. J. (2020). Ethics and governance for digital disease surveillance. *Science*, *368*(6494), 951–954. https://doi.org/10.1126/science.abb9045

Morley, J., Cowls, J., Taddeo, M., & Floridi, L. (2020). Ethical guidelines for COVID-19 tracing apps. *Nature*, *582*(7810), 29–31. https://doi.org/10.1038/d41586-020-01578-0

Mozur, P., Zhong, R., & Krolik, A. (2020). In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags. *The New York Times*. Available at: https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html

Myers, E. M. (2020). *Compounding Health Risks and Increased Vulnerability to SARS-CoV-2 for Racial and Ethnic Minorities and Low Socioeconomic Status Individuals in the United States*. *Preprints* 2020, 2020040234. https://doi.org/10.20944/preprints202004.0234.v1

National Security Commission on Artificial Intelligence (NSCAI). (2020). *Privacy and Ethics Recommendations for Computing Applications Developed to Mitigate COVID-19* (White Paper Series on Pandemic Response and Preparedness, No.1). Available at: https://drive.google.com/file/d/1m0AT21dS2XJ6JIGMgo7SuLSLveWIO8WK/view

NHSX. (2020). *NHS COVID-19 App*. Available at: https://www.nhsx.nhs.uk/covid-19-response/nhs-covid-19-app/

Niedzwiedz, C. L., O'Donnell, C. A., Jani, B. D., Demou, E., Ho, F. K., Celis-Morales, C., Nicholl, B. I., Mair, F. S., Welsh, P., Sattar, N., Pell, J. P., & Katikireddi, S. V. (2020). Ethnic and socioeconomic differences in SARS-CoV-2 infection: Prospective cohort study using UK Biobank. *BMC Medicine*, *18*(1), 160. https://doi.org/10.1186/s12916-020-01640-8

Oppel Jr., R. A., Gebeloff, R., Lai, K. K. R., Wright, W., & Smith, M. (2020). The Fullest Look Yet at the Racial Inequity of Coronavirus. *The New York Times*. https://www.nytimes.com/interactive/2020/07/05/us/coronavirus-latinos-african-americans-cdc-data.html

Pickett, K. E., & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, *128*, 316–326. https://doi.org/10.1016/j.socscimed.2014.12.031

Ramsetty, A., & Adams, C. (2020). Impact of the digital divide in the age of COVID-19. *Journal of the American Medical Informatics Association*. https://doi.org/10.1093/jamia/ocaa078

Reyes, C. Q. (2020). Pandemia Covid-19 e Inequidad Territorial: El agravamiento de las desigualdades educativas en Chile. *Revista Internacional de Educación para la Justicia Social*, *9*(3). https://revistas.uam.es/riejs/article/view/12143

Riley, W. J. (2012). Health Disparities: Gaps in Access, Quality and Affordability of Medical Care. *Transactions of the American Clinical and Climatological Association*, *123*, 167–174.

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, *10*(1), 3069. https://doi.org/10.1038/s41467-019-10933-3

Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D. T., Greaves, F., Harper, L., Hawe, P., Moore, L., Petticrew, M., Rehfuess, E., Shiell, A., Thomas, J., & White, M. (2017). The need for a complex systems model of evidence for public health. *The Lancet*, *390*(10112), 2602–2604. https://doi.org/10.1016/S0140-6736(17)31267-9

Salathé, M. (2018). Digital epidemiology: What is it, and where is it going? *Life Sciences, Society and Policy*, *14*. https://doi.org/10.1186/s40504-017-0065-7

Seng Ah Lee, M. (2020). *Data Ethics in combating COVID-19 after lockdown.* Cambridge University, Trust & Technology Initiative. Research perspectives. Available at: https://www.trusttech.cam.ac.uk/research-agendas-and-perspectives/covid-19-trust-technology-society-research-perspectives/data

Shao, C. (2020). The COVID trolley dilemma. *The American Journal of Surgery*. [Online ahead of print]. https://doi.org/10.1016/j.amjsurg.2020.05.012

Shildrick, M. (2002). *Theory, Culture & Society: Embodying the monster: Encounters with the vulnerable self.* SAGE Publications Ltd. https://doi.org/10.4135/9781446220573

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., … Nosek, B. A. (2018). Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Silver, L. (2019). Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. *Pew Research Center's Global Attitudes Project.* Available at: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/

Sloane, M. (2019). On the Need for Mapping Design Inequalities. *Design Issues*, *35*(4), 3–11. https://doi.org/10.1162/desi_a_00559

Stahl, B. C., & Wright, D. (2018). Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. *IEEE Security Privacy*, *16*(3), 26–33. https://doi.org/10.1109/MSP.2018.2701164

Stanley, J., & Granick, J. (2020). *ACLU White Paper: The Limits of Location Tracking in an Epidemic.* Available at: https://www.aclu.org/report/aclu-white-paper-limits-location-tracking-epidemic

Subramanian, S. V., & Kawachi, I. (2004). Income inequality and health: What have we learned so far? *Epidemiologic Reviews*, *26*, 78–91. https://doi.org/10.1093/epirev/mxh003

Sweeney, Y. (2020). Tracking the debate on COVID-19 surveillance tools. *Nature Machine Intelligence*, *2*(6), 301–304. https://doi.org/10.1038/s42256-020-0194-1

Taddeo, M. (2015). The Struggle Between Liberties and Authorities in the Information Age. *Science and Engineering Ethics*, *21*(5), 1125–1138. https://doi.org/10.1007/s11948-014-9586-0

Taddeo, M. (2020). The Ethical Governance of the Digital During and After the COVID-19 Pandemic. *Minds and Machines*, *30*(2), 171–176. https://doi.org/10.1007/s11023-020-09528-5

Tatem, A. (2020). COVID-19: Data Uncertainty and Effectiveness of Interventions. *HealthManagement*, *20*(3). Available at: https://healthmanagement.org/uploads/article_attachment/hm3-v20-5may-data-uncertainty.pdf

Turner, B. S. (2006). *Vulnerability and Human Rights*. The Pennsylvania State University Press. Available at: https://books.google.com/books/about/Vulnerability_and_Human_Rights.html?id=NESaCgAAQBAJ

Veale, M. (2020). *Analysis of the NHSX Contact Tracing App 'Isle of Wight' Data Protection Impact Assessment* [Preprint]. LawArXiv. https://doi.org/10.31228/osf.io/6fvgh

Westreich, D., & Greenland, S. (2013). The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. *American Journal of Epidemiology*, *177*(4). https://doi.org/10.1093/aje/kws412

Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C. E., Curtis, H. J., Mehrkar, A., Evans, D., Inglesby, P., Cockburn, J., McDonald, H. I., MacKenna, B., Tomlinson, L., Douglas, I. J., Rentsch, C. T., Mathur, R., Wong, A. Y. S., Grieve, R., … Goldacre, B. (2020). OpenSAFELY: Factors associated with COVID-19 death in 17 million patients. *Nature*, 1–11. https://doi.org/10.1038/s41586-020-2521-4

Zhang, B., Kreps, S. E., & McMurry, N. (2020). *Americans' perceptions of privacy and surveillance in the COVID-19 Pandemic* [Preprint]. Open Science Framework. https://doi.org/10.31219/osf.io/9wz3y

Zohny, H. (2020). Pandemic Ethics: Covid-19 Shows Just How Much of Ethics Depends on (Good) Data. *Practical Ethics,* University of Oxford. Available at: http://blog.practicalethics.ox.ac.uk/2020/03/pandemic-ethics-covid-19-shows-just-how-much-of-ethics-depends-on-good-data/