I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13110

Articial Intelligence against COVID-19: An Early Review

Wim Naudé

APRIL 2020

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13110

Articial Intelligence against COVID-19: An Early Review

Wim Naudé
Maastricht School of Management, RWTH Aachen University and IZA

APRIL 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

Articial Intelligence against COVID-19: An Early Review

Artificial Intelligence (AI) is a potentially powerful tool in the fight against the COVID- 19 pandemic. Since the outbreak of the pandemic, there has been a scramble to use AI. This article provides an early, and necessarily selective review, discussing the contribution of AI to the fight against COVID-19, as well as the current constraints on these contributions. Six areas where AI can contribute to the fight against COVID-19 are discussed, namely i) early warnings and alerts, ii) tracking and prediction, iii) data dashboards, iv) diagnosis and prognosis, v) treatments and cures, and vi) social control. It is concluded that AI has not yet been impactful against COVID-19. Its use is hampered by a lack of data, and by too much data. Overcoming these constraints will require a careful balance between data privacy and public health, and rigorous human-AI interaction. It is unlikely that these will be addressed in time to be of much help during the present pandemic. In the meantime, extensive gathering of diagnostic data on who is infectious will be essential to save lives, train AI, and limit economic damages.

JEL Classification: O32, O39, I19, O20

Keywords: articial intelligence, COVID-19, Coronavirus, health, data science, development, technology, innovation

Corresponding author:
Wim Naudé
Technology and Innovation Management (TIM)
RWTH Aachen University
Kackertstraße 7
52072 Aachen
Germany

E-mail: naude@time.rwth-aachen.de

1 Introduction

COVID-19 disease,[1] caused by the SARS-CoV-2 virus, was identified in December 2019 in China and declared a global pandemic by the WHO on 11 March 2020. Artificial Intelligence (AI) is a potentially powerful tool in the fight against the COVID-19 pandemic (Bullock et al., 2020; Petropoulos, 2020). AI can, for present purposes, be defined as Machine Learning (ML)[2], Natural Language Processing (NLP), and Computer Vision applications to teach computers to use big data-based models for pattern recognition, explanation, and prediction. These functions can be useful to recognize (diagnose), predict, and explain (treat) COVID-19 infections, and help manage socio-economic impacts. Since the outbreak of the pandemic, there has been a scramble to use AI, and other data analytic tools, for these purposes, see e.g. Broad (2020); Hollister (2020) and Taulli (2020).

In this paper, I provide an early, rapid review of this AI scramble, discussing the actual and potential contribution of AI to the fight against COVID-19, as well as the current constraints on these contributions. The paper aims to draw quick take-aways from a fast expanding discussion and growing body of work in order to serve as an input for rapid responses in research, policy and medical analysis. The cost of the pandemic in terms of lives and economic damage will be terrible; at the time of writing, great uncertainty surrounded estimates of just how terrible, and of how successful both non-pharmaceutical and pharmaceutical responses can be. Improving AI, one of the most promising data analytic tools to have been developed over the past decade or so, so as to help reduce these uncertainties, is a worthwhile pursuit. Encouragingly, data scientists have taken up the challenge[3].

The key take-aways are as follows. I conclude that AI has not yet been impactful against COVID-19. Its use is hampered by a lack of data, and by too much noisy and outlier data. Overcoming these constraints will require a careful balance between data privacy and public health concerns, and rigorous human-AI interaction. It is unlikely that these will be addressed in time to be of much help during the present pandemic. Instead, AI may "*help with the next pandemic*" (Heaven, 2020). In the meantime, gathering extensive diagnostic data on who is infectious will be essential to save lives and limit economic damages (Baldwin, 2020; Bloom et al., 2020; Dewatripont et al., 2020).

The paper is structured as follows. In section 2 the actual and potential contributions of AI against COVID-19 is discussed. In section 3 the constraints are examined. Section 4 concludes.

2 Actual and Potential Contributions of AI against COVID-19

There are six areas where AI can contribute to the fight against COVID-19: i) early warnings and alerts, ii) tracking and prediction, iii) data dashboards, iv) diagnosis and prognosis, v) treatments, and cures, and vi) social control.

[1] For a "user guide to COVID-19" see Galeotti and Surico (2020).
[2] Within ML an important class of techniques that is frequently used in the fight against COVID-19 is known as Deep Learning, see LeCun et al. (2015).
[3] Which implies that the shelf-life of this paper is likely to be brief.

2.1 Early warnings and Alerts

The case of the Canadian-based AI model, *BlueDot*[4], has already become legendary. It illustrates that a relatively low-cost AI tool (BlueDot was funded by a startup investment of around US$ 9 million) can out-predict humans in spotting infectious disease outbreaks. According to accounts, *BlueDot* predicted the outbreak of the infection at the end of 2019, issuing a warning to its clients on 31st of December 2019, before the World Health Organization did so on 9th of January 2020 (Kreuzhuber, 2020). Bogoch et al. (2020), a group of researchers working with *BlueDot,* listed the top 20 destination cities where passengers from Wuhan would arrive in the wake of the outbreak. They warned that these cities could be at the forefront of the global spread of the disease.

While *BlueDot* is undoubtedly a powerful tool, much of the publicity it has received contain some exaggeration and some undervaluation of the role of human scientists. First, while *BlueDot* sounded an alarm on 31st December 2019, another AI-based model, *HealthMap*[5], at Boston Children's Hospital (USA), sounded an alarm even earlier, on 30 December 2019. Moreover, According to Associated Press[6], only 30 minutes after this, a scientist at the Program for Monitoring Emerging Diseases (PMED) issued an alert. While the AI-based model was faster by only 30 minutes, it, however, attached a very low level of significance to the outbreak. In essence, it required human interpretation and providing context to recognize the threat. In fact, even in the case of *BlueDot*, humans remain central in evaluating and interpreting its output, as Kamran Khan, Founder of *BlueDot*, explained in a podcast[7]. It is therefore correct to stress that human input, across disciplines, is needed for the optimal application of AI (see e.g. Hollister (2020)).

2.2 Tracking and Prediction

AI can be used to track (including nowcasting) and to predict how the COVID-19 disease will spread over time and over space. For instance, following a previous pandemic, that of the 2015 Zika-virus, Akhtar et al. (2019) developed a dynamic neural network to predict its spread. Models such as these will, however, need to be re-trained using data from the COVID-19 pandemic. This seems to be happening now. At Carnegie Mellon University, algorithms trained to predict the seasonal flu, are now being re-trained (Hao, 2020). And as I discuss below, various initiatives are under way to collect training data from the current pandemic.

Various problems bedevil the accurate forecasting of the pandemic, see e.g. Hao (2020); Rowan (2020) and Lazer et al. (2014). These include a lack of historical and unbiased data on which to train the AI; panic behavior which leads to "noise" on social media; and the fact that the characteristics of COVID-19 infections differ from those of previous pandemics. It is not only the lack of historical data but also the problems with using "big data", e.g., harvested from social media, that have shown to be problematic. Here, the pitfalls of big data and AI in the context of infectious diseases, as was illustrated in the infamous failure of *Google Flu Trends*,[8] remain valid. Lazer et al. (2014) referred to these as "big data hubris and algorithm dynamics". For instance, as the infection continues to spread and the social media traffic around it accumulates,

[4] See https://bluedot.global
[5] see http://www.diseasedaily.org/about
[6] Read: http://dailym.ai/3avyCTK
[7] Watch https://www.youtube.com/watch?v=V6BpKSGquRw&feature=youtu.be
[8] See https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/

so the amount of noise accumulates, which has to be filtered through before meaningful trends can be discerned. Generally, and this is also bad news for AI forecasting models in other fields, including economics and finance, since for any prediction algorithm that rely on past behaviour, a global outlier event with its mass of new and unprecedented data, such as COVID-19, can be described as by Rowan (2020) does as "*the kryptonite of modern Artificial intelligence*". As a result he concludes that over the near future "*many industries are going to be pulling the humans back into the forecasting chair that had been taken from them by the models*".

One way to deal with big data hubris and algorithm dynamics is through content moderation on social media. The major social media platforms such as *Google (YouTube)* and *Facebook* have started to use AI more intensively to do content moderation, including checking for fake news (Ortutay and Klepper, 2020), due to their being affected by a reduction in human staff resulting from lockdown measures (Heilweil, 2020). Relying more on AI for content moderation has laid bare the fact that AI is still doing a poor job of it. *YouTube* is reported to have admitted that using AI more extensively in content moderation is "error-prone" (Newton, 2020). This again illustrates the importance of human input to, and direction of, AI.

As a result of a lack of data, noisy social media, big data hubris, and algorithmic dynamics, AI forecasts of the spread of COVID-19 are not yet very accurate or reliable. Hence, so far, most models used for tracking and forecasting do not use AI methods. Instead, most forecasters prefer established epidemiological models, so-called SIR models, the abbreviation standing for the population of an area that is *S*usceptible, *I*nfected, and *R*emoved. For example, the Institute for the Future of Humanity at Oxford University provides forecasts of the spread of the virus based on the GLEAMviz epidemiological model[9]. *Metabiota*[10], a San Francisco-based company, offers an *Epidemic Tracker*[11] and a near-term forecasting model of disease spread. Crawford, an Oxford University mathematician, provides a short and concise explanation SIR-models in an YouTube video[12].

The Robert Koch Institute in Berlin uses an epidemiological SIR model that takes into account containment measures by governments, such as lockdowns, quarantines, and social distancing prescriptions[13]. A similarly extended SIR model, taking into account public health measures against the pandemic and using data from China, has recently been pre-published by Song et al. (2020) and made available in *R* format. The Robert Kock Institute's model has been used earlier in the case of China to illustrate that containment can be successful in reducing the spread to slower than exponential rates - see Maier and Brockmann (2020).

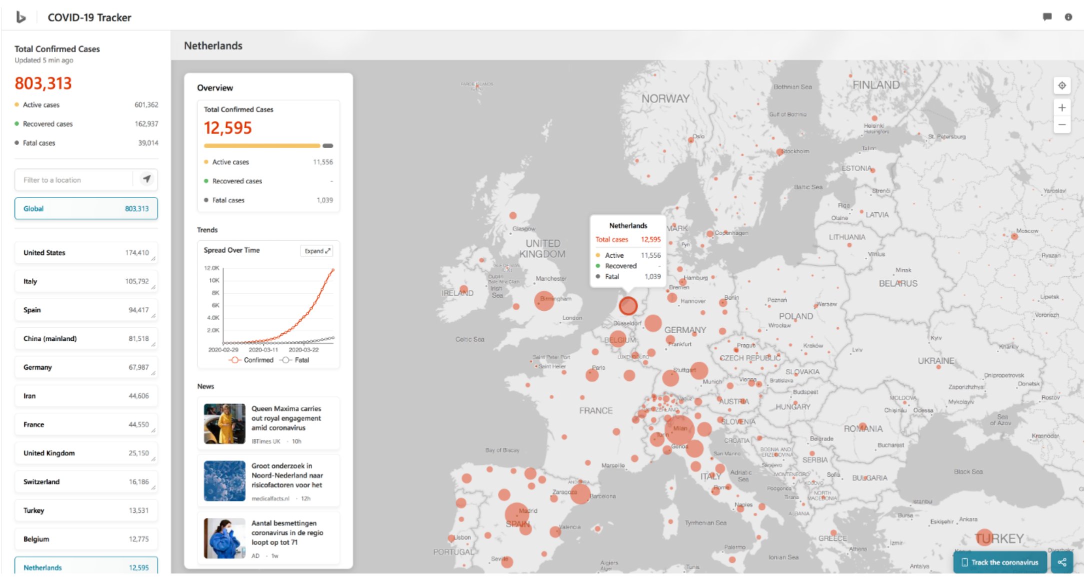
Tracking and predicting the spread of COVID-19 are valuable data inputs for public health authorities to plan, prepare, and manage the pandemic. And to evaluate where they are on the epidemiological curve and whether they succeed in flattening it. It can also provide rough reflections on the possible success of measures taken to reduce or slow down the spread. For example, the Robert Koch Institute made a forecast that the number of infections in the Netherlands will reach 10,922 by 28 March 2020. At this date, according to John's Hopkins University's CSSE, the total number of infected patients in the Netherlands was lower than predicted, at 8,647. This may strengthen arguments that the government's approach is helping to reduce the growth in infections.

[9] See http://www.gleamviz.org
[10] See https://www.metabiota.com
[11] See https://www.epidemictracker.com
[12] Watch it at https://youtu.be/NKMHhm2Zbkw
[13] Their model is explained at http://rocs.hu-berlin.de/corona/docs/forecast/model/

2.3 Data Dashboards

The tracking and forecasting of COVID-19 have caused an industry of data dashboard creation for visualization of the actual and expected spread. *MIT Technology Review*[14] has produced a ranking of these tracking and forecasting dashboards. They rank the top dashboards to be those of *UpCode*, *NextStrain*, the Johns Hopkins' *CSSE*, *Thebaselab*, the *BBC*, the *New York Times*, and *HealthMap*. Other notable dashboards include Microsoft Bing's *COVID-19 Tracker* - See *Figure 1*.

Figure 1: Microsoft Bing's COVID-19 Tracker



Note(s): Screenshot of Bing's COVID-19 Tracker, 31 March 2020.

While these dashboards give a global overview, an increasing number of countries, including emerging economies, already have their own dashboards in place; for instance, South Africa established the *COVID 19 ZA South Africa Dashboard*[15] which is maintained by the Data Science for Social Impact Research Group at the University of Pretoria.

To facilitate the production of data visualizations and dashboards of the pandemic, *Tableau* has created a *COVID-19 Data Hub* with a COVID-19 Starter Workbook[16]. Sarkar (2020) provides a Python script to illustrate how one could extract data from the New York Times's COVID-19 dataset and create data visualizations of the progression of the infection. Makulec (2020) calls for responsible visualization of COVID-19 data, listing *Ten Considerations when Visualizing COVID-19 Data*.

2.4 Diagnosis and Prognosis

Fast and accurate diagnosis of COVID-19 can save lives, limit the spread of the disease, and generate data on which to train AI models. AI may provide useful input in this regard, in particular in making diagnoses based on chest radiography images. According to a recent review of AI applications against COVID-19 by Bullock et al. (2020), studies have shown that

[14] See https://www.technologyreview.com/s/615330/best-worst-coronavirus-dashboards/
[15] See https://tinyurl.com/wsw5c89
[16] See https://www.tableau.com/covid-19-coronavirus-data-resources

AI can be as accurate as humans, can save radiologists' time, and perform a diagnosis faster and cheaper than with standard tests for COVID-19. Both X-rays and Computed Tomography (CT) scans can be used. Rosebrock (2020) offers a tutorial on how to use Deep Learning to diagnose COVID-19 using X-ray images. He makes the point that COVID-19 tests are "*in short supply and expensive*," but that "*all hospitals have X-ray machines.*" Maghdid et al. (2020) has proposed a technique using mobile phones to scan CT images.

Several initiatives are underway in this regard. Wang and Wong (2020) developed *COVID-Net*, which is a deep convolutional neural network (see e.g. Rawat and Wang (2017)), which can diagnose COVID-19 from chest radiography images. It has been trained on open repository data from around 13,000 patients with various lung conditions, including COVID-19. However, as the authors indicate, it is "by no means a production-ready solution", and they call on the scientific community to develop it further, in particular to "improve sensitivity" (Ibid, p.6) . Chen et al. (2020b) published an Deep Learning model (not yet peer-reviewed, however) to diagnose COVID-19 from CT scans, concluding that "*The deep learning model showed comparable performance with an expert radiologist, and greatly improve the efficiency of radiologists in clinical practice. It holds great potential to relieve the pressure off frontline radiologists, improve early diagnosis, isolation, and treatment, and thus contribute to the control of the epidemic*". (Ibid, p.1).

Other initiatives include that of researchers at the Dutch University of Delft who released an AI model for diagnosing COVID-19 from X-rays at the end of March 2020. This model, labeled CAD4COVID, is described on their website[17] as "an artificial intelligence software that triages COVID-19 suspects on chest X-rays images". It relies on previous AI models developed by the university of diagnosis of tuberculosis.

The potential is not yet carried over into practice, although it has been reported that a number of Chinese hospitals have deployed "AI-assisted" radiology technologies[18]. Radiologists elsewhere have expressed their concern that there is not enough data available to train AI models, that most of the available COVID-19 images come from Chinese hospitals and may suffer from selection bias, and that using CT-scans and X-rays may contaminate equipment and spread the disease further. Indeed, the use of CT scans in European hospitals has dropped after the pandemic broke, perhaps reflecting this concern (Ross and Robbins, 2020).

Finally, once the disease is diagnosed in a person, the question is whether and how intensively that person will be affected. Not all people diagnosed with COVID-19 will need intensive care. Being able to forecast who will be affected more severely can help in targeting assistance and planning medical resource allocation and utilization. Yan et al. (2020) used Machine Learning to develop a *prognostic prediction algorithm* to predict the mortality risk of a person that has been infected, using data from (only) 29 patients at Tongji Hospital in Wuhan, China. And Jiang et al. (2020) presents an AI that can predict with 80 percent accuracy which person affected with COVID-19 may go on to develop acute respiratory distress syndrome (ARDS). The sample that they used to train their AI system is, however, small (only 53 patients) and restricted to two Chinese hospitals.

In conclusion, the application of AI to diagnose COVID-19, and to make a prognosis of how patients may progress, has spurred much research effort but is not yet widely operational. It is probably correct as Coldeway (2020) concludes, "*No one this spring is going to be given a coronavirus diagnosis by an AI doctor*". It also seems that comparatively less effort is on using

[17] Available at `https://www.delft.care/cad4covid/`

[18] See for instance this report on Imaging Technology News :`https://tinyurl.com/qtclguo`.

AI for very early diagnostic purposes, for instance, in identifying whether someone is infected before it shows up in X-rays or CT scans, or on finding data-driven diagnostics that have less contamination risk.

2.5 Treatments and Cures

Even long before the COVID-19 outbreak, AI was lauded for its potential to contribute to new drug discovery, see e.g. Coldeway (2019); Fleming (2018); Segler et al. (2018) and Smith (2018). In the case of COVID-19, a number of research labs and data centers have already indicated that they are recruiting AI to search for treatments for and a vaccine against COVID-19. The hope is that AI can accelerate both the processes of discovering new drugs as well as for repurposing existing drugs.

For example, Google's *DeepMind*, a firm famous for is *AlphaGo* game-playing algorithm,[19] has used AI to predicted the structure of the proteins of the virus information that could be useful in developing new drugs. However, as *DeepMind* makes clear on its website[20], "*we emphasize that these structure predictions have not been experimentally verified...we can't be certain of the accuracy of the structures we are providing*".

Beck et al. (2020) report results from using Machine Learning to identify that an existing drug, *atazanavir*, could potentially be repurposed to treat COVID-19. And Stebbing et al. (2020), working with Benevolent AI, a UK AI startup, identified *Baricitinib,* used to treat rheumatoid arthritis and myelofibrosis, as a potential treatment for COVID-19.

It is not very likely that these treatments (in particular a vaccine) will be available in the near future, at least to be of much use during the current pandemic. The reason is that the medical and scientific checks, trails, and controls that need to be performed before these drugs will be approved, once they have been identified and screened, will take time - according to estimates up to 18 months for a vaccine (Regalado, 2020). See also Vanderslott et al. (2020) for an explanation of the process that a potential anti-Covid-19 drug will have to go through.

2.6 Social Control

AI has been, and can further be used, to manage the pandemic by using thermal imaging to scan public spaces for people potentially infected, and by enforcing social distancing and lockdown measures (Rivas, 2020). For example, as described by Chun (2020) in the *South China Morning Post*, "*At airports and train stations across China, infrared cameras are used to scan crowds for high temperatures. They are sometimes used with a facial recognition system, which can pinpoint the individual with a high temperature and whether he or she is wearing a surgical mask.*"

Chinese firm *Baidu* is one of the producers of such infrared cameras that uses computer vision to scan crowds. It is reported that these cameras can scan 200 persons per minute and will recognize those whose body temperature exceeds 37,3 degrees (Dickson, 2020). Thermal imaging has however been criticized as being inadequate to to identify from a distance a fever in people who are wearing glasses (because scanning the inner tear duct gives the most reliable indication)

[19] For a description of AlphaGo's general reinforcement learning algorithm, see Silver et al. (2018).
[20] See `https://tinyurl.com/wtsdagu`

and because it cannot identify whether a person's temperature is raided because of COVID-19 or some other reason (Carroll, 2020).

Moreover, as Chun (2020) further reports, "*This system is also being used to ensure citizens obey self-quarantine orders. According to reports, individuals who flouted the order and left home would get a call from the authorities, presumably after being tracked by the facial recognition system*". This type usage is not limited to China. An AI-based computer vision camera system scanning public areas has been used to monitor whether people in the UK city of Oxford keep to the social distancing measures of the government[21]. A USA computer vision-based startup is already offering "social distancing detection" software, which uses camera images to detect when social distancing norms are breached, after which it will send out a warning (Maslan, 2020). In an extreme case, the Israeli government has approved cyber-monitoring by its security services to identify and quarantine people that may be infected[22].

Whereas using AI to predict and diagnose COVID-19 is hampered due to lack of historical training data, AI tools such as computer vision and robots are not. Therefore, we are more likely over the short term to see this type of AI being used and used moreover for social control. Related technologies, such as mobile phones with AI-powered apps or wearables that harvest location, usage, and health data of their owners, are also more likely to be employed. According to Petropoulos (2020) such apps can "*enable patients to receive real-time waiting-time information from their medical providers, to provide people with advice and updates about their medical condition without them having to visit a hospital in person, and to notify individuals of potential infection hotspots in real-time so those areas can be avoided*".

The fear is that once the outbreak is over, that erosion of data privacy would not be rolled back and that governments would continue to use their improved ability to survey their populations- and use the data obtained in the fight against COVID-19 for other purposes. As Harari (2020) warns "*Even when infections from coronavirus are down to zero, some data-hungry governments could argue they needed to keep the biometric surveillance systems in place because they fear a second wave of coronavirus, or because there is a new Ebola strain evolving in central Africa, or because...you get the idea*".

In section 4 I will return to these concerns.

3 Constraints: Too Much, and Too Little, Data

AI has the potential to be a tool in the fight against COVID-19 and similar pandemics. However, as Petropoulos (2020) concludes, "*AI systems are still at a preliminary stage, and it will take time before the results of such AI measures are visible*". And Bullock et al. (2020) in their review of the use of AI against COVID-19 conclude that "*very few of the reviewed [AI] systems have operational maturity at this stage.*"

It has been shown in this paper that the current use of AI is actually constrained by, on the one hand, by a lack of data, and on the other hand, by too much data. There is not sufficient historical data (yet) on which to train AI models, not enough open datasets and models to work on, but also potential problems of big data hubris, non-adjustment of algorithms, and a outlier data and a deluge of scientific findings, which all need to be shifted and evaluated before

[21] As reported in `https://tinyurl.com/us6fhka`.
[22] As reported by the BBC at `https://www.bbc.com/news/technology-51930681`.

offering concrete diagnostic and treatment options.

In contrast, where AI is easier to use, such as in surveillance, we are likely to see more effort but with potential adverse longer-term consequences for privacy and related human rights concerns (Ienca and Vayena, 2020). In what follows, I will deal in more detail with these matters.

First, as far as the need for more data is concerned, more new training data is clearly needed on COVID-19; more openness and sharing of information is required, and more collaborative and multidisciplinary research is necessary to improve the ability of AI. Most of the publications reporting on diagnostic tools or treatments through AI tend to use small, possibly biased, and Chinese based samples. More diagnostic testing needs to be done if the tracking and forecasting of the pandemic is to improve, and as will be argued below, the world economy is to be re-started. In all of these furthermore, the role of humans in interacting with and steering AI is necessary and perhaps even more important than ever.

So far, there has been promising progress with a number of notable activities recognizing the importance of building and sharing existing datasets and information about the epidemic. One of the first has been the *World Health Organization*'s (WHO) Global Research on Coronavirus disease database[23], which also provides links to other similar initiatives. One of these is the open access data of the *GISAID Initiative* (formerly the Global Initiative on Sharing All Influenza Data).

One of the most ambitious of these focusing on AI, is perhaps the joint initiative between Semantic Scholar, the Allen Institute for Artificial Intelligence, Microsoft, Facebook, and others, to make openly available the *COVID-19 Open Research Dataset (CORD-19)* which contains around 44,000 scholarly articles which are now available for data mining.[24] *Kaggle*, a data science competition platform, has issued a data competition based on this data, a *COVID-19 Open Research Dataset Challenge.*

Other similar initiatives includes that of Elsevier that has made publicly available in its *Novel Coronavirus Information Center* early-stage and peer-reviewed research on COVID-19 and to around 20,000 related articles on ScienceDirect, as well as the full texts for data mining [25], as well as of The Lens, that has made available all its data on patents in what it calls the *Human Coronavirus Innovation Landscape Patent and Research Works Open Datasets* to support the search for new and repurposed drugs[26].

Other notable new data-gathering and open innovation initiatives include that of The University of California, Berkeley, the University of Illinois at Urbana-Champaign, and C3.ai who established the C3.ai Digital Transformation Institute[27]. This Institute has launched a Call for Proposals for *AI Techniques to Mitigate Pandemic.* These should deal amongst others with "Applying machine learning and other AI methods to mitigate the spread of the COVID-19 pandemic" and "Data analytics for COVID-19 research harnessing private and sensitive data".

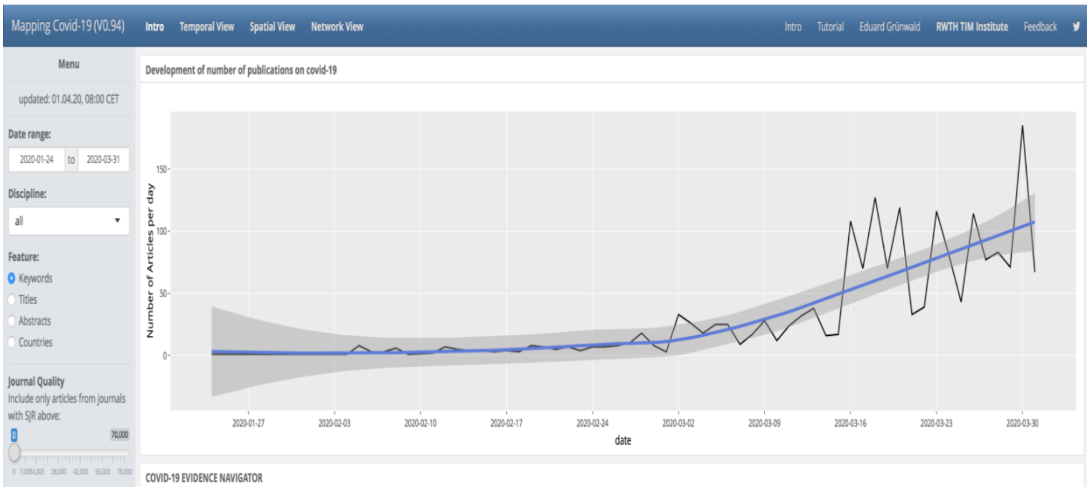
It is not only the large tech companies, publishers, and universities that are promoting open access to data and scientific literature on COVID-19, but also smaller startups and NGOs. For example, Newspeak House - a UK based independent residential college - has started writing

[23] Accessible at `https://tinyurl.com/rdkr4c7`.
[24] Available at `https://pages.semanticscholar.org/coronavirus-research`.
[25] Available at `https://www.elsevier.com/connect/coronavirus-information-center`.
[26] Available at `https://about.lens.org/covid-19/`.
[27] See `https://tinyurl.com/vw76xjv`.

a crowdsourced "Coronavirus Tech Handbook"[28]. And Chen et al. (2020a) published the first public COVID-19 Twitter dataset.

It is not only a lack of data that constrains AI applications, but also, perhaps paradoxically, too much data. As was noted, as the pandemic progresses and the issue dominates the news and social media, too much big data noise and outlier observations are created, and algorithms will be overwhelmed this was the lesson from the Google Flu Trends' failed initiative. Content curation and algorithmic adjustment, both involving human common sense, become especially valuable in such a context. Furthermore, scientists will need to deal with the deluge of scientific papers and new data being generated and shift through these. More than 100 scientific articles on the pandemic now appear daily (185 on 30 March 2020). This potential information overload is, however, where data analytic tools can play an important role. An example of an initiative in this regard is the *COVID-19 Evidence Navigator* by Gruenwald et al. (2020) which provides computer-generated evidence maps of scientific publications on the pandemic, daily updated from PubMed - see Figure 2.

Figure 2: The COVID-19 Evidence Navigator



Note(s): Screenshot of Gruenwald et al.'s COVID-19 Evidence Navigator, 1 April 2020.

4 Concluding Remarks

AI is not yet playing a significant role in the fight against COVID-19, at least from the epidemiological, diagnostic and pharmaceutical points of view. Its use is constrained by a lack of data and by too much noisy and outlier data. The creation of unbiased time series data for AI training is necessary. A growing number of international initiatives in this regard is encouraging; however, there is an imperative for more diagnostic testing. Not only for providing training data to get AI models operational but moreover for more effectively managing the pandemic and reducing its cost in terms of human lives and economic damage.

At the time of writing, the significant efforts of all affected countries have non-pharmaceutical: to shut down their economies through lockdowns, enforcing social distancing, and canceling events. These measures seem, for now, to have succeeded in slowing down the spread and saving lives (McNeil, 2020; Flaxman et al., 2020). However, whether these measures are sustainable for more than a couple of weeks is doubtful. According to Ferguson et al. (2020) from the Imperial

[28] See https://coronavirustechhandbook.com/home.

College COVID-19 Response Team, "*The major challenge of suppression is that this type of intensive intervention will need to be maintained until a vaccine becomes available, given that we predict that transmission will quickly rebound if interventions are relaxed*".

More diagnostic testing will be helpful to eventually halt the pandemic, limit the economic damage from lockdowns, and avoid a rebound once restrictions are relaxed. Dewatripont et al. (2020) make a case for extensive diagnostic testing[29] of the population to allow people to return to work only if they are not infectious, to place in quarantine those who are. They also call for more randomly sampled tests in order to improve our estimates of the proportion of the population with the virus that remain asymptomatic. At present, we just do not know how many people are infected (Britt, 2020). In essence, it may be, as Li et al. (2020) suggests, that 86 percent of all infections are undocumented. If this is the case, then the danger of a rebound of the pandemic is highly likely - and economic recovery even further delayed. Thus, overcoming limited data in terms of who is infectious is critical.

Clearly, data is central to whether AI will be an effective tool against future epidemics and pandemics. The fear is, as I already mentioned, that public health concerns would trump data privacy concerns. Governments may want to continue the extraordinary surveillance of their citizens long after the pandemic is over. Thus, concerns about the erosion of data privacy are justified.

A full discussion of the legal and ethical dimensions of data management falls outside the scope of this article. Two excellent recent commentaries[30] are, however, those of Ienca and Vayena (2020) and Marcus (2020). In short, given the public health threat posed by the pandemic, the European GDPR (Article 9) allows personal data collection and analysis, as long as it has a clear and specific public health aim. Flexibility to gather and analyze big data promptly is essential in combatting the pandemic, even if it may require that the authorities collect more personal data than many people would feel comfortable with. Therefore, it is crucial that the authorities take particular care in their handling of such data and their justifications and communications to the public at large. The danger is that the people could lose trust in government, which will, as Ienca and Vayena (2020, p.1) pointed out, "*make people less likely to follow public-health advice or recommendations and more likely to have poorer health outcomes*".

Finally, although AI's impact has so far been rather limited, the pandemic and the policy responses to it may accelerate the digitalization of the economy, including the move towards greater automation of human labor. As such, the innovations in AI technology that may be an outcome of the present crisis, may require faster progress in laying down appropriate mechanisms for the governance of AI.

[29] The proper management of such extensive testing is another matter - one that will require careful consideration, see e.g. McNamara (2020).

[30] See also the "Statement on the processing of personal data in the context of the COVID-19 outbreak"' by the European Data Protection Board, available at: `https://tinyurl.com/r4r4ycj`.

References

Akhtar, M., Kraemer, M., and Gardner, L. (2019). A Dynamic Neural Network Model for Predicting Risk of Zika in Real Time. *BMC Medicine*, 17(171):https://doi.org/10.1186/s12916–019–1389–3.

Baldwin, R. (2020). COVID-19 Testing for Testing Times: Fostering Economic Recovery and Preparing for the Second Wave. *VOX CEPR Policy Portal*, 26th March.

Beck, B., B.Shin, Choi, Y., Park, S., and Kang, K. (2020). Predicting Commercially Available Antiviral Drugs that may act on the Novel Coronavirus (2019-nCoV), Wuhan, China through a Drug-Target Interaction Deep Learning Model. *bioRxiv*, 2 February(https://doi.org/10.1101/2020.01.31.929547).

Bloom, N., Bunn, P., Chen, S., Mizen, P., and Smietanka, P. (2020). The Economic Impact of Coronavirus on UK Businesses: Early Evidence from the Decision Maker Panel. *VOX CEPR Policy Portal*, 27th March.

Bogoch, I., A.Watts, Thomas-Bachli, A., Huber, C., Kraemer, M., and Khan, K. (2020). Pneumonia of UnknownAaetiology in Wuhan, China: Potential for InternationalSspread via Commercial Air Travel. *Journal of Travel Medicine*, 27(2):1–3.

Britt, R. (2020). Nobody Has a Clue How Many People Have COVID-19. *Medium*, 27 March.

Broad, W. J. (2020). A.I. Versus the Coronavirus. *The New York Times*, March 26th :(https://www.nytimes.com/2020/03/26/science/ai-versus-the-coronavirus.html).

Bullock, J., Luccioni, A., Pham, K. H., Lam, C. S. N., and Luengo-Oroz, M. (2020). Mapping the Landscape of Artificial Intelligence Applications against COVID-19. *ArXiv*, https://arxiv.org/abs/2003.11336v1.

Carroll, J. (2020). Coronavirus Outbreak: Can Machine Vision and Imaging Play a Part? *Vision Systems Design*, 19th March.

Chen, E., K.Lerman, and Ferrara, E. (2020a). COVID-19: The First Public Coronavirus Twitter Dataset. *arXiv*, 16 March(https://arxiv.org/abs/2003.07372v1).

Chen, J., Wu, L., J.Zhang, Zhang, L., Gong, D., Zhao, Y., Hu, S., Wang, Y., Hu, X., Zheng, B., Zhang, K., Wu, H., Dong, Z., Xu, Y., Zhu, Y., Chen, X., L.Yu, and Yu, H. (2020b). Deep Learning-Based Model for Detecting 2019 Novel Coronavirus Pneumonia on High-Resolution Computed Tomography: A Prospective Study. *medRxiv*, 1 March(https://doi.org/10.1101/2020.02.25.20021568).

Chun, A. (2020). In a time of coronavirus, Chinas investment in AI is paying off in a big way. *South China Morning Post*, 18 March.

Coldeway, D. (2019). Molecule.one uses machine learning to make synthesizing new drugs a snap. *TechCrunch*, 3October.

Coldeway, D. (2020). AI and big data wont work miracles in the fight against coronavirus. *Techcrunch*, 26 March.

Dewatripont, M., M.Goldman, Muraille, E., and Platteau, J.-P. (2020). Rapidly Identifying Workers who areImmune to COVID-19 and Virus-Free is a Priority for Restarting the Economy. *VOX CEPR Policy Portal*, 23 March.

Dickson, B. (2020). Why AI Might be the most Effective Weapon we have to fight COVID-19. *The Next Web*, 21 March.

Ferguson, N., Laydon, D., Nedjati-Gilani, G., N.Imai, and et al (2020). Impact of Mon-Pharmaceutical Interventions (NPIs) to Reduce COVID- 19 Mortality and Healthcare Demand. *Imperial College London*.

Flaxman, S., Mishra, S., A.Gandy, and et al (2020). Estimating the Number of Infections and the Impact of Non-Pharmaceutical Interventions on COVID-19 in 11 European Countries. *Imperial College London*, doi: https://doi.org/10.25561/77731.

Fleming, N. (2018). Computer-Calculated Compounds: Researchers are Deploying Artificial Intelligence to Discover Drugs. *Nature*, 557:S55–S57.

Galeotti, A. and Surico, P. (2020). A Uer Guide to COVID-19. *VOX CEPR Policy Portal*, 27 March.

Gruenwald, E., Antons, D., and Salge, T. (2020). COVID-19 Evidence Navigator. *Institute for Technology and Innovation Management, RWTH Aachen University*.

Hao, K. (2020). This is How the CDC is Trying to Forecast Coronaviruss Spread. *MIT Technology Review*, 13 March.

Harari, Y. (2020). The World After Coronavirus. *Financial Times*, 20 March.

Heaven, W. (2020). AI Could Help with the Next Pandemic - But Not With This One. *MIT Technology Review*, 12 March.

Heilweil, R. (2020). Facebook is flagging some coronavirus news posts as spam. *Vox Recode*, 17 March.

Hollister, M. (2020). AI can help with the COVID-19 crisis - but the right human input is key. *World Economic Forum*, 30 March.

Ienca, M. and Vayena, E. (2020). On the Responsible Use of Digital Data to Tackle the COVID-19 Pandemic. *Nature Medicine*, 27th March(https://doi.org/10.1038/s41591-020-0832-5).

Jiang, X., Coffee, M., Bari, A., Wang, J., Jiang, X., Huang, J., Shi, J., Dai, J., Cai, J., Zhang, T., Wu, Z., He, G., and Huang, Y. (2020). Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. *Computers, Materials and Continua*, 63(1):537–551.

Kreuzhuber, K. (2020). How AI, Big Data and Machine Learning can be used against the Corona virus. *ARS Electronica Blog*, 19 March.

Lazer, D., Kennedy, R., King, G., and Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176):1203–1205.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521:436–444.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020). Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV2). *Science*, 16 March(DOI: 10.1126/science.abb3221).

Maghdid, H., Ghafoor, K., Sadiq, A., Curran, K., and Rabie, K. (2020). A Novel AI-enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study. *ArXiv*, https://arxiv.org/abs/2003.07434.

Maier, B. and Brockmann, D. (2020). Effective Containment Explains Sub-Exponential Growth in Confirmed Cases of Recent COVID-19 Outbreak in Mainland China. *medRxiv*, 20 February(ttps://doi.org/10.1101/2020.02.18.20024414).

Makulec, A. (2020). Ten Considerations Before You Create Another Chart About COVID-19. *Medium: Towards Data Science*, 11 March.

Marcus, J. (2020). Big Data Versus COVID-19: Opportunities and Privacy Challenges. *Bruegel Blog*, 23 March.

Maslan, C. (2020). Social Distancing Detection for COVID-19. *Medium*, 30 March.

McNamara, A. (2020). Behavioural Scientists Very Concerned About Mass Coronavirus Testing. *BBC Science Focus*, 1 April.

McNeil, D. (2020). Restrictions Are Slowing Coronavirus Infections, New Data Suggest. *The New York Times*, 30 March.

Newton, C. (2020). The Coronavirus is Forcing Tech Giants to Make a Risky Bet on AI. *The Verge*, 18 March.

Ortutay, B. and Klepper, D. (2020). Virus outbreak means (mis)information overload: How to cope. *AP News*, 22 March.

Petropoulos, G. (2020). Artificial Intelligence in the Fight against COVID-19. *Bruegel*, 23 March.

Rawat, W. and Wang, Z. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, 29:23522449.

Regalado, A. (2020). A Coronavirus Vaccine will take at least 18 monthsif it works at all. *MIT Technology Review*, 10 March.

Rivas, A. (2020). Drones and Artificial Intelligence to Enforce Social Isolation During COVID-19 Outbreak. *Medium: Towards Data Science*, 26 March.

Rosebrock, A. (2020). Detecting COVID-19 in X-ray Images with Keras, TensorFlow, and Deep Learning. *PyImageSearch*, 16 March.

Ross, C. and Robbins, R. (2020). Debate Flares Over Using AI to Detect Covid-19 in Lung Scans. *Stat*, 30 March.

Rowan, I. (2020). What Happens to AI When the World Stops(COVID-19)? *Medium: Towards Data Science*, 31 March.

Sarkar, T. (2020). Analyze NY Times Covid-19 Dataset. *Medium: Towards Data Science*, 30 March.

Segler, M., Preuss, M., and Waller, M. (2018). Planning Chemical Syntheses with Deep Neural Networks and Symbolic AI. *Nature*, 555(https://doi.org/10.1038/nature25978):604– 610.

Silver, D., Hubert, T., Schrittwieser, J., I.Antonoglou, Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., and Hassabis, D. (2018). A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go Through Self-Play. *Science*, 362(6419):1140–1144.

Smith, S. (2018). 6 Things We Learned about Artificial Intelligence in Drug Discovery from 330 Scientists. *BenchSci Blog*, 19 September.

Song, P., Wang, L., Zhou, Y., He, J., Zhu, B., Wang, F., Tang, L., and Eisenberg, M. (2020). An Epidemiological Forecast Model and Software Assessing Interventions on COVID-19 Epidemic in China. *medRxiv*, 3 March(https://doi.org/10.1101/2020.02.29.20029421).

Stebbing, J., Phelan, A., I.Griffin, C.Tucker, Oechsle, O., Smith, D., and Richardson, P. (2020). COVID-19: Combining Antiviral and Anti-Inflammatory Treatments. *The Lancet*, 20:400–401.

Taulli, T. (2020). AI (Artificial Intelligence) Companies That Are Combating The COVID-19 Pandemic. *Forbes*, 28 March.

Vanderslott, S., Pollard, A., and Thomas, T. (2020). Coronavirus Vaccine: Here are the Steps it will Need to Go Through During Development. *The Conversation*, 30 March.

Wang, L. and Wong, A. (2020). COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images. *arXiv*, 22 March(https://arxiv.org/abs/2003.09871).

Yan, L., Zhang, H.-T., Xiao, Y., Wang, M., C.Sun, Liang, J., Li, S., M.Zhang, Guo, Y., Xiao, Y., Tang, X., Cao, H., Tan, X., Huang, N., amd A. Luo, B. J., Cao, Z., H.Xu, and Yuan, Y. (2020). Prediction of Criticality in Patients with Severe Covid-19 Infection Using Three Clinical Features: A Machine Learning-Based Prognostic Model with Clinical Data in Wuhan. *medRxiv*, 3 March(https://doi.org/10.1101/2020.02.27.20028027).