# AI-Powered Immediate Response to Pandemics

## Summaries of Top Initiatives

March 2022

GPAI | THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

*This report was developed by Experts and Specialists involved in the Global Partnership on Artificial Intelligence's project on 'AI-Powered Immediate Response to Pandemics'. The report reflects the personal opinions of the GPAI Experts and Specialists involved and does not necessarily reflect the views of the Experts' organisations, GPAI, or GPAI Members. GPAI is a separate entity from the OECD and accordingly, the opinions expressed and arguments employed therein do not reflect the views of the OECD or its Members.*


**Acknowledgements**

This report was developed in the context of the 'AI-Powered Immediate Response to Pandemics' Project, with the steering of the Project Co-Leads and the guidance of the Project Advisory Group, supported by the GPAI Responsible AI (RAI) Working Group. The GPAI RAI Working Group agreed to declassify this report and make it publicly available.

Co-Leads:
**Paul Suetens**[*], Project Co-Lead, KU Leuven
**Michael O'Sullivan**[*], Project Co-Lead, University of Auckland
**Margarita Sordo-Sanchez**[*], Brigham and Women's Hospital, Harvard Medical School

Project Advisory Group:
**Cyrus Hodes**[**], Blockchain Web Services
**Alan Paic**[†], OECD

The report was written by: **Samuel Curtis**[‡], The Future Society; **Adriana Bora**[‡], The Future Society; **Rui Wang**[‡], The Future Society; **Nicolas Miailhe**[*], The Future Society; **Sacha Alanoca**[‡], The Future Society; **Johanna Hedlund**[‡], ETH Zurich; **Florian Scheidl**[‡], ETH Zurich; **Dinithi Bowatte**[‡], Massey University - Te Kunenga Ki Pūrehuroa; **Stephanie Long**[‡], McGill University; **Mahta Ramezanian Panahi**[‡], Mila and University of Montreal; **Irene Tenison**[‡], Mila and University of Montreal; **Rishav Raj Agarwal**[‡], University of Waterloo and Akasha Imaging; and **Joseph J. Russell**[‡], CH4 Global, with the contribution of the following GPAI Experts: **Michael O'Sullivan***, Project Co-Lead, University of Auckland; and **Margarita Sordo-Sanchez***, Brigham and Women's Hospital, Harvard Medical School.

GPAI would like to thank the following individuals for their support in preparing the report: **Lewis Hammond**[†], **Bruno Kunzler**[†], **Kristy Loke**[†], and **Abdijabar Mohamed**[†], each of The Future Society, and **Karine Perset**[**] and **Francesca Sheeka**[†] of the OECD AI Policy Observatory. Finally, GPAI would like to acknowledge the tireless efforts of colleagues at the International Centre of Expertise in Montréal on Artificial Intelligence (CEIMIA) and GPAI's Responsible AI Working Group. We are particularly grateful for the support of **Ed Teather**, **Lama Saouma**, **Stephanie King**, and **Arnaud Quenneville-Langis** from CEIMIA, and for the dedication of the Working Group Co-Chairs, **Catherine Régis** and **Raja Chatila**.

* Expert of GPAI's Responsible AI Working Group
** Observer at GPAI's Responsible AI Working Group
† Invited specialist
‡ Contracted parties by the CofEs to contribute to projects



**Citation**

GPAI 2022. AI-Powered Immediate Response to Pandemics: Summaries of Top Initiatives, Report, March 2022, Global Partnership on AI.

# CONTENT



*These summaries have been reviewed and approved by the developers of the respective initiatives.

# Introduction

In 2021, the Immediate Response Project Steering Group of the AI & Pandemic Response Subgroup commissioned The Future Society (TFS) to develop an updated and upgraded catalog of AI initiatives with the potential to combat COVID-19 and other future pandemics. The research builds upon last year's report, *Responsible AI in Pandemic Response*.

An impact assessment has been conducted yielding a subset of initiatives that show promise in terms of their potential to scale, to identify those that could benefit most from partnership to deliver on their promise.

The research began in collaboration between The Future Society, the OECD, and GPAI by applying the OECD Framework for the Classification of AI Systems to classify AI initiatives based on their technical characteristics. Those associated with the development of AI systems created or repurposed to aid in COVID-19 response were invited to complete a survey shared in a public announcement. A total of 66 initiatives were identified via TFS desktop research and survey responses.

The Immediate Response Project Steering Group then worked with TFS to build upon the OECD Framework to develop a more impact-focused set of criteria, including:

- Background of the initiative (name, sources, objective/purpose)
- Origin (including organization(s), locality)
- Categorization (type of approach / AI method)
- Scope (domain, target users/operators and beneficiaries, geographic coverage)
- Data (description of the dataset in use including demographics, target population, size, collection timeframe, any public access links)

The 66 AI systems have been classified using this framework to create the Living Repository. This is being shared in an open 'work-in-progress' format in reflection of the immediate needs of the pandemic for those that may find it useful as a resource.

Using these classifications, the Immediate Response Project Steering Group conducted assessments of initiatives' intrinsic scalability and their potential to mitigate the current and future pandemics, to narrow the 66 identified initiatives into a shortlist of 26, a subset of which have been selected as candidates for potential partnerships with the AI & Pandemic Response Subgroup and GPAI more widely.

In this document, we are pleased to share summaries of the 26 initiatives. They include AI and data systems that have been developed to:

- compare how different combinations of border control strategies, home isolation, and testing at varying levels of vaccine coverage affect the risk of an infected traveler causing a community outbreak of COVID-19;
- predict the distances and angles between pairs of proteins' amino acid residues;
- support the screening phase of literature reviews and topic meta-analyses;

- determine the effectiveness of non-pharmaceutical interventions (NPIs) on COVID-19
- facilitate decision-making processes by deriving insights from a large array of near real-time data from various official and unofficial sources;
- aid policymakers' and health officials' decision-making across various topics related to the COVID-19 pandemic;
- identify individuals who are at the greatest risk of heightened vulnerability to COVID-19, based on individuals' pre-existing medical conditions;
- help policymakers simulate pandemic spread dynamics for different government control measures;
- allow users to navigate a large corpus of coronavirus-related literature;
- automatically check statistical facts about the coronavirus;
- provide users with personalized daily COVID-19 "risk scores" associated with regular activities;
- aggregate forecast data on key COVID-19 outcomes, such as cases, deaths, and hospital admissions;
- organize both structured and unstructured COVID-19 data into a knowledge graph that can be navigated and queried to retrieve information;
- illustrate how the COVID-19 pandemic could develop under different national guidelines throughout the pandemic;
- provide a country-level risk modeling framework intended to assist the government and individuals in making informed decisions;
- conduct an automatic, efficient, and detailed evaluation of the severity of COVID-19 in chest CT scans;
- automatically determine the most effective non-pharmaceutical intervention (NPI) strategies to contain the spread of COVID-19;
- quickly and accurately detect the presence of COVID-19 in thoracic CT scans;
- model the spread of COVID-19 based on the prevalence of mask-wearing in a population;
- analyze non-contrast thorax CT scans for COVID-19 pathologies;
- identify, track, and analyze events associated with COVID-19 via mentions on online news articles and social media posts;
- facilitate desktop research on topics related to COVID-19;
- query relationships between biomedical concepts, based on associations derived between terms in unstructured biomedical literature and experimental data;
- aggregate and clean various sources of US pandemic-related raw data to produce COVID-19 "indicators" for "nowcasting" (situational awareness) and short-term forecasting;
- allow users to view current occupancy rates of hospitals across the US and recommendations for intra-state patient transfers based on current occupancy rates.

We intend that the analysis will be used to help inform the Immediate Response Project Steering Group's partnerships approach in 2022, but should also provide a useful tool and model for the critical evaluation of AI initiatives within the ongoing and in future pandemics.

*Researchers tried to make these summaries comprehensive and accurate with publicly available information, but we acknowledge that they may contain errors or details that are outdated. Drafts of each summary were sent via email to the developers of each initiative for their review and approval; the summaries that have been reviewed and approved are marked with an asterisk (*) in the Table of Contents. If you are a developer of one of these initiatives and would like to correct, update, or add information, please contact the International Centre of Expertise in Montreal on Artificial Intelligence ("Centre d'expertise international de Montréal en intelligence artificielle"; CEIMIA) at info@ceimia.org.*

# Summaries

## A statistical model to predict the risk of COVID-19 infection from international arrivals to New Zealand

In early November 2021, researchers from the Department of Physics at the University of Auckland, the School of Mathematics and Statistics at the University of Canterbury, Te Pūnaha Matatini, and Manaaki Whenua Landcare Research—all based in New Zealand—released a non-peer-reviewed publication, *Effect of vaccination, border testing, and quarantine requirements on the risk of COVID-19 in New Zealand: a modeling study* [1]. In this study, funded by the New Zealand Ministry of Business, Innovation and Employment COVID-19 Programme, the Department of the Prime Minister and Cabinet, and Te Pūnaha Matatini, the researchers employ a stochastic branching process model to compare how different combinations of border control strategies, home isolation, and testing at varying levels of vaccine coverage affect the risk of an infected traveler causing a community outbreak of COVID-19 in New Zealand. The researchers aim to provide an evidentiary basis for policy strategies that allow for the safe adjustment of travel restrictions by comparing risk reduction from available policy options. Researchers involved in the development of the model were reported to have advised Prime Minister Jacinda Ardern throughout the COVID-19 pandemic [2], so while it is not explicitly stated in their paper, the model may have been used to advise government decisions in New Zealand.

The model incorporates a set of assumptions about New Zealand's vaccination program, the extent to which vaccination prevents infection, the transmission dynamics of SARS-CoV-2, as well as the accuracy of testing and contact tracing statistics. The values for these parameters are obtained from large span of scientific papers whose findings correspond to the context of this study [1]. For example, the reproduction number ($R_0$) used in this model represented the values associated with the SARS-CoV-2 Delta variant—the most common variant circulating at the time of publication. For the mean generation time and incubation period of the SARS-CoV-2 Delta variant, the researchers acknowledge uncertainties in the available literature and addressed this by performing their own sensitivity analysis with a shorter generation time and incubation period. The model was run on structured data pertaining to the number of cases detected in past arrivals to New Zealand, the reported incidence of COVID-19 in the country of origin from these travelers, and the previous and estimated future numbers of travelers. Sources for these statistics include the New Zealand Ministry of Health through Environmental Science Research New Zealand's EpiSurv database [3], the New Zealand Ministry of Health's COVID-19 2021 Vaccine dataset [4], and New Zealand FluTracking data [5]. These data are all openly accessible, but the aggregated dataset used in this study is not publicly available.

The model itself is a mixed-effects time series model that couples a stochastic age-structured model for the transmission of SARS-CoV-2 virus through a partially-vaccinated population with a simple statistical model for testing international travelers (the probability of testing positive as a function of time since exposure), to generate simulations where international travelers seed putative community transmission events under given sets of interventions [1]. This was based on a model developed by Steyn *et al*. [6] but with a policy-oriented application intended to assist New Zealand's government in gradually relaxing border controls as different stages in the nation's vaccination program are reached. Interventions investigated include vaccination requirements, combinations of pre-departure testing and post-arrival symptom screening/testing using either rapid antigen or PCR tests, post-arrival self-isolation, and

different vaccination rates in the population. In their paper, the researchers use the model to run 100,000 simulations for each combination of interventions, where each is initialized with one infected traveler to determine the transmission potential of the infected traveler and a list of resulting infections in the community [1]. The study highlights three primary outputs from the model that are relevant to decision-makers: 1) the transmission potential of infected travelers under the interventions provided relative to the absence of interventions; 2) the proportion of simulations in which an infected traveler causes transmission in the community, causes community transmission that is never detected, causes transmission that reaches five infections and causes an outbreak (defined in the paper as fifty or more infections); and 3) the number of travelers who would be expected to cause an outbreak (as defined above).

As of January 7th, 2022, this research remains confined to an article preprint (not yet formally peer reviewed), and the model itself is not publicly accessible. Because the use of this model is currently constrained to this study about COVID-19 transmission in New Zealand, data used to run the model has primarily come from public data about COVID-19 cases in and coming to New Zealand. Furthermore, the researchers note that parameters incorporated into the model are themselves specific to the context of New Zealand. For example, some assumptions about transmission dynamics have been catered to populations "that have not yet experienced large-scale epidemics" [1], like New Zealand and Australia. Thus, to adapt this model for other localities, some of these values would need to be tuned to reflect any given area's characteristics. In general, for the model to be used in the future, some parameters would need continuous updating to match developing literature about COVID-19, especially as new variants emerge, affecting not only inputs for the transmission dynamics of the virus but also infection rates in different geographical contexts. If appropriately adapted to the specificities of different localities and their conditions at a given time, the model could be useful for policymakers weighing decisions about travel policies, as it allows them to compare relative risk reductions provided by different combinations of mitigation strategies at different levels of vaccine coverage.

*More information can be found in the Living Repository.*

## AlphaFold*

AlphaFold is a deep learning model trained to make accurate predictions of the shapes of proteins [7]. AlphaFold is a product of DeepMind, an artificial intelligence research laboratory based in the United Kingdom. In March 2020, AlphaFold's developers shared the predicted structures of six under-studied proteins associated with SARS-CoV-2—created with AlphaFold—with the scientific community.

In their blog post announcing the release of COVID-19-related protein structure predictions, DeepMind researchers shared their desire to contribute to the scientific community's interrogation of the functions of viruses and for their model to serve as a hypothesis generation platform for future experimental work in developing therapeutics [8].

Since its first announcement in 2018, the AlphaFold model has gone through numerous stages of development. When announcing the release of the six SARS-CoV-2-associated protein structure predictions, DeepMind referred to their model entered in the 13th biennial Critical Assessment of Protein Structure Prediction (CASP13), dubbed "AlphaFold v1.0" [8]. This version of AlphaFold was trained on publicly-available data consisting of approximately 170,000 protein structures from the professionally-curated Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) [9] and large, open-access databases of protein sequences derived from genome sequencing projects, such as UniProt [10] The sequences of proteins associated with SARS-CoV-2 were also obtained from UniProt.

AlphaFold v1.0's model consists of two stages: (1) a two-dimensional dilated convolutional residual network that takes an amino acid sequence and, using training data, outputs the prediction of distance and torsion between amino acid residues; and (2) a differentiable model that performs gradient descent using the output of the first stage to optimize the 3-dimensional shape of a protein towards its lowest energy potential (in other words, closest to equilibrium) [7].

Following the release of the structure predictions associated with SARS-CoV-2 in March 2020, DeepMind submitted a validated, redesigned version of AlphaFold ("AlphaFold v2.0") to be assessed at CASP14. This iteration replaced the convolutional neural network in the prior model with a transformer-based architecture—the "Evoformer"—which treats the prediction of protein structures as a graph inference problem in 3D space, processing inputs through repeated layers of a neural network block to produce an array that represents the inputs in a lower dimension. This new model had the best performance by a significant margin at CASP14 held in 2020 [11]. AlphaFold v2.0 was described by its developers in the CASP conference in November 2020 and published in July 2021 [11].

Upon the publication of the paper describing AlphaFold v2.0's architecture, Deepmind also openly released its source code, trained weights, and an inference script to the research community [12]. It also partnered with European Molecular Biology Laboratory's European Bioinformatics Institute to release the AlphaFold Protein Structure Database, which includes structural predictions of all of the 20,000 proteins in the human proteome, as well as those from other biologically significant organisms, such as E. coli, yeast, drosophila, and mice.

The developers identified limitations in predicting parts of the human proteome, such as proteins that are unable to be accurately modeled with single-chain structure prediction, and must be modeled in complex or in the cellular milieu. They also note a bias towards the

human proteome for health and medicinal research, while other biologically, medically, or economically important organisms are underrepresented.

The open-source nature of this initiative's source code and the model's noteworthy accuracy in predicting protein structures suggests a high potential for impact in pandemic response, as protein structure prediction is critical for understanding viral biology and pharmaceutical design.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## ASReview against COVID-19*

ASReview is a versatile machine learning tool designed to support the screening phase of literature reviews and topic meta-analyses. Given a database with records to be screened and a minimal set of labels, the system orders the other records based on its assessment of their relevance [13]. Then, the active learning cycle starts: the most relevant record is shown to the user, who provides a label that is then used to train a new model, and the next most likely relevant record is shown to the user. The ASReview framework employs active learning, supporting multiple classifiers, feature extraction techniques, and query and balance (to compensate for underrepresented data [14]) strategies. The initiative was developed at the University of Utrecht, Netherlands, and was originally funded by the Innovation Fund for IT in Research Projects at Utrecht University.

In response to the COVID-19 pandemic, in June 2020, an extension for ASReview was developed for screening COVID-19-related literature [15]. The COVID-19 extension bridges access to two sets of data: 1) the CORD-19 database [16] maintained by the Allen Institute for AI, and, 2) a collection of preprints related to COVID-19 sourced from multiple preprint servers, maintained by researchers at Leibniz Information Centre for Economics and Utrecht University [17]. The CORD-19 database, which is openly accessible, contains full texts and metadata of over 800,000 coronavirus-related research papers published between 1996 and today [18]. It is updated on a weekly basis and is openly accessible. The second database, which is also openly accessible, contains over 66,000 COVID-19-related preprints published since January 2020 and is updated on a monthly basis.

To use the tool, the user initially chooses the database to use—e.g., meta-data containing titles and abstracts of scientific papers—and specifies the active learning model containing a classifier, a feature extraction method, a query strategy, and a balance strategy [19]. The toolbox includes multiple built-in options for each of the above-mentioned elements, or users can add their own models. To start the active learning cycle, the user chooses a set of publications that are relevant and irrelevant to their research question (the model already works with one relevant and one irrelevant record). Next, the system selects the most likely relevant record, which the user is asked to label as relevant or irrelevant. Using this label, the model re-trains and continues the cycle by selecting the next record for the user to label. This is repeated until a user-specified criterion is satisfied. Finally, the model returns a list containing labeled and unlabeled entries, the unlabeled ones sorted according to the model's assessment of their relevance. A technical log file is available, in which every decision of the model is logged, including all probabilities plus labeling decisions for every record in the data and for each iteration of the model—making the procedure both transparent and reproducible.

The tool can also be used for literature reviews on other topics. In this case, the user creates the database and imports it into the software. The remainder of the active learning cycle works as previously described.

In a 2021 *Nature Machine Intelligence* publication, the developers of ASReview performed four simulation studies to test the model's performance on systematic reviews conducted between 2012 and 2020 [20]. These each contained between 2,500 and 8,900 publications and inclusion rates between 0.66-4.84%. The active learning cycle involved presenting one random relevant and one irrelevant record to the user. After sorting the remaining records according to relevance, the system is evaluated based on the following set of performance metrics: the work saved oversampling at a given level X ("WSS@X"), which measures the reduction of screening effort at the cost of missing (1-X)% of relevant items; and the fraction

of relevant records ("RRF10") after screening the top 10% of the sorted records. A randomly sorted list of publications serves as a control. Their results demonstrated the utility of the tool: when screening studies according to the sorted list, on average 95% of relevant studies were found after screening 8-33% of the studies. Moreover, the RRF10 ranged between 70 and 100%.

The ASReview software and the COVID-19 extension are both freely accessible with an Apache 2.0 license [15], [21]. The software appears to be updated at least quarterly, and the underlying data are updated on a weekly-to-monthly basis. As ASReview returns a list of publications ordered according to their relevance according to the model, it is more interpretable than comparable systems that simply assign the labels "relevant" and "irrelevant;" however, the choices made by the underlying classifiers may be difficult to interpret (after participating in the active learning, the user only sees an ordered list as an output, according to the predictions by the model based on the labeling decisions of the screener). This limitation notwithstanding, the tool constitutes a potentially impactful approach to accelerating scientific research, particularly in COVID-19-related research.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## Bayesian hierarchical semimechanistic model*

This initiative developed a Bayesian hierarchical semimechanistic model to determine the effectiveness of non-pharmaceutical interventions (NPIs) on COVID-19 transmission [22]. The initiative was developed by researchers at the University of Oxford, Australian National University, the Quantified Uncertainty Research Institute, Harvard University, the University of Bristol, the University of Manchester, the London School of Hygiene and Tropical Medicine, the London School of Economics and Political Science, the University of Cambridge, Tufts University, and Imperial College London. The initiative is presented in an academic paper published in *Science* in February 2021 [22].

The rationale behind this initiative was to provide an alternative to simulation studies, which tend to make strong assumptions that are relatively difficult to validate, by developing a data-driven, cross-country model that compares national interventions to the subsequent numbers of cases or deaths within those respective regions.

NPI data were collected across 41 countries from January 22nd to May 30th, 2020. To mitigate errors, all NPI data were entered independently by two of the authors, using primary sources, and then manually compared with two public datasets: the Epidemic Forecasting Global NPI [23] and the Oxford COVID-19 Government Response Tracker [24]. Data on confirmed COVID-19 cases and deaths were obtained from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [25]. To prevent bias, data were pre-processed by neglecting COVID-19 case numbers before a country had reached 100 cases, and fatality numbers before 10 deaths [22].

The model presented in this paper was built upon the semimechanistic Bayesian hierarchical model developed by Flaxman *et al* [26], which estimated the effects of NPIs on COVID-19 transmission in Europe. Similar to Flaxman's approach, this model used COVID-19 case and death data to make a 'backward' inference of the number of new cases for each country, which was then used to infer daily reproduction numbers [22]. The reproduction number and the occurrence of NPIs were then used to estimate NPI effects. To account for cross-country variations in effectiveness, reporting, and fatality rates, as well as uncertainty in the generation interval and delay distributions, researchers utilized a Markov chain Monte Carlo (MCMC) sampling algorithm [27] to infer posterior distributions of each NPI's effectiveness.

The researchers found that NPIs demonstrated highly consistent trends across countries [22]. For instance, closing both schools and universities was consistently highly effective at reducing COVID-19 transmission, as was banning gatherings of 10 people or fewer, whereas targeted closures of face-to-face businesses with a high risk of infection, such as restaurants, bars, and nightclubs, had a small-to-moderate effect. They furthermore found that when most NPIs were already in place, stay-at-home orders had only a small additional effect; thus, by using effective interventions, some countries could effectively control COVID-19 spread while avoiding stay-at-home orders.

The researchers point out numerous limitations of their approach, such as an inability to factor in country demographics, regional differences in interpretations or implementations of NPIs, and a lack of data on some NPIs not captured in this study, which may restrict the feasibility of scaling up the tool [22]. Some of these limitations may resolve with time, however, if NPIs were to become more standardized across larger geographies, and as

more COVID-19 case and death data becomes available before and after the implementation of various NPIs.

A more recent study by the same team focused on Europe's second wave shows that business closures, educational institution closures, and gathering bans had smaller effects on reducing transmission compared with the first wave, likely due to organizational safety measures and individual protective behaviors which made various areas of public life safer [28].

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## BlueDot

BlueDot is a software as a service (SaaS) designed to facilitate decision-making processes by deriving insights from a large array of near real-time data sourced from various official and unofficial sources [29]. The BlueDot platform's algorithm aggregates hundreds of datasets and deploys statistical models to detect anomalies, such as potential disease outbreaks, and predict how and where diseases might spread. BlueDot is a successor to BioDiaspora, a scientific research program studying how the world's population is connected through commercial air travel, launched in 2008 by Dr. Kamran Khan, a physician and infectious diseases specialist [30]. BioDiaspora received its first significant capital financing from Horizon Ventures in 2014 and was renamed BlueDot. In 2018, BlueDot launched its early warning system, Insights. Today, BlueDot is maintained by a team of more than 80 employees with expertise in health science and data analytics [30].

The BlueDot platform offers two core components: Insights, which sends clients near-real-time infectious disease alerts [31]; and Explorer, a cloud-based GIS platform that integrates diverse datasets, including air travel and disease surveillance data [32]. During the COVID-19 pandemic, within Insights, BlueDot has offered two products called "COVID Data Suite," for the "latest intelligence about COVID-19, including sub-national data and epidemic curves," and "COVID Focus Reports," which "deliver curated research on COVID-19 that examines where the pandemic is heading" [33].

Every fifteen minutes, an algorithm pulls structured health data, such as official data from the Center for Disease Control or the World Health Organization, as well as unstructured data, including worldwide movements of commercial flight travelers, human, animal and insect population data, and climate data from satellites. and local information from journalists and healthcare workers [34]. In total, approximately 100,000 online articles are identified each day, spanning approximately 65 languages and over 150 diseases and syndromes around the world. The datasets produced are not publicly accessible.

BlueDot's specialists have reportedly developed a taxonomy so that relevant keywords could be scanned efficiently, and then machine learning and natural language processing can be applied to train the system and automatically classify, so that only a handful of cases have to be flagged for human experts to analyze [34]. Upon detecting anomalies in data, such as an accumulative occurrence of unexplained symptoms, users are provided an alert of these warning signs via the Insights feature [33]. Users are able to track in near real-time disease spread, hospitalization as well predictions on the interactive dashboard provided by the Explorer feature. No further information pertaining to the data, model, training, nor testing has been disclosed.

The early-warning system has proven effective in various disease outbreaks, including the ongoing COVID-19 pandemic. On December 31st, 2019, anomalous reports of undiagnosed pneumonia in Wuhan, China were detected on the Insights feature. Two hours after the risk detection, an alert was sent to BlueDot customers [5]. Based on air travel data such as from global airline ticketing, the Insights feature identified 20 cities that might be the first to be affected by the outbreak. Of these, 12 were later found to be among the first cities that reported early cases [4]. These alerts preceded both a warning sent by the US Centers for Disease Control on January 6th, 2020, and the World Health Organization on January 9th, 2020 [35]. By the end of January 2020, BlueDot developers had published two separate articles in the *Journal of Travel Medicine*—one on January 10th, 2020 [36] and one on January 27th, 2020 [37]—warning of the global spread of this "pneumonia of unknown

aetiology” based on commercial flight data. The BlueDot system has demonstrated similarly prescient forecasts of earlier viruses, including Ebola in 2014 [38] and Zika in 2016 [39].

BlueDot offers all of the aforementioned services to licensed clients; website advertising is tailored to those in the business, government, and healthcare sectors [30]. In addition to the advantages that BlueDot has conferred in terms of COVID-19 spread forecasting, the founder emphasizes the system's potential to track future anomalies that could be a sign of the next high-risk disease–potentially breaking out before the COVID-19 pandemic has subsided [40]. However, lack of public access to the platform, data, and models not only make it difficult to assess BlueDot’s potential scalability and impact, but is itself a barrier to the platform’s scalability and impact.

*More information can be found in the Living Repository.*

## C-19 Index*

The C-19 Index is an open-source, AI-based predictive model designed to identify individuals who are at the greatest risk of heightened vulnerability to COVID-19, based on individuals' pre-existing medical conditions [41]. The C-19 Index was developed by researchers at ClosedLoop.ai, a private healthcare software company, and one researcher affiliated at Healthfirst, a New York-based health insurance company. A research article describing their models was uploaded to medRxiv in March 2020 [42] and was published in the Journal of Medical Artificial Intelligence in December of 2020 [41].

In their research article, the researchers note that identifying who is most vulnerable to COVID-19 complications or death is not straightforward; however, patterns that were emerging in data from Wuhan and the US (in early 2020) suggested that the risk of death increased with age, for those who have diabetes, heart disease, blood clotting problems, or have shown signs of sepsis. Researchers believed that building predictive models based on these known risks could be useful for outreach campaigns targeted to those most at risk of severe COVID-19 complications [41].

Researchers used data from two different datasets to train their models: the Center for Medicare & Medicaid Services Limited Data Set for 2015 and 2016 [43], and a medical claims dataset containing 2.5 million Healthfirst insurance beneficiaries. Each dataset represented different US demographics: the former contained data for those over the age of 65 or disabled who receive Medicare, while the latter contained data from overall healthier adults enrolled in Medicaid. Cohorts were created from each data set, and then the resulting cohorts were combined, such that the combined cohort had an age profile consistent with the overall US population [41].

Three different models, which output a person's "C-19 Index" score—the percentile risk of near-term severe complications from an upper respiratory infection, were then trained on the combined cohort's data: (1) a "survey risk factors" logistic regression model that outputs a person's percentile risk score based on responses to a web-based survey; (2) a "diagnosis history model," which train gradient-boosted trees in a time-delayed fashion, allowing the model to use current claims data by simulating the 3-month delay in claims processing that usually occurs in practical settings; and (3) an "expanded feature model," a model built within ClosedLoop—a software system for creating machine learning models—that uses additional engineered features from peer-reviewed studies (not disclosed in their publication). The key differences between each model are the number of features each employs, and thus, their ease of implementation.

The validation dataset contained 14,000 COVID-19 cases in New York City from February 2020 until mid-May 2020. The logistic regression used the fewest features and delivered the lowest performance, with an AUROC (Area Under the Receiver Operating Characteristics—a measurement of a model's ability to distinguish between classes) of .731. In comparison, both the diagnosis history model and the expanded features models obtained AUROCs of .810.

The C-19 Index has already been utilized by at least two healthcare organizations: Medical Home Network, an Illinois-based accountable care organization [44], and Healthfirst, the aforementioned New York-based health insurance company. However, the authors note several limitations of their study which impact the feasibility of using this approach on a larger scale: no real COVID-19 cases were used in the model's training, the approach relied on claims data instead of clinical data, and data excluded those under 18 years of age.

Therefore, moving forward, possible technical enhancements could be to validate the proxy outcome and determine their validity based on COVID-19 data, to build models on COVID-19 vulnerability on COVID-19 data (without having to use other upper respiratory diseases as proxies), and to test on data of those under 18 years of age.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## CAIAC*

CAIAC (Collective and Augmented Intelligence Against COVID-19) is a prototype of a knowledge management platform developed to aid policymakers' and health officials' decision-making across various topics related to the COVID-19 pandemic [45]. The project was launched in July 2020 as a collaboration between Stability.ai, The Future Society, and the Stanford Institute for Human-Centered AI (HAI), with support from Patrick J. McGovern Foundation [46].

The objective of CAIAC is to assist users with decision-making by providing them with the means to navigate authoritative and up-to-date information with a combination of information representation and query tools [47]. The prototype was developed with three initial COVID-19-related use cases in mind: contact tracing, targeting aid towards marginalized groups, and addressing the "infodemic." The platform prototype integrates a number of AI systems into three functionalities: 1) knowledge graphs, in which actors and initiatives (for each use case) are represented by nodes; 2) informative briefs pertaining to the key questions identified within each use case; and 3) a query interface that allows users to search for information within the CAIAC dataset.

CAIAC relies on a combination of structured and unstructured data. Several hundred documents—academic articles, preprints, publications by public health authorities, and news media—from each use case were annotated by humans with related expertise (at least an undergraduate degree in a related field) [47]. Human annotators used software called Hypothesis [48] to parse COVID-19-related literature and extract key entities, topics, statistics, and quotations. Using a bespoke Named Entity Recognition (NER) model with software called Hivemind [49], these thousands of extracted strings were delegated to humans to assign labels (such as "individual," "organization," "quotation," etc.). Data were also collected via expert interviews: within each use case, interviews were conducted, with permission, with 10 domain experts—those highly cited or in positions of authority in authoritative medical and/or scientific institutions. These interviews were transcribed with Otter AI [50] and then labeled in a similar manner to the documents above.

These structured data are visualized on three knowledge graphs—one for each use case—using Kumu software [51]. Informative briefs provide users with a walkthrough of each graph and key questions (those which had been identified as significant in literature or in expert interviews) pertaining to each use case. Users are also able to execute search queries with a chatbot interface, enabled by IBM Watson Discovery [52], which provides users with excerpts from the data most likely related to their query. The details pertaining to the Natural Language Processing (NLP) model(s) employed by IBM Watson Discovery are not publicly available. Work has been ongoing to replace the IBM Watson Discovery-powered chatbot with a language model capable of reliably accurate Natural Language Generation (NLG) output in response to COVID-19 related queries. In this regard, a 6 billion parameter GPT-J (an open-source alternative to GPT-3) language model, was fine-tuned with the COVID-19 Open Research Dataset (CORD-19) and the aforementioned data sets [47]. As of January 7th, 2022, the text generated by this model has not been demonstrated to be reliably accurate and, and as such, an interactive version is not publicly accessible, but interested parties can access an API with a key available upon request by developers [53].

The prototype of CAIAC is publicly accessible [45]. A goal of CAIAC was to explore methods to systematically collect, structure, and make navigable data in domains where literature is quickly emerging [47]. The data collection and labeling processes, as described, are labor-intensive; in order to be scaled across other domains, these tasks could be replaced by automated named entity recognition (NER) methods. Furthermore, the development of CAIAC demonstrated that even some of the most powerful language models, such as GPT-J fine-tuned with COVID-19-related research, are not yet capable of providing reliably accurate summaries from the large body of disparate data sources. Indeed, language models (such as OpenAI's GPT-3) are only just beginning to demonstrate the competency to summarize bodies of text from a single source, such as a book [54]. Thus, improvements in information management and NLG methods would likely be necessary before CAIAC could feasibly be scaled up for other use cases.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## Compositional Cyber-Physical Epidemiology of COVID-19

"Compositional Cyber-Physical Epidemiology of COVID-19" is a framework for simulating the progression of the COVID-19 pandemic based on the enforcement of national non-pharmaceutical interventions (NPIs) [55]. Using estimates for the effect of NPIs on the reproduction number ($R_0$) [56], the approach allows policymakers to simulate pandemic spread dynamics for different government control measures. A paper describing the framework, by researchers at the Department of Electrical and Software Engineering and the Faculty of Medical and Health Sciences at the University of Auckland, was uploaded to medRxiv in May 2020 [57] and was published in *Nature Scientific Reports* in November 2020 [55].

For their model, the authors fitted parameters based on COVID-19 case data provided by the COVID-19 Data Repository [25] (maintained by the Center for Systems Science and Engineering at Johns Hopkins University) for New Zealand between March and September 2020, and for Italy between February and August 2020. The dataset is openly available, and the authors have shared the subset of data they worked with on GitHub [58].

In this framework, disease spread dynamics are modeled as a cyber-physical system, composed of parameterized ordinary differential equations (ODEs), representing the disease spread, and a discrete controller that changes the parameters of the system, representing government interventions [55]. The main parameters of the model are the reproduction number ($R_0$) and transition probabilities between epidemiological states, as defined in the Susceptible, Exposed, Infectious, or Recovered (SEIR) model. The tool does not rely on training; instead, its parameters are inferred from data. For New Zealand, $R_0$ was derived from a formula that takes into account interventions—each with their own weight—that are employed at every phase of lockdown (0-4). For Italy, which lacks such a systematic intervention strategy, the authors divided the transmission trajectory (time) into four phases reflecting the stringency of measures taken by the Italian government in those periods, and fitted the reproduction value for each of these phases using Least Squares curve-fitting with the SEIR model. From here, the tool allows the user to specify a set of control measures that change pandemic dynamics and are switched based on the state of the system. For example, when a critical threshold of new cases per day or percentage of occupied intensive care units (ICUs) is reached, more stringent policy measures will be implemented, which is encoded in the model by a lower reproduction number. Thus, the tool implements a state machine that automatically switches the parameters of the ODE model based on the values of the system.

The model's performance was measured with respect to COVID-19 Data Repository for Italy between February and August 2020 and for New Zealand between February and September 2020. To test the model's performance, the authors implemented control measures that mimic the interventions taken by governments of Italy and New Zealand, in addition to hypothetical interventions, in order to assess the differences between different restriction strategies. It was observed that this model accurately captured the course of the pandemic in New Zealand and Italy over the amount of time for which data existed (approximately 170 days). After that point in time, authors examined how different levels of stringency—reflected in the controller—would impact future outcomes. These tests demonstrated that a controller that reflects New Zealand's tiered lockdown strategy results in better economic outcomes than a "simple" controller (e.g. locking down entirely until infections reach zero or vaccines are available). It also demonstrated that a tiered strategy performs better than an oscillating strategy—alternating between no lockdown and complete lockdown based on active infections.

This framework is openly accessible on the developers' GitHub page and was last updated in September 2020 [58]. The framework as presented relies heavily on accurate parameter estimates, such as the reproduction number in the SEIR model. The authors note that due to regional differences in implementation, population density, mobility, spatial heterogeneity, economic factors, and cultural characteristics, it can be difficult to reliably estimate the effects of government interventions on this figure [55], [59]. However, they also point out that the framework need not be limited to the SEIR model—it could be extended to other continuous models that can be captured through a series of ODEs, such as those demonstrated by CovidSIM [60], SIDARTHE [61], or microscale modeling. Thus, for any particular geography, research would be required to determine how this approach could be best utilized.

*More information can be found in the Living Repository.*

## CoronaCentral

CoronaCentral is a dashboard that allows users to navigate a large corpus of coronavirus-related literature for SARS-CoV-2, MERS-CoV, and SARS-CoV, where articles can be accessed and searched using predefined categories and terms. The portal was developed by Jake Lever and Russ B. Altman, based at the Department of Bioengineering at Stanford University, and supported by the Chan Zuckerberg Biohub and a National Library of Medicine Grant [62]. The earliest commit to the dashboard's GitHub repository is June 2020 [63], and the work was published as an article in Proceedings of the National Academy of Sciences of the United States of America (PNAS) in April 2021 [62].

Articles from the COVID-19 Open Research Dataset (CORD-19) [16] and a selected set of articles containing relevant coronavirus keywords PubMed are processed on a daily basis [62]: URLs of the latest CORD-19 dataset is fetched from the dedicated CORD-19 releases page [18] and metadata CSV files are extracted; PubMed file URLs are fetched from the FTP listing [64] and the XML files are downloaded then filtered for documents that appear to be coronavirus-related. Both PubMed and CORD-19 files are combined into a JSON format. The data are fed through a pipeline that cleans the data, removes duplicates, merges the data, performs categorization, and then uploads data to a MySQL database [62]. Full details are published on the author's GitHub [63], and as of January 7$^{th}$, 2022, CoronaCentral has processed 274,043 articles [65].

To mine the downloaded data for topic, article type, and typical identifier (e.g. author(s), title, journal, year), a Bidirectional Encoder Representations from Transformers (BERT) natural language processor was employed as a machine learning model, trained by a supervised learning approach. An initial 1000 randomly selected articles were manually evaluated to produce a draft list of topics and article types. This was then adjusted to provide better coverage, where further topics were added (e.g. long haul, contact tracing) later in the pandemic. Cross-validation using a 75%/25% training/validation split was used to evaluate BERT-based document classifiers [66]. Topics and article types were predicted together using the title and abstract as input: multi-label classifiers were implemented using ktrain and HuggingFace models [66]. The best BERT model was evaluated on a held-out test set of 500 articles [67]. Optimal parameters were 32 epochs, a learning rate of $5^{-5}$, and a batch size of 8. This had a macro precision of 0.805 and a macro recall of 0.76 [67]. The macro F1 score, which is based on a balance of precision and recall with 1 being the best value, was 0.774 [67].

The data are presented through an online dashboard where users are presented with interactive graphs depicting trends in publications by coronavirus species (SARS-CoV-2, MERS-CoV, and SARS-CoV) and recent articles that are trending based on altmetric scores. Other interactive graphs present publications by topic, location, article type, source (peer-reviewed versus preprint servers), drug (non-vaccine), vaccine, risk factors, symptoms, genetic variation, and viral lineage [65]. Users can execute a query for articles or browse by category (e.g. epidemiology, treatment) and subcategory (e.g. contact tracing, vaccines). The project authors note that topics ranking highest in terms of frequency of publication (e.g. Clinical Reports ranking highest at over 15,000 articles) are overtaken in the rankings by altmetric score (e.g. Transmission ranking highest with over 15 articles in top 100 altmetric score). This suggests that in the context of a dynamic research environment, such as a pandemic, users may wish to rely on an alternative means by which to prioritize research, rather than more canonical citation indices [62]. User feedback is possible via a form on the website, to flag issues and contact the authors, for instance, in the event of mistaken categorization and missing files.

Because a large number of articles are added every month (approximately 10,000), it is reasonably foreseeable that new topics will emerge. The project's authors note that as the breadth of coronavirus literature grows, they may need to add new topics as research focuses shift [62]. Monitoring of trending articles will help identify and verify that topic drift does not noticeably reduce machine learning quality. Presumably, the model will need retraining as the pandemic evolves. The project's GitHub page lists regular updates up until at least October 2021; formal re-releases of new versions may be necessary as an assurance of ongoing integrity of the portal.

*More information can be found in the Living Repository.*

## CoronaCheck

CoronaCheck is an online tool for automatically checking statistical facts about the coronavirus based on data in the COVID-19 Data Repository [25] maintained by the Center for Systems Science and Engineering at Johns Hopkins University [68]. Given a claim provided by the user, the system assesses the claim's validity and provides users with an explanation for how their statement was assessed [69]. This tool is the result of a collaboration between Professor Paolo Papotti's lab [70] at EURECOM and Professor Immanuel Trummer's lab [71] at Cornell University in March 2020 [68].

CoronaCheck is a COVID-19-oriented, proof-of-concept instantiation of a generalized system called 'Scrutinizer,' described in a paper published in *Proceedings of the VLDB Endowment* in September 2020 by the same researchers [72]. The CoronaCheck instantiation involves an algorithm that uses pre-trained language models, in particular, Facebook's cross-lingual language model (XLM) [73, p.], with task-specific fine-tuning to verify single statistical claims submitted by users against authoritative "ground truth" datasets.

CoronaCheck relies upon three categories of data: those used to fine-tune a language model, those provided by users (claims to be evaluated), and those that serve as the source of ground truth. Fine-tuning data consisted of 3 million true claims based on the ground truth dataset, generated based on a set of template sentences, which were syntactically varied to produce robust classifiers in a process described on their Github [74]. Usage data, consisting of text inputs provided by the user, are collected to improve classification in a process described below. Usage data is supplemented by personal data, which is provided to Google Analytics in accordance with CoronaCheck's privacy policy [75]. Ground truth data are obtained from the Center for Systems Science and Engineering at Johns Hopkins University's COVID-19 Data Repository, which aggregates statistics on COVID-19 cases, deaths, and recovery, as well as national, regional, and state-level population [25]. These data are sourced from the World Health Organization and numerous Centers of Disease Control among other institutions and have been updated daily since January 21, 2020 [25].

CoronaCheck processes a user's input claim in four steps: First, the input—a user's statistical claim—is mapped to real-valued vectors using pre-trained embedding functions [68]. Next, the embeddings are passed to classifiers, which match them to a set of dataset queries. The query results are processed and the results are displayed to the user, containing an assessment of the claim and explanations of the results. If the classifiers have low confidence (with respect to a predefined threshold), the system invokes human fact-checkers by generating questions based on the submitted claims, using the classifiers and language models. Based on the classifiers' rankings, the algorithm provides multiple answer options to the human fact-checkers, who answer the question, thereby updating the classification result. This way, new training samples are created, which can improve classifiers' performance.

The accuracy and efficiency of CoronaCheck were tested in comparison to three other state-of-the-art query generator solutions—Tapas [76], TabFact [77] AggChecker [78]—used to translate text claims to database queries. CoronaCheck outperformed all three in its ability to support its claims and its accuracy of claim verification. Furthermore, it took less time to train and to compute verifications [72]. Tests were also performed to assess the amount of time required for domain experts to verify claims with CoronaChecker versus manual methods (using a search engine), and it was demonstrated that, overall, it

took experts less than half of the time to verify a claim with the assistance of CoronaCheck than without it [72].

CoronaCheck is freely accessible on its dedicated page on the EUROCOM site [69], where a terms of use statement notes that the tool should only be used for educational and academic research purposes, not for medical guidance. It furthermore notes that it relies upon publicly-available data from multiple sources that do not always agree. In addition to English, CoronaCheck can also be used in Italian, French, German, Spanish, Arabic, and Traditional Chinese; however, the performance of the tool in these languages has not been presented. Developers also note that an API is available upon request. The code for Scrutinizer is accessible on the developers' GitHub page [74], and its technical documentation [72] is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

*More information can be found in the Living Repository.*

## COVI*

COVI is a research project that culminated in the development of an AI-enabled contact-tracing mobile application. This application aimed to provide users with personalized daily COVID-19 “risk scores” associated with regular activities (such as taking public transportation and socializing with friends), using smartphones to monitor proximity and accumulate relevant statistics. These scores were based on users’ demographic and health profiles [79]. The project commenced in May 2020, led by researchers at Mila (a research institute specializing in artificial intelligence), Canada, with affiliations including the University of Ottawa, Universite de Montreal, The Alan Turing Institute, University of Oxford, University of Pennsylvania, McGill University, Borden Ladner Gervais LLP, The Decision Lab, HEC Montreal, Max Planck Institute, Libeo, and the University of Toronto.

The primary objective of the mobile application was to help members of the general public make informed, risk-reducing decisions in a manner that preserved individual privacy [80]. COVI extends beyond a contact-tracing mobile app by combining contact-tracing information with other user data (e.g., user demographics, health information, symptoms) to predict daily personal risk factors for each user. In addition, COVI translates these personal risk scores into recommendations based on public health guidelines. Finally, collected data is used to define epidemiological models and intervention simulations, which could then be shared with public health officials to help them preempt the resurgence of the virus and inform reopening strategies [80].

COVI was developed to utilize a variety of data from users, all of which would be obtained by consent. Upon opening the app for the first time, users would be provided with an overview of how the app works and the privacy implications of sharing data with COVI. It then would ask for consent for the collection, use, and disclosure of IP-based geolocation history, random “contact” IDs (generated when a phone is within 2 meters of another phone with COVI installed), and users’ current risk levels — all necessary for the app to function optimally [80]. If a user were to start presenting symptoms or be diagnosed with COVID-19, they could report accordingly. Then, contacts made with that user within the past 14 days would be notified, and the symptoms/diagnosis would be factored into the computation of the contacts’ risk scores [80].

COVI also asks for consent for the collection and use of data pertaining to a user’s age, sex, health conditions, active symptoms, ongoing relevant behavior, coarse geographical location, and app analytics information; all of these data (except analytics information) would be fed into the application’s risk assessment function, which would compute locally on the user’s device [80]. Data remains on a user’s device unless the user opted to allow COVI Canada to receive encrypted, pseudonymized data packets and heat-map information (in aggregated form), which would be used by COVI’s underlying ML model and assist in epidemiological research by government-related or research-related third parties [80].

Once collected, data can be used to train deep learning ML models to predict contagiousness risks, and to fit an epidemiological model [80]. COVI deploys architectural scaffolding for deep learning around a Transformer architecture, which draws upon information pertaining to demographics, behaviors, health conditions, symptoms, and contact with other users [81] to dynamically refine the ML model. In addition, the data shared by the app users also enables AgentSim, an agent-based simulator that offers flexibility in designing contact-tracing and epidemiological simulations, to identify new

patterns and specific parameters (such as distance, sex, and age) to model how the virus spreads [82].

Mila was aiming for COVI to be endorsed for use by the Canadian government, in early June of 2020. Despite putting forward a demonstrated effort towards building a "privacy-conscious app" [79] and fostering public trust by explaining the rationale behind their app design decisions in their white paper and in numerous public appearances (newspapers, television and radio), the Canadian government decided to endorse a different application that collected less personal data, citing privacy concerns by provincial and territorial leaders [83].

COVI's code and documentation remain accessible on Mila's website (hosted on GitHub [81] and arXiv [80]) as open-source, with a non-exclusive and royalty-free license, "should [others] wish to deploy an AI-enabled health app inspired by our approach" [79]. The open-access nature, human-centric privacy protocols, and consensual use of encrypted, pseudonymized user data suggest that this tool has a high potential to scale to other geographies.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## COVID-19 Forecast Hub

The COVID-19 Forecast Hub is a central repository for forecast data on key COVID-19 outcomes, such as cases, deaths, and hospital admissions across the United States. It also aggregates forecast data from a large community of academic and corporate prediction modeling teams to build short-term (i.e., one- to four-week) ensemble forecasts, which have been demonstrated to outperform individual forecasts [84], [85], and can be explored on an interactive visualization hosted on its site [86]. The COVID-19 Forecast Hub was created by the Reich Lab of the University of Massachusetts Amherst in March 2020 and has been maintained since then by researchers at the University of Massachusetts Amherst, Iowa State University, Carnegie Mellon University, as well as several international research groups [87]. Research has been funded by the U.S. Centers for Disease Control and Prevention and the US National Institutes of Health [87], with some contributors receiving additional funding from public health, academic, or philanthropic institutions [88]. Notably, the U.S. Centers for Disease Control and Prevention rely on forecasts from the COVID-19 Forecast Hub for their official communications on the COVID-19 outbreak's trajectory [85], [89].

The forecast data collected by the COVID-19 Forecast Hub includes daily and weekly deaths, hospitalizations, and incident cases from all 50 US States, Washington DC, and the 4 territories, as well as aggregated values at the national level, submitted by 82 independent modeling teams (as of January 7$^{th}$, 2022) [85]. In developing forecasts, the Reich Lab utilizes U.S. COVID-19 death, hospitalization, and incident case data reported by the Center for Systems Science and Engineering at Johns Hopkins [90] to develop their forecasts [85], but other modeling teams may rely on other sources for statistics. Forecast data are organized by date, target horizon/end date, location, and type of prediction (point vs. quantile forecast) [85]. The Reich Lab has made these forecast data freely available on Github [91] and a Zoltar API [92]. Additionally, the researchers developed numerous R packages [93], [94] to facilitate the retrieval of forecasts for analysis.

The COVID-19 Forecast Hub uses an ensemble model incorporating forecasts of key COVID-19 outcomes from multiple different models into a single, combined ensemble forecast [95]. Models eligible for inclusion into the ensemble model were built using a wide range of approaches (statistical, machine learning), and were submitted as *point predictions* and/or *quantile predictions* over four-week horizons [85], [88]. Point predictions refer to the single best "prediction" with no uncertainty [88], whereas quantile predictions store predictive distributions, i.e., predictive medians and quantiles [96]. The ensemble is a "median" forecast, meaning that it takes into account all component forecasts equally, as opposed to more "sophisticated" ensemble methods, such as using trained ensemble datasets optimized with weighted interval scores. The ensemble forecasts, as well as all of the forecasts that comprise it, can be explored on an interactive visualization hosted on the COVID-19 Forecast Hub website [86].

To evaluate the performance of their ensemble model, researchers compared the predictions of their ensemble model to 23 individual models that comprised the ensemble (within the testing time frame) to "ground truth" statistics provided by the Center for Systems Science and Engineering at Johns Hopkins [85], [90]. This testing revealed that the ensemble model consistently outperformed the individual models that comprised it [85]. Further testing also revealed the median ensemble method was more consistently accurate than weighted ensemble methods, leading the team to rely on the former [97].

As of January 7th, 2022, the COVID-19 Forecast Hub has collected over 100 million rows of forecast data and continues to collect more [1]. An advantage of the COVID-19 Forecast Hub's ensemble forecast approach is the relative ease of transferring such an approach to new geographical contexts: in areas where accurate COVID-19 outcome data exist and more than one ensembling team is producing COVID-19 forecasts, ensemble models may be constructed. In fact, this ensemble method can, and has, been adapted in other contexts, for example, the European Covid-19 Forecast Hub [98] and the German and Polish COVID-19 ForecastHub [99].

*More information can be found in the Living Repository.*

## COVID-19 Hospital Capacity Management

COVID-19 Hospital Capacity Management is a publicly-accessible online dashboard that allows users to view current occupancy rates of hospitals across the US and recommendations for intra-state patient transfers based on current occupancy rates. It also provides an interactive tool to modify model parameters (eg. patient type, transfer budget, transfer distance threshold) to obtain more customized recommendations within any state or hospital system within the US [100]. The tool was developed by a team from Johns Hopkins Center for Systems Science and Engineering and Malone Center for Engineering in Healthcare and is affiliated with the Center for Data Science in Emergency Medicine and the Department of Civil and Systems Engineering at Johns Hopkins University [100]. The initiative was first publicly announced on October 27, 2020 [101], and the team's first related academic preprint was uploaded to arXiv on November 6, 2020 [102].

In their preprint, the developers state that the motivation behind their effort was to minimize resource shortages, which would, in turn, improve the overall quality of patient care and prevent early discharges and cancellations of elective surgeries [102]. They note a few instances of patient transfers occurring ad-hoc in the COVID-19 pandemic, but remark that treating this issue at a more protracted, system-level—across hospitals, counties, and states—will spur more efficient resource use. They also recognize the alternative approach of hospitals individually and reactively responding by creating surge capacity, but point to studies suggesting that such an approach can lead to a reduced quality of care compared to hospitals working in coordination to make use of existing resources.

For data on past hospital occupancy and COVID-19 hospitalizations, the dashboard relies on statistics provided by the US Department of Health and Human Services [103]. To make future projections, the team uses the US Center for Disease Control's county-level forecasts of COVID-19 cases [89]—an ensemble of models from many forecasting teams, which the researchers behind COVID-19 Hospital Capacity Management then disaggregate to the hospital level.

To make recommendations for patient redistribution, the researchers constructed a series of linear optimization (linear program and mixed-integer linear program) models to solve a multi-period demand problem: "given a set of nodes and time periods, with nominal demand (ie. COVID-19 patients) at each node during each period and fixed capacity at each node, determine the optimal quantity of demand to transfer between each pair of nodes during each time period" [102]. To better reflect real constraints, they extended the model by adding several parameters to each node, including the type of patient (ICU vs acute care), the per-transfer hospital budget, the total transfer budget, percentage of capacity reserved for COVID-19, transfer distance limits, and lengths of stay, among others. On the COVID-19 Hospital Capacity Management site, the output of the models—recommended intra-state patient transfers—are presented on dynamic graphics for each US state. In their paper, the researchers also presented an analogous multi-period method to model critical redistribution (rather than patient transfers), however, this was not presented on their dashboard.

The online dashboard remains accessible and up-to-date with data updated weekly, and the source code publicly accessible on their GitHub repository [104]. Limitations to scaling up this approach include the accessibility and quality of data pertaining to present hospital capacity, the aforementioned related parameters, and the accuracy of forecasts of COVID-19 hospitalizations within a given geographical region.

*More information can be found in the [Living Repository](#).*

## COVID-19 Take Control Simulator

COVID-19 Take Control Simulator is an interactive epidemic simulation web application developed by researchers at Te Pūnaha Matatini, led by Professors Alex James, Michael Plank, and Audrey Lustig, and hosted by the Centre for eResearch at the University of Auckland in New Zealand. Funded and supported by Te Pūnaha Matatini and Manaaki Whenua Landcare Research, the application was first released and announced on Te Pūnaha Matatini's website on April 24th, 2020 [105].

Created as an educational tool for public use, the COVID-19 Take Control Simulator was designed to illustrate how the COVID-19 pandemic could develop under different national guidelines throughout the pandemic [106]. These guidelines ranged from the most lenient precautions taken at Alert Level 1 to the most stringent social distancing measures enforced at Alert Level 4 [107]. Developers intended that users would understand that "the power to control COVID-19 is in each of our hands" [106] interacting with the app.

The web application has two interactive pages: the 'R calculator' page, in which users can calculate their reproduction number ($R_{0f}$) based on the degree to which they observe physical distancing, maintain personal hygiene, practice contact tracing, self-isolate, as well as the outside weather; and a 'Simulator' page, in which users can compute simulations of reported cases and COVID-19-infected individuals based on R values in different Alert Levels and the duration that restrictions are maintained at every Alert Level [106]. These pages run on data describing 1,214 probable or confirmed cases of COVID-19 in Aotearoa up to the 7th of April, 2020, publicly available via the Ministry of Health New Zealand [108].

The model for the 'R calculator' is described in a paper by the developers [109], in which $R_0$ is a function of the variables described above. The 'Simulator' page utilizes a stochastic, continuous-time branching process model similar to Davies *et al*. [110] to simulate reported and infected cases [109]. A set of key assumptions are built into this model, including values for the reproduction number ($R_0$), the viral incubation period, prevalence of undetected cases, and hospitalization rate, among others [109]. On the Simulator's "About" tab, developers cite a number of their papers demonstrating COVID-19 modeling, but it is not clear which of these, if any, describe the values for the parameters and distributions for the model on the Simulator [106].

The COVID-19 Take Control Simulator free public use under a GNU General Public License v3.0 (GNU GPLv3) [106]. The creators note that the Simulator is intended for research and educational purposes only and that it should not be used for decision-making. It is feasible that a simulator of this sort could be developed for other localities; this would require adapting the model with accurate data, parameters, and distributions to suit the locality in question. It should also be noted that COVID-19 Take Control Simulator does not appear to have been updated since 2020. Key developments in the COVID-19 pandemic have occurred since then, such as the emergence of the SARS-CoV-2 Delta and Omicron variants, as well as vaccine accessibility. These developments would influence the parameters used by this model, and such a tool would need to be consistently updated to reflect the characteristics of the pandemic as it changes over time.

*More information can be found in the Living Repository.*

## COVIDcast*

COVIDcast is a site that acquires or identifies raw data sources and extracts from them US COVID-19-related signals ("indicators"), which are intended to inform decision making by a broad range of users—public health authorities, the healthcare industry, the public and private sectors, epidemiologists, data journalists, and the general public [111]. COVIDcast was created by the Delphi Research Group at Carnegie Mellon University—one of the two US Centers for Disease Control and Prevention's (CDC) Influenza Forecasting Centers of Excellence [112]—with support from Amazon, the US CDC, Change Healthcare, the US Defense Threat Reduction Agency, Facebook, Uptake, Optum, and Google.org [111]. In addition to data from the aforementioned sources, the group also uses its own indicators to create forecasting models at the state and county levels. COVIDcast was launched in April of 2020 and its first academic paper was published on June 25th, 2021 [113].

The Delphi Research Group's motivation is to develop the theory and practice of epidemic tracking and forecasting. In doing so, they procure data streams that reflect epidemic (or pandemic) activity, define relevant indicators, and make them available for public consumption. They then use these indicators for "nowcasting" (situational awareness) and short-term forecasting [111].

COVIDcast extracts and shares over one hundred COVID-related indicators, categorized into "public behavior," "early indicators," and "late indicators." Public behaviors include the frequency of bar and restaurant visits, extracted from SafeGraph data; people's pandemic-related attitudes and behaviors (e.g. mask-wearing and vaccination), obtained via a COVIDcast-administered survey promoted on Facebook; and frequency of online searches pertaining to COVID-19 symptoms, extracted from Google data. Early indicators include COVID-19-related doctor visits provided by partnering health system organizations, such as Change Healthcare; and COVID-19 symptoms present in individuals or communities, obtained via the US COVID-19 Trends and Impact Survey (US CTIS) [114]. Late indicators include COVID-19 antigen test positivity rates, extracted from raw data provided by Quidel; hospital admissions data, extracted from data provided by partnering health system organizations; and COVID-19 cases and deaths, provided by Johns Hopkins University and USAFacts. Even though COVIDcast did not launch until April of 2020, they were able to collect data retrospectively beginning in February of that year, and they continue to collect and update data on a daily basis.

The Delphi Research Group aggregates and cleans the data, uses them to extract informative signals, displays some of them on the COVIDcast site through an interactive dashboard, and makes all of them available via a web-based API. On the dashboard, users may browse indicators' daily trends or explore correlations between indicators at the U.S. state- or county-level. Using their indicators, the Delphi Research Group implemented quantile auto-regression-based forecasting time series models for deaths at the state and cases at the county level, and submitted forecasts to the publicly-accessible COVID-19 Forecast Hub [115]. The mortality models were built using population data and COVID-19 case and death counts. The case models submitted were built using population data, case county, and two early indicators (doctor's visit rates and self-reported symptom rates) [115]. These models are included in an ensemble model developed in collaboration with the Reich Lab at the University of Massachusetts, and accessible at the COVID-19 Forecast Hub [116]. Similar models demonstrate the usefulness of other COVIDcast indicators for case forecasting [113].

The COVIDcast dashboard is updated with new data nearly daily, and indicators are publicly accessible via an API [117]. The Delphi Research Group also tracks, stores, and makes available (via their API) all revisions made to each data source [118]. Many of which are revised daily. This is essential for building statistical forecasting models, as these must be trained on versions of the data available in real time.

As of July 2021, the API was reported to have been accessed by "thousands of users every day, requesting hundreds of thousands of pieces of information" [119]. COVIDcast's indicators have reportedly been used by numerous organizations responding to the pandemic, including COVID Act Now, COVID Exit Strategy, DeepCOVID, and the Institute for Health Metrics and Evaluation (IHME). Furthermore, the COVID-19 Forecast Hub's multi-group ensemble models, which integrate Delphi Research Group's forecast models, serve as the basis of the US CDC's COVID-19 forecasting communications [111].

In terms of technical scalability, some COVIDcast data sources may be difficult to acquire for some non-U.S. locations. For instance, indicators that rely on proxy measurements for behaviors in the U.S.—mobility data collected via smartphone activity, survey data obtained via surveys on social media sites, and health care data obtained from health system organizations—may be unsuitable in areas where smartphones are not as widely used, social media use or literacy are less pervasive, and health care infrastructure is relatively weak. On the other hand, the statistical methodology for extracting informative indicators for raw data sources is portable, and Delphi makes all of its algorithms and code publicly-available [120]. Similarly, Delphi's forecasting models do not rely on all of COVIDcast's indicators; the state-level autoregressive mortality models, for instance, use only COVID-19 case counts and death counts as predictors. Such data are readily available for many locations outside the U.S. from Johns Hopkins University and WHO (as this reporting is mandated for the 194 WHO member states by the International Health Regulations of 2005) [121]. Thus, an autoregressive time series model of this type may be more easily repurposed and tested for other geographies.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## CT Pneumonia Analysis

"Automated Quantification of CT Patterns Associated with COVID-19 from Chest CT" is a research paper that proposes a deep learning-based tool for conducting an automatic, efficient, and detailed evaluation of the severity of COVID-19 in chest CT scans [122]. The research paper was first uploaded as a preprint to arXiv on April 2nd, 2020 [123], before being published in the journal *Radiology: Artificial Intelligence* on July 29th, 2020 [122], and is presented as a product on the Siemens Healthineers website as CT Pneumonia Analysis [124]. Researchers involved in this paper are affiliated with Hôpital Foch in France, Feinstein Institutes for Medical Research in the United States, Siemens Healthineers in Germany and France, University Hospital Basel in Switzerland, and Vancouver General Hospital in Canada [122]. All of the authors either have been employed or were partially supported by Siemens Healthineers at the time of publication [122].

The tool uses deep learning methods to assess regions on CT scans with ground-glass opacities (tissues presenting increased attenuation) and consolidation (increased densities) and outputs scores quantifying the severity of infection [122]. Models were trained on a total of 9,749 chest CTs collected from multiple institutions in the U.S, Europe, and Canada between 2002 and April 2020, which were de-identified and approved by their respective ethics committees with waived consent [122]. 200 chest CT volumes were used for testing—100 of which were control volumes and 100 which were COVID-19-positive [122]. The data are not publicly accessible, but information about demographic distribution is presented in the associated publication [122].

The proposed tool employs two models—a lobe segmentation model and an abnormality segmentation model—followed by several mathematical computations to quantify the severity of abnormalities [122]. The lung and lobe segmentation model is a reinforcement learning model, 'Deep Image-to-Image Network' (DI2IN) [125], that takes as an input chest CT scans and generates masks that segment the lungs and lobes in 3D. The DI2IN model was trained on 8,087 CT scans from patients with various diseases, and then fine-tuned on 1,136 CT scans from three groups of patients deemed useful for the abnormality model—those presenting interstitial lung disease, pneumonia, and COVID-19—to improve lung segmentation robustness over the infected areas [122]. The abnormality segmentation model consists of a Dense U-Net, based on Ronneberger *et al*. [126], which transfers 3D chest CT volumes to a segmentation mask of the same size. This model was trained on 901 chest CTs and tested on the aforementioned 200 chest CTs [122]. The quantification of severity is performed by calculating the percentage of opacity (PO), percentage of high opacity (PHO), lung severity score (LSS), and lung high opacity score (LHOS) from the abnormality segmentation masks. PO and PHO represent the percentage of the volume of lung affected by the disease and by severe disease, respectively, and LSS and LHOS represent the percentage of affected lobe and percentage of high opacity in the lobe, respectively. LSS is computed to measure the extent of lung involvement across each lobe and LHOS is computed to measure the extent of abnormalities with severe disease. These scores can be used by medical practitioners to assess the severity and progression of the infection and for early detection.

Numerous correlation and regression analyses of the predictions with the "ground truth" (manual annotations of lesions, lungs, and lobes) demonstrated the abnormalities of ground-glass opacity and consolidations segmented and quantified had a high correlation to the ground truth annotations [122]. It was furthermore observed that there were very few false positives segmented in subjects with no pathologic findings [122]. A $X^2$Contingency test demonstrated that there was no significant difference between ground truth and

predicted metrics (p-value of <.001) [122]. Furthermore, Pearson correlation coefficients, measuring the linear correlation between predictions and ground truth, were 0.92 for PO, 0.97 for PHO, 0.91 for LSS, and 0.90 for LHOS [122].

According to the authors of the paper describing this tool, their work was the first system to evaluate chest CT patterns associated with COVID-19 infections. The tool was awaiting a patent for commercial use at the time of publication [122], but is now offered as a free add-on to paid subscribers of Siemens Healthineers' "*syngo*.via Frontier" post-processing research platform [124], [127]. Such a tool could be useful for assessing the severity of abnormalities in diseases that present COVID-19-like ground-glass opacity and consolidation. The authors noted, however, that the presence of other abnormalities in or around the lung, like pleural effusion, may pose a challenge to the algorithm [122]. They also noted the need for further calibration of the system: while their study evaluated quantification in COVID-19 and healthy cases, it did not evaluate the system's capability to differentiate between COVID-19 and other viral pneumonia or interstitial lung diseases [122]. Such a test would be necessary for the tool to be deployed for diagnostic purposes.

*More information can be found in the Living Repository.*

## Evolutionary Surrogate-Assisted Prescription (ESP)*

Evolutionary Surrogate-Assisted Prescription (ESP) is a machine learning technology designed to automatically determine the most effective non-pharmaceutical intervention (NPI) strategies to contain the spread of COVID-19. ESP was developed by researchers at Cognizant, a for-profit company in the United States, and the University of Texas at Austin. The proof-of-concept for ESP was described in an article uploaded to arXiv in May 2020 [128] and published in *IEEE Transactions on Evolutionary Computation* in April 2021 [129]. An interactive dashboard demonstrating the technique is available online [130], hosted by Cognizant's Evolutionary AI team.

Pandemic modeling efforts frequently employ traditional epidemiological approaches, such as compartment models, to predict the spread of the disease using a few approximated parameters. These methods are imperfect, however, due to the uncertainty underpinning their parameters and the unpredictability of some factors, such as NPIs, across cultural and economic environments. In contrast, ESP uses near real-time data collected on pandemic interventions, their economic impact, and disease transmission to train a long short-term memory (LSTM) neural network capable of predicting how NPIs affect the course of the pandemic without needing to estimate parameters [129].

ESP utilizes the COVID-19 Government Response Tracker dataset [24], maintained by Oxford University Blavatnik School of Government, which contains 23 different *indicators,* including containment measures, economic factors, health system descriptors, and vaccine policies, among others. It also includes data representing different NPIs characterized by measures of stringency, in which historical data has been encoded for over 180 countries. These data have been collected since March 2020, and have been updated regularly, though some economic and health system policies have not been updated since August 2021. ESP developers acknowledge that the dataset is noisy, given discrepancies in how different countries detected and reported case data, how their varying testing policies affected case detection, and how they implemented NPIs in different ways [129]. They also note that the dataset contains some accidental repetitions, missing days, and other mistakes. Despite these challenges, the developers highlight that there are still enough data to effectively train their model to make useful predictions.

In this data-driven modeling approach, ESP applies a LSTM neural network model trained on the countries' case numbers and NPI data (such as border closing and containment) to predict how the pandemic will unfold under various NPIs. A Predictor model ($P_d$) takes a decision as its input and predicts outcomes for that decision [129]. A second model, a Prescriptor ($P_s$), is then created, which takes case information for the previous 21 days as its input and outputs actions that can optimize outcomes within that context. The Prescriptor model is built using neuroevolution [131]–[133], evolved using the Predictor model. The Predictor is thereby the surrogate for the development of the Prescriptor, i.e. the decision policy output, and that the Prescriptor is not restricted by a finite training dataset or its capacity to evaluate real-world characteristics. The Predictor instead serves as a fitness function that can be queried regularly and efficiently [129]. If actions prescribed by the Prescriptor are enforced in the real world, data about their real-world outcomes can be fed back into the model. ESP is therefore a continuous black-box optimization process for adaptive decision-making whose algorithm operates as an outer loop involved in constructing its own Predictor and Prescriptor models. On the online demonstration of this tool, users are allowed to view case predictions across countries and regions based on real-time data (or view historical data "counterfactuals") and Prescriptor settings (NPIs lifted versus maxed out) [130].

To evaluate the performance of their LSTM model, it was compared to a suite of baseline machine learning regression models, including linear regression, random forest regression, support vector regression, and feed-forward neural network regression [129]. Each method was trained independently 10 times on a training dataset containing 14-days of 20 countries' case data leading up to May 6, 2020, and test data consisted of data from May 7 through May 20, 2020. The LTSM model outperformed all of the baseline models across all four metrics: normalized case mean absolute error (MAE), raw case MAE, mean rank (ranks the methods in terms of case error for each country, and then averages over countries), and 1-step $\hat{R}_n$ MAE (the loss the models were explicitly trained to minimize).

The tremendous amount of data that have become available has, according to authors, enabled the ESP approach to be feasible for the first time [129]. As a phenomenological data-driven model, ESP does not explain how given outcomes are produced, but because it can be fed significant amounts of relevant data, it can provide accurate predictions of them. It is therefore a tool that has primarily been designed with decision-makers in mind, to provide them with a greater capacity to evaluate candidate non-pharmaceutical interventions using predictive models. The authors state that this AI system has not been designed to replace human decision-makers, only to augment them. Additionally, because the dataset is not specific to one locality, it can be used by decision-makers across different geographical contexts to inform policies specific to their geographical contexts and goals. Indeed, the developers highlight how ESP can make creative suggestions, such as alternating between NPIs, that achieve "maximal effect and minimal cost." Authors stress that the model does rely heavily on accessible, accurate, up-to-date data being available, especially as data comes to reflect changes in the course of the pandemic (i.e. as new variants emerge, and as different nations enforce different NPIs).

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## GeoSpark Analytics Hyperion COVID-19 Live Dashboard

*Disclaimer: Details pertaining to the technical specifications of this tool were not publicly available; the information below was obtained from blog posts on the GeoSpark Analytics site. Furthermore, whereas the tool was accessible to the general public in November 2020, at some point before October 2021, access was restricted to those with an ArcGIS account associated with the Geospark Analytics team.*

The Hyperion COVID-19 Live Dashboard is a dashboard that uses machine learning to identify, track, and analyze events associated with COVID-19 via mentions on online news articles and social media posts [134]. This tool was developed by GeoSpark Analytics, a private computer software company, in partnership with Esri, a private geographic information systems software supplier (most famous for their product, ArcGIS). The dashboard was publicly announced on a blog post in April 2020 but was built upon the GeoSpark Analytics Hyperion platform, which was developed before the COVID-19 pandemic [135].

Specific information pertaining to the data or models used by the Hyperion COVID-19 Live Dashboard is not publicly available. A blog post describing the Hyperion platform, upon which the dashboard was built, describes three functionalities: (1) categorizing disparate forms of information into classes of activities using a machine learning model that learns patterns in unstructured data to automatically recognize and categorize data from social media, news media and other sources into themes such as social unrest, conflict and terrorism; (2) modeling patterns of human activity by evaluating news, social media, and other information in the location of the anomaly; and (3) continuously assessing levels of "stability," by comparing current activity against long-term trends (which are used to define "normalcy") within geographic regions [136]. In a separate blog post, they illustrate how their dashboard has been built with these functionalities [135]. However, they do not describe exactly what data are utilized to perform these tasks; details pertaining to the machine learning models employed are also not available.

Geospark Analytics claims that their technology detected anomalous activity levels in Wuhan, China, and categorized them as a disease outbreak on December 31st, 2019, eight days before the WHO announced concern over the pneumonia outbreak [136]. In April 2020, they claimed that their dashboard had been viewed more than 15,000 times in the previous month and that it had been integrated into other applications [135], but did not describe in detail what these applications were.

As recently as November 2020, the dashboard was accessible to the general public on their website. As of October 2021, however, it appears that users need to have an ArcGIS account associated with the GeoSpark Analytics organization to access the dashboard.

*More information can be found in the Living Repository.*

## IBM COVID-19 Deep Search

The IBM COVID-19 Deep Search platform organizes both structured and unstructured COVID-19 data into a knowledge graph that can be navigated and queried to retrieve information. The Deep Search platform was developed by IBM Research Europe, based in Zürich, Switzerland, in the first half of 2020, by members of the Scalable Knowledge Ingestion group [137].

IBM's Deep Search access page states that the purpose of the platform is to allow scientists and academics to "unlock the knowledge" of published unstructured and structured data pertaining to COVID-19 [138]. Users can do so by either navigating the knowledge graph manually or building query workflows to extract specific answers from the data.

Deep Search incorporates data from various unstructured and structured data sources. Unstructured data is obtained via the COVID-19 Open Research Dataset (CORD-19) [139], a large resource of scientific papers on COVID-19 and related historical coronavirus research sourced from PubMed Central (PMC) [140], the World Health Organization (WHO) COVID-19 Database [141], as well preprints from bioRxiv, medRxiv, and arXiv [16, p. 19]. Structured data included pharmaceutical and genetic databases from DrugBank [142] and Genbank [143], as well as clinical trials from Clinicaltrials.gov [144] and the World Health Organization International Clinical Trials Registry Platform [145]. In total, the platform is claimed to have ingested 158,524 COVID-19-related papers from the aforementioned sources (as of October 4th, 2021) [146], and the resulting knowledge graph contains approximately 4 million nodes and 50 million edges [138].

Deep Search is an integration of two IBM technologies: Corpus Conversion Service (CCS) and Corpus Processing Service (CPS). The development of both tools preceded the COVID-19 pandemic (IBM notes their "extensive use" in the materials science, automotive, and energy industries) but was combined and made accessible to support pandemic response. CCS is a cloud-based platform that allows users to convert PDFs or bitmap documents into a structured representation of the original data. CCS parses documents (using optical character recognition to parse images), applies ML models on parsed documents to assign semantic labels to content, and reassembles documents into a machine-readable data format, such as JSON [147]. CPS then integrates this data into a knowledge graph, allowing users to navigate structured data manually on the knowledge graph interface, execute queries for specific information, and delve more deeply into specific topics by accessing source documents [138].

Access to Deep Search is granted to scientists and academics. Those interested in using the tool may apply for access on the IBM site [148], and according to IBM, there are 647 registered users, as of October 18, 2021.

According to IBM, CCS is capable of ingesting 100,000 PDF pages per day on a single server with an accuracy above 97%. This capability to structure, parse, and navigate large amounts of scientific data suggests a high potential for Deep Search to scale into areas of research with large amounts of published (or preprinted) research and accessible data.

*More information can be found in the Living Repository.*

## icolung*

icolung is a cloud-based software that uses AI to analyze non-contrast thorax CT scans for COVID-19 pathologies [149]. It generates reports highlighting lung lesions caused by COVID-19 and quantifies volumetric lung involvement and lesion severity [150], allowing intensive care physicians to make more informed diagnoses and treatment decisions. icolung was developed by icometrix, a company specializing in medical imaging solutions, with support from the radiology labs at Universitair Ziekenhuis Brussel, Katholieke Universiteit Leuven, Vrije Universiteit Brussel, and Interuniversity Microelectronics Centre (imec) [151]. The tool was initially created as a part of the Belgian *pro bono* icovid initiative until it received funding by European Union's Horizon 2020 research and innovation program, the Flemish Government's "Onderzoeksprogramma Artificiële Intelligentie Vlaanderen" (AI Research program), and Research Foundation Flanders (FWO) later at some point in 2020 [151]. icolung software has been in development since March 2020, was CE certified in April 2020 [151], and FDA certified in May 2020 [152]. As of September 2021, efforts towards its development are continued by the interdisciplinary icovid consortium, which received funding from the EU's Horizon 2020 programme [153]. Beyond the continuous technical improvement and addition of new features to the software, the icovid project's explicit goal is to demonstrate the clinical value of icolung, by measuring the impact on clinical decision making, and to deploy it on a larger scale.

icolung comprises deep learning modules for voxel-level lesion segmentation and also for probabilistic diagnosis of COVID-19. The latter module was trained, validated, and tested using a collection of 1,419 scans (795 COVID-19-positive, 624 COVID-19-negative) from a "pool" of centers—10 European and Latin American clinics before September 2021 (exact dates not disclosed) [154]—complemented by 150 NSCLC-positive (non-small cell lung cancer) CT scans acquired from the Lung Image Database Consortium (LIDC) and the Image Data Resource Initiative (IDRI) [155] before November 2019. Out of these, 797 scans were used for training, 49 for validation, and 98 for testing [154]. Data from an independent center containing 219 COVID-19-positive and 256 COVID-19-negative samples were kept separate as an independent test set. The lung lesions segmentation module was trained and validated on 61 manually annotated scans from the pooled centers, and evaluated on 45 held-out test scans. The datasets from clinics have not been made openly accessible [154].

The deep learning architecture underlying icolung consists of two main stages: lesion detection followed by probabilistic COVID-19 classification. First, images are preprocessed using a pre-trained image segmentation module, based on the 3D U-net architecture architecture [156], which automatically labels areas of CT scans as regions of interest, followed by a multi-class lesion segmentation approach, which uses a DeepMedic-like [157] to identify patterns suggestive of COVID-19. Inside the diagnostic module, a conditional variational autoencoder (CVAE) module is employed to learn COVID-19 related patterns from the segmentations, classify each segmentation, and to assign a likelihood that the scan stems from a COVID-19 patient [154]. When integrated into a hospital picture archiving and communication system (PACS), an analysis can be performed on CT scans within 10 minutes [152]. Lesions on the CT scan are color-coded and their volume is mapped to a severity score, enabling the practitioner to analyze the lesions visually and make an informed diagnosis.

When evaluated on a test set that incorporated data from the aforementioned pooled centers, the diagnostic model achieved a sensitivity (that is, the proportion of COVID-19-positive individuals it accurately diagnosed) of 94% and a specificity (the

proportion of COVID-19-negative individuals that were correctly diagnosed) of 82% [154]. When tested using data from an independent medical center, the model achieved a slightly higher sensitivity, 96%, but significantly lower specificity, 59% [154].

icolung is currently CE-certified, FDA-permitted, GDPR-compliant, and offered *pro bono* worldwide [151]. It has been adopted by over 150 clinics, and the icovid collaborative is striving to distribute the software to 800 medical clinics worldwide [151]. A favorable aspect of the system is that it allows for interpretability via visualizations, as explained above, and that it offers condensed representations of the detected abnormalities, in the form of severity scores, allowing for rapid interpretation. The code for training the deep learning models is hyperlinked in the paper [154]; however, as of December 13, 2021, this link is inaccessible. The tool's robustness across patient demographics and COVID-19 variants have yet to be empirically validated; to expand this tool's uptake across demographic groups and the COVID-19 pandemic life cycle, it will be necessary to demonstrate robustness in these metrics. In fact, the icovid consortium is presently running a reader study, which will, through the format of a controlled trial with several radiological expert users, quantify unequivocally the benefit of icolung towards clinical image interpretation and decision making.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

## Johns Hopkins US Risk Model

The Johns Hopkins US Risk Model is a county-level COVID-19 risk modeling framework intended to assist the US government and individuals in making informed decisions. The project was announced in September 2020 by researchers at Johns Hopkins Center for Systems Science and Engineering, with funding from the US National Science Foundation, National Institute of Allergy and Infectious Diseases, and NASA [158].

In a blog post announcing this initiative, researchers shared that the goal of their modeling is to identify at-risk populations and to learn the locations and attributes of those that are most exposed to the risk of infection and death from COVID-19 [158]. To this end, researchers claim to have constructed their risk-modeling framework using a "flexible approach" that would allow them to model different risk indicators for different use cases [158]; however, a more in-depth explanation of how this was achieved is not provided.

In building the risk modeling framework, the initiative relied on US epidemiological, mobility, and demographic data from numerous sources [158]. Epidemiological data are drawn from the Johns Hopkins COVID-19 Data Repository, which aggregates authoritative, publicly-available COVID-19 case, death, and recovery rates from across the globe at various levels of granularity—from country-wide to city-wide, depending on availability of data [25]. Mobility data were sourced from mobile phone usage data and provided by SafeGraph. It appears, however, that while SafeGraph provided social-distancing metrics for free at the peak of the pandemic, such data have now been wrapped into their Weekly Patterns product, for purchase [159]. Both population and health indicators were gathered from the US census (population totals, demographic percentages, and age breakdowns), County Health Rankings (smoking percentages, poverty, and chronic disease), and the Definitive Healthcare Dataset published by ESRI (Statistics on hospital beds and availability) [158].

Models developed for forecasting COVID-19 risks at the local, state, and national levels use different statistical methodologies, such as multiple linear regression, logistic regression, random forest regression/classification, and curve fitting [158]. Researchers explored techniques that could further improve predictive capabilities, such as ensemble approaches, input clustering, and deep learning [158]. They claimed to have modeled several different aspects of the outbreak, including cases and deaths over different time horizons, case and death curves' deviations from current trends, case and death rates per person, risk categories based on time-dependent rates of change, and categorical epidemiological classifications [158]. Details describing which sets of data were used for any particular model were not disclosed.

This initiative's site displays a map that visually compares projected quantiles of new cases in each county during the first two weeks of August 2020 (output from the model) to observed cases reported, with striking similarities [158]. However, it is unclear whether this initiative is still under development, whether it is or was used and by whom, and what the process is for obtaining access to the model or its predictions. It is difficult to assess the technical scalability of this tool for numerous reasons, including the ambiguity concerning the licensing of this tool, the minimum amount of data required for any particular model, the accessibility of its data (namely, mobility data which is no longer available for free), and the quality of the required datasets at different geographic scales from across the globe.

*More information can be found in the Living Repository.*

## LitCovid

LitCovid is an open-source, curated database equipped with advanced search and filtering features designed to facilitate desktop research on topics related to COVID-19. This tool was developed by researchers from the Text Mining/Natural Language Processing (NLP) Research program of the U.S. National Library of Medicine/National Institutes of Health's (NLM/NIH) National Center for Biotechnology Information, and funded by the NLM/NIH Intramural Research Program. After it was announced in *Nature* in March 2020 [160], the first full-length paper covering its features and methods was published in *Nucleic Acids Research* in November 2020 [161].

LitCovid's curation workflow consists of four steps repeated on a daily basis: new papers are retrieved, curated, annotated, and indexed. In the first step, also referred to as *Document Triage*, candidate articles are retrieved from PubMed [162] using a broad set of query keywords set of keywords (such as "coronavirus," "ncov," or "SARS-CoV-2") to retrieve all possible relevant articles. Human curators then classify articles with the assistance of machine learning: an ensemble of machine learning models, including support vector machines (SVMs), using bag-of-words features, and convolutional deep neural networks (CNNs), using word embeddings, provide human curators with a score indicating the likelihood of an article's relevance to COVID-19. With these scores, human curators (with training in biomedical data sciences) make final determinations; those deemed relevant populate the LitCovid database [163]. This curation process has been streamlined in a newly-developed online system, LitSuggest [164].

Following curation, articles are annotated with the assistance of a deep learning model that integrates the embeddings produced by BioBERT [165], a language model pre-trained biological with biomedical corpora, with manually crafted features (such as publication types). This model outputs probability scores representing the likelihood that that article pertains to each of eight category labels: general information, mechanism, transmission, diagnosis, treatment, prevention, case report, or epidemic forecasting. Human curators review the probability scores and article content before approving or denying labels. Geotagging is performed using a tool from spaCy [166], an open-source software library for natural language processing. The tool recognizes named entities, which in this case are geographical statements. The origin tag is then obtained by assigning the geographical statement to its respective country using dictionaries. In addition to geographical tags, an open-access, web-based text mining service PubTator [167] provides annotation of drugs and chemical mentions using deep learning techniques. To make the database searchable, the documents are indexed with the help of a web-based, open-source framework by Solr [168], which assigns an index and attributes to each document and processes search queries.

Each of the three methods in the curation workflow has been tested (relevance classification, topic assignment, and geotagging) by comparing the performance of the ML algorithms to human-annotated "ground truth" data. For each task, performance was reported with precision, recall, and micro-F1-scores. The document classification model was trained on ~64,000 documents and tested on 16,000 documents, achieving a micro-F1 score of 0.99 [161]. The topic assignment model was trained on ~32,000 documents taken from the LitCovid database—and therefore considered COVID-19-relevant—as of August 12$^{th}$, 2020. A micro-F1-score of 0.81 was achieved on an independent test set of ~8,000 documents. Geotagging did not require additional training, and on a test set of 361 articles, a micro-F1 score of 0.94 was achieved. While all articles used are freely accessible, the training and test sets used are not publicly available. The authors state that the low

performance of the topic assignment task might be caused by papers that lack abstracts and full texts, where only a title has been provided. This holds for ~40% of the papers within the LitCovid database [161].

Furthermore, the authors compared the set of papers found on LitCovid to those found directly with PubMed's native search function (searched in March 2020). The LitCovid search found 593 papers—136 more than PubMed's native search function—and 446 papers were found by both search methods [161]. The papers that were only found by one search engine were manually checked for COVID-19-relevance. All 147 papers that were only found by LitCovid were proven to be COVID-19-relevant, whereas all 11 papers that were only found by the PubMed search function were not COVID-relevant. This suggests that LitCovid can find literature with both higher precision and better coverage [161].

The initiative demonstrates how machine learning can increase scalability and efficiency in literature curation processes. A curated database is of high value for data-driven discovery and evidence-based medicine, in particular at the time of a pandemic, when there is an acute need to collect and disseminate accurate and relevant scientific findings. By July 2020, the database had been accessed more than ten million times [161]. As of January 7th, 2022, LitCovid is open-source, updated daily, and all papers within the database are freely available [163]. It should be noted, however, that the tool is only available in English, and the database only contains publications in the English language.

*More information can be found in the Living Repository.*

## n*f*erX

n*f*erX is a cloud-based platform that allows users to query relationships between biomedical concepts, based on associations derived between terms in unstructured biomedical literature and experimental—primarily Single Cell RNA-sequencing (scRNAseq)—data [169]. This tool was developed by n*f*erence and Janssen, by researchers funded by these two companies. It was first announced in March 2020 following n*f*erence joining the COVID-19 Healthcare Coalition in the USA, a private sector collaboration with Mayo Clinic, MIT Faculty, coordinated by the not-for-profit organization MITRE [170].

Data underlying the n*f*erx platform came from two types of sources. First, a corpus of over 100 million biomedical documents—including PubMed publications, grants, preprints, patents, and clinical trials—was developed for deriving associations from phrases in unstructured literature. Second, a corpus of experimental data was compiled, including publicly-available bulk RNAseq human and mouse single-cell RNA-seq data (source undisclosed), bulk RNA-seq data obtained from Gene Expression Omnibus (GEO) [171] and the Genotype Tissue Expression (GTEx) portal [172], immunohistochemistry (IHC) data from the Human Protein Atlas [173], as well as tissue proteomics datasets from the Human Proteome Map [174]. Much of these data come from open-access databases; however, the compiled corpuses ingested by n*f*erx are not publicly available.

Using all of these data, local and global association scores for terms are obtained with natural language processing (NLP) methods [175]. The local score is calculated in an unsupervised learning fashion (exact method undisclosed), which captures the strength of association between two concepts based on the frequency of their co-occurrence normalized by the frequency of each concept throughout the corpus [169], [175]. To compute global scores, a word2vec model is employed, in which all words and phrases are projected in a high-dimensional vector space of word embeddings, to construct 'neighborhoods' of concepts [169]. In this embedding, the concepts' proximity (or *cosine similarity*) captures their association in the literature.

In their recent publication in eLife, the developers of n*f*erx present how their tool was used to perform a hypothesis-free expression profiling of the ACE2 receptor, an enzyme to which the viral spike protein of SARS-CoV-2 (and a number of coronaviruses) binds [169]. This profiling indicated that ACE2 may be heavily expressed in tongue keratinocytes, olfactory epithelial cells, airway club cells, and respiratory ciliated cells [169]. This work also identified the gut as a potential hotspot of COVID-19, as small intestine enterocytes share a maturation-correlated transcriptional signature [169].

Overall, n*f*erX provides an innovative approach for deriving associations between biomedical concepts. This includes SARS-CoV-2 targets, but also extends beyond COVID-19, coronaviruses in general, and their target tissues. Systematic knowledge management tools of this sort are of increasing importance as biomedical literature grows intractably. Authors note that novel transformer models, such as BERT [176], could serve a complementary function to word2vec models, by enabling greater contextual sensitivity [169]. They also note that as protein expression can vary widely based on demographic factors and patient pathologies, more comprehensive scRNAseq from across different demographics and pathologies pertinent to COVID-19 will be required to reveal "under-appreciated fingerprints of coronavirus transmission patterns, tissue tropism, and mortality" [169].

*More information can be found in the Living Repository.*

## RADLogics Deep Learning CT Image Analysis

RADLogics Deep Learning CT Image Analysis is an AI-assisted tool designed to quickly and accurately detect the presence of COVID-19 in thoracic CT scans [177]. The tool was developed by RADLogics, a healthcare software company based in New York, USA and Tel Aviv, Israel [178] with support from Tel-Aviv University, Affiliated Taizhou Hospital of Wenzhou Medical University, Mount Sinai Hospital, and The University of Maryland School of Medicine [177]. The first academic article associated with this tool was uploaded to arXiv on March 10, 2020 [177].

RADLogics Deep Learning CT Image Analysis was developed early in the COVID-19 pandemic to respond to the growing need to quickly evaluate large numbers of thoracic CT scans for COVID-19 detection, measurements, and the tracking of disease progression.

The model was trained on 50 thoracic CT scans of patients in China, collected between January and February 2020, which were diagnosed by a radiologist as suspicious for COVID-19 [177]. The cases were extracted by querying a cloud picture archiving and communication (PACS) system for cases that were referred for laboratory testing. Each 2D slice was annotated as normal (n=1036) versus abnormal (n=829).

The Deep Learning CT Image Analysis tool consists of two subsystems that analyze thoracic images at a 3- and 2-dimensional level [177]. Subsystem A is a 3D analyzer for nodules and focal opacities, implemented with off-the-shelf software. Subsystem B detects coronavirus abnormalities using a 2D deep-learning model built on a deep convolutional neural network architecture with ResNet-50 (pre-trained using the ImageNet dataset [179]). Each subsystem makes predictions independently and the overall classification (the “corona score”) is computed based on the ratio of slices determined to be COVID-19-positive out of the total slices of lung images from the outputs from each subsystem.

The Deep Learning CT Image Analysis tool’s classification accuracy was tested on 107 thoracic CT scans—56 COVID-19-positive patients confirmed by RT-PCR, and 51 patients without any abnormal findings in a radiologist’s report—and achieved an AUC of 0.996 (95%CI: 0.989-1.00) [177]. Since this article was uploaded to arXiv in March of 2020, it is not entirely clear whether or how often the tool has been updated, or to how many hospitals the tool has been deployed.

More training and validation would seem beneficial for assessing the transferability of this tool. The training dataset could use a wider variety of clinical data, a larger-scale validation could be conducted and peer-reviewed, and the tool’s capability to distinguish between COVID-19 pneumonia and non-COVID-19 pneumonia could be evaluated (as it is not covered by the available article). Furthermore, instructions for accessing the Deep Learning CT Image Analysis tool are not public-facing; it appears as though interested users need to contact RADLogics directly for access. Access to other RADLogics medical imaging tools requires purchasing credentials to install or access (via the cloud) RADLogics’s patented workflow software; we assume this to be the case for this tool as well.

*More information can be found in the Living Repository.*

## Trove

Trove is an open-source framework that allows for a low-cost and privacy-preserving method of training of Named Entity Recognition (NER) models for patient records, developed by the researchers at the Center of Biomedical Research at Stanford University California. The initiative was first published in *Nature Communications* in April 2021 [180] and was funded by the National Library of Medicine (NLM).

NER models enable automated extraction of relevant information and terms out of a text document. Trove's approach is to couple pre-trained language models such as BioBERT [165] with ontology-based weak supervision, rather than learning with hand-labelled data, to perform NER tasks. Researchers demonstrated this approach to be time-saving, less expensive (by not requiring domain experts to annotate or update corpora), and achieves task performance close to—albeit slightly lower—than fully-supervised learning frameworks using hand-labelled data [180]. Another prominent benefit of this approach, pertinent to pandemic response, is that by using medical ontologies rather than patient notes for training, issues related to the storage and use of individuals' health records are avoided, and thus labelling rules can be shared or made publicly available.

Labels were are sourced from open-source medical ontologies [4], primarily the 2018AA release of the Unified Medical Language System (UMLS) Metathesaurus [181]—cleaned to remove non-English and zoonotic source terminologies—as well as 2019 SPECIALIST abbreviations [182], Disease Ontology [183], Chemical Entities of Biological Interest (ChEBI) [184], Comparative Toxicogenomics Database (CTD) [185], the seed vocabulary from AutoNER [186], ADAM [187], and word-sense abbreviation dictionaries used by the clinical abbreviation system CARD [188].

Models were trained, validated, and tested using six datasets, comprising 4,705 documents in total, that are either publicly available as benchmark datasets or were made available to the authors upon data use agreements with the data owners [180]. Training data were automatically (weakly) labelled and contained 1,725 documents. The test and validation datasets were labelled by hand and contained 1,488 and 1,492 documents, respectively. In addition, the authors performed a COVID-19 case study, approved by the Stanford University Administrative Panel on Human Subjects Research, using a data set of 796 emergency department notes from patients undergoing COVID-19 testing at Stanford Health Care beginning in March 2020. These clinical data are not publicly available as they contain information that could compromise patient privacy but can be obtained upon request from the paper's corresponding author [180].

Models' performance in two categories of medical tasks were analyzed: (1) NER, and (2) span classification, "where entities are identified a priori and classified for cue-driven attributes, such as negation or document relative time, i.e., the order of an event entity relative to the parent document's timestamp" [180]. In NER tasks, weakly supervised models trained solely using labels sourced from ontologies scored 3.9-14 F1 points lower than fully-supervised models, but adding task-specific rules brought the weakly-supervised within 1.3–4.9 F1 points (4.1%) of models trained on hand-labeled data. In span tasks, the weakly supervised model task-specific rules scored 3.4–13.3 F1 points lower than fully-supervised models. The performance of Trove was also measured in comparison with existing weakly supervised NER methods on a disease- and chemical-recognition benchmark task [180]. Starting from a pre-trained language model that is fine-tuned on biological text (*BioBERT*), Trove performed best amongst the models that were trained using weak supervision solely on ontologies and dictionaries (by 1.7 F1 points). When

introducing task-specific labelling rules, Trove outperformed every other weakly-supervised model (by at least 1.9 F1 points) and achieved performance close to systems based on fully-supervised training with hand-labelled data. In the COVID-19 case study, a model trained with weak supervision outperformed a supervised baseline model with hand-labelled training data by 2.3 F1 points.

Python code for Trove is available on Github [189]. The authors note that their task-specific labeling functions were not exhaustive—they only reflect low-cost rules easily generated by domain experts—and that increasing the number of task-specific labelling rules might further improve the model's performance. They also suggest that incorporating data augmentation or multitask learning could mitigate the need to engineer task-specific rules. It should be noted that this framework is English-language specific and its performance in non-English settings have yet to be evaluated.

*More information can be found in the Living Repository.*

## Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations*

*Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations* is a preprinted research article [190] and an associated online, interactive, agent-based simulation [191] that models the spread of COVID-19 based on the prevalence of mask-wearing in a population. The researchers involved are affiliated with the Hong Kong University of Science and Technology, the International Computer Science Institute, Ecole de Guerre Economique, the University of Cambridge, Manifold research, University College London, ELU AI Ltd, the Royal Free Hospital, London, and the Population Research Institute at The Family Federation of Finland. Both the research article and online masking simulator became accessible in April 2020.

The objective of this study was to evaluate the effectiveness of mask-wearing in preventing the spread of COVID-19 with new theoretical models and empirical data-analysis techniques. Researchers aimed to build a base of evidence to support the urgent implementation of universal masking in regions that had not yet adopted it as policy or as a broad cultural norm.

The research article presents two models for predicting the impact of universal face mask-wearing upon the spread of the SARS-CoV-2 virus during the pandemic: a stochastic, dynamic, network-based, compartmental, susceptible-exposed-infectious-recovered (SEIR) approach; and an individual agent-based modeling (ABM) Monte Carlo simulation [190]. For the former approach, researchers used a SEIR model implemented on a stochastic dynamic network, rather than a deterministic SEIR model, as it more closely represented interactions between individuals in a large population. Parameters were tuned to model different degrees of social distancing, lockdown stringency, and mask-wearing, and the empirical characteristics of COVID-19 spread as documented in the “SEIRS+” COVID-19 notebooks [192]. For the latter approach, researchers created a square wraparound two-dimensional environment, within which a population of individuals could exist in one of four SEIR states. The wraparound feature allowed the environment to represent an arbitrarily large space, giving more accurate dynamics without boundary effects from small spaces. Parameters were tuned to best approximate known COVID-19 dynamics, and the impact of masking was modeled by allowing for variation in mask-wearing and *mask characteristics,* with mask transmission rate (T) and mask absorption rate (A) denoting the proportion of viruses that are stopped by masks during exhaling (transmission) versus inhaling (absorption), respectively [190].

The SEIR and ABM predictive models demonstrated that: (1) near-elimination of COVID-19 transmission when least 80% of a population is wearing masks, versus minimal effect on transmission when only 50% or less of the population is wearing masks, and (2) a significant impact when universal masking is adopted early (by day 50 of a regional outbreak), versus minimal impact when universal masking is adopted late (after day 50) [190].

To validate their models, the researchers compared their results with what little (at the time) historical macro-scale empirical data were available. They collected a data set describing the “degree of success” in managing COVID-19 by countries or regions and by the prevalence or enforcement of universal masking. The dataset contained the number of detected COVID-19 cases from Jan 23 to April 10, 2020, and the characteristics of universal masking culture and/or universal masking mandates or government

recommendations within 38 countries/provinces in Asia, Europe, and North America with similarly high levels of economic development. This empirical data validated the predictive models' findings for the need for universal *and* early masking. It is also noteworthy that a study published in November 2021 by Talic *et al.* [193] furthermore validated these predictions, identifying a 53% reduction in incidence due to mask-wearing.

Since this research involved predictive, simulated models, it would be relatively easy to reproduce: models could be tuned to more accurately simulate COVID-19 spread, considering the far greater amount of empirical data on COVID-19 variant characteristics and on geographical masking culture/mandates/recommendations than when the study was initially conducted in April 2020. There also exists a much greater amount of COVID-19 transmission data against which these models may be validated. It should be noted, however, the degree of uncertainty with regards to the influence that a larger base of evidence would have in changing norms or policies concerning mask-wearing.

**This summary has been reviewed and approved by the initiative's developers. More information can be found in the Living Repository.*

# References


[1] "Effect of vaccination, border testing, and quarantine requirements on the risk of COVID-19 in New Zealand." https://www.tepunahamatatini.ac.nz/2021/11/10/vaccination-border-testing-and-quarantine-requirements/.

[2] "How Te Pūnaha Matatini became part of our Covid-19 response - The University of Auckland." https://www.auckland.ac.nz/en/news/2020/10/28/covid19-modelling-efforts-by-TPM-help-NZ.html.

[3] "Public Health Surveillance." https://surv.esr.cri.nz/episurv/index.php.

[4] "COVID-19: Vaccine data," *Ministry of Health NZ*. https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-data-and-statistics/covid-19-vaccine-data.

[5] "FluTracking incidence calculation methods." https://www.tepunahamatatini.ac.nz/2021/09/20/flutracking-incidence-calculation-methods/.

[6] "A COVID-19 vaccination model for Aotearoa New Zealand." https://www.tepunahamatatini.ac.nz/2021/06/30/a-covid-19-vaccination-model-for-aotearoa-new-zealand/.

[7] A. W. Senior *et al.*, "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, no. 7792, Art. no. 7792, Jan. 2020, doi: 10.1038/s41586-019-1923-7.

[8] "Computational predictions of protein structures associated with COVID-19," *Deepmind*. https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19.

[9] R. P. D. Bank, "RCSB PDB: Homepage." https://www.rcsb.org/.

[10] "UniProt." https://www.uniprot.org/.

[11] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021, doi: 10.1038/s41586-021-03819-2.

[12] "Highly accurate protein structure prediction with AlphaFold | Nature." https://www.nature.com/articles/s41586-021-03819-2.

[13] "ASReview - Active learning for Systematic Reviews," *ASReview*. https://asreview.nl/.

[14] H. He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009, doi: 10.1109/TKDE.2008.239.

[15] *ASReview against COVID-19*. ASReview, 2021. https://github.com/asreview/asreview-covid19.

[16] L. L. Wang *et al.*, "CORD-19: The COVID-19 Open Research Dataset," presented at the ACL-NLP-COVID19 2020, Online, Jul. 2020. https://aclanthology.org/2020.nlpcovid19-acl.1.

[17] "covid19_preprints," *figshare*, Mar. 26, 2020. https://figshare.com/articles/code/covid19_preprints/12033672/18.

[18] "CORD-19 Historical Releases." https://ai2-semanticscholar-cord-19.s3-us-west-2.amazonaws.com/historical_releases.html.

[19] De Bruin, Jonathan *et al.*, "ASReview Software Documentation," Apr. 2021, doi: 10.5281/ZENODO.4287119.

[20] R. van de Schoot *et al.*, "An open source machine learning framework for efficient and transparent systematic reviews," *Nat Mach Intell*, vol. 3, no. 2, pp. 125–133, Feb.

2021, doi: 10.1038/s42256-020-00287-7.
[21] “ASReview,” *GitHub*. https://github.com/asreview.
[22] J. M. Brauner *et al.*, “Inferring the effectiveness of government interventions against COVID-19,” *Science*, vol. 371, no. 6531, Feb. 2021, doi: 10.1126/science.abd9338.
[23] “Datasets.” http://epidemicforecasting.org/datasets.
[24] “COVID-19 Government Response Tracker.” https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker.
[25] CSSEGISandData, *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University*. 2021. https://github.com/CSSEGISandData/COVID-19.
[26] S. Flaxman *et al.*, “Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe,” *Nature*, vol. 584, no. 7820, pp. 257–261, Aug. 2020, doi: 10.1038/s41586-020-2405-7.
[27] M. D. Hoffman and A. Gelman, “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo,” *JMLR*, vol. 15, p. 31, Apr. 2014.
[28] “Understanding the effectiveness of government interventions against the resurgence of COVID-19 in Europe | Nature Communications.” https://www.nature.com/articles/s41467-021-26013-4.
[29] “How AI Spotted and Tracked the Coronavirus Outbreak.” https://www.wsj.com/articles/how-ai-spotted-and-tracked-the-coronavirus-outbreak-11580985001.
[30] “BlueDot | Who We Are,” *BlueDot*. https://bluedot.global/press/.
[31] “BlueDot Insights: Your near real-time infectious disease surveillance system,” *BlueDot*. https://bluedot.global/products/insights/.
[32] “BlueDot Explorer: Better public health surveillance for infectious diseases,” *BlueDot*. https://bluedot.global/products/explorer/.
[33] “BlueDot for Business | Outbreak SaaS for enterprise risk management,” *BlueDot*. https://bluedot.global/business/.
[34] “How Canadian AI start-up BlueDot spotted Coronavirus before anyone else had a clue,” Mar. 10, 2020. https://diginomica.com/how-canadian-ai-start-bluedot-spotted-coronavirus-anyone-else-had-clue.
[35] E. Niiler, “An AI Epidemiologist Sent the First Alerts of the Coronavirus,” *Wired*. https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/.
[36] I. I. Bogoch, A. Watts, A. Thomas-Bachli, C. Huber, M. U. G. Kraemer, and K. Khan, “Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel,” *J Travel Med*, vol. 27, no. 2, p. taaa008, Jan. 2020, doi: 10.1093/jtm/taaa008.
[37] “Potential for global spread of a novel coronavirus from China | Journal of Travel Medicine | Oxford Academic.” https://academic.oup.com/jtm/article/27/2/taaa011/5716260.
[38] S. Boseley, “Only three Ebola-infected travellers per month are likely to fly from west Africa,” *The Guardian*, Oct. 20, 2014. https://www.theguardian.com/world/2014/oct/21/ebola-exit-screening-most-efficient-says-lancet-study.
[39] “How a Toronto company used big data to predict the spread of Zika | The Star.” https://www.thestar.com/news/insight/2016/02/22/how-a-toronto-company-used-big-data-to-predict-the-spread-of-zika.html.
[40] TEDx Talks, *Getting out of this pandemic and staying out of the next one | Kamran Khan | TEDxToronto*, (Dec. 07, 2020). [Video]. https://www.youtube.com/watch?v=TkNl6qRGNjE.

[41] D. DeCaprio *et al.*, “Building a COVID-19 vulnerability index,” *Journal of Medical Artificial Intelligence*, vol. 3, no. 0, Art. no. 0, Dec. 2020, doi: 10.21037/jmai-20-47.
[42] D. DeCaprio, J. Gartner, C. J. McCall, T. Burgess, S. Kothari, and S. Sayed, “Building a COVID-19 Vulnerability Index,” p. 9, Mar. 2020, doi: 10.1101/2020.03.16.20036723.
[43] “Limited Data Set (LDS) Files | CMS.” https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets.
[44] “Illinois ACO deploys AI to identify high risk COVID-19 patients.” https://www.beckershospitalreview.com/artificial-intelligence/illinois-aco-deploys-ai-to-identify-high-risk-covid-19-patients.html.
[45] “CAIAC 1.1,” *CAIAC*. https://caiac.uwu.ai.
[46] “Collective and Augmented Intelligence Against COVID-19 (CAIAC),” *Stanford HAI*. https://hai.stanford.edu/events/caiac.
[47] “Collective & Augmented Intelligence Against COVID-19 (CAIAC) Final Report.” The Future Society, Oct. 2021.
[48] “Hypothesis.” https://web.hypothes.is/.
[49] “Hivemind produces exceptional data,” *Hivemind*. http://172.31.38.141/.
[50] “Otter Voice Meeting Notes.” https://otter.ai.
[51] “Kumu.” https://kumu.io/.
[52] “IBM Watson Discovery | IBM.” https://www.ibm.com/cloud/watson-discovery.
[53] “https://test-stability.forefront.link.” https://test-stability.forefront.link/.
[54] “Summarizing Books with Human Feedback.” https://openai.com/blog/summarizing-books/.
[55] “Compositional cyber-physical epidemiology of COVID-19 | Scientific Reports.” https://www.nature.com/articles/s41598-020-76507-2.
[56] P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen, “Complexity of the Basic Reproduction Number (R0) - Volume 25, Number 1—January 2019 - Emerging Infectious Diseases journal - CDC”, doi: 10.3201/eid2501.171901.
[57] J. W. Ro, N. Allen, W. Ai, D. Prasad, and P. S. Roop, “Compositional Cyber-Physical Epidemiology of COVID-19,” May 2020. doi: 10.1101/2020.04.26.20081125.
[58] *Compositional Epidemiology of COVID-19*. PRETzel, 2020. https://github.com/PRETgroup/ccpe-covid19.
[59] G. Viceconte and N. Petrosillo, “COVID-19 R0: Magic number or conundrum?,” *Infectious Disease Reports*, vol. 12, no. 1, Art. no. 1, Feb. 2020, doi: 10.4081/idr.2020.8516.
[60] N. Wilson, L. T. Barnard, A. Kvalsvig, A. Verrall, M. Baker, and M. Schwehm, “Modelling the Potential Health Impact of the COVID-19 Pandemic on a Hypothetical European Country,” Mar. 2020. doi: 10.1101/2020.03.20.20039776.
[61] G. Giordano *et al.*, “Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy,” *Nat Med*, vol. 26, no. 6, pp. 855–860, Jun. 2020, doi: 10.1038/s41591-020-0883-7.
[62] J. Lever and R. B. Altman, “Analyzing the vast coronavirus literature with CoronaCentral,” *Proc Natl Acad Sci U S A*, vol. 118, no. 23, p. e2100766118, Jun. 2021, doi: 10.1073/pnas.2100766118.
[63] J. Lever, *CoronaCentral Machine Learning*. 2021. https://github.com/jakelever/corona-ml.
[64] “Index of /pubmed/updatefiles.” https://ftp.ncbi.nlm.nih.gov/pubmed/updatefiles/.
[65] “CoronaCentral.” https://coronacentral.ai/.
[66] J. Lever, *CoronaCentral Machine Learning - stepByStep*. 2021. https://github.com/jakelever/corona-ml/blob/8ceb22af50d7277ebf05f2fd21bbbf68c080ed76/stepByStep.md.
[67] J. Lever, *CoronaCentral Machine Learning - Machine Learning Details*. 2021.

https://github.com/jakelever/corona-ml/blob/8ceb22af50d7277ebf05f2fd21bbbf68c080ed76/machineLearningDetails.md.
[68] P. Papotti, “CoronaCheck: Computational Fact Checking for Statistical Coronavirus Claims - https://coronacheck.eurecom.fr,” Mar. 2020.
[69] “CoronaCheck: Computational Fact Checking for Statistical Coronavirus Claims.” https://coronacheck.eurecom.fr/en.
[70] “Paolo Papotti: home page.” https://www.eurecom.fr/~papotti/.
[71] “Projects,” *itrummer.github.io*. https://itrummer.github.io/.
[72] G. Karagiannis, M. Saeed, P. Papotti, and I. Trummer, “Scrutinizer: a mixed-initiative approach to large-scale, data-driven claim verification,” *Proc. VLDB Endow.*, vol. 13, no. 12, pp. 2508–2521, Aug. 2020, doi: 10.14778/3407790.3407841.
[73] G. Lample and A. Conneau, “Cross-lingual Language Model Pretraining,” *arXiv:1901.07291 [cs]*, Jan. 2019, http://arxiv.org/abs/1901.07291.
[74] G. Karagiannis, *Stat claims factchecking project*. 2021. https://github.com/geokaragiannis/statchecker.
[75] “Privacy Policy of CoronaCheck,” *iubenda*. https://www.iubenda.com/privacy-policy/90061002/legal.
[76] J. Herzig, P. K. Nowak, T. Müller, F. Piccinno, and J. Eisenschlos, “TaPas: Weakly Supervised Table Parsing via Pre-training,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, Jul. 2020, pp. 4320–4333. doi: 10.18653/v1/2020.acl-main.398.
[77] W. Chen *et al.*, “TabFact: A Large-scale Dataset for Table-based Fact Verification,” *arXiv:1909.02164 [cs]*, Jun. 2020, http://arxiv.org/abs/1909.02164.
[78] S. Jo *et al.*, “Verifying Text Summaries of Relational Data Sets,” *arXiv:1804.07686 [cs]*, Aug. 2018, http://arxiv.org/abs/1804.07686.
[79] B. Prud’homme *et al.*, “COVI,” *Mila*. https://mila.quebec/en/project/covi/.
[80] H. Alsdurf *et al.*, “COVI White Paper,” *arXiv:2005.08502 [cs]*, Jul. 2020, http://arxiv.org/abs/2005.08502.
[81] “COVI-ML.” Mila. https://github.com/mila-iqia/COVI-ML.
[82] P. Gupta *et al.*, “COVI-AgentSim: an Agent-based Model for Evaluating Methods of Digital Contact Tracing,” *arXiv:2010.16004 [cs]*, Oct. 2020, http://arxiv.org/abs/2010.16004.
[83] “Federal government rules out adoption of Mila Institute’s COVID-19 contact-tracing app,” *The Logic*, Jun. 04, 2020. https://thelogic.co/news/federal-government-rules-out-adoption-of-mila-institutes-covid-19-contact-tracing-app/.
[84] “Community - COVID 19 forecast hub.” https://covid19forecasthub.org/community/.
[85] E. L. Ray *et al.*, “Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S.,” Aug. 2020. doi: 10.1101/2020.08.19.20177493.
[86] “Visualization - COVID-19 Forecast Hub.” https://viz.covid19forecasthub.org/.
[87] “About the Hub,” *COVID-19 Forecast Hub*. https://covid19forecasthub.org/doc/
[88] E. Y. Cramer *et al.*, “The United States COVID-19 Forecast Hub dataset,” Nov. 2021. doi: 10.1101/2021.11.04.21265886.
[89] CDC, “Coronavirus Disease 2019 (COVID-19),” *Centers for Disease Control and Prevention*, Feb. 11, 2020. https://www.cdc.gov/coronavirus/2019-ncov/science/forecasting/forecasts-cases.html.
[90] E. Dong, H. Du, and L. Gardner, “An interactive web-based dashboard to track COVID-19 in real time,” *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, May 2020, doi: 10.1016/S1473-3099(20)30120-1.
[91] *reichlab/covid19-forecast-hub*. The Reich Lab at UMass-Amherst, 2021. https://github.com/reichlab/covid19-forecast-hub/blob/4294b7ee31ca693d7fa976194c76eb8fbb1cf750/README.md.

[92] “Project: COVID-19 Forecasts - Zoltar.” https://zoltardata.com/project/44.
[93] *covidHubUtils*. The Reich Lab at UMass-Amherst, 2021. https://github.com/reichlab/covidHubUtils.
[94] *covidData*. The Reich Lab at UMass-Amherst, 2021. https://github.com/reichlab/covidData.
[95] Z.-H. Zhou, “Ensemble Learning,” in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA: Springer US, 2009, pp. 270–273. doi: 10.1007/978-0-387-73003-5_293.
[96] J. Bracher, E. L. Ray, T. Gneiting, and N. G. Reich, “Evaluating epidemic forecasts in an interval format,” *PLOS Computational Biology*, vol. 17, no. 2, p. e1008618, Feb. 2021, doi: 10.1371/journal.pcbi.1008618.
[97] “Challenges in training ensembles to forecast COVID-19 cases and deaths in the United States - International Institute of Forecasters.” https://forecasters.org/blog/2021/04/09/challenges-in-training-ensembles-to-forecast-covid-19-cases-and-deaths-in-the-united-states/.
[98] “European Covid-19 Forecast Hub.” https://covid19forecasthub.eu/.
[99] “ForecastHub.” https://kitmetricslab.github.io/forecasthub/about.
[100] “COVID-19 Patient Redistribution.” https://covid-hospital-operations.com/home.
[101] C. G. / P. Oct 27 and 2020, “New models can help hospitals stay ahead of COVID-19 surges,” *The Hub*, Oct. 27, 2020. https://hub.jhu.edu/2020/10/27/models-help-hospital-systems-allocate-resources-during-covid-19-surge/.
[102] F. Parker, H. Sawczuk, F. Ganjkhanloo, F. Ahmadi, and K. Ghobadi, “Optimal Resource and Demand Redistribution for Healthcare Systems Under Stress from COVID-19,” *arXiv:2011.03528 [cs, math]*, Nov. 2020, http://arxiv.org/abs/2011.03528.
[103] “COVID-19 Reported Patient Impact and Hospital Capacity by Facility | HealthData.gov.” https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/anag-cw7u.
[104] F. Parker, *COVID-19 Patient Redistribution Site*. 2021. https://github.com/flixpar/covid-patient-redistribution-site.
[105] “New interactive app simulates COVID-19 spread in NZ.” https://www.tepunahamatatini.ac.nz/2020/04/24/new-interactive-app-simulates-covid-19-spread-in-nz/.
[106] A. James, M. Plank, and A. Lustig, “COVID-19 Take Control Simulator.” http://covid19takecontrol.nectar.auckland.ac.nz/covid19_takeControl/
[107] “History of the COVID-19 Alert System,” *Unite against COVID-19*. https://covid19.govt.nz/about-our-covid-19-response/history-of-the-covid-19-alert-system/ .
[108] “COVID-19: Data resources,” *Ministry of Health NZ*. https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-data-and-statistics/covid-19-data-resources.
[109] S. Hendy *et al.*, “Mathematical modelling to inform New Zealand’s COVID-19 response,” *Journal of the Royal Society of New Zealand*, vol. 51, no. sup1, pp. S86–S106, May 2021, doi: 10.1080/03036758.2021.1876111.
[110] N. G. Davies *et al.*, “Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study,” *The Lancet Public Health*, vol. 5, no. 7, pp. e375–e385, Jul. 2020, doi: 10.1016/S2468-2667(20)30133-X.
[111] D. Group, “Delphi Research Group.” https://delphi.cmu.edu/about/.
[112] D. Group, “Delphi National Center of Excellence for Epidemic Forecasting.” https://delphi.cmu.edu/about/center-of-excellence/.

[113] D. McDonald *et al.*, *Can Auxiliary Indicators Improve COVID-19 Forecasting and Hotspot Prediction?* 2021. doi: 10.1101/2021.06.22.21259346.
[114] D. Group, “Delphi’s COVID-19 Trends and Impact Surveys (CTIS).” https://delphi.cmu.edu/covid19/ctis/.
[115] “Model: CMU-TimeSeries - Zoltar.” https://zoltardata.com/model/307.
[116] “Home - COVID 19 forecast hub.” https://covid19forecasthub.org/.
[117] “COVIDcast Epidata API,” *Delphi Epidata API*. https://cmu-delphi.github.io/delphi-epidata/api/covidcast.html.
[118] D. Group, “Accessing Open COVID-19 Data via the COVIDcast Epidata API,” Oct. 07, 2020. https://delphi.cmu.edu/blog/2020/10/07/accessing-open-covid-19-data-via-the-covidcast-epidata-api/.
[119] “COVIDcast API: An open repository of real-time COVID-19 indicators,” *News-Medical.net*, Jul. 21, 2021. https://www.news-medical.net/news/20210721/COVIDcast-API-An-open-repository-of-real-time-COVID-19-indicators.aspx.
[120] *Delphi COVID-19 Monitoring Dashboard*. Delphi, 2021. https://github.com/cmu-delphi/www-covidcast.
[121] “International Health Regulations (2005).” https://www.who.int/publications-detail-redirect/9789241580410.
[122] S. Chaganti *et al.*, “Automated Quantification of CT Patterns Associated with COVID-19 from Chest CT,” *Radiology: Artificial Intelligence*, vol. 2, no. 4, p. e200048, Jul. 2020, doi: 10.1148/ryai.2020200048.
[123] S. Chaganti *et al.*, “Automated Quantification of CT Patterns Associated with COVID-19 from Chest CT,” *arXiv:2004.01279 [cs, eess]*, Nov. 2020, doi: 10.1148/ryai.2020200048#.
[124] “AI COVID-19.” https://www.siemens-healthineers.com/en-us/medical-imaging/diagnostic-imaging/ai-covid-19-algorithm.
[125] D. Yang *et al.*, “Automatic Liver Segmentation Using an Adversarial Image-to-Image Network,” in *Medical Image Computing and Computer Assisted Intervention − MICCAI 2017*, Cham, 2017, pp. 507–515. doi: 10.1007/978-3-319-66179-7_58.
[126] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Cham, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
[127] “syngo.via Frontier.” https://www.siemens-healthineers.com/en-us/medical-imaging-it/advanced-visualization-solutions/syngo-via-frontier).
[128] R. Miikkulainen, O. Francon, E. Meyerson, X. Qiu, E. Canzani, and B. Hodjat, “From Prediction to Prescription: Evolutionary Optimization of Non-Pharmaceutical Interventions in the COVID-19 Pandemic,” *arXiv:2005.13766 [cs]*, Aug. 2020, http://arxiv.org/abs/2005.13766.
[129] R. Miikkulainen *et al.*, “From Prediction to Prescription: Evolutionary Optimization of Nonpharmaceutical Interventions in the COVID-19 Pandemic,” *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 2, pp. 386–401, Apr. 2021, doi: 10.1109/TEVC.2021.3063217.
[130] “Evolutionary AI.” https://evolution.ml/demos/npidashboard/.
[131] K. O. Stanley, J. Clune, J. Lehman, and R. Miikkulainen, “Designing neural networks through neuroevolution,” *Nat Mach Intell*, vol. 1, no. 1, pp. 24–35, Jan. 2019, doi: 10.1038/s42256-018-0006-z.
[132] C. A. Coello Coello, “A Comprehensive Survey of Evolutionary-Based Multiobjective

Optimization Techniques," *Knowledge and Information Systems*, vol. 1, no. 3, pp. 269–308, Aug. 1999, doi: 10.1007/BF03325101.
[133] M. T. M. Emmerich and A. H. Deutz, "A tutorial on multiobjective optimization: fundamentals and evolutionary methods," *Nat Comput*, vol. 17, no. 3, pp. 585–609, Sep. 2018, doi: 10.1007/s11047-018-9685-y.
[134] G. Analytics, "Coronavirus Dashboard." https://knowledge.geospark.io/coronavirus-dashboard.
[135] A. Brownfield, "Not Just Another Coronavirus Map," *Medium*, Apr. 03, 2020. https://blog.geospark.io/not-just-another-coronavirus-map-c0a62e9bd2d8.
[136] M. Franklin, "From Localized Incident Reporting to Global Monitoring — How Hyperion has Processed and Analyzed…," *Medium*, Jan. 31, 2020. https://blog.geospark.io/from-localized-incident-reporting-to-global-monitoring-how-hyperion-has-processed-and-analyzed-840d0e48fdb2.
[137] "IBM's New AI Tool Parses A Tidal Wave of Coronavirus Research," *IEEE Spectrum*, Jul. 01, 2020. https://spectrum.ieee.org/ibms-new-ai-tool-parses-a-tidal-wave-of-coronavirus-research.
[138] "Use deep search to explore the COVID-19 corpus," *IBM Research*, Jun. 05, 2018. https://research.ibm.com/covid19/deep-search/.
[139] "CORD-19 | Semantic Scholar." https://www.semanticscholar.org/cord19.
[140] "Home - PMC - NCBI." https://www.ncbi.nlm.nih.gov/pmc/.
[141] "Global research on coronavirus disease (COVID-19)." https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov.
[142] D. S. Wishart *et al.*, "DrugBank 5.0: a major update to the DrugBank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074–D1082, Jan. 2018, doi: 10.1093/nar/gkx1037.
[143] "GenBank Overview." https://www.ncbi.nlm.nih.gov/genbank/.
[144] "Home - ClinicalTrials.gov." https://clinicaltrials.gov/.
[145] "International Clinical Trials Registry Platform (ICTRP)." https://www.who.int/clinical-trials-registry-platform.
[146] "IBM COVID-19 Knowledge Graph." https://ds-covid19.res.ibm.com/datasources.
[147] P. W. J. Staar, M. Dolfi, C. Auer, and C. Bekas, "Corpus Conversion Service: A Machine Learning Platform to Ingest Documents at Scale," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London United Kingdom, Jul. 2018, pp. 774–782. doi: 10.1145/3219819.3219834.
[148] "Deep Search - COVID-19- IBM Research." https://research.ibm.com/interactive/covid19/deep-search/.
[149] "icolung - quantitative COVID-19 analysis of non-contrast thorax CT scans | icometrix." https://icometrix.com/services/icolung.
[150] D. Smeets, "The role of imaging AI and CT in COVID-19," *icometrix*. https://icometrix.com/resources/the-role-of-imaging-ai--and--ct-in-covid19
[151] "Belgian initiative for AI lung scan analysis in fight against COVID-19 goes European | imec." https://www.imec-int.com/en/press/belgian-initiative-ai-lung-scan-analysis-fight-against-covid-19-goes-european.
[152] icometrix, "FDA permits use of icometrix's AI-based quantification for chest CT in COVID-19." https://www.prnewswire.com/news-releases/fda-permits-use-of-icometrixs-ai-based-quantification-for-chest-ct-in-covid-19-301057564.html.
[153] "AI-based chest CT analysis enabling rapid COVID diagnosis and prognosis | icovid Project | Fact Sheet | H2020 | CORDIS | European Commission."

https://cordis.europa.eu/project/id/101016131.
[154] A. D. Berenguer *et al.*, “Explainable-by-design Semi-Supervised Representation Learning for COVID-19 Diagnosis from CT Imaging,” *arXiv:2011.11719 [cs, eess]*, Sep. 2021, http://arxiv.org/abs/2011.11719.
[155] S. G. Armato *et al.*, “The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans,” *Med Phys*, vol. 38, no. 2, pp. 915–931, Feb. 2011, doi: 10.1118/1.3528204.
[156] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Cham, 2016, pp. 424–432. doi: 10.1007/978-3-319-46723-8_49.
[157] K. Kamnitsas *et al.*, “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,” *Medical Image Analysis*, vol. 36, pp. 61–78, Feb. 2017, doi: 10.1016/j.media.2016.10.004.
[158] M. Marshall and L. Gardner, “Predicting COVID-19 Risk in the U.S.,” Sep. 03, 2020. https://systems.jhu.edu/research/public-health/predicting-covid-19-risk/
[159] “Social Distancing Metrics | SafeGraph Docs,” *SafeGraph*. https://docs.safegraph.com/docs/social-distancing-metrics.
[160] Q. Chen, A. Allot, and Z. Lu, “Keep up with the latest coronavirus research,” *Nature*, vol. 579, no. 7798, pp. 193–193, Mar. 2020, doi: 10.1038/d41586-020-00694-1.
[161] Q. Chen, A. Allot, and Z. Lu, “LitCovid: an open database of COVID-19 literature,” *Nucleic Acids Research*, vol. 49, no. D1, pp. D1534–D1540, Jan. 2021, doi: 10.1093/nar/gkaa952.
[162] “PubMed,” *PubMed*. https://pubmed.ncbi.nlm.nih.gov/.
[163] “LitCovid.” https://www.ncbi.nlm.nih.gov/research/coronavirus/.
[164] “LitSuggest,” *LitSuggest*. https://www.ncbi.nlm.nih.gov/research/litsuggest/.
[165] “BioBERT: a pre-trained biomedical language representation model for biomedical text mining | Bioinformatics | Oxford Academic.” https://academic.oup.com/bioinformatics/article/36/4/1234/5566506.
[166] “spaCy · Industrial-strength Natural Language Processing in Python.” https://spacy.io/.
[167] C.-H. Wei, H.-Y. Kao, and Z. Lu, “PubTator: a web-based text mining tool for assisting biocuration,” *Nucleic Acids Res*, vol. 41, no. Web Server issue, pp. W518–W522, Jul. 2013, doi: 10.1093/nar/gkt441.
[168] “Welcome to Apache Solr.” https://solr.apache.org/index.html.
[169] A. Venkatakrishnan *et al.*, “Knowledge synthesis of 100 million biomedical documents augments the deep expression profiling of coronavirus receptors,” *eLife*, vol. 9, p. e58040, May 2020, doi: 10.7554/eLife.58040.
[170] “nference® Launches First-In-Class Single Cell RNA-Sequencing Resource to Arm Researchers Fighting COVID-19 Pandemic,” Mar. 30, 2020. https://www.businesswire.com/news/home/20200330005606/en/nference%C2%AE-Launches-First-In-Class-Single-Cell-RNA-Sequencing-Resource-to-Arm-Researchers-Fighting-COVID-19-Pandemic.
[171] “Home - GEO - NCBI.” https://www.ncbi.nlm.nih.gov/geo/.
[172] “GTEx Portal.” https://gtexportal.org/home/.
[173] “The Human Protein Atlas.” https://www.proteinatlas.org/.
[174] “Human Proteome Map.” https://www.humanproteomemap.org/.
[175] J. Park, A. L. Marquez, A. Puranik, A. Rajasekharan, M. Aravamudan, and E. Garcia-Rivera, “Recapitulation and Retrospective Prediction of Biomedical Associations Using Temporally-enabled Word Embeddings,” May 2019. doi: 10.1101/627513.

[176] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.
[177] O. Gozes *et al.*, “Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis,” *arXiv:2003.05037 [cs, eess]*, Mar. 2020, http://arxiv.org/abs/2003.05037.
[178] “About Us,” *RADLogics*. https://www.radlogics.com/about-us/.
[179] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
[180] J. A. Fries *et al.*, “Ontology-driven weak supervision for clinical entity classification in electronic health records,” *Nat Commun*, vol. 12, no. 1, p. 2017, Apr. 2021, doi: 10.1038/s41467-021-22328-4.
[181] “UMLS 2018AA Release Available.” https://www.nlm.nih.gov/pubs/techbull/tb.html.
[182] N. C. for B. Information, U. S. N. L. of M. 8600 R. Pike, B. MD, and 20894 Usa, *SPECIALIST Lexicon and Lexical Tools*. National Library of Medicine (US), 2021. https://www.ncbi.nlm.nih.gov/books/NBK9680/.
[183] “Disease Ontology: a backbone for disease semantic integration | Nucleic Acids Research | Oxford Academic.” https://academic.oup.com/nar/article/40/D1/D940/2903651.
[184] K. Degtyarenko *et al.*, “ChEBI: a database and ontology for chemical entities of biological interest,” *Nucleic Acids Research*, vol. 36, no. suppl_1, pp. D344–D350, Jan. 2008, doi: 10.1093/nar/gkm791.
[185] A. P. Davis *et al.*, “The Comparative Toxicogenomics Database’s 10th year anniversary: update 2015,” *Nucleic Acids Research*, vol. 43, no. D1, pp. D914–D920, Jan. 2015, doi: 10.1093/nar/gku935.
[186] J. Shang, L. Liu, X. Ren, X. Gu, T. Ren, and J. Han, “Learning Named Entity Tagger using Domain-Specific Dictionary,” *arXiv:1809.03599 [cs]*, Sep. 2018, http://arxiv.org/abs/1809.03599.
[187] W. Zhou, V. I. Torvik, and N. R. Smalheiser, “ADAM: another database of abbreviations in MEDLINE,” *Bioinformatics*, vol. 22, no. 22, pp. 2813–2818, Nov. 2006, doi: 10.1093/bioinformatics/btl480.
[188] Y. Wu *et al.*, “A long journey to short abbreviations: developing an open-source framework for clinical abbreviation recognition and disambiguation (CARD),” *Journal of the American Medical Informatics Association*, vol. 24, no. e1, pp. e79–e86, Apr. 2017, doi: 10.1093/jamia/ocw109.
[189] *Trove*. som-shahlab, 2021. https://github.com/som-shahlab/trove.
[190] D. Kai, G.-P. Goldstein, A. Morgunov, V. Nangalia, and A. Rotkirch, “Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations,” *arXiv:2004.13553 [physics, q-bio]*, Apr. 2020, http://arxiv.org/abs/2004.13553.
[191] “Agent-Based Model for COVID-19 Masking.” http://dek.ai/masksim/.
[192] R. S. McGee, *SEIRS+ Model Framework*. 2021. https://github.com/ryansmcgee/seirsplus.
[193] S. Talic *et al.*, “Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis,” *BMJ*, vol. 375, p. e068302, Nov. 2021, doi: 10.1136/bmj-2021-068302.