Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: October 31, 2022

Citation: Pastor-Escuredo D and Olmos LE, 2022. The landscape of Artificial Intelligence Models and Applications for Epidemic Outbreaks, Medical Research Archives, [online] 10(10). https://doi.org/10.18103/mra.v10i10.3175

DOI
https://doi.org/10.18103/mra.v10i10.3175

ISSN: 2375-1924

RESEARCH ARTICLE

The landscape of Artificial Intelligence Models and Applications for Epidemic Outbreaks

David Pastor-Escuredo[1,2,3,*], Luis E. Olmos[4]

[1]LifeD Lab, Spain
[2]UCL, UK
[3]UNICEF
[4] Facultad de Ciencias Básicas, Universidad de Medellín, Carrera 87 No. 30-65, 050026 Medellín, Colombia

* david@lifedlab.org

ABSRACT
The epidemiology has recently witnessed great advances based on computational models. Its scope and impact are getting wider thanks to the new data sources feeding analytical frameworks and models. Besides traditional variables considered in epidemiology, large-scale social patterns can be now integrated in real time with multi-source data bridging the gap between different scales. In a hyper-connected world, models and analysis of interactions and social behaviors are key to understand and stop outbreaks. Big Data along with apps are enabling for validating and refining models with real world data at scale, as well as new applications and frameworks to map and track diseases in real time or optimize the necessary resources and interventions such as testing and vaccination strategies. Digital epidemiology is positioning itself as a discipline necessary to control epidemics and implement actionable protocols and policies. In this review we address the research areas configuring current digital epidemiology: transmission and propagation models and descriptions based on human networks and contact tracing, mobility analysis and spatio-temporal propagation of infectious diseases and *infodemics* that comprises the study of information and knowledge propagation. Digital epidemiology has the potential to create new operational mechanisms for prevention and mitigation, monitoring of the evolution of epidemics, assessing their impact and evaluating the pharmaceutical and non-pharmaceutical measures to fight the outbreaks. Epidemics have to be approached from the lens of complexity science as they require systemic solutions. Opportunities and challenges to tackle epidemics more effectively and with a human-centered vision are discussed here.

Keywords: Digital epidemiology, contact tracing, networks, disease propagation, epidemiological impact, infodemics, human mobility, Big Data, epidemiological models, privacy, ethics, resilience.

1. INTRODUCTION

Epidemiology is the field that encompasses the study of the distribution, prevalence and etiology of human diseases [1-3]. Although data and models have always been part of epidemiology [4], the appearance of new sources of Big Data and technology [5] have enabled computational frameworks and opportunities to increase impact and knowledge [2]. In parallel to the appearance and use of new data sources, the growth of Artificial Intelligence (AI) and specially Machine Learning (ML) techniques [6-9], are giving rise to multiple methodologies and applications that can be categorized as the emergent digital epidemiology.

The studies of epidemiology have been grounded in data collected in clinical practice and field work [10]. Based on individual's web searches, researchers have been developing algorithms (e.g. Google flu trends) to monitor and estimate the progression of epidemics. [11-14]. This trend generated the first challenges in methodologies and epistemology of the new field [15-17]. Currently, we count on many sources such as social media, social networks, mobile apps and other services that generate data [18]. The COVID-19 pandemic emergency sped up the adoption of digital tech in all sectors with a slow digital transformation.

For this reason, digital epidemiology seeks to understand the dynamics of patterns, both social and clinical, of people affected by the disease, and the causes of these patterns [18]. According to the definition of the World Health Organization (WHO), epidemiology is the study of the distribution and the determinants of the estates and events related with health and the application to disease control and management and other health challenges. Therefore, epidemiology has a pragmatic dimension aimed to improve response systems against epidemics including prevention, management, mitigation and preparation to future epidemics and waves. Besides, due to its complexity and importance, epidemiology is promoting new research practices and techniques. Digital epidemiology has a wider scope, it is not only about new technology, but mostly about the scope of epidemiology to manage complexity of diseases and their factors: biological, social and, environmental. More data is used and analyzed including data that was not thought of or designed for health applications [18]. In this document, we overview the work areas and the ongoing work along with the most important contributions where COVID-19 has been a disruption point [19].

2. MODELS AND NETWORKS

There are two types of epidemiological models: models based on equations and models based on agents (ABMs). Models based on equations assume homogeneity and similar collective behavior [4]. Progress on computation has enabled ABMs that can model heterogeneity in epidemics, e.g. detailed age-stratification, population density, vaccination coverage, and realistic social networks [20-23]. Both types of models are based on the conceptualization of the disease through different states, being SIR (Susceptible, Infected, Recovered) the most used and extended. There are several extensions to introduce new complexities and details. For instance, the model SEIR inserts the state Exposed that comprises people infected in the incubation process. Each state can be parametrized towards quantifying the transition between states given biological and social criteria inferred from clinical data, surveys and questionnaires. While inputs may greatly differ between models, the outputs typically consist in forecasts of infected cases and deaths, and estimations of epidemiological parameters such as the reproductive number R_0. The reproductive number R_0 is the average number of contagions generated by each person and R is the full distribution of (R_n) in each node. Models integrate diseases characteristics, temporality and volume of the epidemic.

A recent key element in sophisticated epidemiological models is to consider the contact between individuals (contact matrices) as multi-layer networks, so the disease depends on the structure, properties and topology of the contact networks that can be partially modelled using metapopulation approaches [22,23]. Even when all scales are interlinked (from biochemical to social), epidemiology based on networks is useful to predict the spatio-temporal propagation of the disease, and also, to implement social distancing policies. Networks allow modelling the behavioral component of the disease through the network itself and its dynamics: percolation and diffusion.

Considering the topology of the network, a disease can propagate with different speed and strengths [24-26]. Several studies have used “scale-free” networks (networks whose degree distribution follows and power law) to simulate a realistic scenario of how people are interconnected. Assuming an infinite population, for infections that do not confer immunity upon recovery, the spreading process on scale-free networks does not exhibit an epidemiological threshold, and has a large heterogeneity in the behaviour of the network. This is a consequence of the extreme heterogeneity in the connectivity distribution of scale-free networks [24]. Considering finite network sizes, more realistic models reproduce threshold and still heterogeneous behavior [27-29].

Percolation is a process that affects network connectedness by deactivating links. This process removes links uniformly at random with some given probability. After removing a critical fraction of links, the network is fragmented in two or several smaller networks. This process features a phase transition that resembles critical transitions in several contexts [30]. The structure is key for this dynamic phenomenon due to the compartments of the network [31,32]. Through percolation it is possible to model the dynamics of disease spreading, when the epidemiological threshold is passed and size of disease outbreaks [33]. Diffusion is the other dynamic phenomenon that can be studied. Diffusion is the process from which several nodes are reached from one node and depends on the topology of the network and the dynamics of the disease. Predicting diffusion is critical to slow down and stop epidemics [34].

Sophisticated epidemiological models consider contact tracing matrices that enable the reconstruction of the contagion network if they are properly designed and collected. These matrices can be stratified (different age groups and gender groups), based on metapopulations [22,23] or multi-layer (if several flows are labelled, i.e. work travels, home travels, leisure travels, etc) [20]. Recent methodological advances comprise the use of hyper-graphs with links that connect several nodes instead of pair-to-pair links [35]. Thus, networks are used to build risk forecast and propagation forecast systems, including the analysis of transportation hubs [36], confirming that heterogeneity favors the propagation as it is easier to percolate. Consequently, actions to prevent, stop and contain epidemics seek to reduce and make more homogeneous the degree of the nodes in all scales for a given time window, so the epidemic is easier to control [37-39].

Network analysis has witnessed a new revolution during COVID-19 due to the new data sources acquired via Bluetooth and geo-location of mobile devices that enable GPS-based and proximity-based contact tracing to obtain dynamic and high-resolution matrices [40]. Multi-source networks will enable multi-partita networks where interactions between people and locations can be represented more realistically.

Networks are also useful to track and understand recovery and resilience, which in this case is favored by heterogeneity processes of recovery within the network [26,41]. Another application is to understand the interaction among concurrent diseases [42].

3. MOBILITY AND PROPAGATION

Mobility has a direct impact on disease propagation air-borne or vector-borne. The mobility studies have a long tradition but have been hampered by the lack of dynamic and fine-grained data to differentiate types of mobility [43]. In the last decades we have witnessed different types of mobility: tourism, events, business, long-term labor or students' mobility. It is not possible to properly study different layers of mobility through surveys and static data. The analysis of mobility and spatial characteristics of diseases depends on the availability and resolution of longitudinal data [44-46].

Human mobility is multi-scale in temporal and spatial dimensions [44,47-49]. Human mobility is also multilayered depending on the population flows. These layers are interconnected and each of them are propagated through a "social medium" [50]. The structure of mobility has an amplifier effect in the propagation due to diffusion and percolation if it is not properly managed [51]. First studies in mobility as epidemiological factor were focused on the global scale based on demographic data and international mobility statistics [52]. The temporal resolution of data in these studies only allowed studying seasonal variability [53]. The models used were gravitation-driven model and radiation-driven model [54]. However, these models only work under strong assumptions and it is difficult to make them work in epidemiological practice [55].

In vector-borne diseases, such as malaria or dengue, or diseases transmitted through air and water, small scale mobility affects the exposition of people to the disease whereas large-scale mobility affects the introduction, reinsertion and circulation of the contagions and even the global propagation [56-58]. Frequently, diseases induce a systemic change in mobility with hard-to-control impact [59]. For this reason, it has been identified the need to create monitoring mechanisms of mobility based on high-resolution mobile devices data.

Mobile phone data are generated from telecom operators and contain geolocation of calls and connections to the Internet. Also, geolocation services for smartphones allow capturing mobility traces (i.e. Cuebiq, Foursquare, etc). The temporal resolution of this data is very high [60,61]. The data requires an anonymization and aggregation process to preserve privacy [62-64]

High-resolution longitudinal allowed the characterization of "hotspots" and optimize the location of actions to prevent and stop the diseases [65,66]. However, mobile phone data has enabled the revolution of mobility for epidemics [67-73]. Sources and sinks characterization was the first and one the most notable to decipher the structure of propagation and generate risk maps separate from vector density maps [67]. However, a more detailed study based on mobility flows descriptors [49], can help understand the dynamics of risk and super-propagation phenomena [74].

Among the mobility phenomena, cultural events in many regions of the world have been analyzed as high-impact events in epidemics [75]. Long-term mobility analysis and mobility profiles are useful tools to understand the dynamics of the epidemics in a disaggregated way [76]. Disaggregation is necessary to understand the socio-economics of the epidemics [39,77-79] and the relationships with other sectors such as work, tourism [37,80] or agriculture and the rural-urban migration [48]. Finally, mobile phone data and survey data can be integrated to have high spatio-temporal and demographic resolutions [81].

During SARS-CoV-2, the number of applications and use cases of mobile phone data has increased [82], including distancing and lockdown measures: lockdown enforcement, measuring the epidemiological impact of the lockdowns and distancing and evaluating the measures for re-opening [39,83].

4. INFODEMICS

Decision-making during pandemics is key for good response and management and to stop negative effects. The asymmetry of negative impact requires additional actions to avoid systemic risk [84-86]. Decision-making demands the right information with the right timing [87], for this reason, information propagation during pandemics has become a key use for United Nations General Secretary [88]. SARS-CoV-2 has been marked by the spread of fake news and news that generate division [89-91]. When this situation gets more severe in moments of crisis, it becomes an infodemic [92-94].

Several works have been done to study information propagation, especially rumor and fake news, using analogies of disease spreading across complex networks [95-97]. For instance, through network analysis it is possible to discover who are the leaders of the social media and their influence in information propagation [98-100] and quantify viral processes and info spreaders [101]. Beyond information propagation, semantics analysis is a useful tool to classify text. New Deep Learning tools [7] are making this task accurate and scalable [102-104].

Facing the risk of hatred content, it is necessary to highlight the need of propagating positive information, being constructive through the social media and the networks. Information empowers the population to make better decisions at the individual and the collective levels. Information can help people keep their environment safe. Furthermore, it helps building up resilience and increase socio-economic impact driven by Collective Intelligence [105].

Information gathered by citizens helps manage risk and understand the epidemic better [106,107], feeding computational systems and models that deal with probabilities beyond demographic and clinical data. In this sense, new sensors to monitor variables and dynamic changes in the population are necessary. There exist already several tools to classify disease analyzing coughing [108] or problems with the smelling [109]. Finally, new channels between authorities and the population are necessary to build up trust and improve response.

5. ARTIFICIAL INTELLIGENCE-DRIVEN POLICY

5.1. Prediction and prevention

An early and rapid response minimizes and mitigates the impact of the epidemics. Models are principally used to predict the evolution of the epidemics. The prediction is based on the area of influence, the epidemiological curves and R_0. The models are expressed in terms of variables like population density, age and gender, implying limitations in the understanding and the prediction of propagation and impact of the epidemic. Variables like vulnerability, socio-economic inequality or WASH infrastructure are key to having more effective models. Epidemiology complexity increases with the variability and the complexity of the ecosystems where the epidemics propagate, hampering the use of local models. Some efforts have been made to develop clustering strategies to have epidemiological profiles for a geographical model [110-112]. Furthermore, models need to be more dynamic to include and explore changes in public policy, locally and internationally. This implies better data infrastructure (Health Data Spaces) and better models to move towards deployment and production phases. Furthermore, process and model parameters backtracking is necessary to perform causality analysis of test clinical evidence.

Apart from modelling, researchers have applied Machine Learning (ML) tools to predict the behavior of epidemiological curves [113,114], and analyze temporal patterns [115,116]. These frameworks have potential for policy making during distancing, lockdown and reopening. Training these models needs datasets that are not often available [117]. For instance, during the crisis of SARS-CoV-2, researchers have used data from other diseases like flu even when the behavior is very different [118,119]. In other cases, it is necessary to use small datasets [6].

Some diseases have a strong environmental component, for instance, increasing the density of vectors [81,120]. Thus, it is important to integrate environmental data and social models with high resolution and real-time. For diseases where the transmission is mainly from individual-to-individual, it is necessary to model asymptomatic cases and their contribution to the propagation [121,122]. There is new research to identify biomarkers and have clinical studies to control asymptomatic cases and understand different immunity, being necessary to configure model parameters [123-125].

Super-propagation has also become central because it has been observed a great variability in the distribution of R (dispersion k), giving rise to super-propagation events and spots. The role of super-propagators at the individual level is also important and can only be studied through contact-tracing matrices [74].

The society needs new tools to manage systemic risk of epidemics. This implies managing the information better by taking advantage of social systems exploiting complexity. Risk is multi-dimensional and even though the health response is the most important phase, it is necessary to leverage economic risk, social inequalities, drawbacks with rights and freedom and the effects on cognition and psychological state of the population.

5.2. Impact tracking and assessment

Non-pharmaceutical measures have become very relevant including lockdowns, distancing, contact tracing and mobility analysis [126]. The objective of these measures is to reduce R_0 (average of distribution R) being the output of predictive models [127,128].

The strategies of distancing are restrictive depending on the transmission medium, the morbidity and the mortality of the disease [129]. To track the effect of these strategies now we can use other data such as mobile phone data [82], Internet searches and social media [130] and other data generated by mobile devices. However, these kinds of systems are not fully implemented in our world. Some agencies are developing workflows to assess the impact on the most vulnerable populations and have a global understanding of the social dimension of epidemics [76].

One of the new systems is the digital contact tracing based on Bluetooth apps or physical proximity derived from GPS location. There are several architectures, centralized (PEPP-PT) and decentralized (DP-3T) to manage the info about risk of contagion. These techniques allow, given privacy and security mechanisms, generating suitable contact matrices with high disaggregation.

Information curation is another key process to avoid negative effects [131]. Ad-hoc systems are normally better than general digital platforms to ensure a responsible flow of information and data [132,133]. Dedicated chatbots and curation pipelines are part of the new epidemiology [6].

5.3. Evaluation of epidemiological measures

Evaluation of measures is key to improve response systems for the short and long term. Evaluation is key for governance and policy, so the mechanisms have to be truthful and transparent to measure the impacts. COVID-19 pandemic has sped up the innovation in this area due to the severity of the distancing and lockdowns and their socio-economic impact. Deep Learning algorithms have been used to simulate scenarios of the pandemic at a global scale [134] and measure the effects of lockdowns [135]. Mobile phone data has been used to measure the effects of the measures in geographical areas but also in meta-populations and population target groups [39,76,78]. A global challenge is to isolate the effect of each mitigation measure to be quantified and evaluated with data [136-139]. Global pandemic has shown the need to make dynamic policy and update the measures using computational models in nearly real-time [20].

6. CONCLUSION

Epidemics are multi-dimensional: molecular, clinical, social and political. A better understanding of social systems is necessary to create new mechanisms based on collective action and efforts to avoid future pandemics. It has been clear that inequality is an important factor that accelerates disease propagation and their impact making the system more fragile and implying latent systemic risk [86,140]. Super-propagation phenomena has been also shown as a systemic problem for pandemics and we do not have the tools to tackle them [141,142]. Epidemiological policy must act at different levels, from local aid to global governance mechanisms [143]. We still need to progress complex multi-scale systems for acquiring data, diagnostics and delivering aid in real-time.

Assessing measures to stop epidemics still present several epistemological, operational and political challenges. For instance, we should think in experimentation and simulations and isolate factors (pharmacological, non-pharmacological, social, political, economic, etc) to quantify how each measure contributes and also assess the synergies of integrated policy [136,137]. Next challenge is a better international system to control epidemics which implies not only regulatory issues but more technology to be prepared to stop future epidemics [143-145]. Cross-disciplinary research is also key for better policies [105], including crossing molecular and clinical research with technological innovation [6]. Innovation in communication channels and novel dashboards is an important area of research [146]. Artificial Intelligence has to exploit the upcoming 5th Industrial Revolution and 2nd Data Revolution to design better systems for response and decision making in all layers of the society including policy makers and citizens [105] and deliver Collective Intelligence platforms to empower people and amplify collective efforts [147,148].

Ethical issues have arisen because of the exhaustive use of technology in some countries during the SARS-CoV-2 pandemic [145,149,150]. Privacy and rights have suffered an important debate because there was not sufficient reasoning on these topics [143]. Ethical frameworks based on principles are necessary to leverage technologies for sustainable development and emergencies [149]. We still must work on implementing epidemiological policies, technology and mechanisms to fight future epidemics and progress towards a more sustainable and resilient society [151].

REFERENCES

1. Green MD, Freedman DM, Gordis L. Reference guide on epidemiology. *Reference Manual on Scientific Evidence*. 2000;2:638.
2. Salathe M, Bengtsson L, Bodnar TJ, et al. Digital epidemiology. *PLoS computational biology*. 2012;8(7)
3. Wormser GP, Pourbohloul B. Modeling Infectious Diseases in Humans and Animals By Matthew James Keeling and Pejman Rohani Princeton, NJ: Princeton University Press, 2008. 408 pp., Illustrated. $65.00 (hardcover). The University of Chicago Press; 2008.
4. Anderson RM, Anderson B, May RM. *Infectious diseases of humans: dynamics and control*. Oxford university press; 1992.
5. Bell G, Hey T, Szalay A. Beyond the data deluge. *Science*. 2009;323(5919):1297-1298.
6. Bullock J, Pham KH, Lam CSN, Luengo-Oroz M. Mapping the landscape of artificial intelligence applications against COVID-19. *arXiv preprint arXiv:200311336*. 2020;
7. LeCun Y, Bengio Y, Hinton G. Deep learning. *nature*. 2015;521(7553):436-444.
8. Kumar A, Gupta PK, Srivastava A. A review of modern technologies for tackling COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*. 2020;14(4):569-573.
9. Shorten C, Khoshgoftaar TM, Furht B. Deep Learning applications for COVID-19. *Journal of big Data*. 2021;8(1):1-54.
10. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of Epidemiology & Community Health*. 2006;60(7):578-586.
11. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *Journal of medical Internet research*. 2009;11(1):e11.
12. Cervellin G, Comelli I, Lippi G. Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. *Journal of epidemiology and global health*. 2017;7(3):185-189.
13. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *The New England journal of medicine*. 2009;360(21):2153.
14. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457(7232):1012-1014.
15. Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. *Science*. 2014;343(6176):1203-1205.
16. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS computational biology*. 2013;9(10)
17. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PloS one*. 2011;6(8)
18. Salathé M. Digital epidemiology: what is it, and where is it going? *Life sciences, society and policy*. 2018;14(1):1.
19. Vespignani A, Tian H, Dye C, et al. Modelling COVID-19. *Nature Reviews Physics*. 2020:1-3.
20. Aleta A, Martín-Corral D, y Piontti AP, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*. 2020:1-8.
21. Espana G, Cavany S, Oidtman RJ, et al. Impacts of K-12 school reopening on the COVID-19 epidemic in Indiana, USA. *medRxiv*. 2020;
22. Moreno López JA, Arregui García B, Bentkowski P, et al. Anatomy of digital contact tracing: Role of age, transmission setting, adoption, and case detection. *Science advances*. 2021;7(15):eabd8750.
23. Hinch R, Probert WJ, Nurtay A, et al. OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLoS computational biology*. 2021;17(7):e1009146.
24. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Physical review letters*. 2001;86(14):3200.
25. Moreno Y, Pastor-Satorras R, Vespignani A. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*. 2002;26(4):521-529.
26. de Arruda GF, Petri G, Rodrigues FA, Moreno Y. Impact of the distribution of recovery rates on disease spreading in complex networks. *Physical Review Research*. 2020;2(1):013046.
27. Pastor-Satorras R, Vespignani A. Epidemic dynamics in finite size scale-free networks. *Physical Review E*. 2002;65(3):035108.
28. Pastor-Satorras R, Vespignani A. Epidemic dynamics and endemic states in complex networks. *Physical Review E*. 2001;63(6):066117.

29. May RM, Lloyd AL. Infection dynamics on scale-free networks. *Physical Review E*. 2001;64(6):066112.
30. Stauffer D, Aharony A. *Introduction to percolation theory*. Taylor & Francis; 2018.
31. Newman ME. Spread of epidemic disease on networks. *Physical review E*. 2002;66(1):016128.
32. Meyers LA, Newman M, Pourbohloul B. Predicting epidemics on directed contact networks. *Journal of theoretical biology*. 2006;240(3):400-418.
33. McMahon T, Chan A, Havlin S, Gallos LK. Spatial correlations in geographical spreading of COVID-19 in the United States. *Scientific reports*. 2022;12(1):1-10.
34. Colizza V, Vespignani A. Invasion threshold in heterogeneous metapopulation networks. *Physical review letters*. 2007;99(14):148701.
35. de Arruda GF, Petri G, Moreno Y. Social contagion models on hypergraphs. *Physical Review Research*. 2020;2(2):023032.
36. Colizza V, Barrat A, Barthelemy M, Valleron A-J, Vespignani A. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS medicine*. 2007;4(1)
37. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*. 2020;368(6489):395-400.
38. Zhang J, Litvinova M, Liang Y, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*. 2020;
39. Martín-Calvo D, Aleta A, Pentland A, Moreno Y, Moro E. *Effectiveness of social distancing strategies for protecting a community from a pandemic with a data driven contact network based on census and real-world mobility data*. 2020.
40. Ferretti L, Wymant C, Kendall M, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. 2020;368(6491)
41. Boy J, Pastor-Escuredo D, Macguire D, Jimenez RM, Luengo-Oroz M. Towards an understanding of refugee segregation, isolation, homophily and ultimately integration in Turkey using call detail records. *Guide to Mobile Data Analytics in Refugee Scenarios*. Springer; 2019:141-164.
42. Sanz J, Xia C-Y, Meloni S, Moreno Y. Dynamics of interacting diseases. *Physical Review X*. 2014;4(4):041005.
43. Prothero RM. Disease and mobility: a neglected factor in epidemiology. *International journal of epidemiology*. 1977;6(3):259-267.
44. Gonzalez MC, Hidalgo CA, Barabasi A-L. Understanding individual human mobility patterns. *nature*. 2008;453(7196):779-782.
45. Riley S. Large-scale spatial-transmission models of infectious disease. *Science*. 2007;316(5829):1298-1301.
46. Eubank S, Guclu H, Kumar VA, et al. Modelling disease outbreaks in realistic urban social networks. *Nature*. 2004;429(6988):180-184.
47. Candia J, González MC, Wang P, Schoenharl T, Madey G, Barabási A-L. Uncovering individual and collective human dynamics from mobile phone records. *Journal of physics A: mathematical and theoretical*. 2008;41(22):224015.
48. Zufiria PJ, Pastor-Escuredo D, Úbeda-Medina L, et al. Identifying seasonal mobility profiles from anonymized and aggregated mobile phone data. Application in food security. *PloS one*. 2018;13(4):e0195714.
49. Pastor-Escuredo D, Frias-Martinez E. Flow descriptors of human mobility networks. *arXiv preprint arXiv:200307279*. 2020;
50. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*. 2009;106(51):21484-21489.
51. Chang S, Pierson E, Koh PW, et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*. 2021;589(7840):82-87.
52. Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, Vespignani A. Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of computational science*. 2010;1(3):132-145.
53. Balcan D, Hu H, Goncalves B, et al. Seasonal transmission potential and activity peaks of the new influenza A (H1N1): a Monte Carlo likelihood analysis based on human mobility. *BMC medicine*. 2009;7(1):45.
54. Simini F, González MC, Maritan A, Barabási A-L. A universal model for mobility and migration patterns. *Nature*. 2012;484(7392):96.
55. Perrota D. Can Mobile Phone Traces Help Shed Light on the Spread of Zika in Colombia? UNGP. 2020. 2018. https://www.unglobalpulse.org/2018/04/can-mobile-phone-traces-help-shed-light-on-the-spread-of-zika-in-colombia/
56. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, et al. The role of human movement in the transmission of vector-borne pathogens. *PLoS neglected tropical diseases*. 2009;3(7)

57. Lynch C, Roper C. The transit phase of migration: circulation of malaria and its multidrug-resistant forms in Africa. *PLoS medicine*. 2011;8(5)
58. Buscarino A, Fortuna L, Frasca M, Latora V. Disease spreading in populations of moving agents. *EPL (Europhysics Letters)*. 2008;82(3):38002.
59. Meloni S, Perra N, Arenas A, Gómez S, Moreno Y, Vespignani A. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific reports*. 2011;1:62.
60. Blondel VD, Decuyper A, Krings G. A survey of results on mobile phone datasets analysis. *EPJ data science*. 2015;4(1):10.
61. Pulse UG. Big data for development: Challenges & opportunities. *Naciones Unidas, Nueva York, mayo*. 2012;
62. Pulse UG. Mapping the Risk-Utility Landscape: Mobile Data for Sustainable Development and Humanitarian Action. *Global Pulse Project Series no18*. 2015;
63. De Montjoye Y-A, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*. 2013;3:1376.
64. Grantz KH, Meredith HR, Cummings DA, et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nature communications*. 2020;11(1):1-8.
65. Bejon P, Williams TN, Liljander A, et al. Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. *PLoS medicine*. 2010;7(7)
66. Dolgin E. Targeting hotspots of transmission promises to reduce malaria. Nature Publishing Group; 2010.
67. Wesolowski A, Eagle N, Tatem AJ, et al. Quantifying the impact of human mobility on malaria. *Science*. 2012;338(6104):267-270.
68. Bengtsson L, Lu X, Thorson A, Garfield R, Von Schreeb J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. *PLoS medicine*. 2011;8(8)
69. Wesolowski A, Qureshi T, Boni MF, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proceedings of the National Academy of Sciences*. 2015;112(38):11887-11892.
70. Bengtsson L, Gaudart J, Lu X, et al. Using mobile phone data to predict the spatial spread of cholera. *Scientific reports*. 2015;5:8923.
71. Tizzoni M, Bajardi P, Decuyper A, et al. On the use of human mobility proxies for modeling epidemics. *PLoS computational biology*. 2014;10(7)
72. Tatem AJ, Huang Z, Narib C, et al. Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria journal*. 2014;13(1):52.
73. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. American Association for the Advancement of Science; 2020. p. eabc0764.
74. Adam DC, Wu P, Wong JY, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nature Medicine*. 2020;26(11):1714-1719.
75. Finger F, Genolet T, Mari L, et al. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proceedings of the National Academy of Sciences*. 2016;113(23):6421-6426.
76. UNICEF. Magic Box COVID-19 report. UNICEF. 2020. https://www.unicef.org/innovation/magicbox/covid
77. Dueñas M, Campi M, Olmos L. Changes in mobility and socioeconomic conditions in Bogotá city during the COVID-19 outbreak. *arXiv preprint arXiv:200811850*. 2020;
78. Gozzi N, Tizzoni M, Chinazzi M, Ferres L, Vespignani A, Perra N. Estimating the effect of social inequalities on the mitigation of COVID-19 across communities in Santiago de Chile. *Nature communications*. 2021;12(1):1-9.
79. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science*. 2021;372(6545):eabg5298.
80. Belderok S-M, Rimmelzwaan GF, Van Den Hoek A, Sonder GJ. Effect of travel on influenza epidemiology. *Emerging infectious diseases*. 2013;19(6):925.
81. Wesolowski A, Stresman G, Eagle N, et al. Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Scientific reports*. 2014;4:5678.
82. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. American Association for the Advancement of Science; 2020.
83. Ventura PC, Aleta A, Rodrigues FA, Moreno Y. Modeling the effects of social distancing on the large-scale spreading of diseases. *Epidemics*. 2022;38:100544.
84. Taleb NN. The Statistical Consequences of Fat Tails. STEM Publishing; 2019.
85. Taleb NN, Bar-Yam Y, Douady R, Norman J, Read R. The precautionary principle: fragility and black swans from policy actions. *NYU Extreme Risk Initiative Working Paper*. 2014:1-24.

86. Norman J, Bar-Yam Y, Taleb NN. Systemic Risk of Pandemic via Novel Pathogens—Coronavirus: A Note. *New England Complex Systems Institute (January 26, 2020)*. 2020;
87. Greenwood F, Howarth C, Escudero Poole D, Raymond NA, Scarnecchia DP. The signal code: A human rights approach to information during crisis. *Harvard, MA*. 2017;
88. General US. Good communication saves lives. 2020. 2020. https://www.un.org/en/coronavirus/good-communication-saves-lives
89. Lazer DM, Baum MA, Benkler Y, et al. The science of fake news. *Science*. 2018;359(6380):1094-1096.
90. Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*. 2017;19(1):22-36.
91. Bakir V, McStay A. Fake news and the economy of emotions: Problems, causes, solutions. *Digital journalism*. 2018;6(2):154-175.
92. Zarocostas J. How to fight an infodemic. *The Lancet*. 2020;395(10225):676.
93. Vaezi A, Javanmard SH. Infodemic and risk communication in the era of CoV-19. *Advanced Biomedical Research*. 2020;9
94. Hua J, Shaw R. Corona virus (Covid-19)"infodemic" and emerging issues through a data lens: The case of china. *International journal of environmental research and public health*. 2020;17(7):2309.
95. Nekovee M, Moreno Y, Bianconi G, Marsili M. Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*. 2007;374(1):457-470.
96. Miritello G, Moro E, Lara R. Dynamical strength of social ties in information spreading. *Physical Review E*. 2011;83(4):045102.
97. Morales AJ, Borondo J, Losada JC, Benito RM. Efficiency of human activity on information spreading on Twitter. *Social Networks*. 2014;39:1-11.
98. Bodendorf F, Kaiser C. Detecting opinion leaders and trends in online social networks. 2009:65-68.
99. Pastor-Escuredo D, Tarazona C. Characterizing information leaders in Twitter during COVID-19 crisis. *arXiv preprint arXiv:200507266*. 2020;
100. Newman ME, Barabási A-LE, Watts DJ. *The structure and dynamics of networks*. Princeton university press; 2006.
101. Borge-Holthoefer J, Meloni S, Gonçalves B, Moreno Y. Emergence of influential spreaders in modified rumor models. *Journal of Statistical Physics*. 2013;151(1-2):383-393.
102. Popat K, Mukherjee S, Yates A, Weikum G. DeClarE: Debunking fake news and false claims using evidence-aware deep learning. *arXiv preprint arXiv:180906416*. 2018;
103. Ruchansky N, Seo S, Liu Y. Csi: A hybrid deep model for fake news detection. 2017:797-806.
104. Singhania S, Fernandez N, Rao S. 3han: A deep neural network for fake news detection. Springer; 2017:572-581.
105. Luengo-Oroz M, Pham KH, Bullock J, et al. Artificial intelligence cooperation to support the global response to COVID-19. *Nature Machine Intelligence*. 2020:1-3.
106. Leung GM, Leung K. Crowdsourcing data to mitigate epidemics. *The Lancet Digital Health*. 2020;2(4):e156-e157.
107. Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health*. 2020;
108. Imran A, Posokhova I, Qureshi HN, et al. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Informatics in Medicine Unlocked*. 2020:100378.
109. Menni C, Valdes A, Freydin MB, et al. Loss of smell and taste in combination with other symptoms is a strong predictor of COVID-19 infection. *MedRxiv*. 2020;
110. Carrillo-Larco RM, Castillo-Cara M. Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach. *Wellcome Open Research*. 2020;5(56):56.
111. Hu Z, Ge Q, Jin L, Xiong M. Artificial intelligence forecasting of covid-19 in china. *arXiv preprint arXiv:200207112*. 2020;
112. Hartono P. Generating Similarity Map for COVID-19 Transmission Dynamics with Topological Autoencoder. *arXiv preprint arXiv:200401481*. 2020;
113. Liu D, Clemente L, Poirier C, et al. A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models. *arXiv preprint arXiv:200404019*. 2020;
114. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*. 2005;67(2):301-320.
115. Bandyopadhyay SK, Dutta S. Machine learning approach for confirmation of covid-19

cases: Positive, negative, death and release. *medRxiv*. 2020;

116. Huang C-J, Chen Y-H, Ma Y, Kuo P-H. Multiple-Input Deep Convolutional Neural Network Model for COVID-19 Forecasting in China. *medRxiv*. 2020;

117. Fong SJ, Li G, Dey N, Crespo RG, Herrera-Viedma E. Finding an accurate early forecasting model from small dataset: A case of 2019-ncov novel coronavirus outbreak. *arXiv preprint arXiv:200310776*. 2020;

118. Lu FS, Hattab MW, Clemente CL, Biggerstaff M, Santillana M. Improved state-level influenza nowcasting in the United States leveraging Internet-based data and network approaches. *Nature communications*. 2019;10(1):1-10.

119. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proceedings of the National Academy of Sciences*. 2015;112(47):14473-14478.

120. Rajarethinam J, Ong J, Lim S-H, et al. Using human movement data to identify potential areas of Zika transmission: case study of the largest Zika cluster in Singapore. *International journal of environmental research and public health*. 2019;16(5):808.

121. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*. 2020;25(10):2000180.

122. Moghadas SM, Fitzpatrick MC, Sah P, et al. The implications of silent transmission for the control of COVID-19 outbreaks. *Proceedings of the National Academy of Sciences*. 2020;117(30):17513-17515.

123. Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A. The role of biomarkers in diagnosis of COVID-19–A systematic review. *Life Sciences*. 2020:117788.

124. Shi Y, Wang Y, Shao C, et al. COVID-19 infection: the perspectives on immune responses. Nature Publishing Group; 2020.

125. Nicholas GD, Petra K, Yang L, et al. Age-dependent Effects in the Transmission and Control of COVID-19 Epidemics. *Nature medicine*. 2020;

126. Eames KT, Keeling MJ. Contact tracing and disease control. *Proceedings of the Royal Society of London Series B: Biological Sciences*. 2003;270(1533):2565-2571.

127. Dietz K. The estimation of the basic reproduction number for infectious diseases. *Statistical methods in medical research*. 1993;2(1):23-41.

128. Liu Q-H, Ajelli M, Aleta A, Merler S, Moreno Y, Vespignani A. Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of the National Academy of Sciences*. 2018;115(50):12680-12685.

129. Glass RJ, Glass LM, Beyeler WE, Min HJ. Targeted social distancing designs for pandemic influenza. *Emerging infectious diseases*. 2006;12(11):1671.

130. Lampos V, Moura S, Yom-Tov E, Cox IJ, McKendry R, Edelstein M. Tracking COVID-19 using online search. *arXiv preprint arXiv:200308086*. 2020;

131. WHO. Infodemic management - Infodemiology;. WHO. 2020. https://www.who.int/teams/risk-communication/infodemic-management.

132. Singh L, Bansal S, Bode L, et al. A first look at COVID-19 information and misinformation sharing on Twitter. *arXiv preprint arXiv:200313907*. 2020;

133. Mejova Y, Kalimeri K. Advertisers jump on coronavirus bandwagon: Politics, news, and business. *arXiv preprint arXiv:200300923*. 2020;

134. Hu Z, Ge Q, Li S, Boerwincle E, Jin L, Xiong M. Forecasting and evaluating intervention of Covid-19 in the World. *arXiv preprint arXiv:200309800*. 2020;

135. Dandekar R, Barbastathis G. Neural Network aided quarantine control model estimation of global Covid-19 spread. *arXiv preprint arXiv:200402752*. 2020;

136. Lai S, Ruktanonchai NW, Zhou L, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. 2020;

137. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. 2020:1-5.

138. Davies NG, Kucharski AJ, Eggo RM, 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*. 2020;

139. Cowling BJ, Ali ST, Ng TW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *The Lancet Public Health*. 2020;

140. Taleb NN, Read R, Douady R, Norman J, Bar-Yam Y. The precautionary principle (with application to the genetic modification of organisms). *arXiv preprint arXiv:14105787*. 2014;

141. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature*. 2005;438(7066):355-359.
142. Kupferschmidt K. Why do some COVID-19 patients infect many others, whereas most don't spread the virus at all? *Science*. 19/05/2020 2020;
143. Mello MM, Wang CJ. Ethics and governance for digital disease surveillance. *Science*. 2020;
144. Kuhn C, Beck M, Strufe T. Covid Notions: Towards Formal Definitions--and Documented Understanding--of Privacy Goals and Claimed Protection in Proximity-Tracing Services. *arXiv preprint arXiv:200407723*. 2020;
145. Vinuesa R, Theodorou A, Battaglini M, Dignum V. A socio-technical framework for digital contact tracing. *arXiv preprint arXiv:200508370*. 2020;
146. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*. 2020;20(5):533-534.
147. Aleks Berditchevskaia KP. Coronavirus: seven ways collective intelligence is tackling the pandemic. The Conversation. 2020. 2020. https://theconversation.com/coronavirus-seven-ways-collective-intelligence-is-tackling-the-pandemic-133553
148. Pastor-Escuredo D, Tarazona-Lizarraga C, Bachmann A, Treleaven P. Collective Intelligence and Governance for Pandemics. *Covid-19: A Complex Systems Approach*. STEM Academic Press; 2021:223-248.
149. Pastor-Escuredo D. Ethics in the digital era. *arXiv preprint arXiv:200306530*. 2020;
150. Berman G, Carter K, Herranz MG, Sekara V. *Digital contact tracing and surveillance during COVID-19. General and child-specific ethical issues.* 2020.
151. General US. The age of digital interdependence. Report of the UN Secretary-General's High-Level Panel on Digital Cooperation ...; 2019.