Containing Future Epidemics With Trustworthy Federated Systems for Ubiquitous Warning and Response

Dick Carrillo [1,2]*, Lam Duc Nguyen [3], Pedro H. J. Nardelli [1], Evangelos Pournaras [4], Plinio Morita [5], Demóstenes Z. Rodríguez [6], Merim Dzaferagic [7], Harun Siljak [7], Alexander Jung [8], Laurent Hébert-Dufresne [9], Irene Macaluso [6], Mehar Ullah [1], Gustavo Fraidenraich [2] and Petar Popovski [3]

[1] School of Energy Systems, Lappeenranta-Lahti University of Technology, Lappeenranta, Finland, [2] Department of Communications, University of Campinas, Campinas, Brazil, [3] Department of Electronic Systems, Aalborg University, Aalborg, Denmark, [4] School of Computing, University of Leeds, Leeds, United Kingdom, [5] School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada, [6] Department of Computer Science, Federal University of Lavras, Minas Gerais, Brazil, [7] Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin, Ireland, [8] Department of Computer Science, Aalto University, Espoo, Finland, [9] College of Engineering and Mathematical Sciences, University of Vermont, Burlington, VT, United States

In this paper, we propose a global digital platform to avoid and combat epidemics by providing relevant real-time information to support selective lockdowns. It leverages the pervasiveness of wireless connectivity while being trustworthy and secure. The proposed system is conceptualized to be decentralized yet federated, based on ubiquitous public systems and active citizen participation. Its foundations lie on the principle of informational self-determination. We argue that only in this way it can become a trustworthy and legitimate public good infrastructure for citizens by balancing the asymmetry of the different hierarchical levels within the federated organization while providing highly effective detection and guiding mitigation measures toward graceful lockdown of the society. To exemplify the proposed system, we choose a remote patient monitoring as use case. This use case is evaluated considering different numbers of endorsed peers on a solution that is based on the integration of distributed ledger technologies and NB-IoT (narrowband IoT). An experimental setup is used to evaluate the performance of this integration, in which the end-to-end latency is slightly increased when a new endorsed element is added. However, the system reliability, privacy, and interoperability are guaranteed. In this sense, we expect active participation of empowered citizens to supplement the more usual top-down management of epidemics.

Keywords: COVID-19, epidemiological warning system, wireless, privacy-preserving data collection, communication networks

1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic has clearly shown that, in many senses, the world as a whole was not prepared for dealing with a disease of such magnitude. Although a final assessment is infeasible at this point (October, 2020), in which a second wave is creeping back in Europe and is poised to rage across the continent by fall 2020 (Cacciapaglia et al., 2020). The current statistics available about COVID-19 indicate that the most successful policies for monitoring and controlling the virus propagation are employing various digital technologies and connectivity (Ienca and Vayena, 2020; Ting et al., 2020). These technologies can be used in two principal ways: (1) to provide status and predictions of the epidemiological spread and facilitate

actions, such as administration of diagnostic tests or preparation of medical equipment; (2) to implement active policies that facilitate societal processes and safe citizen movement, such as shop schedules, organized delivery of goods, etc. The success of such informed policies is also strongly related to the *timing of their implementation* for the respective place, as virus contagion is a spatiotemporal phenomenon that usually leads to exponential growth in the number of infected individuals (Adam, 2020).

Unfortunately, even the most successful timely interventions are in some ways overreaching, shutting down abruptly most of the economic activities and creating social distancing between citizens. While acknowledging that this is our current societal need to contain the propagation of this pandemic, our main hypothesis is that, in the near future, digital technologies and wireless connectivity can enable a *graceful lockdown* by having the following main roles:

- Facilitate selective lockdowns by allowing flexible transfer of the work, education and play activities online to the desired extent.
- Assist in maintaining physical distancing and tracking the spreading of the disease, while offering digital tools to control the level of lockdown/reopening.
- Provide real-time information for targeted and efficient testing.

Note that these are also important today (October, 2020), when a second wave is forcing several countries to rethink a step-by-step reopening based on the aforementioned points.

This will be achieved based on an integrated solution for the ubiquitous and privacy-preserving data acquisition, explainable predictive methods for contagion risk assessment, and digitized policies for selective lockdown and scheduling of societal activities. Data will be locally acquired *via* online social networks, personal smart phone devices, and Internet-of-Things (IoT) sensors in general. It is foreseen that an effective global solution will be decentralized, yet hierarchical (system-of-systems), to provide fine-tuned coordination among the federated entities to orchestrate actions and then avoid both the human rights cost of overly stringent solutions (Kupferschmidt and Cohen, 2020) and the human life cost of "hands-off" approaches. Beyond more traditional centralized policy-making, such as a decentralized platform, which will be built using trustworthy distributed ledger technology (DLT), should also provide citizens with tools for more direct participation in mitigation measures based on tailored individual incentives.

In this sense, we expect to bring managerial decisions closer to the citizens. With such a federated governance, we attempt to convert the usually passive data acquisition process to an active participatory one. Hence, our aim is to move away from a purely top-down management approach. In addition, the proposed platform aims at achieving the minimal use of personal data, processed with the maximum possible privacy-preservation to eliminate global health risks, while protecting the economy (Asikis and Pournaras, 2020). The COVID-19 pandemic has demonstrated that nations and governments need to be better prepared and coordinated to detect and react to global threats focusing on *early detection, early response,*

moving beyond traditionally passive disease monitoring based on voluntary reporting systems. Hence, a rapid, participatory responsive detection system must be in place to support public health officials. A synergistic relationship with public health officials, policymakers, and citizens' active participation will be critical to align the mandate for public health with the protection of privacy, freedom, and democracy (Helbing et al., 2019; Amat et al., 2020). In this context, one main factor is to design a special set of incentives that would allow the citizens to provide secured anonymized access to their data while actively participating in the crowd platform to support early disease detection, a public information system, and possible mitigation measures. Building up such an incentive structure that maps those trade-offs is a key aspect (beyond purely technical ones) for the success of the proposed federated architecture. Furthermore, this system should be constructed and evaluated rigorously following ethical guidelines as in Ienca and Vayena (2020) and Hagendorff (2020).

Our contribution in this paper is 3-fold:

1. A comprehensive analysis of epidemiological models, data collection, and wireless connectivity is done based on key relevant scientific references on pandemics.
2. A federated global epidemiological warning system is proposed based on DLTs.
3. An experimental setup of the integration between DLT and NB-IoT is used to evaluate the wireless network performance on the IoT infrastructure supporting a remote patient monitoring use case.

The rest of the paper is organized as follows: section 2 presents a discussion of epidemiological models and their limitations. section 3 describes the relevance of wireless connectivity on pandemic scenarios. In section 4, the federated global epidemiological warning system is proposed, the experimental setup of the integration of DLT and narrowband IoT (NB-IoT) is detailed. Section 5 is reserved for some discussions and future perspectives.

2. EPIDEMIOLOGICAL MODELS AND DATA COLLECTION

Biomedical data alone do not contain the information required for preventing and mitigating pandemics. In a highly interconnected globalized society, social interactions, environmental data, spatiotemporal events, and collective non-linear phenomena observed on complex infrastructures such as transport systems, require the mining of heterogeneous pandemic big data that are a result of a complex system of systems processes. In fact, modern disease monitoring systems, as the Global Public Health Intelligence Network (GPHIN) (Mykhalovskiy and Weir, 2006) and the Global Outbreak Alert and Response Network (GOARN) (Davies and Youde, 2016), and event-based approaches that use a combination of web crawlers, artificial intelligence (AI), and public health expertise in the detection of indicators (Bernard et al., 2018; Carter et al., 2018) provide good—although limited—examples of successful platforms to identify latent indicators of an outbreak.

To go beyond those existing solutions, data scale, contextualization, and granularity are key requirements for data quality, which is usually orthogonal to privacy preservation, timely processing and analysis as well as storage and processing cost. In the following section, more details will be provided about existing epidemiological models and epidemic-related data collection.

2.1. Models and Their Limitations

Models for emerging epidemics come in various forms, all relying on differently coarse-grained individuals' data. State-of-the-art population-level models for COVID-19 rely on metapopulation models with an underlying Susceptible-Exposed-Infectious-Recovered (SEIR) disease dynamics (Chinazzi et al., 2020). The metapopulations represent regions surrounding travel hubs and are interconnected through mobility data, for example, data from airlines, public transportation systems, and traffic control systems. The disease dynamics itself accounts for the fact that exposed individuals can be undergoing an incubation period, or even presymptomatic and unaware that they could infect others, as suggested by early case data (Wei, 2020). Note that some models opt for a simpler formulation (Zhang et al., 2020), the Susceptible-Infectious-Recovered (SIR) dynamics which ignore the exposed state and therefore eliminate a potentially important delay in disease progression.

Unfortunately, while SEIR and SIR models can be effective at providing large-scale forecasts and predicting importations at aggregated levels, they are less informative at the individual level, where we expect significant heterogeneity across individuals; both in behavior and disease manifestation (Althouse et al., 2020). One key assumption of these models is that all individuals of a given state follow the same dynamics (where state can be, for example, a combination of age and epidemiological status). However, the importance of superspreading events have highlighted that pandemics are often not driven by the average individual events but by rare, stochastic events, where a synergy between individual behavior and disease manifestation can significantly spread the disease (Wang et al., 2020).

Models informed by data at the individual level attempt to account for stochastic events but are less used for forecasting purposes as the data required are seldom available and hard to connect across multiple scales or populations (Hébert-Dufresne et al., 2020). However, individual contact data can inform network models which are powerful for other purposes, for instance, informing or measuring the impact of interventions and policies, such as school closures and lockdowns (Zhang et al., 2020). Likewise, these network models can be instrumental in guiding real-time contact tracing efforts during epidemics (Firth et al., 2020; Kojaku et al., 2020). One key limitation of this model is that they are often not spatially or temporally explicit. While this is not problematic when they are used to guide local efforts and forecasts over a short period, this limitation precludes their integration with spatial and temporal forecasts from metapopulation models.

Altogether, an integrated modeling framework requires accurate estimation of mobility across regions to inform metapopulation model, as well as individual level data on heterogeneity in contact patterns to guide local interventions. Importantly, these two scales need not be connected in any identifiable way. It is not critical to know the local behavior of individuals connecting different populations. Multi-scale data on the same population, with anonymity assured across scales, would itself be of invaluable improvement to the current reliance on disparate (and often private) data sources on different populations. Doing so could help guide more flexible control policies, beyond blanket measures that ignore the often unknown heterogeneity and adaptive nature of human behavior.

More information is therefore necessary to accurately estimate the probability of epidemics given importation event, forecast their eventual spread, and guide control policies. Any particular importation event has the potential to become the initial point of a widespread contagion in a region not reached by epidemics until that point. Since this event basically depends on one individual, aggregated level statistics provide poor description of the heterogeneity of individual level behavior and mobility. Given this uncertainty and the lack of multi-scale real-time informative system, most governments around the world introduced stay-at-home policies to reduce the mobility of their citizens, with the now famous goal of "flattening of the curve." This expression refers to the process of slowing down the virus spread to keep the need for hospital care below the health care capacity. Different variations of the stay-at-home policy have been introduced, e.g., allowing citizens to move in a predefined radius to walk with their pets and exercise and restricting the number of times per day or the number of people per household that can go out. As previously argued, this is, understandably, a necessary abrupt solution, but one that could be improved by properly utilizing crowd-sourced data and active citizens' participation. Multi-scale data gathering could inform more flexible restriction while accounting for local heterogeneity in behavior and mobility patterns. We argue that this important data collection efforts could be accomplished by trustworthy federated systems.

2.2. Data Collection: Current Activities and Potential Existing Sources

At higher levels, mining existing data sets collected by different types of service providers is an obvious and powerful way to quantify to which extent citizens follow *stay-at-home* policies, thereby providing an essential means of assessing their effectiveness. For example, smart grid systems can provide information about the changes in the energy consumption in different areas suggesting that people, on average, spend more time at home. The network usage information from broadband operators can also be used as an indicator of change in human behavior related to their mobility. Another source of information are mobility data sets from cellular network operators that are already present on the network. Smart appliances, like smart TVs, smart fridges, and smart light bulbs can be used to provide information about the overall time spent at home vs. the time spent outside. Finally, information collected by several smart city applications, such as public transport usage and traffic, could also be leveraged to estimate the change in mobility patterns at the level of places, buildings, and vehicles.

At a lower granularity, proximity-tracing platforms have been proposed to collect data at the individual level, for instance, the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) , often in partnership with mobile service providers and operators. By aggregating proximity data over time, these approaches can follow data protection regulations. Among others, two of the largest big tech companies Apple and Google proposed a joint solution that appears to focus on the protection of user privacy, for example, by keeping people anonymous in the central servers and making data submission voluntary. Different technologies are being proposed to automate and extend contact tracing (e.g., shared databases, GPS traces of confirmed cases, and contact tracing through Bluetooth). The use of Bluetooth-based solutions, either in the form of the third-party dedicated apps or as a feature built into the mobile devices' operating systems, seems to be the most promising solution to date. Although these solutions are a new source of granular data, they cannot cover all the relevant forms of infection, namely, asymptomatic infections, and infections occurring out of proximity through shared surfaces (e.g., doorknobs). In this sense, our hope is to combine the metapopulation and individual tracing models in a citizen science framework to consider more possible routes of infection as well as both proximity (e.g. close contact) and mobility data (e.g., location visited). In addition to this, machine learning methods are expected to combine expert knowledge (e.g., from virologists and sociologists) with global data collected from different sources for fitting powerful predictive models to high-dimensional data. In doing so, the target is to leverage the different straights of different data sources at different granularity for privacy protection and allow for more thorough probabilistic forecasts. Ultimately, this will allow for graceful lockdowns, tailored incentives for citizens' participation, and fine-tuned legitimate interventions.

2.3. Data Privacy and Trustworthiness

Fine-grained data collection approaches raise two major concerns as highlighted in Cho et al. (2020): (1) user privacy; and (2) trustworthiness of the shared information. Clearly, sharing location and health information with third-party entities can lead to the misuse of the data set, for instance, through unwanted advertising and health insurance implications. The impact of fake health information on the contact tracing system should be analyzed to assess its implications in terms of access to testing. All these applications rely on the premise that users will trust and adopt these applications. This brings us to an additional problem that could hamper these approaches. The adoption of new technologies is hard to predict, and it might introduce bias, for example, related to age, wealth, level of education, and state of country development. Citizens' participation in warning and response of epidemic outbreaks requires new incentives that reward the responsible use of citizens' personal data for protecting public health, while penalize and prevent citizens' profiling actions, manipulative nudging, and power misuse.

2.4. Challenges and the Way Forward

Designing new data fusion schemes tailored to fuel data analytics processes for prevention and mitigation of pandemics is an open research question: Which smartphone sensor data can model compliance of *stay-at-home* policies? How such models can be enhanced with social media, smart grid, or transport data? Pandemics require that we revisit and potentially reinvent how data should be managed and how systems should be designed to manage data in a more responsible way (Hagendorff, 2020). More specifically, discovering new ways to turn private sensitive data of citizens into public good, while preventing massive surveillance, profiling, and discriminatory analytics, becomes imperative. To pioneer this capability, new transparent and accountable data commons (Wright, 2019), for instance, personal data stores (Alessi et al., 2018; Chowdhury et al., 2018; Truong et al., 2019), to which citizens have incentives to contribute data, are required. The citizens retain the right for their data to be "forgotten" (Burkell, 2017). Ultimately, they have control and give their consent to how these data are used. These data commons are designed to maximize the use of techniques for privacy preservation (Vergara-Laurens et al., 2016; Asikis and Pournaras, 2020), for example, homomorphic encryption, differential privacy, anonymity, and obfuscation, while citizens are made aware of the privacy risks they experience.

3. WIRELESS CONNECTIVITY AND EPIDEMIC-RELEVANT INFORMATION

The previous section covered different aspects related to the data required to support a fine-grained epidemiological model. It is argued that a ubiquitous system can be employed to manage such data in a privacy-preserving manner. Here, we identify wireless connectivity and IoT devices as a way to collect data following those principles.

Wireless connectivity and IoT devices come in many models, but can, in general, be classified into (1) personal devices, such as mobile phones and earbuds; and (2) unattended connected machines and things, such as surveillance cameras and motion sensors. There are three principal sources of epidemic-relevant data acquired through wireless connectivity: (1) online social networks; (2) personal smartphone and mobile data; and (3) sensory and IoT devices.

3.1. Online Social Networks

In critical events, people tend to use online social networks to post comments about the emergency and learn from other users' comments. As a result, online social networks become a rich source of diverse information that could help to understand the main characteristics of the crises and their potential magnitude in the early stages. For instance, in Li et al. (2020), authors collected comments posted in the Chinese social media channel Weibo during the first month of the COVID-19 epidemic in Wuhan (China) to understand the evolution trend of the public opinion. A similar approach was used on Twitter in Alshaabi et al. (2020) and Rosa et al. (2020). Government entities can use this approach to give more attention to the needs of the public during the beginning of the epidemic and adjust their responses accordingly. Another form of data content, normally provided by explicit consent from the citizens, is the GPS location

data from personal devices. At the application level, the Twitter platform offers various user information, geographical location being one of the most important for a wide range of applications. However, only a few users enable their geolocation as public. To discover the user's geolocation (geographical region or exact geocoordinates), there are mainly two approaches called content-based and network-based algorithms. The first one uses textual contents from tweets, and the latter uses the connections and interactions between users. In Do et al. (2018), a neural network model based on multiview learning by combining knowledge from both user-generated content and network interaction is proposed to infer users' locations.

3.2. Personal Smartphone Devices

This source of information comes in many different forms, ranging from the use of metadata in cellular mobile networks for user localization up to the metadata associated with different applications, such as Twitter. The use of data[1] collected in mobile networks during the COVID-19 pandemic has received significant attention (Oliver et al., 2020). Agglomerated mobile network data can be used to verify if interventions, such as school closures, are effective, and help to understand the geographical spread of an epidemic. Different types of mobile network metadata are collected at different levels of the communication system, and they offer varying levels of information. At the lowest level, localization of mobile network users is possible by evaluating the strength of the wireless signal received at base stations. Depending on how many base stations are connected to the device, the location can be determined at the level of entire cells or down to a few meters using triangulation.

The resolution offered by cellular mobile localization maybe too coarse for detecting citizens' proximity and potential infections by face-to-face interactions. Nevertheless, even low-resolution location data can give insight into behavioral patterns (how much time is spent at home, office, and events) of individuals. Proximity detection can be enhanced by using the metadata from Bluetooth devices, such as beacons and discovery messages (e.g., Ferretti et al., 2020). Wi-Fi is also a technology that is used ubiquitously, and its metadata can provide information on user proximity. For example, proximity can be inferred by comparing the lists of access points that each device can see within a given time interval (Sapiezynski et al., 2017).

Fusing the metadata from these different sources, along with the context information (how many family members are at home and which events) is very relevant for monitoring the epidemic.

3.3. Sensory and IoT Devices

Many applications and devices become important tools to provide epidemic-relevant data. These can be related to surveillance/thermal cameras, drones, and even wearable devices. For example, surveillance cameras can be employed to count the number of people entering and exiting a specific area. In other scenarios, thermal cameras are used with specialized settings to focus on human skin temperature range and infrared spot sensors for individual temperature scanning. A similar application to identify individual or group activities in a given place is based on motion sensors as infrared lasers, or ultrasonic sensors. In this approach, the motion sensors are installed and deployed in a specific environment to recognize different human activities.

Other important source of Internet of Things (IoT) data are the smart healthcare devices, for instance, remote monitoring systems to check body temperature, which is a key sign in the support of homeostasis. Other popular applications are oxygen saturation monitoring based on beat oximetry, electrocardiogram monitoring with a specialized framework to estimate the heart rate, elderly monitoring using Doppler radar to identify risk movements of elderly people, sugar level monitoring, and blood pressure monitoring. These IoT devices provide information that is analyzed by the IoT data to support specific applications.

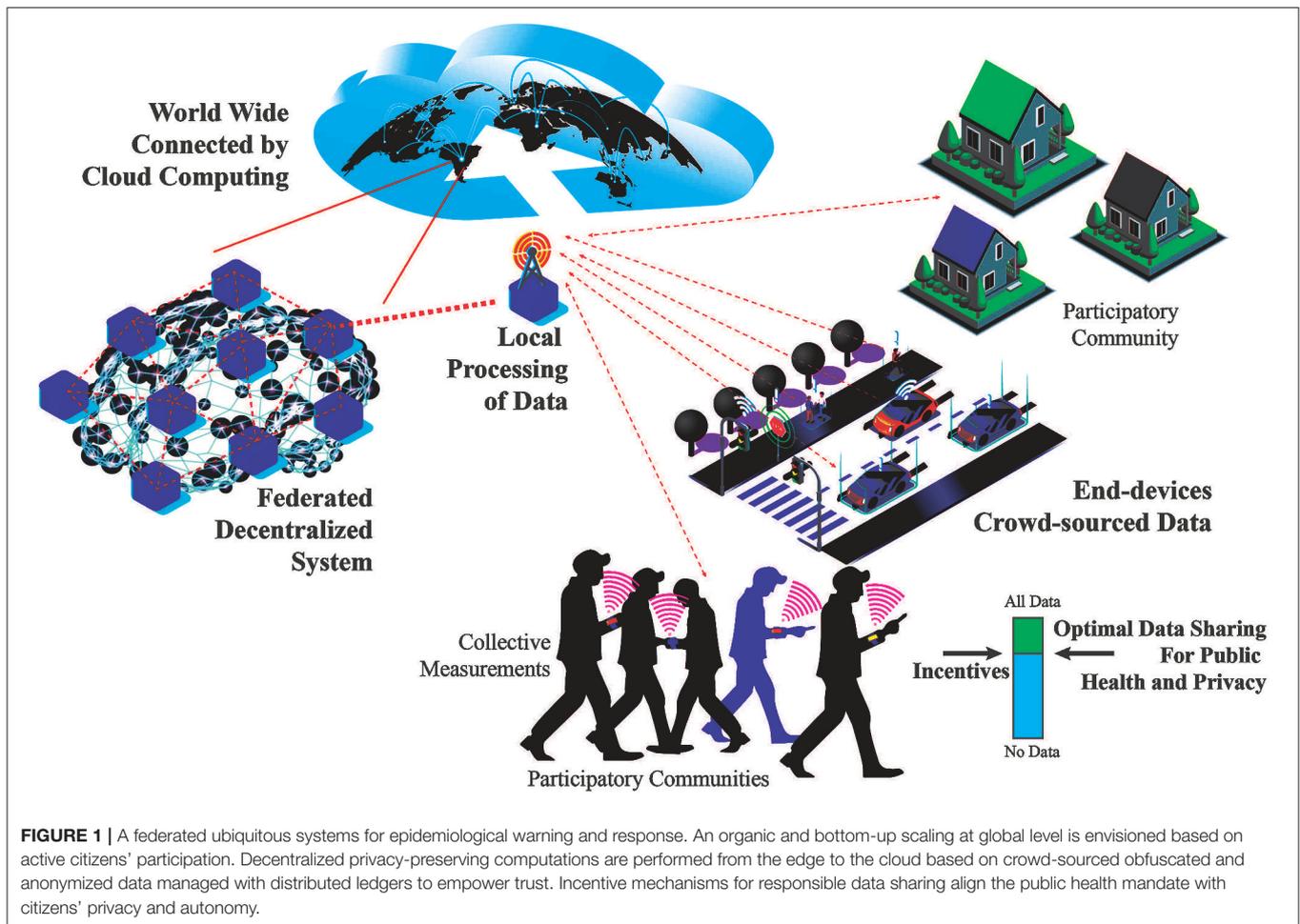# 4. FEDERATED GLOBAL EPIDEMIOLOGICAL WARNING SYSTEM

We envision a federated and decentralized coordination system for epidemiological warning and response. This system is federated by citizen communities that crowd-source data, personal smart phone devices, community-level IoT devices (e.g., LoRaWAN networks) and other computational resources on the edge. It can scale organically and bottom-up to city-level, national-level, and ultimately at global-level so as to coordinate in a socially responsible way the international actions of public health organizations and governments. This scaling requires tailored incentives that align public health policy-making with citizens' privacy and autonomy. Distributed ledgers with secure crypto-economic incentive models running on edge servers and scaling-up on-demand at a global level using wireless connectivity and public cloud infrastructure are the means to support the federated nature of this proposed system. **Figure 1** illustrates our vision.

4.1. Key Components

The proposed federated global epidemiological warning and response system is built upon epidemic-relevant information obtained through wireless connectivity introduced in section 3, and the internet infrastructure including fiber-optic and satellite links. Thus, in addition to the devices that are the data sources (e.g., smartphones, wearables, smart appliances, and cameras) with their applications and existing communication networks, there is a need for a computer infrastructure dedicated to store and process the epidemic-relevant data in a secure and privacy-preserving manner. It is argued that this infrastructure also needs to follow a similar federated organization based on the geographic locations. Each municipality, city, county, or larger neighborhood will rely on edge servers to process their respective data contents. The premise is to keep the computations as local as possible.

However, larger-scale computations related to the interrelations between locations (e.g., mobility, commuting,

---

[1] They are referred to as metadata due to the role played by it in the communication system to which it belongs; here, we will also refer to it as a "data" when there is no danger of ambiguity.

FIGURE 1 | A federated ubiquitous systems for epidemiological warning and response. An organic and bottom-up scaling at global level is envisioned based on active citizens' participation. Decentralized privacy-preserving computations are performed from the edge to the cloud based on crowd-sourced obfuscated and anonymized data managed with distributed ledgers to empower trust. Incentive mechanisms for responsible data sharing align the public health mandate with citizens' privacy and autonomy.

and traveling) require data from these different places. Once again, the federated organization supports this collaborative sharing (up and downstream between the federated entities), but it will probably require computationally more demanding algorithms. It is possible that those computational tasks require more powerful cloud servers, or collaborative parallel computing at the edges. For security reasons, DLTs could be employed to store data from the edge servers at the global level. At the technical level, the following aspects deserve special attention.

4.1.1. Privacy

In this paper, a privacy-by-design approach is proposed with which several critical operations of the federated global epidemiological system can be performed; for instance, decentralized data analytics (Bonawitz et al., 2017; Pournaras et al., 2017), social interactions analysis (Musciotto et al., 2016; Bahri et al., 2018), decentralized planning and resource allocation (Pournaras et al., 2018), and federated learning (Smith et al., 2017; Yang et al., 2019). Such operations integrate in a smart way state-of-the-art techniques, for instance, informational self-determination, homomorphic encryption, differential privacy, obfuscation, and anonymity. Often, privacy may limit the quality of service known as the *system utility*, as the accuracy and

quality of data are deteriorated to hide information content. For instance, predicting the risk of infections for individuals based on an epidemiological network model is a graph-based semi-supervised learning problem. Privacy-preserving semisupervised learning over graphs has been considered in Arai and Sakuma (2011). However, the precise trade-off between (lowering of) privacy protection and learning accuracy is largely open. Pareto optimal trade-offs can be configured and regulated by tuning the parameters of the privacy techniques as previously shown in Asikis and Pournaras (2020). Monetary and other incentives can be used to coordinate data sharing choices in a crowd. However, new types of social incentives are required for such an epidemiological system; for instance, incentives related to well-being, receiving solidarity, and long-term payoffs.

4.1.2. Interoperability

The big data required by the proposed solution have to be interoperable, i.e., the several applications that are providing data to be used to accomplish a specific task have to operate together and share a common "understanding" (Jacoby et al., 2017). This can be achieved by employing standards for health informatics such as ISO TS22220:2011, TS21090:2011, or TS13606. For instance, the data collected from different sources can be used to

predict refinements to patient care or new drug contraindication (Schulz et al., 2019). This key issue, though, has deserved little attention in large-scale epidemiological studies (Moreira et al., 2016; Roos et al., 2017); it is usually assumed that heterogeneous data sources are compatible with each other. In practice, though, the highly heterogeneous data sources lead to poor interoperability, which creates barriers to effectively combat pandemics like COVID-19, as indicated by Legido-Quigley et al. (2020).

4.1.3. User Interfaces

A successful platform also involves a suitable end-user interface Maciejewski et al. (2011). In this sense, data consumption by public health officials and global health agencies will require user-friendly web-based interfaces, using common dashboarding techniques. However, beyond this, the proposed federated platform has to consider citizen participation and the heterogeneity of end-users. Therefore, the following characteristics should be taken into consideration in the design of the platform: it should be (a) explainable/accountable to improve, for instance, awareness and engagement (Geneviève et al., 2019); (b) gamified to engage and incentivize participation (Wazny, 2018; Neto et al., 2020); and (c) customizable for different user groups at an international level.

4.2. Proposed Architecture

The proposed architecture has to articulate the key elements from data acquisition to analytics, following the best practices related to privacy and cyber-security. At the acquisition level, in addition to the existing data retrieval from the web, wireless-connected devices will send data to edge nodes that will be associated with specific regions. Data will be anonymously preprocessed at the edge, including some intelligent detection of anomalies or event detection, and then sent to regional (cloud) servers to run a more complete model that will fuse the geolocated timestamped data and run detection models based on explainable AI approaches combined with mathematical computational methods. In this case, crowd-sourcing models are applied based on hardware as a public good in an approach similar to the DIASPORA social network (Bielenberg et al., 2012), also extending this approach to software and data. Regional models will be associated in *federations* resembling the governance structure of the actual regions under interest, which is customizable to local policies/governance models.

Following this federated and collaborative organization, it will be possible to build organically a *fine-tuned early detection* system at the global level in a decentralized yet hierarchical manner to support *graceful lockdowns*. This would turn the proposed federated architecture into a holarchy, where each level is independent and self-sustained, but it can also be encapsulated at a level to capture new goals (Diaconescu et al., 2018).

This system will rely on DLTs to guarantee a trustful system without a responsible third party and minimize the risk of data manipulation (Nguyen et al., 2020). Distributed ledgers could also be used to facilitate different (crypto-economic) incentive mechanisms, such as token curated registries. Its main objective is to build a crowd-sensing trustworthy platform based on privacy-preserving methods to detect potential harmful symptoms in a specific region in almost real-time and flag them to relevant authorities. This alarm needs to be accurate, reliable, and explainable. It will also require a user-tailored interface that could empower citizens by providing detailed explanations and easy-to-access information, while being a tool for policymakers (from the city level to global organizations) designing the correct interventions (a variety of incentives, sanctions, and other persuasive measures) given the specific context of the epidemic in some specific region in a given period of time.
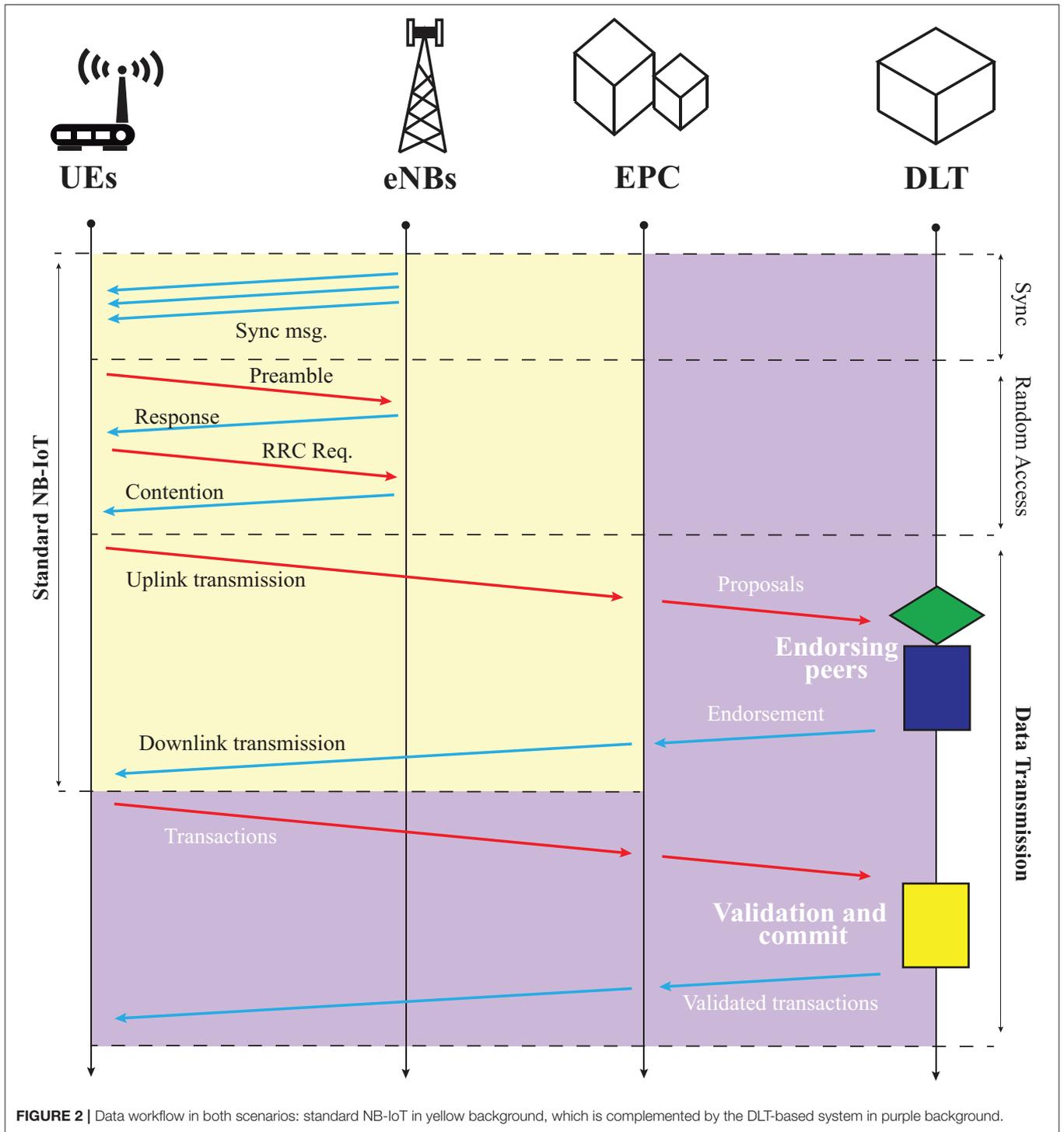
4.3. Representative Use Case

To obtain insights from a practical implementation, we define a specific use case that matches the premises of the proposed architecture. This key industry use case focuses on remote patient monitoring. Here, many issues in treating chronic patients could be reduced or resolved through more efficient patient care. In the case of pandemics, such as COVID-19, the remote monitoring of diagnosed patient can represent a key difference to avoid the virus acceleration. Besides that, the application should guarantee privacy and security on the gathered data. To guarantee people participation, the system can provide incentives, such as bonus or tax compensations. In the context of the proposed federated global epidemiological warning system, the remote patient monitoring is a representative use case, in which the integration between DLTs and IoT devices plays a key role. We analyze this integration through an experimental setup based on NB-IoT, which is one of the most representative low-power wide-area network (LPWAN) technologies in cellular networks. Considering that the proposed architecture requires massive end devices deployment, low power consumption, superior coverage, and worldwide deployment, NB-IoT becomes a key solution to support the federated global epidemiological warning system. Some details of the integration setup are described in the next sub-section, emphasizing data workflow and end-to-end (E2E) latency. We consider latency because the experimental results indicate that it is the most sensible network parameter when DLTs are built in the top of the connectivity system.

4.4. Experimental Setup and Results
4.4.1. Setup Description

In this section, we provide brief explanation of the integration between the NB-IoT network and the DLT-based system, which is illustrated in **Figure 2**.

In a conventional NB-IoT system, the uplink data generated by the user equipments (UEs) is transmitted though specialized packet messages and routed toward an edge data center to be stored and processed (yellow background in **Figure 2**—representing a standard-NB-IoT). At this point, the monitoring system has no control on the collected data, so modification, corruption and losses may occur. Conversely, in our DLT-enabled NB-IoT setup, the uplink data generated by the UEs is transmitted to a randomly chosen group of endorsing peers of Hyperledger Fabric, a type of DLT used in our study, as transaction proposals (check the purple background in
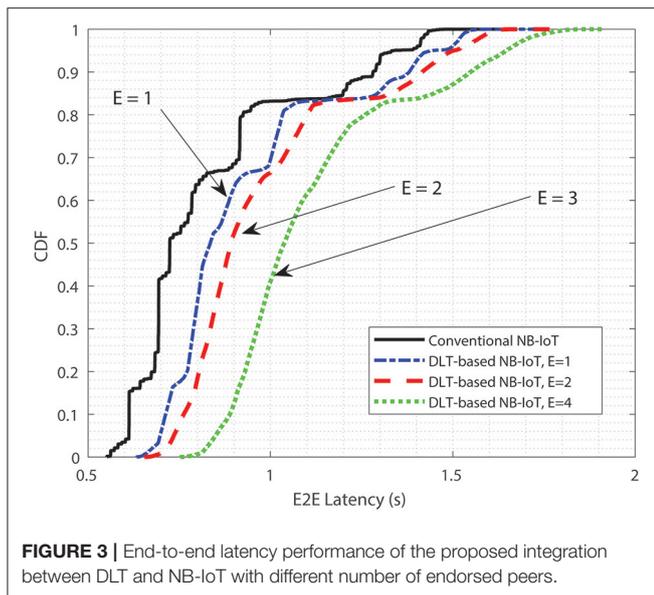
FIGURE 2 | Data workflow in both scenarios: standard NB-IoT in yellow background, which is complemented by the DLT-based system in purple background.

Figure 2). Then, each of the peers signs the transaction using the elliptic curve digital signature algorithm (ECDSA) and adds the signature before returning the signed message back to the IoT devices (UEs).

The peers that provide an endorsement of the ledger send an update to the application, but do not immediately apply the proposed update to their copy of the ledger. Instead, a response is sent back to the IoT device to confirm that the transaction

proposal is correct, has not been previously submitted to the ledger, and has a valid signature. Therefore, the security increases with the number of endorsing peers. In addition, smart contracts can be executed to update or query the ledger.

Then the IoT device broadcast the confirmed transaction proposals along with the confirmation to the ordering service. The received transactions are ordered chronologically to create blocks. These transaction blocks are delivered to all peers for

FIGURE 3 | End-to-end latency performance of the proposed integration between DLT and NB-IoT with different number of endorsed peers.

validation. Then, the peers append the block to the ledger, and the valid transactions are committed to the current state database. Finally, a confirmation message is transmitted back to the IoT devices to report that the submitted transaction has been immutably published to the ledger. The importance of confirmation is analyzed in Danzi et al. (2020).

4.4.2. Results
We evaluate the performance of an integrated DLT and IoT system with *permissioned Hyperledger* Fabric and NB-IoT. Our experimental setup is based on Hyperledger Fabric v1.4, NB-IoT development kits SARA EVK N211 and one NB-IoT Amarisoft eNB station. We compare the E2E latency of conventional NB-IoT with DLT-based NB-IoT in various scenarios. The number of endorsing peers is varied from one to four peers. The IoT device transmits packet every minute, and the total number of messages transmitted is around 1,000.

Figure 3 shows that the end-to-end (E2E) latency is computed based on the total uplink and downlink transmission latency and validation latency in the ledger. We observe that when adding a new endorsing node to the same channel, this will increase the endorsing latency. In comparison with conventional NB-IoT system, the DLT-based NB-IoT performs at a higher latency due to the validation process from ledger to verify transaction as well as building block. This is a trade-off between the DLT-based system and the conventional IoT system. DLT-based IoT system guarantees the trust; however, the latency is slightly degraded.

5. DISCUSSION AND FUTURE PERSPECTIVE

5.1. Latency Analysis on Practical DLT and NB-IoT Integration Setup
A federated global epidemiological warning system was proposed. Based on it, a specialized use case focused on remote

patient monitoring was considered. As one critical piece of this use case is the integration of DLTs and IoT devices, we implemented an experimental integration setup between DLTs and NB-IoT.

In this study, we conclude that the latency increased when the number of endorsing peers is increased. Results indicate that in average the latency is increased 0.25 s per endorsed peer that is added. However, this increases the guarantees in terms of secrecy, privacy, and interoperability.

The experimental setup also reveals additional challenges, such as the performance of DLT-based system with massive number of connections, as well as a trade-off between privacy, interoperability, and system performance.

5.2. Federated Global Epidemiological Warning System
Although the proposed architecture is designed to be fair among all citizens, few important practical points still deserve attention. Good quality data usually follow the wealth distribution over all scales of the proposed federation. In other words, good quality data are very likely to be less available in poorer regions of a city. The same is valid for poor regions within a country and poor countries. This existing gap needs to be considered when building incentives for participation, otherwise the proposed solution has the potential of reinforcing inequality (Gianfrancesco et al., 2018).

Another current issue we cannot neglect is the public perception of wireless technologies. In this sense, part of the general public perceives IoT as insecure (Zubiaga et al., 2018) and fifth generation (5G) as a health hazard (Hasan, 2019) to the point of claiming an astonishing causal relation between 5G and COVID-19, which has caused destruction of base stations across the United Kingdom. Beyond these conspiracy theories that are hard to combat, there exist legitimate concerns of anonymized data not being anonymous (Rubinstein and Hartzog, 2016), and of novel surveillance techniques introduced in times of crisis that are maintained for monitoring (legally or illegally) populations, e.g., post-9/11 surveillance in the United States (Landau, 2013).

All these have to be carefully addressed from the early stages of the system development. Furthermore, it has to be clear that neither *technology* nor *data* can prevent another outbreak on their own, but can only provide the extremely valuable tools to enable the holy grail of controlling epidemics: *early detection, early response* along all relevant actors within the federated organization. In other words, technology and data detect and identify potential harms and suggest actions and reactions, but the final diagnostic and further interventions are due to the responsible institutions within the federated structure and citizens' active participation. Our proposal answers those challenges based on informational self-determination as the way to build trustworthy and secure public infrastructure that will enable *graceful lockdowns* as advocated here. In this sense, the proposed solution introduces a more balanced management strategy, moving away from purely top-down approaches toward a participatory system where

the citizens are active. We further expect that, even without being implemented, the high-level architecture introduced here could offer important technological suggestions to decision-makers of how to start smoothly resuming activities after lockdowns.

6. CONCLUSIONS

The federated system warning response envisioned for epidemiological warning and response is a potential enabler to scale a global-level coordination between public health organizations and governments. As it is based on DLTs, secure crypto-economic incentive models can be used to motivate citizens and institutions. Simultaneously, privacy and interoperability are assured.

A specific use case of the proposed system is based on the integration of DLTs and a popular IoT transport technology, such as NB-IoT. A laboratory setup implementation of this integration was proposed and detailed in this paper. It indicated that there is a trade-off between privacy and some key network key performance indicators (KPIs), such as latency. This trade-off should be studied and analyzed in terms of interoperability and system performance in order to maximize the benefit of the proposed trustworthy federated system.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because: The dataset generated on this manuscript is a collection of measurements that are part of a private network. If it is the case, we can try to deal with the setup owner. Requests to access the datasets should be directed to dick.carrillo.melgarejo@lut.fi.

AUTHOR CONTRIBUTIONS

DC: architect of federated global epidemiological warning system, analysis of Proof of concept and results on NB-IoT, and general management of the publication content. LD: proof of concept and results on NB-IoT and DLT. PN: state of the art analysis used in the Introduction and Systemic definitions. EP: data analysis description on sections Epidemiological Models and Data Collection, Wireless Connectivity, and Epidemic-Relevant Information, and Federated Global Epidemiological Warning System. PM: Relevant description of Epidemiological models in section, Epidemiological Models and Data Collection. DR: communication network description. MD: data analysis and communication network description. HS: communication network description and systemic definition of Federated system. AJ and LH-D: Data Analysis models and models used in Chapter II. IM: communication network description and systemic definition of federated system and data analysis methodologies. MU: wireless communication network description on IoT networks. GF: wireless communication overview on IoT networks. PP: system contribution in all aspects of the paper, since the architect definitions, until explanation of proof of concept results. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Adam, D. (2020). Special report: the simulations driving the world's response to COVID-19. *Nature* 580, 316–318. doi: 10.1038/d41586-020-01003-6

Alessi, M., Camillo, A., Giangreco, E., Matera, M., Pino, S., and Storelli, D. (2018). "Make users own their data: a decentralized personal data store prototype based on ethereum and ipfs," in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)* (Split: IEEE), 1–7.

Alshaabi, T., Minot, J., Arnold, M., Adams, J. L., Dewhurst, D. R., Reagan, A. J., et al. (2020). How the world's collective attention is being paid to a pandemic: COVID-19 related 1-gram time series for 24 languages on Twitter. *arXiv preprint* arXiv:2003.12614. doi: 10.1371/journal.pone.0244476

Althouse, B. M., Wenger, E. A., Miller, J. C., Scarpino, S. V., Allard, A., Hébert-Dufresne, L., et al. (2020). Superspreading events in the transmission dynamics of sars-cov-2: opportunities for interventions and control. *PLoS Biol.* 18:e3000897. doi: 10.1371/journal.pbio.3000897

Amat, F., Arenas, A., Falcó-Gimeno, A., and Muñoz, J. (2020). Pandemics meet democracy. Experimental evidence from the COVID-19 crisis in spain. *SocArXiv.* doi: 10.31235/osf.io/dkusw

Arai, H., and Sakuma, J. (2011). "Privacy preserving semi-supervised learning for labeled graphs," in *ECML PKDD 2011. Lecture Notes in Computer Science, Vol 6911* (Berlin: Springer).

Asikis, T., and Pournaras, E. (2020). Optimization of privacy-utility trade-offs under informational self-determination. *Future Generat. Comput. Syst.* 109, 488–489.

Bahri, L., Carminati, B., and Ferrari, E. (2018). Decentralized privacy preserving services for online social networks. *Online Soc. Netw. Media* 6, 18–25. doi: 10.1016/j.osnem.2018.02.001

Bernard, R., Bowsher, G., Milner, C., Boyle, P., Patel, P., and Sullivan, R. (2018). Intelligence and global health: assessing the role of open source and social media intelligence analysis in infectious disease outbreaks. *J. Public Health* 26, 509–514. doi: 10.1007/s10389-018-0899-3

Bielenberg, A., Helm, L., Gentilucci, A., Stefanescu, D., and Zhang, H. (2012). "The growth of diaspora-a decentralized online social network in the wild," in *2012 Proceedings IEEE INFOCOM Workshops* (Orlando, FL: IEEE), 13–18.

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., et al. (2017). "Practical secure aggregation for privacy-preserving machine learning," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, TX), 1175–1191.

Burkell, J. (2017). The future of privacy lies in forgetting the past. *Eur. Data Prot. L Rev.* 3:435. doi: 10.21552/edpl/2017/4/4

Cacciapaglia, G., Cot, C., and Sannino, F. (2020). Second wave COVID-19 pandemics in Europe: a temporal playbook. *Sci. Rep.* 10:15514. doi: 10.1038/s41598-020-72611-5

Carrillo, D., Nardelli, P. H. J., Pournaras, E., Rodríguez, P. M. D. Z., Dzaferagic, M., Siljak, H., et al. (2021). *Containing Future Epidemics withTrustworthy Federated Systems for UbiquitousWarning and Response.* preprint on webpage at: https://lutpub.lut.fi/bitstream/handle/10024/160869/Early_warning_based_on_mobile_connectivity_and_IoT.pdf?isAllowed=y&;sequence=3.

Carter, D., Stojanovic, M., and de Bruijn, B. (2018). Revitalizing the global public health intelligence network (GPHIN). *Online J. Public Health Inform.* 10:e59. doi: 10.5210/ojphi.v10i1.8912

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science* 368, 395–400. doi: 10.1126/science.aba9757

Cho, H., Ippolito, D., and Yu, Y. W. (2020). Contact tracing mobile apps for covid-19: privacy considerations and related trade-offs. *arXiv preprint* arXiv:2003.11511.

Chowdhury, M. J. M., Colman, A., Kabir, M. A., Han, J., and Sarda, P. (2018). "Blockchain as a notarization service for data sharing with personal data store," in *2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)* (New York, NY: IEEE), 1330–1335.

Danzi, P., Kalor, A. E., Sorensen, R. B., Hagelskjær, A. K., Nguyen, L. D., Stefanovic, C., et al. (2020). Communication aspects of the integration of wireless IoT devices with distributed ledger technology. *IEEE Netw.* 34, 47–53. doi: 10.1109/MNET.001.1900180

Davies, S. E., and Youde, J. R. (2016). "Gphin, goarn, gone" the role of the world health organization in global disease surveillance and response," in *The Politics of Surveillance and Response to Disease Outbreaks* (London, UK: Routledge), 119–132.

Diaconescu, A., Porter, B., Rodrigues, R., and Pournaras, E. (2018). "Hierarchical self-awareness and authority for scalable self-integrating systems," in *2018 IEEE 3rd International Workshops on Foundations and Applications of Self* Systems (FAS* W)* (Trento: IEEE), 168–175.

Do, T. H., Nguyen, D. M., Tsiligianni, E., Cornelis, B., and Deligiannis, N. (2018). "Twitter user geolocation using deep multiview learning," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6304–6308.

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., et al. (2020). Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science* 368:eabb6936. doi: 10.1126/science.abb6936

Firth, J. A., Hellewell, J., Klepac, P., Kissler, S., Kucharski, A. J., and Spurgin, L. G. (2020). Using a real-world network to model localized covid-19 control strategies. *Nat. Med.* 26, 1616–1622. doi: 10.1038/s41591-020-1036-8

Geneviève, L. D., Martani, A., Wangmo, T., Paolotti, D., Koppeschaar, C., Kjelsø, C., et al. (2019). Participatory disease surveillance systems: ethical framework. *J. Med. Intern. Res.* 21:e12273. doi: 10.2196/12273

Gianfrancesco, M. A., Tamang, S., Yazdany, J., and Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern. Med.* 178, 1544–1547. doi: 10.1001/jamainternmed.2018.3763

Hagendorff, T. (2020). The ethics of ai ethics: an evaluation of guidelines. *Minds Mach.* 30, 99–120. doi: 10.1007/s11023-020-09517-8

Hasan, S. F. (2019). "The 5G debate in New Zealand–Government actions and public perception," in *International Conference on Ad Hoc Networks* (Queenstown: Springer), 3–12.

Hébert-Dufresne, L., Althouse, B. M., Scarpino, S. V., and Allard, A. (2020). Beyond r0: heterogeneity in secondary infections and probabilistic epidemic forecasting. *J. R. Soc. Interface* 17:20200393. doi: 10.1098/rsif.2020.0393

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., et al. (2019). "Will democracy survive big data and artificial intelligence?," in *Towards Digital Enlightenment* (Zurich: Springer), 73–98.

Ienca, M., and Vayena, E. (2020). On the responsible use of digital data to tackle the COVID-19 pandemic. *Nat. Med.* 26, 463–464. doi: 10.1038/s41591-020-0832-5

Jacoby, M., Antonić, A., Kreiner, K., Łapacz, R., and Pielorz, J. (2017). Semantic interoperability as key to IoT platform federation. *Lecture Notes Comput. Sci.* 10218, 3–19. doi: 10.1007/978-3-319-56877-5_1

Kojaku, S., Hébert-Dufresne, L., Mones, E., Lehmann, S., and Ahn, Y.-Y. (2020). The effectiveness of backward contact tracing in networks. *arXiv preprints* arXiv–2005. doi: 10.1038/s41567-021-01187-2

Kupferschmidt, K., and Cohen, J. (2020). Can China's COVID-19 strategy work elsewhere? *Science* 367, 1061–1062. doi: 10.1126/science.367.6482.1061

Landau, S. (2013). Making sense from Snowden: what's significant in the NSA surveillance revelations. *IEEE Secur. Privacy* 11, 54–63. doi: 10.1109/MSP.2013.90

Legido-Quigley, H., Asgari, N., Teo, Y. Y., Leung, G. M., Oshitani, H., Fukuda, K., et al. (2020). Are high-performing health systems resilient against the covid-19 epidemic? *Lancet* 395, 848–850. doi: 10.1016/S0140-6736(20)30551-1

Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T., et al. (2020). Characterizing the propagation of situational information in social media during covid-19 epidemic: a case study on weibo. *IEEE Trans. Comput. Soc. Syst.* 7, 556–562. doi: 10.1109/TCSS.2020.2980007

Maciejewski, R., Livengood, P., Rudolph, S., Collins, T. F., Ebert, D. S., Brigantic, R. T., et al. (2011). A pandemic influenza modeling and visualization tool. *J. Vis. Lang. Comput.* 22, 268–278. doi: 10.1016/j.jvlc.2011.04.002

Moreira, J., van Sinderen, M. J., Pires, L. F., and Costa, P. D. (2016). "Improving semantic interoperability of big data for epidemiological surveillance," in *International Conference on Interoperability for Enterprise Systems and Applications, I-ESA 2016* (Guimaraes: International Society for Technology in Education), 314–324.

Musciotto, F., Delpriori, S., Castagno, P., and Pournaras, E. (2016). "Mining social interactions in privacy-preserving temporal networks," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (San Francisco, CA: IEEE), 1103–1110.

Mykhalovskiy, E., and Weir, L. (2006). The global public health intelligence network and early warning outbreak detection. *Can. J. Public Health* 97, 42–44. doi: 10.1007/BF03405213

Neto, O. L., Cruz, O., Albuquerque, J., de Sousa, M. N., Smolinski, M., Cesse, E. Â. P., et al. (2020). Participatory surveillance based on crowdsourcing during the rio 2016 olympic games using the guardians of health platform: descriptive study. *JMIR Public Health Surveill.* 6:e16119. doi: 10.2196/16119

Nguyen, L. D., Kalor, A. E., Leyva-Mayorga, I., and Popovski, P. (2020). Trusted wireless monitoring based on distributed ledgers over NB-IoT connectivity. *IEEE Commun. Magaz.* 58, 77–83. doi: 10.1109/MCOM.001.2000116

Oliver, N., Letouzé, E., Sterly, H., Delataille, S., De Nadai, M., Lepri, B., et al. (2020). Mobile phone data and COVID-19: missing an opportunity? *arXiv preprints* arXiv:2003.12347.

PEPP-PT. *Pan-European Privacy-Preserving Proximity Tracing*.

Pournaras, E., Nikolic, J., Omerzel, A., and Helbing, D. (2017). "Engineering democratization in internet of things data analytics," in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)* (Taipei: IEEE), 994–1003.

Pournaras, E., Pilgerstorfer, P., and Asikis, T. (2018). Decentralized collective learning for self-managed sharing economies. *ACM Trans. Auton. Adapt. Syst.* 13, 1–33. doi: 10.1145/3277668

Roos, M., Martin, E. L., and Wilkinson, M. D. (2017). "Preparing data at the source to foster interoperability across rare disease resources," in *Rare Diseases Epidemiology: Update and Overview* (Cham: Springer), 165–179.

Rosa, R. L., De Silva, M. J., Silva, D. H., Ayub, M. S., Carrillo, D., Nardelli, P. H. J., et al. (2020). Event detection system based on user behavior changes in online social networks: case of the COVID-19 pandemic. *IEEE Access* 8, 158806–158825. doi: 10.1109/ACCESS.2020.3020391

Rubinstein, I. S., and Hartzog, W. (2016). Anonymization and risk. *Wash. L. Rev.* 91:703.

Sapiezynski, P., Stopczynski, A., Wind, D. K., Leskovec, J., and Lehmann, S. (2017). Inferring person-to-person proximity using WiFi signals. *Proc. ACM Interact. Mobile Wearable Ubiq. Technol.* 1, 1–20. doi: 10.1145/3090089

Schulz, S., Stegwee, R., and Chronaki, C. (2019). *Fundamentals of Clinical Data Science* (Cham: Springer Nature).

Smith, V., Chiang, C.-K., Sanjabi, M., and Talwalkar, A. S. (2017). "Federated multi-task learning," in *Advances in Neural Information Processing Systems* (Long Beach, CA), 4424–4434.

Ting, D. S. W., Carin, L., Dzau, V., and Wong, T. Y. (2020). Digital technology and COVID-19. *Nat. Med.* 26, 459–461. doi: 10.1038/s41591-020-0824-5

Truong, N. B., Sun, K., and Guo, Y. (2019). "Blockchain-based personal data management: from fiction to solution," in *2019 IEEE 18th International Symposium on Network Computing and Applications (NCA)* (Cambridge, MA: IEEE), 1–8.

Vergara-Laurens, I. J., Jaimes, L. G., and Labrador, M. A. (2016). Privacy-preserving mechanisms for crowdsensing: survey and research challenges. *IEEE Intern. Things J.* 4, 855–869. doi: 10.1109/JIOT.2016.2594205

Wang, L., Didelot, X., and Yang, J. (2020). Inference of person-to-person transmission of covid-19 reveals hidden super-spreading events during the early outbreak phase. *Nat. Commun.* 11, 1–6. doi: 10.1038/s41467-020-18836-4

Wazny, K. (2018). Applications of crowdsourcing in health: an overview. *J. Glob. Health* 8:010502. doi: 10.7189/jogh.08.010502

Wei, W. E. (2020). Presymptomatic transmission of SARS-CoV-2-Singapore, January 23–March 16, 2020. *MMWR. Morb. Mortal. Wkly. Rep.* 69, 411–415. doi: 10.15585/mmwr.mm6914e1

Wright, S. A. (2019). "Privacy in iot blockchains: with big data comes big responsibility," in *2019 IEEE International Conference on Big Data (Big Data)* (Los Angeles, CA: IEEE), 5282–5291.

Yang, Q., Liu, Y., Chen, T., and Tong, Y. (2019). Federated machine learning: concept and applications. *ACM Trans. Intell. Syst. Technol.* 10, 1–19. doi: 10.1145/3298981

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., et al. (2020). Changes in contact patterns shape the dynamics of the covid-19 outbreak in China. *Science* 368, 1481–1486. doi: 10.1126/science.abb8001

Zubiaga, A., Procter, R., and Maple, C. (2018). A longitudinal analysis of the public perception of the opportunities and challenges of the internet of things. *PLoS ONE* 13:e0209472. doi: 10.1371/journal.pone.0209472