

The Case for Cross-Border R&D in Digital Epidemiology

Proposal for an International Research Consortium

The global Covid-19 response has revealed the need for sovereign nations to share epidemiological data of all types even as they struggle to develop aligned approaches to digital epidemiology that protect data privacy. The proposed R&D consortium of member nations should bring together researchers and data scientists with, as necessary, government regulators and government “data diplomats” to address pressing and specific challenges in epidemiology which have been identified as national priorities. The collaboration should focus on solving pressing cross-border R&D problems in digital epidemiology—not creating data sharing guidelines, but rather creating specific solutions that can inform guidelines—working in collaboration with policy makers and regulators.

The COVID-19 pandemic has revealed a common need of sovereign nations for more effective digital epidemiology. Public health authorities and nationally funded bioscience research communities need the ability to rapidly analyze multiple types of digital information on a disease outbreak and respond across multiple geographies. Further, in virtually every country, private enterprises are a critical element of a nation’s public health response to disease outbreaks. This means that effective digital epidemiology is also a common need for domestic health care companies, multinational medical device companies, and multinational pharmaceutical companies.

The COVID-19 pandemic has also revealed that effective digital epidemiology is not just a common need but also a shared need. The health security¹ of one nation is irrevocably tied to the public health of virtually every other nation. Unlike traditional national security, it is not just the public health of a nation’s allies that matters but — because of global economic integration and inherently porous borders — public health in nations regarded as strategic competitors or those with malign intent are also critical to national health security.

Infectious disease outbreaks are inherently cross-border and risk mitigation, public health responses, and medical responses all depend on access to, and analysis of, information from other nations and localities. No nation can address these borderless problems in isolation

¹ Health security resides in effective public health and medical care for a nation’s population and overlaps substantially with, but is distinguishable from, biosecurity which relates to the ability to defend against biological weapons deployed by malign actors or strategic competitors in acts of war or terrorism.

and COVID-19 has highlighted long-standing and unresolved cross-border data privacy and policy differences.

There is an immediate need for technologically advanced liberal democracies to lead in mounting new international R&D collaborations in digital epidemiology that engage the largest possible population of nations without undermining the social compact (consent of the governed) that is unique to the liberal democracies.

Digital Epidemiology, Public Health, and Privacy in Liberal Democracies

Adopting a broad definition of digital epidemiology, the immediately apparent R&D needs span (1) the rapid digital transformation of traditional epidemiology; (2) new approaches based in data mining and artificial intelligence and working with viral sequences, human genomic information, and medical records; and (3) new approaches to traditional health and environmental data as well as approaches that draw health-relevant information from social or digital media, commercial transactions, transport information, and cell phone locational data.

Many countries are currently focused on similar analytical problems in health security; every robust national solution requires international collaboration, certainly with regard to data and almost certainly with regard to analytical approaches; and establishing common, or workably aligned, data privacy standards is on the critical path for virtually every liberal democracy.

While the rights and data privacy of an individual may seem unimportant in comparison to the risk of global pandemics, a critical feature of public health in liberal democracies is the willingness of individuals to voluntarily act in the interest of the community – from curbing or changing personal activities (e.g., reducing travel) to active participation in vaccination programs, participating in clinical trials, or sharing phenotypic and genotypic information with health care researchers in both the public and private sectors.

Public health in liberal democracies depends on citizens trusting that their privacy will be protected, that digitization can benefit their lives, and that data sharing is not simply a means of financializing health or de-prioritizing individuals. This is all the more so among societal groups who have historically not received quality and affirming health care on par with others in their nations.

International R&D Consortium in Digital Epidemiology with Privacy by Design

Establishing working international R&D collaboration in digital epidemiology is like developing cross-border high speed rail (digital data) on top of a rail system where individual countries still have different track gauges and where there is massive bureaucracy over passenger identification at each border (traditional health data). As a result, international collaboration on public health matters and on data privacy is hardly unplowed ground.

<https://www.brginstitute.org/project-description>

3 May 2021

First, there are national and international public health entities (e.g., US CDC and WHO) that have considerable interest in and experience with health data and health security. The same is true of a wide range of public and private enterprises that engage in, or fund, health research from the US NIH to global pharmaceutical and medical device companies. Second, cross-border digital (and data) issues are the subject of a number of promising initiatives such as those under the aegis of the UN and the GPAI:

- UN roadmap for digital cooperation <https://www.un.org/en/content/digital-cooperation-roadmap/> and;
- GPAI data governance working group <https://gpai.ai/> and <https://gpai.ai/projects/ai-and-pandemic-response/gpai-ai-pandemic-response-wg-report-november-2020.pdf>

An R&D collaboration can, however, be much more focused than the activities of either public health entities or international digital governance/sharing initiatives. An R&D consortium can be a vehicle to bring together researchers and data scientists with, as necessary, regulators and “data diplomats” to address pressing and specific challenges in epidemiology. Effective sovereign-to-sovereign R&D collaboration in digital epidemiology – an international R&D consortium – needs to be grounded in three key features:

Governing mechanisms and operating procedures that treat member nations as the client for R&D services: Global health and national public health is crowded with research and operational not-for-profit (NGOs) and private, for-profit multinational companies (MNCs). While both of these communities are critical, the world has not yet reached a point of stateless collaboration. Blunt geopolitics (jockeying for position and advantage) and genuine adherence to different principles and interpretations of the rights of the governed and methods of governance (e.g., data privacy regulations or antitrust protections) still shape international collaboration in global and public health.

Because sovereign nations have unaddressed shared and common needs for R&D the proposed consortium in digital epidemiology should be funded by, and responsive to, nations. Given the role of private enterprises and not-for-profits in national public health activities, there should be room for such enterprises at the table, but the controlling governance body should be made up of representatives of supporting governments. This R&D consortium may need to start with only a few participating nations, but those nations should commit to eventually engaging the largest possible population of nations without undermining the social compact (consent of the governed) that is unique to the liberal democracies.

Mechanisms to address national data privacy issues: As a vehicle to pursue common and shared R&D challenges in digital epidemiology, an R&D consortium will need mechanisms to address both intra-jurisdictional and cross-border data privacy and regulatory issues. Regulators in all nations attend to national laws and international agreements and take their responsibility to citizens seriously. R&D in digital epidemiology is not, of course, only a matter of sharing personal health or DNA data; contact tracing data from cell phones, for example, will raise very different concerns about intrusive surveillance than targeted data on patient responses to a drug. Therefore, a significant but important challenge in organizing

the consortium is how to instantiate the capacity to navigate multiple national regulatory processes for specific R&D activities, perhaps by having regulatory personnel assigned specifically to the consortium from member governments.

Strategy for results and privacy by design: An R&D consortium comprised of sovereign members will not resolve competing national data privacy laws and regulations and should not be expected to do so. The strategy for such a consortium should be to focus on best-available solutions to pressing problems identified by member nations and their representative health enterprises, not creating data sharing guidelines.

A second critical feature of the research strategy for the consortium is to pioneer approaches and solutions in digital epidemiology founded in privacy by design. This requires creating approaches that enable data sharing while strengthening security and respecting and incorporating international privacy norms. Adopting this as a feature of the consortium's research strategy is a significant break with the past as international privacy norms have usually been regarded as a barrier rather than a part of the solution. The challenges are inherently cross-border and so must be the solutions.

There are a number of promising new approaches to solving privacy issues. One illustrative approach is the concept of "taking the algorithm to the data", allowing data to be held and protected separately (not pooled). Similarly, the Broad Institute Terra platform uses the concept of federated computing. This approach to privacy by design allows the agglomeration necessary for big-data and cross-border approaches by accessing, but not pooling, multiple data sets. The trend toward doing more computing at the edge of a networked system (that is, locally rather than in a remote data center or the cloud) can also support privacy. It can involve computing within a mobile or other end-user device, or it can involve dedicated intermediate (edge) systems. This approach has the benefits of local processing and a reduction in the volume of what is communicated, conserving bandwidth. Alternatively variety of statistical techniques may be used to modify data in ways that protect privacy and also support analyses of aggregated data. While not a panacea, the concept of "taking the algorithm to the data" gives a glimpse of technical solutions to solve cross-border data privacy concerns.

Starting from this position, an international R&D consortium could deliver immediate value to sovereign nations and health care enterprises by: (1) providing the best-available answers to select pressing questions while highlighting and proposing solutions to the limitations imposed by current data sharing and access agreements; and (2) applying privacy-by-design approaches to intra-jurisdictional and cross-border digital epidemiology with the goal of improving on the current best-available answers.

Also, the consortium has an opportunity to demonstrate its value by delivering early "wins" in improved economic performance and public health of member nations.

The Potential Economic Benefits of Collaborative R&D in Digital Epidemiology

China is the only major economy to emerge from 2020 with positive GDP growth and one element of China's relative public health and economic success during the pandemic has been government's access to all types of domestic data. This highlights the immediate public health value (and thereby economic value) of being able to use a wide range of data from a large population – 1.4 billion people in China, equaling the combined populations of US, Europe and the Middle East North Africa region (MENA) – for creative digital epidemiology.

The point, of course, is not to argue that liberal democracies should abandon individual rights and consent of the governed but rather to drive home the point that there are very real and shared economic and public health benefits to rapidly resolving health data collaboration among liberal democracies.

There are important long-term economic benefits also. Existing and successful mechanisms of public and private sector data sharing in genomics and virology have been highlighted in recent months by the multinational (and multi-corporate) press to develop and test COVID-19 vaccines.

Further, new generation cross-border digital epidemiology has the potential to reach substantially “upstream” from viral sequencing and vaccine development to early identification of a disease outbreak as well as “downstream” into prediction of the path of an outbreak and the design and evaluation of different mitigation measures. In short, new generation digital epidemiology is an important tool in preventing and mitigating both health crises and the economic disruption that inevitably accompanies them.

Further, the same systems of data sharing (or approaches that enable data access for analysis without revealing individuals' data) have long been understood as potentially providing significant long-term health and economic value. Predictive medicine, empirically-based treatment, health management, and disease management are increasingly important aspects of health care – and therefore the focus of considerable interest and effort by health care enterprises and companies – worldwide. The COVID-19 pandemic may be the immediate stimulus driving advances in digital epidemiology but the long-term economic value of the liberal democracies “getting this right” will be huge.

Conclusion

There are knotty technical questions — in both analysis and data — and very real national cybersecurity vulnerabilities in establishing a multinational research consortium in digital epidemiology. In both health and economic terms, the benefits of a shared R&D enterprise will greatly outweigh the costs, but national governments need to become comfortable with the risk-return tradeoffs of the necessary sovereign-to-sovereign agreements in such an enterprise.

As with any new approach or start-up, there are advantages to starting with a small team and an exploratory stance. One approach is for the ministers of science, health, and security of the G7 nations establish a working group, comprised of representatives of governments, health care enterprises, and researchers from all seven member nations. The pressing task for such a working group is to design an R&D enterprise that can serve the health security needs of all liberal democracies, while protecting the rights and privacy of entities and individuals. Starting with the G7 is not a final step but rather an approach for working out, among a small group of liberal democracies, a template that should ultimately allow the participation of a much larger number of countries.

Digital epidemiology is a pressing area of need where specific solutions are required for specific projects. In virtually every liberal democracy data privacy is a fundamental human right that has to be addressed and safeguarded. The proposed consortium would allow governments to collaborate immediately on cross-border problems while supporting and enabling new privacy protecting approaches.

ENDNOTES

1. The work plan for the project on Global Innovation and National Interests includes the development of, and advocacy for, proposals for public-private international R&D consortia. Six members of the [Project Working Group](#) led the development of this case statement and proposal: Marjory Blumenthal, David Delpy, Bruce Guile, Stephen Johnson, Steven Koonin, and Tadaaki Taniguchi. Bruce Guile and Stephen Johnson led in drafting this case statement and proposal.
2. This proposal is additive to, and consistent with, the March 31, 2021 Statement of the Science Academies of the Group of Seven (G7) nations "[Data for International Health Emergencies: governance, operations and skills](#)": If there are differences, they lie in the idea that an R&D consortium would press for more coordinated and urgent action at a governmental level, and a process where specific expedited projects (not solely health emergencies) are used to inform regulatory and technical solutions to data issues.