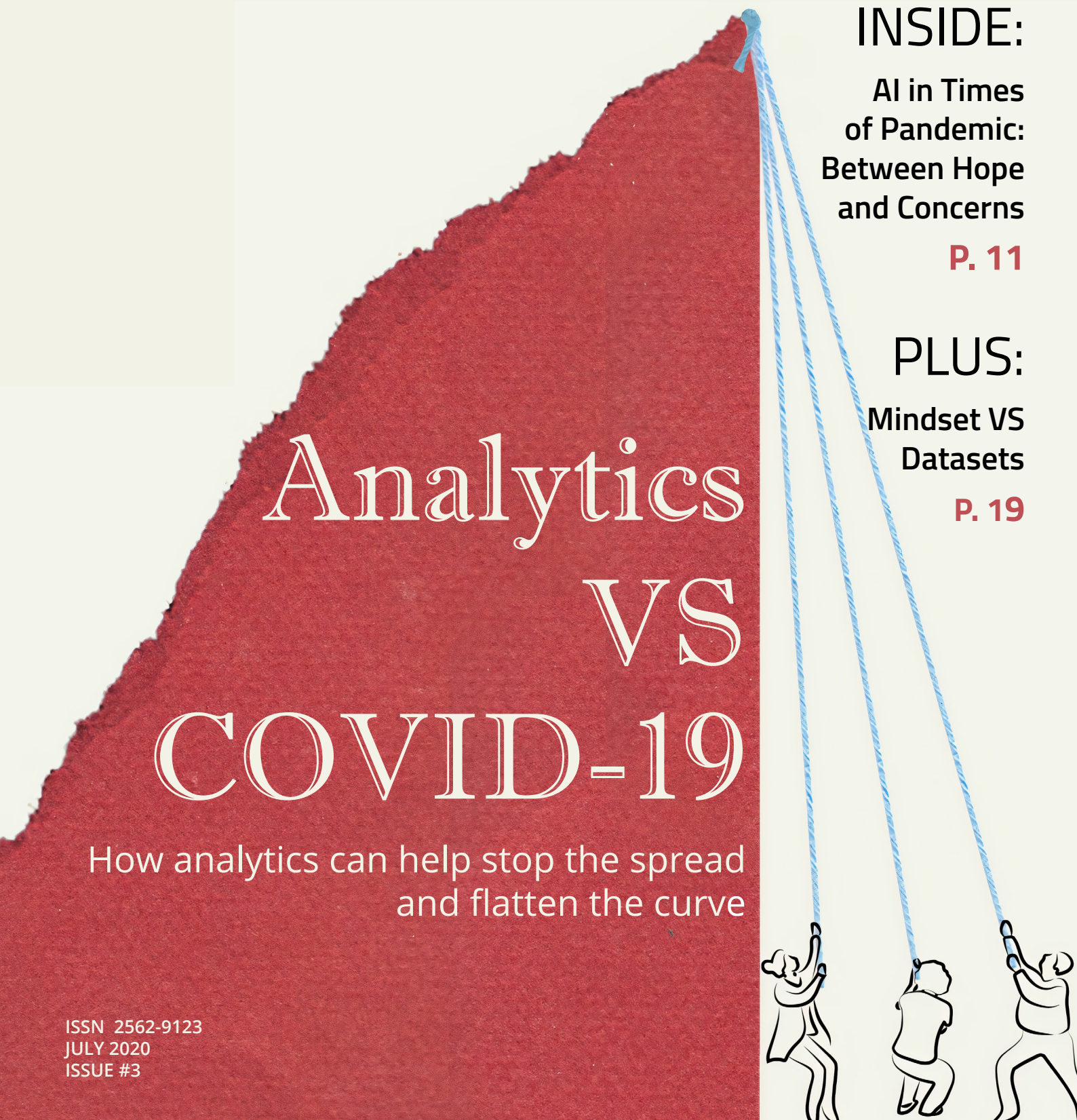
ANALYTICS IN GOVERNMENT QUARTERLY

FOR GOVERNMENT DECISION MAKERS



INSIDE:

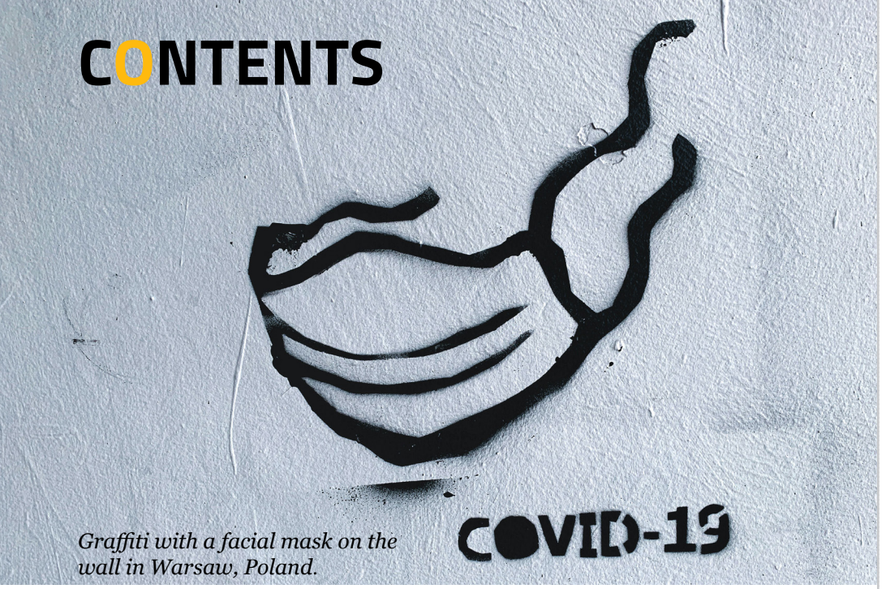
PLUS:

Analytics VS COVID-19

How analytics can help stop the spread and flatten the curve

ISSN 2562-9123
JULY 2020
ISSUE #3

CONTENTS



Graffiti with a facial mask on the wall in Warsaw, Poland.

FEATURES

COLUMNS

ANALYTICS IN GOVERNMENT QUARTERLY

EDITOR'S LETTER

In this issue, we explore the use of analytics that include visualization and artificial intelligence in helping the public to understand and cope with emergency situations such as pandemics. Hubert Laferrière explores the delicate issue of privacy versus public protection related to contact tracing apps. One the one hand, these apps could be useful in the fight against the spread of viruses in general. On the other, they raise critical privacy concerns and, recent findings suggest that they might not be as effective as simpler and less invasive analog methods.

Kevin Kells explores the use of visualization in presenting virus-related information to the public. One of the critical issues is the degree to which the human brain is wired to understand the implications of exponential growth. Kevin discusses the importance of providing easily understandable data to illustrate credible evidence to enable collective action in stopping the spread of the virus. Tara Holland builds on these ideas to discuss the organizational aspects of data packaging and delivery exploring the close link that should exist between policy makers and data scientists. We have all seen the different responses to pandemic management from countries across the globe: some politicians who are guided by data and others who are guided by other considerations. Betty Ann Turpin explores the broader issues of cause and effect of pandemics. Alex Ramirez provides a unique perspective on the many different voices that should be heard when deploying analytics in government. Rami Abielmona discusses the application of AI in emergency situations.

Ultimately, the use of advanced technologies can play a role in helping us to manage policy responses to pandemics as well as other critical social issues. But the willingness of policy makers to follow data-driven guidance that can help balance the economic, social and health-related impact of their decisions is not something that any algorithm can predict. We hope you enjoy the readings in this issue.

Gregory Richards, Ph.D.
Managing Editor

Corporate

Analytics in Government Quarterly magazine is published 4 times per year by the Government Analytics Research Institute, a consortium of the University of Ottawa, Carleton University, the University of Quebec en Outaouais, SAS and the Institute on Governance. The institute conducts research with government organizations who are experimenting with the introduction of analytics of all forms. Professors and students work on proof of concepts, testing of algorithms as well as examining the organizational practices needed to fully integrate analytics into business processes.

All opinions expressed herein are those of the contributors and do not necessarily reflect the views of the publisher or any person or organization associated with the magazine.

General Inquires

Letters, submissions, comments and suggested topics are welcome, and should be sent to agq@governmentanalytics.institute or visit our website http://governmentanalytics.institute

Subscription Information

You can subscribe to the magazine online at http://governmentanalytics.institute/magazine

ISSN 2562-9123

By Kevin Kells, Ph.D.

Visual Analytics as a Key to Help the Public Understand Pandemic Data

Visualizing near-real-time data has played an important role in helping the public understand the situation of the COVID-19 pandemic and to help persuade members of our communities to choose behaviors that are better for them and, crucially, better for everyone in the community. Visualizing health system and case data is obviously important for health professionals and government authorities responding to the pandemic and for entrepreneurs and industry who can mobilize to fill gaps and needs. But for the public, situational awareness provided by audience-appropriate presentation of pandemic data is also very important.

A visual understanding of the situation reduces feelings of panic and clarifies the logic of shelter-in-place or stay-at-home rules. It also helps maintain trust in government institutions by demonstrating, with data, how individual sacrifices are paying off by establishing control over the virus spread. This trust is important, as eventually an isolated public, exhausted and frustrated, will need to follow the corresponding instructions as the pandemic comes to an end—be it a medical or a social end.[1]

Tell a Navigable Story with a Dashboard

Data Science offers a series of steps when dealing with a problem we're trying to understand and solve: Describe, Diagnose, Predict, and Prescribe, usually undertaken in that order, with each subsequent step increasing in complexity and generally decreasing in the absolute certainty of the results.[2] This article will focus on the "Describe" step, where the factual data is collected, prepared, and then presented to the audience. Pandemic "models" that forecast the future spread of the pandemic virus, foresee future health system capacity, and simulate possible response scenarios fall under the "Diagnose", "Predict" and "Prescribe" steps.

The goal of the Describe step is to provide clarity to the audience and to help them gain insights. One important aspect of this work asks where to obtain the data, how to ensure it is reliable, correct, complete, and up-to-date, and to annotate data sources, assumptions, caveats, and known inaccuracies. The other important aspect is: "How can we present this data to the intended audience in a way they can understand it?" Clarity and insights are human results we want to obtain, so it is about more than just having good data, but presenting it appropriately. In this case, the audience is the general public, so the way we present the data —both the content we choose and the design—must be simple, intuitive, and familiar. The common denominator for the general public is to tell a story visually and to let the audience interact and discover further details on their own.

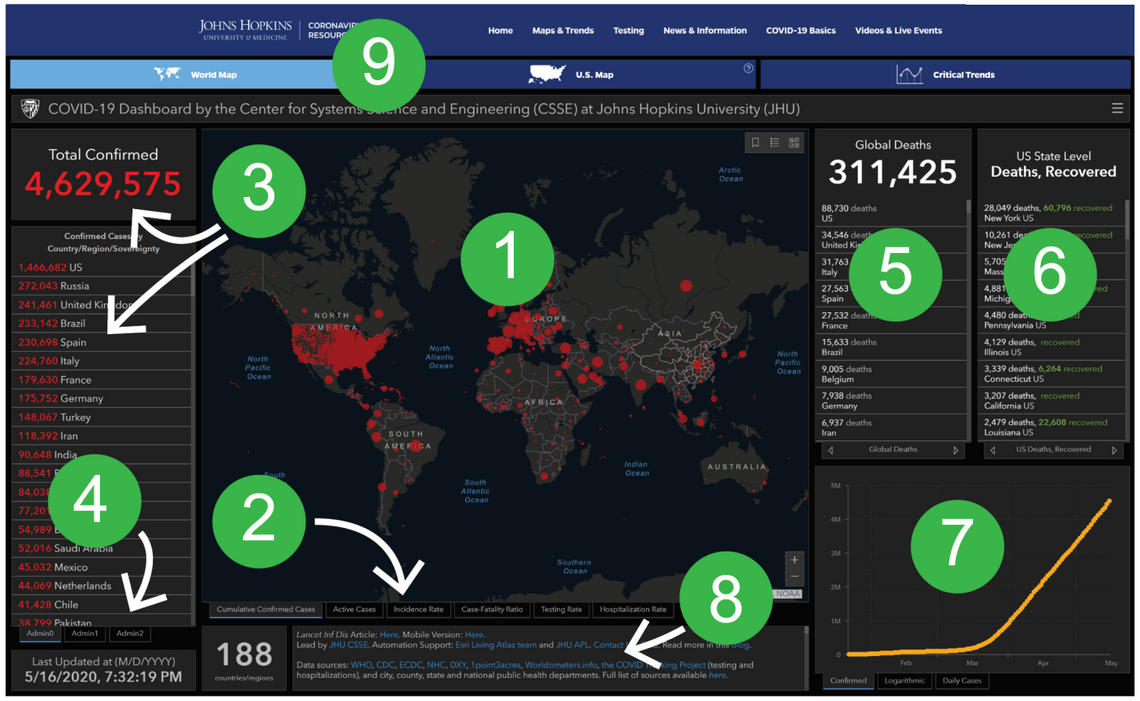
The Coronavirus Resource Center web page dashboard *(See the image below)* hosted by the Johns Hopkins Centers for Civic Impact provides an example of design elements that have made this presentation very successful in the public and media.

What makes this web page a good dashboard are: A) The entire contents are visible on a single, non-scrolling page, and B) A readable summary of all the important data are represented at once in about seven dashboard regions.[3]

Aesthetics are important given the public audience. An evaluation of any data presentation should include a review of the aesthetics so that audience engagement is maximized via different means of human perception. The public audience responds readily to pleasing aesthetics of a presentation, including graphics, font choice, and color scheme; these are important factors in the success of the Johns Hopkins presentation.

The elements on the Johns Hopkins page are:

1. A movable, zoomable



The Coronavirus Resource Center web page developed by John Hopkins University.

1 Gina Kolata, May 10, 2020, "How Pandemics End", NY Times

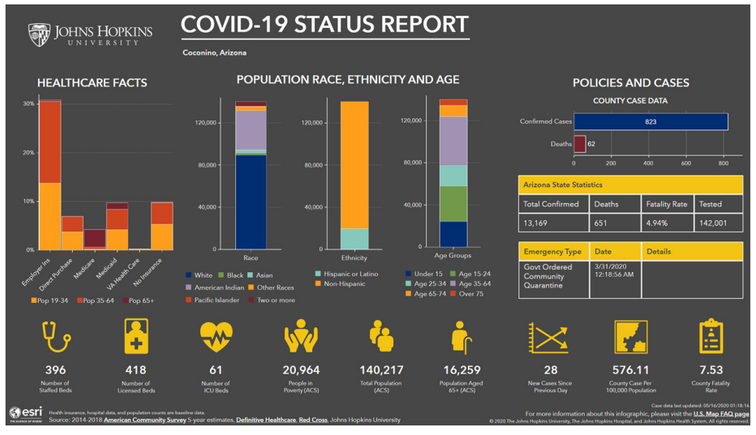
2 Boyer, A. & Bonnin, G. (2016). "Higher Education and the Revolution of Learning Analytics"

3 Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information". Psychological Review.

world map with cumulative confirmed cases visually presented;
2. The tabs at the bottom of the map permit the user to select different geographic datasets;
3. The total confirmed worldwide case count with the confirmed cases by geopolitical subdivision;
4. Different geopolitical subdivision schemes are available;
5. Global deaths and recovered cases;
6. US State level deaths, hospitalizations, and deaths;
7. Global confirmed cases and daily new cases over time;
8. Data sources, caveats, and explanations of terminology;
9. Alternative scope (US instead of Worldwide data) and a Critical Trends presentation.

By clicking into a particular geographic area for more details, we find evidence that the dashboard is continually evolving and improving.

The graphic below is improved over the version. A multi-bar graph and a pie chart were replaced each by stacked bar charts. Pie charts are generally deemed poor graphical elements for presenting data because they are hard to read, they don't have a scale, and because the brain is not as good at comparing angles as it is at comparing lengths. Web-based data visualization is easily changed, reconfigured, and improved. Even user behavior can be taken directly into account to make improvements. For example, if users continually click through to navigate to one certain detail, the display of that detail can be migrated to the main page to increase convenience and reduce mouse clicks needed to find that detail. Any major effort such as the Johns Hopkins COVID-19 dashboard should be accompanied by a process to receive, review, prioritize, and promptly take action on feedback and suggestions that are received both internally and from external users. The feedback-and-improvement process is an integral part of the effort; the setup of this process should be given due attention when organizing and managing such a data visualization effort.



The Coronavirus Resource Center status report.

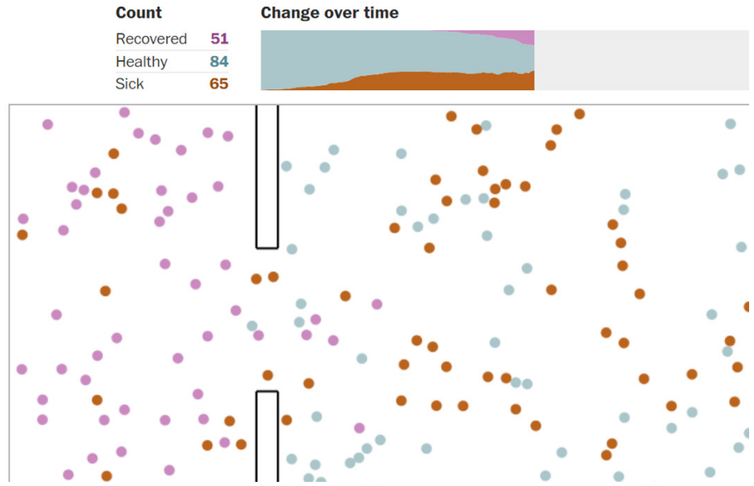
Pandemic Dynamics and Exponential Thinking

As a final example of pandemic data visualization for the public audience we present visual explainers of pandemic dynamics as animated simulation tools. This animated production from the Washington Post, offered in a dozen languages, is graphical and animated. And fun—just what is needed to prompt self-directed engagement of a general, public audience in an educational activity.[4 5]

As members of the general public, we are used to thinking linearly compared with exponentially. When told, "the number of cases increases by 1,000 every three days," we are able to grasp that fairly well. When told, "the number of cases doubles every three days," which is the exponential growth characteristic of an unchecked pandemic, the implications are not readily grasped by the linear thinking process, nor is the crucial urgency of timely decision-making always readily

4 Villagrasa, S., Fonseca, D., Redondo, E., & Duran, J. (2014). "Teaching Case of Gamification and Visual Technologies for Education". Journal of Cases on Information Technology.

5 Hamari, J., Koivisto, J., & Sarsa, H. (2014). "Does Gamification Work? - A Literature Review of Empirical Studies on Gamification."



The Washington Post simulitis spreads in a town of 200 people.

apparent. In the linear case, one month later, the number of cases goes from 1,000 to 11,000. In the exponential case, it goes from 1,000 to 1 million.

If we can avert 1,000 cases from spreading the contagion, in the linear case, we'll have 1,000 fewer cases after a month. In the exponential case, the same reduction of cases is achieved if we wait until the end of the month to avert them. But if we can avert those same 1,000 at the beginning of the month, we reduce by 500,000 the number of cases at month's end, solely due to the nature of exponential growth.[6]

Engaging the Public Audience with Analytics

Communicating data to the general public requires simplicity, aesthetics, interactivity, and the creation of a visual story that is intuitive and familiar. In contrast to academic and professional publications, the burden of good communication is much more on the communicator, where successful conveyance is a function more than of having good data, but of communicating it well.

A dashboard approach to presenting data such as the Johns Hopkins Centers for Civic Impact's Coronavirus Resource Center web page offers the public a single location where they can turn to obtain the most frequent details they seek on one main web page that requires no scrolling. Careful selection of about seven dashboard elements, continually improved and adapted to reflect feedback received, permits users to interactively navigate in a familiar way to find details. The aesthetically pleasing, visual representation of the data captures the public user's attention and leads to improved understanding by encouraging their engagement.

Additional tools such as animated simulations provide the general public interactivity and the element of fun, important to encouraging exploration and engagement.

This pandemic has bridged the sphere of individual behavior and liberty and the sphere of community life and health in a very stark way that is a new experience for most of us. The exponential nature of pandemics makes it harder for the general public to understand the logic of certain government decisions such as lock-downs, though they may be mathematically wise. Reflecting the data back to the public as a story they can understand and explore in a way to help bridge the communication gap that often exists.

6 Siobhan Roberts, 2020-03-13, "The Exponential Power of Now", NY Times Science

About the Author

Kevin Kells, Ph.D., has worked as an R&D Engineer in software systems in the Financial and Semiconductor industries in Switzerland, Silicon Valley, and Ottawa, and currently works with real-time data and news feed systems at a major market news and data company in New York City. He has extensive experience in non-profit management, both in the area of human systems and IT systems. He received his Ph.D. from the Swiss Federal Institute of Technology (ETH), Zurich in computer simulation of semiconductor devices and holds an MBA with areas of focus in entrepreneurship and business analytics from the University of Ottawa, Telfer School of Management.

Artificial Intelligence and Pandemics: How Analytics Can Help

How data-driven decision-making can flatten the curve at all stages of an outbreak

By Tara Holland

New York governor Andrew Cuomo is emerging as a leader in the fight against the Covid-19 pandemic that is crippling the world. His state has been the hardest hit in the U.S., possibly one of the hardest-hit regions in the world, but his staunch resistance to prematurely opening up the state economy has resonated with people.

"I operate on facts and on data and on numbers and on projections," he told reporters in March. "Not, 'I feel,' 'I think,' 'I believe,' 'I want to believe.' Make the decisions based on the data and the science."Those could be the words of a data scientist.

We are swimming in the very data that we need to combat this pandemic and its public health impact: prediction, treatment, allocation of resources, recovery. How can we harness the power of that data to stem the tide of this pandemic, and prepare for the next?

Checking the Spread

It's estimated that 80 percent of new human diseases have "jumped species." As human population grows largely unchecked, we come into increasing contact with animals that, though possibly asymptomatic themselves, harbor viruses that can spread to humans with deadly results. With volumes of detailed data on population, migration patterns and cultural practices, artificial intelligence can model those interactions to predict the hot spots for interactions with unfamiliar creatures.

We also have huge amounts of data from airlines, customs, and railroads from which we can forecast with great accuracy how disease might spread. The power of analytics technology has grown by leaps and bounds, even since the Ebola outbreak of 2014. After the fact, researchers used AI and travel data to remodel the spread of the virus with astonishing accuracy, right down to the county level in Texas.

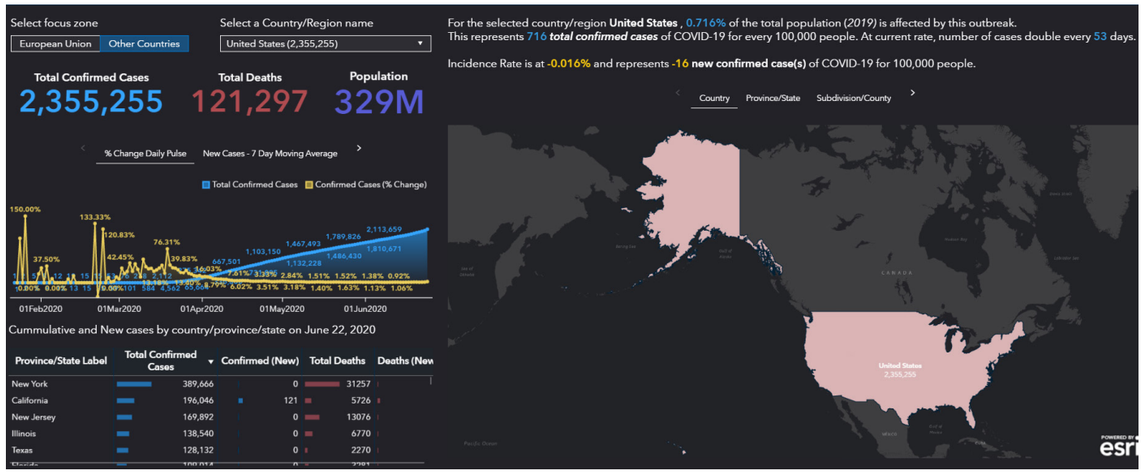
Mining social media content with natural language processing can reveal travel patterns and anomalous illness outbreaks to supplement that travel data, and can do it much faster than traditional forecasting methods.

We can also model how different responses—or lack thereof—can influence the spread of an outbreak. What-if analyses can help direct policy decisions and resource allocation to mitigate the spread. Knowing more, and knowing it sooner, is key to a pandemic response.

Shaping the Response

Policymaking supported by data scientists is still too rare. A key to shaping the response to a public health emergency is presenting a huge volume of data in a fashion that government officials can understand. Consider the number of data sources that must be interpreted and integrated into pandemic response:

1. Public health: Where is testing being done? What is the location of confirmed cases? What are regional recovery rates? If there's a noticeable difference among regions, why?
2. Medical resources: Where are the clinicians? How many hospital beds are available, and where? Where are there ventilators and other critical equipment? Where should supplies be directed?
3. Social Determinants of Health: Covid-19 has had an uneven impact on the population. Where are the pockets that are suffering heavily? Who is suffering from housing and food insecurity? Current social programs data can point the way.
4. Crime: Where are the "hot spots" of escalating crime during the outbreak? How do we direct front-line law enforcement resources?
5. Essential workforce: How many nurses, personal support workers, first responders, army reservists, and others wanting to help, and where? How can they be better deployed and supported?
6. Emergency Benefits: Real-time data on provincial and federal measures being rolled out to support



A dashboard visualize the data related to COVID-19.

families, communities and local businesses who face financial crisis. Understanding commerce and employment by region and industry can help flesh out the pandemic impact picture.

That's a lot to take in. A well-designed dashboard that integrates that data visually, in a user-friendly format, helps decision-makers understand the impact of an outbreak. Interactive visuals can demonstrate the impact of government intervention, whether related to commerce, public safety and health, and public welfare. A clear, comprehensive visual can guide policy-makers to shape their response.

SAS has an online tool, based on its Viya platform, that allows interaction with collected Covid-19 data visually to predict how it can be influenced.

Treatment and Recovery

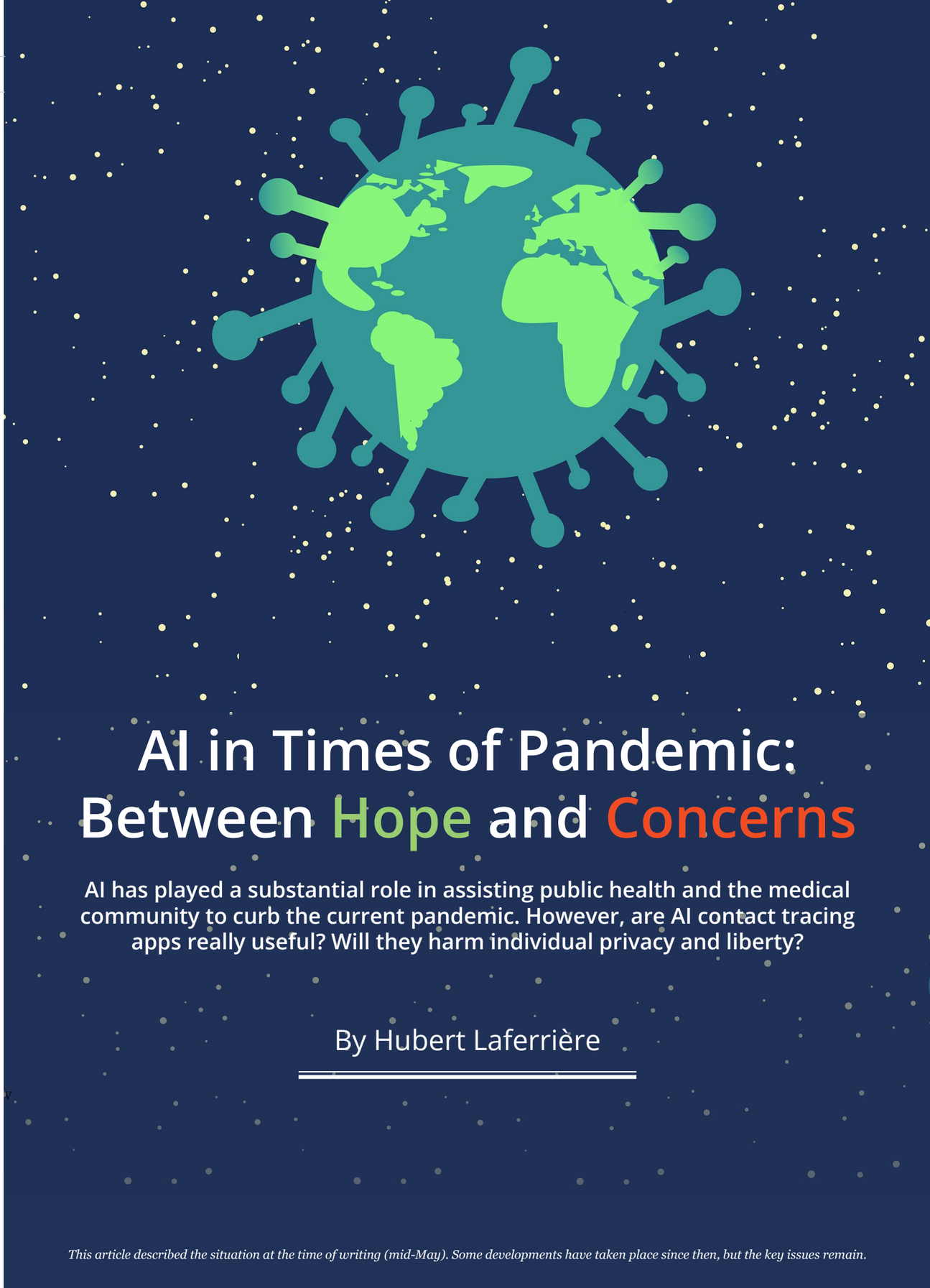
Deep learning—another AI technology—comes into its own when developing treatments for viral outbreaks. By using imagery of Coronavirus patients—X-rays, ultrasounds, and the like—as a data input, deep learning technology can create models that speed diagnosis, analyze the effectiveness of existing treatments, and even point in the direction of new, more effective treatments. Researchers still aren't sure how Covid-19 is killing us. New York Times columnist Charlie Warzel listed 48 "known unknowns" about the virus, and it's not a comprehensive list. That's a field ripe for the application of AI technologies; integrating and modeling huge volumes of data can lead to answers.

AI can also point researchers in the direction of effective viral medications and vaccines by digesting volumes of data from similar outbreaks. Recently, Australian researchers created the first AI-developed vaccine, saving years of time and millions of dollars.

When this pandemic is contained, AI technologies will speed the recovery and preparedness of policymakers for the next outbreak. A data-driven response has a better likelihood of preventing or quickly containing such a pandemic. We should strive for that.

About the Author

Tara Holland, Principal, Global Government Practice at SAS Institute. Being a part of the SAS global government team, She is able to bring analytics solutions and expertise from around the world to governments across Canada. Her role is to connect and collaborate the best resources from government clients, partners and all business units within SAS.



AI in Times of Pandemic: Between Hope and Concerns

AI has played a substantial role in assisting public health and the medical community to curb the current pandemic. However, are AI contact tracing apps really useful? Will they harm individual privacy and liberty?

By Hubert Laferrière

This article described the situation at the time of writing (mid-May). Some developments have taken place since then, but the key issues remain.

Over the past weeks, media have been reporting on the development of digital applications that would make it possible to monitor Coronavirus transmission chains by collecting travel and interaction data from individuals using AI technology. These apps, known as peer-to-peer contact tracing apps, are fuelling discussions on the issue of individual privacy and liberty. Attitudes of circumspection and mistrust are displayed regarding the potential effects of these apps when used in the field of public health. Some are raising concerns about potential harmful side effects of this technology on the democratic institutions and culture, including freedom of choice and movement, some are qualifying the apps as "invasive surveillance measures". [1]

AI is Stepping up against Covid-19

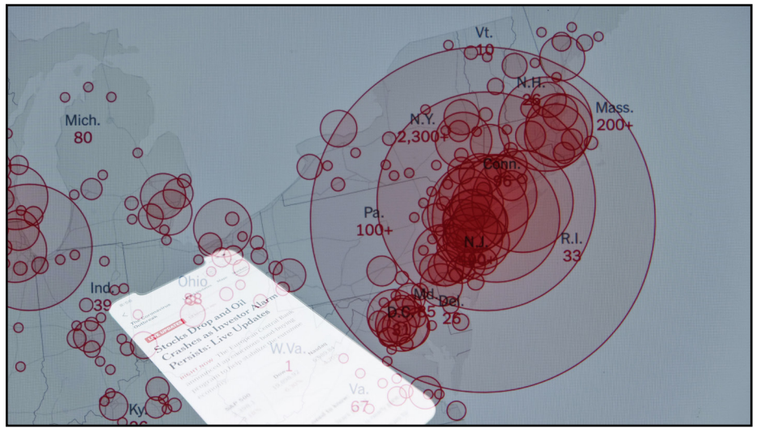
There is however no doubt that AI is and will play a substantial role in assisting the medical community to curb the current pandemic and find a cure. For example, an AI technology has identified preliminary signs of the pandemic. On December 30, 2019, BlueDot, a Canadian start-up firm, using machine learning (ML) to monitor outbreaks of infectious diseases around the world, alerted various governments, hospitals, and businesses to an unusual bump in pneumonia cases in Wuhan, China. Nine days later, the World Health Organization officially flagged what we have all come to know as COVID-19.[2]

Last year, the U.S. National Institutes of Health (NIH) confirmed that "Machine Intelligence (MI) is rapidly becoming an important approach across biomedical discovery, clinical research, medical diagnostics/devices, and precision medicine."[3] Experts are now using MI to study the SARS-COV-2, test potential treatments, diagnose individuals, analyze the public health impacts, detect early signs of virus explosion, monitor public safety, and more.

Last February, scientists at the Massachusetts Institute of Technology and Harvard used a deep learning neural network to identify a molecular compound different from most antibiotics, that was discovered from the Drug Repurposing Hub and demonstrated it's effectiveness against a bacterium that no existing drug can eliminate. The discovery has implications for fighting the SARS-CoV-2: finding a novel use for an existing compound, and using neural nets in place of familiar chemical definitions.[4]

AI has played a substantial role in assisting public health and the medical community.

1 Fox Cahn, A., Domino, A., (2020-04-06), Tracking Everyone's Whereabouts Won't Stop COVID-19, Fast Company.
2 Heaven, W.D., (2020-03-12), AI could help with the next pandemic—but not with this one, MIT Technology Review,
3 Cutillo, C.M., Sharma, K.R., et alii, (2020-03-26), Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency., Nature Research Journal- NPJ Digital Medecine.



Scientists from Carnegie Mellon considered an early version of COVID Voice Detector, an app that would analyze a patient's voice to detect an infection. There is an "AI powered stethoscope" (Eko) helping doctors to directly treat patients with "wireless auscultation" of the heart and/or lungs. This is helpful when practitioners wear significant amounts of protective gear. Medicago, a Québec city-based private firm, is already working on a vaccine using ML algorithms to measure 10,000 to 15,000 molecules. In a nutshell, AI initiatives and activities are buzzing in life sciences and technology areas.

AI and COVID -19 Contact Tracing

The use of contact tracing applications aims at supporting the public health authorities by accelerating their efforts to target the transmission chain of the Coronavirus and individuals by assisting them to better manage social distancing, particularly useful with the lift of lockdown measures.

MIT Technological Review has documented 25 individual, significant automated contact tracing efforts globally (as of May 7).[5] Some countries, mostly using proposed solutions developed by Google and Apple, have indeed deployed the applications (France, United Kingdom, Germany, China, Taiwan, Singapore, Australia, New-Zealand, Switzerland, Israel, etc.). Interestingly, these apps are essentially government mandated, whether their use is imposed or voluntary.

Tracing applications seem to trigger a high level of concerns regarding privacy and personal information protection including the management of data. Most of the countries have stressed the importance of maximizing measures to protect personal information. However, people are concerned with the government tracking them as well. In France, the Parliament approved the tracking app which was launched on June 2. The responsible minister indicated that StopCovid is a tool at the service of citizens which does not represent a threat to their freedoms, emphasizing its "voluntary use", "pseudonymization", "temporary" duration and "transparency".[6]

Here in Canada, researchers at the Montreal Institute for Learning Algorithms (MILA) are designing a smartphone app for contact tracing, COVI is a peer-to-peer AI-based tracing app. "This could be an important ingredient in the fight," said Joshua Bengio, Scientific Director at MILA, in an interview in Science Business.[7] But he recognized the use of the COVI to collect information may put in contradiction intersecting people's trajectories and privacy. "The government and the public would not accept an app that doesn't match our cultural expectations. I don't think the government will force people to use something like this, so it has to be something people accept".

At the provincial and federal levels, discussions are underway. The government of Alberta encourages its citizens to use the AB TraceTogether application, first used by the Singapore authorities. Some

4 Stokes, J.M., Yand. K., et alii, (2020 -02-20), A Deep Learning Approach to Antibiotic Discovery, Cell, Volume 180, Issue 4.
5 O'Neill, P.H., Ryan-Mosley, T., Johnson, B. (2020-05-07), A flood of coronavirus apps are tracking us. Now it's time to keep track of them, MIT Technology Review.
6 Laurent, Corinne (2020-05-28), Traçage numérique : l'application StopCovid validée par le Parlement, La Croix.
7 Kelly. É. (2020-04-02). Science in overdrive: Researchers are inspired and exhausted by scale of COVID-19 challenge, Science Business.

ARTIFICIAL INTELLIGENCE

Montreal-based MILA, an AI research institute, is ready to offer governments across Canada a contact tracing App.

provinces may follow Alberta; the Québec government seems not ready to deploy such an application. Prime Minister Justin Trudeau says the federal government will soon be recommending to Canadians to download a COVID tracing app.[8]

Are AI Contact Tracing Apps Really Useful?

Singapore's experience suggests the app is not a panacea and must first serve as a support for professionals assigned to the crucial contagion tracing task. A developer of the TraceTogether application has warned health authorities against any form of "technological triumphalism".[9]

In Iceland, where the app penetration rate in the population is among the highest in the world, a senior figure in the country's COVID-19 response indicated the real impact of the app has been small, when compared to manual tracing techniques like phone calls. In this context, the technology appears to be close to useless and not the game changer everyone expected.[10]

Chief public health officer Dr. Theresa Tam expressed concerns about the mistakes that could be made by contact tracing apps if not developed properly, particularly the chance of false positives, thus alarming unnecessarily a whole bunch of people.[11]

Contact tracing apps will not be deployed in Belgium, where authorities prefer traditional methods derived from well-established public health practices. In the U.S., mandatory application-based contact tracing appears not to be on the table.

"Une fausse bonne idée"?[12]

Jason Millar, Canada Research Chair in the Ethical Engineering of Robotics and AI at the University of Ottawa, asserted in Policy Options[13] that a well-intentioned but unproven application might not produce the desired outcomes; it could make things worse. In addition to significant privacy concerns, Millar's key concern is that contact tracing apps can reinforce existing social biases, thus stigmatizing locations and communities even if anonymizing measures are stronger. The apps still provide users with information about high-risk locations and therefore stigmatization may not be stopped. In South Korea, some members of the LGBT community provided false information for fear of being personally identified with traditional tracing practices.

In commenting on the Alberta TraceTogether app deployment for Radio-Canada, Julie Paquette, professor of ethics at Saint-Paul University, has recognized that the Alberta version app is less intrusive than the Singapore original, but risks about privacy still exist. Will the data be removed from the application and erased since they were collected for a specific purpose?[14] In addition to privacy, her underlying question appears crucial: What is the likelihood that the app will not be used for other purposes? Paquette's main concern is about the potential ongoing use of the apps outside the emergency context, triggering a shift in purpose in the intent: Will these exceptional measures shift to a new surveillance system and to a more general acceptance of this type of surveillance? Generalizing the use of tracing apps could lead to a fall sense of security and open the door to the development of a surveillance culture, as surveillance could become the norm and an integral component of social relations.

In reaction to MILA's COVI app, Jocelyn Maclure, President of the Québec Commission on ethics in science and technology, expressed his concerns about technological solutions applied to complex and persistent social problems, an approach he labelled as "techno-solutionism". At the same time, he outlined that critics of the solutions tend to focus exclusively on the risks associated with the use of technology, setting aside some benefits or improvements generated by the solutions. He recognized that more sophisticated technological tools could be beneficial for the public health authorities but he called for robust measures to mitigate the ethical risks inherent in the deployment of such technologies: if it is up to experts in epidemiology, infectious diseases, and public health to define the needs in this area, it is up to independent ethics experts to address the ethical risks.[15]

In fact, we understand little about how the tracing apps could affect society and human attitudes. While there is no doubt that AI contributes and will contribute to the efforts to tackle current issues generated by the pandemic, one can hope that some boundaries will be drawn during and after this pandemic nightmare.

About the Author

Hubert Laferrivère was the Director of the Advanced Analytics Solution Centre (A2SC) at Immigration, Refugees and Citizenship Canada. He had established the A2SC for the Department of IRCC and led a major transformative project where advanced analytics and machine learning were used to augment and automate decision-making for key business processes.

8 Aiello, Rachel,(2020-05-22), PM Trudeau: Feds will soon 'strongly' recommend contact tracing app, CTV News.
9 Péloquin, T., Thibodeau, M., (2020-04-20), Géolocaliser la pandémie, une fausse bonne idée, La Presse.
10 Johnson, B., (2020-05-11), Nearly 40% of Icelanders are using a covid app—and it hasn't helped much MIT Technology Review.
11 Boynton, John, (2020-05-03), Consent for coronavirus tracing apps must be 'meaningful', Canada's privacy watchdog says, Global News.
12 Title borrowed from Péloquin, T., Thibodeau, M., op.cit.
13 Millar, J. (2020-04-15), Five ways a COVID-10 contact-tracing app could make things worse, Policy Options.
14 Seker, F., (2020-05-09), L'application AB TraceTogether téléchargée 140 000 fois depuis son lancement, Radio-Canada.
15 Maclure, Jocelyn, (2020-05-19), Entre le technosolutionnisme et le catastrophisme, La Presse

Expanding Big Data Measures and Machine Learning to Manage Pandemics

By Betty Ann M. Turpin, Ph.D.

The world has pulled together remarkably well to combat the outbreak. To understand this pandemic, nations have turned to public health and infectious disease experts. As we struggle to gain control, a primary and understandable focus has been on the direct causal effects related to contracting and transmitting the virus. Based on information releases and expert opinions, there is a high degree of consensus that nations will have to manage this outbreak for at least 18-24 months in the hopes of it subsiding and stabilizing, as researchers try to find a cure.

Globally, countries are reporting data such as the number of cases, recovery and death volumes, and even indirect data such as flight travel patterns. However, each country is using its own methodology for identification, data gathering, analysis, and reporting. Time has not permitted the development of a global data cloud for data sharing. Big data sources would yield valuable insights but will require global harmonization and standardization of the data methodology. A good example of cooperation can be found in the Canadian Public Health Surveillance initiative that collects information related to 29 specific diseases and threats to public health in sentinel systems to track and forecast health events and determinants.[1] Currently Canada, through the Public Health Agency of Canada, is coordinating with the provinces and territories, and using the current flu tracking system, to track and update on the virus. The measures are comprised of 7 components[2]: geographic spread, laboratory confirmed detections, syndromic, outbreak, severe outcomes, strain characterization and antiviral resistance testing, and vaccine monitoring. These components are lagging indicators/ measures that are also used to generate forecasting data.

Some countries have begun to consider the secondary impacts of COVID-19, such as deaths due to transplant surgeries not being possible, lower than average blood supply so that lives cannot be saved, starvation rates due to increased lack of access to food, to name just a few. However, capturing this information is not easy and is certainly resource intensive, and again varies across nations. These are consequences of the disease occurrence but suggest variables that could be used in prediction and management of pandemics.

So, if we are to look at the "health" of a nation, during the peak and recovery periods, and even the prevention of COVID-19 re-occurrence, we need to know why specific interventions work (i.e., what are the underlying mechanisms) and for whom and under what circumstances produce the greatest benefits. To do this, we must look beyond direct causal effects and lagging indicators and consider an indirect-relationship model, one that includes mediating and moderating variables that influence uncontrolled and controlled variables. Using an indirect-relationship model can improve the understanding of the pandemic and thus its management.

As shown in Figure 1, a mediator as an intervening variable that is affected by the intervention (line A), which in

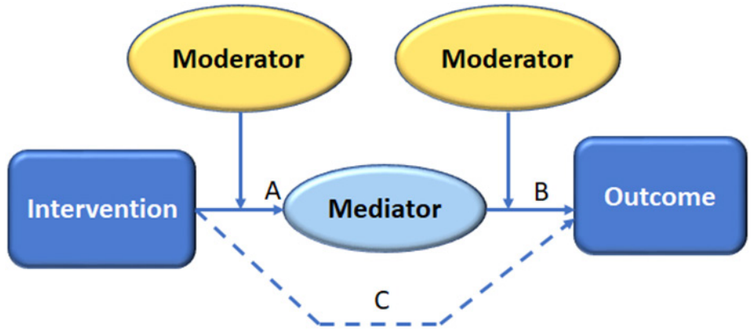


Figure 1: An Indirect Casual Model[3]

1 https://www.canada.ca/en/public-health/services/surveillance.html.
2 https://www.canada.ca/en/public-health/services/diseases/flu-influenza/influenza-surveillance/about-fluwatch.html
3 Donaldson, S.I. Mediator and Moderator Analysis in Program Development. In Sussman (Ed.), Handbook of Program Development for Health Behaviour Research and Practice (pp. 470-500).

turn affects the outcome that is desired (line B). Conceptually, mediator models assume that the intervention causes changes in the mediator variable, and the mediator variable then causes changes in the outcome. But other factors (other variables) become apparent. Moderators are qualitative (e.g., sex, race, class) or quantitative (e.g., the level of reward) variables that affect the direction and/or strength of the relation and relate to personal situations, psychological profiles, and work differently for different types of people.

A moderator variable (vertical arrows) affects the direction (+/-) or strength of the relationship between the intervention and the mediator, or the mediator and the outcome. Line C is an unknown variable that may/not affect the outcome.

For example, consider the current-day interventions adopted in most countries. Physical distancing is the intervention, as the assumption has been this intervention will lower transmission rates. The mediator is social behavior. Thus, the desired outcome is the reduced spread of the virus. So, to identify moderators, we ask the questions: 1) what might affect the relationship between the intervention and the mediator, 2) what might affect the relationship between the mediator and outcome. In terms of the first, age is a factor and may be a stronger relationship for older adults (50+ years) and less strong for young adults, access (to stores, parks, etc.) and wearing facial masks are the second moderators. But at this point in time, because we do not collect or integrate this information into the data analysis, we can only assume they contribute to a positive outcome.

Importantly there can also be/unknown/unanticipated variables that impact (line C) the outcome. Such as defiance of the distancing recommendation, refusal to use a tracking/tracing app on one's phone, etc. To identify variable C, it would be helpful to consider research work related to unintended consequences.[4] Additionally, this indirect casual model shows just one outcome, but there be many, and so for each outcome, this model should be applied.

We know this intervention and variables have already been applied in most countries and is working.

We also know that many surveillance systems track moderator A-type variables, and most of these are lagging indicators. The challenge to expanding surveillance is to identify other moderator B variables and then build these into the AI system. The focus should be on interventions that yield the best and strongest effects, that is, can moderate the outcome. A second challenge is knowing what these variables are/may be. We do not always know, so applying the unintended consequences approach can help.

With the synchronicity of multiple data sets and moderator variables, machine learning analysis capability can enhance predicative potential of pandemic occurrences and insight into what effective interventions can predict, prevent, and manage the viruses.

About the Author

Betty Ann M. Turpin, Ph.D., C.E., President of Turpin Consultants Inc., is a freelance management consultant, practicing for over 25 years, has also worked in the federal government, in healthcare institutions, and as a university lecturer.

Her career focus is performance measurement, data analytics, evaluation, and research. She is a certified evaluator and coach.

4 Morrell, Jonathan A. (2005). Why Are There Unintended Consequences of Program Action, and What Are the Implications for Doing Evaluation? American Journal of Evaluation.

Mindset Versus Datasets

More than having the latest tools or following the latest trends, the use of analytics in government is about becoming a more agile organization that can easily pivot when necessary. Knowing when and how to turn around is therefore essential to transform our practices. To know when a pivot is needed, managers must have a way to become aware of changes in their environment. These changes are usually signaled with small variations in daily trends. Capturing this is possible by tracking a vast number of additional details, that is data, in daily operations.

Ask the Right Question

Data can reveal much about such changes and the associations between variables, highlight valuable information, discover patterns, trends and signals that are not generally visible to managers. Automated data capture, known as data mining, has been available since the last quarter of the 20th century. There is a vast amount of resources associated with data mining. The aspect that I have found missing in most of them, is how to ask the right questions.

Questions are a way to test our own assumptions. Questions are anchored in the framework from which they emerge. That is why in research the essential first step is to write down an inquiry question. These questions must emerge from the extant literature, otherwise they are considered atheoretical.

In organizations, questions emerge during our daily interactions with our colleagues. A comment made by one, becomes an insight by another. The only way that can happen is by becoming detached from our routines. At some point, looking at the same data, we become "blind" to seeing more than we already think we see. It is a comment by someone else, looking at the data, that allow us to "see" the data anew.

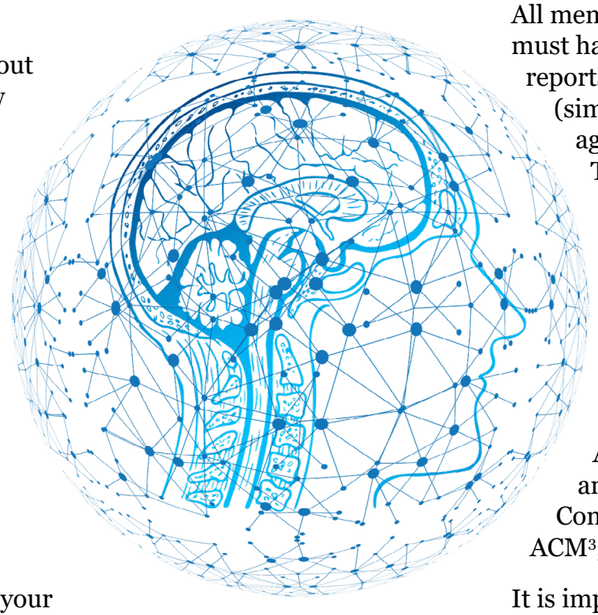
What are the strategies that we can develop to be sensitive and

responsive to such insights into our daily performance? For one, remember that having a diverse teamwork is essential when working in Analytics projects. Second, benchmark from what others are doing. Three, document what you are doing. Finally, schedule regular meetings to evaluate your progress. Asking the right questions requires good team synergy.

Team Importance

Much has been said about teamwork. We all know teamwork is valuable and I would like to talk about the type of teams that will help organizations move forward on their Analytics projects. To really allow your team to move forward, the team needs to include different "voices". In particular, the team must include someone who understands your business needs, knows your organization's mission, shares your organization's vision and business model. This voice will ask questions regarding goals and objectives, i.e., how is this project compatible with our vision? How will this project help us achieve our mission? How can we use the data to measure our Key Performance Indicators? How will this data help us solve our problems? How can it be used to pursue some new opportunities? How can it help us fulfill directives and regulations?

We also need a 'voice' that understands the technology available and if it is proficient in manipulating data into information. This voice will guide the team to capitalize on the invested tools and use those tools effectively and efficiently. This voice also will know which data to use and how to gather it. The main goal



of this voice is to empower the team to use the tools the organization has and/or evaluate if additional tools are needed.

The third voice is that of a data scientist, someone with a good background in statistics and analytics, research design, data mining techniques, programming languages, including AI and machine learning. This voice will participate in generating report. Therefore, it is expected that communication and visualization skills are excellent. It is also expected the voice can talk to the organizational side of the analysis. That is, make clear to the team and management what the data are saying. In sum, this voice will enhance the creative component of the team.

All members of the team must have the time to read reports on what other (similar) businesses/agencies are doing. There are specialized magazines that monthly or quarterly present a collection of best practices. Besides Analytics in Government Quarterly, you can read INFORMS' Analytics Magazine[1] and ORMS Today[2] , Communications of the ACM[3], among others.

It is important to know what others in the industry are doing. Look for stories in newspapers, trade press or mainstream media. Technology vendors sometimes post in their sites customer stories that showcase what they are doing. If possible, attend conferences and participate in events where advances in the use of analytics in your line of work are presented.

1 https://pubsonline.informs.org/magazine/analytics
2 https://pubsonline.informs.org/magazine/orms-today
3 https://dl.acm.org/magazine/cacm

Keep an Analytics Project Diary

Use a diary to keep track of the project's progress. This diary is a narrative of the project, i.e., who participated, what was agreed, what questions were asked, what data was suggested, who was responsible to gather, clean and transform the data, who was responsible for its analysis, who was responsible for running models, who is responsible for evaluating them, etc. The essence of this dairy is about ensuring team insights are captured and shared. In addition, it can reflect the "thinking" behind the decisions and data analysis that the team undertakes.

Include in the diary any pertinent facts, arranged in chronological order, how questions were answered, how problems were solved, how useful it turned out. Include a summary of key points of a discussion but avoid unnecessary details that are not relevant to the project. This diary will be the team's memory.

Monitor Progress

Schedule progress meetings along the lifetime of the project. Review if the scheduled tasks were successfully performed, document advances and set clear and measurable objectives in between meetings.

Each of these issues is a way to help decision-makers to expand the team's mindset by allowing the dataset to set direction. This is how fact-based decision-making is implemented. Sometimes we have an idea on what is going on in the organization, and when the data shows that is the case, a level of confidence evolves that can impact on making decisions. We have to learn to trust the data. To do so, it is necessary to trust the team, know what best practices are in your field, have a historical recount of the project, and the assurance that in the evaluation meetings, the established criteria was met. Analytics is a process, but it is definitely not devoid of effective of team insight and expertise. To know what you want from your data, you need to ask the right questions and more often than not team members contribute to this, foster a collaboration between business and technology expertise, and let the dataset change your mindset in a way that moves your business/mandate forward, as effectively as possible.

About the Author

Alex Ramirez, Ph.D. is an Associate Professor in Information Systems at the Sprott School of Business – Carleton University. He has worked in education for over 30 years. He obtained his Ph.D. from the Molson School of Business – Concordia University in Montreal, a master's degree from Syracuse University in the U.S. and a BSc. High Honours from ITESM, Mexico's top private university.

Monitoring a flower growth stages.

By Rami Abielmona, Ph.D.

Analytical Emergency Management

Imagine a mass gathering for a planned protest is being held in a major North American city near downtown buildings and the surrounding streets. Both municipal and law enforcement planning has been completed, and officers have been briefed and assigned positions. A manned helicopter has been stationed for overhead surveillance. Road closures along the planned route are in place and the population has been notified of the protest to control traffic and bystanders. A command post is established and is receiving information from officer radios, network chat as well as video and imagery from surveillance aircraft and officers in the field. Incoming data is digitized, analyzed, and displayed on the command centre's Artificial Intelligence and Machine Learning (AI/ML) enabled Common Operating Picture (COP). Monitoring of open source, Twitter, Instagram and weather reports is established, while the location and the open source messages are displayed on the COP.

At a certain moment, a disturbance occurs in the crowd caused by a person (e.g., running through the crowd) or a group intending to infiltrate the mass gathering in an attempt to transform the demonstration into a violent event. What transpires next is quite disorderly: the situation devolves rapidly when surprise transforms into chaos and a vast amount of data begins to be transmitted from multiple sources and soon overwhelms the command centre's ability to weed out the noise and focus on what is relevant. Important questions have be to be answered in a rapid and focused manner: which of the data points carry useful information, what can be safely ignored for now, what is the most optimal way of redeploying assets and what information services can be called upon to deescalate the situation and maintain order and control?

Using a data-oriented architecture and Big Data Analytics (BDA), the perpetrators can be detected through image and video analytics from the pre-positioned static Electro-Optical (EO) / Infrared (IR) camera sensors as well as Full Motion Video (FMV) and thermal sensors on board aerial assets (e.g., helicopters, UAVs). These are processed using image and video analytics to produce indicator tracks focusing on the perpetrators and helping commanders weed out the noise. The suspects can then be tracked and classified using decision-support system software that provides ML-based anomaly detection identifying the suspects as being anomalous to the background scene (i.e., the mass demonstration). This subject anomaly can be due to their speed, shape, movement, behaviour and overall pattern-of-life (POL) which has been detected as being substantially different from the norm (e.g., a group of people moving in the opposite direction of the crowd, a person intersecting through the demonstration, a group of people congregating together and moving within the crowd).

Retroactive Analytical Scheme

From an analytics standpoint, emergency management requires the adoption of all four known types of analytics (described below), in addition to retroactive analytics as a fifth analytic type. A retroactive analytical scheme allows an autonomous system to learn over time, solving the short-term view of reactive systems, and to respond in real-time to world events, solving the calculative rationality of proactive systems. This can also be used to provide recommendations to better prepare for future situations by modeling and simulating the possibilities in a retroactive manner. This affords for better intelligence assessment and therefore mission planning, through understanding what could have happened, and how it could have unfolded within an emergency situation. Here is a brief description of each analytic below:

1. **Descriptive Analytics:** answers the question of "What happened?" by extracting spatiotemporal information and fusing (i.e., correlating) multiple data sets and sources to provide insights into the past. This will allow for exploration into the past and discovery of exactly how these emergency events and situations transpired;
2. **Diagnostic Analytics:** answers the question of "Why something happened?" by data-drilling and POL recognition to provide insights into the present. This will allow authorities

Black Lives Matter protest in Montreal.

to explore the present and identify why emergency events and situations evolve in a certain manner;

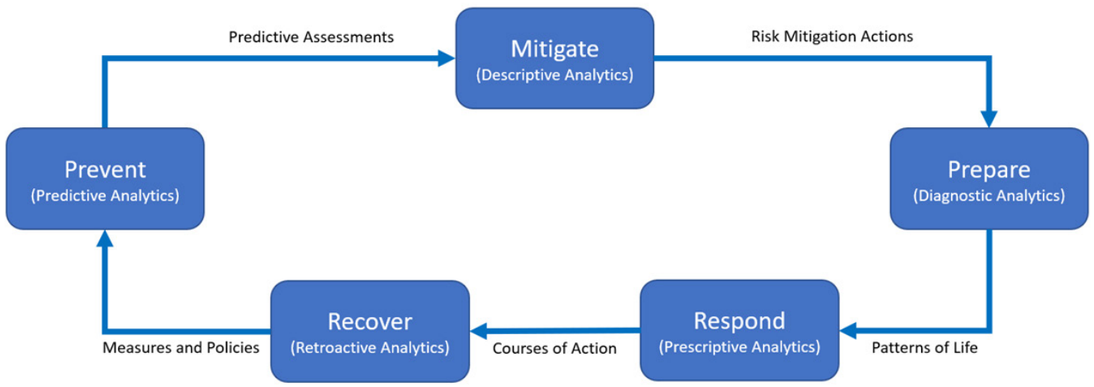
3. **Predictive Analytics:** predicts "hat will most likely happen" by learning past and current tendencies and forecasting future trends (i.e., through models of potential outcomes). This will strengthen the preparedness capabilities by allowing authorities to explore the future and identify how emergency events and situations most likely will evolve;

4. **Prescriptive Analytics:** answers the question of "What should be done?" by running multi-objective optimization (MOO) within a modeling and simulation environment to provide recommended responses to better deal with the future. This will strengthen the mission-critical capabilities by allowing authorities to explore the various courses of action while attempting to mitigate future emergency risks or take advantage of promising trends; and

5. **Retroactive Analytics** answers the question of "What could have been done?" by scenario and mission planning through modeling and simulation, in order to provide recommendations to better prepare for similar emergency situations in the future. This will strengthen the decision-making capabilities by allowing the authorities to retroactively explore the past and figure out how to better deal with emergency situations in the future.

Analytical Emergency Management

Authorities need to make accurate, timely and rapid decisions at different stages of the emergency management spectrum within a complex and information intensive environment. This five-step process, called Analytical Emergency Management (AEM), typically involves various trends, transitions and tempos while mapping quite suitably to the above BDA. As we can see from the below figure representing the closed loop AEM cycle, officials initially need to be able to prevent emergency situations by predicting them through historical training of models using predictive analytics. If an event actually transpires or a situation develops, officials need to be able to mitigate the situation by rapidly detecting it and making rapid decisions for adapting to, eliminating and reducing the risks to the human population using descriptive analytics. As the situation unfolds, a near real-time, accurate picture is required as to the cues and patterns required to prepare to respond to the emergency and manage its consequences through diagnostic analytics. Subsequently, as the emergency situation evolves (devolves), officials need to actively respond to the situation and support decisions about how to allocate limited and/or secure additional resources are to be taken through prescriptive analytics. Finally, when the situation comes to an end, mitigation and response measures, as well as policies, are analyzed to quickly recover from the emergency and restore conditions to an acceptable level, all the while analyzing and implementing lessons learned and improving first responder training through retroactive analytics.



The Analytical Emergency Management framework

AI-Powered AEM

AI/ML finds itself embedded throughout the AEM cycle. Whether it is the optimal placement and rapid redeployment of assets (such as law enforcement vehicles) and sensors (such as first responder imagery), as well as the real-time exploitation of data sources (such as social media feeds), or the automated path planning for all platforms and on-board sensors as well as the AI-based response generation for the various first responder assets (e.g., fire, EMS, etc.) provides an optimal Course of Action (COA) for the emergency response, AI/ML is a driving factor for each facet of AEM.

An AI-powered AEM cycle within an emergency management environment is essential for the efficient protection and overall safety of the population and first responders alike. The smart COP provides an efficient multi-sensor solution for the highly demanding situational assessment

performed during the AEM cycle. The persistent real-time optimized situational picture provides the decision maker with a host of intelligence products with which to make an informed decision in a time critical fashion for an effective mitigation of a rapidly developing situation. AI/ML approaches are thus required to cope with the vast quantity of sensor data that are to be exploited and to provide optimized recommendations to first responders. This concept also allows for the integration of incidence policies, emerging communications sources, radio and other location-based services, open/closed source and specific responder communication services and, based on all of this information, optimizes, presents risk factors and recommends COAs to incidence commanders through the AI-powered COP.

AEM provides the opportunity to integrate AI/ML into a multi-layer analytical framework, moving away from user-centric solutions to ones that are more autonomous in nature and deployable across a variety of responder domains (e.g., fire, police, search-and-rescue, health, etc.). It is obvious that, with the recent increase in mass public disturbances and emergencies coupled with the advent of BDA and the proliferation of AI/ML capabilities, emergency management has to evolve, in the very near future, to become more analytical, asymmetric and adaptive. Our safety and security would greatly benefit from that.

About the Author

Rami Abielmona, Ph.D., P.Eng., SMIEE, is the Vice President of Research & Engineering at Larus Technologies Corporation and is responsible for all research and development of software, hardware and products, as well as the management and direction of the research team. He was one of the recipients of the Ottawa Business Journal (OBJ) Top 40 Under 40 Award in 2011, was named as the Part-Time Professor of the Year at both the Faculty of Engineering and the University of Ottawa in 2012 and was a recipient of the NSERC Synergy Award for Innovation (for Small and Medium-Sized Companies) in 2016.

ANALYTICS IN GOVERNMENT QUARTERLY