HARNESSING COMMERCIAL DATA FOR PUBLIC GOOD

Design reflections on the Pandemic Intelligence: Trusted Exchange Network (PI:TEN)

A Solutions Workshop report from the Schwartz Reisman Institute for Technology and Society at the University of Toronto in collaboration with MaRS Discovery District and the Toronto Region Board of Trade

UNIVERSITY OF TORONTO

SCHWARTZ REISMAN INSTITUTE FOR TECHNOLOGY AND SOCIETY

MaRS

TORONTO REGION BOARD OF TRADE

CONTENTS

REPORT AUTHORS:

Gillian K. Hadfield, Director, Schwartz Reisman Institute for Technology and Society, Professor of Law and Professor of Strategic Management, University of Toronto

Jovana Jankovic, Research and Communications Specialist, Schwartz Reisman Institute for Technology and Society, University of Toronto

Jamison Steeve, Senior Advisor, Policy, Strategy, and Solutions, Schwartz Reisman Institute for Technology and Society, University of Toronto

Benjamin Wald, Postdoctoral Fellow, Schwartz Reisman Institute for Technology and Society, University of Toronto

Date of publication:
May 13, 2021

Workshop design by
Quinn Davidson

Photography by
Jenna Muirhead

Production by
Daniel Browne

1. EXECUTIVE SUMMARY

Can we use our data footprints as a new tool for combatting COVID-19?

The **MaRS Discovery District (MaRS)**, in collaboration with the **Toronto Region Board of Trade (TRBOT)**, is contemplating creating a tool called the **Pandemic Intelligence: Trusted Exchange Network (PI:TEN)**. PI:TEN aims to leverage already-existing data reflecting mobility and interactions in public settings so that businesses and other organizations can work with public health through secure methods to inform, facilitate, and improve COVID-19 safety measures. While this tool may be helpful in combating the pandemic, and possibly future public health issues, it raises issues of usability and privacy.

With these possibilities and concerns in mind, MaRS and TRBOT reached out to the team at the **Schwartz Reisman Institute for Technology and Society (SRI)** to convene a Solutions Workshop that could wrestle with the challenges and opportunities associated with PI:TEN. Using its subject matter expertise and design thinking approach, SRI was able to convene a diverse and talented group of experts to explore the risks and assess the possibilities associated with PI:TEN during a three-day workshop. While limited in its scope, we at SRI hope this report is useful both to MaRS and TRBOT as they move forward, and to the broader public who wish to engage on these issues.

PI:TEN aims to leverage existing data to inform and improve COVID-19 safety measures.

Overall, workshop participants were impressed with MaRS's desire for input on PI:TEN. Some participants lamented the absence of such a tool when the pandemic began and hoped that this type of instrument could be built for future public health crises. The details of PI:TEN and the design suggestions from the workshop are discussed in detail in the body of this report. Our key design recommendations are:

a. **Make PI:TEN easy for individuals to use.** Whether it's learning from the rollout of the Canadian COVID Alert app or reducing the number of consents and notifications, usability by participants is key to increasing number of participants and allowing PI:TEN to achieve its public health goals. MaRS and TRBOT were encouraged to demonstrate the utility of the proposed system by using synthetic data in an effort to show that PI:TEN is beneficial and safe. With respect to consent, PI:TEN should consider ways in which a single consent to use information could be obtained from participants before a positive COVID-19 result activates contact tracing protocol.

b. **Ensure obstacles to data partner participation are minimized.** Even if PI:TEN is built, data partners may not participate. Efforts must be made to reduce the associated costs, labour, and risks that data partners would incur as a result of joining this effort. The desire to do public good cannot be the only incentive. Rather than placing all the onus on data partners, PI:TEN could consider centralizing the collection and processing of data. This would obviously require significant oversight and governance.

c. **Engage public health.** For this project to be successful, the information gathered by PI:TEN must be provided to public health authorities (both at municipal and provincial levels) in a form, time, and manner that will facilitate timely decisions and help experts battle COVID-19. PI:TEN's plan to create "exposure network graphs" would produce data linking one exposure event with another. As each event would contain the time and location of exposure, as well as the number of people exposed, this information could be helpful in enabling public health to better understand, track, and predict the spread of COVID-19.

d. **Detail how PI:TEN will comply with legal frameworks for data collection and use.** The issues of oversight, consent, and de-identification are central for PI:TEN's success. Using a novel approach to governance and information, PI:TEN would be wise to work with government and privacy commissioners to gain clarity around the oversight and legal obligations required to move forward and gain public trust.

e. **Design technical aspects of the project for maximum efficiency.** Decisions made at the outset by MaRS should be done in a way that maximizes the speed, efficiency, and privacy of information. For example, facilitating database linkages through the creation of unique trace IDs in advance of a positive COVID-19 case could reduce time needed for effective contact tracing.

The team at SRI would like to thank MaRS and TRBOT for asking us to lead this workshop, as well as all the participants for their time and intellectual generosity. In addition, MaRS wishes to thank the Ivey Foundation for their generous support towards this initiative. We hope that this report can contribute to ongoing, important thinking around policy relating to public health responses to the ongoing pandemic, as well as broader questions around harnessing commercial data for public good.

2. INTRODUCTION

Are we using all the tools, technologies, and data available at our disposal to support public health in the fight against COVID-19?

As a third wave of COVID-19 washes across the Greater Toronto Area, citizens, businesses, and governments are struggling for answers as to what to do next. While hope can be found in the increasing availability of vaccines, the reality of another lockdown is hitting weary Ontarians hard. In the absence of detailed data and information, decision makers and individuals alike are left with blunt instruments of protection as their only response to ensure safety: extreme social distancing, self-isolation, and stay-at-home orders.

Since the pandemic began, some members of the public and public health experts have called for an increase in contact tracing, more detailed information on the spread of the virus, and an increased use of technology to battle COVID-19. There is a belief amongst some that restrictions and lockdowns could be more granular and controlled, facilitating safer movement in high-density locations through improved data access, information sharing, and faster contact tracing. Countries around the world explored the utility of contact tracing apps, including here in Canada with the COVID Alert app. These efforts have proven unsuccessful in our country with minimal uptake of the app, leaving the desire for better information and tracing unfulfilled.

In an effort to fill the gap, MaRS is collaborating with TRBOT to create a tool called the Pandemic Intelligence: Trusted Exchange Network (PI:TEN). PI:TEN aims to leverage already-existing data reflecting mobility and interactions in public settings so that businesses and other organizations can work with Toronto Public Health through secure methods to inform, facilitate, and improve COVID-19 safety measures. If the project is implemented successfully, proponents hope that PI:TEN can facilitate targeted re-openings across Toronto's Financial District Pilot Zone, with the potential for use in other locations.

A complex project comprising novel technological and legal thinking, as well as vital multi-stakeholder engagement, PI:TEN aims to leverage existing data—collected, for example, when someone uses a loyalty, transit, or building access card in a public setting—to achieve a public good without sacrificing privacy or transparency. This approach is laudable, but it raises a number of questions. There are risks associated with issues regarding consent, usability, appropriateness, governance, and technology that require careful consideration and deliberation. That is why MaRS and TRBOT came to SRI to convene an SRI Solutions Workshop, bringing together a diverse group of experts to surface and address the issues and possibilities associated with PI:TEN. As an institute working at the intersection of technology and society, we at SRI were intrigued by the opportunity to work with the issues surrounding PI:TEN so that we might explore what is possible, both in ethical and practical terms. We were happy to engage on this project and this report captures the process and output from our workshop.

3. ABOUT PI:TEN

A proposed system to improve contact tracing and exposure alerts in Toronto’s Financial District Pilot Zone—while preserving privacy.

a. What is PI:TEN?

PI:TEN is a proposed system to allow existing data collected by different organizations about people’s location and movements in public settings in Toronto’s Financial District Pilot Zone to be combined and used for contact tracing and exposure notification in a privacy-respecting manner. It is designed to integrate insights from multiple data partners, and to expand as new data partners are incorporated into the program. Examples of potential data partners are public transit agencies, COVID-19 tracing apps used in cafes and restaurants, and building management software. Participants are individuals who choose to share their data in a privacy-protected way with PI:TEN, and who will receive notifications of potential exposures arising from their public interactions in Toronto’s Financial District Pilot Zone.

PI:TEN’s goal is to leverage existing data collection to permit rapid exposure notification and contact tracing that is faster than the spread of the disease.

The current proposal calls for PI:TEN to be operated by a non-profit corporation. MaRS is proposing to launch PI:TEN as a pilot project in the Toronto Financial District that would operate only as long as the Public Health Agency of Canada, Public Health Ontario, and Toronto Public Health consider COVID-19 to be a public health emergency. Once PI:TEN ends, all information collected would be deleted. PI:TEN’s governance model, legal instruments, and technology, however, would remain available for rapid deployment in future public health emergencies. The roll-out of the program would be accompanied by a public education campaign and information sessions. PI:TEN would complete a privacy impact assessment, and update this assessment each time a new data partner is added to the program. MaRS believes that the combination of a sound governance model with appropriate legal instruments will ensure the responsible oversight of data.

b. Why PI:TEN?

Private companies across the Financial District today regularly collect data on customer activities and locations through smartphone apps, customer loyalty programs, transit passes, building access cards, etc. In response to COVID-19, other organizations have begun collecting new data, such as check-ins required by restaurants and cafes and health passes required by schools and daycare centres. This data already exists, managed under existing legal and ethical duties to respect the privacy of Ontarians. While being mindful of the ways data aggregation can affect privacy, PI:TEN is envisioned as a way of putting this data to work for the public good.

The main goals of PI:TEN are:

1. **To leverage existing data collection to permit rapid exposure notification**—in particular, to allow for contact tracing that is faster than the spread of the disease.

2. **To gather insight that will allow public health officials to better understand the spread of COVID-19**, allowing for more effective and more targeted interventions.

3. **To increase confidence in public health measures** so that people are aware of which activities are safe or pose risks, with a view towards facilitating partial re-openings and minimizing restrictions when safe.

c. PI:TEN's entry into the field

Despite the existence and roll-out of vaccines for COVID-19, the team behind PI:TEN believe their proposal can still provide significant value. Several participants in our workshop asked why MaRS was taking this project on now. As vaccines are coming online, some felt it would be better for organizations like MaRS and TRBOT to focus their efforts on public communication, vaccination, and testing.

First, MaRS believes that, even with vaccination rates climbing, there will remain a period of time during which large sections of the population will remain at risk. Their stance is that a rapid roll-out of a system like PI:TEN could aid in the control of COVID-19 while vaccination programs take effect.

Second, despite the emergence of vaccines, there is a chance that COVID-19 will be with us for a long time to come. A lack of vaccine access, especially in developing countries, produces risk of new rounds of infection in Canada. Even where vaccines are available, some people will refuse to use them, and others will be unable to be vaccinated or will be vulnerable to infections despite the vaccine. There is the potential that new variants will be more resistant to the vaccine than current versions of COVID-19. All of these factors raise the possibility that COVID-19 could become a chronic, rather than acute, emergency, with the need to periodically respond to localized outbreaks or new variants. The team behind PI:TEN believes that having a system developed and ready for deployment could be a valuable tool in these situations.

Third, COVID-19 has shown us that the risk of new infectious diseases is likely to be ongoing, and that rapid response is critical to avoiding repeated widespread lockdowns. Several of the workshop participants lamented the fact that this type of system did not exist in March of 2020. As seen by the SARS crisis in 2003, Toronto is a potential hot-spot for new diseases due to the volume of international travel to and from the city. Building the legal, technical, and institutional infrastructure for a program like PI:TEN might leave us better prepared to rapidly and effectively respond to the next potential pandemic.

d. The PI:TEN journey

If the PI:TEN model reviewed at the workshop were to be implemented, it would operate as follows:

1. **Launching PI:TEN.** PI:TEN establishes itself as a single purpose, not-for-profit entity. PI:TEN completes a privacy impact assessment per the guidance of the Information and Privacy Commissioner of Ontario (IPC). The first goal for PI:TEN is to achieve critical mass in the number of data partners required to function as a robust trace and notification tool. Participating data partners reach out to their customers to become participants in PI:TEN. Participants would need to consent to the use of their information for each data partner, and would also select whether they would be notified of a possible COVID-19 exposure by their data partners or by Toronto Public Health. PI:TEN would aid the efforts of data partners through a robust public education campaign.

2. **Activating PI:TEN.** The PI:TEN system is activated when a person living near or working in the Financial District Pilot Zone tests positive for COVID-19. A Toronto Public Health case manager asks a person infected with COVID-19 to provide consent to PI:TEN to trace their steps over the previous 15 days. This would be in addition to Toronto Public Health efforts to assemble an itinerary of the infected person's whereabouts. PI:TEN then works with the infected person and the relevant data partners to create a validated trace request. The data partners search their own databases for the time and locations where the infected person was present, creating a 15-day location itinerary, and then

identify others in their database who have consented to participate in PI:TEN and who may have been exposed to the infected person. These are known as first-order contacts. Data partners then use the itineraries of first-order contacts to search for additional people in their database who have consented to participate in PI:TEN and who may have been exposed. These are known as second-order contacts. People who may have been exposed are notified either by the data partner or Toronto Public Health, depending on their prior notification choice.

3. **Learning from PI:TEN.** Data partners *do not* share the identities of first- and second-order contacts with PI:TEN. They only share the identities of contacts who requested notification by public health with Toronto Public Health. However, each data partner creates an anonymized exposure network graph for each data partner trace code, recording the quantity and nature first- and second-order exposures. This exposure network graph is encrypted and shared with PI:TEN, who then aggregates all of the exposure network graphs from different data partners for each trace code, creating a graph that shows where and when an infected individual exposed others, and how many were exposed. The identity of the individuals exposed is not collected, but the number and location of exposures is. This is then shared with public health, to aid in the understanding and prevention of COVID-19. PI:TEN believes these exposure network graphs would be useful in helping public health understand the spread of COVID-19 and predict potential super-spreader events. The hope is that this tool could move faster than the virus.

All PI:TEN-related data held by data partners and by PI:TEN itself would be purged on a rolling 30-day schedule—e.g. data sent to or generated by PI:TEN that is older than 30 days would be automatically be erased. The exposure network graphs sent to public health would be retained by public health for use in the study and prevention of COVID-19.

For a summary visualization of the proposed system, see Figures 1 and 2 on the next page. For a complete visualization, see the PI:TEN Detailed Systems Map at **bit.ly/2P71KuR**.

LAUNCH OF PI:TEN

Fig. 1

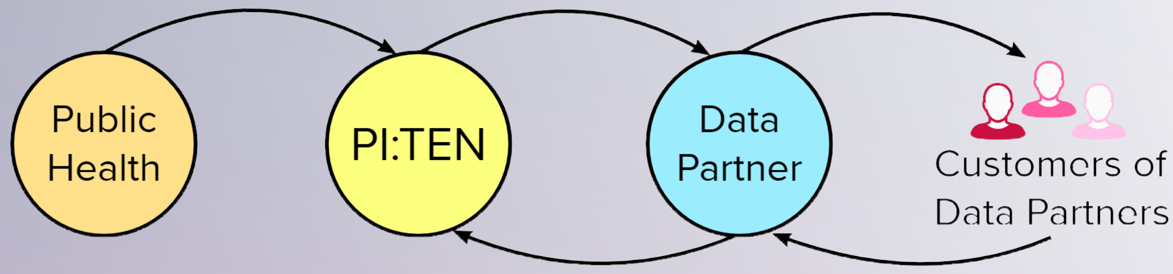


Fig. 2

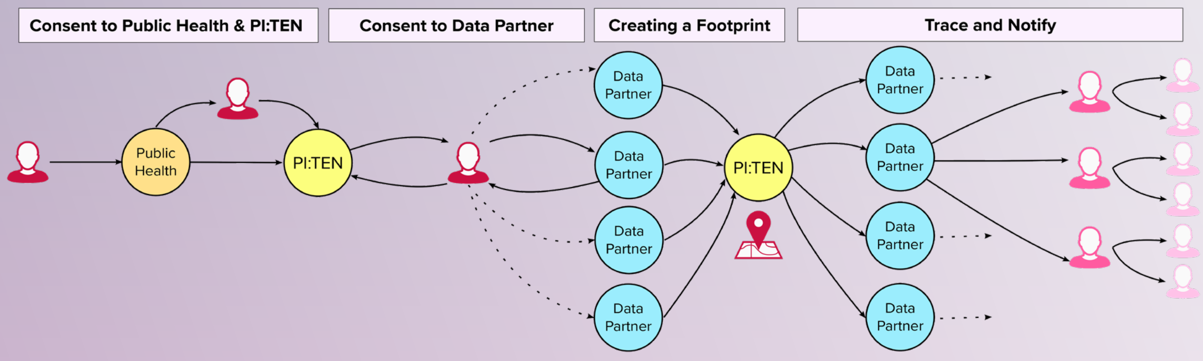
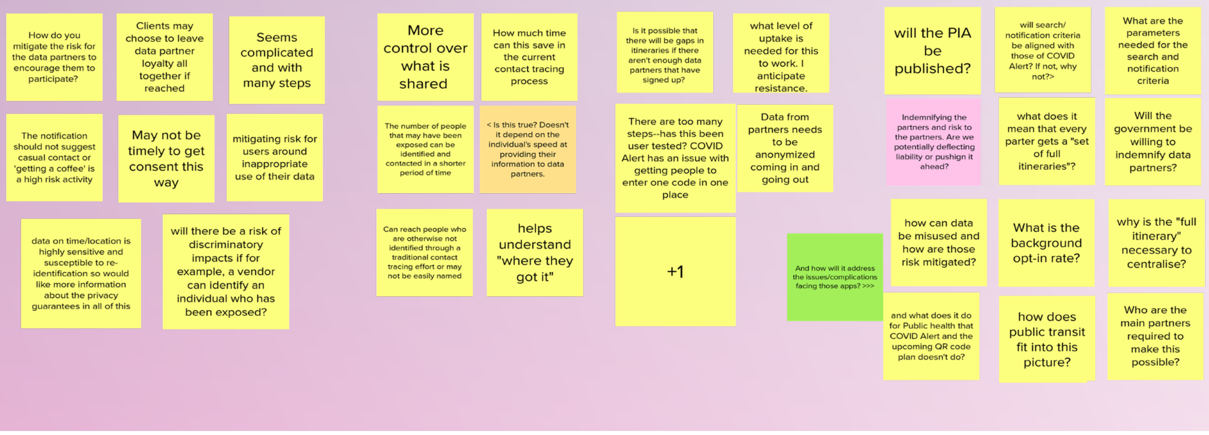


Fig. 3



How do you mitigate the risk for the data partners to encourage them to participate?

Clients may choose to leave data partner loyalty all together if reached

Seems complicated and with many steps

The notification should not suggest casual contact or 'getting a coffee' is a high risk activity

May not be timely to get consent this way

mitigating risk for users around inappropriate use of their data

data on time/location is highly sensitive and susceptible to re-identification so would like more information about the privacy guarantees in all of this

will there be a risk of discriminatory impacts if for example, a vendor can identify an individual who has been exposed?

More control over what is shared

How much time can this save in the current contact tracing process

The number of people that may have been exposed can be identified and contacted in a shorter period of time

< Is this true? Doesn't it depend on the individual's speed at providing their information to data partners.

Can reach people who are otherwise not identified through a traditional contact tracing effort or may not be easily named

helps understand "where they got it"

Is it possible that there will be gaps in itineraries if there aren't enough data partners that have signed up?

what level of uptake is needed for this to work. I anticipate resistance.

There are too many steps--has this been user tested? COVID Alert has an issue with getting people to enter one code in one place

Data from partners needs to be anonymized coming in and going out

+1

And how will it address the issues/complications facing those apps? >>>

will the PIA be published?

will search/ notification criteria be aligned with those of COVID Alert? If not, why not?>

What are the parameters needed for the search and notification criteria

Indemnifying the partners and risk to the partners. Are we potentially deflecting liability or pushign it ahead?

what does it mean that every parter gets a "set of full itineraries"?

Will the government be willing to indemnify data partners?

how can data be misused and how are those risk mitigated?

What is the background opt-in rate?

why is the "full itinerary" necessary to centralise?

and what does it do for Public health that COVID Alert and the upcoming QR code plan doesn't do?

how does public transit fit into this picture?

Who are the main partners required to make this possible?

Fig. 1-2 (top, middle): PI:TEN Overview Systems Map. Detailed Map is available at bit.ly/2P71KuR.
Fig. 3 (bottom): Workshop participants engaged in real-time discussion via online collaboration boards to articulate potential risks, value, functionality, and questions.

4. SRI DESIGN PROCESS

Developing innovative solutions to crucial questions through a human-centred design-thinking lens.

MaRS asked the Schwartz Reisman Institute at the University of Toronto to convene a workshop to explore not only the risks, value, and feasibility of this initiative, but also the potential answers to outstanding questions about the design and implementation of PI:TEN.

SRI Solutions Workshops employ a human-centred design-thinking lens to curate and stimulate conversations between experts that draw out key insights and troubleshoot efficiently. This iterative method clearly defines problem statements and meaningful challenges, captures key insights and potential solutions, and lays the foundation for prototyping ways to meet needs, overcome obstacles, and find new answers to persistent questions. Rather than an open-ended ideation brainstorm, SRI's Solutions Workshops aim to convene experts to develop advice and recommendations based on specific examples and questions, providing a clearer explanation of the problem and possible considerations for a way forward.

SRI Solutions Workshops lay the foundation for prototyping ways to meet needs, overcome obstacles, and find new answers to persistent questions.

SRI conducted a three-day workshop during which we convened 21 participants, including experts in law, computer security, disease modelling, public health, commerce, civil liberties, data analytics, public policy, digital governance, and related fields for a series of structured and guided solutions-building sessions. A list of the participants can be found in the appendix of this document.

Over the three days (two days for virtual discussions, and an intervening day used to consolidate emerging issues and questions), the SRI PI:TEN workshop tackled as many questions, risks, and alternatives as possible within the time available. Our goal was to explore the issues surrounding PI:TEN on Day 1, consolidate insights on Day 2, and explore how we might move forward on solutions on Day 3. What follows is a summary of the results.

Day 1: Exploration

a. Breakout groups

SRI worked with the MaRS team in advance to develop a systems map (discussed above) that outlined in detail the design choices MaRS made in their prototype. Participants received the systems map in advance of our workshop. To begin the workshop, we broke our group of 21 participants into four groups with a mix of expertise in each group. An SRI facilitator joined each group to help them walk through the process and capture the conversation and issues raised. SRI decided to keep MaRS representatives out of these breakout groups so that the conversation would be driven by the participants. Using the systems map as a starting point, the conversation on Day 1 was meant to stress test the PI:TEN concept by identifying the potential risks of the design choices, explore the value that was being created, and test the functionality of the proposal. Looking at PI:TEN in this manner allowed participants to generate both clarifying and conceptual questions to ask MaRS in a structured Q&A session. A sample online collaboration board from the workshop can be found on the previous page.

b. Questions and answers

At the end of our first day together, participants engaged in an extensive Q&A session with representatives from MaRS. Overall, participants seemed intrigued by the concept and promise of PI:TEN, but had pointed questions about some of the design and implementation decisions. Participants also appreciated MaRS's commitment to engage with experts and take advice on board at an early stage of development.

At the end of the session, MaRS was provided with an extensive list of questions to answer in advance in the second day of the workshop. All of the questions and answers are provided in the appendix of this report.

Several themes emerged from the Q&A session.

First, a number of participants asked whether PI:TEN was necessary at this stage of the pandemic, whether MaRS's efforts could be better utilized on more low-tech solutions, and how PI:TEN was different from the COVID Alert app. On this last point, MaRS stated that PI:TEN differentiates itself from the approach taken in the COVID Alert app by aggregating, and putting to use, data already being collected by other organizations (data partners), identifying second-order contacts, and providing useful information to public health that can help trace the spread of COVID-19.

A second area of focus in the Q&A was the level of connectivity between PI:TEN and Toronto Public Health. MaRS clarified that they would want to work directly with Toronto Public Health to ensure that the type and format of the information collected would be of use to Toronto Public Health.

Third, there were many questions raised about the identity, viability, and privacy challenges faced by the data partners. While MaRS does not, as of the writing of this report, have any data partners actively signed on to the PI:TEN initiative, they are hoping to partner with mobility platforms, transit and ride-sharing platforms, and building and office owners. With the amount of location information flowing back to these data partners under the current design, concerns were raised regarding privacy, despite the information being anonymized. Furthermore, given the risks, costs, and labour, and friction being placed on the data partners in the name of meaningful consent, some workshop participants questioned whether or not these partners would be willing to participate. In the absence of strong incentives, many data partners may not be motivated by good corporate citizenship and may opt out, thereby preventing PI:TEN from meeting its adoption goals.

In debates around use of data, it is often easy to identify privacy risks. What is more challenging to articulate is the value and public benefit generated.

Day 2: Consolidation and translation

On the second day of our workshop, the SRI team took the comments, notes, and input from Day 1 and attempted to consolidate and translate the emerging themes, issues, and sticking points raised by participants. Our goal was to read these back to the group on Day 3 so as to create a common understanding of PI:TEN, and set the groundwork for generating possible solutions to some of the challenges raised.

Clear articulation of value is essential to ensure uptake and build trust—especially when initiatives involve private sector data partners who may not always be thought of as protectors of the public interest.

We began by attempting to capture the perceived value created by PI:TEN. SRI has found that in debates around the use of data, it is often easy to comprehend and identify privacy risks. What is often more challenging to articulate is the value and public benefit that is also generated. In the case of PI:TEN, participants and proponents pointed out significant value creation, including:

- Better/faster/more comprehensive contact tracing, to slow COVID-19 transmission;
- Better public health understanding of how COVID-19 spreads, through PI:TEN's exposure network graphs;
- Improved government response with more targeted public health interventions;
- Possibility of greater public willingness to return to work, transit, and the downtown core;
- Greater sense of public contribution to public health solutions;
- Building a useful information infrastructure for future public health crises;
- Building an infrastructure for future health interventions.

Should PI:TEN move forward, it will be essential to engage with the public in an effort to educate, inform, and demonstrate the risks and benefits of the system to all potential users. PI:TEN is an untested approach to managing a pandemic, and the public—as well as some of our workshop participants—would need to see how and why this approach could perform an important role that other initiatives (e.g. temperature checks, mask mandates, vaccines) cannot fill. As with any initiative designed to further public health or the public good, clear articulation of value will be essential to ensure uptake and build trust. This will be especially true in the case of PI:TEN, as it will involve private sector data partners who may not always be thought of as protectors of the public interest. PI:TEN will need to explore ways during the pilot phase to build the trust necessary, and demonstrate the system can deliver on the promised public benefit.

Having established a framework of potential benefits for PI:TEN, SRI drew upon workshop participants' observations to articulate a number of risks. These risks ranged from the possibility of privacy breaches to the usability of the PI:TEN system as proposed. The risks identified included:

- Low uptake by individuals due to numerous consent points, poor understanding of risks and benefits, and/or the need for those diagnosed with COVID-19 to visit multiple websites to provide consent while ill;
- Low uptake by data partners due to cost, risk, and labour;
- Inability or unwillingness of public health to participate;
- Failure to induce desired behaviour change—namely increased confidence and willingness to return to work and other activities in the Financial District;
- Unapproved use of data;
- Data breach;
- Low value to public health in the current pandemic;
- Persistence of information past the authorized date;
- Excessive notification that would dissuade users from participating.

Emerging challenges

To help set up our solutions-oriented discussion on Day 3, the SRI team helped frame some of the challenges that were raised on the first day of our discussion. There were three key areas identified; each is discussed in detail below.

a. Usability challenges

For PI:TEN to be successful, it needs widespread adoption by individuals and a significant group of data partners to sign on to the program. Furthermore, for it to achieve its maximum desired benefit, it must produce information that public health can quickly and reliably act on. There are a number of challenges in improving usability for all stakeholders.

i. Usability by participants

For PI:TEN to be successful, it needs widespread adoption by individuals and a significant group of data partners.

Existing tools, such as the COVID Alert app, have seen minimal uptake. Workshop participants pointed out that PI:TEN will need to improve on this uptake rate. If the low uptake of the COVID Alert app is linked to the need to download an app, then PI:TEN has an advantage. However, if it is due to privacy concerns, then the wider consent net cast by PI:TEN might hamper adoption. Some of our participants with deep expertise in privacy noted that the COVID Alert app was possibly overdesigned for privacy, and in doing so overestimated the demand for an app with limited public benefits (providing only personal exposure notification and only with respect to exposure to others who have activated the app). In this sense, PI:TEN might aim to make its system both easy for participants to use and easy for them to see how their choice to contribute data is helping to fight the pandemic.

PI:TEN's consent-heavy model has the potential to make the proposed system less useable for participants. There are multiple points of consent needed before someone can take full advantage of the system. The original consent to be included in the program is run through the data partners, which means people will need to consent individually to each data partner for that data to be useable by PI:TEN. Each data partner will have different data and probably different language and processes around consent. Since much of the value of the system arises from the ability to combine data from multiple sources, this means that individuals will need to give their consent several times at the very start of the process. Then, after a positive COVID-19 diagnosis, an individual must again consent to each data partner providing data to PI:TEN. As previously mentioned, this asks a lot of someone who may be seriously ill and is likely to be stressed and anxious. It also introduces a troubling potential for excluding those who are less comfortable with, or have limited access to, technology and internet connection.

Workshop participants also raised concerns about whether the source of exposure alerts might also play into the uptake and usability of PI:TEN. Being informed of a potential exposure by a commercial data partner rather than public health may lead people to take that information less seriously. Even if information is provided directly by PI:TEN, there may not be the same level of trust in the advice offered as if the message came from public health. Moreover, this approach may simply feel more privacy-intrusive—"Why is my coffee shop telling me about my exposure risk? What else do they know about me?"—than if the information comes from a public entity like PI:TEN or public health authorities.

ii. Usability by data partners

Widespread participation by data partners is equally crucial to the success of PI:TEN, and in fact may be more crucial than citizen uptake in earlier stages of the project's development. Making participation as easy and costless as possible would increase the number of data partners contributing their data.

One potential friction point is the format of the data being provided. Different data partners are likely to store their data using different formats. Many will have data that is incomplete or has errors in it. Requiring data partners to reformat data into a common form would create significant expense and difficulty, so PI:TEN will need to be able to accommodate the different data storage and formatting systems of different data partners.

Another potential friction point is the labour required by data partners to update their apps or other digital tools in order to work with PI:TEN. This likely means sending out alerts to users through apps, email, or other means (e.g. public signage in businesses). Questions arise: Would each data partner draft their own notification language, or terms and conditions? Is there a way to standardize the notification that comes from all data partners so as to reduce the associated costs and labour? Could PI:TEN provide those resources?

Independent oversight creates accountability, generates trust, and is key not only for privacy law, but for increasing participation from the public.

The same questions and challenges exist in the area of consent, as each data partner is seeking consent from each participant, leading to multiple touch points for the participants. Is there a way to standardize the structure of the consent so all data partners can make use of the same consent and terms?

Data partners are also likely to worry about the privacy and liability risks raised by their participation. This may lead them to want long and complex consent forms, which works against the citizen usability described above. The government indemnifying some of this liability would likely incentivize participation, but might also lead to moral hazard in the form of lax data security.

iii. Usability by Toronto Public Health

Workshop participants raised concerns over the form of the data provided to public health, both locally and provincially. One of the advantages of PI:TEN is that it would provide valuable information to public health, and therefore help to both track and understand the spread of COVID-19. But for this benefit to be realized, data must be provided in a form that is useable for public health. This is especially important because public health resources are already stretched very thin by the COVID-19 pandemic. The concept of the exposure network graph could support public health's desire to further understand how COVID-19 spreads. This concept will need to be further refined with public health so that it will not require extraneous efforts to use the information.

b. Legal challenges

There are three key legal challenges raised by the PI:TEN proposal: oversight, consent, and de-identification.

On the issue of oversight, PI:TEN is envisioned as a not-for-profit entity that is not engaged in commercial activities. Nor is PI:TEN contemplating acting as a health data custodian. As a result, neither of the central pieces of privacy legislation (Personal Information Protection and Electronic Documents Act and Personal Health Information Protection Act) would appear to apply. PI:TEN has said that they will present their proposal to the Information and Privacy Commissioner of Ontario and include relevant pieces of the privacy legislation in their articles of incorporation. Several workshop participants noted that this planned oversight might not be enough to establish public trust or meet legal obligations. PI:TEN should work with the Government of Ontario and with the privacy commissioners to ensure that there is independent oversight of PI:TEN that creates accountability and generates public trust. Not only is this key for privacy law, but it will assist in increasing participation from the public. There may be a need for the government to bring in new legislation for these types of proposals.

The second key issue raised by the proposal is that of consent. Consent is one of the ways in which the sharing of personal information can be authorized. It is a central concept to many privacy regimes. The PI:TEN model seeks consent from participants at multiple points in an effort to empower the individuals and protect their privacy, but questions remain. Will the consent be "meaningful and informed" (per privacy legislation) and, since there is a lack of clarity as to which piece of privacy legislation may apply, under what legal authority is the consent being sought? Even if these questions are resolved, multiple points of consent create a burden on the participants which may dissuade them from joining PI:TEN.

There are ways in which PI:TEN could potentially make consent less cumbersome than it appears under the existing model, but these risk undermining its meaningfulness. PI:TEN is encouraged to work closely with the office of the Information and Privacy Commissioner of Ontario to explore ways in which consent could be achieved in a swift and effective way.

Even without any personal identifiers, detailed location and time data is largely understood within the privacy community to be "identifiable."

Third, there is the question of "de-identification" and the privacy issues that arise with it. While the PI:TEN model plans to use what they call "anonymized exposures," it is not clear if this will live up to the obligations to protect "personal information" or to "de-identify" contained in privacy legislation. This is particularly the case because of the privacy risks associated with possessing detailed location and time data. Even without personal identifiers, this type of information is largely understood within the privacy community to be "identifiable." This leads to a number of questions around how PI:TEN will identify and manage privacy risks:

- The PI:TEN model proposes working with "anonymized exposures," but a detailed itinerary of locations/times will not be truly anonymous, since it will often be easy to reidentify an individual given even a small amount of outside information about them.
- Even if the itinerary is linked to a trace code and no other identifiers, PI:TEN will still be handling what is defined as "personal information" in privacy legislation. Is that part of the PI:TEN consent model?
- When PI:TEN asks data partners to conduct contact tracing, this means the partners are engaging a new use of their own data while also integrating new data that includes personal health information. This offloads privacy risk to the data partner.

It should be noted that location data was a frequent and recurring obstacle to agreement in global conversations about COVID-19 apps. A variety of early proposals for such apps aimed to collect location data as well as the Bluetooth "handshakes" that, for example, the Canadian COVID Alert app uses. But critics highlighted the extreme sensitivity of location data.

Furthermore, workshop participants raised the question of whether data partners learn new private information about individuals because of PI:TEN. It is imperative to understand that this may implicate privacy rights and obligations. For example, a data partner could learn that an individual was infected with COVID-19 if any entity sends them a request to share data for that specific individual with PI:TEN. This may amount to personal health information, the handling and disclosure of which is governed by Ontario's Personal Health Information Protection Act. Data partners may come under new duties and individuals may face new privacy risks that did not exist prior to PI:TEN. PI:TEN should seek further clarity on this aspect as they move forward and consider alternative models that minimize the likelihood that data partners learn health information.

c. Technical challenges

Most of the technical concerns raised about PI:TEN centred on the difficulty for data partners in setting up their systems for participation. Under the current PI:TEN proposal, data partners seem to bear the majority—or all—of the risk, cost, and labour by virtue of participating.

Data is diverse. It varies in type, granularity, format, and completeness, and can be accessed, processed, and analyzed in different ways depending on its characteristics.

For example, PI:TEN may work with widely varying types of businesses as data partners. These businesses may each have a proprietary manner of formatting their data. Some data may lend itself to export and analysis easier than others. Some databases may not be linked to a particular type of business (e.g. credit cards retain data from all types of businesses—their vendors). As part of incentivizing data partners to participate, PI:TEN can't necessarily depend on them to reformat or reconfigure their datasets to be more easily consolidated.

A related issue is the linkage of disparate databases. If public health generates an anonymous trace code—which is the case under the proposed PI:TEN model—then how might each of the data partners connect that trace code to an individual in their records? The current model would place an obligation on the participant, after a positive COVID-19 diagnosis, to go to each data partner's website and enter the trace code. This is a cumbersome model and places a significant onus on the participant, thereby likely reducing usability and utility. PI:TEN should consider finding a technical solution that reduces the burden on the participant.

Finally, questions about data de-identification remain. As mentioned in the section above on legal challenges, PI:TEN as currently conceived may not meet the legislative requirements of data de-identification. The use of location-based data, even when "personal information" is removed, is ripe for "re-identification attack." A technical solution to this could be to use a privacy-preserving analysis technique (or algorithm) which reveals the statistical information about the data being used, but does not reveal information about particular individuals.

Data is diverse. It can be accessed, processed, and analyzed in different ways depending on its characteristics.

Day 3: "How might we…?"

SRI pulled the insights from the discussion on Day 1 and the comprehensive Q&A document from MaRS together to converge on revised design challenges for the second round of our "in-person" virtual workshop. Using "How might we…?" as a prompt, SRI developed six questions to act as the jumping-off point for our structured brainstorm. The reason we use "How might we…?" questions is because they encourage participants to focus on solutions rather than further problem identification or critique.

SRI asked the workshop groups to consider the following "How might we…?" questions:

1. How might we increase participation without reducing meaningful consent?

2. How might we demonstrate the actual public benefits of PI:TEN to participants?

3. How might we reduce data breach risk without reducing public health benefits?

4. How might we achieve PI:TEN's objective to produce rapid reliable exposure network graphs without sharing new (personal) information with data partners?

5. How might we ensure full compliance with limitations on data use by data partners?

6. How might we ensure full compliance with limitations on data use by PI:TEN?

These questions were meant to stimulate conversation and explore possible pathways forward on some of the largest issues raised on Day 1 of the workshop. Each group was asked to distill their discussion of the "How might we…?" questions down to feedback that could be conveyed to PI:TEN by filling in the statement: "__________ is a key option or factor for PI:TEN to consider in their next design iteration."

Participants were encouraged to have more than one recommendation. Given time limitations, the group was able to tackle four of the "How might we…?" questions. Here are the results:

1. How might we increase participation without reducing meaningful consent? Key options or factors for PI:TEN to consider in their next design iteration include:

- Trade-off between ease of consent and meaningfulness.
- Consent standardization.
- Checking in on continued consent after the initial agreement.

2. How might we demonstrate the actual public benefits of PI:TEN to participants? Key options or factors for PI:TEN to consider in their next design iteration include:

- Effective storytelling.

3. How might we reduce data breach risk without reducing public health benefits? Key options or factors for PI:TEN to consider in their next design iteration include:

- The number of points of linkage/re-identification.
- Centralized, one-way flow of information to PI:TEN.

4. How might PI:TEN achieve its objective to produce rapid reliable exposure network graphs without sharing new (personal) information with data partners? Key options or factors for PI:TEN to consider in their next design iteration include:

- Articulating what stakeholders are learning—not sharing.
- Data vs. notification and intelligence.

5. KEY DESIGN PRINCIPLES

From ease of use to strengthening partnerships and compliance with legal frameworks, five key principles can help inform PI:TEN's design.

During our workshop, several key design principles emerged that SRI hopes can be useful to MaRS as they attempt to move this project forward. These principles represent the key takeaways from our time together working on the thinking behind PI:TEN. Overall, participants in the workshop lauded MaRS for engaging a diverse group of experts and having them "show their homework" in an effort to improve PI:TEN. While questions remain about the implementation of such a tool, many participants identified the possible utility of such an approach, both for the immediate public health needs, as well as those that may arise in the future. The workshop led us to believe that MaRS should consider the following:

a. Make PI:TEN easier for individuals to use.

Participation and uptake in existing COVID-19 eradication efforts has been remarkably low across the board, and communication from governmental agencies has often been vague, contradictory, and difficult to understand. PI:TEN would do well to learn from these past mistakes, and make participation seem worthwhile, safe, simple, and justified. There was a lively discussion that explored the possibility of demonstrating the utility of the PI:TEN model with synthetic data, through which the public could see the public health benefits and privacy protections available. If the public sees the proposed system as safe and beneficial, they are more likely to participate.

> **The streamlining and standardizing of consent is crucial—it could be achieved through methods like advance consent, layered consent, and an escalating model of consent.**

The current proposal requires numerous points of consent. While this has been done to maximize privacy and transparency, these same goals may be accomplished in other ways without deterring participation. The streamlining and standardizing of consent is crucial—it could be achieved through methods like advance consent, layered consent, and an escalating model of consent. PI:TEN is also encouraged to reduce the onus on COVID-19 positive individuals by exploring ways to get consent earlier in the process, which can be updated or revoked at any time. The model of organ donation, in which individuals indicate their wish to donate organs when they renew their driver's license, may be a helpful example.

b. Ensure obstacles to data partner participation are minimized.

Data partners face a number of potential disincentives to participating in PI:TEN. PI:TEN could dismantle these obstacles through initiatives like data consolidation, centralizing the collection and processing of data in PI:TEN (rather than requiring data partners to generate itineraries), and assisting with the drafting of legal language such as terms and conditions—especially for smaller businesses who may not have the resources to do this in-house.

With regard to PI:TEN centralizing the collection and processing of data, rather than requiring data partners to generate itineraries: this would require significant oversight and governance over PI:TEN and there are questions as to whether the existing legislation is adequate. But this may be a longer-term option that is stable and productive. PI:TEN could consider building on the model of a trusted computing environment—where computations may be performed on the data in a secure and privacy-protected manner without releasing the raw data, and where all data-processing is transparent and auditable.

In the absence of centralizing more of the risk and data collection with PI:TEN, there are some steps that can be taken to reduce the possibility that data partners will learn something new about their customers through PI:TEN, resulting in a privacy breach. Fake queries, which introduce noise into a dataset, are regularly used in location-based services to reduce the risk of re-identification or misuse of information. PI:TEN may want to consider using the approach.

c. Engage Toronto Public Health.

Commercial data partners are not the only partner with whom PI:TEN must work. A thorough consideration of public health's resources, mandate, responsibility, and computing power will be necessary to ensuring their effective participation as well. For example, PI:TEN could ensure that data provided to public health does not require further processing.

PI:TEN has said that one of the main benefits of this approach versus that of the COVID Alert app is that information shared with public health can allow them to track and learn about the virus. Currently, Toronto Public Health interviews those who have tested positive for COVID-19 in an attempt to recreate that person's itinerary from the previous 15 days. While useful, this is a limited means of tracking exposures. PI:TEN's planned exposure network graphs (see Section 3.d.3) for participants who test positive for COVID-19 would contain data linking one exposure event to another, utilizing the itinerary from data partner information (along with the public health generated itinerary). Each exposure event would have the time, location, and number of people exposed. Such a system could help rapidly identify and notify first- and second-order contacts of a participant who has tested positive for COVID-19. However, the work of PI:TEN must be done in concert with Toronto Public Health to make sure the information being shared is done in a useful way. Both PI:TEN and public health (locally and provincially) are encouraged to discuss the benefits of this type of technology to tackle the current pandemic, and possibly other health emergencies in the future.

d. Detail how PI:TEN will comply with legal frameworks for data collection and use.

Those working on PI:TEN must ensure that what they call "anonymized exposures" match up with definitions in existing privacy legislation. The term "anonymized exposures" is not used in legislation or by privacy experts, but rather the terms "personal information" and "de-identification" are. The "anonymized exposures" contemplated by PI:TEN may not meet the requirements of de-identification if the records can be easily decoded to link back to an identifiable person.

A law and technology expert at the workshop pointed out that a detailed itinerary of locations and times will by definition not be "de-identified" and will often amount to "personal information." Location-based data is one of the most sensitive forms of data, as it is susceptible to "re-identification attack," which is when an attacker can use auxiliary information—such as who was the one person in a coffee shop at a particular point in time—to connect all other anonymous locations in an itinerary to that same individual. To protect against this, MaRS may want to consider the work of SRI researchers who have found that rather than focusing on making a particular set of data private (by removing personally identifying information), the scientific community has discovered that making the computational technique (or algorithm) private provides more meaningful guarantees.

Finally, PI:TEN must continually keep in mind that data breach is a significant risk, from both a legal and technical perspective—not to mention that it's a key concern in citizen-facing and user-facing perspectives. A legal solution to alleviating data partner concerns about liability for data breach will be needed.

e. Design technical aspects of the project for maximum efficiency.

Increasing the usability of PI:TEN is essential to meet the public health goals of the project. With a view toward efficiency, facilitating database linkages through the creation of unique trace IDs in advance of a positive COVID-19 case would help meet the competing requirements of speed, efficiency, use of existing data, and privacy. The work of database linkage among a disparate group of commercial data partners, each with data in their own propriety format, would be challenging according to multiple computer security and data processing experts at the workshop. As one computer security expert put it: “If you want something to be fast, you do the work up-front.”

Finally, with respect to the threat of data breach, computer security experts at the workshop proposed a roadmap or test environment in which a relatively less complex but similar concept to PI:TEN could be tested. Demonstrating the ability to conduct first-order tracing is imperative to guarding against data breach and potential partner liability.

Location-based data is one of the most sensitive forms of data, as it is susceptible to “re-identification attack.”

6. POTENTIAL FUTURE PATHWAYS

Could key parts of the PI:TEN model inspire new directions for public policy?

The efforts being made by MaRS and the Toronto Region Board of Trade to tackle this issue are commendable. They have recognized that there is a vast array of data already being collected in Toronto's Financial District and that there may be a way to put it to use to battle COVID-19. As one workshop participant put it: "This a piece of infrastructure that we should have, and should have had already. We should look into building this now for the next pandemic." Furthermore, the openness, authenticity, and honesty with which MaRS and TRBOT entered into this workshop process demonstrated a willingness to learn and improve on their thinking. We at SRI believe that more organizations need to engage in this type of approach.

Should there be a duty for private corporations to share information that can protect the public?

Canada has recently proposed updates to federal privacy legislation (Bill C-11). Among other things, the updates aim to allow organizations to share data within certain parameters if there is deemed to be societal benefit. This, along with many other privacy legislation updates around the world for a digital and data-rich context, demonstrates there is urgent thinking taking place on the topic. It's likely that Canadian provinces will follow suit with their legislation as well.

On the other hand, privacy experts specializing in civil liberties cautioned whether commercial data with potentially socially-beneficial uses should be used in a democracy. In order for socially-beneficial uses of data to allow for data release, detailed parameters must be set to prevent mis-use of this kind of regulation. Should there be a duty for private corporations to share information that can protect the public? Public discussions have not progressed far enough on this topic.

Experts in law and technology recognized that there is a pressing need for special legislation that sets up a way of managing information, especially information whose disclosure is compelled. This would facilitate the development of things like legal firewalls around data, criminal penalties for misuse, etc. For example, we compel individuals to provide census data—this is data that we've decided is an important public resource and is needed for public decision-making. But we have a statute that protects it. Overall, public trust in Statistics Canada is high.

How might legislation change in future to facilitate the kind of data use that PI:TEN envisions? And how might we minimize data sharing while creating and using exposure network graphs?

Aside from potential future pathways in legal and technical matters, the benefits to public health and epidemiology researchers of a project like PI:TEN are also significant. While PI:TEN's current model sees the project as finite and temporary, there is reason to think that an extrapolated, longer-term, ongoing project of this nature could have tremendous value. One public health researcher pointed out that aggregate data such as PI:TEN's exposure network graphs could be immensely valuable; the primary data could still be protected or destroyed, but the secondary data that comes out of it is well worth retaining.

Experts across all subject areas also agreed on the usefulness of retaining aspects of this type of system past vaccination and past COVID-19.

Having a ready program in place for other similar public health and safety threats would be valuable. If PI:TEN is to be disassembled after its usefulness, how might we retain key parts of the model and the lessons learned from its development and possible implementation?

Perhaps most importantly, longer-term efforts to improve public outreach and education on the contemporary role and use of data are front and centre. Multiple workshop participants noted that, as a society, we must give individuals the option to make an educated, transparent decision about whether data release and use is valuable enough to them and their communities to be facilitated.

There is a pressing need for special legislation that sets up a way of managing information—especially information whose disclosure is compelled.

7. ABOUT

About the Schwartz Reisman Institute for Technology and Society

The Schwartz Reisman Institute for Technology and Society (SRI) was established through a generous gift from Canadian entrepreneurs Gerald Schwartz and Heather Reisman in 2019. SRI is a research and solutions hub dedicated to ensuring that powerful technologies like artificial intelligence are safe, fair, ethical, and make the world better — for everyone. SRI develops new modes of thinking in order to understand the social implications of technologies in the present age, and works to reinvent laws, institutions, and social values to ensure technology is designed, governed, and deployed to deliver a more just and inclusive world. SRI researchers range in fields from law to computer science, engineering, philosophy, political science, and beyond. SRI draws on world-class expertise across universities, government, industry, and community organizations to unite fundamental research on emerging technologies with actionable solutions for public policy, law, the private sector, and citizens alike.

About MaRS Discovery District

MaRS is North America's largest urban innovation hub and a registered charitable non-profit. MaRS supports over 1,400 Canadian science and tech companies, including high-growth start-ups, scale-ups, and innovators of all types who are tackling some of society's greatest challenges in health, cleantech, fintech, and other areas. In addition, MaRS also convenes all members of the innovation ecosystem to drive breakthrough discoveries, grow the economy, and make an impact by solving real problems for real people — in Canada and around the world.

The work carried out by our community is making innovation mean something again, and our purpose is to help them create a better world. Innovators at MaRS are working together to help streamline our healthcare system, reduce the effects of climate change, unlock new cures for disease, envision the jobs of the future, improve mobility in crowded cities, and provide food and clean water for a growing population.

We believe we can be the engine for Canada to lead in the innovation economy, and we drive positive global impact as the partner of choice for entrepreneurs and the innovation community in that pursuit.

*MaRS would like to credit the generous funding of the **Ivey Foundation**.*

About the Toronto Region Board of Trade

The Toronto Region Board of Trade (TRBOT) is one of the largest and most influential chambers of commerce in North America. TRBOT's constant flow of ideas, people, and introductions to city-builders and government officials firmly roots them as connectors for—and with—the business community. TRBOT acts as catalysts for the region's growth agenda, at home and on a global scale with its World Trade Centre Toronto franchise. Backed by more than 13,500 members, TRBOT advocates for policy change that drives the growth and competitiveness of the Toronto region, aiming to have Toronto recognized as one of the most competitive and sought after business regions in the world.

8. APPENDIX - A

Workshop attendees

Name	Organization
Lisa Austin	Faculty of Law, University of Toronto
Elise Belzil	Ontario Ministry of Health
Fred Carter	Office of the Information and Privacy Commissioner of Ontario
Jan DeSilva	Toronto Region Board of Trade
Effie Gournis	Toronto Public Health; Dalla Lana School of Public Health, University of Toronto
Brianna Guertin	Ontario Ministry of Health
Gillian Hadfield	Schwartz Reisman Institute for Technology and Society, University of Toronto
Asif Khan	GroundLevel Insights Inc.
Jerry Koh	MaRS Discovery District
David Lie	Faculty of Law, Dept. of Electrical and Computer Engineering, University of Toronto
Peter Loewen	Munk School of Global Affairs & Public Policy, University of Toronto
Jackie Lu	Mozilla Foundation
Peter MacLeod	MASS LBP
Brenda McPhail	Canadian Civil Liberties Association
Chad Molleken	ThinkData Works
Stephen Moscicky	Ontario Municipal Employees Retirement System
Dr. Christine Navarro	Toronto Public Health; Dalla Lana School of Public Health, University of Toronto
Alex Ryan	MaRS Discovery District
Ashleigh Tuite	Dalla Lana School of Public Health, University of Toronto
Lewis Wynne-Jones	ThinkData Works

Workshop participation does not indicate authorship or endorsement of this report.

8. APPENDIX - B

Q&A with Jerry Koh, MaRS Discovery District

Workshop participants posed a variety of questions to Jerry Koh, Director, Systems Innovation and MaRS Discovery District. Koh's answers are transcribed in full below.

Q: How would PI:TEN incentivize people to sign up? Can the benefits to individuals be more emphasized—placed in the foreground more?

A: We have yet to focus on the means and channels we would use to incentivize people. At a high level, there will be:

- An engagement workshop series where citizens will be engaged in learning about PI:TEN and providing feedback, including how we would incentivize other citizens to participate.
- An information and education campaign, including signup posters at target venues.
- Working with data partners to identify the most effective way to utilize their channels to push information and request consent; e.g. app platform, webpage, email, SMS.

We have budgeted more for citizen communication and engagement than the cost of the technology, governance, and legal instruments.

The value proposition for individuals includes:

- Keeping yourself safe and, if you are exposed, being able to take the appropriate action as early as possible to give yourself the best chance of recovery.
- Protecting your family, friends, and co-workers from being exposed to you, so that you won't become the source of their infections.
- To protect your community, and to help public health protect your community, businesses, and the local economy.

But we also know that, due to prolonged COVID-19 restrictions, keeping everyone else around you safe is not necessarily the top priority for many Canadians, even the responsible and educated ones.

We had considered approaches like rewards and loyalty programs. But we were also concerned about the implications of motivating people in the wrong ways and that they would be signing up for the wrong reasons without understanding the implications to their private information—something we often accuse big commercial platforms of.

We will be conducting rapid user research to determine the best ways to incentivize adoption. How might we best do this?

Q: How would you test the usability of PI:TEN at various steps?

A: We have not conducted usability testing at this point. This will be a major priority once we can access funding. Usability testing here will be different, in the sense that there are no PI:TEN apps. The experience is spread across multiple platforms. So we know it is even more vital to design a cohesive experience that is also highly informative and easy to navigate. How might we do this? What are better alternatives?

Q: How would you field-test all the components of PI:TEN?

A: At a high level, we will organize components as nesting sandboxes. We will start with dummy data to:

- Validate the value proposition and usability for individuals, communities, public health, and businesses.
- Test the viability of governance and legal instruments to uphold privacy while delivering value.
- Test technology and resource feasibility with public health and data partners.

We will then proceed to tightly-scoped field tests with a limited number of individuals and data partners in the Financial District Pilot Zone. How might we do this? What are better alternatives?

Q: What are the search criteria used to identify and notify contacts?

A: PI:TEN will work with public health to determine the search criteria. At its simplest, PI:TEN's data partners will be searching for contacts who are within unsafe distances for an unsafe period. This may be augmented with additional contexts such as whether the exposure is indoor or outdoor. Each data partner will have different capabilities. We will work with public health and the data partners to determine the appropriate equivalent search criteria per data partner. These search criteria can also be dynamically adjusted as public health learns from emerging evidence.

PI:TEN and its data partners can also notify users with more relevant instructions and information based on the nature of an exposure, and guide users to determine what to do based on, for example, their last test result and time, vaccination status—things known to themselves (and not PI:TEN or data partners).

Q: Can you clarify what happens with anonymized exposure entries? Would no contract tracing happen in this case?

A: The anonymized exposures will also be used for contact tracing.

Q: How is PI:TEN different from the COVID Alert app and the upcoming QR code platform?

A: The key difference is that PI:TEN is not an app, but rather a platform designed to obtain intelligence (the results of the specific pre-agreed analysis, from multiple sources of data that are already being collected) and to automate actions: notify people of their exposures. This will include the new apps that have sprung up in the past year, including the COVID Alert app and the upcoming QR code platform.

By integrating intelligence from multiple data partners, PI:TEN can rapidly conduct contact tracing to identify people who may have been exposed to a COVID-19 case (first-order contacts) and people who may have been exposed to first-order contacts (second-order contacts). This means that people can be notified faster than the virus can spread. With enough data, PI:TEN can also trace backwards to identify possible sources of infections. To our knowledge, these second-order tracing and backwards tracing capabilities do not exist in any other platforms or solutions, including the COVID Alert app.

The way PI:TEN and its data partners identify contacts (the search criteria used) will also be different from the COVID Alert app. Each data partner will have different capabilities. We will work with public health and the data partners to determine the appropriate equivalent search criteria per data partner.

One of the major features of the COVID Alert app is its privacy feature—how it stores data in deidentified ways in a decentralized manner. But this also means its value is only for individuals, and that public health can get virtually no useful information from it. PI:TEN will notify individuals of potential exposures by using intelligence derived from multiple data sources. It will also help public health build intelligence about how COVID-19 is spreading and how to better protect communities.

The COVID Alert app requires significant adoption for it to be effective. Comparing its number of downloads to the population (age 15 and above) shows approximately 20 per cent adoption. PI:TEN does not require any app installation. But it will require users to provide consent, through push notifications or other channels.

Q: How do we know that PI:TEN will be valuable for public health?

A: The ability to rapidly identify and notify second-order contacts is unique to PI:TEN and critical to containing COVID-19 spread. This is due to the nature of COVID-19, which may be contagious while asymptomatic. By the time first-order contacts are confirmed to be positive, they may have infected second-order contacts who would have, in turn, infected third-order contacts and so on.

We have also been working with Toronto Public Health to identify the kind of information that will be valuable. The concept of the exposure network graph will support public health needs to further understand how COVID-19 spreads. We'll need to further refine with public health to ensure it will not require extraneous efforts in order to be used, and to validate first with dummy data before moving to limited field data.

Q: What are the data parameters of the exposure network graphs? What is the estimated time (graph calculation) it will take to get to this point?

A: The graph contains data linking one exposure event to another. Each exposure event will contain time, location, and number of people exposed.

Our preliminary estimate is that this will take less than a day for PI:TEN. Although the data volume may be large, PI:TEN only has to perform a simple key replacement on the results returned by data partners before consolidating and returning to public health. We are confident this can be further optimized.

Q: I would like to hear more about why this approach is needed rather than low-tech interventions like investing in traditional contact tracing, or having companies in the area give paid sick days to workers. I can see that this could fill some gaps, but need to hear more.

A: Traditional contact tracing alone is not viable when case counts are high and public health resources are constrained. As we have seen in Toronto, starting in October 2020, Toronto Public Health had to prioritize contact tracing for select cases. COVID-19 cases can be contagious even when asymptomatic. To beat the rate of infection, effective tracing and notification would require conducting first- and second-order contact tracing. This would likely require the recruitment of people beyond healthcare professionals—who are needed elsewhere. This would mean potentially sacrificing the security of private information even if money was not a problem.

Paid sick leave is a vital and much-needed tool. However, employees would need to present with symptoms to know that they have to take sick leave and self-isolate. COVID-19 cases can be contagious even when asymptomatic. By the time an employee takes sick leave, they may already have infected their co-workers.

PI:TEN is not a silver bullet to replace traditional contact tracing and paid sick leave. It is meant to work alongside these interventions and help make them significantly more effective.

Q: What do we mean when we say "public health"?

A: In the Financial District Pilot Zone, we are primarily referring to Toronto Public Health—local public health units have the primary response mandate. Public Health Ontario will also be interested in the information provided by PI:TEN to inform policy.

Q: What is the relationship between PI:TEN and public health?

A: PI:TEN will be operated by an independent non-profit that will have a partnership agreement with public health. The non-profit will have no other purpose than to operate PI:TEN in order to help public health work with data partners to protect people.

Q: Does PI:TEN require IT resources from the province for case and contact management (CCM) integration?

A: Yes. We believe the effort will be minimal given that CCM is built on an enterprise customer relationship management (CRM) platform designed to be extended. If this is not possible, we can also build a standalone platform that public health can operate separately from CCM.

Q: Why is the consent giving process so complicated? Why not a simpler approach? For example, organ donation is a simple, up-front opt-in. This would also make it easier to communicate and promote PI:TEN. When someone is confirmed to be COVID-19 positive, they are likely to not be in the right state of mind to go through the laborious steps of giving consent.

A: This is one of the major design dilemmas. To address the privacy concerns of drawing upon intelligence from multiple sources, we elected to go the route of

- Obtaining consent for only as much data as needed at that point. At launch we only want permission to check people's digital footprint and notify them if they may have been exposed without exchanging information about them across data partners. When Cam [an example test subject] confirms his COVID status, we are asking for permission to round up and share his digital footprint across data partners.
- Ensuring that people understand how their data will be used, not just by public health and PI:TEN but also by each data partner. Each data partner will have different types of data and will use different approaches to identify contacts.

We love the idea of the organ donor opt-in approach. For this to work in the context of PI:TEN, one party would have to be capable of linking users' identities across multiple platforms. For example, PI:TEN would need to have the capability to link together Cam's identities on Rogers, Presto, RBC, Uber. How might we best address this privacy risk? How might we better balance usability and privacy protection?

Q: It may not be possible to get retroactive consent from a privacy perspective. The current design appears to rely on multiple individuals providing consent for PI:TEN to look backwards.

A: In the current design, all users participating in PI:TEN will provide consent at launch to allow data partners to compare their digital footprint to exposure events and notify them. This is forward-looking, i.e. data before launch will not be used.

When Cam confirms his COVID status, he provides consent again, to share his digital itinerary with data partners for the contact tracing. This digital itinerary is created after Cam's consent, based on data created after the launch.

Q: Who are the data partners?

A: To be transparent, no one has signed on the dotted line to becoming data partners. We currently have some COVID platforms that have agreed in principle to work with us. The enterprises are cautiously open to exploring.

The major types of data partners are:

- Restaurant/Business Check-in/out platforms. These are the apps and platforms that let people "check in" and "check out" of restaurants, typically by scanning a QR code. The data that will be queried are the times users walk into the restaurant and when they leave, and the location of the restaurant. Some platforms will also have more granular information (e.g. table or area).
- Building/Office/Campus platforms. These are apps and platforms that let enterprises, campuses and building managers monitor the movement of people (e.g. Siemens Enlighted, Thrive Health). They typically use security gantries and checkpoints, sensors, Wi-Fi, RF beacons and other sensing technologies. They also typically include self-assessments for symptoms. Some will also include the time and status of the last COVID test, and even vaccination status. The data that will be queried is the time that users interact with the checkpoints and their movement throughout the building and offices.
- Mobility platforms. This refers to the platforms of the mobile companies, like Bell, Rogers, Telus. The data that will be queried is the time range and movement of the mobile devices. One shortcoming of mobility platforms comes up with high-density areas, with multi-story buildings. This is where the previous platforms (e.g. restaurants, building platforms) can help.
- Transit and ride-sharing platforms. This refers to platforms such as Presto, TTC, Uber, Lyft, etc. The data that will be queried is the time range and ride.
- Banks and payment platforms. This refers to the payment platforms like Interac, Visa, Mastercard, Strip, and others who are conducting transactions on behalf of banks. This can also include food ordering apps like UberEats, Ritual, etc. The data that will be queried is the time and location of the transaction, with the caveat that the locations of some of transactions may not be available.

Q: Given the connection to healthcare, and PHIPA implications, would public health solicit consent on behalf of the private partners?

A: No. With this current design, public health will not solicit consent on behalf of the data partners. This is because each actor working with PI:TEN is asking for consent for specific purposes, within their responsibility and mandate. Public health would not be able to take responsibility for how third parties are handling the data.

In addition, there are obligations that the person must fulfill in Toronto, under Section 22 of the Health Promotion and Protection Act. This can become confusing, to navigate between the individual's legal obligation to provide information to Toronto Public Health for case management and contact management, versus the individual volunteering to contact tracing using existing data collected by third parties.

Q: How do you know that it would be feasible to link the data?

A: PI:TEN does not join data across data partners. Instead, it depends on each data partner's capability to assemble Cam's digital itinerary of time and locations. In some cases, it will be direct mapping, while others will provide equivalent data. The digital itineraries are then merged into one and sent back to each data partners as search parameters. Then each data partner uses its capabilities to identify potential contacts and notify them. While this is not insignificant, this type of data and analytics is also fundamental to most business intelligence needs. Because the data partners are collecting different types of data, we will have to work with public health to translate the search criteria into what each data partner can work with.

Q: How will PI:TEN ensure that the data collected will be deleted?

This is one of the key design questions. This is where we believe a solid combination of the governance model and legal instruments will be needed to ensure the oversight of data, in addition to the technology platform design.

For example, can the non-profit that will operate PI:TEN have legally binding obligations in its incorporation documents, and its agreements with public health and data partners? In addition, PI:TEN will be transparent wherever possible without endangering privacy and security. It will be open to audits. The Privacy Impact Assessments will be published.

For individual data partners, they will abide by the partnership agreements and their data policies.

Q: Is it realistic to have PI:TEN in place while it is still useful? Given that we are now rolling out vaccines, that they may be available to all as early as May. Are there other valuable uses, for example with other contagious diseases like measles?

A: In terms of timing versus vaccination, it may be September before we reach herd immunity. There is also the challenge of anti-vaxxer communities. PI:TEN can help ensure infections in these communities can be quickly contained. In terms of the technology, we believe we can get the prototype up in a month. We are also confident of getting at least two data partners on board. The challenge is to get the governance and legal instruments ready within the same timeframe.

One of the cornerstones of our design is that PI:TEN would have a limited lifespan. When COVID is no longer a public health crisis, it will be shut down. We felt this was critical, given the extensive privacy concerns it raises. While the technology can be used for other purposes, we would need to redesign the governance model and the legal instruments to ensure that it is not abused. Once we can build and validate the first version of PI:TEN, if there were sufficient urgency and demand for the same kind of system to be set up for a different public health challenge, we would go through similar rounds of consultations. We would not augment the existing system.

What happened in Singapore is a good case study for potential overreach. When the Singapore system was set up, they assured citizens it would only be used for the narrow purpose of fighting COVID. But then they started allowing the police to use the same platform. This can undermine public confidence in any future public health surveillance platforms. We believe in being very cautious, in being singularly purposed with PI:TEN, and to sunset it when it is no longer needed, or when the number of cases is too small that the risk of re-identification becomes too high.

Q: How do you know that PI:TEN will be effective? Do the majority of people comply with isolation when advised by public health? Or do we have a compliance challenge?

A; PI:TEN builds on the Swiss chess concept, that to effectively detect exposures, to identify and notify contacts, the key is not to put our bets in one app. Instead, if we can utilize data that is already collected, we would be able to create an extensive detection network.

As mentioned earlier, it does not have an adoption problem the way that the COVID Alert app does. But the consent approach will also pose an adoption challenge. The privacy considerations put into the COVID Alert app are extensive, while PI:TEN is relatively untested.

Given that users will be directly engaged by the data partners to provide consent, we think the barrier of adoption may be lower compared to the COVID Alert app, which depends on the user to install the app.

The other key consideration of effectiveness is whether PI:TEN can change the behaviour of individuals when they receive notification of exposure. There have been limited findings shared on the compliance rate in Toronto from one-off research and media investigations. Like the COVID Alert app, this is unknown. Public health restrictions have also had limited success. PI:TEN can support sending context-specific notifications that can be made more relevant to how the individual was exposed, and more relevant instructions. We think this would make notifications more compelling.

The truth is, there is no precedence with what we are doing. While we can rationalize how effective PI:TEN will be, we will ultimately have to prototype, build, and test it to be sure. Around the world, South Korea and Singapore have taken similar approaches. But their approaches would be incompatible with Canada's value and privacy system. This is why we want to do a pilot in the Financial District Pilot Zone, to generate data and evidence, and to be able to test it in a controlled environment.

Q: What is the scale of the pilot required for it to be viable, and to demonstrate its value? What is the minimum number of data partners and users required?

A: We believe that we will be able to demonstrate effective results at the controlled scale provided by the Financial District Pilot Zone. We believe that going to a large regional scale is just too risky.

For the pilot, we will focus on people living and working in or near the Zone. We will also include select regions where large numbers of people are commuting into the Zone. We are not confident of the minimum number of consenting people we need. Using what we know about COVID tracing apps, we would need to aim for the number of unique individuals providing consent to the data partners, reaching 60% to 80% of the target population.

The complication here is that the current number of people coming into work is also significantly lower compared to pre-COVID times. The Pilot Zone has already undertaken modelling to determine safe targets for occupancy, and the number of people returning to the Zone in phases. We will work with these targets as the denominator. The additional complication is the large number of people who would usually commute into the Zone. We would not want to scale up to large numbers before becoming confident with PI:TEN. The early phases of reopening will focus on people who already live in or near the Zone, who utilize active transport options like walking or cycling. This will allow us to scale up the scope of the pilot gradually, matching the scale of the reopening.

For data partners, we will begin with two to three restaurant and building management platforms, that we aim to have consistent rollout across the Zone. Based on what we know about COVID, these platforms will cover venues and circumstances that are of the highest risks, outside of large private gatherings.

Q: What do we know about the population who would be expected to use this, and are they the same people who haven't downloaded the COVID app already?

For the pilot, we will focus on people living and working in or near the Zone. We will also include select regions where large numbers of people are commuting into the Zone. We are uncertain about the overlap with people who haven't or choose not to download the app.

Q: What does data partner on-boarding look like?

A: We will work with the new data partner to understand their data capabilities and determine if they can plug into one of the existing interface models, and develop a new interface if it is not possible. PI:TEN will also provide playbooks and support the data partner in determining the best approach to obtain consent and support the education campaigns necessary. The new data partner will need to undertake a Privacy Impact Assessment for the scope of PI:TEN partnership.

Q: What have you considered in terms of the feasibility and costs/viability for data partners? The data may be there but they may not be easily accessible and easily applicable to contact tracing and notification.

A: The key data analytics required is to match time range and location range. The data partners we will be starting with are COVID tracing platforms, so they will be ready. We think this type of analytics is also common on platforms that support business analytics. However, the effort will not be insignificant. In some cases, the analytics will be direct, while others will use workarounds to provide equivalent results. We recognize that there may be significant costs to data partners. We will work with public health and the data partners to translate the search criteria into what each data partner can work with. We will work to minimize the burden on the data partners.

Q: Will PI:TEN be subjected to an ethics review?

A: We have not considered an ethics review. That said, our principle is to be as transparent, trustworthy, and accountable as possible. If undergoing the ethics review will help, we will do it.

PI:TEN will engage a reputable third party to conduct a Privacy Impact Assessment as per the guidelines of the Information and Privacy Commission of Ontario. The Privacy Impact Assessment will also be published.

The Privacy Impact Assessment will be updated whenever a new data partner is onboarded. This is because each data partner added to PI:TEN potentially increases the privacy risks by an order of magnitude. This is our brute force approach. We would welcome alternatives.

Q: How will people be able to know whether a place they are visiting is part of PI:TEN, and whether their use of services affects them and/or the network?

As part of the Financial District Pilot Zone, there will be posters and other mediums prominently displayed. We are also considering the use of **DTPR** communication iconography to inform people and allow them to manage their participation in PI:TEN. In addition, PI:TEN and data partners will also regularly remind users of their participation, of what they have consented to and the option to withdraw consent.

Q: Has the research team developing this talked to CDS, who have been grappling with many of the same issues?

A: No. We would appreciate a warm introduction to the Canada Digital Service, as well as the Ontario Digital Service!

We have spoken to the Salesforce volunteers who worked on the COVID Alert app, to learn from their experience.

Q: How granular is consent? Will it include the ability to set limits/conditions or to revoke?

A: In our current design, consent is simply opting in or out with public health, with PI:TEN or with each of the data partners. Users will be reminded regularly of what they have consented to and given the option to revoke.

Q: What is the default notification level maximum or minimum?

A: To be determined. How might we do this?

Q: Would property owners write this into leasing agreements, compelling retailers to participate?

A: In the context of the Financial District Pilot Zone, participation is voluntary. The Zone coalition is also considering a form of a business standard that participating businesses can display prominently, similar in concept to DineSafe.

Q: If the towers could be opened up safely, would it be worth it to subsidize access to PI:TEN?

A: We would love it if the answer is yes! The building managers are all considering building management platform solutions that would help support a safe return to work in their buildings. PI:TEN helps extend the surveillance between the buildings and to the whole district. We have suggested this to select building managers. Fingers crossed.

Q: The "anonymized exposures" —a detailed itinerary of location/time—will not be "anonymous."

A: The anonymized exposures collects time and location from multiple individuals as discrete unrelated data points. We believe the data can remain anonymous with a sufficiently large number of individuals, combined with geomasking. If there are too few individuals, each with long detailed itineraries, it would be possible to re-identify. This would be one of several predetermined thresholds that will halt PI:TEN.

Q: PI:TEN is handling "personal information" even if the itinerary is linked to a trace code and no other identifiers. Consent for this will be needed too—is that part of the consent model? (And consent is not the only obligation, which goes back to wanting details about accountability and oversight.)

A: Yes. But we have yet to determine the right accountability and oversight model. How might we do this?

Q: When PI:TEN asks a data partner for data about an individual, that data partner learns something new about the individual—do you understand that this is a privacy issue too?

A: Yes. And this would be part of the consent required by PI:TEN and each data partner. Since each data partner has different data and analytics capabilities, what they learn will also be different. This is why we went with the cumbersome approach of the user providing consent to each data partner separately.

Q: When PI:TEN sends a full itinerary to each data partner PI:TEN is giving them additional information that the data partner did not previously have, and it would be easy for the data partner to learn who this itinerary pertains to—that is a big privacy risk. How do they propose to mitigate this?

A: We propose to mitigate the risk through legal partnership agreements and technology requirements that the data cannot be used for other purposes, and will be deleted on a rolling 30-day basis. Where it is feasible, the data will also stay within a partition separate from the rest of the data partner's network.

Every new data partner added will trigger an update of PI:TEN's Privacy Impact Assessment. The new data partner will also be required to conduct a Privacy Impact Assessment. Each data partner will also have their existing data and privacy policies to comply with.

Q: When PI:TEN asks a data partner to do contact tracing, this is a new use of their data and also integrates new data, which offloads privacy risk to the data partner. What is the thinking around this?

A: Every new data partner added will trigger an update of PI:TEN's Privacy Impact Assessment. The new data partner will also be required to conduct a Privacy Impact Assessment. Each data partner will also have their existing data and privacy policies to comply with.

Q: Regarding the relationship between this and what Lisa Austin and David Lie proposed in the "Safe Sharing Sites" paper—they were generally opposed to sharing data with so many different parties. The whole point was to create methods to avoid this.

A: We think the data shared between data partners is kept to a minimum to achieve the first- and second-order trace and notification function and to return valuable pandemic intelligence on how exposure events relate to one another. The data that is shared is also covered by informed consent. We agree this still presents risks.

How might we achieve the functions or equivalent outcomes without sharing data between the data partners? Is it possible?

Q: The location data point was a real sticking point in global conversations about COVID-19 apps. There were a variety of early proposals for such apps that would collect location data as well as the Bluetooth "handshakes." But people were very critical of those proposals because of the sensitivity of location data, even though most of those proposals were more privacy-protective than PI:TEN appears to be, and even though the apps were always proposed as consent-based.

A: Agreed! We hope PI:TEN can be assessed with a different lens. PI:TEN itself does not collect new data. It enables existing data and analytics to be used to fight COVID. This is the same data we have "consented" to be used in large parts to sell things to us. But is this a difference that matters enough? What if we focus only on the integration of apps created specifically for COVID?

8. APPENDIX - C

Participant commentary

The SRI team is proud of the group we assembled to stress test PI:TEN. As a result of the expertise in the virtual room, the diversity of opinions, and the intellectual generosity of participants, we were able to have a thoughtful and iteration-forwarding discussion in a short period of time. Often, in these types of workshops, participants are not able to articulate all of their thoughts for full discussion amongst the group. As a result, those opinions are not necessarily reflected in the report relating the workshop's findings. We at SRI wanted to provide participants with an opportunity to voice their opinion publicly as part of our final report. As a result, we gave all participants the opportunity to provide a 250-750 word response to be included in the Appendix. Our hope is that this approach will increase the transparency, quality, and completeness of this report, and we hope to be able to use this method in future workshops.

Lewis Wynne-Jones, *ThinkData Works*:

For the first few weeks of COVID-19, I didn't wear a mask. Most of us didn't. Initially, masks weren't required in the stores that remained open, and seemed to many of us a novelty—something worn by some people out of an abundance of caution. It's hard to remember now, but the masks, the sanitizer bottles in the entrance of every store, the two metre distance between us: these were not instant precautions, but learned ones. These protective measures caught on because studies suggested that they were some of the best ways to mitigate the spread of the virus. It wasn't intuition, it was data.

Data has been instrumental during the COVID-19 pandemic. From Johns Hopkins daily case rates to genomic sequencing, we have benefited from a data-rich environment. When this data is applied to public policy (and embraced by our communities) we see a drop in cases, more beds in the ICU, and the beginnings of economic recovery. When it is ignored, we see the opposite.

The data that has been most impactful, to date, is the reactive data: information generated because of the pandemic. These are the case counts, government measures, and sequencing data that exist within the context of the pandemic and our response to it.

We've been less successful using the other type of data that could be leveraged to mitigate the impact of the pandemic. This is the data generated every day as we move around our cities, commute to work, and buy our morning coffees—the passive data that has become the heartbeat of an increasingly connected society. The benefit of harnessing this data can't be overstated. If you can piece together a person's movements through an area, you can develop a framework for alerting people who may have been exposed to COVID-19, in real time. Contact tracing like this would have a hugely beneficial effect, leading to faster containment, quicker response, and more detailed information about how the virus spreads.

So why hasn't this happened? The problem is not with the availability of the data. Between smartphones, sensors, and things like transit passes and loyalty cards, the data we need to build this kind of solution exists in abundance. But as we've learned from open data initiatives globally, available does not equal accessible. For good reasons, the general public can't access this information. Sometimes the data is maintained and protected by the city, other times it's owned by the business where I've shopped, or the company that runs the app I'm using.

At a time when data governance is top of mind for any organization that generates or maintains data, more restrictions on its use, not less, will be preferred. For citizens who are increasingly concerned with their digital footprint, safeguards around consent and appropriate use of their personal data need to be established. The problem is not technological. The problem is social, ethical, and logistical.

A system for trusted data exchange, with willing data partners and widespread adoption, would have an immediate and lasting effect on pandemic response. Although COVID-19 is far from over, we have to look forward to see what types of solutions we can build, using the resources we have, to prepare for the future. The barrier to entry needs to be low, users need to be protected, and the benefit needs to be real.

As the new census rolls out across Canada, we should be reminded of the contract we engage in with our government every few years. The census works because there is a fine balance struck between trust and benefit. As data increasingly becomes the natural resource generated from every activity we do, we need to find that same balance in order to maximize its impact. This may not immediately result in contact tracing or a seamless information exchange between citizens, businesses, and public health. In order to protect our privacy, it may not impact us individually. As an experiment in public well-being, the initial result may be that we gain better insight into how to protect at-risk communities, prevent overcrowding at hospitals, and create better policy.

We wear masks now because the data showed us that it benefits not the individual but the community around us. In order to make good on the data that exists in the public domain and within the servers of the organizations that own it, we should focus our attention on communal actions that have downstream impact, rather than individual outcomes. In order to get more data into the hands of people who can use it, we need to find the balance that the census finds, a trade-off between individuation and security. Whatever course of action we take, data will surely be at the centre of our strategy, granular enough to be useful, and aggregated enough to be safe.

Brenda McPhail, *Canadian Civil Liberties Association*:

The PI:TEN project, and the Schwartz Reisman Institute workshop that led to this report, opens up a set of questions about privacy, data, and public good uses that are timely and complicated. The Canadian Civil Liberties Association (CCLA) appreciates the opportunity to take part in the conversation, and to comment on the subsequent report. We would like to emphasize two key points, and one larger concern.

First, at its core, this project is about convincing people that data they are allowing to be collected for one purpose can (or should) be used for another, public good, purpose. It is important in this context to be clear about the nature of the data partners, and types of data, being proposed, including but not limited to:

- partners who have customer loyalty data (data collected with consent);
- transit data (data generated through a payment system, which while technically optional, has been socially engineered to strongly promote use);
- access code data (not really data collected voluntarily in the first instance, unless entering a home or workplace is considered optional); and
- provincially-mandated contact tracing data (data collected by regulation).

Each of these data categories raises different issues to be addressed when considering what informed consent means, and how to meaningfully acquire it. While the report accurately reflects the conversation in the workshop about reducing the consent burden on individuals who might choose to participate in PI:TEN, CCLA commends MaRS for their initial commitment to require consent for each data partner (rendered more important because of the different data types and forms of collection identified above). We strongly believe that no less than granular, repeated, and fully informed consent would be fit for the project's purpose. It will be a burden on individuals and create friction within the system, and given the legal challenges identified in the report it will be difficult, but it is necessary. The discussion of regulatory reform to alter this necessity was intriguing, but far from current reality.

The second point, addressed in this report but worthy of highlighting, is the very real problems inherent in a model that gives partner organizations information about individual customers' positive COVID-19 status or exposure, and a role in notifying people about such exposures. It is mentioned in the report that this "implicates" privacy rights. CCLA would say this has the potential to trample on them. Data partners should absolutely not receive information about a customer's positive COVID-19 status and the model must be adjusted accordingly. This report also notes that it may "feel more privacy intrusive" to be notified of a potential COVID-19 exposure by your local coffee shop than by a public health professional; we should be clear, it is, in fact, more intrusive, and also puts data partners in a role they are likely unqualified to fulfill.

The pandemic has given rise to a plethora of plans to leverage technology and data to assist in the fight to protect public and economic health—some successful, some not. Those that involve some form of surveillance, whether of movement, contacts, or location, have been the most contentious because they place privacy rights at risk, which in turn opens up the door for individualized discriminatory impacts based on the sharing and use of personal health information. The PI:TEN project is another such plan, although, to their credit, MaRS has demonstrated a strong understanding that human rights must also be protected if a true "public good" purpose is to be achieved. What this workshop and the subsequent report have demonstrated, and what CCLA would most like to stress, is that as a society we need more difficult, down-and-dirty, multi-faceted, and multi-stakeholder conversations about the concept of repurposing data for public good, and ultimately, we need a coherent regulatory structure in place to provide the right protections and safeguards. PI:TEN has a difficult, and possibly impossible task ahead to adequately address the issues of privacy, legality, accountability, oversight, security, and utility identified in the report, all of which are important and necessary in order to proceed.

HARNESSING COMMERCIAL DATA FOR PUBLIC GOOD

Design reflections on the Pandemic Intelligence: Trusted Exchange Network (PI:TEN)

A Solutions Workshop report from the Schwartz Reisman Institute for Technology and Society at the University of Toronto in collaboration with MaRS Discovery District and the Toronto Region Board of Trade

May 13, 2021

srinstitute.utoronto.ca
hello@torontosri.ca

@TorontoSRI