Ontario’s COVID-19 Data Platforms, or 
Si vis pacem, para bellum

Michael Brudno
University Health Network (UHN)
Department of Computer Science, UofT
Vector Institute
Ontario

- Biggest province in Canada (40% by population, ~12 million people)
- World class Computational Science
  - UofT ranked top 10 in world
  - Vector Institute founded by Geoff Hinton (Turing award 2020) et al
  - Health computing platform (HPC4Health) enables secure computing on biomedical datasets
Ontario

• Biggest province in Canada (40% by population, ~12 million people)
• World class Computational Science
  ○ UofT ranked top 10 in world
  ○ Vector Institute founded by Geoff Hinton (Turing award 2020) et al
  ○ Health computing platform (HPC4Health) enables secure computing on biomedical datasets

COVID-19 Data Platform (Ontario Health Data Platform)

• CONNECTED Multiple existing components, BUILT on world-class expertise, ENABLED data-driven decisions
ICES uses *de-identified and linked* Ontario population-based admin data to generate evidence that stimulates improvements in health and health care for Ontario.
Institute for Clinical and Evaluative Sciences

ICES Data Repository: De-identified and Linkable

Provider/Facilities
- Physicians
- Hospitals
- Complex care
- Long-term care
- Homes
- Home care

Health Services (IKN)
- Physician claims
- In-patient hospital discharge abstracts
- Emergency and ambulatory care abstracts
- Prescription drug claims (65 and over)
- Home care claims
- Rehab claims
- Long-term care visits
- In-patient mental health data

People & Geography (IKN)
- People in Ontario since 1985
- Unique individual anonymous ID
- Postal Code Conversion/Geographical Population Estimates
- Canada Census Profiles
- Death register

*Special Collections (IKN)
- Registries (Cancer, Stroke, CCN, *Birth outcomes)
- Federal immigration register
- *First Nation Metis
- Developmental Disabilities
- *HIV
- Primary Collected Data & Other!!! (IKN)

Derived chronic conditions (IKN)
- (using linked data)
  - Diabetes
  - Hypertension
  - COPD
  - CHF
  - AMI
  - Asthma
  - IBD

IKN = unique algorithm based on Ontario health card number
HIGH-PERFORMANCE COMPUTING FOR HEALTHCARE

Our mission is to make high-performance computing accessible to healthcare providers. Together we are building the engine that will help make personalized medicine and diagnostics a reality.

MISSION STATEMENT
Computing for Health

From genomics to medical imaging, almost every discipline in healthcare is dealing with a "Data Deluge" of information. Translating this into something that will ultimately benefit patients requires massive amounts of computation and storage in an environment that is fast, secure, and run with efficiency. HPCH4Health is a consortium of health providers who are working together to build the next-generation of compute engine for clinical research.

PARTNER MEMBERS
Working Together

Two of the biggest hospitals in Canada, the Hospital for Sick Children and the Princess Margaret Cancer Centre at UHN, have decided to join together and build a framework to handle their future research computing and data needs. HPCH4Health will be accepting new member health research institutes to use this platform, and will also work with others to deploy the system we have built elsewhere.

OUR GRATITUDE
Thanks to Our Supporters

By working together, we have built a model for a computing infrastructure that is much bigger than what we could have built on our own. We would like to thank our internal funders, including the SickKids Foundation and the Princess Margaret Cancer Foundation. A large portion of our equipment has been purchased through grants from the Canada Foundation for Innovation and operational money has been provided by Compute Canada, both of whom we also thank.
The HPC4Health: Basic Facts

- Joint “elastic compute cloud” built by Hospital for Sick Children and UHN and launched in 2015
- A node in the Compute Canada supercomputing infrastructure
- Capable of processing large genomics or other types of data sets (“Big Data”)
- Enterprise level governance, management and technical support
- Capability to handle both research AND clinical data processing
The HPC4Health in a Nutshell
ICES + HPC4Health = HAIDAP: Health AI Data Platform
# HAIDAP + COVID = OHDP (Ontario Health Data Platform)

<table>
<thead>
<tr>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>Federated high-performance computing environment</strong> for secure, accurate, and privacy-protective linkage of large health data sets to allow for data analytics and research to support Ontario’s response to COVID-19</td>
</tr>
<tr>
<td>• Ministry of Health custodianship of datasets to enable research, including clinical and administrative data</td>
</tr>
<tr>
<td>• Protect the <strong>privacy and confidentiality</strong> of personal health information through innovative strategies and effective data stewardship</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Current Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Launched on <strong>July 15, 2020</strong> and has supported 49 research and analytics projects</td>
</tr>
<tr>
<td>• Informs <strong>critical pandemic planning</strong> including the vaccination strategy, containment strategies, preparation for pandemic waves, support for vulnerable populations, and promotion of health equity.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Future Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Evolution into an ongoing government-sponsored research and analytics data platform supporting a health data ecosystem</td>
</tr>
</tbody>
</table>
OHDP links large health datasets from a variety of sources to create an unprecedented volume of rich, connected data.
Federated Platform

OHDP DATA SOURCES
DIRECT IDENTIFIERS ARE REMOVED AND/OR TOKENIZED

FEDERATED OHDP COMPUTING ENVIRONMENTS
SECURE, HIGH-PERFORMANCE COMPUTING PLATFORMS

OHDP DATA ACCESS
SAFE, PRIVACY-PROTECTIVE RESEARCH AND ANALYTICS

ACCESS
OHDP-I
ACCESS TO ADDITIONAL DATA UNDER ICES

AUTHORITY:
PHIPA S. 45 HEALTH SYSTEM MANAGEMENT, EVALUATION, MONITORING;
PHIPA S. 44 RESEARCH: REB REQUIRED
TRUSTED RESEARCHER / AUTHORIZED USER

Ontario
Federated Platform

<table>
<thead>
<tr>
<th>OHDP DATA SOURCES</th>
<th>DIRECT IDENTIFIERS ARE REMOVED AND/OR TOKENIZED</th>
</tr>
</thead>
</table>

FEDERATED OHDP COMPUTING ENVIRONMENTS

<table>
<thead>
<tr>
<th>OHDP-Q</th>
<th>OHDP-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>• HIGHER COMPUTING POWER AND CAPACITY</td>
<td>• ACCESS TO ADDITIONAL DATA UNDER ICES</td>
</tr>
<tr>
<td>• ACCESS TO DATA UNDER MINISTRY OF HEALTH</td>
<td>ACCESS</td>
</tr>
</tbody>
</table>

OHDP DATA ACCESS

<table>
<thead>
<tr>
<th>AUTHORITY: PHIPA S. 44 RESEARCH; REB REQUIRED</th>
<th>AUTHORITY: PHIPA S. 45 HEALTH SYSTEM MANAGEMENT, EVALUATION, MONITORING; PHIPA S. 44 RESEARCH; REB REQUIRED</th>
</tr>
</thead>
</table>

TRUSTED RESEARCHER / AUTHORIZED USER

Ontario
What does OHDP offer to researchers?

• High-performance computing to support big data analysis and machine-learning related to COVID-19.

• Expedited access to computing resources and data, including clinical and administrative data, with costs covered in full by the Ontario government, if certain requirements are fulfilled.
What does OHDP offer to researchers?

• High-performance computing to support big data analysis and machine-learning related to COVID-19.

• Expedited access to computing resources and data, including clinical and administrative data, with costs covered in full by the Ontario government, if certain requirements are fulfilled.

What does OHDP want from researchers?

• Projects that are aligned to COVID-19 Health Research Priorities of the Ministry of Health. Timely, real-world insights that can support health, social and business policy priorities, such as support for vulnerable populations, clinical and public health interventions, and re-opening strategies.

• Engagement from the research community in the co-design of an Ontario big data platform that is modern, integrated, effective, and researcher-friendly
How has OHDP been used

• Building models to predict health outcomes
• Epidemiological analysis
• Driving policy
How has OHDP been used

- Building models to predict health outcomes
- Epidemiological analysis
- Driving policy
How has OHDP been used

- Building models to predict health outcomes
- Epidemiological analysis
- Driving policy
How has OHDP been used

• Building models to predict health outcomes
• Epidemiological analysis
• Driving policy

**COVID-19 in Toronto neighbourhoods: How postal code data highlights vaccine inequities**

By Nick Westoll
Posted April 7, 2021

Laura Rosella, a scientist with the ICES, told Global News she and others have been crunching the real-time statistics in an effort to help “change our response.” She said the group receives COVID-19 and vaccination data from a variety of health sources and categorizes that data based on where people live.

Rosella said two major things stood out to her after doing a high-level look at Toronto. She noted the city, as well as Peel Region, particularly saw higher incidences of COVID-19 with lower vaccination rates.
How has OHDP been used

- Building models to predict health outcomes
- Epidemiological analysis
- Driving policy

WATCH: Ontario ramps up vaccine rollout in hot spots in 114 'highly impacted' areas, including for those 18 and older

Mobile teams and pop-up clinics to begin targeting hot spots neighbourhoods, on-the-job clinics also in the works
What is next?

- Simplifying project onboarding without compromising privacy/security
- Allowing researchers to bring + connect external datasets
- Connecting data across Canada
- Connecting with international efforts
Next step: connect OMICS
CanDIG for human genomes (also on HPC4Health)

National network for (human) genomics data

Based on trusted local nodes (one in QC, two in ON, one in BC).

Query sent to sites in the federation, which make local authorization decisions

Data for response to show to user is ephemerally brought back to user’s home site and computer
Key lessons

- Existing expertise must be in place to move forward quickly
- Maximally utilize existing centers of excellence
- Getting the right datasets together is the hardest part
- Policy expertise is more important than technical expertise
  - Policy drives technology, not vice-versa
- Researchers should be consumers, not drivers
Acknowledgments

- HPC4Health: Carl Virtanen, Rob Naccarato, Amanda Silva
- ICES: Michael Schull, Charles Victor
- MOH: Christine Sham, Roman Rex
- Indoc (Ken Evans, Tom Gee)
- CanCOGen (Lisa Strug, Guillaume Bourque, Steven Jones)
- CanDIG (Jonathan Dursi, Guillaume Bourque, Steven Jones)