Access
Opportunity
Connection

Compute • Calcul
Ontario

2017 - 2018
Annual Report

Compute • Calcul Ontario

Table of Contents

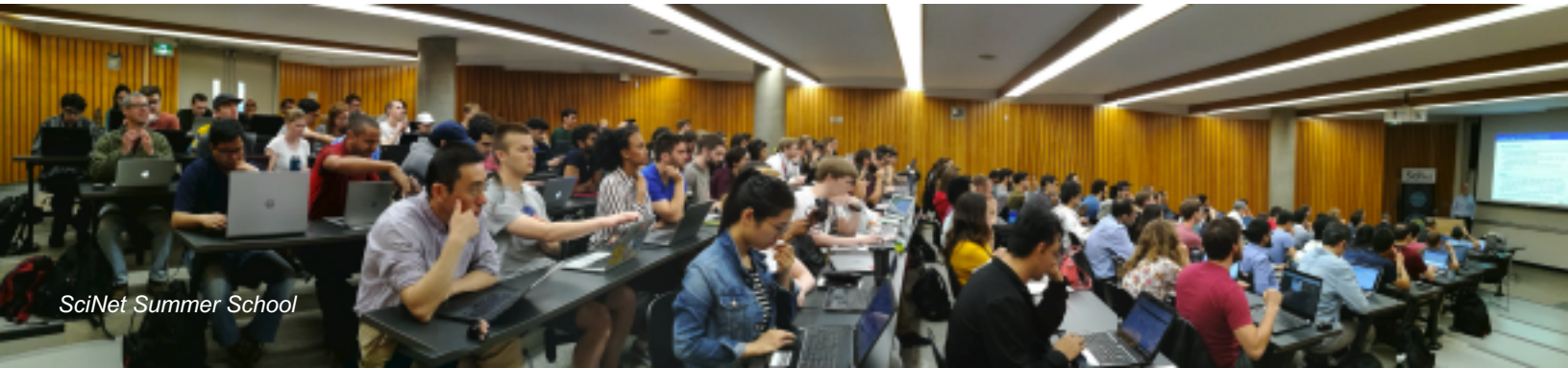
Introduction

Compute Ontario is pleased to present its 2017-18 Annual Report. In addition to providing an overview of its achievements, this report also profiles a series of research impact stories. These stories are intended to illustrate examples of research made possible through access to digital research infrastructure (DRI). DRI is a general term that includes a robust network, advanced research computing (ARC), data management, research software, and highly qualified personnel (HQP) to maintain Canada's DRI resources and train researchers in advanced computing techniques. Compute Ontario's work, and the work of our consortia and partners, supports all of these DRI components either directly or in collaboration with others in the ecosystem.

This past year, Compute Ontario achieved progress on many of its strategic goals. Two of its most notable achievements include the delivery of a technology investment study and the completion of our highly qualified personnel study. These two initiatives involved months of discussions and research under the auspices of two provincial advisory groups comprised of subject matter experts. The insights gained from each study take us one step closer to realizing our vision of driving DRI to accelerate research and enhance competitiveness in the global marketplace for a more prosperous Ontario.

This report discusses our accomplishments within headings aligned to Compute Ontario's strategic priorities. These strategic priorities are to:

1. Inform Policy.
2. Support Digital Research Infrastructure Needs of Researchers.
3. Increase the Supply of Highly Qualified Personnel.
4. Connect DRI Communities.



SciNet Summer School

Reporting on Priorities

1.0 Inform Policy

Compute Ontario serves as the voice of our community, coordinating and advocating for key strategies that enhance digital research infrastructure (DRI). We ensure that Ontario keeps pace with advancements in DRI, and is well-positioned to support an evolving research community that helps to drive Ontario's economy. In 2017-18, Compute Ontario initiated a number of activities to inform policy development in Ontario and nationally.

In 2017-18, Compute Ontario undertook province-wide consultations to develop a position on the national coordination of DRI. The resulting position paper was formally published in fiscal year 2018-19, and concluded that to realize a world-class DRI ecosystem in Canada, the national approaches to a strategy must:

- Reflect principles which promote pan-Canadian ideals for research and implement a 21st-century DRI ecosystem that is modern and sustainable.
- Embody a philosophy which is:
    - researcher centred
    - universal, equitable, and accessible
    - forward-looking, adaptive and nimble
    - collaborative and integrated
    - sustainable.
- Recognize the importance of a federated approach to coordination which applies those principles through appropriate governance tools and processes.
- Recognize and clearly define the appropriate roles and accountabilities for all participants in the ecosystem.
- Apply approaches to funding that are grounded in the same set of principles, and foster appropriate investments nationally, regionally, and across sectors.
- Respect the existing diversity of expertise within the DRI sector.

The President and CEO was asked in late winter 2018 to present the paper to senior leadership at the Ministry of Research, Innovation and Science (now Ministry of Economic Development, Job Creation and Trade) in support of an Ontario perspective on a national DRI Strategy.

In addition to Compute Ontario's position paper, the Leadership Council on Digital Research Infrastructure (LCDRI) was formed at the request of ISED to provide recommendations to the federal government on how to improve the DRI ecosystem in Canada. The LCDRI was a group comprised of leaders from the digital research community in Canada. LCDRI formed working groups to develop recommendations on several topic areas subsumed within DRI. Regional organizations across Canada like Compute Ontario were asked to participate and appoint two representatives from each region.

Ontario's two representatives were from the Centre for Advanced Computing (CAC) at Queen's University, and SOSCIP -- a consortium of academic and industry participants. The working groups were instrumental in framing the Council's recommendations on ways in which DRI could be better supported and coordinated nationally. LCDRI's final recommendations to ISED were submitted in the fall of 2017. The 2018 Canadian federal budget referred to the work of LCDRI as being instrumental in identifying the $573 million for DRI, which was announced in that document.

2.0 Promote Digital Research Infrastructure

Compute Ontario, in collaboration with a pan-Canadian network of partners, actively promotes a strong national DRI federation. In so doing, we strengthen the position of Ontario's academic research community. As Canada succeeds, Ontario succeeds.

In 2017-18, Compute Ontario worked with its consortia and ORION – an organization that oversees an ultra-high-speed fibre-optic network dedicated to enabling research, education, and innovation (RE&I) in Ontario – to develop a proposal to support urgently-needed DRI: 100G networking, and test a new platform for artificial intelligence enabled health research (HAIDAP).

The Ministry of Research, Innovation and Science (now Ministry of Economic Development, Job Creation and Trade) awarded ORION and Compute Ontario $10 million to undertake this capacity-building exercise. This is new funding not previously earmarked for DRI in Ontario and was distributed by ORION and Compute Ontario to strengthen the DRI ecosystem in the Province.

One of Compute Ontario's strategic priorities is to: *Coordinate and support the advanced computing needs of Ontario's academic research community and other key stakeholders, in partnership with sites across the province that deliver and manage these services.* In that regard, in 2017-18, Compute Ontario worked with the University of Waterloo and the University of Toronto to launch two new supercomputers, Graham and Niagara, and initiated a technology investment study.

2.1 Graham Launch

In May 2017, Compute Ontario supported the launch of supercomputer Graham at the University of Waterloo. Through active participation in the procurement process and participation in planning, funding, national site selection competitions, and the organization of the actual media launch, Compute Ontario played a strong advocacy and supportive role in the launch of supercomputer Graham.

Graham provides expanded resources for researchers across the Canada who are working on a broad range of topics. These include, but are not limited to, artificial intelligence, genomics, and advanced manufacturing. With its extraordinary computing power and a storage system of more than 50 petabytes (50 million gigabytes), Graham supports researchers who are collecting, analyzing, or sharing immense volumes of data. The supercomputer is managed by SHARCNET – one of the consortia that Compute Ontario supports through funding, policy advocacy, and planning processes. More about the research that Graham enables can be found on pages 14 and 15 of this report.

2.2 Niagara Launch

The second supercomputer unveiled in 2017-18 was Niagara at the University of Toronto, installed and operated by the university's high-performance computing team, SciNet. Compute Ontario staff played a supportive role in the procurement processes, funding allocation and technical reviews that resulted in Niagara. As well, Compute Ontario played a lead role in the development of the media strategy and launch of Niagara that involved dignitaries, high-profile researchers such as Nobel Laureate John Polanyi, and politicians from provincial government.

Niagara is the most powerful computer system in Canada. It can run a single job across all 60,000 cores using a high-performance network which efficiently connects all of the compute nodes. Niagara was designed from the outset to be energy efficient and consider its environmental impact.

More about the research that Niagara enables can be found on pages 16 and 17 of this report.

Niagara launch event. Pictured L to R: Nizar Ladak, Dr. W.R. Peltier, Reza Moridi, Dr. Roseann O'Reilly Runte,and Dr. Vivek Goel

2.3 Provincial Study for Technology Investment

As an outcome of the 2016-17 strategic retreat and in response to a call to better define future infrastructure needs, Compute Ontario launched a project to develop a Technology Investment Strategy for Ontario. An advisory committee of 10 DRI experts was formed in the fall of 2017 with diverse representation from universities, research hospitals, ARC sites, and industry. The Phase 1 report -- including recommendations and data gathered from online surveys and in-depth interviews -- was completed in June 2018. Phase 2 is now underway and will include impact analyses of past DRI investment as well as a proposed future investment strategy.

Obtaining information on infrastructure needs is of particular importance as the number of DRI users in our province continues to increase — growing by 14% from last year to 2,533 people and representing 36% of the national user base.

Overall, researchers in Ontario and their teams accounted for over a third of the total national usage of Compute Canada resources. Ontario's usage of CPUs grew by 10% while the increased capacity resulting from newly installed GPUs resulted in a 60% increase in GPU usage.

Together, this information about users and data will support improved planning in the short and long terms to better serve the needs of Ontario's growing number of researchers who rely on DRI to address some of the world's most pressing problems.

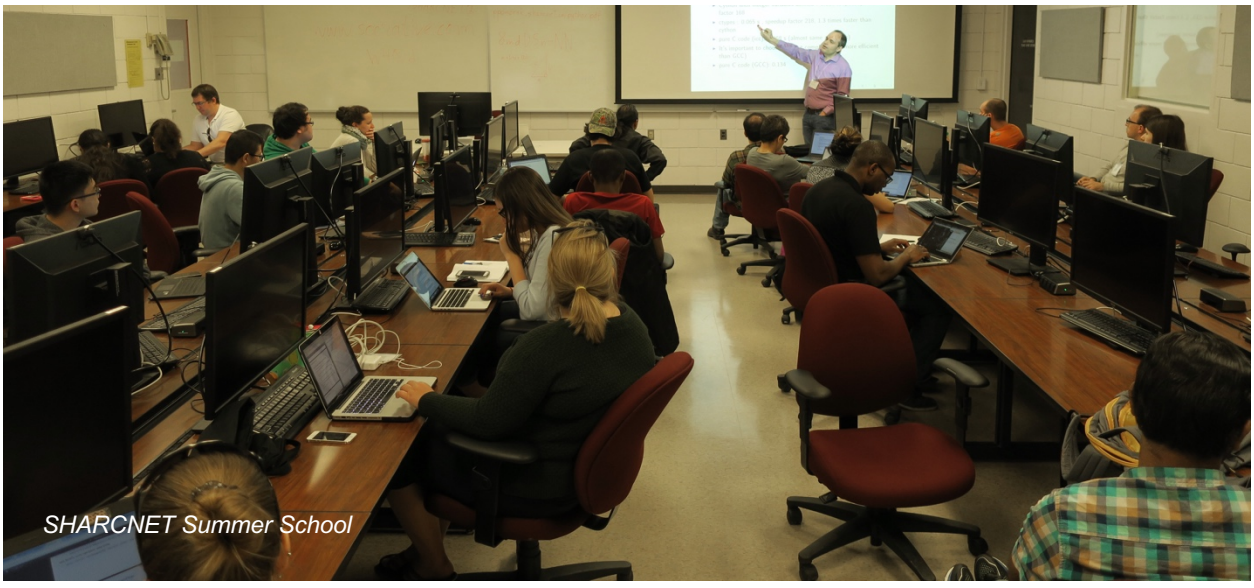
3.0 Increase Highly Qualified Personnel and Access to Data

Nothing is more fundamental to the success of a DRI ecosystem than the people who support it. Ontario's ability to increase the supply of highly qualified personnel (HQP) is predicated upon developing strategies that foster and promote the technical skills required to operate DRI systems. In that regard, Compute Ontario established a 12-member provincial advisory committee, including members with expertise in science and technology, the social sciences and humanities, public policy, and drawn from private industry, to advise in the development of an HQP study.

The study provided important insights concerning the challenges of attracting, training, and retaining people with skills in big data research in both the academy and the private sector.

The committee developed a standardized definition of HQP in DRI as: *Individuals with advanced degrees at the master's level and above, or equivalent experience, who are undertaking or facilitating research and development that applies advanced computational approaches and technologies in any research domain*.

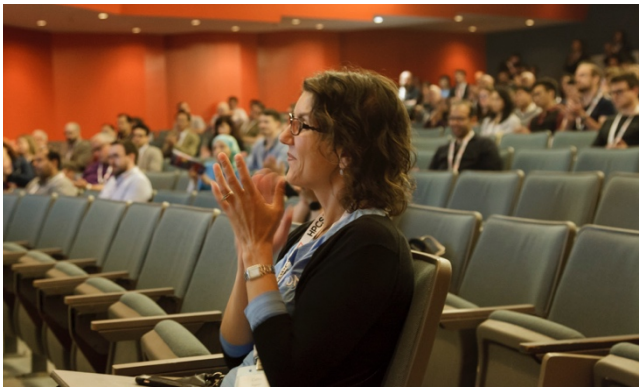
This definition was shared with the Vector Institute for Artificial Intelligence and helped inform their definition of HQP in AI. This creates common standards for Ontario organizations to support mutually beneficial goals. The final recommendations from the committee will guide the articulation of strategies to be implemented by Compute Ontario in 2018-19. This work will also provide baseline evidence for government and industry initiatives for increasing highly qualified personnel in DRI.



SHARCNET Summer School

4.0 Connect DRI Communities

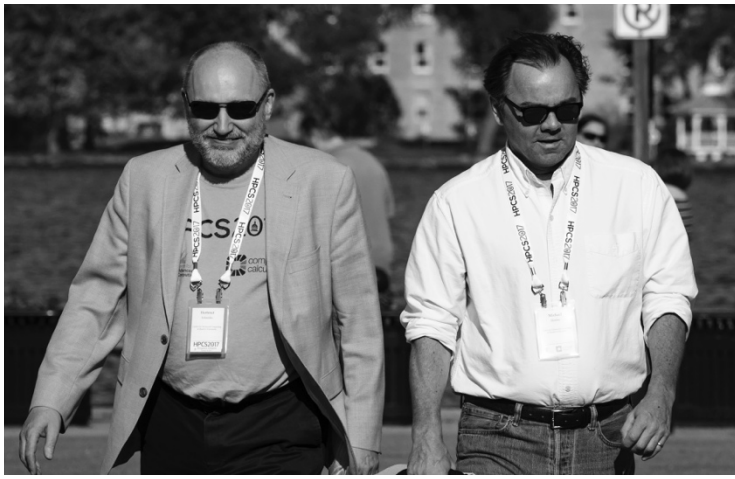
Another of Compute Ontario's strategic priorities is to operate as a focal point for connecting communities and constituents throughout Ontario's DRI ecosystem. Several initiatives in 2017-18 allowed Compute Ontario to bring together community leaders to celebrate successes, discuss best practices, and examine common issues.



For example, in partnership with the Centre for Advanced Computing, Compute Ontario co-sponsored the 2017 High-Performance Computing Symposium (HPCS) in Kingston, Ontario. This event brought together 319 delegates from across Canada, which included some of Canada's most-celebrated academics and users of DRI, such as Dr. Carolyn McGregor, Dr. W.R. Peltier, and Nobel Laureate Dr. Art McDonald. HPCS provided an opportunity for students to present posters, and for researchers from a diverse range of disciplines and career stages to come together to share research, best practices, new technologies, and to learn more about industry opportunities and partnerships.

As well, in 2017-18 Compute Ontario entered into several Memoranda of Understanding (MOU) with different organizations to help clarify the ways in which Compute Ontario works with these organizations. These MOU outlined collaborative opportunities for Compute Ontario and the organizations to work together. In 2017-18, Compute Ontario initiated or completed MOUs with:

- Centre for Advanced Computing
- HPC4Health
- SciNet
- SHARCNET
- Centre for Addiction and Mental Health
- Institute for Clinical Evaluative Studies
- Ontario Institute for Cancer Research



Photos from HPCS 2017 in Kingston, Ontario

Research Computing Impact

The Centre for Advanced Computing

Diabetes Research Gets a Security Upgrade

Almost 2.4 million Canadians live with diabetes, and the complications that may accompany this disease can be life threatening. Diabetes Action Canada works to transform health outcomes of people living with diabetes. One project that supports this mission is a secure environment in which researchers can run advanced analysis on health datasets to study how to better treat people with diabetes and prevent complications.

Conrad Pow, Senior Project Manager at Diabetes Action Canada, along with partners at the Centre for Advanced Computing, have delivered a proof-of-concept national diabetes repository. This repository holds de-identified patient-level data to conduct research using a secure environment for storage and analysis.

“We’re trying to answer questions as granular as what is an individual’s chances of lower-limb amputation or blindness as a direct result of diabetes,” says Mr. Pow. “A completely secure, remotely accessible repository allows researchers anywhere in Canada to request access to our datasets and run computational analysis.”

“Placing privacy first enables amazing projects,” stated Chris MacPhee, Assistant Director (Operations) at the Centre for Advanced Computing. “We’re Canada’s first Personal Health Information Protection Act (PHIPA)-compliant academic computing site and continue to provide secure solutions to amazing projects such as Diabetes Action Canada. Whether ending diabetes, curing cancer or improving mental health – we’re privileged to be part of dozens of medical projects to increase the quality of life of Canadians through platforms trusted by our partners, researchers, and patients.”

Before the national diabetes repository was operational, researchers typically had to request very specific cuts of anonymized data. This data would only be sent when a contractual obligation for a data destruction process was outlined. Now, researchers can request access to the repository by submitting a proposal which is then assessed by a research governing committee. Once a project is approved and has undergone a research ethics review, Diabetes Action Canada can grant access to project-specific anonymized datasets (which cannot be removed from the secure system) and a cloud environment for researchers to run analyses. The patient data is kept secure and a researcher cannot re-use a dataset without explicit permission.

“At Centre for Advanced Computing, we were able to leverage technological infrastructure, servers, secure connections, and technical expertise,” said Pow. “In the future, we’d like to add other datasets related to genomics and imaging, and CAC has

the computing power to do that rather than us having to spend research money on a large infrastructure build. It's scalable technology for scalable research."

Protecting Irreplaceable Cultural Artifacts Using Modern Big Data Strategies

Professor Susan Lord of Queen's University and Director of the Vulnerable Media Lab (VML) is working with her collaborators Dr. Rosaleen Hill (Director, Art Conservation) and Dr. Dylan Robinson (CRC Indigenous Arts) to preserve and remediate media archives created by several of Canada's diverse cultural communities.

Dr. Lord and numerous Queen's colleagues, including Drs. Robinson, Bertrand, and Hill, and members from Queen's Library and Archives, together with partners across the country, are collaborating on an SSHRC-funded project called Archive Counter-Archive, led by Dr. Janine Marchessault at York. The Centre for Advanced Computing (CAC) at Queen's University is a key partner in the VML and the Archive Counter-Archive project.

"The objective of this project is to work with groups across Canada to preserve community-based archives and help them make their histories available to people across the country and internationally," says Dr. Lord. "We are using modern digital tools to develop a sustainable platform to share and build sustainable data management systems to protect legacies. We aim to develop methods and processes to ensure this media and art history is preserved and made available according to culturally specific and ethically driven forms of access, thus engaging in new conversations about cultural heritage."

This kind of preservation requires massive storage resources that must be specially managed and curated, a problem that Mr. MacPhee and the team at the CAC are uniquely able to solve. "Our job is to make sure Dr. Lord doesn't have to worry about technology and can 'think big' with this project. What originally started out as 10 terabytes of storage has scaled up two or three orders of magnitude as the project has progressed over the years.

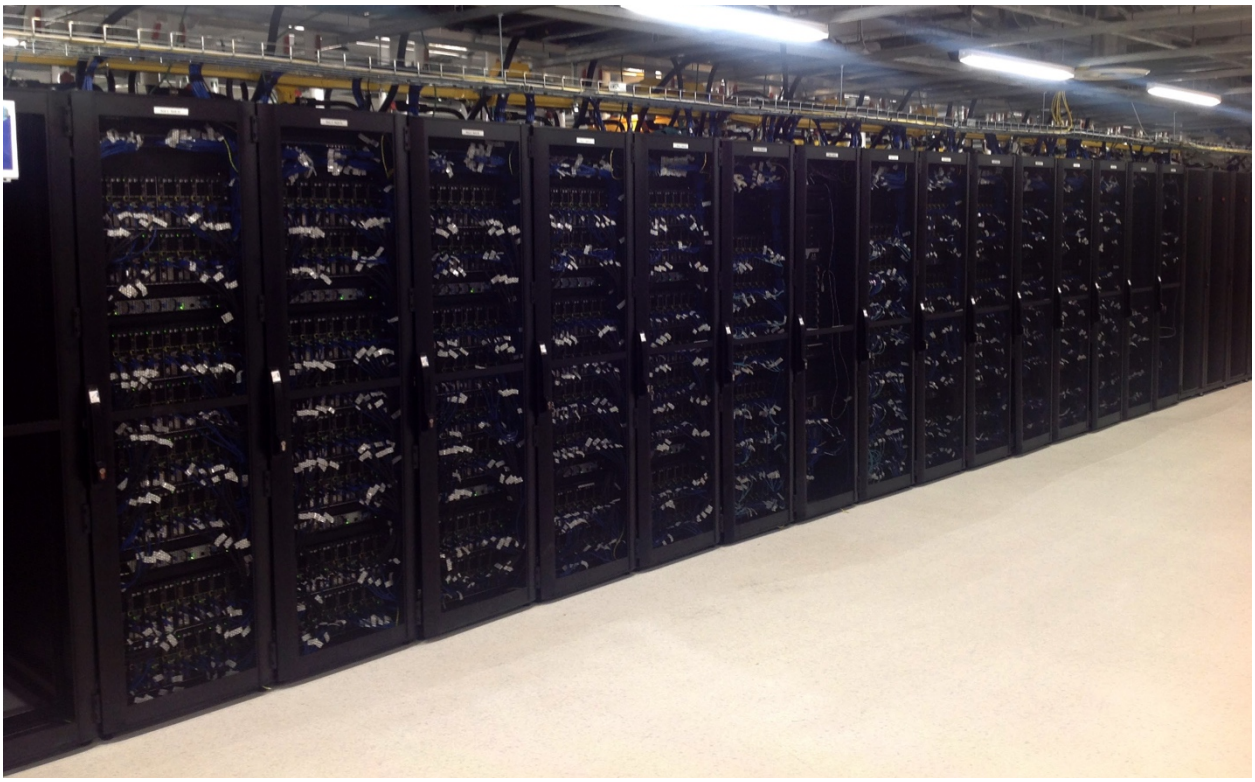
It means that this vulnerable media, which we can never get back if destroyed, can actually be archived. It's also exciting for us when we find an advanced computing solution to support research not traditionally associated with high-performance computing."

Working with an international advisory group including stakeholders from Indigenous and LGBTQ2 communities, Dr. Lord and the Centre are developing new practices for storage and curation that, more than standard security, include considerations for sensitivity of data and respect for the communities whose cultural histories are being stored. "We can host medical data, no problem," Mr. MacPhee notes. "And while rightfully there are policies and standards for how that data is handled, this initiative is

very different. It's not just about keeping data in a secure environment, it's ensuring cultural information is protected in a way that is meaningful to each community. It's about treating these archives with the respect they deserve."

As this project moves forward, the Vulnerable Media and Archive Counter-Archive partners are working with CAC and the Agile Humanities Group to create a platform that will be accessible online to students and researchers across Canada and beyond, which they hope will allow for conversations and allow communities to communicate about the work that has been done and is yet to be done.

"The technology, the storage, and platform sustainability give this project the necessary infrastructure," says Dr. Lord. "We need to preserve Canada's rich cultural history for future generations. And we can only do that guided by ethically driven protocols generated from affected communities and working with folks like Chris, and Jeremy Heil and Heather Home from Queen's Archives."



Data centre servers

HPC4Health

CanDIG provides one-stop analysis platform for secure genomics data

With strict controls over the use of and access to genomics data, how do we create a system to help researchers better understand the role genes play in manifesting into deadly diseases and cancers?

The Canadian Distributed Infrastructure for Genomics, or CanDIG, may be the solution. CanDIG is a national computation infrastructure for genomics analysis which aims to remove the barriers doctors and researchers face in discovering, exploring and analyzing the large genomics datasets that are increasingly generated in healthcare settings.

“Health data is sensitive and privacy controls are very high, as they should be. It’s hard to get data to cross provincial boundaries,” explains Dr. Jonathan Dursi, architect of CanDIG and research associate at the Hospital for Sick Children (SickKids) in Toronto. "Both federal and provincial regulations exist around the access and use of health data collected in hospitals and patient care.

“We are building a platform where people can perform analysis over the federated data by sending the analysis to the data without actually exposing or sharing the data explicitly. We’re trying to make sure researchers can understand and ultimately solve problems that are too complex for any one hospital, institution or province to tackle on their own,” said Dr. Dursi.

Through CanDIG’s completely distributed platform, data providers have complete control over who can access datasets and how much data they can access. The data has federated analysis built on top of application programming interfaces, or APIs, so it can be analyzed without being copied.

The goal is to enable researchers to analyze genomic data on a Canada-wide scale, connecting researchers with both the data and the compute power required to support breakthroughs in genomics research and genetic sequencing. The platform has wide-ranging implications. Not only will CanDIG help doctors and researchers better understand, treat, and possibly prevent disease, it will create a new approach to delivering secure, innovative and collaborative research with large genomics datasets.

“What makes CanDIG stand out on an international scale is the ability to federate the data across multiple jurisdictional boundaries. We’re building an infrastructure for a federation of health data that the rest of the world is going to need in the next five to 10 years."

With sites in Toronto, Vancouver, and Montreal, CanDIG is a collaboration between HPC4Health directors, Dr. Michael Brudno, professor at the University of Toronto and

Carl Virtanen, Director and Research Lead, UHN Digital; Guillame Bourque, an associate professor at McGill and Director of Bioinformatics at the McGill University & Genome Quebec Innovation Center (MUGQIC); and Steve Jones, Associate Director, Canada's Michael Smith Genome Science Centre and professor at UBC and SFU.

CanDIG is funded by CFI's Cyberinfrastructure Initiative which supports research data infrastructure projects that create customized, shared and integrated data resources capable of enabling leading-edge research on significant scientific, social and economic questions. The consortium is part of the Global Alliance for Genomic & Health Initiative (GA4GH), which is facilitating industry partnerships with major companies such as Google, Microsoft and Amazon, as well as other research centres.

Accelerating Health Research in the Data Safe Haven

A new partnership will help researchers and doctors apply cutting-edge technologies such as machine learning and high-performance computing to gain valuable insights from Ontario's population-wide datasets.

The Institute for Clinical Evaluative Services (ICES), HPC4Health, and Compute Ontario launched a pilot initiative called the Ontario Data Safe Haven (ODSH), a secure private cloud that will provide access to a wide range of datasets dating back to 1991, including as many as 13 million linkable health records.

"It's the first time we've been able to access data held by ICES in a secure, high-performance computing environment," says Dr. Michael Brudno, who is leading the ODSH pilot project. Brudno is a professor in the Department of Computer Science at the University of Toronto. He is also the Director of the Centre of Computational Medicine at the Hospital for Sick Children and Scientific Director of the HPC4Health Consortium. Brudno's focus is the use of high-performance computing to analyze the datasets of large numbers of electronic medical records (EMRs).

"The goal is to conduct analysis on the data and learn. For example, we can look at anonymized data of patients who have had frequent hospital visits and investigate what makes them similar and different from those in the general population. This can help us better prepare and plan future healthcare processes and policies.

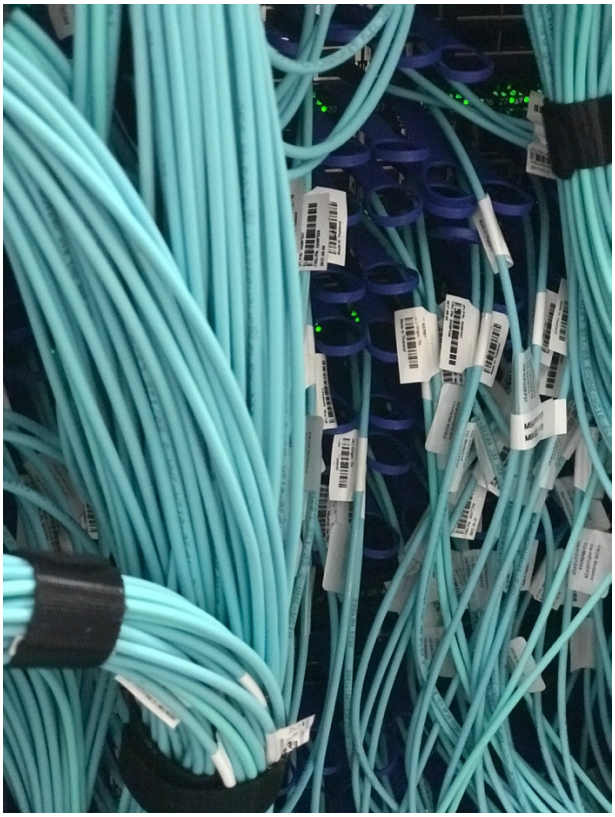
"Data Safe Haven is proof of the concept that you can bring sensitive, protected data sets into an environment where you can link them with other sensitive data and perform computations," he says.

The ICES Data Repository consists of publicly funded administrative health services records for the Ontario population. The data captured include patient-level administrative data from both health and non-health government ministries and agencies, and clinical data captured from hospitals. Other data include research studies

and biometric and biologic data that broadly represent the Ontario population. This pilot project will empower researchers to post, access and analyze data and build linkages that could lead to insights around more efficient, economical and appropriate health care for Ontarians. “Health care resources are precious and need to be properly managed across the province. By identifying the heaviest users of the system, we can make sure their needs are satisfied adequately and free up resources. Hopefully, this will result in shorter wait times at the doctor’s office and in hospital waiting rooms and emergency departments,” explains Brudno.

“Compute Ontario, one of the executive sponsors of this project, wants to better understand how computing can be applied to sensitive data. [Compute Ontario’s CEO and President] Nizar Ladak has been phenomenal in supporting us, both financially and as a member of our steering committee.”

The collaborative team includes Michael Schull, Alison Paprica, Carl Virtanen and Rob Naccarato. The project is funded by Compute Ontario, the Province of Ontario, the Canadian Foundation for Innovation, the SickKids Foundation, the University Health Network and Princess Margaret Cancer Foundation, and ICES.



Data centres can require over 40 kilometres of fibre optic network cables.

SHARCNET

Graham is Helping to Uncover More Than Secrets of the Brain

Terms like Artificial Intelligence and machine learning conjure images of machines that can think. This is obviously not the case as scientists are still puzzled as to how the human brain makes judgements or triggers even the simple act of blinking.

One neuroscientist from the University of Waterloo has made significant progress in deciphering how those basic brain functions might occur. Dr. Chris Eliasmith is the director of the Centre for Theoretical Neuroscience where the globally renowned SPAUN (Semantic Pointer Architecture Unified Network) brain model was developed. SPAUN is the world's largest functional brain model, simulating 6.5 million neurons, and is used to understand how the human brain works and how neurons communicate with one another.

The large-scale brain model's capabilities are outstanding with a digital eye for visual input, a simulated arm used to draw its responses and the ability to pass basic elements of an IQ test. Eliasmith and his team are accessing supercomputer Graham to program software specifically written to describe how each neuron functions within the brain model. Before Graham, it took Eliasmith 9,000 seconds to simulate one second of response.

SPAUN is running 10 times slower than real time; a 900% improvement in performance with the help of Graham's computational capacity. "This increase in speed will accelerate us towards building better neuromorphic computers and robots that are more natural for human interaction. Neuromorphic hardware can help connect models like SPAUN to the world and to give these models a body," says Eliasmith, who is currently working with several neuromorphic chips.

Being able to fuse these models with high-performance computers enables the myriad of innovations we hear about daily, such as assisting the development of life-saving medicine by identifying effects of pharmaceutical drugs on cognitive behaviour and improving artificial intelligence by understanding the algorithms of the human brain. The research conducted on SPAUN powers the type of neural analysis behind innovations such as the autonomous vehicle.

To learn more about SPAUN and the work for Dr. Eliasmith, visit the website for University of Waterloo's Centre for Theoretical Neuroscience.

Leading Graham User Develops New Systems to Reduce CO2 Impact

In recent years the increase of carbon dioxide (CO2) in the atmosphere has been well documented. Carbon dioxide is a heat-trapping gas, commonly associated with global climate change and is a key area of study for a researcher with the University of Waterloo who is designing next-generation tools in advanced manufacturing and energy sectors.

Dr. Luis Ricardez-Sandoval, an associate professor and Canada Research Chair in Chemical Engineering, uses supercomputer Graham to build innovative computational tools that will provide new insight and accelerate clean power-generation technologies. One aim is to mitigate CO2 emissions by improving the design and operational management of industrially relevant chemical processes.

Dr. Ricardez-Sandoval and his team of researchers are leveraging Graham's superior processing capabilities to retrieve new data by using multiscale modeling techniques which involve the coupling of systems that occur at different spatial and time scales.

By identifying these new occurrences and reactions, Dr. Ricardez-Sandoval is able to understand how fine (micro) scale events impact CO2 emissions at a larger (macro) scale.

"We are trying to find rare events that seldom occur in a system and for that we require a lot of computational power. Exploring and investigating reactions is time-consuming. My appointment to the Canada Research Chair has granted me more access to SHARCNET and Graham," explains Dr. Ricardez-Sandoval.

The rate at which Graham has been able to generate results has significantly advanced Dr. Ricardez-Sandoval's research. He is developing materials that can reuse the CO2 captured to generate new valuable chemicals and fuels (in a process known as CO2 utilization). "We're able to design materials that will allow us to improve CO2 capture for certain systems, and therefore reduce or curb global warming," says Ricardez-Sandoval.

His team generates results at a pace that has never been achieved until now, and which will improve the efficiency and optimization of systems economically, environmentally and socially. More details about Dr. Ricardez-Sandoval's research can be found at his research group website.

SciNet

Canada's Most-Powerful Research Computer Enables Ground-Breaking Research in its First Weeks of Use

Officially unveiled in March 2018, Canada's most powerful academic research computer made waves in the research community for the types of problems researchers can now solve. Located at the University of Toronto (U of T), the system is open to all Canadian university researchers.

For Niagara's first test, U of T's Richard Peltier ran a "heroic calculation" – one that shed new light on how the world's oceans physically function. "By devoting the entire machine, not only a portion of it, to this one calculation – that's why it's 'heroic,'" says Dr. Peltier, a U of T Professor of physics and Scientific Director of SciNet. "This is pure, curiosity-driven research."

"These kinds of calculations, which use the entire machine, weren't possible before," adds Dr. Daniel Gruner, Chief Technical Officer for SciNet.

"With the full power of 60,000 interconnected cores running a single, enormous calculation – this is a scale of inquiry that is truly profound."

The calculation was undertaken in partnership with researchers at the University of Michigan and the Jet Propulsion Lab at Caltech. Niagara will give researchers the computing power they need to study and find solutions to some of the world's biggest challenges. It will enable large-scale computation and simulation required for artificial intelligence, climate change research, ocean modelling, genomics, astrophysics, and other disciplines using big data research to fuel discovery.

Niagara is the fastest computer system in the country and is able to run a single job across all 60,000 cores thanks to a high-performance network which inter-connects all the nodes.

Training from ARC Experts Fuels Discovery of AI Methods to Map the Cosmos

When George Stein and Philippe Berger, doctoral candidates at University of Toronto (U of T), were first taught neural network computational methods, they were skeptical that it would work to create simulations of the universe in their specialized field of computational cosmology. “We build huge cosmological experiments to survey and observe more of the universe than ever seen before,” says Stein. “To help understand these observations we need to simulate what we expect to see. We’d always taken a problem, made approximations about the gravitational evolution in our simulated universes, and run those using the biggest computer you can get time on. That’s always been the way, and you need bigger and bigger computers to solve these equations as our observations get more complex.”

“Both the observational techniques [new types of telescopes] and simulations which study the large-scale structure of the Universe are becoming increasingly HPC/ARC intensive. For example, many fundamental questions in cosmology are most easily or straightforwardly addressed through direct simulation, but it takes weeks on the world’s largest supercomputers to simulate a single galaxy — forget the billions that this generation of galaxy surveys will observe,” says Berger.

University of Toronto’s high-performance computing centre, SciNet, offers training for students across the country who are interested in bringing computational methods into their research. Applications analyst and plasma physicist Erik Spence is the instructor for the Neural Networks courses that SciNet offers. When asked about the difference between traditional ARC and Artificial Intelligence, Spence states: “You need both. You definitely need your old-school calculations before you can even think about using a neural network. When I was teaching George and Phil, they had a perfect problem that fell into the framework really nicely.”

“We completely by-passed the massive computations usually needed for these types of simulations and got straight to the answer,” says Stein. “We were actually able to create the cosmological simulations a thousand times faster,” Stein added.

The benefit of this accelerated speed, of course, is that the simulations take less time, power and resources. In this case, Stein and Berger needed cosmological maps used to aid telescopes that look into the far reaches of the galaxy. For some uses, the maps can be relatively approximate.

Both Stein and Spence were clear, however, that using neural network or artificial intelligence methods for all aspects of research simply doesn’t work. Spence notes that “If all you need is the end of the research, then this method is perfect. If any of the data in between in order to understand the process of how you got to the end – which is often the case — you need the big systems, and it takes more time. Neural networks are inherently handicapped by the fact that you need big datasets to train them against.”

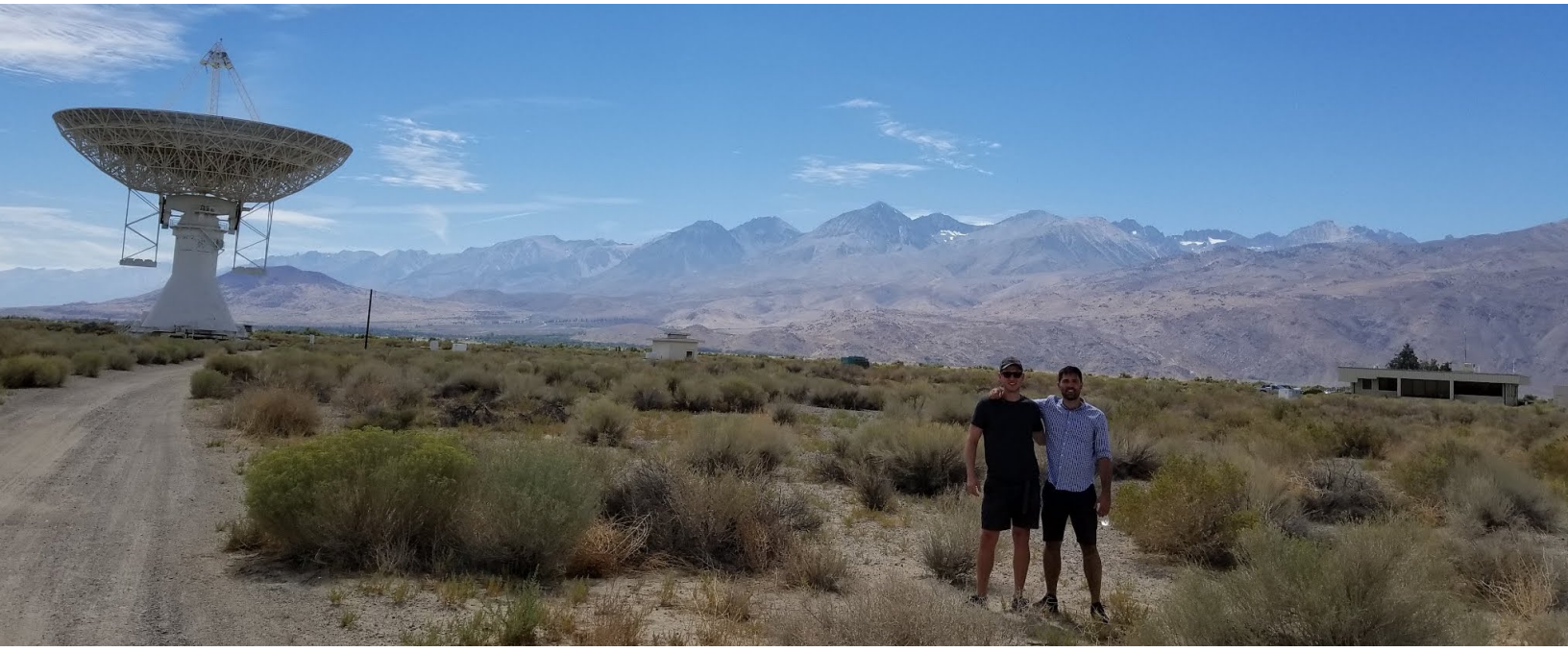
“Neural networks are an incredibly accurate approximation method, and we have shown this is the case even for gravity which is highly non-linear. Understanding why this is the case is currently a major focus of machine learning. Gravity is a great example to study since the exact answer can be simulated (although expensively). This provides a solid reference to compare to the output of the Neural Network and which helps in trying to understand its inner workings,” says Berger.

Both Berger and Stein have taken several courses with SciNet and appreciate the depth that it has brought to the way they approach their research. “When I started my PhD, I barely knew how to code. I had the option to take a few SciNet courses as modules for course credit towards my program. These first covered introductory coding skills, then scientific programming in C++ and Python, and gave me a solid foundation on which I could build. I was later able to take more advanced high-performance computing courses and the Neural Networks course that Erik taught,” says Berger.

For all of the Ontario consortia, including HPC4Health, Centre for Advanced Computing, SciNet, and SHARCNET, training the next generation of researchers is foundational to their work, in addition to maintaining some of Canada’s most powerful computing systems.

In Ontario in 2017-18, the highly technical and knowledgeable instructors and staff delivered 25,321 teaching hours to 9,201 students.

To read the paper resulting from this research, visit https://arxiv.org/abs/1805.04537.



George Stein and Philippe Berger in California

Funding

For the period April 1, 2017 to March 31, 2018, which represented the third operating year of Compute Ontario, the Ministry of Research and Innovation provided funding in the amount of $1,58-million. There were also eligible carry forward funds in the amount of $158-thousand resulting in a total operating budget of $1,738-million for the fiscal year 2017-18. All funding was received on schedule.

All staffing was in place during 2017-18, and several key initiatives were undertaken. The outcomes of these initiatives are highlighted in the key achievements section of this report. Over 90% of the available budget was spent in 2017-18, with $150-thousand to be carried forward into the next fiscal year, which is within the eligible carry-forward limit.

Financial Report

A report on Compute Ontario finances is included in Table 1. Compute Ontario completed the year with expenditures totalling $1,587,311. It did not spend the entire $1,738-million that was allocated in 2017-18, which included the original budget of $1,580-million plus eligible carry-forward from 2016-17 of $158-thousand. The difference of $150,689 will be carried forward to next fiscal year as it fits within the eligible 10% carry-forward limit set by the Ministry.

The commentary on individual line item variances are noted in the Table, with the most significant variances highlighted below:

- Salary savings of $68-thousand were realized from turnover of an administrative position in January 2018. The position was subsequently re-filled in April 2018.

- The Board and Corporate Secretary function, which had previously been contracted to a third party, is now being fulfilled by permanent Compute Ontario staff, which resulted in budget savings of $30k.

- The budget for conferences, meetings and events was underspent by $22k as staff were busy preparing for the inaugural Ontario Advanced Research Computing Congress (OARCC) event in May 2018.

- The special studies category was overspent by $100-thousand, largely due to the HQP study being deferred from the previous fiscal year. Special studies expenditures also included Compute Ontario's Technical Study and the balance of expenditures related to the Data Safe Haven project.

- The media and materials category was overspent by $38-thousand largely due to Compute Ontario's rebranding initiative which required additional design work and website updates.

Table 1 – Compute Ontario Financial Variance Report

For the period ending March 31, 2018

Expense Category	Budget 2017-18	Actuals Mar. 31, 2018	Variance Mar. 31, 2018	Comments
Salaries, Wages & Benefits	$980,805	$912,346	$68,459	turnover savings
Staff Training	$19,500	$5,761	$13,739	priority on policy development and governance
Board & Corporate Secretary	$30,000	$0	$30,000	consulting contract ended prior fiscal; function managed via existing CO staff
Conferences, Meetings & Events	$80,000	$57,547	$22,453	conferences, events, annual symposium; staff attendance at conferences
Premises	$66,432	$71,319	($4,887)	rent based on MaRS contract
Special Studies	$286,751	$387,058	($100,307)	balance of Data Safe Haven, HQP Study (deferred from prior fiscal), Tech Study, consortia-funded proposals
Professional Services	$10,000	$22,769	($12,769)	strategic planning & facilitation; additional administrative and communications support
Legal, HR, Procurement	$15,000	$12,588	$2,412	legal advice for CO as required
Accounting Financial	$0	$0	$0	Western donating backend services as in-kind
Insurance	$15,000	$5,173	$9,827	CGL, D&O, property insurance
Audit	$7,500	$13,613	($6,113)	using Western University's auditor; additional audit work requested
IT	$29,612	$27,785	$1,827	small hardware, software, phones, mobiles, IT support, cloud services, annual certificates
Media & Materials	$25,000	$62,845	($37,845)	creative services, media monitoring & advertising, CO rebranding, web updates
Office Administration	$14,400	$8,507	$5,893	supplies, printing, courier, postage
Total Expenses	**$1,580,000**	**$1,587,311**	**($7,311)**	
Carry Forward	$158,000	$0	$158,000	eligible c/f from 2016-17
Total Budget	**$1,738,000**	**$1,587,311**	**$150,689**	eligible c/f from 2017-18

Schedule of Expenditures

COMPUTE ONTARIO / CALCUL ONTARIO

MINISTRY OF RESEARCH AND INNOVATION FUNDING

Year ended March 31, 2018

KPMG LLP
140 Fullarton Street Suite 1400
London ON N6A 5P2
Canada
Tel 519 672-4800
Fax 519 672-5684

INDEPENDENT AUDITORS' REPORT

To the Management of Compute Ontario / Calcul Ontario and the Minister of Research and Innovation

We have audited the accompanying schedule of expenditures of Ministry of Research and Innovation (the "Ministry") funding of Compute Ontario / Calcul Ontario and notes comprising significant accounting policies and other explanatory information for the year ended March 31, 2018 (the "Schedule"). The Schedule has been prepared by management based on the financial reporting provisions in Section F of the agreement dated April 1, 2015 between Compute Ontario / Calcul Ontario and the Ministry (the "Agreement").

Management's Responsibility for the Schedule

Management is responsible for the preparation of the Schedule in accordance with the financial reporting provisions in the Agreement, and for such internal control as management determines is necessary to enable the preparation of the Schedule that is free from material misstatement, whether due to fraud or error.

Auditors' Responsibility

Our responsibility is to express an opinion on the Schedule based on our audit. We conducted our audit in accordance with Canadian generally accepted auditing standards. Those standards require that we comply with ethical requirements and plan and perform the audit to obtain reasonable assurance about whether the Schedule is free from material misstatement.

An audit involves performing procedures to obtain audit evidence about the amounts and disclosures in the Schedule. The procedures selected depend on our judgment, including the assessment of the risks of material misstatement of the Schedule, whether due to fraud or error. In making those risk assessments, we consider internal control relevant to the entity's preparation of the Schedule in order to design audit procedures that are appropriate in the circumstances, but not for the purpose of expressing an opinion on the effectiveness of the entity's internal control. An audit also includes evaluating the appropriateness of accounting policies used and the reasonableness of accounting estimates made by management, as well as evaluating the overall presentation of the Schedule.

We believe that the audit evidence we have obtained is sufficient and appropriate to provide a basis for our audit opinion.

Opinion

In our opinion, the schedule of expenditures of Ministry of Research and Innovation funding of Compute Ontario / Calcul Ontario is prepared, in all material respects, in accordance with the financial reporting provisions in Section F of the agreement dated April 1, 2015 between Compute Ontario / Calcul Ontario and the Ministry.

Basis of Accounting

Without modifying our opinion, we draw attention to note 2 to the Schedule, which describes the basis of accounting. The Schedule is prepared to provide information to The Ministry of Research and Innovation. As a result, the Schedule may not be suitable for another purpose.

Restriction on Use

Our report is intended solely for the Management of Compute Ontario / Calcul Ontario and The Ministry of Research and Innovation and should not be used by parties other than Compute Ontario / Calcul Ontario and The Ministry of Research and Innovation.

KPMG LLP

Chartered Professional Accountants, Licensed Public Accountants

London, Canada

September 25, 2018

COMPUTE ONTARIO / CALCUL ONTARIO

Notes to Schedule of Expenditures

March 31, 2018

1. Project Description:

Compute Ontario / Calcul Ontario signed an agreement dated April 1, 2015 with The Ministry of Research and Innovation (the “Ministry”) for the operating costs of Compute Ontario (the “Project”).

Under the agreement, The Ministry of Research and Innovation will provide a maximum of $8,500,000 cash contribution to the Project for the project period, which ends on March 31, 2021. The project objective is to explore the merits of developing a high performance computing and big data strategy for the province.

The schedule of expenditures of Ministry of Research and Innovation funding of Compute Ontario / Calcul Ontario (the “Schedule”) presents Compute Ontario / Calcul Ontario’s portion of eligible expenditures incurred for this Project during the reporting period.

2. Basis of accounting:

The Schedule has been prepared in accordance with the financial reporting provisions contained in Section F of the Agreement dated April 1, 2015 between Compute Ontario / Calcul Ontario and The Ministry of Research and Innovation.

The Schedule has not been prepared in accordance with the presentation principles or the presentation of all the financial statements and related note disclosures required by Canadian accounting standards for not-for-profit organizations.

The Schedule was prepared to assist the Organization to meet the reporting requirements of the Ministry. The schedule is intended solely for the use of Compute Ontario / Calcul Ontario and the Ministry of Research and Innovation. Accordingly, readers are cautioned that the Schedule may not be suitable for another purpose.

3. Significant accounting policy:

Expenditures:

Expenditures are eligible under the Project if they were incurred in the period of the Project and are directly related to the completion of the Project.

Board Members & Committees

Fiscal year 2017-18 saw Compute Ontario achieve a full board complement with the appointments of Drs. Carolyn McGregor and Charmaine Dean. There was also greater movement towards more formal governance and planning procedures as demonstrated through the implementation of an integrated board and sub-committee work plan. The move towards more formal mechanisms of monitoring will allow the board to better support strategic oversite of the organization.

Mark Daley
Board Chair

Ranil Sonnadara
Vice Chair

Sylvain Charbonneau
Director

Charmaine Dean
Director

Vivek Goel
Director

Warren Keith Helland
Director

Shannon MacDonald
Director

Carolyn McGregor
Director

Catherine Middleton
Director

Daniel Sinai
Director

Salim Teja
Director

Susan Ursel
Director

Dereck Whitmell
Director

Compute Ontario
2017-18 Annual Report

This report was published by:

Compute Ontario
661 University Ave.
MaRS West Tower, Suite 1140
Toronto, ON M5G 1M1
computeontario.ca | info@computeontario.ca