

Day 1 - Tuesday, May 27

9:30 AM – Keynote: AI's Role in Digital Research

Speaker: Dr. Sidney Shapiro, Business Analytics Professor and AI Researcher, University of Lethbridge

Location: BMO Amphitheatre

Description:

Artificial intelligence (and more recently, Generative AI) is rapidly transforming research across disciplines in Canada—shaping how studies are designed, how data is analyzed, and how knowledge is produced. This evolution is bringing about fundamental changes to research workflows, methods, and infrastructure. At the same time, it raises critical questions about how public research institutions can engage with AI meaningfully and sustainably, particularly in the context of growing disparities in digital research capacity.

This keynote will examine the current and emerging role of AI in Canadian research, with a focus on infrastructure challenges and strategic considerations. While private industry has advanced rapidly in developing generative AI technologies using vast computational resources, academic institutions face structural limitations that affect their ability to keep pace. Limited access to high-performance computing, specialized engineering expertise, and integrated systems continues to restrict what public institutions can build, implement, or study.

As generative AI becomes both a method for research and a subject of research itself, the question becomes: what infrastructure and strategy are needed to support this work in the academic context?

The talk will highlight how AI is being used as a research tool to support discovery, enhance analysis, and enable new forms of scholarly communication. It will also address the increasing significance of generative AI as an object of research, with active inquiry taking place across domains such as ethics, linguistics, education, and computational science. This dual role presents a set of unique challenges for research planning, funding, and institutional coordination.

Key issues to be addressed include uneven access to advanced computing infrastructure, gaps in workforce readiness, and the growing need for national coordination around shared resources such as compute, data, and software. The session will also explore how institutional and policy frameworks can support responsible AI adoption, with attention to governance, training, and ethical oversight.

The keynote will conclude by outlining strategic priorities for building sustainable AI capacity in Canada's research ecosystem. These priorities extend beyond technical infrastructure to include talent development, interdisciplinary collaboration, and investment models that reflect both the promise and the complexity of AI in research.

11:00 AM – Strengthening Global Training and Skills Development Partnerships

Speakers: Catherine Di Vita (Digital Research Alliance of Canada), Kathryn Unsworth (Australian Research Data Commons)

Location: BMO Amphitheatre

Description:

In early 2024, a proposal to advance bilateral collaboration around national training strategies and frameworks between the Alliance and the Australian Research Data Commons (ARDC) was introduced. The agreement was formally signed in February 2025, effective through to December 31, 2026, launching the start of a two-year knowledge and staff exchange pilot.

This pilot is in the form of a series of bilateral staff exchanges both in Canada and Australia and involves each host country showcasing their respective work in-country. The objective of the pilot is to evaluate the efficacy and impact of on-site international staff exchanges, documenting the lessons learned along the way. By comparing and contrasting the skills development landscapes of ARDC and the Alliance, the pilot aims to identify points of convergence and opportunities for collaboration.

Additionally, the pilot leads will share and document their experiential knowledge about training, skills, and workforce development from their respective countries (research sectors) and leverage the knowledge and understandings from their local training communities.

In this presentation, we will provide an overview of each organization, an outline of the areas of priority identified in the first exchange, along with an update of the progress made thus far. We will then discuss how we plan to develop and further the work with a focus on international stakeholder relationships to expand skills and workforce development in our two jurisdictions.

Finally, we invite all stakeholders and interested audience members to provide any feedback on this exciting journey of international collaboration. Your insights and experiences are invaluable to the success of this pilot. Please participate in our interactive poll to share your thoughts on skills-related topics that currently resonate with you. Additionally, we welcome questions during Q&A time and/or post-presentation.

11:15 AM – Optimizing Training Reporting in Canada's DRI Ecosystem

Speaker: Ines Hessler, CTO, ACENET

Location: BMO Amphitheatre

Description:

Over the past decade, training has become an increasingly important part of our mission, with demand continuing to grow. To enhance our offerings and better leverage the data we collect, we launched a pilot project aimed at streamlining and standardizing training data collection and reporting processes. Since we don't yet have full confidence in which metrics will best meet the Alliance's national reporting needs, this project also includes a collaborative component to align our approach with broader Federation expectations.

11:30 AM – User Certification: Let's Do It!

Speaker: Sergey Mashchenko, High Performance Computing Technical Consultant, SHARCNET

Location: BMO Amphitheatre

Description:

With the current setup, users of the national systems (both PIs and sponsored users) can log in to the clusters and start using them the moment their account is approved. No training is currently required. This occasionally creates issues for other users of the system, if/when a new user inadvertently misuses a valuable shared resource.

Here in SHARCNET, we propose that the proper solution for the above issues would be an introduction of obligatory New User training, enforced via a User Certification mechanism.

This workshop will allow the audience to test the principal component of the user certification mechanism—self-paced user certification courses. We hope a fruitful discussion will follow, focusing both on the technical (Moodle, CCDB, cluster setup) and political (can we/should we force this on new users? the best ways to sell it to users?) aspects of the proposal.

11:45 AM – Redefining Training Discovery: An Introduction to Explora, the New Training Portal

Speakers: Catherine Di Vita (Digital Research Alliance of Canada), Ramses van Zon (SciNet), Chris Want (University of Alberta)

Location: BMO Amphitheatre

Description:

The DRI Workforce offers a wide array of training across multiple disciplines, including ARC, RDM, and Cybersecurity. However, these training efforts are scattered and lack a centralized platform for discovery and communication. To address this concern The National Training Discovery Portal is projected to be released at the end of April.

The National Training Discovery Portal aims to address the challenges of locating and registering for training events, as well as accessing training materials, which have been reported as difficult and inefficient. By developing a centralized portal, this project seeks to enhance the visibility and accessibility of upcoming regional and local training events across the country, specifically in the areas of RDM, ARC, RS, and cybersecurity. The initial prototype was created by Chris Want at the University of Alberta. This project has now been taken on by the National Training Coordination Council.

The portal will serve as a comprehensive discovery tool for both current and past training events and materials, replacing the existing ARC training calendar. Ultimately, the project will improve communication of training opportunities and significantly enhance the overall learner experience for researchers and professionals nationwide.

The team has decided to take a phased project approach to the development of the training portal. The first phase, projected release in April, will aggregate events being held by training providers across the nation for discovery in a single searchable platform. The second phase focuses on building a repository of training materials.

In this presentation, we aim to discuss the development process of phase one, identifying challenges and solutions employed. We will also showcase the portal and discuss plans for phase two development and implementation. The 1.0 launch will rely heavily on user feedback for continuous improvement and to inform the development of phase 2. So, we will then open the presentation for feedback from the community.

Stream 1: Advancing HPC

1:00 PM – The Future of HPC in the Age of AI

Speaker: Roman Baranowski, DRI Software Architect, ARC UBC

Location: BMO Amphitheatre

Description:

In this talk /panel discussion I/we would like to talk/discuss about the architectural differences between AI and 'typical' HPC based workflows and try to address the challenges we are facing and find a path forward. We as a Alliance Community have to be ready so let's talk....

1:30 PM – Kubernetes as a Natural Evolution of HPC

Speaker: Shaun Bathgate, University of Victoria

Location: BMO Amphitheatre

Description:

Traditional high-performance computing (HPC) platforms have long powered

groundbreaking, data-intensive research, leveraging specialized schedulers and containerization tools like Apptainer (formerly Singularity). Increasingly, however, research computing demands extend beyond traditional batch processing, such as supporting interactive workflows, short-lived services, and rapidly evolving collaborative projects. Kubernetes, initially developed to orchestrate cloud-native applications at scale, is uniquely suited to address these emerging needs.

This session introduces Kubernetes not as a replacement, but as a natural and complementary evolution of traditional HPC. By clearly distinguishing each platform's strengths, attendees will gain insight into how HPC excels in orchestrating large-scale, massively parallel batch jobs, while Kubernetes provides a robust environment for managing interactive workloads, microservices, and dynamic, event-driven processes.

Through practical examples, we will demonstrate how Kubernetes can seamlessly extend existing HPC capabilities, bridging workflow gaps without disrupting established practices. Attendees will discover how Kubernetes simplifies the deployment of interactive applications and ephemeral services, areas historically challenging for traditional HPC schedulers, while allowing HPC teams to maintain their primary focus on computational research tasks.

Further, the talk will highlight that adopting Kubernetes within research computing does not entail increased complexity for end-users. With a well-managed Kubernetes cluster, infrastructure details are abstracted away by platform teams, freeing researchers to concentrate solely on creating container images and running their workloads.

By illustrating how Kubernetes and HPC environments naturally complement and enhance each other, attendees will leave equipped with new ideas on how integrating these platforms can improve flexibility, scalability, and overall efficiency in research computing for Canadian post-secondary institutions and their research partners.

1:45 PM – Beyond MPI and OpenMP: Teaching Parallel Programming in Modern Research Computing

Speaker: Alex Razoumov, Simon Fraser University

Location: BMO Amphitheatre

Description:

For decades, MPI and OpenMP have been the standard for teaching distributed- and shared-memory parallel programming in HPC. While these basics are still introduced in introductory HPC courses and researchers are directed to MPI/OpenMP/CUDA courses taught elsewhere in the Federation, the focus over the past decade has shifted to higher-level frameworks and languages.

This transition began in 2017 with *Introduction to High-Performance Computing in Chapel*, which later became the foundation for HPC Carpentry's Chapel course.

Designed for novice HPC users, it provides high-level abstractions for task and data parallelism. In 2024, the course was expanded to include GPU computing, unifying parallelism across multicore processors, multiple nodes, and GPUs.

In this session, we will discuss these developments and highlight the benefits of modern tools for teaching parallel programming in research computing.

2:00 PM – Research Computing and Data Storage at Scale: 10 Years, 140+ Webinars

Speaker: Alex Razoumov, Simon Fraser University

Location: BMO Amphitheatre

Description:

Over the years, we have delivered numerous research computing webinars on specialized topics not typically covered in our training courses. Many of these webinars focus on tools that can greatly enhance workflows for researchers running large numerical simulations or managing extensive datasets.

This session will highlight key tools at the intersection of High-Performance Computing (HPC) and Research Data Management (RDM) that researchers are encouraged to integrate into their work, such as in-situ visualization, lossy 3D data compression, distributed storage of large datasets, and DAR (Disk ARchiver).

Stream 2: Supporting Excellence

1:00 PM – An Open-Minded Discussion About Communication to the Research Community

Speaker: Marie-Helene Burle, Simon Fraser University

Location: MB 3.435

Description:

The Alliance has two important audiences: stakeholders and researchers. Stakeholders because without money nothing can be done; researchers because without them what's the point? While communication to stakeholders is mature, communication to the research community is lacking. I invite you to a roundtable to discuss this.

1:30 PM – Know Your Clients: Through the Traces of Their Work

Speaker: Mark Hahn, SHARCNET-McMaster, Compute Ontario

Location: MB 3.435

Description:

Understanding the researchers' needs is critical to providing the infrastructure that enable them to maximize their contributions to their respective research areas. When researchers use DRAC resources they leave traces about their needs and successes. This starts with their user and group account details but extends out into their job records and research outputs. What can we tell about the needs and successes of our researchers by aggregating data from the various information sources that we currently have (e.g. user profiles, allocations, job records, publications, support tickets, etc) and what else could we learn by implementing new data collections that would give us more fine grained access to other key properties of research production (e.g. storage use, software utilization, job performance metrics, training registrations and outcomes, etc).

1:45 PM – Optimizing Resource Utilization in HPC: Tackling Waste in the Alliance Ecosystem

Speakers: Moïra Dion and Hélène Gingras, Calcul Québec

Location: MB 3.435

Description:

This talk will address the challenges of resource waste in the Alliance ecosystem, focusing on inefficiencies in CPU, GPU, and memory usage on Narval. Due to growing demands and varying user experience levels, inefficient resource utilization has led to delays in job scheduling and reduced research productivity.

We initiated a three-month project to optimize intervention processes, quantify the financial impact of waste, and identify the types of jobs causing inefficiencies. The project highlights the importance of proactive engagement and standardized workflows in reducing resource waste and ensuring equitable use of shared computing resources.

2:00 PM – Monitoring GPU Utilization and Waste Management Best Practices

Speaker: Nikolai Sergueev, Calcul Québec/Université de Montréal

Location: MB 3.435

Description:

Stream 3: Platforms for Progress**1:00 PM – The Colibri Initiative: Towards Accessible Cloud Services**

Speakers: Sarah Cameron-Pesant and Lydia Vermeyden, Calcul Québec

Location: MB 3.445

Description:

Un besoin important a été identifié dans la communauté de recherche - notamment en sciences humaines et sociales, mais aussi dans plusieurs autres disciplines - pour des logiciels libres prêts-à-l'emploi qui facilitent la collaboration interinstitutionnelle.

L'initiative Colibri est menée par Calcul Québec en vue de développer des nouveaux services infonuagiques qui soient plus accessibles pour les chercheuses et chercheurs qui n'ont pas les compétences techniques et/ou les ressources nécessaires à l'utilisation de nos services actuels de type infrastructure-en-tant-que-service. L'objectif de cette initiative est, à terme, de démocratiser l'accès à certains logiciels libres en soutenant davantage les chercheuses et chercheurs dans leur installation, configuration et maintenance.

Un premier projet pilote a été mené en 2023-2024 (CRSH Connexion 2022, Meurs *et al.*) dans le cadre de l'Initiative conjointe visant le renforcement de la capacité de gestion des données de recherche" (<https://www.sshrc-crsh.gc.ca/results-resultats/recipient-recipiendaires/2022/cg-sc-may-2022-fra.aspx>). Pendant ce projet pilote, nous avons fourni des machines virtuelles préinstallées avec les outils Mattermost et Nextcloud à quelques groupes de recherche. L'expérience s'est avérée positive et riche en apprentissages.

La présentation se déroulera en trois parties :

1. Nous aborderons notre vision pour des services infonuagiques plus accessibles basée sur les apprentissages que nous avons faits et les défis que nous avons rencontrés pendant le projet pilote.
2. Nous vous présenterons plus en détail le travail réalisé à date, pour fournir notre premier service accessible: Nextcloud.
3. Nous vous inviterons à partager votre expérience en termes de services infonuagiques accessibles et des besoins des groupes de recherche avec lesquels vous travaillez en termes d'outils clés-en-main, etc.

1:30 PM – Does Your DMP Template Need a Refresh? A Cyber-Aware Approach to RDM

Speaker: Luc Letarte, Research Cybersecurity and Compliance Specialist, UBC

Location: MB 3.445

Description:

Data Management Plans (DMPs) – they seem simple, but are they? How can we

navigate the complexities of data security? How can we design a one-size-fits-all template for diverse research needs while ensuring requirements are met?

In today's digital age, the requirements for DMPs have evolved, with a heightened focus on cybersecurity, privacy, and compliance. This session will present updates to modernize the UBC DMP template with data security in mind, as well as meeting the increasing demands of funding agencies and regulatory bodies.

1:45 PM – DMP Assistant: Stable Present, Evolving Future

Speaker: Marcus Closen, Product Lead, Digital Research Alliance of Canada

Location: MB 3.445

Description:

DMP Assistant is the Canadian solution for data management planning. It provides a national, bi-lingual platform for Canadian researchers. This presentation considers the past, present, and future of DMP in Canada and the potential of greater collaboration between platforms like DMP Assistant and other DRI assets in the space.

2:00 PM – Adapting Open-Source: A Learning Opportunity for Research Support and Developers Alike

Speakers: Danica Evering and Richie Motorgeanu, McMaster University

Location: MB 3.445

Description:

More researchers than ever are accessing our digital research infrastructure and services: Advanced Research Computing, Research Data Management, and Research Software Development. Using our new README Creator tool as a jumping off point, this interactive talk discusses the challenges and benefits of developing and adapting open-source software to develop tools that make researchers' lives easier.

We'll start with setting the table about research support in Canada and noticing issues that researchers are having. The Tri-Agencies' new Research Data Management Policy lays the groundwork for researchers soon being required to deposit data into a research data repository: an online location for long-term data storage. But if data is going to be correctly understood two years from now (either by yourself or by other researchers) you need a README file. This important documentation provides details about your files, how they're organized, how the data was collected, and how it can be reused. However, researchers often fail to include a README file. Although a text-based README template is one solution, researchers could benefit from helptext and links out to support.

We'll go into the details of how we translated an open-source tool into something institutionally specific; in our case, how we adapted an open-source software citation file format app into a README generator that guides users through the README creation process. We will discuss how we have benefited from having a multidisciplinary team working together—across pillars, as research support staff and student workers—to **address researcher needs** and dig deeper into some of the **challenges and learning opportunities that we encountered along the way**. We will emphasize making sure students have opportunities to ideate, contribute to, and develop technical solutions for projects makes for hands-on learning of future qualified professionals.

We'll save some time at the end to of our 15 minutes to start identifying researcher needs in our different groups that could be addressed with open-source tools, training materials, or documentation. We'll also open the floor to the audience to chime in with any experiences they've had either working with colleagues from other disciplines or adapting open-source tools, materials, and documentation.

2:45 PM – Research Data Management Demystified: Your Questions, Our Answers!

Speakers: Amanda Tomé, Marcus Closen, Tristan Kuehn, Natalie Williams, Lee Wilson

Location: BMO Amphitheatre

Description:

The AMA (Ask Me Anything) style interactive session about research data management (RDM) practices and services will facilitate discussion, dispel myths about RDM practices, and foster mutual understanding between colleagues. The goal is to create greater connection and understanding between the different functional areas of the DRI ecosystem, to identify avenues for collaboration between RDM and other areas, and to understand gaps in knowledge that need further investigation.

3:15 PM – What Really Happens at FRDR? An Overview of Curation and Preservation Services

Speaker: Natalie Williams, Digital Research Alliance of Canada

Location: BMO Amphitheatre

Description:

Have you ever gone looking for a file and couldn't find it? Have your postdocs or students struggled to complete a secondary analysis using a dataset created by another trainee who has since left the lab?

This session will discuss how curation and preservation activities prevent these issues, enabling secondary use, reproducibility, and extending the life of research data.

3:45 PM – Building an Inter-Institutional and Cross-Functional Research Data Management Community: From Strategy to Implementation

Speakers: Anneliese Eber (University of Waterloo), Jennifer Abel (University of Calgary), Michael Steeleworthy (Wilfrid Laurier University)

Location: BMO Amphitheatre

Description:

With the release of the Tri-Agency research data management (RDM) policy in 2021, all Canadian research institutions and hospitals eligible for federal Tri-Agency funding are required to have created and published an institutional RDM strategy.

This session presents key recommendations from a national SSHRC-funded workshop hosted by the University of Waterloo, which synthesized rich discussions into actionable recommendations relevant to funders, government agencies, academic consortia, university administration, researchers, and practitioners.

4:00 PM – Laying the Foundation: The Pilot National Research Data Management Jumpstart

Speaker: Jennifer Abel, University of Calgary

Location: BMO Amphitheatre

Description:

Research data management (RDM) training in Canada has largely been based around 'one-shot' sessions focused on Data Management Plans (DMPs) and data deposits, often missing key concepts of computational reproducibility essential to maintaining data integrity throughout a project's lifecycle.

This session discusses the process, outcomes, and lessons learned from piloting a week-long, 20-hour national research data 'Jumpstart' training program aimed at graduate students and early-career researchers.

4:15 PM – Has Magic? False: File Format Signature Development for Research Data Formats

Speaker: Amanda Tomé, Digital Research Alliance of Canada

Location: BMO Amphitheatre

Description:

Understanding file formats is one of the first steps in determining preservation capabilities and risks to long-term access. However, digital preservation identification tools fail to identify many file formats associated with research data, creating a barrier to an important step in the preservation process.

This session describes ongoing work by the digital preservation team to develop file format signatures to identify file formats found in the Federated Research Data Repository (FRDR).

Day 2, May 28

9:00 AM – Keynote: Health Sciences in the Digital Age

Speaker: Dr. Guillaume Bourque, Professor in the Department of Human Genetics and Director of Bioinformatics at the McGill Genome Center; Board Director, Digital Research Alliance of Canada

Location: BMO Amphitheatre

Description:

High-throughput technologies are transforming biomedical research by enabling the detailed characterization of individual genomes and cellular molecular processes at unprecedented resolution. When paired with advances in machine learning and artificial intelligence, these technologies hold immense promise for driving innovation in health sciences and improving healthcare outcomes. However, realizing this potential requires overcoming significant challenges in data management, software infrastructure, and computational resources. Here, I will present key initiatives addressing some of these challenges. The Global Alliance for Genomics and Health (GA4GH)—a not-for-profit alliance of over 500 organizations—is developing technical standards, policy frameworks, and tools to promote the responsible, voluntary, and secure use of genomic and related health data worldwide. I will also present SecureData4Health, a secure cloud-based infrastructure designed to support the analysis and sharing of genomic and health data. It builds on existing capabilities in Ontario through HPC4Health and adds a new compute node in Québec to enhance national capacity. Finally, I will introduce the Pan-Canadian Genome Library (PCGL), a landmark initiative unifying Canada's human genome sequencing effort. PCGL establishes a centralized, federated data management system aligned with international standards and respectful of jurisdictional and cultural constraints around human genetic data. It supports both retrospective and new projects, ensuring long-term archival, accessibility, and integration of genomic, clinical, and phenotypic data. Together, these efforts lay the foundation for a robust, secure, and collaborative ecosystem that will accelerate genomic research and improve health outcomes across Canada and beyond.

10:00 AM – Humanities and Social Sciences Digital Research Infrastructure in Canada: Current State and Future Directions

Speakers: Alyssa Arbuckle, Research Infrastructure Grants Officer; James MacGregor, Director, Research Infrastructure & Development

Location: BMO Amphitheatre

Description:

In Canada, digital research infrastructure (DRI) has evolved over decades following policy developments, asserted researcher needs, and strategic priorities. This infrastructure, however, serves different disciplines in different ways. DRI in the humanities and social sciences comprises various key areas, including the twin streams

of publishing and libraries and related areas of research data management, preservation, and persistent identifiers (PIDs). In 2024, the Canadian Research Knowledge Network and Érudit embarked on a collaborative project to survey and map digital research infrastructure in Canada, specifically pertaining to these disciplines. The methodology included a scoping review of comparable reports and analyses from different international jurisdictions; research into humanities and social sciences DRI; and a set of conversations with 22 key stakeholders in the Canadian humanities and social sciences DRI landscape. The resulting landscape analysis surveys 11 key DRI organizations connected to the humanities and social sciences (to varying degrees), as well as an additional 23 related initiatives.

Most of those who engaged in conversation consider DRI as the tools, technologies, hardware, software, and people who facilitate digital research. Some respondents emphasize the “invisible but critical” nature of DRI, while others are steadfast in their assertion of how vast DRI is, or the importance of considering DRI throughout the whole research lifecycle, from the conception of an idea to its eventual publication and preservation as a research output. Although this landscape analysis specifically focused on the humanities and social sciences, most stakeholders suggest that the *concept* of DRI is discipline agnostic at its core. In *practice*, however, DRI works differently for different disciplines, who have varying, unique methodological and community needs—acknowledging the overlaps and connections between humanities and social sciences DRI and STEM DRI.

This presentation will focus on areas of confluence in humanities and social sciences DRI that have been influential, with a specific focus on the collaborative approach to integrating PIDs into Canadian DRI through the Canadian PID Advisory Committee (CPIDAC), coordinated by the Alliance with others. We will also outline the challenges that DRI organizations and initiatives in the humanities and social sciences currently face. Overall, this presentation will provide an overview of landscape analysis findings and point to future directions for the Canadian humanities and social sciences DRI community.

10:15 AM – Qualitative Supports and Software for Humanities and Social Sciences (HSS) Researchers

Speakers: Lina Harper, Humanities and Social Sciences Analyst, McGill; Jérémie Dion, Coordinator of the LibreQDA project and PhD candidate in STS at UQAM

Location: BMO Amphitheatre

Description: A high level, introductory overview of qualitative research and needs of HSS researchers. We look at the goals and challenges of qualitative methods, proprietary software, and conclude with a preview of open source software. LibreQDA (in beta) is being developed with the support of Calcul Quebec.

Stream 1: Secure Foundations

11:00 AM – Sensitive Research Computing with SciNet4Health

Speakers: Shawn Winnington-Ball, Manager, Information System Security, SciNet, University of Toronto; Yohai Meiron, Scientific Applications Analyst, SciNet, University of Toronto

Location: BMO Amphitheatre

Description:

We introduce a secure computing enclave at SciNet High-Performance Computing Consortium. Codenamed S4H, this environment is already available to groups at the University of Toronto as a pilot project. S4H aims to meet researchers' needs for hosting and working with sensitive data, which SciNet's main cluster, Niagara, does not accommodate. We'll explain how S4H data are encrypted at rest and access is hardened, talk about the difficulties of providing isolation for different research groups on a shared system, and explore the different components that make it possible, such as key management and containerization mechanisms. We will also focus on our adoption of the Cybersecurity Maturity Model Certification (CMMC) framework, describing our journey deciphering the complexities of the NIST 800-171 control set, crafting Plans of Action and Milestones for compliance gaps, and internal and external assessments to verify compliance.

11:15 AM – Secure Data for Health (SD4H): A Secure Digital Research Cloud

Speakers: Nathalie Aerens, SD4Health Platform Manager, McGill; Pierre-Olivier Quirion, SD4H Tech Lead, McGill - CQ - SD4H

Location: BMO Amphitheatre

Description:

The future of Canada's digital research in health and genomics is relying on robust privacy and security safeguards. Data security has always been a concern for human genomics research, and since its progressive but intensive usage of the Canadian DRI infrastructure, different levels of pressure have been put on us, Alliance members, to make our infrastructure secure.

Achieving robust security for human health projects is no longer a distant goal. New requirements, like the recent update to the NIH Genomic Data Sharing Policy, have made it a priority for research groups previously less concerned with this aspect. The old approach of neglecting security demands is not viable for many new research projects.

SD4H is based on an OpenStack cloud infrastructure managed jointly by McGill's Canadian Center for Computational Genomics (C3G) and Calcul Québec. Advanced Research Computing, Research Data Management, and Research Software services are developed and deployed on this infrastructure, providing privacy and security for its users.

11:30 AM – Federated Identities for Scientific Platforms Using CILogon

Speaker: Darren Boss, Senior Research Computing Analyst, University of Victoria

Location: BMO Amphitheatre

Description: CILogon is used to simplify access to platforms via the EduGain identity federation. It enables easy user authentication and sharing of enriched identity data across multiple applications. CILogon is already being used in platforms like DMP Assistant, and the Cloud Connect pilot with more platform integrations in the planning phase.

11:45 AM – API Security

Speaker: Ryan McDonald, Cybersecurity Manager, UVic / Arbutus Site

Location: BMO Amphitheatre

Description: APIs are the backbone of modern applications enabling integration and communication between systems. However, with great connectivity comes great responsibility. This technical talk will delve into API security, exploring the risks exposed via APIs, the best practices for securing APIs, and how to implement API security measures based on work to secure our community cloud APIs. Join us to learn how to fortify your APIs against common threats.

12:00 PM – Controlled Access Management (CAM) for Research Data Initiative

Speaker: Victoria Smith, Data Governance & Ethics Lead, Digital Research Alliance of Canada

Location: BMO Amphitheatre

Description: This presentation will provide a brief overview of the Controlled Access Management (CAM) for Research Data Initiative, including outputs from Phase 1, and current and planned work. Applying the principles of open science in the context of sensitive and restricted access research involve challenges on different levels.

Stream 2: Innovating Systems

11:00 AM – MonarQ Taking Flight

Speaker: Lydia Vermeyden, Director for the Development of New Research Services, Calcul Québec

Location: MB 3.435

Description: 2025 is the international year for quantum science and technology, marking 100 years since the first official publication of quantum mechanics. At Calcul Québec we are celebrating with the launch of our quantum services on MonarQ, our 24 qubit quantum computer. But what is quantum computing, and is it really that different

from managing the classical systems, software and data that we know and love? What happens if classical and quantum systems are connected together? We will be discussing these questions while sharing our extraordinary journey of launching quantum computing services at Calcul Québec, and the amazing team that has made it possible. After all, successfully facilitating the research of tomorrow is not just about the latest technology, it's about the people that make it happen.

11:15 AM – PennyLane-Calcul Québec: Streamlining Quantum Algorithm Research on MonarQ

Speakers: Samuel Richard & Felix-Gabriel Boucher-Luneau, Calcul Québec

Location: MB 3.435

Description: Come hear the details of the Arbutus Cloud renewal, the exciting capabilities it's bringing forward, and an exploration of how the rapidly-changing global landscape has put a sudden spotlight on technology autonomy at every level.

11:30 AM – Streamlining Magic Castle Clusters Creation at Calcul Québec Using Terraform Cloud

Speaker: Maxime Boissonneault, Director, Research and Teaching Support, Calcul Québec

Location: MB 3.435

Description: Calcul Québec manages about a hundred Magic Castle clusters on a yearly basis, mostly over the fall and winter seasons. Most of these are for our own training workshops, but some of these are also to support graduate academic classes.

Handling this many clusters with different configurations is a lot of work. To allow Calcul Québec team to do this work, we have streamlined the creation of these clusters. This is done using a few tricks which I will demonstrate in this talk.

The first trick is using a common Git repository to host all cluster configurations. We have abstracted much of the terraform configuration, to hide the complexity and let our team specify only useful parameters, avoiding most mistakes.

The second trick is using Terraform cloud to hold configuration parameters, run terraform, and host the terraform state. This allows to share the load between multiple people within our teams.

During this talk, I will first ask the audience to choose some parameters of the cluster, which I will then launch. While the cluster is being provisioned automatically, I will explain what is happening behind the scene.

11:45 AM – A Home Among the Clouds

Speaker: Jeff Albert, Manager and Architect, Advanced Research Computing Infrastructure, University of Victoria

Location: MB 3.435

Description:

12:00 PM – State of the Accelerators

Speaker: Pawel Pomorski, HPC Analyst, Lead of Accelerators Working Group, SHARCNET

Location: MB 3.435

Description: The Alliance has just experienced a generational jump in the capabilities of its accelerator hardware, with hundreds of new NVIDIA H100 GPUs installed across new Alliance clusters. This session will discuss the impact of these changes, and future prospects for accelerators, as they apply to new hardware purchases that we might consider in the coming years.

We will talk about new features in H100 GPUs and the challenges of utilizing them efficiently. We will also cover new GPUs coming from NVIDIA in the next few years.

Another exciting hardware development are AMD's MI300A APUs (Accelerated Processing Units) that combine AMD CPU and GPU cores in a single chip, allowing both to share a fully unified memory space. Early experiences with the MI300A APUs installed on the Nibi cluster will be discussed.

We will move on to discuss upcoming developments in accelerators in general, focusing on those dedicated to deep learning AI calculations that go beyond the current paradigm of using modified GPUs (Graphics Processing Units) towards hardware designed specifically for AI calculations. Among various new developments to mention are wafer-scale processors that build very large integrated circuit networks. Another option are FPGAs (Field Programmable Gate Arrays) which offer a low power alternative for some AI applications. Finally, we will discuss new developments in neuromorphic computing, where hardware is inspired by the structure and function of the human brain, and thus uses physical artificial neurons to perform computations.

Stream 3: Empowering Research

11:00 AM – UseGalaxy Canada Platform Update

Speaker: Carol Gauthier, Project Manager/Developer, Université de Sherbrooke/Calcul Québec

Location: MB 3.445

Description:

Galaxy (galaxyproject.org) is a versatile, open-source platform designed to democratize data-intensive research through its intuitive graphical interface. By removing the barriers of programming and systems administration, Galaxy empowers researchers to conduct

complex computational analyses.

Originally focused on genomics, its domain-agnostic nature has established it as a leading workflow management system in bioinformatics.

The international UseGalaxy initiative provides globally accessible, pre-configured Galaxy environments through a network of servers, including UseGalaxy.org, UseGalaxy.eu, and UseGalaxy.org.au. Building upon this foundation, the UseGalaxy.ca team has successfully launched the Canadian branch in 2024.

Hosted on the Beluga Cloud, thanks to an Alliance RPP, UseGalaxy.ca features a robust infrastructure comprising 500 virtual CPUs and 2.3 TB of RAM. The platform has experienced consistent growth, now exceeding 500 users. Plans are underway to officially integrate UseGalaxy.ca into the international UseGalaxy consortium within the next year, further solidifying its role in the global research community.

Furthermore, UseGalaxy.ca has partnered with developers of UseGalaxy.org from Johns Hopkins University and of UseGalaxy.org.au from Melbourne University to implement the application service within the Alliance Cloud Connect Project (ACCP). This presentation will provide a concise technical overview of UseGalaxy.ca's components, detail its integration with the ACCP, present an update on development, implementation, and usage statistics, and outline the platform's future roadmap.

11:15 AM – CBRAIN: A Web-Based Distributed Computing Platform for Collaborative Research

Speaker: Bryan Caron, Director, Operations and Development (CBRAIN and NeuroHub) and Co-PI of NeuroHub, McGill University

Location: MB 3.445

Description:

CBRAIN (<https://cbrain.ca>) is an open source, web-based, collaborative research software platform designed to address major challenges in big data research. CBRAIN allows scientists to launch large-scale big data analyses using advanced scientific tools through an easy to use web-based user interface, removing the steep learning curve and pitfalls associated with a complex command-line environment. It facilitates data-sharing and provenance tracking of results obtained from large-scale data processing through computationally intensive scientific pipelines using high-performance computing (HPC) clusters and cloud resources. CBRAIN is used in numerous major neuroinformatics projects and is a pillar component of NeuroHub, a core platform of McGill University's Healthy Brains, Healthy Lives initiative.

In production since 2009, CBRAIN is a mature and flexible orchestration infrastructure seamlessly connecting scientists to a network of advanced computing and data resources.

CBRAIN allows users to easily launch and manage the execution of complex computational pipelines and choreographs all associated data transfers across a distributed network of resources. CBRAIN is integrated with OpenID Connect compliant systems for identity management, authenticating users with their institutional Identity

Provider and authorizing access to CBRAIN services. With CBRAIN users are able to upload their own data, create datasets and share that data. Fine-grained access controls enable users to securely store and share data according to a study's data management policy. A wide variety of open and controlled access datasets including those from the Canadian Open Neuroscience Platform, the Human Connectome Project, OpenNeuro and the UK Biobank are available, avoiding unnecessary data duplication. CBRAIN supports a wide range of HPC and cloud-based data storage types, with interoperability extending to multi-modal data management systems such as LORIS, through which users can create cohorts of interest for analysis in CBRAIN.

Using Apptainer containers and the Boutiques descriptive tool framework, CBRAIN makes available over 160 pre-configured analysis pipelines for neuroimaging and genetics. Any tool that can be containerized can be integrated within CBRAIN. CBRAIN allows users to select from a range of shared or dedicated lab / HPC / cloud resources. CBRAIN is able to save results of the pipeline back to a dataspace on any selected shared or authorized storage location. With the CBRAIN API users can programmatically perform all functions of the web UI. As such CBRAIN supports interactive data exploration and analysis through Python and R Jupyter notebooks allowing users to retrieve data from CBRAIN, analyze and visualize it, and then store the results back into CBRAIN.

CBRAIN now has over 2000 users across 193 locations in 59 countries, providing researchers in neuroinformatics, genomics and other areas the ability to easily collaborate, manage and analyze data at scale, masking the complexities of working with HPC and cloud resources. While CBRAIN is primarily used for applications in brain science, it is a completely general platform that could be used in other settings.

11:30 AM – OSSECMon: High-Performance Linux Security Monitoring with eBPF

Speaker: Gabriel Lapointe, Security Analyst, McGill University

Location: MB 3.445

Description: Premièrement, nous allons expliquer le contexte dans lequel OSSECMon a été créé (Nous avons remarqué des manquements à certains niveaux concernant, entre autres des principes de base de sécurité, ex; la non répudiation et cela nous a mené à développer une solution qui n'existait pas encore). Par la suite, nous ferons un bref survol de la technologie principale qui est utilisée par le programme (*eBPF*) et expliquerons comment nous l'utilisons dans ossecmon suivi d'un bref survol sur des outils complémentaires utilisés comme *inotify*, *yara* et *openscap*. Le tout sera bouclé en expliquant comment OSSECMon peut être utilisé afin de se conformer à certaines normes comme le NIST, ISO, etc. Le temps va probablement nous manquer mais le tout sera clôturé par une période de questions.

11:45 AM – Interactive HPC Computing with Open OnDemand

Speakers: Grigory Shamov, Site Lead/HPC Specialist, University of Manitoba; James Willis, Scientific Applications Analyst, SciNet

Location: MB 3.445

Description: In this talk, we will introduce Open OnDemand, a web-based interface designed to provide easy access to High-Performance Computing (HPC) resources. Terminal-based interfaces can be challenging for new users with limited experience, resulting in a steep learning curve.

Open OnDemand aims to make HPC more accessible by offering an intuitive graphical interface that simplifies the process of submitting, monitoring and managing jobs. We will explore the key features of Open OnDemand, including web-based access, job management, file management and support for interactive applications like Jupyter Notebooks, RStudio, and VS Code. Additionally, we will demonstrate the SciNet and Grex Open OnDemand portal and discuss their deployment and use cases.

12:00 PM – From Bash Scripts to GitOps: Automating Kubernetes Deployments with ArgoCD

Speaker: Shaun Bathgate, Senior Advanced Research Computing Systems Administrator, University of Victoria

Location: MB 3.445

Description:

Research teams commonly manage Kubernetes deployments using imperative approaches such as bash scripts or GitHub Actions. Although effective initially, these methods quickly become complex, difficult to maintain, and error-prone as workloads scale or spread across multiple clusters. This brief, focused talk introduces GitOps and ArgoCD as powerful alternatives to imperative methods, emphasizing how they simplify and automate Kubernetes management.

The session begins by clearly defining GitOps, emphasizing its core practice of managing infrastructure declaratively through version-controlled repositories. Participants will see how this approach significantly reduces manual interventions, increases transparency, and enables easy rollbacks to stable states.

Through concise, practical examples, the talk demonstrates the advantages of ArgoCD's automated synchronization. A side-by-side comparison with traditional scripting methods clearly illustrates how GitOps simplifies and stabilizes daily operational tasks. The audience will observe firsthand the simplicity and clarity of declarative Kubernetes manifests and continuous reconciliation.

Additionally, the talk briefly touches upon secure secret management using tools such as SealedSecrets, highlighting secure best practices within the GitOps model without diving deeply into technical configuration specifics.

By condensing core GitOps concepts into a compact demonstration, the session provides attendees with a clear, actionable introduction to GitOps and ArgoCD. Participants will leave understanding how adopting GitOps principles can significantly

streamline Kubernetes operations, enhance collaboration, and reduce deployment complexity in research environments.