













The arctic knowledge-based system: Science gateway integration for petascale arctic data processing and geospatial feature prediction

Andrew Wilcox^a ^{*}, Meisam Shayeghmoradi^b , Stephen Miller^c , Ian Nesbitt^d ,
Saisri Pogalla^c , Aymane Ahajjam^e , Walker McKee^c , Sheridan Parker^f ,
Matthew Johnson^c , Aaron Bergstrom^c , Naima Kaabouch^b , Timothy Pasch^f 

^a Department of Earth System Science & Policy, University of North Dakota, Grand Forks, 58201, ND, USA

^b Artificial Intelligence Research Center, University of North Dakota, Grand Forks, 58201, ND, USA

^c Computational Research Center, University of North Dakota, Grand Forks, 58201, ND, USA

^d National Center for Ecological Analysis and Synthesis, University of California Santa Barbara, Santa Barbara, 93101-5509, CA, USA

^e School of Electrical Engineering and Computer Science, University of North Dakota, Grand Forks, 58201, ND, USA

^f Department of Communication, University of North Dakota, Grand Forks, 58201, ND, USA

ARTICLE INFO

Keywords:

Science gateway
Remote computing
Knowledge-based systems
Permafrost analytics
Machine learning

ABSTRACT

Science gateways have become essential platforms that integrate computational resources, data services, and workflows for domain researchers, enabling artificial intelligence-driven (AI) analyses at scale. Building on this paradigm, we introduce the Science Gateway component of the Arctic Knowledge-Based System (A-KBS), designed to advance AI-assisted modeling of permafrost dynamics and other Arctic geospatial processes. The A-KBS provides researchers with a unified portal to configure and execute multi-horizon prediction tools for active layer thickness, ground deformation, wildfire occurrence, freeze/thaw states, soil and air temperature analyses, and to run global scale geospatial HPC workflows leveraging data from across the circumpolar Arctic. This system orchestrates workloads through Kubernetes-based (K8s) containerized environments, Globus Data Transfer/Compute services, and distributed computing tools such as Slurm, ParSL, and Ray.io. Its web portal is deployed on the University of North Dakota's (UND) virtualized, load-balanced K8s cluster with cloud migration enabled by Rancher, while Python-based AI functions authored in JupyterHub are executed on remote systems through Globus Compute Endpoints. Current development has integrated the A-KBS with the UND high-performance computing (HPC) Talon cluster and Amazon Web Services-managed K8s resources, with a roadmap in place to extend this integration to other HPC environments including the San Diego Supercomputer Center's Expanse System. By coupling scalable infrastructure with a suite of existing AI-driven workflows for environmental prediction tasks, the A-KBS accelerates Arctic science, strengthens cryospheric research, and supports decision-making through its integration with cyberinfrastructure surrounding the DRP (Defense Resiliency Platform Against Extreme Cold Weather) initiative.

1. Introduction

Environmental transformation driven by the loss of permafrost is occurring across landscapes in Arctic and sub-Arctic regions (Marchand et al., 2018; Park et al., 2016). As these perennially frozen soils begin to thaw, the process can lead to slope collapses, differential ground deformation, and the rapid formation of thermokarst lakes, posing significant challenges for critical transportation and energy infrastructure in the region (Hjort et al., 2022; Streletskiy et al., 2023). To address these growing challenges, scientists and decision-makers require powerful tools to accurately model and predict these environmental shifts and their potential impacts.

A science gateway is a highly specialized iteration of web portal with the unique capability of providing teams/communities of researchers with access to powerful, highly tailored tools, datasets, workflows, and computational resources (Wilkins-Diehr, 2007; Lynch et al., 2010). These gateways transcend conventional access barriers by enabling intuitive, location-independent interaction with high-performance computing (HPC) infrastructures and advanced analytics. Their platform-agnostic architecture ensures broad accessibility, allowing engagement irrespective of a user's technical expertise or geographic location (National Science Foundation, n.d.a; Lynch et al., 2010; Gesing et al., 2024). The National Science Foundation's (NSF)

* Corresponding author.

E-mail address: andrew.wilcox@ndus.edu (A. Wilcox).

<https://doi.org/10.1016/j.acags.2026.100322>

Received 16 October 2025; Received in revised form 5 January 2026; Accepted 15 January 2026

Available online 4 February 2026

2590-1974/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

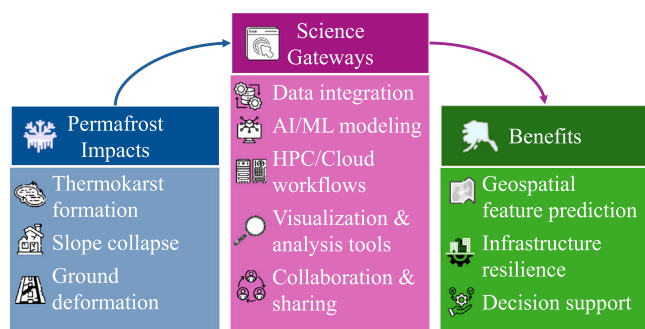


Fig. 1. Science gateways provide Arctic researchers with the tools to respond to permafrost-related challenges. By integrating data, modeling, and cloud workflows, these platforms deliver results that inform decision-makers.

Science Gateway Catalog lists numerous portals spanning diverse scientific disciplines, a testament to their growing importance in scientific research (Wilkins-Diehr, 2011; Science Gateway Community Institute, n.d.).

To support decision-making tools leveraged towards the advancement of Arctic sciences, we have developed the Arctic Knowledge-Based System (A-KBS), a high-performance computational environment that orchestrates distributed workflows through Parsl and Globus Compute, linking the gateway to multiple supercomputing systems through multi-user Globus endpoints. As a component of this development we have implemented a science gateway facilitating research related to the artificial intelligence-assisted (AI) modeling components of permafrost dynamics and other Arctic geospatial processes. Fig. 1 provides an overview of how science gateways mediate between permafrost-related challenges and actionable outcomes. The A-KBS is a one-stop computational platform, allowing domain researchers to configure and use AI-assisted multi-horizon prediction tools for a variety of critical variables in Alaska and beyond. This includes tasks such as forecasting active layer thickness, ground deformation, wildfire occurrence, and soil and air temperatures.

Pan-Arctic research spanning approximately 21 million km² across the circumpolar permafrost region (Obu, 2021) presents computational challenges that exceed the capacity of conventional desktop or single-server workflows. The A-KBS gateway was designed to process high-resolution satellite imagery, synthetic aperture radar (SAR) data, digital elevation models, in-situ observations, and high-resolution LiDAR point clouds of the Dalton and Steese Highways gathered by the researchers and their collaborators in Alaska. By integrating with national HPC resources such as SDSC's Expanse system, the platform enables researchers to scale analyses to match the scope of Arctic environmental transformation without requiring expertise in HPC system administration. This scalable architecture also supports the operational demands of the Defense Resiliency Platform Against Extreme Cold Weather (DRP) initiative (<https://drp.dataone.org/>), where decision-makers in Alaskan field conditions require model outputs within minutes for mission planning and other time-sensitive operations. Furthermore, as the vast amounts of remote sensing data (Liljedahl et al., 2019) in existing Arctic databases continue to expand with input from recent satellite missions such as NISAR (Kellogg et al., 2020), a petascale-ready gateway ensures the A-KBS remains viable infrastructure for future Arctic research.

The purpose of this work is to report on the creation mechanisms and development processes for the A-KBS Science Gateway. This research has the potential to benefit those working within the field of cryospheric sciences and assist others in creating their own science gateways for various scientific domains.

This paper is organized as follows. Section 2 reviews science gateways within Arctic and cryospheric sciences, highlighting existing platforms and infrastructure relevant to the development of the A-KBS.

Section 3 describes the architecture of the A-KBS. Section 4 presents the deployment and validation of the system. Section 5 examines current challenges and outlines potential directions for future development. Finally, Section 6 summarizes the paper's contributions and discusses the broader implications of science gateways for Arctic research.

2. Science gateways and data portals in arctic research

Science gateways, such as the A-KBS gateway, have emerged as vital cyberinfrastructure components that bridge the gap between complex computational tools and domain-specific Arctic-focused research scientists and teams (Kalyanam et al., 2019). Across research communities, shared information and collaborative platforms enable advances in accessibility and synergistic efforts (Becciani et al., 2015), leading to greater scientific understanding of complex environmental processes including transformations within Arctic permafrost and other cryospheric-related geospatial research and analytics.

As an example of a related science gateway, the Permafrost Discovery Gateway (PDG) is an NSF-funded Arctic earth science gateway launched in 2019 by the University of Alaska Fairbanks and partners to provide cyberinfrastructure for the exploration of permafrost dynamics (Liljedahl et al., 2019, 2021). Tools within the PDG include Clowder for data management, web-based visualization tools, and access to computational resources for image analysis, leveraging satellite imagery from the Polar Geospatial Center, Landsat, and Sentinel. Researchers have used the portal on projects such as hydrological mapping in permafrost landscapes (Liljedahl et al., 2024). Although new AI-driven tools for risk analysis are currently in development, the PDG does not yet have predictive tools built into its infrastructure (Tempesta, 2025).

Similarly, the University at Buffalo's Ghub Project is an open-access gateway that accelerates ice sheet and glacier research (Tulenko et al., 2025). Launched in 2020 and supported by the National Science Foundation (NSF) and EarthCube, Ghub is powered by the HubZero platform (<https://hubzero.org>). While the A-KBS gateway will examine permafrost processes on land, tools within Ghub focus on ice sheet observation, projection modeling, and remote sensing. Research involving Ghub has included hosting derived datasets and tools, such as a resampled elevation, slope, and roughness dataset of the Greenland Ice Sheet and crevasse detection tools (Tulenko et al., 2025; Sperhac et al., 2021; Mann-Bruch, 2022). However, a core limitation of Ghub is that it functions as a repository for data and user-created tools, rather than a unified environment for direct computational work (Sperhac et al., 2021).

Several data-focused portals also play key roles in Arctic research. The Polar Geospatial Center (PGC) at the University of Minnesota provides extensive polar imagery and data, including 8 PB of sub-meter imagery and DEMs. It offers a web-based platform, FRIDGE, for dataset discovery and automated processing workflows on national cyberinfrastructure available through HPC platforms like XSEDE (now known as ACCESS-CI) and systems such as Blue Waters (Dickson et al., 2024; Pundsack et al., 2016; Morin, 2016). While the PGC offers a robust platform for data discovery and access, it primarily serves as a data portal, lacking a unified computational environment that integrates predictive modeling (of Minnesota, 2025). The Circumarctic Environmental Observatories Network (CEON) strengthens long-term observations by connecting existing research stations and networks, making diverse environmental data available through its interactive mapping platform (Tweedie and Webber, 2003; Webber, 2004). The Barrow Area Information Database (BAID) hosts over 9600 research site records for northern Alaska and provides interactive web-mapping tools for data exploration and download (Cody et al., 2012). Finally, the NSF's Arctic Data Center (ADC) serves as the primary repository for the NSF's Office of Polar Programs, preserving diverse datasets to support reproducible science and data use (National Science Foundation, n.d.b; National Science Foundation, n.d.c; Jones et al., 2021). However, the ADC is not designed as a computational environment, rather it hosts



Fig. 2. A broad overview of the A-KBS, showing its function as a science gateway connecting the DRP to external supercomputing resources for AI-assisted modeling and predictive analytics.

data and provides access to datasets via DOIs. While these platforms excel at data preservation and discovery, many of them do not provide direct access to HPC resources for data processing and predictive analytics.

The platforms reviewed above have advanced Arctic research considerably, yet each addresses only a subset of the capabilities required for end-to-end AI-driven geospatial analysis. Rather than functioning as a repository or a collection of standalone tools, the A-KBS gateway provides an integrated environment where researchers can access and process distributed Arctic datasets, configure prediction workflows, and dispatch computations to HPC systems without managing the underlying infrastructure. This design responds directly to the demands of geoscience for the circumpolar Arctic, where the expansion of permafrost data (Liljedahl et al., 2019), varying feature resolutions, internationally distributed repositories, and the computational intensity of machine and deep learning geospatial prediction necessitate tightly integrated cyberinfrastructure.

3. System design and architecture

The A-KBS gateway integrates interactive user tools, automated workflow orchestration, distributed AI training, and secure data transfer via Globus into a unified, web-accessible platform, connecting the A-KBS/DRP to HPC resources. A broad overview of the A-KBS can be seen in Fig. 2, and the system's architecture and components are discussed in detail below.

3.1. Core enabling technologies

The A-KBS leverages multiple modern cyberinfrastructure technologies to bridge the gaps identified in previous work and provide a robust computational platform for Arctic research.

- **Globus:** Globus services simplify the movement of scientific data across diverse storage systems, making them well-suited for integration into science gateway architectures (Chard et al., 2016). Globus Connect enables high-performance data transfers, while Globus Compute allows users to run Python functions on remote systems using Globus endpoints. This functionality removes many barriers to HPC access, such as complex credential management and user interaction with schedulers (Chard et al., 2016; Bauer et al., 2024; Ananthakrishnan et al., 2024). Within the A-KBS gateway, Globus was chosen to support the transfer of

large volumes of data needed for circumpolar research. Furthermore, by automating authentication and job orchestration between the K8s-based gateway and remote HPC resources, Globus supports time-sensitive Arctic mission planning. This automation is critical in field scenarios where connectivity and time constraints preclude manual management of credentials or scheduler interactions.

- **Kubernetes:** Kubernetes (K8s) has emerged as a powerful tool for deploying scalable, fault-tolerant architectures for science gateways. Its use in orchestrating JupyterHub environments on HPC resources has been shown to serve hundreds of users simultaneously (Zonca and Sinkovits, 2018). Due to the high resource demands of AI/ML applications, K8s's container orchestration and microservice management capabilities complement HPC workload managers like SLURM and Slinky, making it valuable in hybrid HPC-Cloud environments (Zhou et al., 2021; Decker et al., 2025). This architectural choice is particularly important for the A-KBS gateway as Arctic research workflows are often campaign-driven, with computational demands spiking sharply during seasonal data releases or field-based modeling efforts. K8s provides the necessary elasticity to scale resources in response to these surges while ensuring optimal resource utilization during lower-demand periods.
- **JupyterHub:** The JupyterHub integration within the A-KBS gateway provides a secure, multi-user platform for scientific data analysis and computation (Zonca and Sinkovits, 2018), accessible through the user frontend dashboard. It allows authenticated users to launch and manage notebook instances, where they can author, edit, and execute code locally or submit jobs to high-performance resources through a configured Globus Compute Endpoint. The AI-based Arctic decision support tools accessible through the A-KBS portal (e.g., active layer thickness, wildfire risk, freeze/thaw state classification, and air and soil temperatures) were developed and validated in prior work by the authors and their collaborators (Ahajjam et al., 2025c,a,b, 2024; Chance et al., 2024). While the specialized decision support interface for these tools is the subject of a forthcoming publication, they are hosted on the K8s cluster and accessible through the A-KBS portal, with deeper integration currently under development. JupyterHub is integrated as the primary interactive development environment. Here, researchers can leverage the underlying AI models as modular building blocks to modify existing algorithms, process large-scale geospatial datasets in parallel, or author new specialized Arctic workflows that can be dispatched to remote HPC resources via Globus.

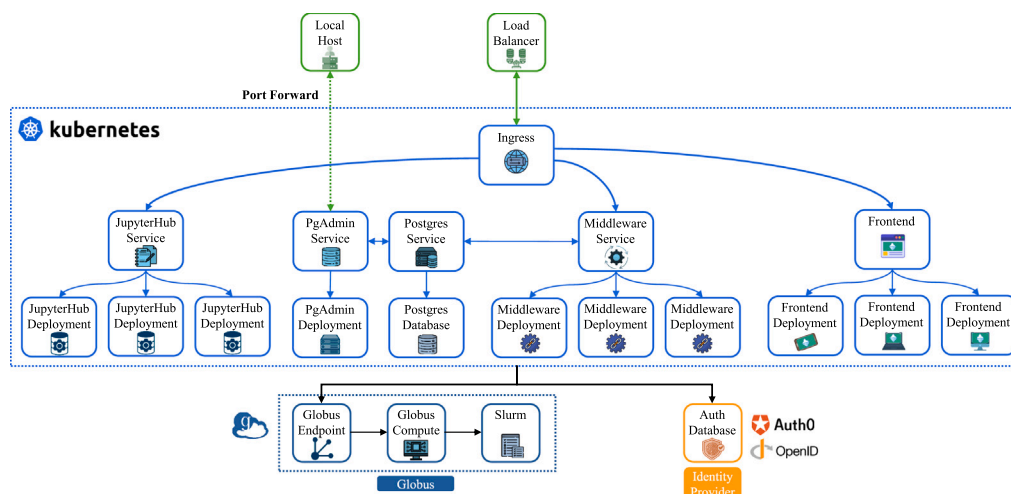


Fig. 3. A detailed overview of the A-KBS's 3TA architecture. The figure highlights key components, including the frontend, middleware API, and a K8s-managed backend that integrates with the IdP for secure access, while using Globus for scalable data and workflow management with external resources.

3.2. System architecture and components

The A-KBS is designed for deployment on a K8s-based three-tier (3TA) architecture to enable scalable, AI-enhanced modeling of Arctic geospatial processes. Its headless design separates backend logic and data services from presentation layers, with core domain logic encapsulated in a middleware API (Salesforce Research, 2022). This promotes scalability and flexibility for integrating with a variety of interfaces, including the DRP. An overview of the A-KBS architecture is shown in Fig. 3.

The gateway's 3TA enables each component to run as an independent deployment to support replication and load balancing. External traffic enters through a load balancer which then distributes requests across the K8s cluster, an ingress controller then routes incoming HTTP/HTTPS traffic to appropriate services according to pre-defined rules. Core logic is managed by the headless middleware API that coordinates interactions among all system components within the A-KBS.

3.2.1. A-KBS frontend and middleware

The science gateway frontend, built with the Next.js React framework, provides a unified web-based interface for users, as seen in Fig. 4. From the dashboard, authenticated users of the A-KBS can launch and monitor JupyterHub notebook sessions, execute code, visualize and analyze data, and submit jobs to Globus Compute endpoints via the middleware API. This single environment streamlines workflow execution and data access for science gateway users. Fig. 5 shows the gateway's dashboard and additional resources accessible through the frontend interface.

The middleware, built with the NestJS framework, serves as the headless application façade. It exposes both GraphQL and REST APIs to both the frontend and approved external services, managing requests for Jupyter Notebook access, executing CRUD operations on the PostgreSQL database, and facilitating large-scale data exchanges with Globus. This decoupled design is critical, as it allows multiple clients, including web, mobile, and external systems, to interact with the AKB's backend services without being dependent on a specific frontend.

3.2.2. PostgreSQL database and pgAdmin

The PostgreSQL database serves as the platform's primary relational data store, providing ACID-compliant storage for all application services. PostgreSQL was chosen due to its reliability, PostGIS support (<https://postgis.net/>), full-text search, and support for geospatial data. All interactions with the database occur securely through the

middleware's APIs, ensuring consistent enforcement of access controls and data consistency within the A-KBS. To protect data integrity, PostgreSQL allows for automated backup schedules with tested restore procedures to ensure availability and disaster recovery.

For database administration, pgAdmin provides a secure, web-based interface for monitoring the PostgreSQL database used by the A-KBS. Accessible to authorized administrators, pgAdmin supports schema management, query execution, and performance monitoring. The pgAdmin webservice also enables verification, scheduling, and restoring of database backups through the interface, ensuring the A-KBS's long-term reliability.

3.2.3. Globus integration and JupyterHub

The JupyterHub integration within the A-KBS provides a secure, multi-user environment for data analysis and computation using Python-based notebooks, allowing users to connect to external compute resources. Using K8s-backed spawners, the system ensures secure multi-user isolation and persistent sessions within JupyterHub, allowing users to resume previous work. This integration facilitates an efficient workflow with data transfer through Globus endpoints that quickly moves input/output to the notebook interface.

The A-KBS uses Globus services to create a secure, scalable framework for offloading computational tasks and managing data transfers between the platform and external resources, which then executes the tasks using frameworks including Parsl and Ray.io. It leverages Globus Compute to enable custom JupyterHub notebook kernels to offload code execution to designated Globus Endpoints, such as remote HPC clusters and supercomputing centers. The system reduces the complexity of a backend workload manager by using Slurm to automatically schedule jobs submitted through Globus Compute. All user interactions, including data transfers and monitoring job progress, occurs directly from the A-KBS frontend, with Globus handling complex processes such as user authentication, fault recovery, and performance optimization in the background. Globus also provides Globus Search capabilities (University of Chicago, 2026), which allows for the indexing and querying of data based on geospatial metadata, streamlining the discovery of relevant research assets.

3.2.4. Identity provider and security

The Identity Provider (IdP) within the A-KBS delivers centralized authentication and authorization for access across system components. Using OAuth 2.0, OpenID Connect, and Security Assertion Markup Language (SAML) protocols, the IdP enables single sign-on (SSO) through the A-KBS frontend. Authenticated users can then access JupyterHub,

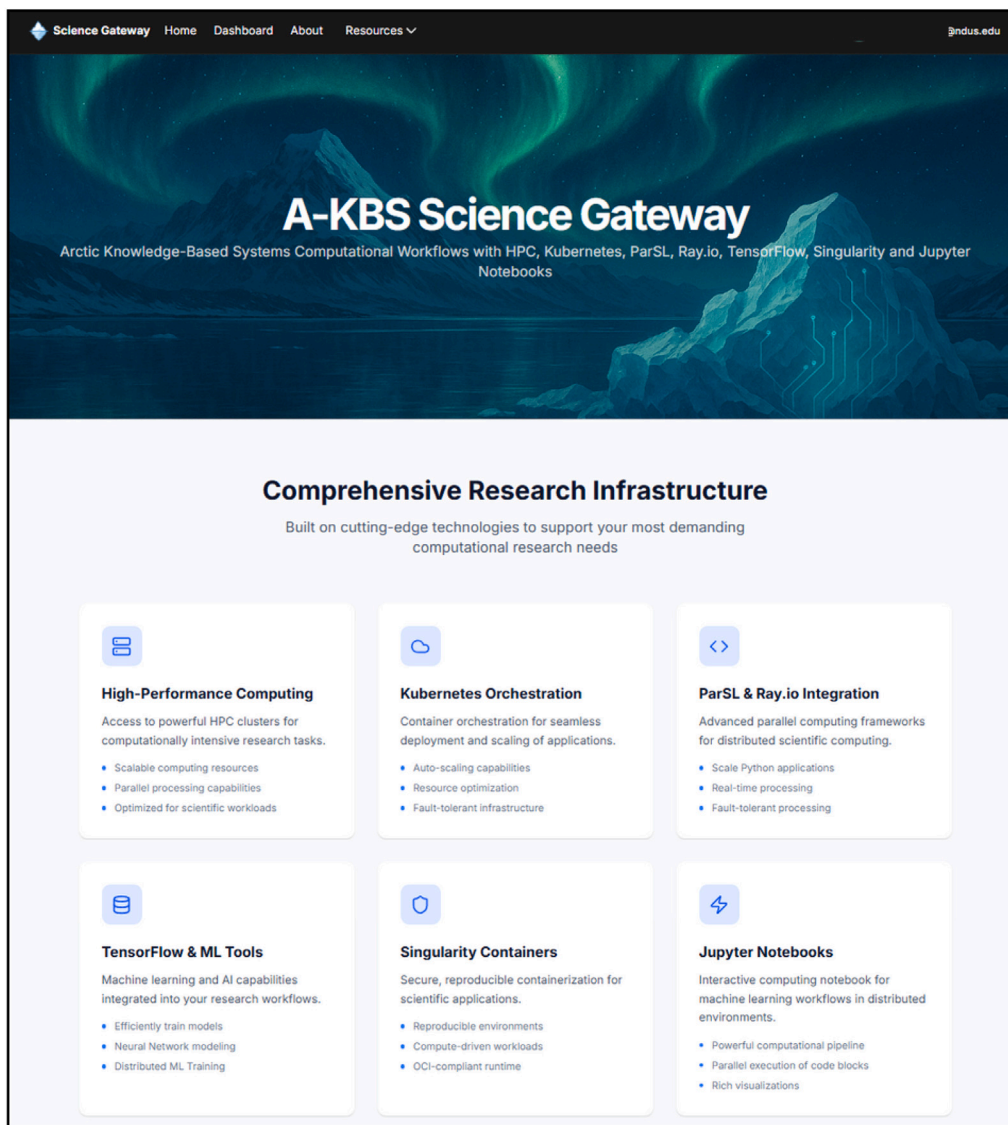


Fig. 4. The A-KBS web portal landing page, illustrating the user dashboard from which researchers can access JupyterHub, different data, and submit jobs to external computing resources.

Globus workflows, and other services. The system issues and validates JSON Web Tokens (JWTs), with Role-Based Access Control (RBAC) to restrict or allow access to specific system features based on user roles or group membership.

3.2.5. Deployment, scalability, and observability

The A-KBS system is deployed on a Rancher-managed K8s cluster, providing a scalable and centralized platform for managing all system components. Key services such as the middleware, science gateway frontend, and JupyterHub are packaged as K8s deployments (Madupati, 2021), with multiple replicas to support load balancing and high availability. The K8s deployments are managed using Helm charts. The architecture supports horizontal scalability, allowing each core service to scale independently by adjusting the number of replicas in its K8s deployment. User traffic is managed by the Ingress controller, which handles Transport Layer Security (TLS) and Secure Sockets Layer (SSL) termination in addition to routing, while an external load balancer distributes incoming traffic across the cluster and provides Domain Name System (DNS) resolution for public access to system services.

As described in previous sections, security is integral for the deploy-

ment. SSL/TLS is enforced at the ingress level in addition to JWT-based authentication used validation across services within the A-KBS. Internal services that do not require public access are not exposed outside the cluster and utilize a ClusterIP service configuration, with K8s network policies enforcing restricted inter-service communication and component isolation. The A-KBS system uses K8s-native monitoring and logging tools for observability, performance insight, and operational transparency. Metrics related to application health, resource usage, and cluster performance are collected using Prometheus and visualized with Grafana dashboards. For logging, a centralized pipeline is implemented with the ELK stack (Elasticsearch, Logstash, Kibana) to ensure log ingestion, storage, and analysis.

4. Results: A-KBS system implementation

This section presents the outcomes of A-KBS development and initial deployment, detailing the core capabilities, scalability, and reliability of the platform. The results confirm that the gateway's design effectively addresses the complexities of scientific workflows and provides a robust environment for Arctic sciences.

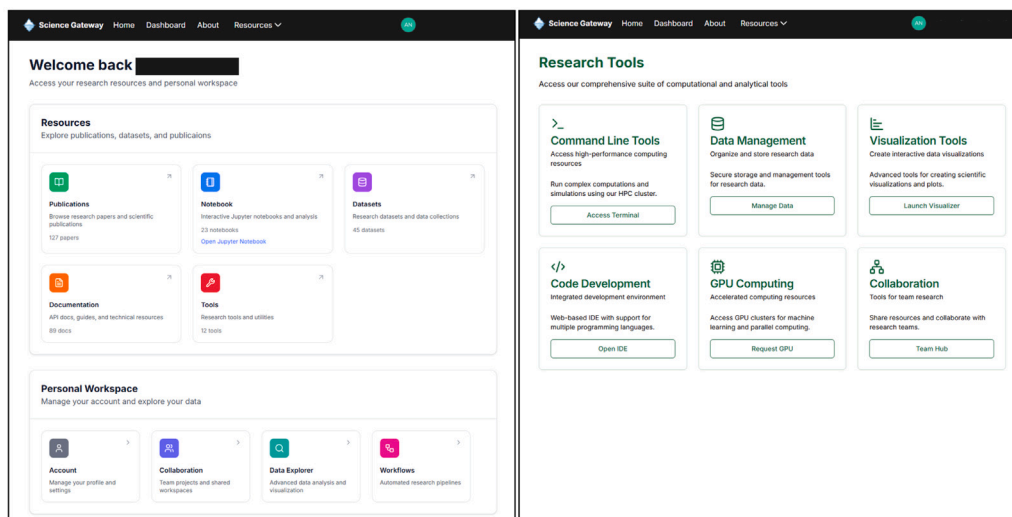


Fig. 5. A-KBS Science Gateway user interface. The dashboard (left) provides authenticated users with access to research resources including publications, datasets, and documentation. The Research Tools page (right) enables users to launch JupyterHub notebooks, manage data, access HPC and GPU resources, develop code, and visualize results within an integrated environment.

4.1. System capabilities and operational validation

The deployment and initial operation of the A-KBS demonstrated the platform's core capabilities and operational readiness. The K8s-based 3TA proved to be a capable framework for integrating disparate services. Core functionalities were validated, including user authentication via IdP, management of JupyterHub notebooks, and execution of remote tasks. This deployment confirmed the AKBS's foundational stability.

To further validate the platform's capacity, we executed a multi-horizon active layer thickness (ALT) prediction model developed by Ahajjam et al. (2025c). This workflow employs a weighted ensemble of Extra Trees, Bagging Regressor, and CatBoost models to predict ALT across the circumpolar permafrost region using thaw depth measurements from 115 Circumpolar Active Layer Monitoring (CALM) (Brown et al., 2000) network sites spanning 1962–2024. The model integrates geospatial features including topographic variables (elevation, slope, latitude, longitude) derived from AW3D30, vegetation indices (EVI) from MOD13Q1, and surface pressure over perennially frozen ground from ERA5-Land. Predictions are generated at same-year, +1, +2, and +5 year horizons. The pipeline executed data ingestion, 5-fold cross-validation, model training, and test set inference across all horizons in under 5 min on the A-KBS deployment, orchestrated by the University of North Dakota's (UND) K8s cluster and executed on Amazon Web Services K8s resources (Elastic Kubernetes Service), with inference requiring less than 1.5 s. The ensemble achieved R^2 values exceeding 0.80 across all prediction horizons (see Table 1, Fig. 6). As the A-KBS gateway integrates additional federated HPC resources, these workflows will support even broader analyses, such as pan-Arctic inference across millions of grid cells and the processing of massive LiDAR and NISAR datasets, while maintaining the rapid turnaround required for operational decision-making.

This success underscores the gateway's ability to handle workloads while shielding the researcher from technical overhead. By leveraging Globus services and integrating them with the JupyterHub interface, the platform streamlined the process of offloading computationally intensive tasks during the test. This eliminated the need for users to manually interact with the Slurm workload manager or manage secure shell connections. The A-KBS's access to computational resources minimized technical barriers, allowing researchers to focus on their workflows.

Table 1

Active layer thickness (ALT) prediction performance and execution time across prediction horizons. The weighted ensemble combines Extra Trees, Bagging Regressor, and CatBoost models using inverse-RMSE weighting.

Horizon	Model	RMSE	R^2	Train. Dur.	Inference Dur.
Same Year	Extra Trees	24.946	0.802	10.98	0.088 s
	Bagging Regressor	26.362	0.779	65.22	0.326 s
	CatBoost	25.493	0.793	5.75	0.002 s
	Ensemble	24.241	0.813	–	–
+1 Year	Extra Trees	24.494	0.809	7.73	0.064 s
	Bagging Regressor	26.920	0.770	54.00	0.323 s
	CatBoost	26.730	0.773	5.30	0.002 s
	Ensemble	24.328	0.812	–	–
+2 Year	Extra Trees	24.488	0.809	5.45	0.044 s
	Bagging Regressor	27.192	0.765	51.16	0.262 s
	CatBoost	24.779	0.805	9.35	0.002 s
	Ensemble	24.308	0.812	–	–
+5 Year	Extra Trees	24.654	0.807	3.39	0.026 s
	Bagging Regressor	27.324	0.763	59.12	0.345 s
	CatBoost	26.023	0.785	8.96	0.002 s
	Ensemble	24.574	0.808	–	–

4.2. Scalability and interoperability

The platform's modular design and use of containerized services also proved effective in achieving horizontal scalability and interoperability. The headless middleware layer, built with Nest.JS, successfully coordinated interactions between the A-KBS frontend and backend components using the GraphQL and REST APIs. This approach demonstrated the system's flexibility to handle growing user demand and integrate with new external resources. The reliance on standards-based tools like Globus and K8s further validated the architecture's ability to adapt to different computing infrastructures, from on-premise HPC clusters to future cloud environments.

4.3. Data reliability and unified workflow

The A-KBS's initial deployment demonstrated two key successes: the unification of complex services and enhanced data reliability during operation. An important outcome is the implementation of the IdP and SSO, eliminating a point of friction and allowing researchers to use

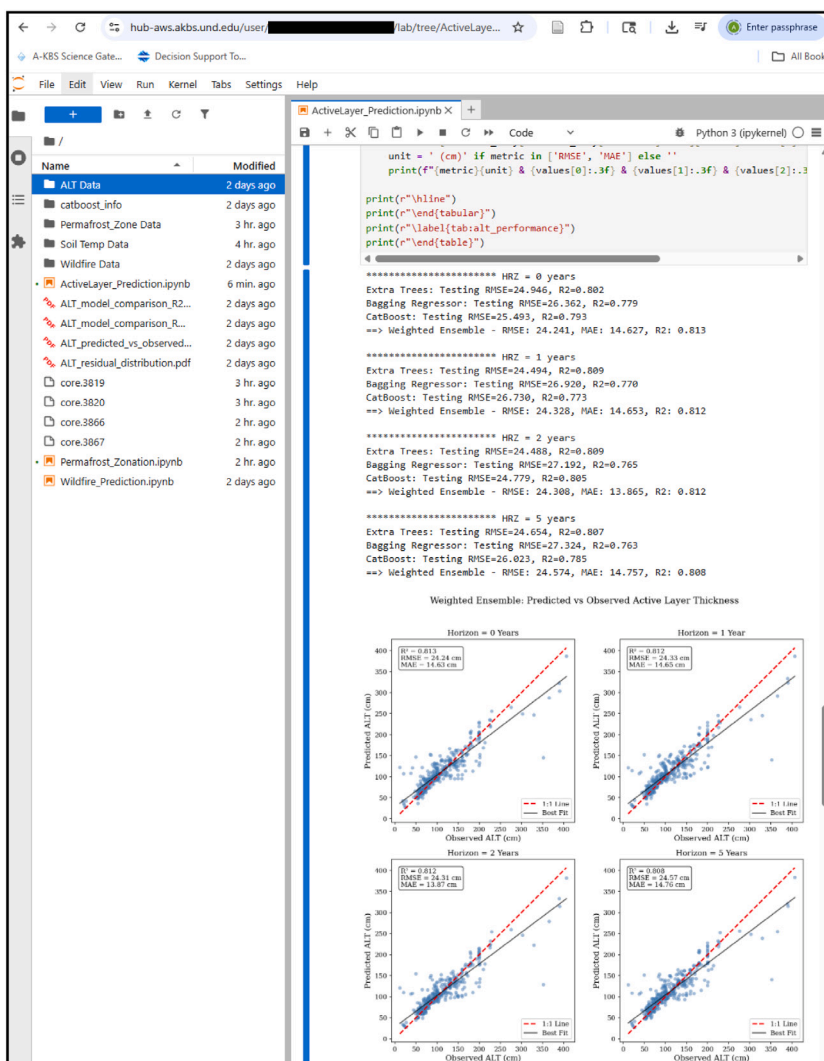


Fig. 6. Example of the multi-horizon active layer thickness (ALT) prediction workflow executed through the A-KBS Science Gateway, illustrating interactive model execution and visualization as the gateway automatically abstracts to remote HPC resources.

their institution credentials. Concurrently, the PostgreSQL database, managed securely through the middleware, provided a reliable data store. The automated backup and monitoring procedures confirmed that the platform can maintain data integrity and availability to support research into complex Arctic geospatial processes and cryospheric sciences.

5. Discussion, challenges, and future research

This section discusses the broader challenges encountered during the development of the A-KBS and outlines key areas for future research to enhance its capabilities and long-term use.

5.1. Remaining technical challenges

Despite the solutions implemented, several technical challenges remain, highlighting a broader need for continued research in science gateway development.

- **Scalability to Petascale:** While our approach addresses many scalability issues, bottlenecks remain when supporting highly concurrent AI workloads. Future research is needed on more efficient

workload distribution and adaptive scheduling for data-intensive Arctic petascale modeling.

- **API and Data Interoperability:** Integrating diverse data sources in formats like NetCDF and HDF5 remains a challenge. While flexible ingestion pipelines were created, future work will focus on automatic data organization and building frameworks for rapidly adapting to data schema changes.
- **Database Reliability at Petascale:** Scaling the A-KBS's PostgreSQL database to petascale workloads presents directions for new research into adaptive data partitioning, query optimization for AI-driven workloads, and federated schema harmonization.
- **Security and Interoperability:** A significant challenge lies in consistently applying RBAC across a variety of different computing systems and platforms. This is particularly difficult when aligning with different security and data policies from multiple institutions within a single, unified login system such as the A-KBS.

5.2. Future research directions

Future research will prioritize:

- Developing AI-driven tools that automatically compose and configure analysis pipelines to exactly meet user goals.
- Developing cross-domain interoperability frameworks to seamlessly support multidisciplinary research and facilitate data exchange.
- Creating dynamic personalization engines that tailor tool and content recommendations based on user roles, history, and domain.
- Developing a framework to deploy and scale A-KBS endpoints across a variety of HPC environments, including ACCESS-CI systems like Expanse.

6. Conclusion

In this paper, we have detailed the design, implementation, and deployment of the A-KBS Science Gateway, a platform developed for AI-assisted modeling of permafrost dynamics and other geospatial processes. The A-KBS architecture leverages distributed and HPC backends to provide a single-point access system. By removing the complexities of underlying cyberinfrastructure, such as manual job scheduling and data management on platforms such as the UND Talon HPC cluster, the A-KBS lowers technical barriers for domain scientists. Its deployment and operational validation confirm its stability, scalability, and foundational reliability as a tool for enhancing research into permafrost dynamics and cryosphere-related sciences.

The A-KBS establishes a scalable foundation for future development. By linking advanced scientific workflows to decision-support frameworks, the gateway highlights a key role in translating complex research into actionable insights for national security and Arctic infrastructure resilience. Future development will focus on a range of capabilities that further broaden the platform's utility. This includes enhancing tools for processing highly concurrent, petascale workloads, facilitating the platform's deployment on HPC resources such as Expanse. The A-KBS demonstrates the capacity of science gateways to accelerate data workflows, and facilitate the processing of global-scale data leveraging orchestrated, containerized applications, providing researchers and collaborators with the tools required for accelerated post-processing of Arctic data leveraging AI/ML workflows for enhanced insight and decision support tools in both online and connectivity constrained environments.

CRedit authorship contribution statement

Andrew Wilcox: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Meisam Shayeghmoradi:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Stephen Miller:** Writing – original draft, Software, Methodology, Conceptualization. **Ian Nesbitt:** Writing – review & editing, Software, Methodology, Conceptualization. **Sheridan Parker:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Aaron Bergstrom:** Writing – review & editing, Supervision, Software, Project administration. **Naima Kaabouch:** Writing – review & editing, Supervision, Project administration. **Timothy Pasch:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors thank Saisri Pogalla for her technical contributions to the A-KBS gateway implementation.

DISTRIBUTION A: Approved for Public Release. Distribution is Unlimited. This material is based upon work supported by the Broad Agency Announcement Program and the Cold Regions Research and Engineering Laboratory (ERDC-CRREL) under Contract No. W913E524C0017.

Data availability

The authors do not have permission to share data.

References

- Ahajjam, A., Allgaier, M., Chance, R., Chukwuemeka, E., Putkonen, J., Pasch, T., 2025a. Enhancing prediction of wildfire occurrence and behavior in alaska using spatio-temporal clustering and ensemble machine learning. *Ecol. Inform.* 85, 102963.
- Ahajjam, A., Putkonen, J., Chukwuemeka, E., Chance, R., Pasch, T.J., 2024. Predictive analytics of air temperature in alaskan permafrost terrain leveraging two-level signal decomposition and deep learning. *Forecasting* 6 (1), 55–80.
- Ahajjam, A., Soaper, M., Chance, R., Chandler, J., Pasch, T., 2025b. Predicting freeze-thaw states in alaska permafrost landscapes using climate reanalysis and machine learning. *Int. J. Appl. Earth Obs. Geoinf.* 140, 104576.
- Ahajjam, A., Wilcox, A., Soaper, M., Chance, R., Pasch, T., 2025c. Multi-horizon active layer thickness prediction across circum-arctic permafrost regions using geospatial machine learning. *J. Geophys. Res.: Mach. Learn. Comput.* 2 (4), e2025JH000969.
- Ananthakrishnan, R., Babuji, Y., Baughman, M., Bryan, J., Chard, K., Chard, R., Clifford, B., Foster, I., Katz, D.S., Hunter Kesling, K., Janidlo, C., Mello, R., Wang, L., 2024. Enabling remote management of faas endpoints with globus compute multi-user endpoints. In: *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing*. ACM, Providence RI USA, pp. 1–5. <http://dx.doi.org/10.1145/3626203.3670612>, URL <https://dl.acm.org/doi/10.1145/3626203.3670612>.
- Bauer, A., Pan, H., Chard, R., Babuji, Y., Bryan, J., Tiwari, D., Foster, I., Chard, K., 2024. The globus compute dataset: An open function-as-a-service dataset from the edge to the cloud. *Future Gener. Comput. Syst.* 153, 558–574. <http://dx.doi.org/10.1016/j.future.2023.12.007>, URL <https://www.sciencedirect.com/science/article/pii/S0167739X23004703>.
- Becciani, U., Sciacca, E., Costa, A., Massimino, P., Pistagna, C., Riggi, S., Vitello, F., Petta, C., Bandieramonte, M., Krokos, M., 2015. Science gateway technologies for the astrophysics community. *Concurr. Comput.: Pr. Exp.* 27 (2), 306–327. <http://dx.doi.org/10.1002/cpe.3255>, URL <https://onlinelibrary.wiley.com/doi/10.1002/cpe.3255>.
- Brown, J., Hinkel, K.M., Nelson, F., 2000. The circumpolar active layer monitoring (CALM) program: research designs and initial results. *Polar Geogr.* 24 (3), 166–258.
- Chance, R., Ahajjam, A., Putkonen, J., Pasch, T., 2024. Artificial intelligence for predicting arctic permafrost and active layer temperatures along the alaskan north slope. *Earth Sci. Inform.* 17 (6), 6055–6073.
- Chard, K., Tuecke, S., Foster, I., 2016. Globus: Recent enhancements and future plans. In: *Proceedings of the XSEDE16 Conference on Diversity, Big Data, and Science At Scale*. In: XSEDE16, Association for Computing Machinery, New York, NY, USA, pp. 1–8. <http://dx.doi.org/10.1145/2949550.2949554>, URL <https://dl.acm.org/doi/10.1145/2949550.2949554>.
- Cody, R.P., Kassin, A., Gaylord, A., Brown, J., Tweedie, C.E., 2012. BAID: The Barrow Area Information Database - an Interactive Web Mapping Portal and Cyberinfrastructure for Scientific Activities in the Vicinity of Barrow, Alaska, vol. 2012, pp. GC21A–0936, ADS Bibcode: 2012AGUFMGC21A0936C. URL <https://ui.adsabs.harvard.edu/abs/2012AGUFMGC21A0936C>.
- Decker, J., Metje, S., Kunkel, J., 2025. Running kubernetes workloads on rootless HPC systems using slurm. *Cloud Comput.*
- Dickson, J., Kelleher, C., Bakker, J., Cloutier, M., Dalbotten, D.M., Im, D., Hossaini, A., Johnson, R., Klassen, J., Linde, S., 2024. Polar geospatial center: High-resolution monitoring and geospatial/remote sensing training and support for arctic and antarctic researchers. In: *AGU Fall Meeting Abstracts*, vol. 2024, pp. C23A–0429, Issue: 429. URL <https://ui.adsabs.harvard.edu/abs/2024AGUFMC23A.0429D/abstract>.
- Gesing, S., Pierce, M., Marru, S., Zentner, M., Huff, K., Bradley, S., B. Cleveland, S., R. Brandt, S., Ramnath, R., Kee, K., Dahant, M., M. Villegas Martínez, B., Contreras Sepulveda, W., J. Sánchez Mondragón, J., 2024. Science gateways and AI/ML: how can gateway concepts and solutions meet the needs in data science? In: *Critical Infrastructure - Modern Approach and New Developments*. IntechOpen, pp. 1–21. <http://dx.doi.org/10.5772/intechopen.110144>, URL <https://www.intechopen.com/chapters/86501>.

- Hjort, J., Streletskiy, D., Doré, G., Wu, Q., Bjella, K., Luoto, M., 2022. Impacts of permafrost degradation on infrastructure. *Nat. Rev. Earth Environ.* 3 (1), 24–38, URL <https://www.nature.com/articles/s43017-021-00247-8>.
- Jones, M.B., Dozier, J., Schildhauer, M.P., Budden, A.E., 2021. Advancing arctic research and education through data preservation and reuse at the arctic data center. NSF Award. 20, 42102, ADS Bibcode: 2021nsf....2042102J. URL <https://ui.adsabs.harvard.edu/abs/2021nsf....2042102J>.
- Kalyanam, R., Zhao, L., Song, C., Biehl, L., Kearney, D., Kim, I.L., Shin, J., Villoria, N., Merwade, V., 2019. MyGeoHub—A sustainable and evolving geospatial science gateway. *Future Gener. Comput. Syst.* 94, 820–832. <http://dx.doi.org/10.1016/j.future.2018.02.005>, URL <https://linkinghub.elsevier.com/retrieve/pii/S0167739X17310087>.
- Kellogg, K., Hoffman, P., Standley, S., Shaffer, S., Rosen, P., Edelstein, W., Dunn, C., Baker, C., Barela, P., Shen, Y., et al., 2020. NASA-isro synthetic aperture radar (NISAR) mission. In: 2020 IEEE Aerospace Conference. IEEE, pp. 1–21.
- Liljedahl, A.K., Jones, B.M., Brubaker, M., Budden, A.E., Cervenec, J.M., Grosse, G., Jones, M.B., Marini, L., McHenry, K., Moss, J., Morin, P.J., Nitze, I., Soliman, A., Wind, G., Witharana, C., 2019. Permafrost discovery gateway: A web platform to enable discovery and knowledge-generation of permafrost big imagery products. In: EPIC3AGU Fall Meeting 2019, San Francisco, USA, 2019-12-09-2019-12-13 San Francisco, USA, AGU. AGU, San Francisco, USA, URL <https://epic.awi.de/id/eprint/50806/>.
- Liljedahl, A., Jones, M., Budden, A., Witharana, C., Cervenec, J., Jones, B., Jones, C., Marini, L., Walker, L., Hasan, A., 2021. The permafrost discovery gateway. In: AGU Fall Meeting Abstracts, vol. 2021, pp. C55E–01, URL <https://ui.adsabs.harvard.edu/abs/2021AGUFM.C55E..01L/abstract>.
- Liljedahl, A., Nitze, I., Witharana, C., Manos, E., Jones, M., Li, W., McHenry, K., Perera, A., Cohen, J., Nicholson, T., Marini, L., Thiessen-Bock, R., 2024. Permafrost discovery gateway: Exploring hydrological changes through big geospatial data (invited). 104, 440910, ADS Bibcode: 2024AMS...10440910L. URL <https://ui.adsabs.harvard.edu/abs/2024AMS...10440910L>.
- Lynch, V., Chen, M., Cobb, J., Kohl, J., Miller, S., Speirs, D., Vazhkudai, S., 2010. Neutron science TeraGrid gateway. *J. Phys.: Conf. Ser.* 251, 012097. <http://dx.doi.org/10.1088/1742-6596/251/1/012097>, URL <https://iopscience.iop.org/article/10.1088/1742-6596/251/1/012097>.
- Madupati, B., 2021. Kubernetes: Advanced deployment strategies* technical perspective. Preprint on Zenodo.
- Mann-Bruch, K., 2022. Hubzero® reaches 20th anniversary. URL https://www.sdsc.edu/news/2022/PR20220817_hubzero.html.
- Marchand, N., Royer, A., Krimmer, G., Roy, A., Langlois, A., Vargel, C., 2018. Snow-covered soil temperature retrieval in Canadian arctic permafrost areas, using a land surface scheme informed with satellite remote sensing data. *Remote. Sens.* 10 (11), 1703. <http://dx.doi.org/10.3390/rs10111703>, Number: 11 Publisher: Multidisciplinary Digital Publishing Institute. URL <https://www.mdpi.com/2072-4292/10/11/1703>.
- of Minnesota, U., 2025. Introduction to user services – polar geospatial center. URL <https://www.pgc.umn.edu/guides/user-services/introduction/>.
- Morin, P.J., 2016. The polar geospatial center: Community and facility support. NSF Award. Number 1559691. *Dir. Geosci.* 15 (1559691), 59691, URL <https://ui.adsabs.harvard.edu/abs/2016nsf....1559691M/abstract>.
- National Science Foundation, (n.d.a) Science Gateways | ACCESS Support. n.d. URL <https://support.access-ci.org/tools/science-gateways>.
- National Science Foundation, (n.d.b) NSF Arctic Data Center The primary data and software repository for NSF Arctic research. URL <https://arcticdata.io/>.
- National Science Foundation, (n.d.c) NSF Arctic Data Center The primary data and software repository for NSF Arctic research. URL <https://arcticdata.io/about/>.
- Obu, J., 2021. How much of the earth's surface is underlain by permafrost? *J. Geophys. Res.: Earth Surf.* 126 (5), e2021JF006123.
- Park, H., Kim, Y., Kimball, J.S., 2016. Widespread permafrost vulnerability and soil active layer increases over the high northern latitudes inferred from satellite remote sensing and process model assessments. *Remote Sens. Environ.* 175, 349–358. <http://dx.doi.org/10.1016/j.rse.2015.12.046>, URL <https://linkinghub.elsevier.com/retrieve/pii/S0034425715302686>.
- Pundsack, J.W., Nguyen, C.T., Porter, C.C., 2016. The polar geospatial center: Community and facility support. NSF Award. 15, 59691, ADS Bibcode: 2016nsf....1559691M. URL <https://ui.adsabs.harvard.edu/abs/2016nsf....1559691M>.
- Salesforce Research, 2022. State of commerce 2nd edition.
- Science Gateway Community Institute, sciencegateways - Gateways Central Welcome. n.d. URL <https://sciencegateways.org/gateways-central-welcome#/home>.
- Sperhac, J.M., Poinar, K., Jones-Ivey, R., Briner, J., Csatho, B., Nowicki, S., Simon, E., Larour, E., Quinn, J., Patra, A., 2021. Ghub: Building a glaciology gateway to unify a community. *Concurr. Comput.: Pr. Exp.* 33 (19), e6130. <http://dx.doi.org/10.1002/cpe.6130>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpe.6130>.
- Streletskiy, D.A., Clemens, S., Lanckman, J.-P., Shiklomanov, N.I., 2023. The costs of arctic infrastructure damages due to permafrost degradation. *Environ. Res. Lett.* 18 (1), 015006, Publisher: IOP Publishing. URL <https://iopscience.iop.org/article/10.1088/1748-9326/acab18/meta>.
- Tempesta, J., 2025. Collective action for a changing arctic. URL <https://www.nceas.ucsb.edu/news/collective-action-changing-arctic>.
- Tulenko, J.P., Goliber, S.A., Jones-Ivey, R., Quinn, J., Patra, A., Poinar, K., Nowicki, S., Csatho, B.M., Briner, J.P., 2025. Brief communication: Enabling open cryosphere research with ghub. *EGU sphere* 2025, 1–12, Publisher: Copernicus Publications Göttingen, Germany. URL <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-894/>.
- Tweedie, C.E., Webber, P.J., 2003. CEON: A circum-arctic environmental observatories network. URL https://www.researchgate.net/publication/252673549_CEON_A_Circum-Arctic_Environmental_Observatories_Network.
- University of Chicago, (2026). Geospatial Search. URL https://docs.globus.org/api/search/guides/geo_search/.
- Webber, P., 2004. Circumarctic environmental observatories network. CEON. background document for the meeting of the senior arctic officials of the arctic council, selfoss, iceland, may 4-5, 2004.. URL <http://hdl.handle.net/11374/598>. Publisher: Circumarctic Environmental Observatories Network (CEON).
- Wilkins-Diehr, N., 2007. Special issue: Science gateways—Common community interfaces to grid resources. *Concurr. Comput.: Pr. Exp.* 19 (6), 743–749. <http://dx.doi.org/10.1002/cpe.1098>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpe.1098>.
- Wilkins-Diehr, N., 2011. A history of the TeraGrid science gateway program: a personal view. In: Proceedings of the 2011 ACM Workshop on Gateway Computing Environments. ACM, Seattle Washington USA, pp. 1–12. <http://dx.doi.org/10.1145/2110486.2110488>, URL <https://dl.acm.org/doi/10.1145/2110486.2110488>.
- Zhou, N., Georgiou, Y., Pospieszny, M., Zhong, L., Zhou, H., Niethammer, C., Pejak, B., Marko, O., Hoppe, D., 2021. Container orchestration on HPC systems through kubernetes. *J. Cloud Comput.* 10 (1), 16. <http://dx.doi.org/10.1186/s13677-021-00231-z>, URL <https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-021-00231-z>.
- Zonca, A., Sinkovits, R.S., 2018. Deploying jupyter notebooks at scale on XSEDE resources for science gateways and workshops. In: Proceedings of the Practice and Experience on Advanced Research Computing: Seamless Creativity. PEARC '18, Association for Computing Machinery, New York, NY, USA, pp. 1–7. <http://dx.doi.org/10.1145/3219104.3219122>, URL <https://dl.acm.org/doi/10.1145/3219104.3219122>.