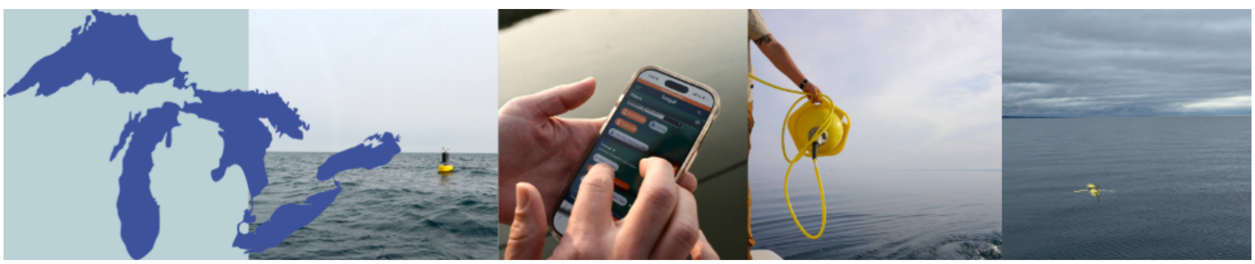
great lakes
observing system



GLOS Integrated Operations and Data Management Manual

Published : April 30, 2026

Introduction

Standard Operating Procedures and Governance Framework for the Great Lakes Observing System

The Great Lakes Observing System (GLOS) operates a comprehensive, integrated regional system that connects observing infrastructure, data management, user engagement, and technology development to deliver reliable, real-time environmental intelligence across the Great Lakes. This Integrated Operations and Data Management Manual serves as the foundational governance and operations framework guiding how GLOS designs, implements, and sustains this system in alignment with U.S. Integrated Ocean Observing System (IOOS) requirements and NOAA data management policies.

At its core, GLOS functions as a coordinating backbone for a distributed, multi-sector observing network, supporting over 40 million people who rely on the Great Lakes for drinking water, transportation, recreation, and economic activity. The system integrates data from buoys, gliders, radar systems, remote sensing platforms, and partner-provided datasets into a unified cyberinfrastructure that enables seamless data ingestion, quality control, analysis, and public dissemination through platforms such as Seagull.

This manual outlines the **core components** of the GLOS regional system:

- **Observing System Operations** define how physical assets are deployed, calibrated, maintained, and sustained in partnership with academic, federal, tribal, and private-sector contributors to ensure long-term, high-quality environmental data collection.
- **Data Management and Cyberinfrastructure (DMAC)** establishes standardized workflows for ingesting, processing, quality control (including QARTOD implementation), archiving, and distributing data in near real-time, ensuring alignment with IOOS standards and FAIR data principles.
- **User Engagement and Governance** ensures that system priorities are continuously informed by stakeholder input through structured engagement, feedback loops, and transparent decision-making processes.
- **Software Development and Delivery** defines agile, user-centered approaches to building and maintaining digital platforms, including web and mobile applications, APIs, and data services that translate complex datasets into actionable insights.
- **Quality Management Planning (QMP)** provides a system-wide framework for ensuring data integrity, operational reliability, and continuous improvement across all GLOS activities.

Together, these components form a **scalable, cloud-based, and standards-driven architecture** that supports both operational decision-making and long-term scientific research.

GLOS also maintains established policies and practices to ensure the responsible management of data beyond environmental observations, including the handling of user information associated with its digital platforms. Data privacy, protection, and appropriate use of analytics are governed through formal policies and operational controls, ensuring transparency and accountability in how user data are collected, managed, and utilized. These practices are documented in GLOS' publicly available privacy policy and are implemented across all relevant systems and applications.

Looking ahead, GLOS continues to advance its capabilities to meet evolving regional and national needs. This includes expanding observing coverage to address critical data gaps, modernizing cyberinfrastructure to support increasing data volumes and advanced analytics, and exploring emerging technologies such as artificial intelligence and machine learning for predictive modeling and risk assessment.

By integrating these domains into a single operational framework, GLOS ensures that its regional observing system remains **trusted, transparent, user-driven, and future-ready**. This manual not only documents how the system operates today, but also establishes a clear, adaptable foundation for innovation, collaboration, and long-term sustainability across the Great Lakes region.

---

1. Observing System Operations, Maintenance, and Strategic Evolution

The GLOS Observing System is a distributed, partner-driven network designed to deliver reliable, high-quality environmental observations across the Great Lakes basin. The system integrates a diverse portfolio of in situ and remote sensing platforms operated by academic institutions, federal and state agencies, tribal and First Nations partners, and private organizations.

GLOS plays a central coordination role by sustaining high-priority observing assets, integrating partner data streams, and ensuring that observations are aligned with regional user needs and national IOOS priorities. The observing system is designed to balance operational reliability, long-term data continuity, and strategic expansion, ensuring that the network remains fit-for-purpose as environmental conditions, technologies, and user requirements evolve.
This section defines the operating procedures for observing system operations, asset management, and lifecycle maintenance, while also outlining the strategic framework used to sustain and enhance the system over time.

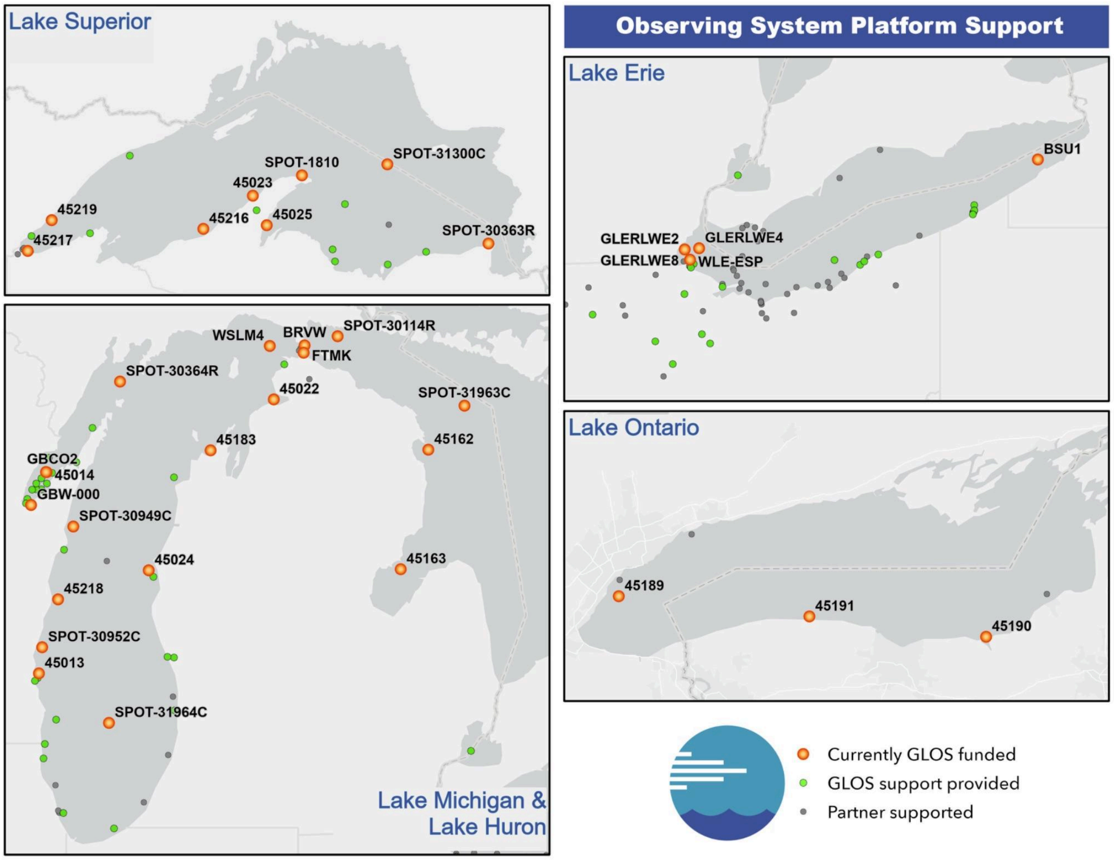


Figure 1: Maps of shoreline and off coast platforms either currently GLOS funded (large, hollow, orange dots), GLOS supported (small, hollow, green dots), or partner supported (smallest, gray dots).

1.1. Observing System Outcomes

GLOS maintains and evolves the observing system to deliver measurable outcomes that support science, operations, and decision-making across the Great Lakes region.

Core Outcomes

- **Sustained, year-round observations:** The system is designed to provide continuous environmental data across the Great Lakes basin, including expansion into winter operations where observational gaps have historically limited situational awareness.
- **Support for public safety and maritime operations:** Observations of waves, winds, currents, water levels, and temperature directly support safe navigation, search and rescue operations, and hazard forecasting for agencies such as the U.S. Coast Guard and National Weather Service.
- **Water quality and ecosystem monitoring:** The observing system supports monitoring of harmful algal blooms (HABs), water quality parameters, and ecosystem indicators critical to drinking water protection, fisheries management, and coastal resilience.

- **Long-term environmental records:** Sustained observations contribute to long-term time series datasets that are essential for climate monitoring, model validation, and trend analysis.
- **Modeling and forecasting integration:** Observational data are delivered in near real-time to support NOAA operational models and regional forecasting systems, improving predictive capabilities across the basin.

1.2 Observing System Architecture and Components

The GLOS observing system is composed of multiple platform types and partner contributions that together form an integrated regional observing enterprise.

Core Observing Platforms

- **Buoys:** Provide continuous measurements of meteorological and oceanographic variables including wave height, wind speed, temperature, and water quality parameters. These platforms form the backbone of the real-time observing network.
- **Underwater gliders:** Mobile platforms that collect vertical profiles of physical and biological variables across large spatial areas. Increasingly used for fisheries management and ecosystem monitoring.
- **High-Frequency Radar (HFR):** Shore-based radar systems that provide spatially continuous surface current data, supporting applications such as oil spill response, search and rescue, and coastal circulation analysis.
- **Shore-based and coastal systems (including cameras):** Fixed installations that provide visual observations and support emerging applications such as AI-driven wave analysis, shoreline monitoring, and ice detection, in addition to standard meteorological observations like air temperature and wind speed and direction.

Emerging and Supplemental Technologies

- **IoT-enabled sensors and low-cost platforms:** Expanded spatial coverage and deployment in previously underserved areas, including nearshore and community-based locations.
- **Remote sensing and satellite-derived datasets:** Integrated to complement in situ observations and provide basin-wide context for environmental conditions.
- **Cabled observatories (future opportunity**): Support high-frequency, high-resolution data collection for advanced research and operational applications.

Partner-Contributed Observations

- GLOS integrates data from **70+ regional partners**, including federal agencies, academic institutions, tribal organizations, and private sector contributors.
- Partner data are incorporated through standardized ingestion pipelines and made accessible through the GLOS DMAC system (see DMAC data ingestion section).

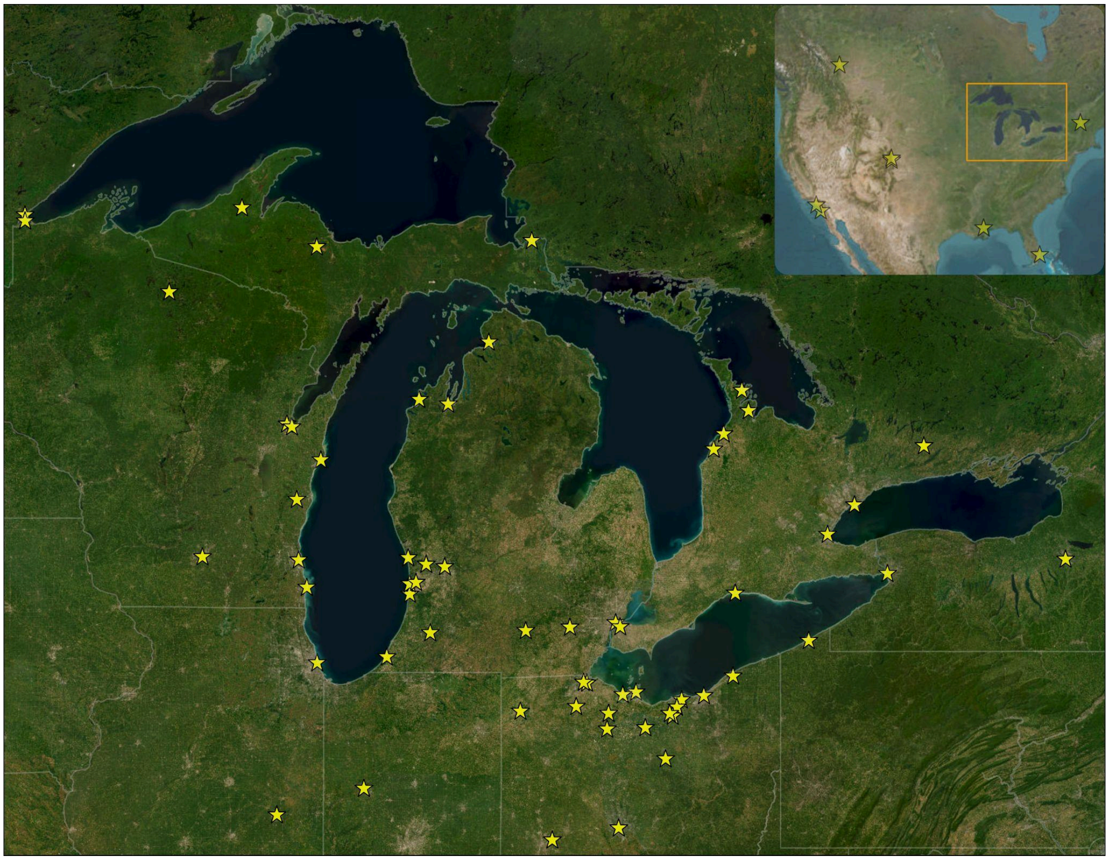


Figure 2: Locations of GLOS partner organizations

1.3 Observing System Operations and Maintenance

GLOS implements standardized procedures to ensure that observing assets operate reliably, generate high-quality data, and remain aligned with user needs.

Deployment and Recovery Operations

- Observing assets are deployed and recovered on seasonal or project-specific schedules, with careful coordination between GLOS and partner organizations.
- Deployment planning includes site validation, logistics coordination, and pre-deployment system checks by the platform managers in coordination with the GLOS staff.
- Recovery operations include data retrieval validation, physical inspection, and preparation for maintenance or redeployment.

Calibration and Validation Procedures

- Instruments are calibrated in accordance with manufacturer specifications and established scientific best practices.
- Calibration occurs:
    - Prior to deployment (baseline calibration)

    - During deployment (where feasible)
    - Post-recovery (validation and drift assessment)
- Ground-truthing activities, including water sampling and comparative measurements, are used to validate sensor performance wherever possible.

Operational Monitoring

- Real-time monitoring systems track data availability, transmission continuity, and system health.
- Automated alerts are being implemented to notify GLOS and data providers of data interruptions, anomalies, or equipment failures.
- Routine data checks are performed to ensure consistency and reliability prior to downstream processing (see DMAC QA/QC procedures).

Maintenance and Troubleshooting

- Preventative maintenance is conducted during scheduled recovery periods, including cleaning, sensor replacement, and system upgrades.
- Reactive maintenance is triggered by operational alerts or partner reports and prioritized based on system criticality.

Coordination with DMAC

- Observing operations are tightly integrated with DMAC workflows to ensure seamless data ingestion, quality control, and dissemination.
- Detailed ingestion, processing and QA/QC procedures are described in the DMAC section (See DMAC).

1.4 Asset Management and Lifecycle Tracking

GLOS maintains comprehensive records of all observing assets to support operational efficiency, accountability, and long-term system sustainability.

Inventory Management

- An internal inventory system tracks all equipment, including sensors, platforms, and supporting infrastructure. Inventory records include asset type, ownership, deployment status, and location and are used for grant reporting.
- Detailed Metadata is collected for each platform according to the IOOS metadata profile which is then shared with DMAC for further processing. (See GLOS_Metadata form.)

Instrument History and Lifecycle Tracking

- Each asset is associated with a lifecycle record documenting:
    - Deployment history
    - Maintenance activities

    - Instrumentation records
- Lifecycle tracking supports long-term performance evaluation and replacement planning.

Partner Responsibilities

- Roles and responsibilities for asset ownership, maintenance, and operation are clearly defined through agreements with observing partners. GLOS data providers are responsible for submitting metadata and providing timely deployment/recovery updates.
- GLOS provides oversight and coordination to ensure consistency across the distributed network.

1.5 Partner Coordination and Network Governance

The observing system operates as a collaborative network, with GLOS providing coordination, oversight, and support to ensure consistent performance and alignment with regional priorities.

Partner Engagement and Coordination

- Regular communication with observing partners includes:
    - Scheduled check-ins (monthly or quarterly)
    - Annual coordination meetings
    - Ad hoc issue resolution and technical support
- GLOS facilitates collaboration between partners and end users to ensure that observing assets meet real-world needs.

Training and Capacity Building

- GLOS supports partner training in deployment, maintenance, and data management practices (e.g. the May 2023 Seagull Workshop in Toronto). Workshops and guidance materials are provided to ensure consistency with IOOS standards and best practices.

Data Integration Support

- GLOS Observing and DMAC teams collaborate to onboard new data providers, ensuring compatibility with ingestion pipelines and metadata standards. Technical support is provided to partners to streamline data integration and improve data quality.

Shared Responsibility Model

- Observing system operations are distributed across partners, enhancing resilience and reducing single points of failure. GLOS maintains oversight to ensure network-wide consistency and performance.

1.6 Strategic Approach to Sustaining and Enhancing the Observing System

GLOS employs a structured, adaptive strategy to ensure that the observing system remains effective, sustainable, and responsive to evolving needs.

Guiding Principles

- **Balance continuity and innovation:** Sustain long-term datasets while incorporating new technologies and approaches.
- **User-driven prioritization:** Align observing investments with documented user needs and critical data gaps.
- **Leverage partnerships:** Expand system capacity through collaboration and co-investment with regional partners.
- **Scalability and adaptability:** Design the system to accommodate new data types, technologies, and use cases.

1.7 System Expansion and Build-Out Strategy

GLOS maintains a long-term vision for expanding the observing system to address critical gaps and emerging priorities.

Priority Expansion Areas

- **Winter and year-round observations:** Expand deployments to address seasonal data gaps and improve safety and forecasting during winter months.
- **Water quality and ecosystem monitoring:** Increase real-time monitoring of HABs, pathogens, and other water quality indicators.
- **Nearshore and coastal coverage:** Improve spatial resolution in high-use and high-risk coastal areas.
- **Integration of new technologies:** Evaluate and adopt emerging technologies such as AI-enabled imaging, IoT sensors, and remote sensing platforms.

---

2. Data Management and CyberInfrastructure

The GLOS Data Management and Cyberinfrastructure (DMAC) system provides end-to-end data services supporting the ingestion, quality control, management, dissemination, and archival of observational and modeled data across the Great Lakes. The system is built on a cloud-native, scalable architecture (Seagull platform) designed to support real-time operations, interoperability with national systems, and long-term data stewardship.

DMAC operations are guided by IOOS DMAC standards, FAIR data principles, and NOAA data management policies, ensuring that data are accessible, reliable, and usable for a wide range of stakeholders including researchers, operational agencies, and the public.

2.1 Data Ingestion and Integration

GLOS operates a flexible, event-driven ingestion framework capable of integrating heterogeneous data streams from internal observing assets, partner systems, and external model providers in near real-time.

Standard Procedures

1. **Multi-source Data Ingestion**
   - GLOS ingests data from GLOS-funded platforms, regional partners, and federal sources (e.g., NOAA, USGS), supporting both observational and model datasets. Data sources include APIs, secure FTP, secure http JSON POST endpoints, and automated harvesting workflows.
   - Supports heterogeneous formats and transmission methods (real-time streams, batch uploads).

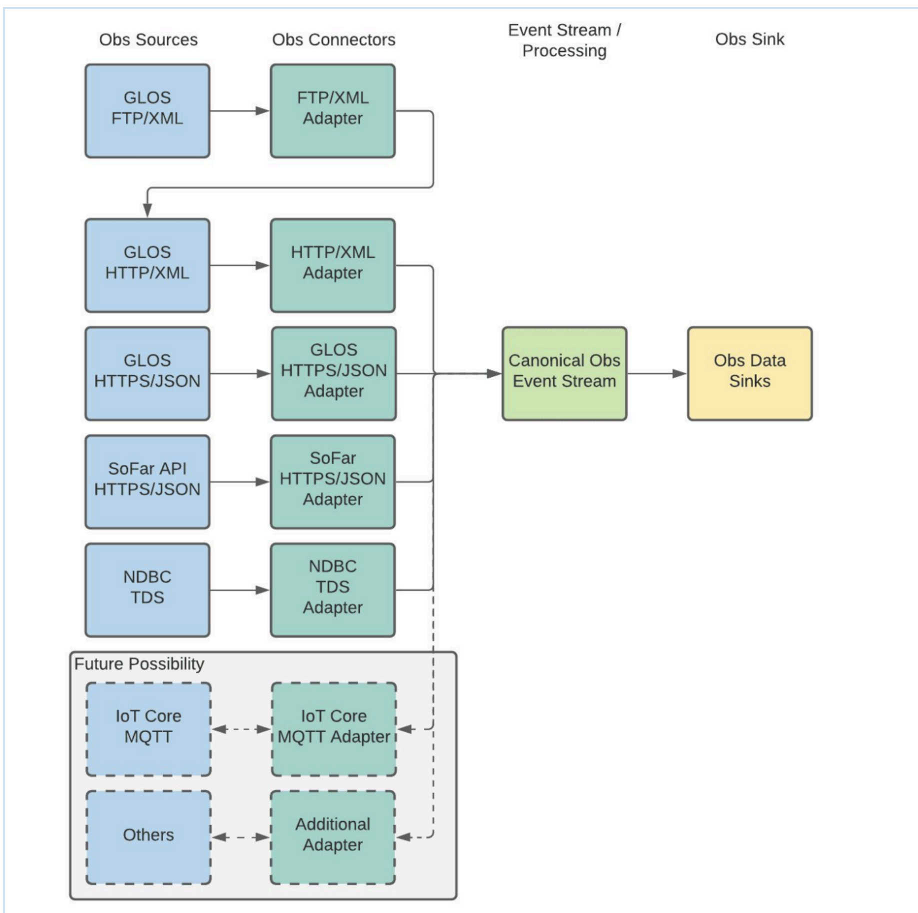


Figure 3: Demonstration of various near real time obs ingestion

2. **Event-driven streaming architecture**

- Data ingestion is implemented through streaming pipelines that enable near-real-time processing. These pipelines support high-frequency data flows and ensure minimal latency between observation and availability.

3. **Format validation and standardization**
    - Incoming data are validated against accepted formats (e.g., JSON), deduplicated and transformed into standardized internal representations aligned with community conventions (e.g., CF conventions).
4. **Metadata association at ingestion**
    - All ingested datasets are linked to metadata records at the point of entry, ensuring traceability, discoverability, and downstream interoperability.
5. **Scalable ingestion for emerging technologies**
    - The ingestion framework is designed to accommodate new and future data types including radar systems (HFR, IRaMS), webcams, IoT sensors, and cabled observatories, supporting increasing data volume and velocity.
6. **Integration of external model outputs**
    - Operational forecast models (e.g., hydrodynamic and wave models) are ingested, processed, and made available alongside observational datasets to support integrated data products.

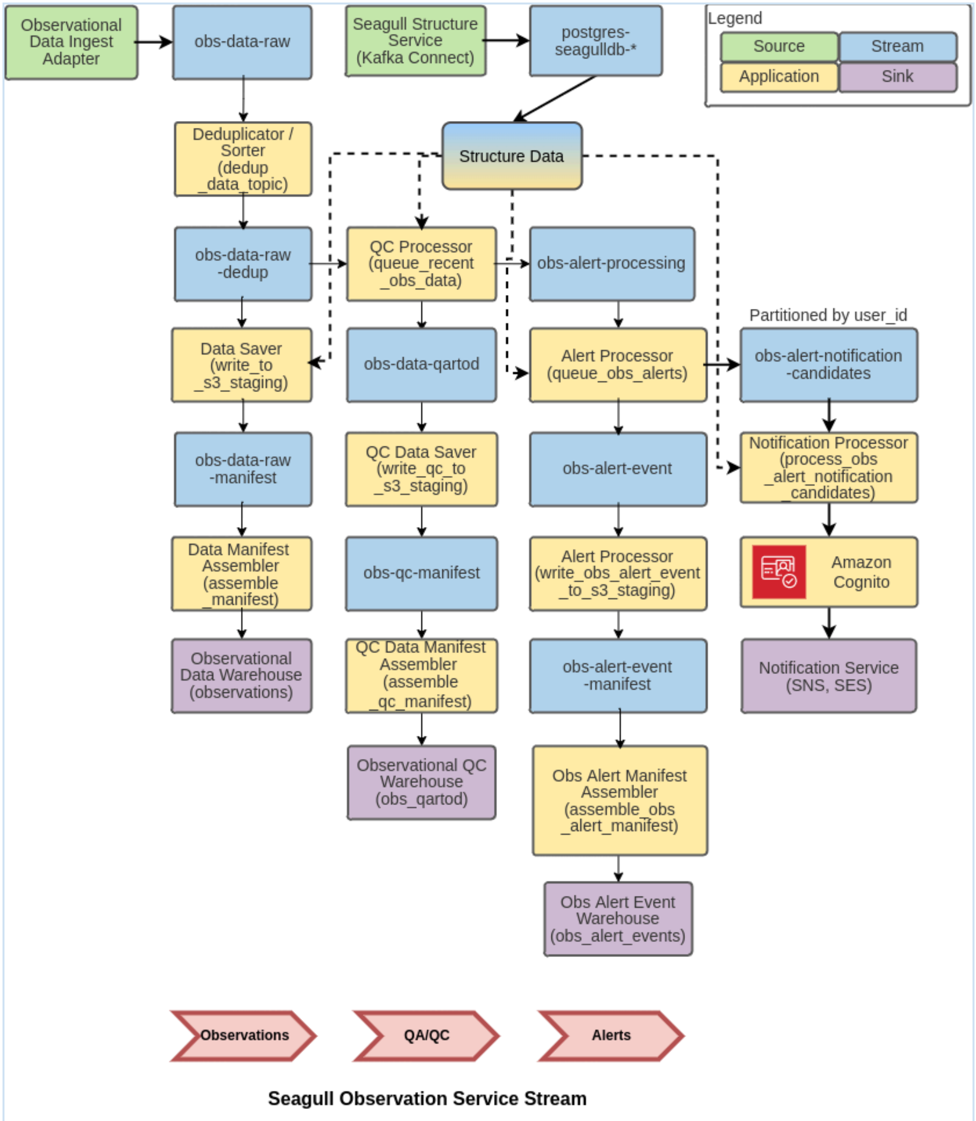


Figure 4: Seagull Platform's Observational Event Stream

2.2 Data Quality Control Procedures (QARTOD Compliance)

GLOS applies a comprehensive Quality Management Plan (QMP) incorporating automated and manual QA/QC processes aligned with IOOS QARTOD standards to ensure reliability, consistency, and usability of all distributed datasets.

Standard Procedures:

- **Implementation of QARTOD standards**
  - All applicable observational datasets are subjected to Quality Assurance/Quality Control of Real-Time Oceanographic Data (QARTOD) tests, including gross range, spike, rate-of-change, and flat-line checks.
- **Automated QC Processing**
  - **Stage 1:** Syntax and format validation during ingestion

- **Stage 2:** Automated [QARTOD checks] during processing
- **Stage 3:** Test results are stored alongside observations and linked via unique identifiers for traceability.

- **Real Time Flagging System**
  - Data points are assigned standardized quality flags indicating pass, suspect, or fail conditions, enabling users to assess data usability.
- **Manual review and anomaly detection**
  - GLOS staff and partners conduct routine reviews of data streams to identify anomalies, sensor drift, or systematic errors not captured by automated tests.
- **Feedback and Correction Loop**
  - QC results are shared with platform operators to support timely identification and correction of instrument or transmission issues.

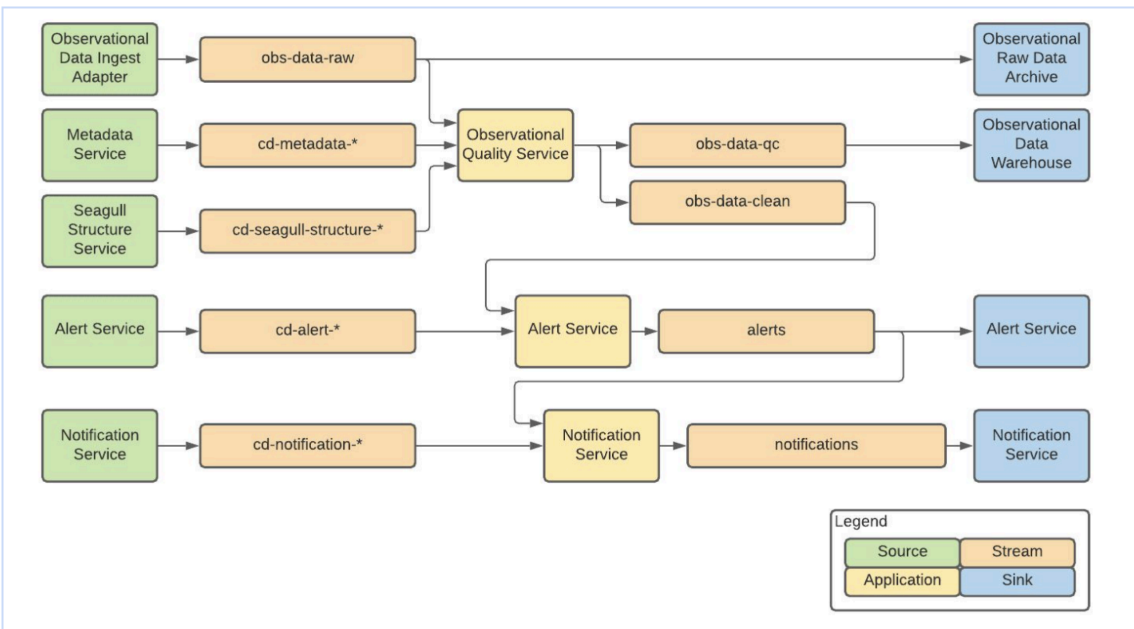


Figure 5: Demonstration of the role of Quality service on Seagull Platform

2.3 Cyberinfrastructure and System Architecture

The GLOS DMAC system is supported by a cloud-native, highly scalable cyberinfrastructure designed for reliability, performance, and operational resilience.

Standard Procedures:

- **Cloud-first architecture**
  - All core services are deployed within a cloud environment, enabling scalable storage, compute, and network capabilities to support growing data volumes and user demand. Services are documented through the Terraform infrastructure-as-code framework

    - With the exception of a part of authentication with Google, all infrastructure is deployed under Amazon Web Services (AWS). Google authentication is referenced in but not implemented via Terraform. Most infrastructure is implemented under the AWS Availability Zone (AZ) us-east-2 (Ohio), with some exceptions going through us-east-1 (N. Virginia).
- **Containerized and orchestrated services**
    - Backend services are deployed using containerization frameworks and managed through orchestration platforms (e.g., Kubernetes), ensuring modularity, high availability, and update-ability.
- **Streaming and distributed processing systems**
    - Data pipelines leverage distributed streaming technologies to handle real-time ingestion and processing at scale.
- **Data storage and management systems**
    - Structured and unstructured data are stored across optimized systems (e.g., relational databases, object storage, mounted cloud file systems), supporting both real-time access and long-term retention.
- **Infrastructure-as-code and automated deployment**
    - Infrastructure is provisioned and managed using infrastructure-as-code tools (e.g. Terraform), enabling reproducibility, version control, and efficient scaling of system components.
- **System monitoring, logging, and alerting**
    - Continuous monitoring of system health, performance, and data flows is implemented through logging frameworks (e.g. AWS CloudWatch) and automated alerting systems to ensure rapid response to issues.
- **Continuity of operations and resilience planning**
    - Backup systems, redundancy strategies, and disaster recovery protocols are maintained to ensure uninterrupted data services.

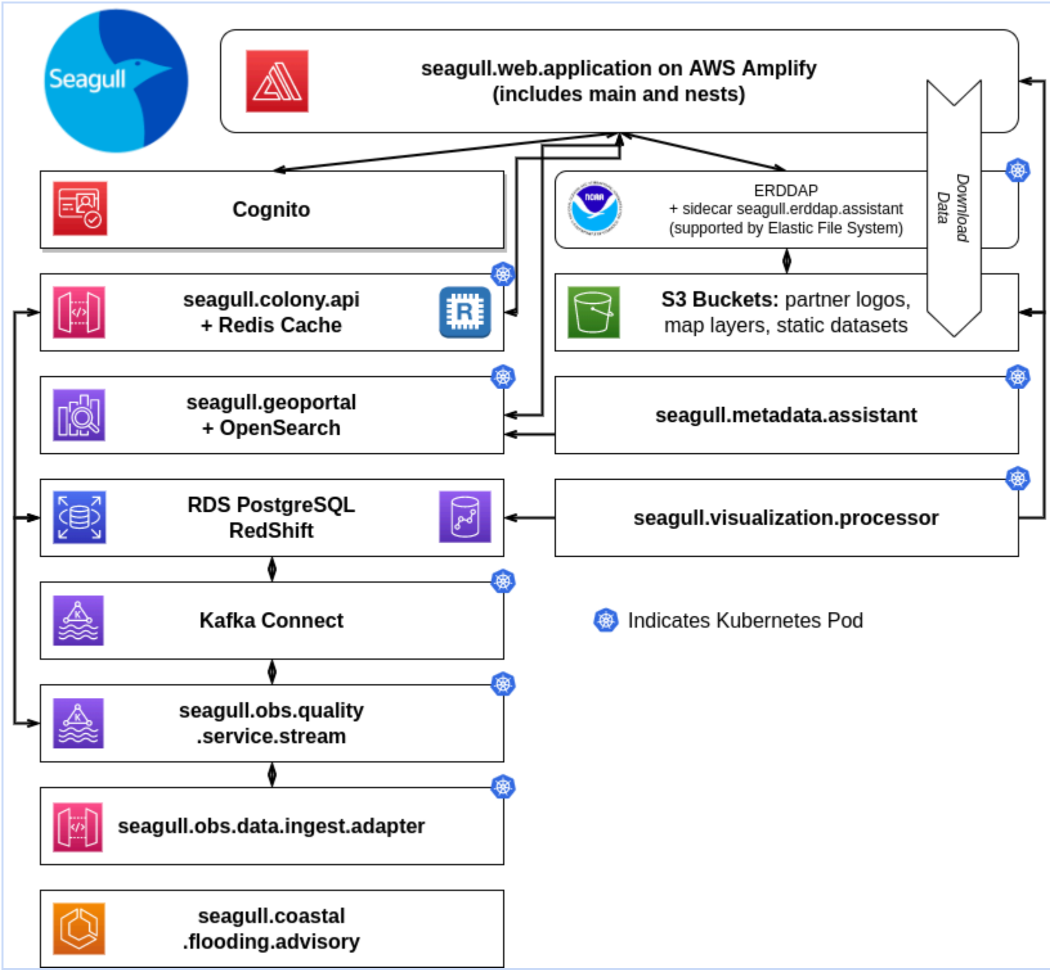


Figure 6: Seagull Platform at a glance

2.4 Data Access and Delivery

GLOS provides standardized, interoperable data access services that enable both human users and external systems to discover, access, and integrate data efficiently.

Standard Procedures

- **Standards-based data services (ERDDAP, APIs)**
  - Data are delivered through interoperable services including ERDDAP and REST APIs, supporting machine-to-machine access and integration with external platforms
    - GLOS Seagull ERDDAP: https://seagull-erddap.glos.org/
    - GLOS Seagull API: https://seagull-api.glos.org/docs
    - FAQ on their usage here: https://glos.org/data/faq/
- **Multi-format data delivery**
  - Datasets are made available in multiple standard formats (e.g., NetCDF, CSV, JSON) to accommodate diverse user needs and applications.

- **Integrated user-facing platforms (Seagull)**
    - Observational and model data are visualized and accessed through web and mobile (upcoming) software applications, providing intuitive interfaces for non-technical users.
- **Metadata-driven discovery systems**
    - Data discovery is supported through catalog services and search platforms aligned with IOOS metadata standards and ISO conventions. (e.g. GLOS geoportal: https://seagull-geoportal.glos.org/)
- **Real-time and near-real-time data availability**
    - GLOS prioritizes rapid data delivery to support operational decision-making, including forecasting, safety, and resource management.
- **Integration with national and global systems**
    - Data are shared with systems such as IOOS, National Data Buoy Center (NDBC), and the Global Telecommunications System (GTS), contributing to broader operational and forecasting networks.

2.5 Data Archiving and Long-Term Stewardship

GLOS ensures long-term preservation and accessibility of datasets through standardized archival processes aligned with national data centers and NOAA requirements.

Standard Procedures

- **Submission to National Archives**
    - Observational datasets are routinely submitted to NOAA's National Centers for Environmental Information (NCEI) through automated pipelines in accordance with established submission agreements. Some Real-time datasets are sent to GTS through NDBC.
- **Compliance with archival standards and formats**
    - Data are formatted and documented to meet archival requirements, including metadata completeness and adherence to accepted data standards. Formal agreements (e.g., NCEI Submission Agreement) govern archival processes.
- **Local and cloud-based redundancy**
    - Data are retained within cloud storage systems and supplemented with backup strategies to ensure data integrity and availability.
- **Tracking and reporting of archival submissions**
    - Submission status, completeness, and success rates are monitored and logged to ensure transparency and reliability of archival processes.
- **Long-term data stewardship practices**
    - GLOS maintains policies for data retention, preservation, and accessibility to support long-term scientific and operational use.

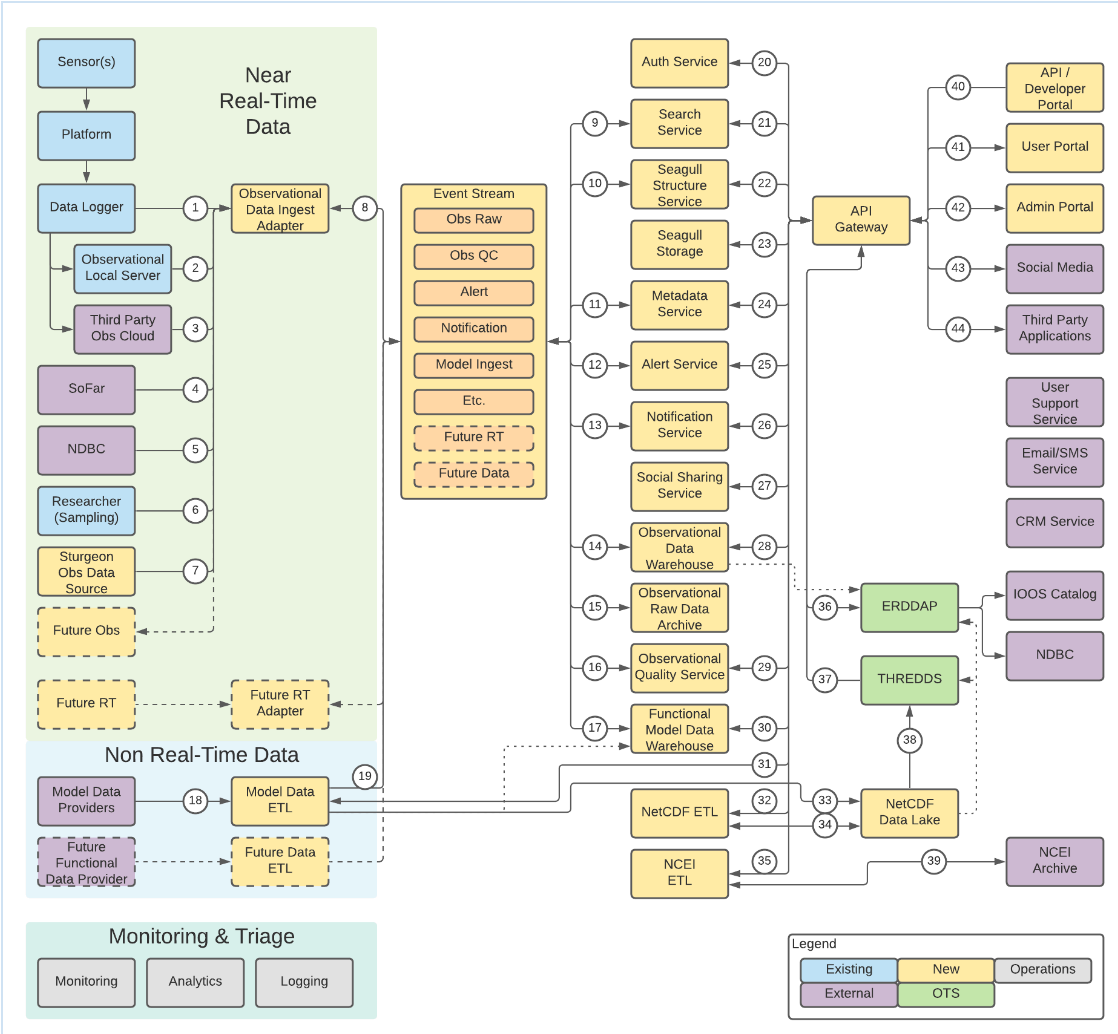


Figure 7: End to End Data flow on Seagull Platform

2.6 Implementation of IOOS/NOAA Data Management Protocols

GLOS adheres to established data policies and standards to ensure open, equitable, and responsible data management across the regional observing system.

Standard Procedures

- **Alignment with NOAA data policies and directives**
  - Data management practices comply with NOAA Administrative Orders and data sharing directives, ensuring consistency with federal requirements. GLOS prioritizes open access and timely data availability, particularly for datasets generated through IOOS funding.
- **IOOS Metadata Profile Compliance**
  - Metadata are aligned with IOOS metadata profile, including platform, sensor, and parameter descriptions, using CF conventions and controlled vocabularies.

- **Implementation of FAIR data principles**
  - Data are managed to be Findable, Accessible, Interoperable, and Reusable through standardized metadata, open access services, and persistent identifiers.
- **Catalog and Discovery Services**
  - Geoportal implementation supports ISO-compliant dataset discovery. Metadata are indexed and searchable across systems.
- **Protocol Adoption and Updates**
  - GLOS evaluates and adopts new IOOS protocols as released. Implementation occurs through iterative development cycles.
- **Interoperability Across Systems**
  - Integration with IOOS DACs, NOAA systems, and external data providers are supported.

2.7 System Evolution

GLOS continuously evolves its DMAC capabilities to support emerging technologies, advanced analytics, and next-generation data products.

Standard Procedures:

- **Scalable analytics environments**
  - Infrastructure supports advanced data processing and analytics workflows, including large-scale data analysis and model integration.
- **Development of derived data products**
  - GLOS generates value-added products by integrating observational and modeled data to support decision-making applications.
- **Exploration of AI and machine learning applications**
  - Opportunities for AI/ML are evaluated across QA/QC, data processing, and predictive analytics, including risk assessment tools for coastal safety.
- **Continuous system modernization**
  - Emerging technologies, cloud-native tools, and distributed processing frameworks are evaluated and integrated to enhance system performance and capability.
- **Alignment with national innovation strategies**
  - Technology development efforts are aligned with broader initiatives, including NOAA's artificial intelligence strategy and IOOS priorities.

---

3. User Engagement and Feedback Integration

GLOS implements a coordinated and continuous engagement framework to ensure that regional observing, data management, and data delivery activities are responsive to the needs of diverse user communities across the Great Lakes. This framework integrates stakeholder input, user feedback, and transparent decision-making processes into all stages of GLOS system design, development, and operations. Engagement activities are conducted through a combination of formal governance structures, direct stakeholder interactions, and user-facing platforms, enabling both structured and real-time input into GLOS priorities.

The procedures below define how GLOS systematically engages stakeholders, collects and integrates user feedback, and ensures transparency in regional system priorities and design. These processes are embedded across observing, data management, and user-facing systems to ensure that GLOS delivers data and information products that are responsive to diverse and evolving user needs across the Great Lakes.

3.1 Stakeholder Engagement and Input Solicitation

Objective: Ensure diverse, coordinated, and continuous stakeholder input into engagement strategies, observing priorities, and data product development.

- **Conduct Recurring Stakeholder Engagement Activities**
    - GLOS organizes regional workshops, surveys and targeted interviews with key user groups. These engagements provide structured opportunities to gather input on user needs, system gaps, and emerging priorities across recreational, maritime, research, and management communities.
- **Integrate Feedback Mechanisms into User Platforms**
    - Feedback tools are embedded within Seagull web and mobile applications. These tools enable continuous, real-time input from users, complementing formal engagement activities and capturing in-context feedback during application use.
- **Engage Governance and Advisory Structures**
    - GLOS engages its Board and advisory groups to review engagement strategies and regional priorities. These formal governance mechanisms provide strategic oversight and ensure alignment with organizational and regional objectives.
- **Coordinate with Regional and Federal Partners**
    - Engagement efforts are aligned with regional, federal, and academic partners. This coordination reduces duplication, strengthens regional consistency, and supports integrated stakeholder engagement across the Great Lakes ecosystem.
- **Engage Emerging User Groups**
    - GLOS proactively engages local stakeholders, Indigenous communities, and community-based organizations. This approach ensures broad representation in shaping system priorities and engagement strategies.
- **Maintain Defined Engagement Frequency**
    - Engagement is conducted continuously (via platforms), quarterly to biannually (via stakeholder interactions), and annually (via structured planning processes). This tiered frequency ensures both ongoing responsiveness and formal incorporation of stakeholder input into planning cycles.
- **Track Engagement Reach and Participation**
    - GLOS monitors engagement through platform usage and stakeholder participation metrics. Current reach (*as of 2025*) includes 250,000+ annual users

and ~4 million annual page views, complemented by recurring stakeholder engagement across the region.

3.2 User Feedback Collection, Assessment, and Response

Objective: Systematically collect, evaluate, and integrate user feedback to guide system improvements, data priorities, and product development.

- **Collect Feedback Through Multiple Channels**
    - GLOS gathers feedback via in-app tools (e.g. UserSnap), surveys, analytics (Google), and direct stakeholder engagement. This multi-channel approach ensures comprehensive coverage of both active and passive user input.
- **Monitor User Behavior Through Analytics**
    - Platform analytics are used to assess usage patterns, feature interactions, and engagement trends. These insights provide quantitative evidence of user needs and system performance at scale.
- **Aggregate and Review Feedback Regularly**
    - Feedback from all sources is consolidated and reviewed on a recurring basis. This ensures that user input is not siloed and is consistently considered across systems and teams.
- **Integrate Feedback into Planning and Prioritization**
    - User input is incorporated into annual planning processes, product roadmaps, and development backlogs. Prioritization balances user demand with technical feasibility, regional priorities, and IOOS guidance.
- **Implement Iterative Product Improvements**
    - GLOS applies iterative development cycles to refine features, data integration, and user experience.This allows for continuous improvement and rapid response to evolving user needs.
- **Validate Enhancements Through Pilot and Phased Rollouts**
    - New features and capabilities are tested through pilot implementations before full deployment. This reduces risk and ensures that changes effectively address user needs.
- **Track Feedback Impact on System Improvements**
    - GLOS links user feedback to implemented changes and system enhancements. This creates traceability between user input and outcomes, supporting accountability and continuous improvement.
- **Monitor System Usage and Feedback Scale**
    - Feedback processes are supported by analytics from 250,000+ users and interactions across 250+ observing platforms. This scale enables data-driven decision-making and prioritization of system enhancements.

3.3 Transparency of Priorities and System Design Decisions

Objective: Ensure that GLOS priorities, system design decisions, and data products are accessible, understandable, and visible to stakeholders and the public.

- **Publish System Information and Priorities Publicly**
    - GLOS shares priorities, project updates, and system developments through its website and user platforms. This ensures that stakeholders and the public have access to current information on system activities and direction.
- **Communicate Updates Through Multiple Channels**
    - Updates are disseminated via reports, newsletters, stakeholder communications, and regional events. This multi-channel approach increases visibility and accessibility of system developments.
- **Document Internal Decision-Making Processes**
    - GLOS maintains documentation of planning, prioritization, and system design decisions. This ensures consistency, accountability, and the ability to trace how decisions are made.
- **Coordinate Across Teams and Partners**
    - Decision-making is coordinated across DMAC, observing, and engagement teams, as well as regional and federal partners. This ensures alignment across system components and avoids fragmented or conflicting priorities.
- **Ensure Public Accessibility of Data and Outputs**
    - Real-time data and system outputs are delivered through publicly accessible platforms. These platforms provide visibility into system performance and outputs, reinforcing transparency.
- **Enable Stakeholder Understanding of Decision Processes**
    - GLOS communicates how user feedback and engagement inform system priorities and design. This helps stakeholders understand the linkage between their input and system evolution.
- **Track Transparency Through Platform Reach**
    - Transparency is supported by platforms delivering real-time data from 250+ observing assets to a broad user base. High engagement levels indicate effective public access to system outputs and information.

---

4. Software Development and Delivery

The Great Lakes Observing System (GLOS) develops, operates, and maintains software systems that enable the delivery of data and information products through its Data Management and Cyberinfrastructure (DMAC), including platforms such as the Seagull web application and Seagull Coast mobile application. These systems support the transformation of observational and modeled data into accessible, user-centered products that serve a wide range of stakeholders, including researchers, decision-makers, and the public.

The framework below defines the processes, methodologies, and governance frameworks used to guide software development and delivery across GLOS systems. It establishes a consistent and structured approach to designing, building, deploying, and maintaining software applications, ensuring alignment with user needs, organizational priorities, and national guidance, including NOAA software governance policies and IOOS DMAC standards.

The procedures emphasize iterative and agile development practices, user-centered design, robust quality assurance, and scalable cloud-based deployment. It also outlines coordination mechanisms between GLOS staff and contracted technology partners to ensure efficient execution, transparency, and continuity of operations. Through this approach, GLOS ensures that its software systems remain reliable, adaptable, and capable of supporting evolving data delivery needs and emerging technologies.

4.1 Development Methodology (Agile Framework)

- **Agile and Iterative Development**
  - GLOS follows Agile development methodology to enable adaptive planning, rapid iteration, and continuous improvement. Development is organized into short, iterative cycles that allow for frequent reassessment of priorities and incorporation of user feedback.
- **Sprint Planning and Execution**
  - Work is structured into defined sprints (typically 2 weeks), each with clearly scoped objectives, deliverables, and acceptance criteria. Regular sprint ceremonies including planning, daily stand-ups, and retrospectives ensure alignment across teams and continuous process improvement.
- **Backlog Management and Prioritization**
  - All development tasks, including features, enhancements, and bug fixes, are maintained in a centralized backlog (e.g., Jira). Items are prioritized based on user needs, system requirements, and strategic objectives, ensuring efficient allocation of development resources.

4.2 User-Centered Design and Product Development

- **User Personas and Use Cases**
  - Software design is guided by clearly defined user personas (e.g., recreational users, researchers, decision-makers). Features are developed based on user stories that articulate specific needs and expected outcomes.
- **User Journey Mapping and Interface Design**
  - GLOS employs user journey mapping to design intuitive workflows and interfaces. Design tools (e.g., Figma, Miro) are used to prototype and refine user interfaces prior to development.
- **Feedback-Driven Iteration**
  - User feedback is systematically collected through analytics, surveys, and feedback tools, and is directly incorporated into product development cycles to improve usability and functionality.

4.3 System Architecture and Design Principles

- **Modular and Scalable Architecture**
  - Systems are designed using modular components (e.g., APIs, microservices, data pipelines) to enable scalability, maintainability, and flexibility as system requirements evolve.
- **Cloud-Native Infrastructure**
  - The Seagull platform leverages cloud-native services, including containerization (Docker), orchestration (Kubernetes), and managed cloud services (AWS), to ensure high availability and scalability.
- **Standards-Based Design**
  - Software systems adhere to IOOS and NOAA standards for data formats, metadata, and service delivery (e.g., ERDDAP, NetCDF, ISO metadata), ensuring interoperability and integration with national systems.

4.4 Code Management and Version Control

- **Centralized Version Control**
  - All source code is maintained in centralized repositories (e.g. GitHub), ensuring traceability, collaboration, and version history management.
- **Branching and Release Strategy**
  - Structured branching strategies (e.g. one branch per new feature or bug fix) are used to support stable releases and controlled deployment workflows.
- **Code Review and Quality Assurance**
  - All code changes are reviewed through pull request workflows to ensure code quality, consistency, and adherence to development standards.

4.5 Testing and Quality Assurance

- **Automated and Manual Testing**
  - Testing frameworks include unit testing, integration testing, and system-level validation where applicable. Manual testing complements automated processes for user interface and workflow validation.
- **Pre-Deployment Validation**
  - All features are validated against defined acceptance criteria prior to release to ensure functionality and reliability.
- **Performance and Reliability Testing**
  - Systems are evaluated for performance, scalability, and resilience to ensure reliable operation under varying data loads and user demand.

4.6 Continuous Integration and Continuous Deployment (CI/CD)

- **Automated Build and Deployment Pipelines**
  - CI/CD pipelines are implemented to automate code integration, testing, and deployment processes using tools such as AWS CodeBuild and GitHub Actions.

- **Environment Management**
    - Separate environments (development, staging, production) are maintained to support testing and controlled releases.
- **Release and Rollback Procedures**
    - Deployment processes include version control and rollback mechanisms to minimize operational risk and ensure system stability.

4.7 Monitoring, Maintenance, and Operations

- **System Monitoring and Observability**
    - Infrastructure and applications are continuously monitored for performance, availability, and reliability using cloud-based monitoring tools.
- **Logging and Alerting**
    - Centralized logging and automated alerting systems enable rapid detection and resolution of issues.
- **Issue Tracking and Resolution**
    - Bugs, incidents, and system enhancements are tracked in centralized systems and addressed through structured development workflows.

4.8 Documentation and Knowledge Management

- **Technical Documentation**
    - Comprehensive documentation is maintained for system architecture, APIs, data pipelines, and infrastructure components.
- **User Documentation**
    - User-facing documentation, including guides and help resources, is provided to support effective use of GLOS data products and platforms.
- **Knowledge Transfer and Continuity**
    - Documentation and standardized workflows ensure continuity across internal staff and external development partners, reducing dependency on individual contributors and employees.

4.9 Governance, Compliance, and Security

- **Alignment with NOAA and IOOS Policies**
    - Software development practices align with NOAA software governance requirements and IOOS DMAC guidance, ensuring compliance with federal standards.
- **Security and Access Control**
    - Systems implement secure authentication, role-based access control, and data protection practices consistent with cloud security standards.
- **Open and Reusable Development Practices**
    - Where appropriate, GLOS promotes open-source development and reuse of software components to support collaboration across the IOOS community.

5. GLOS Quality Management Plan (QMP)

The GLOS Quality Management Plan (QMP) establishes the overarching framework through which GLOS ensures the quality, reliability, and usability of data, products, and services delivered across the Great Lakes Observing System enterprise. This plan applies system-wide, spanning observational activities, data management and cyberinfrastructure (DMAC), modeling, and user-facing applications.

GLOS operates as an end-to-end system from data collection to decision-support products and therefore adopts a holistic approach to quality management. This includes standardized quality assurance (QA) and quality control (QC) practices, structured governance, defined roles and responsibilities, and continuous improvement processes.

This QMP aligns with IOOS certification requirements, NOAA data management policies, and FAIR data principles, and is implemented in coordination with subsystem-specific procedures, including:

- Observing System - for detailed quality control procedures in data collection
- Data Management and CyberInfrastructure - for detailed data QA/QC, QARTOD implementation, and data lifecycle management
- Software Development and Delivery - for quality in software design, development, and deployment

5.1 Foundational Quality Principles

GLOS quality management is guided by the following core principles:

- **Commitment to Data Integrity and Reliability -** All data and derived products are managed to ensure they are accurate, timely, and fit for purpose. Quality considerations are embedded from data collection through dissemination and archival.
- **User-Centered Data Utility -** Data systems and products are designed to maximize usability for diverse stakeholders, including recreational users, researchers, and decision-makers. (*See Software Development section for user persona development and product design processes.)*
- **Open Data and Interoperability -** GLOS adheres to IOOS and GEOSS data sharing principles, ensuring data are openly accessible, machine-readable, and interoperable across regional, national, and global systems. *(See Data Management and Cyberinfrastructure for metadata standards and interoperability implementation.)*
- **Standards-Based and Federated Architecture -** GLOS supports a federated data ecosystem, where data providers maintain autonomy while adhering to common standards for integration, access, and reuse.

- **Transparency and Accountability -** Data quality procedures, assumptions, and limitations are documented and made accessible to users wherever possible.

5.2 Organizational Structure and Responsibilities

GLOS implements a structured governance model to ensure accountability and consistency in quality management across all activities.

- **Quality Assurance (QA) Leadership -** A designated QA Manager (or equivalent role) is responsible for maintaining the QMP, defining quality policies, and ensuring compliance with IOOS and NOAA standards. This role operates independently of direct data generation activities to maintain objectivity.
- **Data Management Lead (DMAC Oversight) -** Responsible for ensuring that all datasets entering the GLOS DMAC meet required quality standards, including metadata completeness, QA/QC compliance, and archival readiness. *(Refer to Data Management and Cyberinfrastructure for detailed QA/QC workflows and QARTOD implementation.)*
- **Observations Lead -** Oversees quality assurance for in-situ and remote sensing data collection, including adherence to calibration protocols, deployment standards, and QA/QC reporting requirements.
- **Principal Investigators (PIs) and Data Providers -** PIs and partners are responsible for implementing project-level QA/QC procedures, documenting metadata and methodologies, and ensuring that submitted data meet GLOS metadata and quality requirements.
- **Project and Product Management -** Ensures that quality considerations are integrated into product design, development prioritization, and user experience decisions. *(Refer to Software Development for agile workflows and quality checkpoints.)*
- **Executive Oversight and Review -** GLOS leadership and advisory bodies provide strategic oversight, conduct periodic reviews, and ensure alignment with organizational and stakeholder priorities.

5.3 Quality Assurance and Control Processes

5.3.1 Planning and Documentation

- **Quality Assurance Project Plans (QAPPs) / QA/QC Reports -** All projects are required to define quality objectives, performance metrics, and QA/QC procedures through standardized documentation. These documents serve as the foundation for quality evaluation and compliance.
- **Metadata Requirements -** All datasets must include standards-compliant metadata to support discovery, interpretation, and reuse. Metadata practices align with IOOS, ISO, and CF conventions. *(See Data Management and Cyberinfrastructure for metadata schema and implementation details.)*
- **Standard Operating Procedures (SOPs) -** Observational and data management activities follow established operating procedures, including instrument calibration, maintenance, and data handling protocols.

- **Contractual Quality Requirements -** QA/QC expectations are embedded within contracts and agreements with partners and vendors, ensuring accountability and consistency in data quality practices.

5.3.2 Data Quality Control Procedures

- **Automated Quality Control (Real-Time Data) -** Real-time data streams undergo automated validation, including syntax checks, range checks, spike detection, and rate-of-change analysis as per the IOOS guidelines. These processes generate standardized quality flags. *(Detailed procedures and thresholds are documented in the DMAC SOP appendix.)*
- **Manual Review and Validation -** Data streams are periodically reviewed by subject matter experts to identify anomalies not captured by automated processes.
- **Calibration and Ground Truthing -** Instruments are calibrated pre- and post-deployment, with additional validation through laboratory analysis and comparative measurements where applicable.
- **External Data Integration -** Data from federal and third-party sources are assumed to meet their respective QA standards, but GLOS performs additional validation and monitoring where feasible.

5.3.3 Data Management, Storage, and Archiving

- **Data Integrity and Storage -** Systems are designed to ensure secure, redundant storage of data, with validation checks during ingestion and processing.
- **Archival and Backup -** Data are archived in accordance with NOAA requirements, including submission to national repositories such as NCEI. Backup procedures include regular snapshots and off-site storage.
- **Data Retention Policies -** Original datasets are retained by data providers and GLOS systems for defined periods to support reproducibility and long-term analysis.

5.3.4 Data Validation and Usability

- **Multi-Level Review Process -** Data undergo multiple stages of validation from initial checks by data providers to final validation within the DMAC system.
- **Data Quality Objectives -** Data are evaluated against defined criteria including accuracy, precision, completeness, and representativeness.
- **Data Qualification and Flagging -** Data points are assigned quality flags to indicate usability and confidence levels, ensuring transparency for end users.

5.3.5 System and Infrastructure Quality

- **Infrastructure Reliability and Monitoring -** GLOS maintains a robust cloud-based infrastructure with monitoring for performance, uptime, and data flow continuity. *(See DMAC SOP for infrastructure architecture and monitoring systems.)*
- **Capacity and Scalability Management -** Systems are designed to scale with increasing data volumes and user demand, with proactive capacity planning.

- **Security and Access Control -** Infrastructure includes security measures such as firewalls and authentication systems wherever possible.
- **Software Quality Assurance -** Software systems follow structured development, testing, and deployment practices to ensure reliability and maintainability. (*Refer to the Software Development section for CI/CD pipelines, testing strategies, and release management.)*

5.4 Continuous Improvement and Assessment

- **Annual and Periodic Reviews -** GLOS conducts regular assessments of data quality, system performance, and user satisfaction, typically aligned with annual planning cycles.
- **Corrective Action Process -** Identified issues trigger a formal corrective action process, including documentation, root cause analysis, and resolution tracking.
- **User Feedback Integration -** Feedback from stakeholders including bug reports is incorporated into system improvements, ensuring timely resolution of issues and alignment with evolving user needs. *(See Engagement for feedback collection and prioritization workflows.)*
- **Training and Capacity Building -** Staff and partners receive ongoing training in QA/QC procedures, data standards, and system tools.
- **Adaptive Improvement Framework -** GLOS adopts an iterative approach to improving quality systems, incorporating new standards, technologies, and best practices over time.

---

6. Data Privacy, User Analytics and Data Sharing Governance

GLOS is committed to responsible, transparent, and standards-aligned management of all data under its stewardship, including both environmental observations and user-generated data associated with its digital platforms. This section outlines the policies, procedures, and governance mechanisms used to ensure data privacy, ethical use of analytics, and appropriate data sharing practices in alignment with NOAA data management policy, IOOS guidance, and GLOS's publicly available privacy policy.

6.1 Data Privacy and Protection

GLOS implements established data privacy policies and technical safeguards to ensure that any user-related data collected through its platforms are handled securely, transparently, and in accordance with applicable regulations and best practices.

Standard Procedures

- **Privacy Policy Governance**

    - GLOS maintains a publicly accessible privacy policy that defines how user data are collected, used, stored, and protected across all platforms, including the Seagull web application and mobile applications. This policy serves as the primary reference for users and is reviewed periodically to ensure continued compliance with evolving standards and regulations.
- **Data Minimization Principles**
    - GLOS limits the collection of user data to only what is necessary to support platform functionality, user experience improvements, and system performance monitoring. Personally identifiable information (PII) is not collected unless explicitly required for a defined feature (e.g., account registration, alerts).
- **Secure Data Storage and Access Controls**
    - User data are stored within secure cloud environments (AWS) with role-based access controls (RBAC), encryption at rest and in transit, and restricted access limited to authorized personnel. Access is monitored and periodically reviewed to ensure compliance with internal policies.
- **Authentication and User Account Security**
    - Systems requiring user authentication implement secure login mechanisms, including third-party authentication integrations where applicable. Security best practices such as tokenization, session management, and password protection are enforced.
- **Compliance with NOAA and Federal Data Policies**
    - Data handling practices align with applicable NOAA data governance frameworks, including administrative orders related to data management and software governance.

6.2 User Analytics and Data Use

GLOS uses user analytics to improve system performance, usability, and relevance of data products, while ensuring that such analytics are conducted in an ethical, aggregated, and privacy-conscious manner.

Standard Procedures

- **Purpose-Driven Analytics Collection**
    - User interaction data (e.g., page views, feature usage, session behavior) are collected to understand how users engage with GLOS platforms and to inform product development and prioritization decisions.
- **Aggregation and De-identification**
    - Analytics data are aggregated and de-identified prior to analysis to prevent attribution to individual users. Individual-level tracking is not used for decision-making or external reporting.
- **Integration with Product Development Workflows**
    - User analytics are integrated into agile development cycles, informing backlog prioritization, feature enhancements, and user experience improvements for Seagull and related applications.

- **Performance and Usage Monitoring**
  - Analytics are also used to monitor system performance, identify bottlenecks, and ensure reliability and scalability of user-facing applications.
- **Transparency in Analytics Use**
  - GLOS communicates its use of analytics through its privacy policy and ensures that users are informed about data collection practices in a clear and accessible manner.

6.3 Data Sharing and Access Governance

GLOS promotes open data sharing while respecting partner agreements, data ownership, and applicable restrictions. Data sharing practices are aligned with IOOS principles, NOAA policies, and FAIR data standards.

Standard Procedures

- **Open Data by Default**
  - Environmental observational data managed by GLOS are made publicly accessible through standards-based services (e.g., APIs, ERDDAP) whenever permissible, supporting transparency and broad reuse.
- **Alignment with NOAA Data Sharing Policies**
  - GLOS adheres to NOAA data sharing directives, ensuring that data collected with federal funding are made available in near real-time or as soon as logistically feasible.
- **Partner Data Agreements**
  - For externally sourced datasets, GLOS establishes data sharing agreements that define data ownership, usage rights, attribution requirements, and any restrictions on dissemination.
- **Data Licensing and Attribution**
  - Where applicable, datasets are accompanied by clear attribution requirements to ensure proper acknowledgment of data providers and compliance with usage terms.
- **Access Control for Restricted Data**
  - In cases where data cannot be fully open (e.g., partner-restricted datasets), GLOS implements controlled access mechanisms and honors all agreed-upon restrictions.
- **Interoperability and Standards Compliance**
  - Data are shared in interoperable formats and through standardized services to ensure compatibility with IOOS, NOAA, and global data systems.

6.4 Continuous Review and Governance

GLOS maintains an adaptive governance approach to ensure that privacy, analytics, and data sharing practices evolve with technology, policy, and user expectations.

Standard Procedures

- **Periodic Policy Review**
  - Privacy, analytics, and data sharing policies are reviewed periodically and updated as needed to reflect regulatory changes, technological advancements, and organizational priorities.
- **Cross-Functional Oversight**
  - Oversight of these practices is coordinated across DMAC, software development, and leadership teams to ensure consistency and alignment with system-wide objectives.
- **User Trust and Transparency**
  - GLOS prioritizes user trust by maintaining transparency in data practices, providing clear documentation, and ensuring accountability in how data are managed and used.

Appendix A

QARTOD Thresholds

GLOS has worked to incorporate reasonable thresholds for as accurate as possible QARTOD assessments. See the below table for incorporated thresholds. To understand how these numbers are incorporated into the QARTOD assessment, view the ioos_qc github repository at https://github.com/ioos/ioos_qc. Numbers below are in the canonical units per the CF Conventions.

How the tests work, in summary:

- Gross range: if a value falls outside the given span, the value is flagged as such
- Spike: if a value changes in magnitude from one time point to the next greater than a given magnitude, the value is flagged as such
- Rate of change: if a series of values (3 time steps) changes more than a given magnitude, the value is flagged
- Flat line: code can be interpreted a couple of ways - if a series of values does not exceed a particular tolerance range for a certain number of data points (rolling point-to-point) or across a rolling time span in seconds, the series of values is flagged as such

air_pressure (pascals):

- gross range
  - fails outside span of 50,000 - 111,000
  - suspect outside span of 95,000 - 103,000
- spike

    - suspect magnitude of 100
    - fail magnitude of 200
- rate of change
    - fails if it changes more than 1,500 over a time frame
- flat line
    - if within 10 pascals:
        - suspect in 6 data points
        - fails in 12

air_temperature (kelvin)

- gross range
    - fails outside span of 227.55 - 338.75
    - suspect outside span of 255.45 - 310.95
- spike
    - suspect magnitude of 2.2
    - fail magnitude of 3.3
- rate of change
    - fails if it changes more than 3.3 over a time frame
- flat line
    - if within 0.1 degrees
        - suspect in 2 data points
        - fails in 6

dew_point_temperature (kelvin)

- gross range
    - fails outside span of 227.55 - 338.75
    - suspect outside span of 244.25 - 310.95
- spike
    - suspect magnitude of 2.2
    - fail magnitude of 3.3
- rate of change
    - fails if it changes more than 3.3 over a time frame
- flat line
    - if within 0.1 degrees
        - suspect in 2 data points
        - fails in 6

sea_surface_temperature, sea_water_temperature (kelvin)

- gross range
    - fails outside span of 268.15 - 323.15
    - suspect outside span of 273.15 - 308.15
- spike
    - suspect magnitude of 1

- fail magnitude of 2
- rate of change
    - fails if it changes more than 1.5 over a time frame
- flat line
    - if within 0.1 degrees
        - suspect in 2 data points
        - fails in 4

relative_humidity (percent)

- gross range
    - fails outside span of 0 - 100
- spike
    - suspect magnitude of 15
    - fail magnitude of 20
- rate of change
    - fails if it changes more than 20 over a time frame
- flat line
    - if within 0.1 percent
        - suspect in 6 data points
        - fails in 12

relative_humidity (degrees clockwise from north)

- gross range
    - fails outside span of 0 - 360
- spike
    - fail magnitude of 360 (as maximum is 359.999…)
- rate of change
    - fails if it changes more than 360 over a time frame
- flat line
    - if within 0.1 degrees
        - suspect in 2 data points
        - fails in 6

wind_speed (meters per second)

- gross range
    - fails outside span of 0 - 100
    - suspect outside span of 0 - 44.704
- spike
    - suspect magnitude of 6.7056
    - fail magnitude of 8.9408
- rate of change
    - fails if it changes more than 8.9408 over a time frame

- flat line
    - if within 0.1 meters per second
        - suspect in 2 data points
        - fails in 6

wind_speed_of_gust (meters per second)

- gross range
    - fails outside span of 0 - 100
    - suspect outside span of 0 - 44.7
- spike
    - suspect magnitude of 11.2
    - fail magnitude of 17.9
- rate of change
    - fails if it changes more than 17.9 over a time frame
- flat line
    - if within 0.1 meters per second
        - suspect in 2 data points
        - fails in 6