NATIONAL GEOPHYSICAL FACILITY PROPOSAL SUMMARY

1. Executive Summary

The document summarizes a comprehensive proposal for the **NSF National Geophysical Facility (NGF)** operated by the EarthScope Consortium. It presents a vision for advancing geophysical research infrastructure in the U.S. through integrated observational networks, community engagement, and technological innovation. The main goal is to modernize and unify existing systems - such as seismic, geodetic, and geophysical monitoring to better serve Earth science, hazard mitigation, and societal resilience. The proposal emphasizes collaboration, open data, and long-term sustainability.

2. Background and Rationale

This section traces EarthScope's legacy and the success of its foundational programs like USArray, Plate Boundary Observatory (PBO), and decade-scale persistent networks. It identifies challenges that necessitate the NSF NGF:

- Fragmentation among existing facilities
- Outdated hardware and software
- Increasing demand for rapid, open-access data
- Need for cross-disciplinary, multi-hazard integration

The rationale is that NSF NGF will serve as a **unified, interoperable system,** aligning observational data with societal needs (e.g., earthquake early warning, volcanic monitoring, and climate adaptation).

3. Objectives and Mission

The NSF NGF's stated mission is to enable transformative geoscience research and support hazard response through next-generation tools and collaborations.

Key objectives include:

- Maintaining and modernizing national geophysical networks;
- Providing open, real-time data access;
- Expanding workforce training and diversity; and
- Building partnerships with federal, academic, and local agencies.

The vision statement highlights data-driven discovery and public impact.

4. Technical Components

4.1 INSTRUMENTATION

The NSF NGF instrumentation strategy emphasizes efficient operations, technological modernization, and expanded scientific capability across global, regional, and investigator-driven systems. It details organization, management, and specific plans for network operations, PI support, near-surface geophysics, and seafloor geodesy.

NATIONAL GEOPHYSICAL FACILITY PROPOSAL SUMMARY

EarthScope Instrumentation aims to maximize technical performance and efficiency using advanced engineering and team-based management.

Key goals include:

- Operating persistent networks efficiently through upgraded hardware and common platforms;
- Expanding PI instrumentation (including near-surface and seafloor assets);
- Using next-generation AI tools for real-time diagnostics and state-of-health monitoring; and
- Increasing shared use of network sites and permits for multiple sensors.

The Global Seismographic Network (GSN) operates 152 stations worldwide (43 under NSF via EarthScope, 96 via USGS). The NSF NGF addresses GSN operations by:

- Continuing modernization with broadband sensors, satellite telemetry, and low-power, high-fidelity systems;
- Implementing AI/ML-backed diagnostics to enable predictive maintenance and reduce costs;
- Maintaining collaboration with global and domestic partners; and
- Engaging internationally via FDSN, CTBTO, IUGG, and IASPEI forums to align global seismic practices.

The Network of the Americas (NOTA) with ~1,147 GNSS and 270 meteorological stations—monitors crustal deformation across the Western Hemisphere. The NSF NGF addresses regional networks, especially NOTA by:

- Upgrading to multi-constellation GNSS receivers and AI-driven real-time monitoring for improved reliability;
- Collaborating with USGS ShakeAlert for early earthquake warning;
- Co-locating of multiple sensors (meteorological, GNSS, hydrologic) for cross-disciplinary research; and
- Expanding network density and aperture through data-sharing with non-NGF GNSS stations.

NSF NGF supports PI-driven geophysical research by:

- Delivering a broad pool of portable instrumentation (seismometers, UAS sensors, MT systems, GPR, etc.);
- Incorporating new scope to include active-source seismology, seafloor geodesy, and near-surface geophysics; and
- Supporting the entire research development cycle from proposal development to field deployment to data archiving and final publications.

Near-Surface Geophysics (NSG) will be fully integrated into PI instrumentation support by:

- Combining existing SAGE/GAGE near-surface instruments into one pool;
- Conducting community workshops in years 1–2 to refine acquisition plans; and
- Building federated instrumentation partnerships to support equipment sharing.

Sea-Floor Geodesy (SFG) builds on prior NSF and USGS projects using wave gliders, GNSS-A, and pressure sensors by:

- Developing expertise during the transition from the Near-Trench Community Geodetic Experiment;
- Establishing partnerships for engineering and maintenance.

- Supporting FAIR data standards for SFG; and
- Integrating SFG education and logistics into existing NSF NGF systems.

4.2 DATA SERVICES

Data Services provides the foundation for all geophysical data management across the NSF NGF. It outlines how the facility will transition fully to cloud-native architecture, provide AI/ML-ready infrastructure, ensure open access and FAIR (Findable, Accessible, Interoperable, Reusable) data, and sustain the largest multidisciplinary geophysical data archive in the world.

Key goals include:

- Maintaining continuous, uninterrupted data service and availability;
- Fully refactoring systems to cloud-native architecture (post-migration to AWS);
- Leveraging cloud scalability for AI/ML research sandboxes; and
- Creating a dynamic management engine and team-based operational model for efficiency.

Data objects are managed via ARCO data containers, integrated data, code, and compute. Metadata standards (e.g., StationXML, miniSEED, RINEX) will be harmonized across domains to enable cross-disciplinary queries via a single API. NSF NGF data services also support partnerships with major hazard monitoring systems like USGS ShakeAlert, which now integrates real-time GNSS data via NSF NGF sensors.

NSF NGF will maintain multidisciplinary data archives, ranging from raw (Level 0) to advanced modeling results (Level 4):

- Examples of Low-level products include: GNSS time series, seismic data, radar imagery.
- Examples of High-level products include: QA/QC dashboards, AI-driven network performance tools, seismic data catalogs, processed GNSS data and hazard datasets (e.g., for USGS ShakeAlert, NOAA weather, and tsunami monitoring).

NSF NGF commits to open-source, infrastructure-as-code development, ensuring transparency and community participation realized through:

- Public GitHub repositories and internal GitLab pipelines;
- Continuous integration, testing, and versioning using Agile workflows and CI/CD pipelines.
- Curated libraries of community tools (e.g., GNSS-IR, GARPOS for seafloor geodesy);
- Containerized software and Jupyter notebooks for interactive analysis.

The NSF NGF will serve diverse users - academic, federal, and commercial through cost-effective, customizable cloud interfaces including identity management and cloud financial operations. EarthScope is adopting a zero-trust security model, ensuring all interactions are authenticated (“never trust, always verify”). Archive backups are maintained in AWS Glacier for deep storage and at NSF-funded sites like Indiana University JetStream2 for recovery. EarthScope’s NSF cybersecurity assessment rated it best-in-class among comparable research facilities.

Through strict metadata and ontological standards, NSF NGF will enable simultaneous use of heterogeneous datasets (e.g., seismic + GNSS + DAS). A queue-based cloud architecture supports real-time normalization of data streams, improving throughput and enabling AI. NSF NGF uses S3 object storage (AWS-compatible and open-source) for scalable, vendor-independent HPC integration.

Finally, the On-Ramp to the Cloud project develops large-scale, multimodal access tools and cyberinfrastructure training. Partnership with 2i2c provides an auto-scaling Jupyter interface for interactive data access, AI experimentation, and education. A planned ambassador program will link developers and researchers, promoting equity, reproducibility, and open data and code.

4.3 ENGAGEMENT

The Engagement group focuses on education, workforce development, and communication, ensuring that NSF NGF not only advances geophysical science but also broadens participation and societal impact.

The Engagement program's mission is to build an inclusive, skilled, and informed geophysics community.

Key goals include:

- Developing innovative, data-driven educational resources using facility data.
- Recruiting and preparing a next-generation workforce in geophysics.
- Engaging the public, policymakers, and scientific community in facility discoveries.
- Integrating engagement expertise across all NSF NGF operations.
- Ensuring equitable, inclusive, and safe work environments.

The Engagement team operates across three main areas: **education, workforce development, and evaluation**. A project life-cycle model (concept → pilot → deployment → dissemination → evaluation) ensures iterative improvement and accountability.

NSF NGF education initiatives aim to make geophysics accessible through real data and modern learning tools.

Core programs include:

- K–12 & Higher Education Resources: Updated web tools (Station Monitor, Velocity Viewer) and cloud-linked data simulations for classroom use.
- NSF NGF Field School (3 levels):
    - Field School 1 - 10-day introductory program for community college students and undergraduates (diverse cohorts).
    - Field School 2 - Month-long advanced experience in geophysical fieldwork.
    - Field School 3 - Specialized graduate-level and faculty training using NSF NGF instrumentation.
- Instrumentation in Higher Education: Faculty development program teaching near-surface methods (GNSS, GPR, seismic, photogrammetry).
- Online Learning and Metrics Tracking: Educators access materials via a unified authentication system with integrated analytics.

Workforce programs span the entire career pipeline from early recruitment to postdoctoral research and technical training:

- Recruitment - Outreach through media, internships, field schools, and partnerships.
- Training & Education - Skill-building workshops, online modules, and short courses.
- Retention & Upskilling - Ongoing technical training and community-building.
- Transitioning - Supporting non-geophysicists to use NSF NGF data and methods.

Geophysics communications enhance NSF NGF visibility and public engagement through:

- Comprehensive communications**:** consistent branding and messaging.
- Digital media presence**:** TikTok, YouTube, Instagram, and Facebook.
- NSF NGF website: integrating support, data access, engagement, and federated infrastructure.
- Annual NSF NGF Community Science Conference: A hybrid event showcasing user research and services, alternating between virtual and in-person.
- Movers and Shakers Program: Six-month mentorship for students, educators, and early-career scientists to co-create and share geophysics stories online.

4.4 FACILITY INNOVATION AND LIFECYCLE MANAGEMENT

The EarthScope NSF NGF proposal outlines an Innovation and Life-Cycle Management (FILM) framework - a dynamic approach that guides research, development, implementation, and retirement of technologies across all NSF NGF components. The goal is to ensure continuous modernization, efficiency, and sustainability for instrumentation, data services, and engagement systems through structured R&D, innovation rollouts, and phased recapitalization.

EarthScope defines Facility Innovation and Life-Cycle Management activities through three stages:

1. **Research & Development (R&D)**: Creation of new technologies or methods not yet implemented in geophysics.
2. **Innovation**: Phased deployment of new practices or technologies developed internally or externally.
3. **Life-Cycle Management:** Ongoing review, renovation, and retirement of operational systems.

Core FI&LM goals include:

- Anticipating and addressing emerging scientific and engineering needs.
- Integrating innovative practices into core facility operations.
- Improving operational performance through modernization.
- Applying structured and transparent management for all NSF NGF systems.

R&D initiatives are central to NSF NGF's future-readiness, focusing on AI, data federation, and predictive maintenance and include:

- AI/ML-Ready Data
  - All NSF NGF data will be cloud-native and AI-accessible. Researchers can use free or low-cost sandbox environments for testing machine learning workflows before scaling to production-level HPC systems.
- The system supports:
  - Real-time streaming and processing of high-frequency data.
  - Development of AI/ML toolchains for data interpretation and event response.
  - Explore the use of large language models (LLMs) to automatically generate outreach materials and rapid event visualizations.
- Federated Data Access
  - The NSF NGF will federate access to external data repositories using common metadata standards.

- AI for Instrument Health
  - AI and machine learning will be applied to monitor and diagnose sensor network health.
  - Integrated metadata (sensor type, power, telemetry, site images) for real-time monitoring.
  - Predictive maintenance through condition-based analytics.
  - Automated diagnostics for operational optimization.

Innovation initiatives include new technologies and engineering frameworks to be implemented during NSF NGF operations:

- New Sensor Data and Metadata
  - Support for emerging technologies like Distributed Acoustic Sensing (DAS), Distributed Temperature Sensing (DTS), and geodetic imaging;
  - Balanced edge computing, real-time streaming, and selective storage for sustainable management.
  - Standardized metadata for next-generation sensors.
  - Expanded DAS community access through federated instrumentation networks and best-practice development.
- Contributed High-Rate GNSS
  - NSF NGF will integrate thousands of non-NGF GNSS stations into its archives, creating a single federated entry point for high-rate data.
  - Uniform metadata, quality standards, and low-latency access via the cloud.
- Integrated Catalog Data Products
  - NSF NGF will host a repository with DOI minting for specialized or “orphaned” datasets (e.g., event catalogs, near-surface data, and AI-enhanced seismic catalogs).
- Common Sensor Platform (CSP)
  - Standardized, modular power and telemetry systems for all NSF NGF sensors.
- AI Chatbot for Learning
  - NSF NGF’s e-learning infrastructure will include a generative AI chatbot that supports student learning and engagement.

Life-cycle management ensures sustainability of all NSF NGF assets through systematic recapitalization, retirement, and modernization.

Activities include:

- With instrument holdings valued at $85–$90 million, NSF NGF will invest ~$4M annually to maintain capability.
- NSF NGF will sustain and modernize key open-source tools like Generic Mapping Tools (GMT) through curated Git repositories and collaborative development.
- Support FAIR-compliant open data and reproducible science.
- Continue fail-safe data backup using deep storage (AWS Glacier and university systems).
- Retirement of outdated formats and cloud optimization
- Legacy engagement platforms (e.g., geodetic velocity viewer) will be refactored for cloud compatibility and cybersecurity compliance.