ORION Network Telescope

Research Data Catalog

20 Years of Global Internet Threat Intelligence

474TB Raw Data • 76TB Processed Events • 475,000 Monitored IPs

Daily: 150GB Captures • Global: 229,000+ Networks Observed

Merit Network Research & Development

March 2026

Research Collaboration Inquiries

research@merit.edu

www.merit.edu

About the ORION Telescope

The ORION Network Telescope monitors approximately 475,000 unused IPv4 addresses, capturing all unsolicited traffic directed toward this darknet space. Since no legitimate services operate on these addresses, every packet represents scanning activity, attack attempts, misconfiguration, or other anomalous Internet behavior.

Since October 2005, we have continuously collected one of the world's largest datasets of Internet threat intelligence - over 474 terabytes of compressed packet captures spanning 20 years. This unique vantage point enables researchers to study global cyber threats, track attack evolution, and develop next-generation security solutions.

What Makes ORION Data Unique

- **Unprecedented Scale:** 474TB raw data + 76TB processed events over 20 years
- **Continuous Operation:** Uninterrupted collection since 2005 with no data gaps
- **Global Coverage:** Traffic from 229,000+ unique /24 networks worldwide
- **Rich Context:** GeoIP, ASN enrichment, automated event detection, DNS data
- **Multiple Formats:** Raw PCAPs for deep analysis, structured JSON for ML/analytics

Available Datasets

Dataset	Size	Time Period	Format
IPv4 Telescope	474TB	Oct 2005 - Present	PCAP (hourly, gzip)
Processed Events	76TB	2005 - Present	JSON (enriched)
IPv6 Telescope	944MB	2016 - Present	PCAP (gzip)
HoneyTrap	375GB	2022 - Present	Active honeypot data

Dataset Details

IPv4 Network Telescope

Raw packet captures from approximately 475,000 monitored IPv4 addresses distributed across multiple subnets (/15, /16, /17, /19 CIDR blocks). Captures all protocols with complete packet headers and payloads when available. Current daily collection averages 150GB compressed, representing billions of packets.

Data includes: Microsecond timestamps, full 5-tuple (src/dst IP/port, protocol), TTL, TCP flags, packet sizes, complete payloads

Typical uses: DDoS backscatter analysis, Internet-wide scanning studies, vulnerability exploitation tracking, longitudinal security research, protocol analysis

Processed Events Dataset

Structured JSON events automatically derived from raw telescope data through our processing pipeline. Each event represents an identified scanning campaign, DDoS backscatter episode, or other significant traffic pattern with rich contextual enrichment.

Enrichment includes: GeoIP country/region, ASN attribution, DNS reverse resolution, behavioral classification, confidence scoring, temporal aggregation

Typical uses: Machine learning model training, threat intelligence feeds, statistical analysis, temporal pattern studies, attack attribution research

IPv6 Telescope & HoneyTrap

IPv6 Telescope: Monitors 2001:48a8:8000::/33 address space, providing crucial insights into IPv6 adoption patterns, emerging IPv6-specific scanning techniques, and nextgeneration threat landscape evolution.

HoneyTrap: Active honeypot system that responds to connection attempts, capturing complete attack payloads, malware samples, exploitation sequences, and post-compromise command execution for deeper threat analysis.

Research Applications

DDoS Attack Studies

- Backscatter analysis for attack volume estimation and victim identification
- Attack vector characterization and amplification technique detection
- Temporal pattern analysis and campaign correlation across time
- Geographic distribution mapping of attack infrastructure

Internet Scanning Research

- Large-scale scanning campaign identification and tracking
- Scan methodology analysis and fingerprinting techniques
- Botnet command and control infrastructure detection
- Targeted vs. opportunistic scanning behavior classification

Vulnerability & Exploit Analysis

- Early detection of vulnerability exploitation in the wild
- Time-to-exploitation measurements for disclosed CVEs
- Zero-day activity detection through anomaly analysis
- Exploit delivery mechanism and payload analysis

Machine Learning & AI

- Threat detection model training with labeled ground truth
- Anomaly detection algorithm development and validation
- Traffic classification and attribution using deep learning
- Predictive modeling of attack campaigns and threat evolution

Longitudinal Studies

- 20-year historical trend analysis of Internet threats
- Protocol adoption patterns and technology shift analysis
- Threat landscape evolution tracking across decades
- Economic and geopolitical event correlation with cyber activity

Future Plans: The National Distributed Network Telescope (NDNT)
The next evolution of ORION's twenty-year dataset is the National Distributed Network Telescope (NDNT)—a nationwide collaborative effort led by Merit Network to extend darknet monitoring across diverse Internet regions through partnerships with research and education (R&E) networks and universities throughout the United States.

NDNT builds upon ORION's foundation by deploying distributed telescope nodes across participating institutions, each contributing unused blocks of IPv4 address space to expand the scope and diversity of monitored networks. This distributed architecture will dramatically increase the visibility of global Internet activity, enabling fine-grained regional analysis of scanning behavior, threat propagation, and emergent attack vectors.

With more than 40 partner institutions already engaged, NDNT will exponentially grow the available dataset from hundreds of terabytes to multiple petabytes over the coming years. Each new node will contribute unique vantage points, enriching the collective dataset with broader geographic and topological coverage. The integration of diverse sensors and enrichment pipelines will enhance data fidelity, reduce bias from single-location observation, and support federated analytics across multiple research domains.

Future enhancements will include:

- **Standardized Data Pipelines** to normalize and aggregate traffic from all telescope nodes into a unified data lake with consistent metadata, timestamping, and enrichment fields.
- **National Data Sharing Integration** with partner platforms such as the Open Science Grid (OSG) and the National Research Platform (NRP) to provide secure, federated access for researchers nationwide.
- **User-Friendly Access Tools**, including a graphical interface for dataset exploration, API-driven query endpoints, and integration with data science curricula to support education and training.
- **Long-Term Research Roadmap** (5–7 Years) focused on expanding IPv6 visibility, incorporating active deception systems (honeynets and honeypots), and developing AI-assisted traffic classification and visualization.

Together, these efforts will transform ORION from a single, Michigan-based vantage point into a national-scale, continuously growing resource for cybersecurity research, policy development, and STEM education—providing an unprecedented longitudinal view of unsolicited Internet activity and the evolution of global threats.

How to Access ORION Data

Data Use Agreement Required

All access requires execution of a formal Data Use Agreement specifying research purpose, data handling procedures, security requirements, privacy obligations, and publication attribution guidelines. We work with researchers to ensure compliance with institutional policies and ethical research standards.

Access Modalities

Secure Research Environment

Isolated virtual machines with direct access to both raw and processed datasets. Includes computational resources appropriate for large-scale analysis, development environments, and standard analysis tools.

Query Interface

SQL and API access to processed events dataset for programmatic queries, enabling integration with external analysis pipelines and automated research workflows without direct infrastructure access.

Data Export

Curated dataset subsets for offline analysis, subject to capacity constraints and specific research requirements. Ideal for researchers needing portable datasets for local analysis or integration with proprietary tools.

Request Process

Step 1: Submit Research Proposal

Contact research@merit.edu with a brief description of your research objectives, methodology, expected duration, and data requirements. Include information about your institution and any relevant publications or projects.

Step 2: Review and Execute Agreement

Work with our team to review the Data Use Agreement and any institutional requirements. We accommodate various institutional approval processes and can coordinate with your legal/compliance teams.

Step 3: Receive Access and Support

Upon agreement execution, receive access credentials, technical documentation, and onboarding support. Ongoing technical assistance available throughout your research.

Contact Information

Research Data Access Requests research@merit.edu

Research Leadership

Pierrette Renée Dagg, Ph.D.

Director of Research, Merit Network prdt@merit.edu

About Merit Network

Merit Network is a nonprofit internet service provider dedicated to serving Michigan's education and research community. For over 50 years, Merit has operated critical internet infrastructure and conducted pioneering research in networking, cybersecurity, and internet measurement.

Our mission combines operational excellence with research innovation, providing both essential connectivity services and unique research datasets that advance the field of cybersecurity.

www.merit.edu