

PHEMI Central Big Data Warehouse

PHEMI Central™ is a production-ready big data warehouse with built-in privacy, data sharing, and data governance. PHEMI Central delivers the scalability and economics of Hadoop with indexing, cataloging, fine-grained access control, and enterprise-grade data management.

Economically integrate all of your data.

The PHEMI Central Big Data Warehouse, built on Hadoop, unlocks your data silos and makes all of your data available for analytic and operational applications. Big data technology allows you to scale to petabytes of data with cluster economics. PHEMI Central adds simplified deployment and out-of-the-box operations and the ability to integrate immediately with existing data sources and analytics tools. PHEMI's solution simplifies every step, from installation to getting your end users up and running.

Index, catalog, and enrich your data for findability and performance.

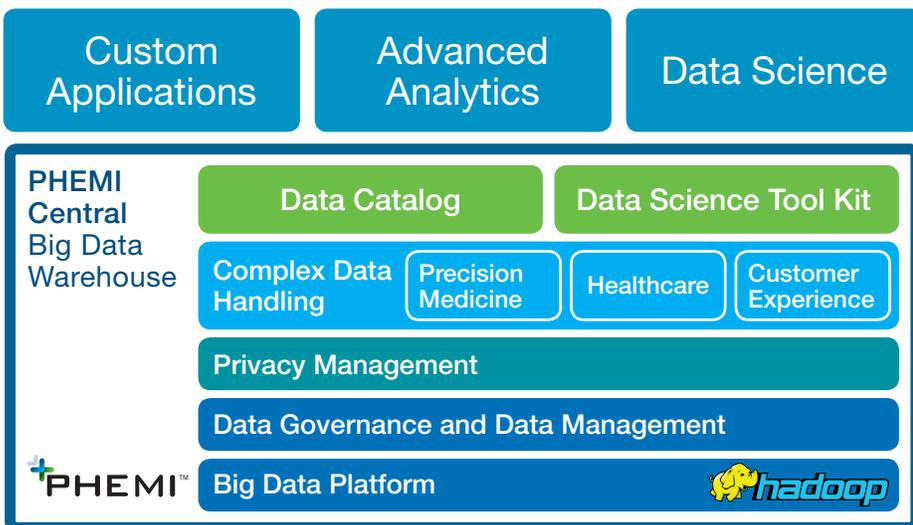
PHEMI Central intelligently curates your data from the moment it arrives. Data is indexed and cataloged on ingest for sub-second lookups. PHEMI Central can cleanse, standardize, encrypt, or otherwise process selected information. Data Processing Functions—programs written in common programming languages that operate right in the datastore—allow you to transform data at any point in time, without having to manage the complexities of distributed computing.

The PHEMI Data Catalog adds the ability to curate virtual datasets for controlled, self-serve, point-and-click access by end users.

Protect and share data with privacy, security, and governance.

PHEMI's privacy, data sharing, and data governance features let you automatically protect private and sensitive data, while being able to share it with collaborators. PHEMI Central controls access to information at the field level, ensuring rightful access at scale. PHEMI Central can anonymize, hide, or redact data for different user authorizations. Access control based on user attributes and metadata enforces rightful access, and advanced data management features like version control, lifecycle management, and data sharing agreement enforcement help you achieve compliance and governance objectives.

PHEMI CENTRAL BIG DATA WAREHOUSE



PHEMI Central leverages three unique innovations

Metadata Framework

Extensible, descriptive, end-to-end metadata enables field-level access control and data management.

DPF Framework

PHEMI Central lets you develop custom programs, called Data Processing Functions (DPFs), that can parse, recognize, extract, cleanse, standardize, encrypt, mask, or otherwise transform selected fields. The complexities of distributed computing are handled automatically by PHEMI's DPF Framework.

Policy Enforcement Engine

Every user request for data is filtered through a rules engine based on metadata and user attributes. PHEMI's privacy and security framework ensures that data remains governance-compliant at all times.

Ease your entry into big data and grow your solution at your own pace.

Easy deployment

PHEMI Central can be deployed at the customer premise, as a managed service, or as a cloud-based service. Unlike Do-It-Yourself solutions, PHEMI Central is a fully integrated product, enterprise-grade and ready to deploy.

PHEMI Central integrates with your existing applications, ETL tools, analytics tools, and existing databases or Enterprise Data Warehouse. You can preserve your investment in existing infrastructure, easing your way into production and expanding incrementally.

Cluster reliability and economics

The system uses low-cost commodity hardware components and Direct Attached Storage (DAS) disk drives to lower the cost of ownership compared to traditional Enterprise Data Warehouse systems. Storage and compute resources scale linearly from terabytes to petabytes.

All data in the system is replicated three times to ensure availability and resiliency. DAS drives can be hot-swapped without impacting performance or data availability. Larger or faster DAS drives and nodes are absorbed into the system and load-balanced automatically.

Sophisticated system and data management

The system provides clear visibility into system health, diagnostics, troubleshooting, capacity, and digital assets under management. System management capabilities can be integrated with existing tools.

PHEMI Central keeps the system secure with user roles that limit what operations a user can perform. Communication links from data sources or to consuming systems can be encrypted using Secure Sockets Layer (SSL) or Transport Layer Security (TLS). The system maintains a complete, tamperproof audit log of operations and data access.

PHEMI Central also provides system and data management capabilities formerly only available in traditional data warehouses. Data retention policies and data sharing agreements are automatically enforced. Data in the system is immutable: the original data cannot be modified and data is only purged from the system based on the configured retention policy. Robust version control and rollback capabilities mean that data is never lost, corrupted, or overwritten.

Visit www.phemi.com for more information.

SPECIFICATIONS	
On-Premise Deployment*	Cloud Deployment*
<p>4 Data Nodes. Each:</p> <ul style="list-style-type: none"> • 8xCore (2.2GHz) • 64 GB RAM • 16 TB Direct Attached Storage <p>3 Master Nodes. Each:</p> <ul style="list-style-type: none"> • 4xCore (2.2 GHz) • 64 GB RAM • 300 GB RAID1 Storage • 1 TB Storage <p>1 Management Node:</p> <ul style="list-style-type: none"> • 4xCore (2.2 GHz) • 64 GB RAM • 300 GB RAID1 Storage • 1 TB Storage <p>10 Gigabit Ethernet Network</p>	<ul style="list-style-type: none"> • Subscribe to PHEMI Central as a managed service running on Amazon Web Services or ClearData for secure healthcare in the cloud. • Cloud service grows from 1 TB storage capacity.
Data Ingest Protocols	Data Export Protocols
<ul style="list-style-type: none"> • SFTP File Transfer • HTTP/HTTPS Manual Upload • REST Web Services API • ODBC/JDBC SQL Interface 	<ul style="list-style-type: none"> • Excel/CSV/TSV Download • REST Web Services API • ODBC/JDBC SQL Interface
Analytics Tools	Data Processing Functions
<p>Supports leading analytics tools, including:</p> <ul style="list-style-type: none"> • R • SAP • SAS • SPSS • Stata • Tableau • Qlikview • Netezza • MySQL 	<ul style="list-style-type: none"> • Excel Reader • CSV Reader • Variant Call Format (VCF) Reader • JSON Reader • XML Reader • Custom DPFs



* All our deployments align with appropriate privacy and security requirements, including Health Insurance Portability and Accountability Act (HIPAA) and Health Information Technology for Economic and Clinical Health (HITECH) Act as well as Canadian federal and provincial legislation.