

## Data Privacy Manager for Azure

The potential of data in healthcare is to slow the ever-increasing costs of care, help providers practice more effective precision medicine, and support preventive self-care. However, the biggest challenges in using healthcare data are privacy and security, connecting the lakes of existing data while collecting and managing new kinds of data.

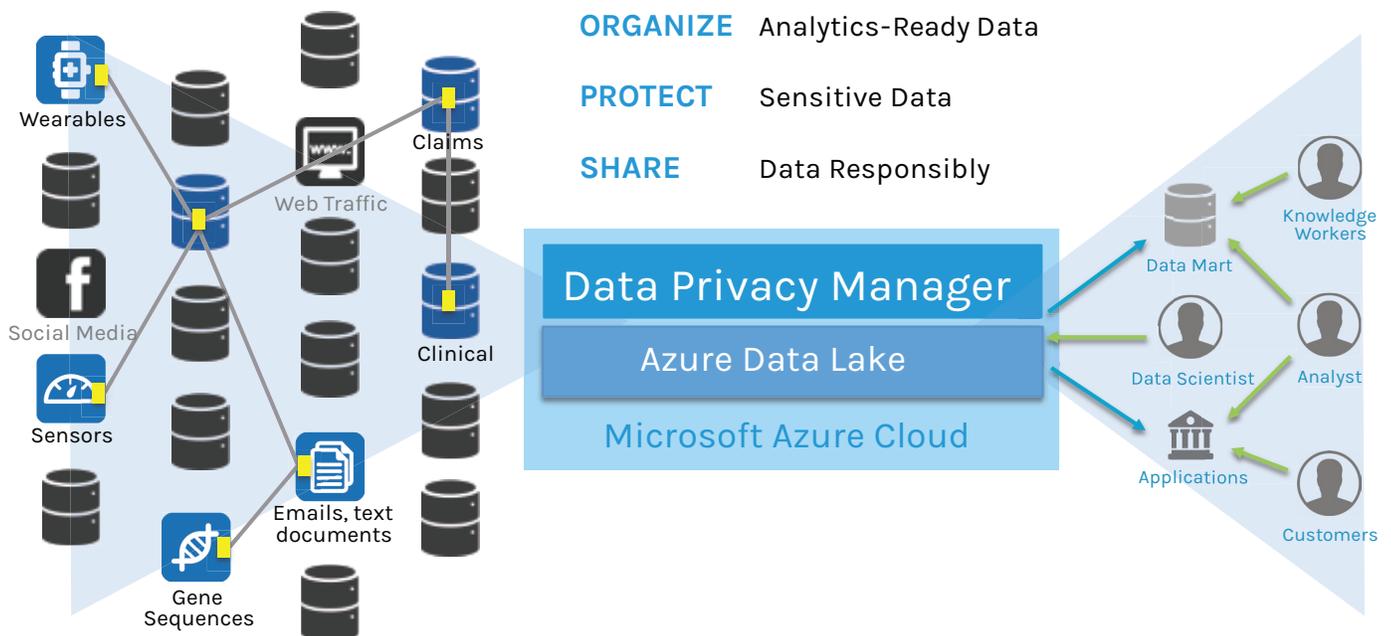
Conventional data protection products simply lock down your data. The sophisticated PHEMI Data Privacy Manager alleviates privacy concerns, making it possible to easily share data responsibly and unlock new and actionable insights. PHEMI Data Privacy Manager is purpose-built for healthcare, providing the highest standards of privacy, security, and governance because it is built on a zero-trust architecture and privacy by design principles.

### Unlock Value From Your Data

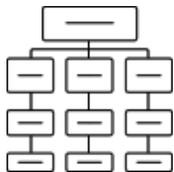
PHEMI Data Privacy Manager simplifies the collection, preparation, and organization of complex, multi-structured data. It manages the full data lifecycle from acquisition and cataloging, through to transformation into analytics-ready data sets. High-performance data pipelines and automatic data governance capabilities mean that analysts spend their time on discovery, not data preparation.

### Share Data Responsibly

- Reduce the complex integration and security effort of open-source software.
- Tag, inventory, de-identify, tokenize, audit and track provenance of your sensitive data.
- Index, catalog, and model complex data enabling advanced analytics, responsive queries, agile data models and self-serve access to analytics-ready data.
- Uniformly and consistently enforce policy-based access rules at scale to control who is allowed to see what data.
- Take control of your redundant, obsolete and trivial data.



## Data Privacy Manager for Azure



### ORGANIZE Analytics-Ready Data

Organize your data lake and inventory your sensitive data so you can avoid getting trapped in a data swamp. Integrate all your complex data at scale and speed so that you have the agility to respond to changing business demands.

### Data Processing

PHEMI was made for complex data.

Whether you're looking for Natural Language Processing, Machine Learning, Predictive Modeling, or Artificial Intelligence, PHEMI can help you pre-process and extract insights from your unstructured data so you can curate a library of clean, analytics-ready data.

### Indexing

PHEMI lets you link, query, build cohorts and analyze data at speed and scale.

Index any column and build data models for fast lookup within your policy-based privacy, security and governance requirements, so people can only see the sensitive data you allow them to see.

### Cataloging

PHEMI dictionaries allow you to quickly describe and search the data under management so you and your users understand exactly what you are working with and where it came from.



### PROTECT Sensitive Data

Secure your data and let users, applications and visualization tools access data within your compliance framework, scaling across sources, users, and organizational boundaries.

### Control

PHEMI lets you tag, inventory and control your sensitive data whether it's in a file or buried in a relational table.

It helps you enforce data usage agreements across departments, business units, partners and licensed third party content.

### De-Identify

PHEMI lets users work with data without breaching privacy guidelines.

It includes a comprehensive library of de-identification algorithms to hide, mask, truncate, group, and perturb data.

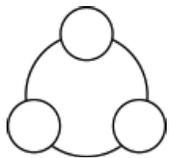
System functions include dataset-specific and system-wide tokenization so you can link and share data without risking data leakage.

### Monitor

PHEMI collects audit logs concerning not only what changes were made to the system, but also data access patterns.

Human and machine-readable de-identification reports are automatically generated to meet your enterprise governance risk and compliance guidelines.

## Data Privacy Manager for Azure



### SHARE Data Responsibly

Scalable data protection solutions need to go beyond encryption and de-identification to allow businesses to responsibly share sensitive data. With PHEMI Data Privacy Manager, you can set policy rules that uniformly and consistently enforce access for users and applications in real-time.

### Selective Access Control

Attribute-Based Access Controls let you easily template and assign access rules to ensure compliance at scale.

Based on user and application authorizations, the system automatically selects the non-materialized view a user, application or visualization tool is allowed to see.

One dataset, one SQL query, and one analytics dashboard can let: Doctors see PHI data but not Payment Card Industry (PCI) or Secret data; Finance see PCI data but not PHI data; European users see Personally Identifiable Information (PII) data, but restrict users in the rest of the world to only see de-identified data.

### Risk-Based De-Identify

Your responsibility to protect data extends well beyond your data center footprint.

The risk of re-identification or unintended use of data is a consistent problem for organizations working with sensitive data.

Data Privacy Manager lets you quantify and control the risk of re-identification.

It lets you go further in controlling data movement and access based on user permissions, data sensitivity, and dataset location, reducing the risk of a data breach.

### Access Data On Your Own Terms

PHEMI lets you access data through a variety of standard protocols and tools.

Whether you need an ODBC Driver to your existing analytics tool; Spark-based data science tools such as Spark MLlib, H2O Sparkling Water, Apache Zeppelin or Cloudera Data Science Workbench; the PHEMI REST API to a custom application; or even exported data to downstream targets.

All access to data is fully audited and governed by Attribute-Based Access Controls.

- Ingest big data in minutes, not hours
- Build datasets in minutes, not months
- Look up data in a fraction of a second
- Cut data sprawl in half
- Share datasets with privacy, security & governance at scale



## Data-Driven Insights

Be agile, and keep your competitive edge by reducing data friction.

Let PHEMI be your single source of truth and trust so you can accelerate the secondary use of data without compromising privacy, security, and governance.



## Optimize the Customer Journey

Streamline how customers interact with your company, optimizing customer pathways to reduce churn, improve customer satisfaction and lower call center volumes without compromising privacy, security, and governance.



## Extract Insights Using Text Analytics

Discover new insights from your emails, social media, surveys and documents including keywords, topics, sentiment and semantic concepts, de-identifying data and linking it to existing datasets.



## Data Library with Privacy, Security & Governance

Ensure everyone is working from a single version of the truth; and reduce the time to access data from months to minutes, without restricting rightful access to sensitive data.



## Harness Precision Medicine Data

Organize, index and link high volumes of genotype/phenotype data for interactive lookups, machine learning and predictive modeling with full privacy, security, and governance.



## Manage Big Data Across Departments

Manage cross-organizational and cross-department data sharing agreements, controlling appropriate data usage and managing the data preparation workflow while preserving data lineage.

## Data Privacy Manager Specifications

### Advanced Analytics

- Apache Zeppelin
- H2O Sparkling Water
- ODBC SQL Driver
- REST Web Services API
- Spark MLlib
- Spark SQL

### Data Ingest

- Avro
- Azure
- CSV, TSV, Excel
- FTP/SFTP
- HDFS
- HL7
- JSON
- Kafka
- NFS
- REST
- SMTP/POP3
- SNMP
- Splunk
- SQL
- Syslog
- Twitter
- VCF (Genomic)
- XML

### Data Export

- Avro
- Azure
- Cassandra
- Couchbase
- CSV
- DynamoDB
- FTP
- HBase
- Hive
- Kudu
- Mongo
- Parquet
- REST
- SMTP
- SQL

### Data Indexing and Modeling

- Custom indexes
- Data Lineage
- Data Dictionary
- Geospatial
- Graph
- Metadata management
- Tabular
- Time Series
- Workflow orchestration

### Privacy and Security

- Attribute-Based Access Controls
- Audit logs, Compliance reports
- Data sensitivity tagging
- De-identification/tokenization algorithms
- Encryption at rest and in motion
- Kerberos, KMS, KTS
- Retention policy enforcement
- Risk-Based De-Identification & filtering
- Version control
- Zones of trust