

Why [tile]DB Carrara for Biopharma

Accelerate breakthroughs with all your data

[tile]DB
DESIGNED FOR DISCOVERY™

- ✓ Organize
- ✓ Structure
- ✓ Collaborate
- ✓ Analyze

At a Glance

Multimodal data, including large-scale, high dimensional multiomics data, is the new frontier. This is where biopharma is hunting for new targets to expand treatment capabilities and improve patient outcomes through earlier disease detection and diagnosis. TileDB Carrara is a foundational data platform designed for discovery. Unlike traditional tabular data warehouses, Carrara's innovative array-based architecture manages both bleeding-edge frontier data and common research data with equal efficiency. TileDB's unified approach adopts a FAIR data approach, eliminates data silos and dramatically reduces infrastructure complexity, equipping research, data and business teams to Organize, Structure, Collaborate on, and Analyze complex biological datasets seamlessly. For biopharma organizations, this means faster discovery, lower infrastructure and people costs, and simplified compliance with industry regulations.

Technical Features

ORGANIZE



- TileDB Catalog enables folder organization with rich metadata & descriptions
- TileDB Filesystem provides direct cloud access without local downloads
- Efficient search functionality across all data and code
- Preview widgets to verify data before use
- Single point of access control and logging

COLLABORATE



- Compliant with SOC 2, HIPAA, and GDPR regulations
- Teamspaces for secure group collaboration
- Granular access policies and comprehensive activity logging
- Views feature allows sharing without physical data duplication
- Functions as a Trusted Research Environment

STRUCTURE



- Array-based architecture transforms complex frontier data into efficient formats
- Specialized vertical offerings:
 - TileDB-SOMA for single-cell analysis
 - TileDB-VCF for population genomics
 - TileDB-Biolmg for biomedical images
- New TileDB Tables feature powered by Apache DataFusion

ANALYZE



- Integrate with popular tools such as Seurat, Scanpy, Bioconductor
- TileDB Workflows for running NextFlow scripts
- Integrated Jupyter notebooks built into the system
- Support for multiple programming languages
- Native distributed computing platform
- Serverless user-defined functions
- Low/no-code dashboard development & deployment
- AI/ML functionality including native support for ML models, built-in vector search, & LLM integrations

Transform your object store into a Trusted Research Environment

Organize

all files into a catalog with metadata; access directly from cloud as-if local.

Structure

complex frontier data into efficient formats for performant downstream analysis.

Collaborate

within a Trusted Research Environment, run federated queries with compliance readiness.

Analyze

common & frontier data with equal efficiency using custom-built multiomics & imaging solutions.

[tile]DB Carrara | The Database Designed for Discovery

AWS S3 Storage

TYPES OF DATA

Sample metadata

Gene models

Compound data

Clinical trial data

Gene-disease association datasets

Experimental metadata

Prevalence & incidence databases

Disease ontologies

Phenotype ontologies

Claim records

Research papers

Audio/video

Imaging data

Bulk/single-cell/spatial tx

Variant data/metadata

Business Benefits

Accelerated Drug Discovery Pipeline

Our unified platform streamlines the entire data pipeline from generation to insight while breaking down silos and enabling secure team collaboration.

Reduced Operational Costs

By consolidating multiple tools into a single platform with efficient data structuring, we minimize both software licensing costs and the need for large engineering teams.

Enhanced Compliance and Security

The platform offers comprehensive security through built-in compliance with SOC 2 Type 2, HIPAA, and GDPR requirements, along with robust audit logging and controlled collaboration features.

Improved Research Team Productivity

Researchers benefit from intuitive data organization, seamless tool integration, and reduced data preparation time through streamlined workflows.

Technical Benefits

Advanced Data Management Capabilities

Our array-based architecture efficiently handles complex biological data types with native support for genomics, transcriptomics, and imaging data, including specialized features for single-cell analysis and population genomics.

Powerful Analysis Infrastructure

The platform provides a built-in distributed computing environment with native AI support and flexible API integration for comprehensive research analysis.

Optimized Storage and Performance

Our cloud-native architecture delivers efficient file operations and advanced metadata management without requiring local downloads.

AI-Ready Data Platform

Enable trusted and accelerated model development through streamlined data governance, improved access to data, and enhanced security.

FEATURES	Status Quo	[tile]DB Carrara
Data Model	Primarily tabular data in traditional databases, or unstructured files in data lakes	Universal array-based architecture that efficiently handles both tabular and frontier data (multiomics, imaging, etc.)
Infrastructure Requirements	Multiple systems needed (databases, data lakes, compute clusters)	Single unified platform with built-in compute capabilities
Data Access	Often requires data downloads or copies between systems	Direct cloud access via S3 without downloads, with virtual file system capabilities
Performance Scaling	Limited by traditional database architectures or file system constraints	Native distributed computing with serverless architecture
Format Support	Usually requires data conversion between formats	Native support for industry-specific formats (VCF, SOMA, etc.)
Collaboration Model	Typically involves creating copies or managing complex permissions	Virtual views and tablespaces without data duplication
Compliance & Security	Often requires additional security layers and auditing tools	Built-in compliance (SOC 2, HIPAA, GDPR) with comprehensive audit logging

Customer Case Studies



Build an enterprise-wide multiomics data mesh to drive AI initiatives.



Diagnose critical rare diseases within hours in the NICU.



Power next generation drug discovery using single-cell data at scale.



Build a FAIR data platform for multiomics data to identify new targets.

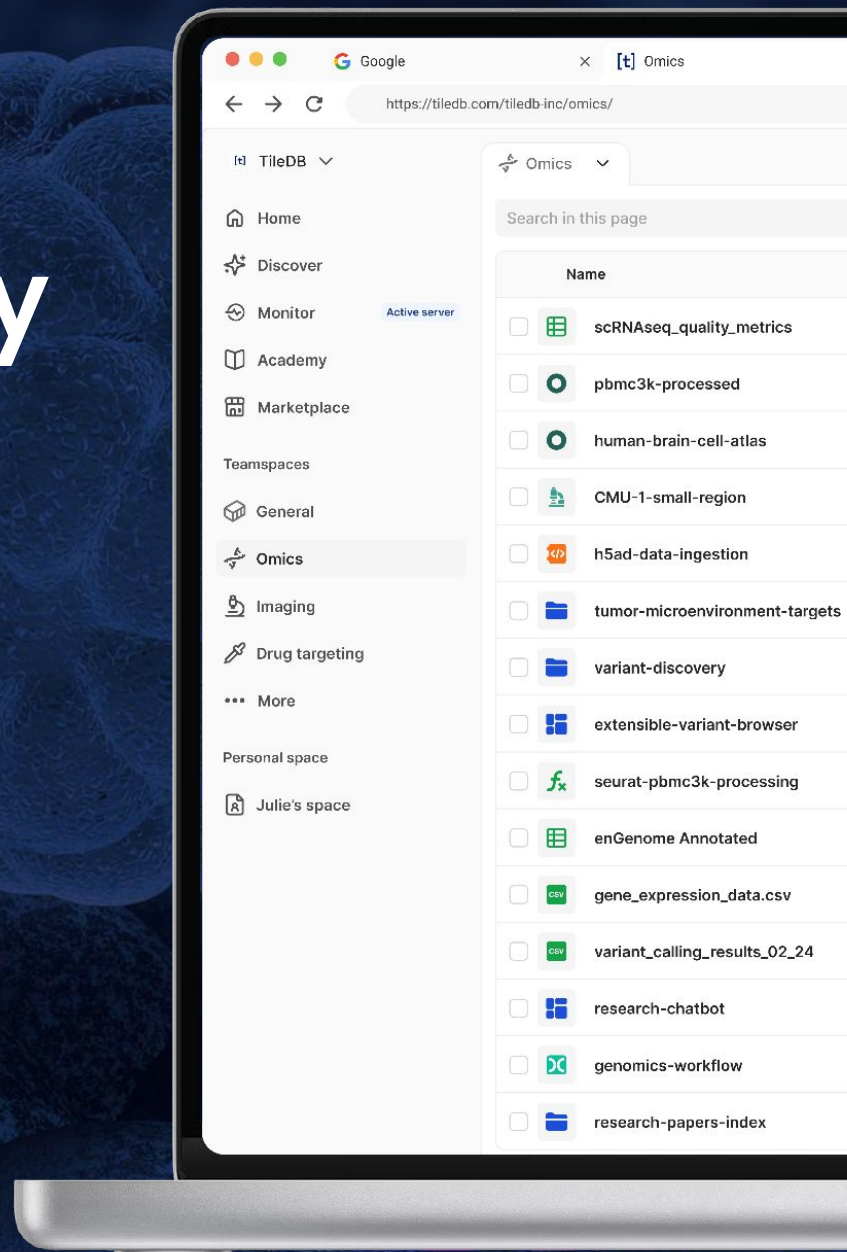
Discover detailed customer case studies on our website at tiledb.com

[tile]DB Carrara

Designed for Discovery

TileDB Carrara modernizes biopharma data management by unifying research data and analytics in a secure, compliant FAIR platform. Its scalable array-based architecture handles both common and frontier data, accelerating drug discovery while reducing costs and complexity.

tiledb.com →



Trusted Customers & Partners



In partnership with:

