

Features and Capabilities

DNAnexus[®] is helping hundreds of biomedical organizations accelerate scientific discovery. With DNAnexus, precision health innovators are able to manage, analyze, and collaborate on multimodal data to harness its full potential and unlock insights that improve patient care. A listing of features and capabilities of the platform is below.

Compute and Storage

Tailored Instance Types	Customize your experience with tailored instance types. DNA nexus instance classes and types provide optimal ratios of memory, CPU and high-speed storage for bioinformatic workloads.
Spark Cluster Support	Leverage Apache Spark™ for enhanced performance — no expertise required. Easily run robust, distributed, and efficiently parallelized computational jobs by launching and managing large scale Apache Spark clusters on-demand.
Scalability	Efficiently manage compute resources by using only what is needed. Scale your environment to meet your data demands with a high-performance NGS analysis platform built to support thousands of nodes and millions of objects. Minimize delays and accelerate turnaround time with exceptional uptime and powerful compute capacity, including parallelizable execution of data analysis.
Management and Traceability	Keep track of data provenance. DNA nexus file management tools enable users to easily upload, organize, search, add metadata, and view a file’s source.
Developer Experience and API and SDK	Develop on DNAnexus using your favorite language. DNA nexus diagnostics toolkit can be used to create apps and debug with ease. Open APIs and SDKs provide the ability to upload and store various file types with the flexibility to choose command line or web interface.
Spend Management	Stay on budget. With DNA nexus , administrators or project owners can set, monitor, and notify users on spending budgets.
Use Your Own Account	Use your AWS & Azure credits. DNA nexus is on Amazon Marketplace and Azure Marketplace so if you have cloud provider issued credits they can be consumed via this mechanism.
Usage Reporting	Understand your spending. Drill down into detailed monthly reports on compute and storage spend to understand where you can further optimize.
Data Archiving	Save money by storing infrequently used files in less expensive storage on-demand. Archive and unarchive files easily. You can search and view metadata even when items are in an archived state.

Regions and Clouds

Global Regions	Access and analyze data globally. Deploy and manage data and analysis in regions across the world with a single user interface.
Adherence to Regional Regulations	Ensure global regulatory compliance. Data and analyses are managed per region ensuring that data stay within regional borders and adhere to region-specific regulations.
Multi-cloud	Experience multi-cloud support. Deploy the same tools and workflows on AWS, Oracle OCI, and Microsoft Azure® regions.
Cloud Service Integrations	Seamlessly integrate applications. Securely leverage cloud services from inside DNAnexus to enable a wide variety of use cases.

Data Management and Query

Structured Dataset Organization	Integrate your phenotypic data with your massive multiomics datasets. Organize genomic, transcriptomic, clinical, imaging and real world data into datasets with refined platform-level control for data access, analysis for multiple data types.
Custom and Complex Dataset Queries	Investigate datasets for deeper understanding. Further examine organized datasets using SQL to construct custom and complex queries. Access datasets from the command line.
Powerful Cohort Browser	Easily handle the scale and complexity of clinico-omics datasets. Browse and explore millions of phenotype fields and trillions of omics. Seamlessly build, combine, and compare multimodal cohorts, and carry out complex analyses - no coding required.
Schema Agnostic Pheno-data Loading and Management	Prepare and organize clinical, medical, and phenotypic data. Ensuring proper data structure is a common challenge when integrating and analyzing data. DNAnexus ingests data into the database and creates a dataset representing data, dictionaries, and codings provided for use with dataset based queries and analyses.
Genomic Annotations	When genomic information is ingested variants can be annotated using industry standards such as NCBI's dbSNP, Ensembl's gnomAD, and SnpE . Variant annotation information can be used in query filtering, or for knowledge exploration.
Structured Somatic Data Store	Explore and analyze somatic data. Store somatic variant data, such as SNVs, CNVs, INDELS, and fusions for exploration or tumor-normal comparisons. Link clinical information to create a clinicogenomics dataset, define cohorts.
Structured Germline Data Store	Explore and analyze germlined data. Store germline variant data such as genotypes and alleles to use for cohort creation and analysis.
Gene Expression Exploration	Translate genes expression into actionable insights. Use tools that facilitate scalable filtering and visualization of bulk gene expression data, enabling a deeper understanding of disease biology.

Security and Compliance

Compliance	Keep your data secure and reduce risk. DNAnexus is compliant with the industry’s strictest security & privacy regulations – including ISO27001, NIST 800 framework, HIPAA, PHI-enabled, GDPR, GxP/ICH regulations, CAP/CLIA, FedRAMP Moderate – and hosts data in your preferred geography.
Organization Administrative Tools	Manage your team’s DNAnexus activity. Leverage administrative tools for continuous visibility and easy management of your team’s DNAnexus organization-level account. Assign admins, update policies, add/remove members, view compute and billing information, and more all in one place.
Data Sharing, Collaboration, Access Control	Collaborate securely and compliantly, across hallways and borders. DNAnexus provides a centralized, secure environment to manage data and pipeline access on the file, project, and user level. Publish your own tools and pipelines to share them with collaborators.
Audit Trail	De-risk your audits. Access a human and machine-readable daily log of all activities (login, upload/download, run analysis, data sharing, etc.) related to users and projects of the organization for seamless auditability and reproducibility.

Data Analysis

Integrated JupyterLab	Directly access projects from JupyterLab. The DNAnexus platform includes embedded JupyterLab, allowing for enhanced collaboration and direct tracking of work in a secure environment. Support for Python, R, and Bash.
Prebuilt JupyterLab Environments	Domain specific pre-installed packages to support multipleomics use cases. Built with data analysts gold standards in mind, the DNAnexus platform provides multiple JupyterLab environments for statistical analysis, machine learning/AI, association studies, and image processing.
Comprehensive Tools Library	Leverage out-of-the-box apps supported by DNAnexus and partners. The DNAnexus Tools Library has many ready-to-use industry standard pipelines, e.g. GWAS algorithms.
Data Visualization	Explore omics and gene expression data with innovative visual applications. Data can rapidly be explored and visualized through a genome browser or your own visualization tools.
Custom Web Apps Support	Build custom web-apps interfaces. Deploy custom web applications within the DNAnexus security boundary. DNAnexus supports web app technologies including R Shiny and NodeJS.

Workflows

Support for Multiple Workflow Languages	Easily move workflows and apps to the cloud. The platform supports industry-standards Docker, Nextflow, and WDL, providing pipeline portability for building and hosting your own analyses within DNAnexus’ cloud-native user interface or command- line interface.
Smart Reuse	Save development time and money. DNAnexus lets users allocate compute resources only to the downstream parts of the pipelines you’re optimizing, allowing for the reuse of outputs from earlier steps that were unchanged. This reduces development costs by eliminating workflow repetition and decreasing the time to iterate on a pipeline code.
Data Streaming and Automation	Automatically stream and process data. DNAnexus allows users to implement end-to-end data automation, allowing data to stream directly from their source all the way through complete analysis workflow execution.
Detailed Job Metrics	Optimize performance and investigate failures. DNAnexus provides .csv and CLI access to detailed job metrics for troubleshooting and optimization.

Available Services and Support

Expert Science and Cloud Support	Solve today’s genomic challenges - faster. Leverage our team of expert scientists with deep bioinformatics and cloud expertise and broad industry experience, to help you achieve your goals.
Expert Quality and Compliance Services	Stretch your quality & compliance resources. Our team helps customers with their medical device submissions, clinical trial protocols and submissions, managing data collaboration from CROs, regulated production diagnostics, and other regulated processes that occur on DNAnexus.
Care and Support	Provide a positive and productive experience. We provide customer service options to enable and empower customers to utilize their DNAnexus products to their utmost capability. We aim to not only provide high quality customer support, but also encourage and promote customer success.
Training and Onboarding	Ensure high return on investment. We will train your team to ensure they have the knowledge and skills to optimize use of DNAnexus. Personalized, hands-on training includes a general introduction, how to use the web user interface, command-line, and how to develop apps and workflows.

ABOUT DNANEXUS

DNAnexus is a leading provider of secure, scalable, and intuitive biomedical data analysis software and bioinformatics applications for the life sciences and healthcare industries. We actively manage and support more than 125 petabytes of complex genomic, multi-omic, and clinical datasets on behalf of a growing network of collaborations with large-scale biobanks, as well as leading pharmaceutical, clinical diagnostic, academic research, and government organizations. Scientists across 48 countries are now using our platform to gain data-driven insights that can advance scientific discovery, accelerate precision medicine, and improve patient care.