

AutoML Assistant: Democratizing and Automating AI/ML for Easier, Faster Insight Generation

Accelerating Insights with AI/ML on DNAnexus Platform

As life sciences organizations generate unprecedented volumes of genomics, imaging, and clinical data, AI and ML have become a strategic priority for uncovering meaningful patterns and accelerating translational discovery. Yet, adoption across the industry remains slow, fragmented, and resource-intensive—hampered by siloed systems, complex workflows, and a shortage of technical expertise—leaving many teams unable to transform data into actionable insights.

The DNAnexus Platform bridges this gap by providing a secure, collaborative, and AI/ML-ready ecosystem that unifies data management, scalable computing, and intelligent automation. By streamlining the end-to-end AI/ML lifecycle, the DNAnexus ecosystem empowers a broad range of users—from data scientists to biologists and clinical researchers—to confidently harness AI/ML, accelerate discovery, and drive the next generation of precision medicine.

INTRODUCING DNANEXUS AUTOML ASSISTANT

AutoML Assistant is the newest innovation in the DNAnexus AI/ML ecosystem, built to democratize AI/ML for life sciences. Building on the foundation of the DNAnexus AI/ML Accelerator—with convenient, scalable tools like ML JupyterLab, MLflow, and Data Profiler—it introduces intuitive, conversational AI-driven workflows. Non-technical users can explore data, build models, and interpret results effortlessly, while data scientists and engineers gain an intelligent copilot for hypothesis testing and model optimization.

Purpose-built for high-impact use cases such as biomarker discovery, cohort comparison, and outcome prediction, AutoML Assistant delivers speed, interpretability, and reproducibility. By empowering both domain experts and technical teams, it redefines how organizations generate insights and advance precision medicine with confidence.



This document introduces the DNAnexus AutoML Assistant. For information on the other DNAnexus components in the DNAnexus Ecosystem, follow these links:

- [DNAnexus Core Platform](#)
- [Apollo & Omics Data Assistant](#)
- [AI/ML Accelerator](#)

Empower Every Researcher

AI/ML Accelerator accelerates ML workflow by supporting developers with quick starts and scalable tasks.

Transform Raw into Research-ready Data

Apollo & ODA manage datasets and cohorts, helping you refine required data with ease and efficiency.

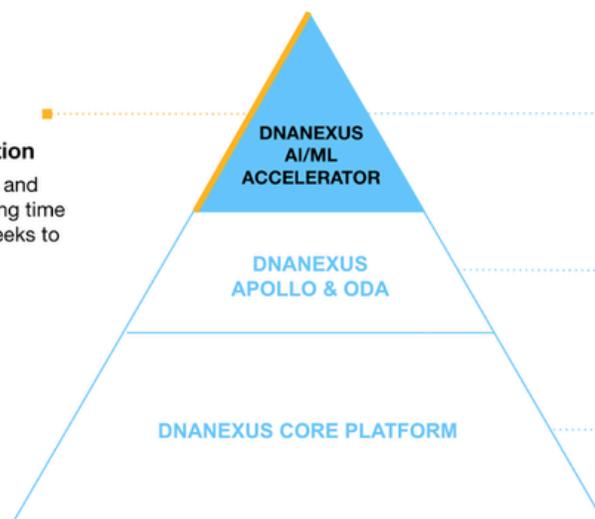
Power Precision Health with Trust & Scale

DNAnexus Platform provide a secure, compliant, and scalable workspace for managing all research data.

AUTOML ASSISTANT (NEW)

Accelerate Insight Generation

AutoML Assistant streamlines and automates ML workflows, cutting time from data to decisions—from weeks to hours.



AUTOML ASSISTANT: YOUR CONVERSATIONAL COPILOT FOR AI/ML INSIGHT GENERATION

AutoML Assistant transforms natural language into end-to-end machine learning workflows, empowering researchers of all backgrounds to generate insights quickly, intuitively, and transparently.

For all users, AutoML Assistant provides:

- **Conversational, GenAI-Powered Guidance:** Enables users to perform cohort analysis, feature engineering, model training, and interpretation—all through natural language.
 - *Example:* "Use two cohorts, Responders and Non-responders, to build a predictive model that classifies whether a cancer patient will respond to the 5-FU treatment."
- **Context-Aware Integration with DNAnexus Apollo:** Intelligently interprets data structures from Apollo Cohorts—created via Omics Data Assistant (ODA) or Cohort Browser—bridging the gap from data selection to insight discovery.
- **Automated ML Pipeline Execution:** Builds, trains, and evaluates models behind the scenes with common ML libraries like PyCaret, PyTorch, Scikit-learn, etc., minimizing manual intervention while maximizing reproducibility.
- **Statistical Analysis & Visual Summarization:** Performs key statistical comparisons and delivers clear, interactive visualizations for transparent, data-driven insights.

For Data Scientists and Developers, AutoML Assistant provides:

- **Reusable Python Artifacts:** Automatically generates dataframes, models, and Python scripts—ensuring results are reproducible, traceable, and ready for further exploration in ML JupyterLab.

Key Benefits for Your Teams

AutoML Assistant brings speed, transparency, and accessibility to AI/ML, empowering both decision-makers and technical experts to explore data and generate insights more effectively.

- **Democratize AI/ML and Foster Collaboration:** Natural language interface and no-code workflows empower clinicians and translational researchers, while developers can reuse artifacts and accelerate prototyping.
- **Streamline Cohort Analysis & Model Development:** Automated data prep and statistical comparisons make it easy to explore cohorts and extract meaningful patterns.
- **Ensure Security, Transparency, & Reproducibility:** Traceability with reusable Python Artifacts and visual outputs ensures insights are explainable, shareable, and easy to replicate.
- **Accelerate Scientific Discovery & Decision Making:** AutoML reduces model development time from weeks to hours, enabling faster, data-driven decisions.

