Palantir

Palantir Foundry

Data / Logic / Decisions

Rapid Hypothesis Generation

Secure Collaboration

Connected Biotech

IMPACT STUDY — AT A GLANCE

At a biotech startup, Foundry rapidly accelerates R&D

Data-driven investigations run over 5TB of EQTL data and 1TB GWAS results

Data Foundation includes:

→ 2.1M Chembl Compounds

→ 110M Pubchem Compounds

→ 20K UniProt Proteins

→ 5 TB Experiment Images

→ 365K of Alpha Fold PDB Files

[See Page 3 for Details]

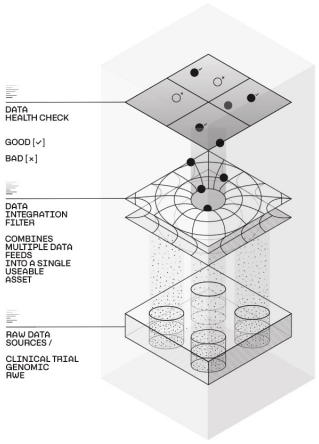
SUMMARY

Palantir Foundry significantly accelerates life sciences research at leading organizations, from the National Institutes of Health (NIH) to biotech startups. Out of the box, the platform dynamically synchronizes, processes, and connects complex assays, bioinformatics data, high resolution images, GWAS studies, and more.

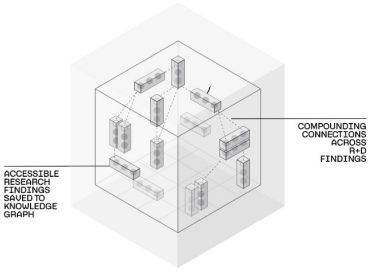
Whether you are Day 0 biotech or a large enterprise, Foundry centralizes this information along with analytics, models, and experimental outcomes into an evolving and trusted foundation. Our customers contextualize and streamline experiments, expedite hypothesis generation, and increase collaboration across their growing organizations.

FOUNDRY PLATFORM OVERVIEW

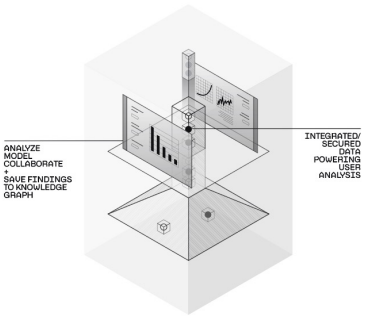
01 —
Make Data Useable



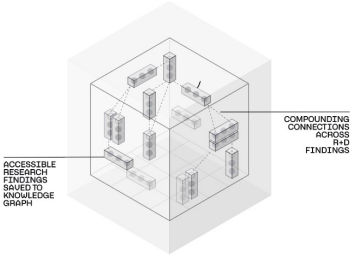
02 —
Protect Data at All Times



03—
Facilitate Collaborative Analytics



04—
Promote Knowledge & Reusability



Copyright © 2021 Palantir Technologies Inc. ("Palantir"). The information in this document is proprietary and confidential. Unauthorized disclosure to any third party is strictly prohibited. The content provided herein is provided for informational purposes only and shall not create a warranty of any kind. All data shown is notional.

Palantir

Connected Biotech

KEY OBJECTIVES

01 Optimize the growing environment of external information to better tailor and support internal research and development

02 Streamline internal collaboration to accelerate insight generation across teams

03 Automate and optimize experimentation, from hypothesis to operational insight

CHALLENGES

Now more than ever, the biotechnology and life sciences sector plays a vital role in navigating global health challenges amidst complex and changing conditions. Rapid advances in information technology along with proliferating information presents new opportunities — but also creates new challenges as companies must synthesize a vast array of in-house and external data (licensed and public), analytics, insights, and models.

To investigate relationships across sources and enable research based on a trusted and foundational source of truth, these organizations seek to connect genomics data, molecular and protein libraries, cell-based assays, multi-format images, experimental metadata, and more. Additionally, to make smart decisions based on research, organizations require robust analytical frameworks and predictive models that can both read from and contribute to this complex network of information.

FOUNDRY PLATFORM

The Foundry platform creates a central brain for the organization — connecting data, logic, and decisions.

The result? Close coordination between research and operations that minimizes drag between processing information and transforming it into value-generating outcomes.

Foundry enables a a unified and commonly managed technology stack, giving both early and later stage companies the opportunity to maximize what makes them uniquely valuable.

FOUNDRY ADVANTAGES

Out-of-the-box data connectors, interoperable with various in-house and cloud architectures — and compatible with CheMBL, PubChem, UniProt, and others

Automated lineage system that tracks usage of synchronized data, so that any research output (e.g., dataset, model or visualization) can be traced to the original version of the raw data — ensuring repeatability, transparency, and auditability

Seamless and automatic support for writing back data to a common ecosystem, which links traditionally siloed user groups such as lab operators, molecule researchers, and data scientists through shared investigations and decisions

Ready-to-use frameworks and interfaces for AI/ML model management, deployment, training, and optimization

Cross-language (Python, SQL, R, Java) coding environments to enable versioned collaboration

Operability in a fully-managed cloud environment that scales up and down automatically with the size of a business — ensuring impact across a company's lifecycle

Readily deployable templates for data pipelines, analytics, models, and user interfaces (low/no code environments), which significantly reduce the technical barrier to entry and burden on in-house development (e.g., OMOP pipelines for common RWE data sources, and a Clinical Trial Harmonization Suite)

World-leading security and permissions infrastructure that caters to multiple levels and types of access control required for collaboration on sensitive research and clinical operations data

Performant processing of high scale data such as genomic analysis supported by Foundry's underlying distributed computing architecture and its ability to connect to custom high performance compute clusters

Automated data integration in any format (e.g., HL7, FHIR, CDISC, genomic file formats such as BAM or VCF, and high-content image files)

Copyright © 2021 Palantir Technologies Inc. ("Palantir"). The information in this document is proprietary and confidential. Unauthorized disclosure to any third party is strictly prohibited. The content provided herein is provided for informational purposes only and shall not create a warranty of any kind. All data shown is notional.

Palantir

Connected Biotech

Impact Study 01 →

Creating a Centralized Knowledge Repository for Use in Hit Identification and Lead Generation

CHALLENGE

To accelerate exploration and discovery of molecules, proteins, and genes, a biotech startup needed to relate high volumes of internal and external data — including millions of dose response data points across disparate compound libraries and terabytes of experimental images.

Handling and analysis of this data was spread across different teams — including hit identification researchers, chemists, data scientists, and lab operators — impeding fast decision-making and raising questions as to how certain conclusions were reached. Additionally, a fragmented layer of tooling made it difficult to replicate experiments and trace data lineage from experiments from beginning to end.

SOLUTION

The customer uses Foundry as a central brain that coordinates exploration, analysis, and decision making across systems and data sources to expedite existing research processes and fuel cross-organizational collaboration.

Using Foundry, our customer has created a platform that (1) centralizes the data and workflows to structure knowledge, (2) performs rapid and effective hit prioritization, and (3) codifies a standardized set of terms for research to ease data retrieval, streamline experiments, and automate analysis.

Foundry has enabled our customer to excel in previously impossible ways →

Conduct high-scale genomic analysis across multiple modalities to identify candidate genes; this includes 5 terabytes of of EQTL data and 1 terabyte of GWAS results

Create a knowledge graph with over 50 unique nodes of information (e.g. connecting genes to associated proteins and targets, along with experimental activity)

Enrich a compounding data asset by integrating a large variety of data sources using entity resolution, including:

2.1M
Chembl compounds

110M
Pubchem compounds

365K
Alpha Fold PDB files

20K
UniProt proteins

5TB
Internal experiment images

Develop low and zero code workflows for scientific users to rapidly assess candidate genes using computationally intensive methods

Prioritize compound libraries for enhanced high through-put screening

Create a digital twin of research to easily iterate on, explore, and prioritize molecules that are the most promising

Establish a secure and accessible environment for researchers to identify and evaluate novel opportunities, maximizing cross-pollination of ideas whilst maintaining a high-degree of security and audit-ability

Validate and quality control against experimental standards

Copyright © 2021 Palantir Technologies Inc. ("Palantir"). The information in this document is proprietary and confidential. Unauthorized disclosure to any third party is strictly prohibited. The content provided herein is provided for informational purposes only and shall not create a warranty of any kind. All data shown is notional.

Palantir

Connected Biotech

Impact Study 02 →

Producing a Resilient Biomanufacturing Supply Chain

CHALLENGE

A client developing novel cell therapies needed a supply chain that was both flexible and highly controlled to account for raw material shortages, vendor backlogs, and other shocks, without compromising quality. Teams across the company operated in siloed systems with little visibility into the downstream impact of any system changes.

Making effective resource allocation decisions also proved difficult due to disparate inventory and operational records. They required the ability to dynamically re-route resources as needed along with visibility into their internal testing team's future capacity to better understand when additional vendor services were needed.

SOLUTION

Single source of truth for end-to-end visibility

Automated alerts, both threshold rules and ML-model based

Scenario analysis tools to mitigate disruptions

Within Foundry, this client has integrated disparate ERP systems into a single source of truth. They migrated paper and email based manufacturing processes into an electronic batch record. The granular details are then recorded by technicians or integrated from IoT sensors, such as environmental monitoring. Both in-process and release assays are tracked, and equivalent data from CMO vendors is imported to achieve end-to-end visibility, spanning from raw material sourcing to cell modification, expansion, and quality assurance.

Automated alerts, both threshold rules and ML-model based, ensure that troubling patterns in a batch are immediately brought to the attention of manufacturing teams in time to intervene. When deviations are severe enough to warrant discarding a batch, the lineage of shared materials (media, equipment, etc.) allows for confident predictions to be made regarding other batches which may have been affected. Regardless of their technical ability, Foundry's suite of analytical tools enables subject matter experts to carry out such analysis with clear visibility into the process at hand.

As batches progress into inventory or are held for investigation, projections of product supply automatically adjust to inform planning. Scenario analysis tools allow the supply chain teams to experiment with adjusting raw materials and staff time allocations across programs to limit the impact of disruptions and adapt to changes in clinical demand. All of these scenarios and plans are created from inventory data collected through live connections with ERP systems. Projections can be written back to such systems, taking advantage of existing infrastructure investments while avoiding silos and enabling collaborative planning and visibility by teams across the organization.

Copyright © 2021 Palantir Technologies Inc. ("Palantir"). The information in this document is proprietary and confidential. Unauthorized disclosure to any third party is strictly prohibited. The content provided herein is provided for informational purposes only and shall not create a warranty of any kind. All data shown is notional.