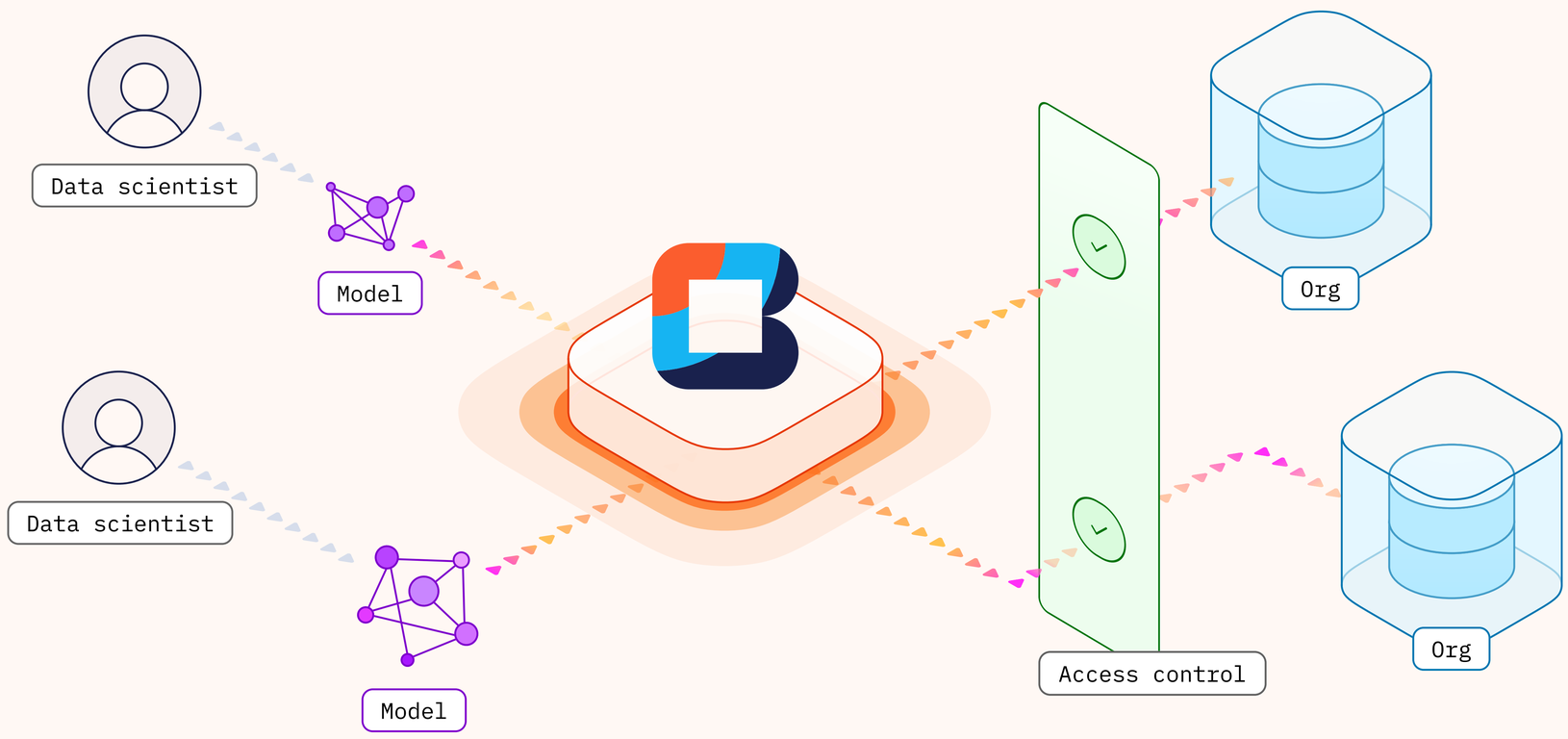
About Bitfount

Data is critical to advancing our understanding and delivery of healthcare, but datasets are vastly underutilised due to barriers to collaboration stemming from privacy, security, or commercial sensitivity, as well as technical barriers to the creation, transfer, and maintenance of multiple up-to-date copies of a dataset.

Bitfount remedies this by providing a flexible, easy-to-use platform for privacy-preserving data collaboration through Federated Data Science. Federated Data Science enables data teams, analysts, and researchers to gather statistical insights as well as develop machine learning models on data which they don't have access to in raw form.

It also enables custodians of sensitive data to share the benefits of their data, without giving up control or privacy. This is achieved by sending encrypted analysis and AI code to data where it already lives, behind the custodian's firewall, and retrieving the results of analysis queries or machine learning tasks without the need to transfer raw data to Bitfount or any third parties.

The Bitfount platform can be easily installed and operated via a no-code desktop or web application, and via an extensible open source Python SDK.



How it works

Think of Bitfount as the infrastructure connecting all parties involved in your data collaboration. In addition to Bitfount's federated execution layer, it also contains a powerful governance layer, providing usage-based access controls which can be configured to govern the way in which users collaborate.

Datasets connected to Bitfount are accessed via Pods (Processors of Data). A Pod is the local service which handles processing of data science tasks on the datasets configured within it, allowing insights to be gained whilst data remains behind the custodian's firewall.

Users collaborate using Bitfount Projects. Projects contain a collection of settings that determine the who, what, and how of the data collaboration taking place within them.

There are three key user groups

Project owner

Responsible for bringing together collaborators, setting the project's terms & conditions, and selecting the data science task(s) to be carried out within the project.

Data scientist

Develops AI models and data science tasks for projects, involving machine learning algorithms, analytics queries, and other analysis protocols.

Data custodian

Owns the data. Connects one or more datasets and links them to Projects for analysis. No patient data is transferred away from the data custodian's systems.

Use cases

Clinical trial feasibility
Analysis of trial feasibility with varying recruitment criteria using site data, including EHR and imaging data.

Clinical trial pre-screening
Identification of eligible patients based on EHR and image biomarkers.

Post-market surveillance
Distributed real world evidence based on analysis of EHR and imaging data.

Operational research
Federated benchmarking and audit across multiple sites for operational- and quality-of-care improvements.

Biomarker development
Leveraging distributed data to build new AI-based biomarkers.

Biomarker validation
Distributed evaluation of AI-based biomarkers.

Population research
Federate your demographic research across remote sites.

Features

Easy setup

- Point and click desktop and web based applications
- Other install options including Open Source Python SDK and Docker image
- SaaS platform for data discovery, access control, account administration and more
- Custom integrations with PACS and EHR systems available
- Regular automated software updates provided

Secure

- Single Sign-On (SSO) as standard (support for SAML, OAuth & OIDC)
- Compatible with organisational IT firewalls (no incoming connections required) and proxies
- Granular Role-Based Access Control (RBAC)
- Built-in governance workflow for setting up projects and requesting and granting permissions
- Secure Multi-Party Computation (SMPC) protects individual contributions
- Bitfount handles all fully-encrypted communication and orchestration
- Differential privacy for mathematically prescribed disclosure control (ML and SQL)

Flexible

- Machine Learning and SQL and other data science operations supported
- Algorithms such as federated training, evaluation, inference, SQL and much more
- Built-in privacy layer with techniques including differential privacy and secure aggregation.
- GPU support for fast model training
- Connect data from cloud and on-prem sources
- Support for tabular and image data
- Webhooks for easy workflow integration
- Organise your data using filtered views