

How a
Top 10
Pharmaceutical
Company

**Centralized Imaging Data to
Accelerate Clinical Research
and AI Development**

The drug-development lifecycle generates scores of valuable imaging data. Pharmaceutical companies can leverage that data to accelerate R&D, make better-informed decisions about trial efficacy and develop AI models that can improve time to market. But enabling data access, standardization and analysis in a timely manner is another story.

The Challenge

A leading global pharmaceutical company faced this issue with years of accumulated images, spanning multiple therapeutic areas including oncology, ophthalmology and neurology. Their data were scattered across disconnected systems, each with its own naming conventions, formats and access workflows. These silos created significant inefficiencies, as teams struggled to locate datasets and researchers repeated work others had already completed.

Then there's the compliance and governance risk that accompanies collaborating with external partners. This often requires extensive manual work to standardize data, as metadata can be manipulated or stripped during transit, while variation in imaging data modality, instrumentation, acquisition type and data presentation requires specialized domain expertise to process.

With datasets stretching to terabytes across cohorts, manual data management was cost-prohibitive, time-intensive and prone to error. Preprocessing tasks for a single DICOM exam for analysis could take hours when performed manually, delaying downstream research and adding significantly to operational costs. They needed a way to scale their imaging data management strategy so they could discover secondary insights from cross-study analyses while accelerating their AI and machine learning ambitions with consistent, high-quality datasets.

The Goal



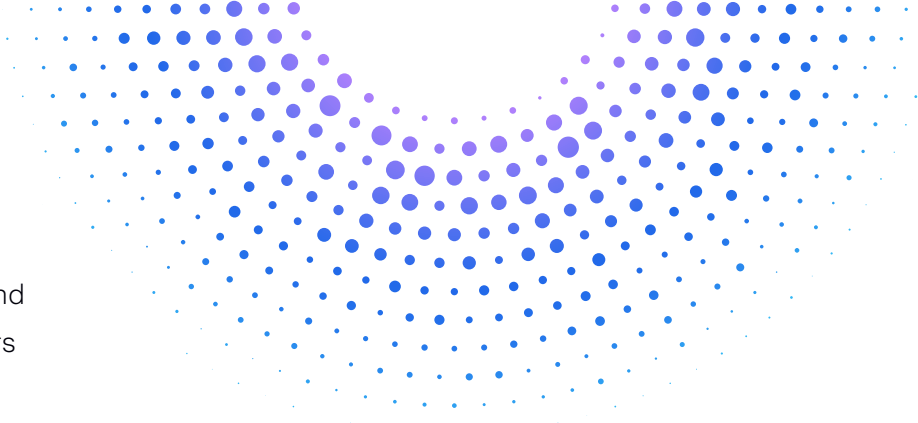
Centralized imaging data management for secondary analysis and AI development

The Solution

The company needed a centralized imaging platform that could sit atop its secure cloud infrastructure and harmonize ingestion, curation, preprocessing and analysis for all imaging modalities. They engaged Flywheel to implement the Flywheel imaging data management and analysis platform on Amazon Web Services (AWS), bringing their imaging data together in one location. This enabled collaborators and stakeholders across the globe to access and analyze data as soon as it was ingested while at the same time unlocking legacy data that wasn't collected with reusability in mind. They could do this without sacrificing data privacy, given Flywheel's built-in, role-based access controls.

Flywheel became the company's centralized imaging repository, with harmonized data access points for early exploratory work and AI dataset curation. The platform's cloud-scale architecture allowed new curation processes to be introduced over time as the company's research needs evolved.

In addition, Flywheel automated company-specific workflows for de-identification, metadata extraction, quality checks and standardized indexing with Flywheel Gears, containerized algorithms that can be easily added to the



platform. From ingestion to segmentation and QC, Gears helped them save countless hours on data processing, with modality-specific workflows for MRI, CT, PET, OCT and Fundus imaging. This also significantly reduced the variability that had previously slowed down research.

The Outcome

Accelerating AI Validation and Clinical Decision-Making

With Flywheel's extensible architecture and automation capabilities, the company was able to test and validate AI models much faster than before, and at a previously unachievable scale. Leveraging Flywheel's Reader Studies module accelerated ground truth data collection for training and fine-tuning models. The company has developed several AI models and has now received FDA clearance for one of them.

Beyond building out their AI capabilities, they found significant additional value in using the platform to support ongoing clinical trials more directly. By harmonizing imaging data from CROs, reading centers and internal groups into a single accessible environment, teams were able to perform more timely and comprehensive analysis to support clinical trials while saving time and funds from third-party data requests that could take weeks to arrive. With better access and more comprehensive, preprocessed datasets, the team will also be able to make faster and more

How Flywheel Achieves FAIR Data Principles

Findability is enabled with automated metadata extraction and classification, creating an easily searchable collection in which all aspects of each dataset — and every process run across those datasets — are automatically indexed and immediately available for discovery.

Accessibility is ensured through centrally managed authentication protocols, a web-based interface, and open REST APIs that enable programmatic access and containerized algorithm deployment via Flywheel's software development kit.

Interoperability is achieved by aligning data to existing R&D workflows and automating de-identification for consistent anonymization across departments.

Reusability is built in through standardized curation and quality controls, along with full provenance, ensuring data can be reliably applied, reapplied and shared.

informed decisions at each stage of clinical trials, translating into further cost savings by diverting resources toward more successful trials.

In addition, the platform enabled secondary endpoint analysis to build a richer picture of how patients were responding to treatment. This included automating quantification of full tumor burden rather than a limited set of target lesions, providing a more complete and sensitive view of disease activity.

And with data FAIRification across the enterprise achieved, they've laid the groundwork for a new era of AI-enabled drug development. The platform now supports data from more than 240 clinical trial datasets and counting, representing more than 100,000 patients across multiple therapeutic areas.

An All-in-One Solution for Maximizing the Value of Imaging Data

Overall, the company was able to dramatically compress their imaging data curation timelines. Tasks that previously required months of manual effort were reduced to a matter of weeks; curation that once took hours per acquisition could now be completed in minutes. With that time freed up, researchers could focus on more high-level work, like developing models and accelerating decisions. These advances transformed imaging from a bottleneck into a strategic accelerator.

With a reliable and scalable infrastructure in place and a new AI model already cleared, the company is well-positioned for the future of AI-driven clinical trial support and staying competitive in the race to develop the next life-changing therapy.

Key Achievements at a Glance



- 240+ datasets and 100,000+ patients represented in a unified imaging environment
- More than 200 million images ingested, with 150 million+ automated jobs run across global teams
- Several AI models tested and validated using Flywheel to create analysis-ready datasets

Visit flywheel.io

[Schedule a demo](#) or [get in touch with a Flywheel representative.](#)