



---

**IMPLEMENTATION OF  
INSTRUMENT RAW FILE  
MANAGEMENT SOLUTION**

**A CASE STUDY**

---

# CONTEXT BEFORE ZIFO'S INVOLVEMENT

## Customer Overview

Our customer was one of the top global Bio-pharmaceutical companies researching and developing therapies in the areas of cancer, HIV/AIDS, cardiovascular disease, diabetes, hepatitis, rheumatoid arthritis and psychiatric disorders.

## Problem Statement

Customer's Translational Medicine function operated in 5 different sites across the globe. Many of the labs supporting Translational Research were also situated at different locations and operating at varying levels of maturity. There were significant differences in the SOPs, policies, workflows, applications and instruments followed by every lab. This resulted in data (such as analytical data) critical to Translational Research getting fragmented across different storage systems limiting accessibility and increasing the risk in data quality. This prompted the Customer to develop a large file management solution built on enterprise IT data standards and using leading storage technologies that satisfied the below needs:

- Ability to store all type of data
- Ability to store large files up to 50 GB
- Ability to secure data using IAM
- 99.999999% durability
- Adopt a platform-based approach for LIMS and raw file management solution
- Flexible cost options

Moreover, the Customer was in the process of acquiring another large Bio-pharmaceutical company. This meant the existing systems and processes had to become structured and extensible to accommodate new laboratories and ways of working.

## Objectives

Key value drivers to be delivered by a unified process for raw data management:

Increased  
collaboration

Improved data quality,  
availability and  
accessibility

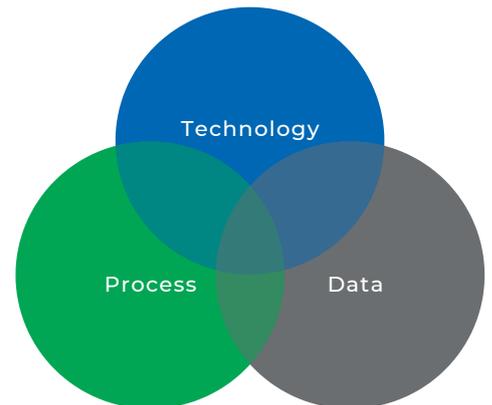
Cost reduction with  
applications retirement  
and licenses reduction

# ZIFO'S METHODOLOGY

Zifo's Scientific Informatics Consultants performed a three-fold analysis focusing on Process, Data and Technology to understand the current state challenges and define future-state capabilities.

The analysis focused a number of critical factors related to Large file management, notably:

- Data security
- Accessibility and Scalability
- Cost options
- Existing infrastructure of Customer
- Access and security restrictions of Customer



Future-state solution design was driven based on the following foundational principles::

- Solution should leverage key enterprise IT components (LIMS, sample management, project/requests management, lab instrument files and analysis data)
- The solution should respond faster to changing landscape and environment
- The solution should facilitate access and metadata management by leveraging well-defined governance processes and policies
- The solution should provide comprehensive audit trails, versioning and traceability

Once the future-state design was approved by the Customer, Zifo Consultants developed a tailored solution for large file management leveraging Industry Best Practices and leading technology solutions .

# SOLUTION

Need	Proposed Solution	Reason
A single solution to configure and manage lab workflows	Sapio Sciences Exemplar LIMS	Highly and quickly configurable, works with both SQL and Oracle, is secure and flexible with respect to hosting
A platform to manage raw files (metadata, file access and archiving)	iRODS as a middleware to integrate with Data Lake	Comprehensive solution to policy-based data management across geographically separated users
A low-cost, multi-tier storage system for Raw Files generated during sample analysis	Amazon S3	Globally (and Customer's data policies recommended) means of storage and exchange. Also, Exemplar is hosted by Sapio on AWS
A tool to automatically copy the files physically to AWS in the background	SmartSync	Full-featured solution to back up and synchronize data to AWS



1. Analyst performs experiment in Exemplar and gets assigned unique Experiment ID
2. Analyst performs experiment to process and analyze bio-samples through analytical techniques such as FACS, single cell RNA Sequencing, ELISA, LCMS
3. Raw data from instrument stored in pre-defined folder in Instrument Workstation; tagged with unique Exp. ID
4. SmartSync moves files from Workstation to appropriate storage hierarchy in S3 bucket
5. iRODS monitors all updates to S3 hierarchy
6. iRODS creates Audit Log of all updates
7. Exemplar LIMS performs the following: Retrieves iRODS log file, parses AWS S3 File Movement log file, parses unique Experiment ID and location of Raw Data files from iRODS Audit Log file

# IMPLEMENTATION HIGHLIGHTS & BENEFITS

## Implementation Highlights:

- Zifo's solution design was approved by the Customer Data Architects, Data Governance and Security teams
- Automated management of approved result data through a new UI for results uploaded into Exemplar LIMS and transfer to S3 after approval
- The solution was developed and deployed for user testing and feedback in under 5 months
- Some facts:
  - Number of sites involved: 3
  - Number of instruments integrated: 60
  - Number of Workflows which supported the integration: 18

## Benefits:

One of the most significant benefits to the Customer was the end result of a standardized, global process for large file management which provided:

- a solution built on enterprise IT data standards
- a secure and reliable storage space for large files of any data type
- means to collaborate easily
- improved data quality and accessibility

Other benefits included reduced human error and effort and reduced cost incurred through procurement of licences of multiple applications.

# LET'S BUILD TOGETHER

*Contact our expert to explore how we could help you*



Raj is a Scientific Application Analyst at ZIFO RnD Solutions and is based out of USA. His focus is around AWS, Exemplar, Java, SQL and Solution Designing.

He has helped our customers with:

1. Data migration from existing DB to Exemplar-recommended data model
2. In-House(Customer's) Exemplar hosting on Unix server
3. External integration with Exemplar with AWS S3 using Java for samples, raw data and analysis results
4. External integration with Exemplar with repository that provides REST API services via JSON using Java for samples

He can be reached at [rajarshi.s@zifornd.com](mailto:rajarshi.s@zifornd.com)

