

# End-to-End and Fully Integrated Clinical Development Platform with eDC/Labs etc. for Data Management, Medical Review, Statistical Analysis and Adaptive Design

Peter Wang, Janssen R&D; Johnson and Johnson  
Sunil Vindiyala, Janssen R&D; Johnson and Johnson

## ABSTRACT

*LSAF (SAS® Life Science Analytics Framework) and SPACE are Janssen's end-to-end, and fully integrated clinical development platform that pulls in raw data off eDC plus dozens of Central/Local labs by schedules into LSAF -- where data managers all over the world, leverage the automated workflows to check, correct, and validate clinical raw data in order to ensure its completeness and accuracy on regular daily and sometimes hourly basis. Data then feeds into Medical Review/Safety Portal for additional/in-depth analysis per protocol and pre-set standards. And it's feed further into AWS hosted SAS and R GRIDs for programming and statistical analysis. Pinnacle 21 and Adaptive Design are fully integrated into LSAF based on "real-time" clinical evidence that allows our clinical team to adjust development protocol timely.*

*It took Janssen all together about 10 years to build this fully integrated platform. We started in 2013 with our 1<sup>st</sup> integrated LSAF platform with eDC, Labs for data management, and the Medical Review was then integrated in 2014. Pinnacle 21 and Adaptive design functions were added 2 years later by 2016. SAS and R GRID were then fully integrated in 2021.*

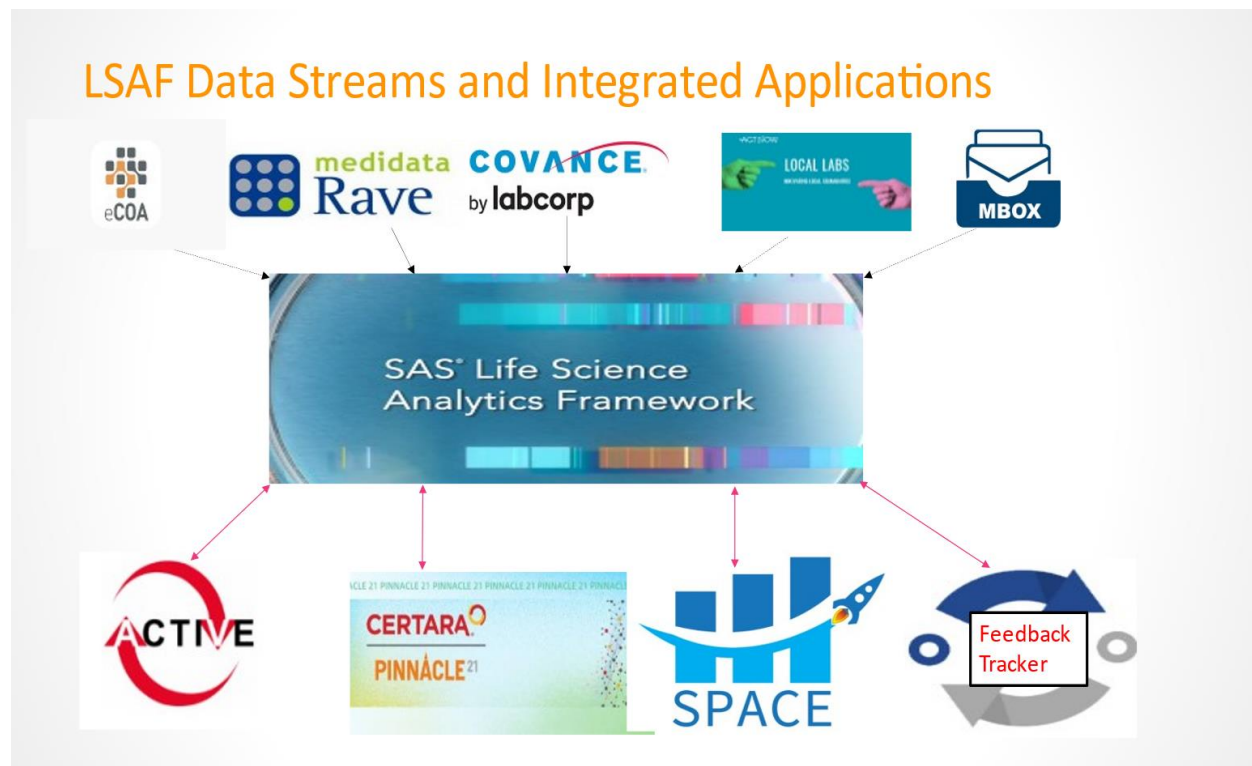
## INTRODUCTION

There are many streams of data throughout the clinical development process. This includes data captured by sites using eDC systems such as RAVE; data off central and local labs, and additional ones like PK etc. Over the past 10 years, we have consolidated a couple dozens of individual clinical development applications into a fully integrated platform with LSAF and SPACE etc. LSAF functions as our clinical data repository that facilitates seamless data flow for the cleaning and data management processing. It autoloads from RAVE, labs, CROs with built in workflows while adhering to all the GxP requirements. Timely and ongoing clean data enable Janssen to shorter timelines for analytical processing. LSAF is powered with SAS engine; and all its components of a data store provides flexibility for our 5000+ global users 24x7. LSAF also provides integrated data flow, security and versioning and audit history etc.

SAPCE and its SAS GRID architecture gives us a high-performance computing environment for clinical programmers and statisticians to generate aDAM data seamlessly. Integrating the environments together has provided the benefit of eliminating data duplication and redundancy thus facilitating both accuracy and efficiency in the process of clinical development. LSAF is an End-to-end and fully

integrated clinical development platform that pulls in raw data from eDC and dozens of Central/Local labs using scheduled jobs. Data managers complete all the work leveraging automated workflows to check, correct, and validate clinical raw data to ensure its completeness and accuracy on regular daily and sometimes hourly basis. Data then feeds into Medical Review/Safety Portal for additional/in-depth analysis per protocol and pre-set standards. And it also feeds into AWS hosted SAS and R-Grid environments for programming and statistical analysis. Pinnacle 21 and Adaptive Design are fully integrated into LSAF based on "real-time" clinical evidence that allows our clinical teams to adjust development protocol timely. It took us about 9 years to build this fully integrated platform that facilitated integration with eDC, Labs clinical data management and Medical Review and was up and running by 2014. Pinnacle 21 and Adaptive design functions were added 2 years later by 2016. SAS and R GRID were then fully integrated in 2020. There are over 5000 global users and the environment is supported 24x7. There are many successful stories with excellent outcomes, and we plan to share what we consider "areas for improvement" and our "lessons learnt" as well.

## DEEP DIVE INTO LSAF AND SPACE

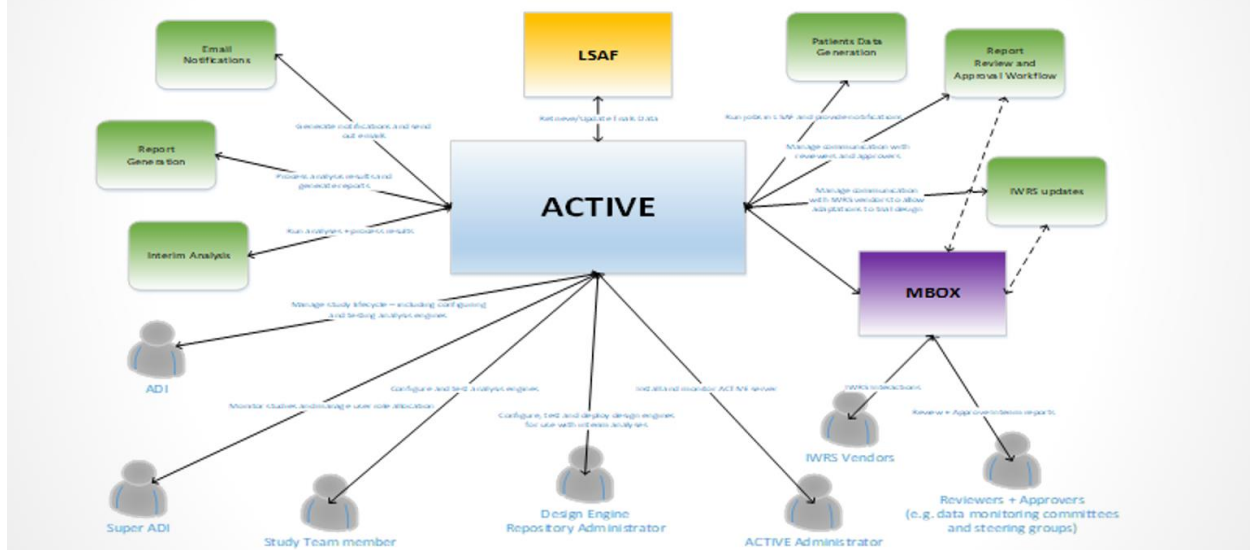


SAS Lifesciences Analytics Framework (LSAF) is an environment that provides an integrated system for aggregation, integration, transformation, managing, analyzing, reporting, and reviewing clinical research information. LSAF provides version control and audit history features for research content. LSAF is designated as global data repository for all our clinical trials. Global access by external partners, built in workflows, and role based controlled access allows for multiple, flexible outsourcing/partnership

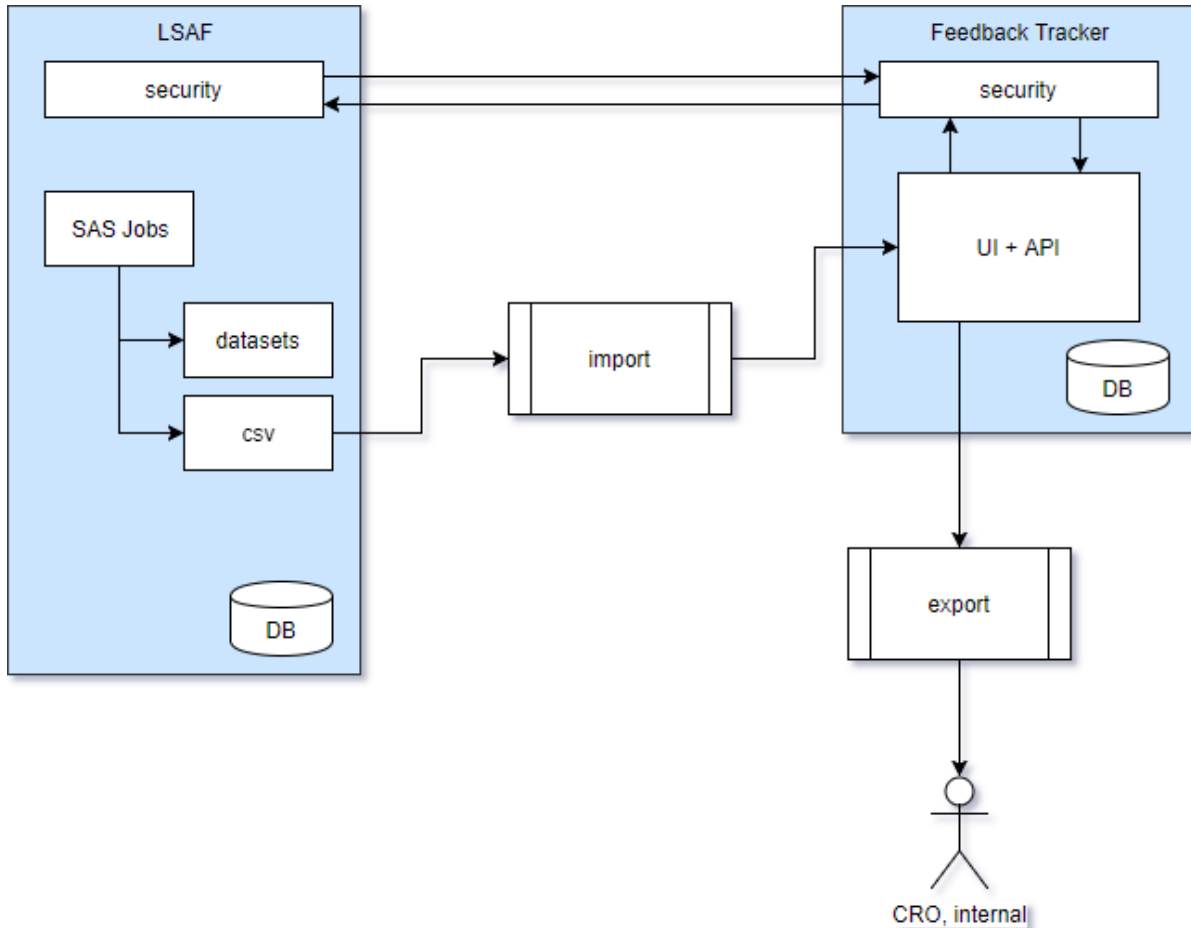
models.

LSAF highlights include: Global data repository for clinical trials, worldwide access – Accessible over secure web from anywhere, Data storage and processing in the form of SAS datasets, Provides Personal workspace for every user, licensed for unlimited users, Integrated Development Environment (IDE) for SAS program development, Workflows, Job scheduler, Detailed audit trail features, Access for external partners, Extensible using JAVA and SAS Macro API's and extensions modules to autoloading data from different sources using sFTP and HTTPS. LSAF is an elastic n-tier architected solution that provides opportunities for easy expansion nodes at each of the presentation, mid-tier and SAS work layers. Currently the system has 3000+ active users and 1200 active trials, and it is used by Janssen internal and external users across the globe 24x7. LSAF acts as the central data repository of all clinical data with each stream eCOA, Raw data from RAVE, lab data from Covance and local labs and CRO data from MBOX all automated using sFTP or HTTPS protocols. The cleaning and conversion of data are handled using the LSAF job of jobs feature and flows through a defined workflow called automated data flow without user interaction. The feature was developed using the power of SAS engine which is built into LAF. Data conversion to a Janssen standard called Data review model or SDTM is also automated using LSAF workflow mechanism. Once the data is converted, it is exposed for consumption to down stream systems. One of the examples is the Adaptive trial design system call Active that provides opportunity for early detection and trial design changes. This has been extremely beneficial both in terms of study duration and cost savings. ADI management of study workflow Study Team configuration of analyses – in the setting up of new studies and in any subsequent changes to these studies, whilst in production to curate critical issues uncovered has made excellent business cases.

## Deep Dive 1. Adaptive Design (ACTIVE) System Architecture



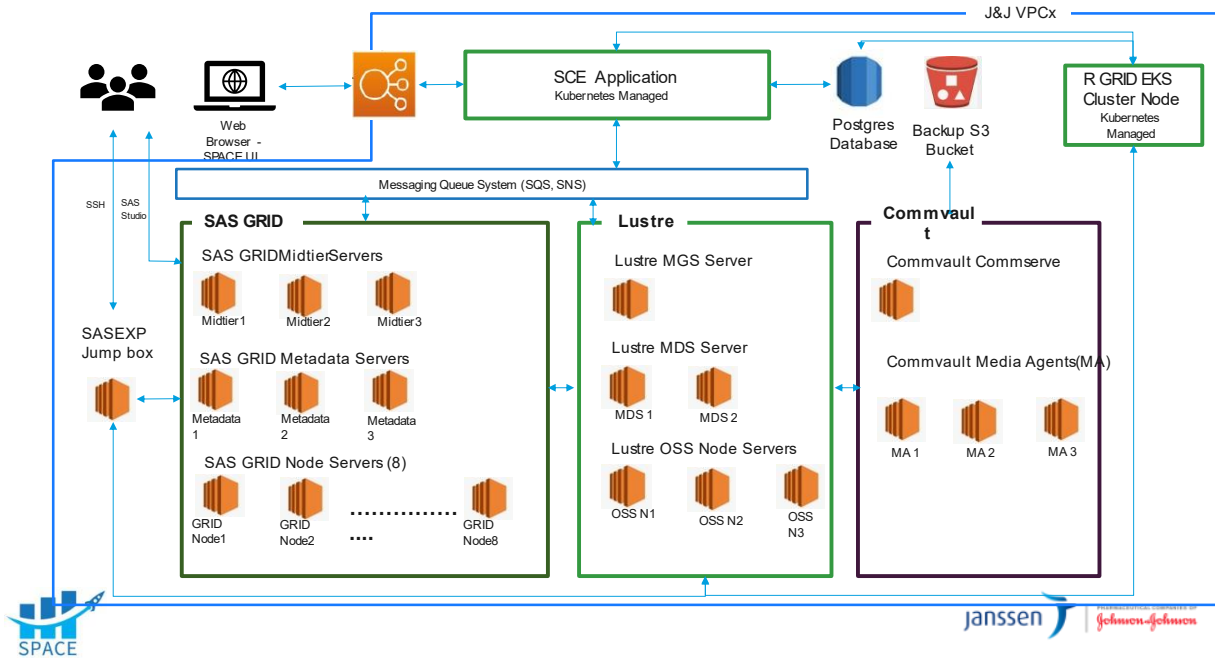
Generation of subject data in a standard format for use with analysis engines, analysis of interim data via trial execution design analysis which can be run using predefined design engines available from a central repository made adaptive design strategy seamless. The integration also facilitates external analyses, run offline by the user via a manual extraction and upload of inputs and results to a shared MBOX location for review.



Deep dive 2: Feedback Tracker. It is another key integration into our LSAF environment. It facilitates detection of data issues early and an integrated issue resolution process. The efficient follow-up of detected/reported issues, along with seamless interaction between the different parties for query resolution, along with up-to-date statistics on the progression of the issue resolution has made identifying data issues and query management quite efficient and effective.

## SPACE

## Architecture – SPACE Platform



SPACE is a web-based custom solution that is integrated under the hood with both SAS GRID and R-GRID. The GRID mechanism provides a high-performance computing for process DRM/SDTM data to run study analysis and aDAM conversion. SPACE is integrated with LSAF using extensible Java APIs and does not require data to be copied over nor requires manual interaction as the process is scheduled to look for new data to streamline the process further. We achieved our objective for providing an integrated environment with end-to-end clinical data processing.

## CONCLUSION

Janssen's fully integrated clinical development platform continues to progress and has worked quite well for us in many areas, including:

1. Stable and scalable platform with reliable performance. While scaling up with exponential growth for users, studies, and job runs, median job run is stable around 15 seconds; and we are running over 230,000 jobs monthly.
2. WEB based platform allowing onboarding external users.
3. No limit on number of licensed users
4. Robust workflows
5. One platform to store the data and perform data integrations, transformations, and cleaning.
6. Integrated analytical engine (SAS)
7. LSAF APIs are very robust and well designed for integrations.

8. Integrations, e.g., RAVE, ACTIVE, PE21, SPACE
9. Excellent data browsing (individual datasets content) Robust job scheduling capability.
10. Auditing
11. Flexible version control
12. Automated download and data push
- 13 Cloud based (AWS) GxP complaint Platform at scale
14. Robust and scalable compute tires (SAS GRID, R GRID & Python)
15. Multiple computing engines with containerization (open source)
16. Robust workflows
17. Simple folder structure
18. Easy Integration etc.

As they are work in progress, we are working for further improvements across the following areas.

1. Security model
2. Folder structure
3. Large volume of Data handling, esp. for Biomarker/Genomics data
4. Performance challenges, esp. with large study with tens of thousands of subjects.
5. Data Flow Automation
6. Global Operational Support Team management etc.

Thank you.

## CONTACT INFORMATION

1. Peter Q. Wang [pwang@its.jnj.com](mailto:pwang@its.jnj.com)
2. Sunil Vindyal [svindyal@its.jnj.com](mailto:svindyal@its.jnj.com)