

Lessons Learned from Teaching 250+ Life Science Analytics (LSAF) Classes to Our Colleagues at Janssen Research and Development

Margie Merlino and Jeanne Keagy, Janssen Research and Development

ABSTRACT

In 2014 Janssen purchased SAS® Drug Development (SDD), later renamed Life Sciences Analytics Framework (LSAF), from SAS® Institute and successfully implemented the system into their Data Management (DM) processes for uploading and validating SDTM datasets. The following year, Janssen expanded the use of LSAF to their Data Analysis (DA) processes for creating ADaM datasets, TFL's and ad-hoc analyses and simulations with study data.

Prior to the use of LSAF, Janssen's DA group used Windows SAS and an in-house developed user interface to manage the activities related to clinical trial deliverables and ad-hoc analyses. Using LSAF for both DM and DA had many advantages such as both groups could access one central repository to view all study data, interim and database locks. However, despite the success of LSAF within DM, the DA transition to the LSAF environment did not go well.

From 2016 through 2018, Jeanne and Margie taught over 250 LSAF classes to their co-workers at Janssen Pharmaceutical Research company. Jeanne conducted both the DM and DA training and Margie conducted DA training. What they will present are some of the technical and not-so-technical lessons learned from teaching students from a variety of technical and scientific backgrounds how to develop their SAS programs and jobs in the cloud-based LSAF environment.

INTRODUCTION

The decision to transition the Data Analysis (DA) activities at Janssen Research and Development from a Windows SAS to the cloud-based LSAF environment was quite momentous but carefully reviewed. LSAF was already used by Janssen's Data Management (DM) associates and their CROs for uploading and storing the SDTM data for most studies. A proof of concept was conducted in which the DA outputs (ADaM datasets and TFL's) for a Phase 1 study were successfully developed in LSAF. The LSAF product included functionality for version control and audit trail reporting that did not exist in Janssen's Windows SAS environment.

As far as hardware considerations, the LSAF environment was hosted by SAS® Institute so Janssen would no longer have to maintain servers to run Windows SAS, and they would no longer have to purchase and pay maintenance fees on desktop licenses for each SAS programmer. This was critical because Janssen's current SAS servers were outdated and the version of the Windows operating system was soon to be out of maintenance.

In preparation for the move to LSAF by statistical programmers and biostatisticians, training was developed and conducted, starting approximately one year prior to the planned cutover date to the new environment. Training classes were held between 2-3 times per week during 2016 and through the first quarter of 2017. What follows are the challenges encountered during training and the lessons learned from those challenges.



WHO MOVED MY FOLDERS?

In his book “Who Moved My Cheese?” author Spencer Johnson provides a humorous metaphor that considers the way in which different personalities cope with change. And while the target audience for this book is presumably those coping with the loss of a job or a relationship, abrupt changes to routine tasks can create stress and a sense of loss as well. That is what happened when the statistical programmers and biostatisticians attended LSAF training and got their first look at the new SAS development environment. They could not find their study folders and files because an entirely new folder structure was introduced. And there were other differences in the new environment that caused concern.

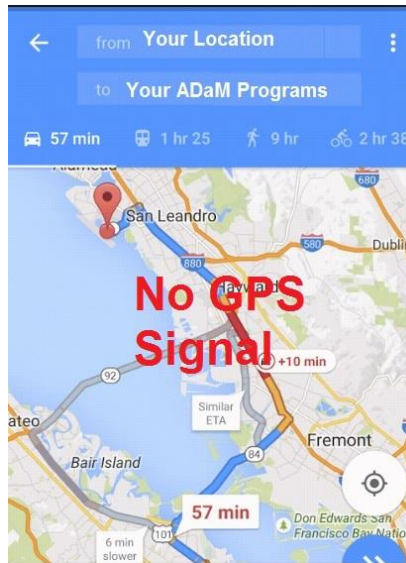


WHERE IS MY STUDY?

In their Windows SAS environment, the Janssen study teams use the same folder structure/hierarchy for every study and the folder names consist of the compound, protocol, and purpose of the submission (interim analysis, periodic safety update report, clinical study report, etc.). This is true across all therapeutic areas so that anyone who is familiar with the structure can quickly navigate and find what they’re looking for within any compound. The folder structure for a lupus study will have the same organization as that for an Ebola study and as that for a cancer study. Anyone requiring the objects (programs, outputs, documents) for a study need only know the compound name in order to find the correct folder.

As stated in the introduction, the Data Management group was already using LSAF and consequently the SDTM data was already being stored in LSAF. In the past they had downloaded the SDTM data to the windows environment to make it available to the statistical programmers and biostatisticians. But going forward, the statistical programmers would do their SAS development in LSAF where their SDTM data was already located. Hence, SDTM downloads would no longer be required because everyone would be working in the same environment. This was one of the “pluses” of moving to LSAF!

Here's the rub: the folder structure for the SDTM data in LSAF is not the same as the folder structure that the statistical programmers used in Windows. There is no compound name identifying each study folder, instead there is a random 4-digit number, representing the clinical development plan. Since the SDTM data was already stored in LSAF, it was decided to keep the folder structure intact to minimize rework. The folders for the stat programmers had to be reorganized to work with the Data Management structure already in place. Thus, the Data Analysis study teams had difficulties finding their "cheese", i.e. their study folders. No big deal, right? Wrong!



Imagine you are blindfolded and dropped off in a new town and told you must find the grocery store. You have no GPS signal and there's no one to ask. This was how the Data Analysis study teams felt about the prospect of looking for their study folders and files in LSAF. Not only were they to learn how to develop and execute programs in LSAF, they had to figure out a way to find their studies in LSAF.

I DON'T KNOW WHAT I DON'T KNOW

Data Analysis study teams had to learn how to develop jobs in the new LSAF environment. Jobs are wrappers for the SAS programs providing the programs with all input and output folder locations and job parameters. The Data Management group used the same few well-tested jobs over and over to process data transfers, but this was an entirely new concept to the Data Analysis group. When the Data Analysis group started using LSAF to develop jobs they didn't know exactly which folders they would really need so they sometimes specified more folders than they really needed. This was not a good practice because it caused some LSAF contention issues.

I CAN'T SEE YOU

Another change that took place in the move to LSAF was no longer being able to see other's SAS programs in development. You see, the LSAF environment is divided into two storage areas:

1. The Repository. All objects (files and folders) for all studies are stored in the repository and the access is read-only.

2. The Workspace. Each LSAF user has their own workspace and they alone can create, view and modify the files in their individual workspace. All development and unit testing are done in the workspace, then “checked in” to the Repository where others can see it.

By contrast, in the Windows environment everyone can see everyone else’s programs, whether in development or in production. This makes it easy to copy programs or code examples across studies. This is particularly true for the ad-hoc analyses performed because those typically would not be moved to the production environment. Consequently, new LSAF Repository/Workspace architecture was cumbersome for them.

LESSON LEARNED

The Data Analysis associates should have been introduced to the new folder structure prior to the LSAF training. In every training class we discussed the same question: “How will I find my study in LSAF?” Similarly, the Data Analysis associates would have benefited from an introduction to the concept of a “job” and how it’s used, prior to their LSAF training. Finally, the LSAF architecture of Repository/Workspace should have been shared prior to the training although it’s a bit of an abstract concept until you’re actually working on and manipulating files in LSAF.

TIMING IS EVERYTHING

The LSAF classes in the US were scheduled multiple times each month at Janssen’s locations in New Jersey and Pennsylvania, plus there were Web conferencing training sessions available for remote workers outside of the NJ-PA area. It was mandatory that associates attend the training in order to get access to the LSAF production environment, and the training took a full day to complete.

The “drop dead” date for the decommissioning the Windows SAS servers was published and known to everyone. Additionally, the individual cutover dates for every clinical trial moving to the LSAF environment were published so all Data Analysis associates knew their personal deadlines for completing the LSAF training.



There’s an expression “If you want something done, ask a busy person to do it”. This could be said for the individuals who attended LSAF training because all were working on studies with hard submission deadlines. Consequently, they signed up for training as early as possible to avoid conflicts with future submission dates.

Unfortunately, there were many issues encountered when migrating studies from the Windows to LSAF environment and as a result the date to decommission the Windows SAS servers got pushed out multiple times. Those individuals who attended training early on had little recollection of the specifics of the training because they didn’t get to use LSAF right away, and many needed to attend training a second (and third!) time so they would be ready when their studies were migrated to LSAF.

LESSON LEARNED

Ideally the LSAF training should have been scheduled “just in time” so that it was fresh when the associates needed to use it. But with the migration dates and Windows SAS decommission date slipping multiple times, that would have been difficult to plan.

Alternatively, Janssen could have developed **refresher** training, such as YouTube videos, targeted for those who had already attended the full day classroom or Web conferencing training. And this should have been done as soon as the original Windows servers decommission date was delayed. With the refresher training, individuals could select the topics they needed to review and watch when it was convenient for them. There is a steep learning curve to the LSAF navigation, and the easiest way to review that navigation is to watch it being done.

TOO MUCH INFORMATION

In their article “How to Conduct an Effective Training Session (Simplify Compliance LLC)”, the authors present 12 steps to accomplish the goal of making training enjoyable and engaging. The first step is:

1. **Tell trainees what you’re going to cover.** Introduce your session with a brief overview of the training subject’s main points.

The LSAF training was conducted in a single full-day session. The thought was that it was better to get it done in a day to minimize the time that the attendees had to spend away from their clinical trial work. Below are the main points that the agenda covered (take a deep breath):

1. What is LSAF?
2. Why LSAF?
3. New programming process in LSAF
4. New folder structure
5. Differences between Windows SAS and LSAF SAS (essentially the differences between wlatin1 and UTF-8 encoding)
6. LSAF jobs (batch submission of programs)
7. How to logon
8. How to navigate
9. Hands-on exercise



The hands-on exercise alone took three hours to complete. It was comprehensive, that is, it started with creating a new program in an individual's workspace, then covered navigation between the workspace and repository, check-in & check-out from the repository, version control, job creation, job execution and audit trail history. While the classes were kept relatively small to allow for a good instructor-student ratio, the amount of material covered was mind-numbing.

LESSON LEARNED

In the 12-step process we referenced earlier, there was an item further down the list of 12 steps:

9. Reuse successes in the training and incorporate those techniques **in future sessions**.

We should have embraced that bolded text and developed future sessions, i.e. modular training. It would have been easy to break up the agenda items into separate modules so each **future session** picked up right where the previous training left off.

DID ANYTHING GO RIGHT??

You would think at this point that the LSAF training was a failure. Actually, the Data Management LSAF training was and continues to be successful. And despite the pitfalls, the Data Analysis associates were satisfied with their training, just not the LSAF development environment. Below are some of the highlights of what went well:

LATHER, RINSE, "REPEAT"

The Data Management training focused on how to use objects that had already been created in LSAF to process data transfers. The instructions shared in the training specified how to "repeat", i.e. schedule, the programs and jobs already developed, optimized and verified in LSAF. The training primarily targeted one role, the Data Managers, so the attendees tended to be a more homogeneous group compared to the Data Analysis training attendees.

NAVIGATION CHEAT SHEET

The trainers encountered the same learning curve for the LSAF navigation as their Data Analysis students. There were hands-on exercises built into the PowerPoint slides used for the training classes, but it was difficult to follow so the trainers created a "cheat sheet" for themselves to get up to speed on the LSAF navigation. They realized that this would be helpful to the students during and after the training and that cheat sheet was shared and became a key part of the training.

LOCATION, LOCATION, LOCATION

For various reasons, the training and use of LSAF was met with more success outside the US. The pilot study for LSAF was conducted in Europe so they had more "skin in the game" and that resulted in more enthusiastic use of LSAF.



The associates in China were also pleased with the use of LSAF, although it had more to do with geography than the training. They experienced better response time in LSAF, which is cloud-based, than

with the Janssen's Windows SAS servers. That was due most likely to the physical location of the Windows SAS servers which were located on the East Coast of the US.

CONCLUSION

In 2017 Janssen decided to halt its plans to move its Data Analysis activities to LSAF. However, the lessons learned described in this paper are being incorporated into the training plans for the new statistical computing environment currently being developed at Janssen. The Data Management training was improved over the last few years to include several of the lessons learned covered in this paper such as modular training, refresher/recorded training sessions, slides available to all on the LSAF Portal and "just-in-time" training.

REFERENCES

Johnson, S. 1998. *Who Moved My Cheese*. New York, NY: Putnam Adult.

Simplify Compliance LLC. "How to Conduct an Effective Training Session." Accessed April 12, 2019. Available at <http://trainingtoday.blr.com/article/how-to-conduct-an-effective-training-session/>.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Margie Merlino
Janssen Research and Development
mmerlin2@its.jnj.com

Jeanne Keagy
Janssen Research and Development
jkeagy1@ITS.JNJ.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.