

SPACE – Launching GSK Biostatistics towards new horizons

Eileen Ching, GSK, Collegeville, USA
Caroline Phares, GSK, Collegeville, USA

ABSTRACT

In today's data-centric environment, having a modern fit-for-future technical landscape for statistical computing and data science is a critical success factor in the pursuit of bringing the best medicines and vaccines to patients faster and at a lower cost. Like many in the industry, GSK embarked on a journey to design and implement such a landscape in 2020, in the form of a large and ambitious program called SPACE (Statistical & Programming Analytics Computing Environment). Unlike the traditional approach to programs like this, GSK chose to adopt an agile delivery framework which also meant a shared learning curve. The aim of the program was focused on delivering capabilities for interactive data visualization, a modern Statistical Computing Environment (SCE), and planning, publishing, and sharing of results. SPACE is seen as the key enabler for transforming Biostatistics at GSK. We will share the story of our journey as we are nearing the conclusion of the program, including the challenges we faced, the goals we accomplished, the lessons we learned, and how we are planning to continue the journey.

INTRODUCTION

Prior to 2020, biopharmaceutical research and development (R&D) had undergone significant changes resulting from a combination of technological and scientific advancements. Ever-increasing volume and variety of electronic data were becoming more and more prevalent, and data science and machine learning gained a clear footing in biopharmaceutical research. Technology had also progressed, with a steady move toward cloud computing, agile methodology, DevOps, and the use of open source analytical tools.

It was at this juncture that GSK's Pharma R&D organization decided to launch the SPACE (Statistical & Programming Analytics Computing Environment) program to modernize clinical trials statistical analysis and reporting. The first major change came shortly after the launch of the program, when GSK announced the integration of its Pharma and Vaccines R&D units, and the split of its Consumer Healthcare organization into a separate company. Within a few months, the COVID-19 pandemic struck, and the need to analyze and glean information from data more quickly and efficiently became a necessity for the global biopharma community, not only to meet the demand for COVID-19 vaccines and therapeutics, but also to adapt already ongoing clinical trials to the changes in healthcare brought about by the pandemic. Concurrently, TransCelerate BioPharma, Inc., a nonprofit consortium, embarked on an initiative to develop a framework for Modernization of Statistical Analytics.

These environmental changes and progress in the industry required the SPACE program to quickly adapt and embrace change as a means of delivering the greatest value for the investment and staying at the forefront of progress throughout the lengthy, multi-year program. In this paper, we will depict how embracing change allowed us to adapt to the COVID-19 pandemic, re-evaluate and adapt our plan for delivering a modern SCE, and incrementally build a solution which would allow GSK to take full advantage of a flexible modern data science platform to power clinical trial analysis and reporting workflows, as well as provide farther reaching data science and analytical programming capabilities, thus benefitting additional user groups and use cases across GSK R&D.

OUR JOURNEY

SPACE PROGRAM OBJECTIVES

In early 2020, faced with an ageing SCE and an industry-wide push toward the use of modern biostatistics and data science in clinical development, GSK's Pharma Biostatistics organization, in partnership with GSK's R&D Tech organization, embarked on a three year program called SPACE, to transform Biostatistics capabilities through the implementation of a new platform which would enable process simplification, multiple statistical programming languages, truly scalable compute resources, process automation, integration with many data sources and data types, and compliance with regulatory requirements. The goal was to enable the Biostatistics organization to fully leverage modern technologies and methodologies for more powerful, more efficient analytics and reporting, achieving a fit-for-future transformation of a Statistical & Programming, Analytics Computing Environment to support clinical statistics and programming, non-clinical and research statistics, data science, and real world evidence teams. The main value propositions were to:

- Deliver Business Efficiency
- Provide Enhanced Capability
- Enable Faster Analysis & Reporting
- Simplify Ways of Working

Figure 1 below illustrates in more detail the business changes which SPACE aims to deliver and the benefits:

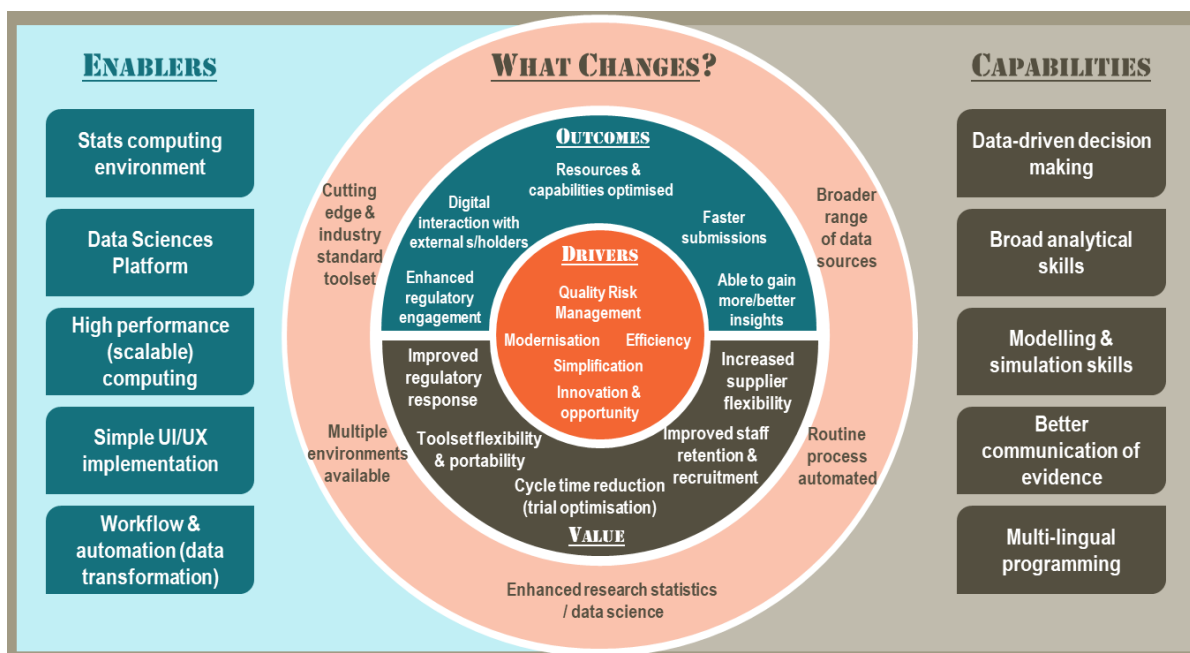


Figure 1

During the pre-launch, analysis stage of the program, the stakeholder scope quickly changed to include Vaccines Biostatistics as a result of the GSK Vaccines and Pharma integration. Although both are part of the same company, Vaccines and Pharma business units have traditionally operated independently from each other with their own systems and processes. In order for the integrated Development organization to achieve efficiency and top level performance, it was critical to harmonize the business processes as well as the tools and technology from the two business units, and SPACE was seen as the vehicle to deliver a single, unified digital data and analytics platform for the new GSK R&D, and to enable our teams to use the right tool for the job.

Achieving world-class performance requires a world-class organization with a modern and agile workforce with a growth mindset. Although not a primary objective, another important goal of SPACE was to help GSK develop and retain our existing talented staff and attract new talent including those new to the industry by giving them the environment where they can become productive and innovate without a steep or extensive learning curve.

SPACE PROGRAM DELIVERY METHODOLOGY AND STRUCTURE

During this time, there were also several changes being launched in GSK's IT organization. R&D Tech, our Business-facing IT group, was initiating a fundamental change, to transform into a product-oriented, DevOps organization structured around the Scaled Agile Framework (SAFe®) [1]. Simultaneously, the global infrastructure group was setting out on an aggressive and disruptive push toward moving infrastructure from on-premise servers to third party cloud providers. These changes involved significant impact to people, processes, and tools.

Given the high level of complexity and the significant scope of SPACE, we decided to work with a number of strategic partners that specialized in different disciplines, such as business analysis, UX design, software development, systems integration, and SAFe® delivery. Based on our plan for delivering a number of products, we organized the program into product-centric multi-disciplinary teams working closely with a group of Business Owners accountable for setting business priorities, articulating business benefits, and driving change management. The program is co-led by a Business Lead and a Tech Product Lead via a collaborative partnership, enabled by the SAFe® Agile Release Train (ART) construct (see Figure 2).

The SPACE Agile Release Train (ART) Schematic

How the ART is constructed indicating the focus horizon each construct operates at

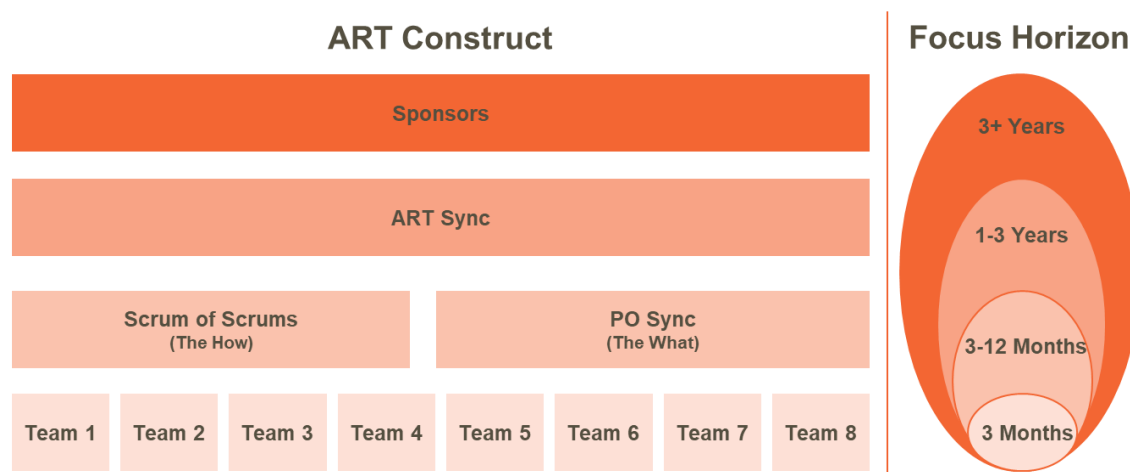


Figure 2

Following the SAFe® model, SPACE products are developed and delivered via Program Increments (PIs) which are 12-week periods during which we plan, do, inspect, and adapt. The aim is to develop and deliver in an agile way with an emphasis on business-centric solutions. Instead of designing a complete solution that we think the users need, we start simple and elaborate the design incrementally, delivery by delivery, PI by PI. This allows feedback and features to emerge based upon value and priority. It means truly understanding business needs and performance goals and being able to fulfill and enable them starting with minimum viable products (MVPs), then continuously enhancing and expanding functionality.

WORKING GLOBALLY IN A PANDEMIC

SPACE was officially launched in February of 2020, just before the global COVID-19 pandemic emergency was declared. In addition to a new set of risks related to illness, the daily lives and work patterns of program team members and stakeholders were disrupted and permanently changed. We formed a global team consisting of members across multiple countries, continents, and time zones, which presented many challenges especially given what the ideal setup should be for teams working in an Agile framework. Virtual collaboration, meetings and whiteboard tools became essential to our ability to communicate, collaborate with each other, and innovate together.

As a result of the COVID-19 pandemic, any possible face to face interactions between team members based on the same site suddenly became impossible. We had to abandon our plans for doing in-person PI planning, and instead resort to virtual multi-day events. Virtual meetings early in the morning and late into the night became the accepted practices. While the fast pace of Agile and not having a co-located team continue to present us with challenges, over time, SPACE team members learned to flex, adapt, and accommodate each other, and became more resilient from Sprint to Sprint, PI to PI.

SOLUTION DIRECTION: BUY VS. BUILD VS. COMBINATION

Presenting multiple solution options is expected as part of the project approval process at GSK. Scanning the market for a commercial off-the-shelf (COTS) and/or software as a service (SaaS) product is typically considered first, prior to considering a custom-built solution. When SPACE was initiated, the program leadership team along with subject matter experts from the Business and IT, conducted a market scan for COTS and SaaS SCE products. These solutions typically address the needs of clinical statistics and programming, and data management. However, the scope of SPACE was broader and included delivering solutions for non-clinical and research statistics, data science, and real world evidence teams as well. There was still a need to evaluate additional solutions to meet those needs and this was done in parallel.

Several well-known COTS and SaaS SCE solutions were evaluated and compared with GSK's current custom-built SCE used in Pharma R&D called HARP. The effort was concluded with a result that no product on the market which could provide GSK Pharma Biostats with the same or better capability than HARP, and that a custom solution would provide a greater value for the investment and would truly transform our business.

With our ways of working forcing us to prioritize our focus by the associated return on investment, we agreed to focus on the creation of a new suite of custom-built products for which we were certain no COTS solutions existed. We began to build these products, and the framework for SPACE-delivered solutions, on GSK's Microsoft Azure cloud

environment, which was available to us at the time. Our framework evolved to include Azure Kubernetes Service (AKS), SQL Server, Azure file storage, Azure Data Lake (ADLSv2), Apache Kafka®, Docker, along with other technologies. We used Azure DevOps (ADO), not only to document our requirements, but also to drive our CI/CD pipelines (see Figure 3), in conjunction with Azure GitHub to manage our code and design documentation, and Selenium Automation for test automation.

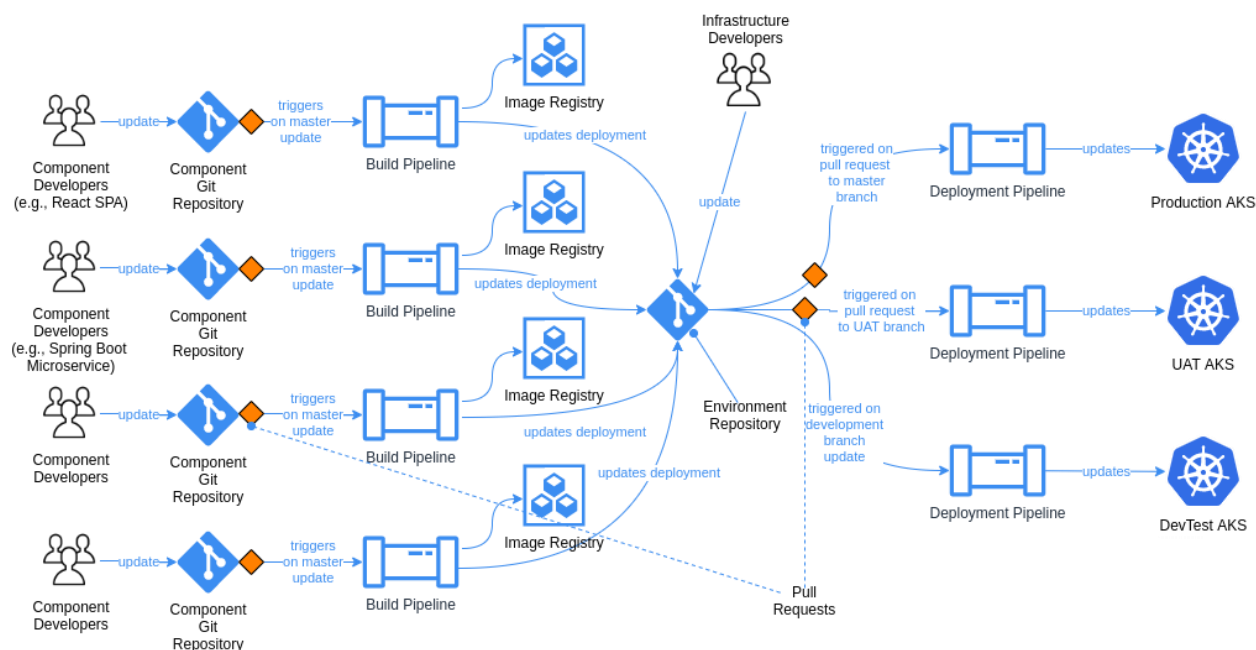


Figure 3

One of the first products we began to develop is called RAPIDO Data Viewer, which provides a library of standard interactive R Shiny visualizations of ADaM data, available for any study. This product provides medical writers and other data reviewers with self-service capabilities (such as dynamic filtering and sub-setting and drill-downs) to explore study data without relying upon statisticians and clinical programmers to make changes to static displays, in addition to eliminating all non-core listings from clinical study submissions. RAPIDO Data Viewer also provides clinical statisticians and programmers with the capability to create and deploy additional standard R Shiny visualizations, which are referred to as RAPIDO Dashboards, for immediate use by medical writers on any study. User access to the RAPIDO Dashboards is based on the role of the individual on that study. RAPIDO Data Viewer makes use of SQL Server to provision data to the R Shiny visualizations in a relational database format in order to optimize performance, since ADaM datasets are typically delivered in sas7bdat format, which is not conducive for interactive review and exploration.

Another custom-build solution we prioritized was focused on efficiencies and metadata-driven automation of planning, reviewing, and publishing of statistical results. The aim of this product, called PULSAR, is to allow clinical scientists and medical writers to collaborate more closely with statisticians and clinical programmers throughout the statistical analysis process, eliminating time and effort-consuming rework resulting from hand-offs at the end of the process, and ultimately minimizing the time between database freeze to the publishing of statistical results and clinical study reporting. PULSAR makes use of containerization for the creation and publishing of bundled PDF outputs, in order to optimize use of cloud compute resources and minimize cost.

Once Agile delivery of RAPIDO DV and PULSAR was established and stable, the main focus was once again directed toward deciding on an SCE solution. We agreed with the sponsors of the program that we would focus on clinical GxP workflows as the priority use case for the SCE solution, but that we would keep line of sight and applicability to other use cases within the scope of SPACE as a factor in decision-making. We conducted a fresh market scan, which resulted in a deep-dive evaluation of four potential solutions, one of which was another look at the Domino Data Lab's Enterprise MLOps data science platform ("Domino"), which had not been marketed as an SCE solution at the time. We were initially impressed with the ability it provides data scientists and statistical programmers to collaborate on a project, using their tool(s) of choice: SAS®, R, and/or Python, a requirement which was of great importance in all user groups in the scope of SPACE. Even though there was no explicit, built-in SCE workflow and data pipeline, the flexibility in potentially building multiple workflows, serving multiple use cases and user groups, was a compelling value driver to warrant consideration as our SCE solution. Based on our strategic direction of a modern, future-proof SCE and the widely accepted definition of SCE within the industry [2], we put together a short list of

must-have requirements representing both the Business and IT, and conducted a Proof of Concept (POC) with Domino to determine how many of the requirements could be met out-of-the-box and whether or not a simple bespoke solution was feasible for those requirements which were not immediately met. In order to determine how much additional value could be achieved through Domino, we also opened the POC to other user groups within the SPACE scope, as well as those outside the scope, who were looking for similar programming and compute capabilities. We simultaneously conducted evaluations of the other three COTS and SaaS SCEs, in a series of deep-dive demonstrations of how the same requirements would be met. If a solution based on Domino could not meet our SCE requirements, we would be able to select an alternative solution for our SCE.

The Domino POC included a set of functional and technical success criteria. All but four of the criteria were met as expected out-of-the box. For those which were not, either a bespoke solution to meet the need which incorporated the Domino workflow, design, and capabilities was determined to be possible, or we identified an alternate out-of-the-box solution within Domino which met the true need which the criterion was intending to fill. These four criteria are listed in Table 1.

Criterion	Solution
Showing dependency/traceability between datasets and between datasets and displays	Domino provides overall traceability between input data, executed code, and output. It also provides the ability to take and label snapshots of the environment, code, input and output data, showing that traceability and full reproducibility. This was found to be sufficient, but we are building a solution which performs GxP execution of code with incorporated automated snapshots and tagging of the environment, ensuring this workflow is conducted in a controlled fashion every time.
Content of audit trail (who, what, when, etc.) that meets GxP regulations	Although user activities are logged, we did not interpret the log to meet the ALCOA standards. Our integrated custom solution will capture the required audit trail.
Ability to build workflows and process automations	We built our own workflows using a combination of Domino integrated with GitHub Enterprise, which provides much more robust capability than we have found to be available in either our current SCEs or those we evaluated. Domino also integrates with Jira, which provides us with flexibility to enhance and evolve our workflows.
Ability to customize the interface (e.g., hiding features/menu items/choices, etc.)	Domino does not currently allow hiding of menu options through configuration; however, we are able to control and restrict users from carrying out specific tasks. We have asked that this feature be incorporated into the Domino Data Lab backlog as a high priority item for GSK, which would enhance the user experience tremendously for certain user groups.

Table 1

After a short and sharp POC, we determined that our SCE solution would be powered by Domino for the following reasons:

- Domino provides capabilities and flexibility beyond traditional SCEs
- Domino provides a uniquely versatile platform which effectively leverages cloud native technology to meet data science and programming needs beyond just analysis and reporting for clinical study submissions.
- Domino includes a rich API interface which allows us to leverage metadata to build automations, effectively manage access and sharing of information, and execute control from outside the system.
- Domino meets the needs of additional use cases beyond clinical study analysis & reporting, such as R Shiny application development and publishing, and other data science and modeling workloads

We especially liked the following Domino features:

- Quick and easy setup of projects
- Flexibility in how to use projects to design workflows
- Ability to run multiple programming languages with consistent look and feel
- Working with SAS® Studio in Domino is easy and fast
- Can have a controlled GxP R environment along with a more dynamic environment for exploratory work
- Versioning and snapshots for full reproducibility of data science activities
- R job performance is comparable to our existing on-prem R environment
- Powerful search functionality that displays results in a useful way
- Granting access and collaboration with others is simple and easy
- Ease of publishing apps (e.g., R Shiny, Markdown)
- API is broad and allows for easy integration and enhancing functionality

- The ability to access different types of data sources, internal and external to Domino

In keeping with the SPACE theme, we named our Domino-powered SCE solution METEOR (Multilingual Environment to Transform, Enable, and Optimize Reporting). METEOR will have three standalone custom components:

1. METEOR Project Portal - allows users with a particular lead role on a study to create a secure Domino analysis project for that study, load the sensitive data into it, and administer access to those who require it in order to fulfill their duties
2. METEOR Mission Control - provides users with the ability to execute their team's developed and quality controlled code to create GxP compliant output in a secure, controlled, and audited manner
3. METEOR Metadata Management - allows users to create and manage their metadata specifications from standard templates, and will allow programs executed within Domino for an analysis, to access the relevant metadata to create outputs. It will also allow metrics and trends to be captured and analyzed for the purposes of continuous improvement and identifying opportunities for efficiency.

The following diagram (Figure 4) illustrates the SPACE ecosystem in the context of the high level business process of preparation, conduct, and dissemination of analysis results.

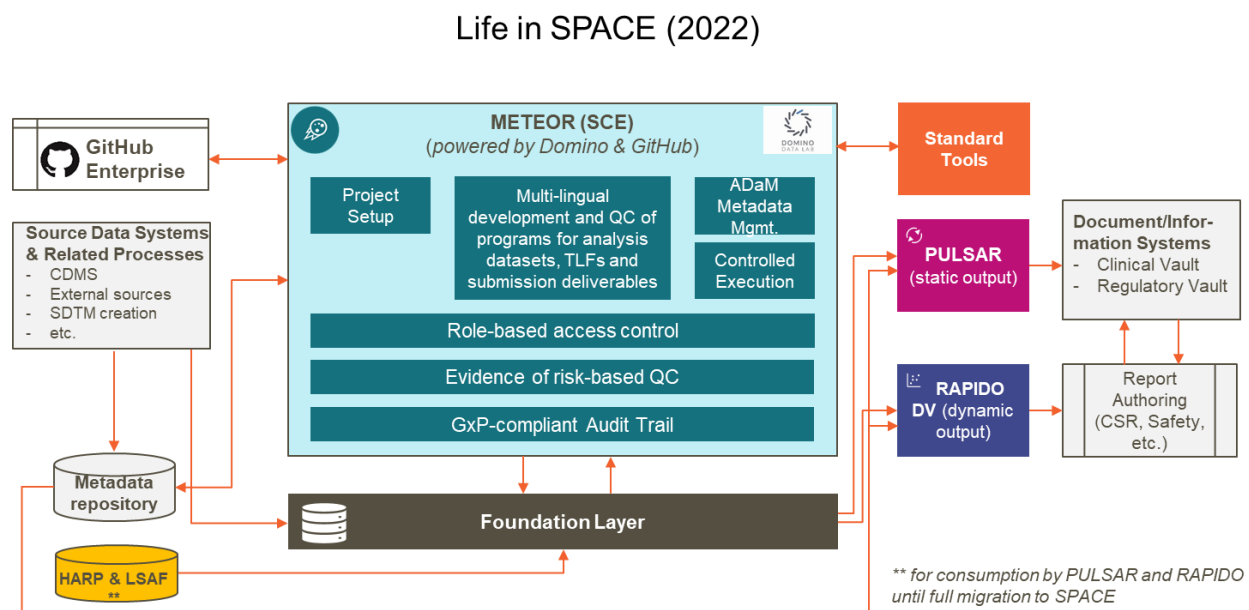


Figure 4

Rather than having been planned from the beginning, our resulting solution design for SPACE has emerged into an extendible, flexible, and versatile platform with standalone modules which follow a particular pattern.

KEY LEARNINGS

THE RELATIONSHIP BETWEEN BUSINESS AND IT

Before we launched SPACE, a Waterfall software development methodology was generally used in projects at GSK R&D, with specific handover events, such as the Business providing an approved set of requirements to IT, and IT providing a fully developed system to the Business for user acceptance testing. This type of relationship often resulted in IT delivering solutions that did not fully meet the business needs or the Business creating their own technical solutions that did not scale or adhere to GSK IT's standards.

Consistent with a Waterfall approach, during the pre-implementation phase of SPACE, GSK engaged a vendor to coordinate and conduct a series of requirements gathering workshops and stakeholder interviews. It was during this time that the program scope increased to include Vaccines Biostatistics, expanding the number and types of stakeholders, the number of existing legacy systems and processes, and the number of technology-related pain points, therefore increasing both the time needed to conduct the requirements analysis and the number of requirements in the list. The resulting requirements were often solution-oriented, tied to the legacy systems, and without clear business problems or business drivers associated with them. Clear measures for how we would know a requirement had been met were also lacking.

It was during this time that SPACE was identified by R&D Tech as an early adopter of SAFe®. The R&D Tech organization had already been structured to be product-oriented, with Product Owners as the accountable party for the vision and roadmap for each product, as well as the product backlog of desired features (requirements). Microsoft Azure DevOps was adopted as the tool for managing this backlog. The challenge with fitting a product-oriented model to a program such as SPACE was that distinct products did not exist yet. The vastness of the SPACE scope, its variety of stakeholders, the multiple legacy products which were in scope for replacement, and the solution design which should be shaped by applying modern, cloud-native architecture patterns, made it difficult to assign immediate ownership to non-existent products. The role of the product-aligned Business Owner also created confusion in terms of ownership and accountabilities. These would eventually, naturally emerge from a true Agile approach to the program. Because there was a substantial learning curve, as well as a reluctance, in some cases, to adopting an Agile mindset and redefining the relationship between Business and IT, there were significant struggles in simply getting off the ground.

Until the relationship between Business and IT could be transformed to one based upon implicit trust and mutual ownership of success and failure, the program was at a standstill. As leaders of the program, we recognized that this trust-based partnership needed to start with us, and only in addressing our teams as a united front could we transform the culture to one which would allow us to achieve the program goals. We, as Business and IT leaders, needed to see ourselves as equally accountable partners in the success or failure of the program. Together, with the support of our Business and IT Sponsors, we aligned on a vision statement, a mission statement, and clear set of Business value drivers which would guide our teams' priorities, in the form of a jointly owned roadmap, and we made the commitment to conduct our program by the principles of Agile methodology. We structured our program organization together, in a way that allowed and encouraged teams to self-organize around their work, and to organize their work around clear priorities.

Throughout our journey together, we have learned that:

1. A true partnership between Business and IT leadership, based on trust and respect, with a common vision, mission, and measurable Business value drivers, permeates all levels of a program and is essential to executing a successful program.
2. Self-organizing teams of Business and IT, guided by clear goals and principles, are essential to efficiency in design and delivery of solutions.
3. Clear goals and laser focus on measurable value in prioritization and decision-making at all levels are essential to a healthy Business-IT relationship.
4. Excellent solution designs emerge from Business and IT working closely together, with equal ownership and stake in the success of the solution.

AGILE DELIVERY CHALLENGES

Because of the Agile transformations initiated in R&D Tech, the Tech members of SPACE were equipped with at least basic level knowledge of Agile and SAFe®, some with a significant level of experience, while others being relatively new to the practice. The Business members of the team on the other hand, had no prior training in Agile and most were accustomed to the traditional Waterfall method of delivering systems and software. The impact of the gap was quite evident as our first couple of PIs did not achieve all the objectives. We therefore made a considerable amount of effort to train the entire team on SAFe® and related tools, added more rigor and structure to the scrums, and enlisted the help of a dedicated Agile coach.

Clarity in roles and responsibilities was key to ensure the team members were confident in their accountabilities and deliverables. As the team was onboarded to SAFe® and learned to scrum, we felt that it was important that everyone understood their role on the ART in addition to the Agile ways of working. In our effort to continuously adapt and improve, we evolved the team structure of SPACE from one that tried to closely follow the theoretical SAFe® model to one that we felt was fit for purpose and practical for our environment and needs, over several iterations (in the true spirit of Agile). The team members shared challenges and best practices with each other during retrospectives so we could take the learnings into the next Sprint or PI. Eventually, for the vast majority, roles and responsibilities became more fluid and natural within the teams, which self-organized around their common goal. It became more common practice that the best person for a task was the one to take it on regardless of their job title, since the success or failure of the team was fully and equally shared amongst the team members.

One lingering challenge we had to overcome was around setting expectations with our users and key stakeholders. Although the SPACE team was trained on Agile, as an organization, GSK was still relatively immature with regard to the idea of incremental delivery. We had to rigorously and continuously emphasize our approach of starting simple with iterative enhancement/refinement and educate our users on Agile, as our user community and stakeholder groups of SPACE consistently had higher expectations than what we were able to deliver in each PI. However, incremental delivery proved invaluable to understanding the true needs of the user community and including their feedback as we built the solutions, rather than waiting until we had delivered too much to efficiently and effectively change. We learned that having a UX Designer on a particular product to conduct usability testing and user feedback

sessions was tremendously useful in not only guiding product strategy in terms of usability and user experience, but the change management strategy as well.

Another major challenge we faced was how to meet regulatory compliance requirements as part of delivering a validated system, with Agile. The GSK IT Quality Management Framework originated from the traditional software development methodology and has not fully adapted to Agile. As a result, a large proportion of the capacity in each PI was allocated to and spent on documentation. This significantly slowed down our pace of delivery because the teams were only able to complete a limited number of features within each PI. As GSK Tech is building more integrated documentation capability into our Agile delivery tools, we are continuing to work with our Computer Systems Validation and Quality experts in IT and Business to simplify our GxP validation and documentation framework.

Finally, beyond tools and processes, one challenge SPACE has experienced since the inception of the program is resourcing, including both IT and Business. For IT, it was mostly around the specific skillset and technical expertise needed for the targeted solutions. For the Business, however, it was the fact that we didn't have many team members whose primary job was working on SPACE. The majority of the business team had a quarter or less of their capacity dedicated to SPACE which presented an ongoing risk to the program as they could at any moment pull back their time on SPACE if demanded by their pipeline project work. We tried to mitigate this risk by increasing the size of the team, but having a smaller but dedicated team would have been the ideal. Furthermore, the limited capacity of the business team overall is compounded by the dispersed locations and multiple time zones, creating significant hurdles to Agile ways of working.

PLANNING FOR AND EMBRACING CHANGE: PROGRAM RESILIENCE

As the SPACE team was embracing multitudes of changes, such as Agile vs. Waterfall, new DevOps tools, new ways of working between Business and IT, different team structures, and changing business processes, we also had to focus on helping our target user community, R&D Biostatistics, understand, prepare for, embrace, and adopt change. The highest priority goal of SPACE is to replace the current systems and tools used for clinical analysis and reporting with a new suite of fit-for-purpose and future-proof products, so GSK maintains license to operate and is able to accelerate our speed of execution without compromising quality and compliance. Although users from the two business units coming together are using different platforms and tools, HARP for Pharma and LSAF for Vaccines, they all had high expectations for what SPACE was going to deliver and when.

The initial steep learning curve of Agile already slowed down our velocity of delivery, and the discomfort with the PI process and associated short time horizon also made the change management planning and execution difficult. As such, there was very little information being communicated out, and when there was a release, the team struggled with disappointment management. The communication gap also plagued the SPACE team internally, as the team grew larger in size and more new members were getting onboarded. With a change in business leadership and acting on the feedback from both SPACE team members and our external stakeholders, we took a number of concrete steps to improve communications internally and communications/engagement externally with those who would be impacted by SPACE. We created a program roadmap that succinctly described the delivery plan and targets to both the program team and our external stakeholders. We implemented a monthly intra-program newsletter focused on program decisions, progress made, best practices, and team member spotlights. With our user community, we started the publication of a monthly newsletter for Biostatistics, so we kept them informed of program progress and engaged in key messages every step of the way. We started posting more updates, stories, and videos about SPACE, on GSK's company communication and employee engagement platform, and held townhall style virtual meetings that gave users the opportunity to learn more. Due to the restrictions imposed by the pandemic, we really had to resort to various modes of virtual communications to bring our users on the journey along with us. Following GSK's REAL (Relevant, Empathetic, Aligned, Lasting) principles and the Prosci ADKAR model [3], we built a new Change Management plan that served as a playbook for us to:

- communicate with purpose and regularity
- be open and transparent
- invite engagement and participation
- combine early awareness with just-in-time training
- monitor and measure

SPACE not only means new ways of working, but also different ways of thinking and to certain extent, major cultural shifts (e.g., risk averse culture needs people to think outside the box and envisage a future less dependent on controls). In anticipation of the SCE MVP deployment and major ramp-up of adoption of other SPACE products in 2022, with the strong support of the SPACE Sponsors, we brought on a larger change management team who are helping us with adoption and sustainability, communications and engagement, and training and capability. No longer a reactive measure, change management is now truly and fully integrated into our PI planning and delivery.

Just like the fact that medicine doesn't change lives until patients know about it, trust it, and take it; SPACE won't change our business until users know about it, trust it, and use it.

PLANNING FOR AND EMBRACING CHANGE: SOLUTION RESILIENCE

Progress in software development over the last ten years has removed design constraints which used to exist in legacy frameworks. Cloud native architecture provides not only greater flexibility, but also allows for easy, quick, low-cost, low-risk deployments of functionality, bringing value to the end user more quickly and frequently. SPACE pushed vigorously to make full use of cloud computing and cloud native architecture, as GSK's IT infrastructure teams were working hard to build services to support the new paradigm. We made the commitment at the beginning of the program to embrace the new architecture patterns, regardless of whether a central services framework was in place. We found that it was the correct decision, because in addition to enabling development to start small and progress, it also made it possible for SPACE to influence greatly what the central services needed to offer. This essentially allowed the central services to be developed out of a true set of prioritized needs, in line with Agile delivery.

Although, initially, these services were non-existent, our business SMEs and user representatives were keen to begin designing and seeing solutions. At this time, the relationship between IT and the Business was strained. The Business, which was much larger in number at the time with respect to the number of developers, embarked on designing full conceptual applications in the form of clickable wireframes. However, IT needed time to start small, with simple, focused application UI designs to meet simple requirements, and spend the majority of effort on developing deployment environments, and building CI/CD pipelines. The IT team also needed to increase its number of developers to meet the demand. This foundation was quite time consuming and effort intensive to build without central infrastructure services ready for an aggressive move to cloud, and without a fully resourced IT team to meet the business demand. Seeing applications in their simplest form without much direct, measurable business value was disappointing to business stakeholders, and created a perception that the program was not delivering. However, once the foundation was established, the speed at which applications could be developed, tested, deployed, and updated, brought pride and confidence in the program. We found that starting simple and small with application UI designs, allowed them to evolve with more valuable feedback from user representatives testing and end users using the applications. We also found that user representatives' ideas often changed as applications evolved, and many of the initial UI designs were not needed in the end.

INDUSTRY TRENDS

Just as cloud native technology was driving IT design, development and deployment of solutions, there was similar technical progress in clinical programming trends within the industry. The use of GitHub, R, and Python were becoming more prevalent. Crowdsourcing and collaboration across this function began and is continuing to drive progress, along with the usage of other languages like R and Python, in addition to SAS®. Interactive visualization of clinical data using tools like R Shiny is seen as a means to reduce clinical submission effort and time. Additionally, the appetite for defining and solving problems using AI/ML is growing, and there is an increasingly closer relationship between clinical research and development. We found that the needs of the clinical programming community have been changing to overlap with the needs of the data science community, and the importance of data science in clinical development has grown even further since the launch of the SPACE program.

Regulatory Agencies and Health Authorities are increasingly open to analysis and reporting of clinical data using the right tool for the job including open-source, as more and more members of the biopharma industry are collaborating together to share best practices and design and implement solutions that will accelerate the delivery of medicines and vaccines. One of these nonprofit collaborative organizations, TransCelerate, recently developed a new framework for modern statistical analytics with adherence to key principles of regulatory compliance. According to TransCelerate, "by developing a foundational framework that increases Health Authorities' confidence in the use of modernized statistical analysis and efficient technology across the industry, an opportunity exists to provide clarity and instill confidence in the use of modern software technologies, including those that generate outputs that support regulatory submissions" [4]. TransCelerate launched the Modernization of Statistical Analytics (MSA) Framework in April 2021 including a set of resources which "were developed by the Modernization of Statistical Analytics (MoA) Initiative in an effort to accelerate the use of novel statistical computing capabilities when engaging with Health Authorities" [5].

Choosing a data science platform, such as Domino's Enterprise MLOps Platform, which provides the capabilities of a modern SCE, has enabled us to fully take advantage of this progress in the industry, and being able to respond to these changes and trends in our SPACE solution designs has been critical to staying relevant and modern, well over two years into the program.

CONCLUSION

Our SPACE journey over the last two years has been filled with trials and tribulations, remarkable accomplishments, and invaluable lessons. We have overcome many challenges and matured together as a team in areas of Agile ways of working, collaboration in global matrixed environment, partnership between IT and Business, setting realistic expectations with our stakeholders, managing mindset change with our users, experimenting with new and modern tools and platforms, and working with industry partners to achieve our common goals.

If we had to distill everything we learned down to a few critical points, they would be:

1. Fundamental understanding of Agile is important but we also need to be flexible and practical in our own implementation.
2. Collaboration and trust between Business and IT are essential to our success.
3. GxP validation framework must leverage and align with Agile in order for the system delivery to benefit from the value proposition of Agile.
4. The multi-dimensional nature of a transformational change requires proactive and robust change management for enabling the mindset shift and building technical capabilities.
5. We must think beyond the traditional SCE to deliver a truly forward-looking and future-proof modern data science platform.

Despite the unexpected from the COVID-19 pandemic, SPACE's launch was fueled by a combination of truly exciting drivers both inside and outside the company. GSK laid out ambitions to reach top-level performance while going through organizational changes. Our diverse user community of statisticians, programmers, and data scientists needed new modern and flexible systems that allowed them to use the right tool for the job and to do their best work. The strategic solution we forged was to provide them with a single and yet versatile platform that would meet both GxP and non-GxP needs. This additional complexity demanded a robust and effective delivery approach, and we embraced Agile as the methodology of choice for SPACE. The biopharmaceuticals industry as a whole was also embarking on new open-source statistical analytics framework to transform research and development. With our continuously evolving delivery model, matrix team collaboration, and strategic thinking, SPACE is poised to harness the readiness and eagerness of a user community already in the midst of significant change to create a new and robust pattern for the next twenty years of R&D at GSK.

REFERENCES

1. Scaled Agile Framework: <https://www.scaledagileframework.com/>
2. PHUSE SCE Whitepaper (2019-2021): https://phuse.s3.eu-central-1.amazonaws.com/Events/2021/US+Connect+2021/SCE_White_Paper_Final_15Dec2021.pdf
3. The Prosci ADKAR Model: <https://www.prosci.com/methodology/adkar>
4. TransCelerate Modernization of Statistical Analytics Initiative: <https://www.transceleratebiopharmainc.com/initiatives/modernization-statistical-analytics/>
5. TransCelerate Modernization of Statistical Analytics Solutions: <https://www.transceleratebiopharmainc.com/assets/modernization-of-statistical-analytics-solutions/>

CONTACT INFORMATION

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

Eileen Ching

Caroline Phares

Email: eileen.l.ching@gsk.com

<https://www.linkedin.com/in/caroline-m-phares/>

Brand and product names are trademarks of their respective companies.