

## **Change Management: A futuristic opportunity for SAS Programmers to build R Library**

Shashikant Kumar, Ephicity Lifescience Analytics, Bengaluru, India  
Mrityunjay Kumar, Ephicity Lifescience Analytics, Bengaluru, India

### **Abstract**

Pharmaceutical industries are rapidly evolving especially after Covid-19 pandemic. A race has begun to develop time-bound and cost-effective treatments for safeguarding patient's life. Retaining talents without compromising quality and cost has been challenging in recent time. Often, we hear the talks, that AI based tool such as ChatGPT gaining popularity in healthcare industry, and it may take over SAS programmers in clinical industry which appears to be a myth.

Leadership could focus on existing SAS Programmers building expertise on open-source software such as R, Python etc. to convert the library of existing standards programs developed over the years using R, Python etc. This novel approach allows SAS Programmers to build expertise using new technologies in more efficient manner and help pharmaceutical companies save cost on licencing which needs renewal on yearly basis. Discussion on some real time strategies would help leaders or SMEs to adapt the change more effectively.

### **Topics for Discussion**

#### **1. Change is Inevitable**

We have seen for past several years how pharmaceutical industries have evolved and adapted to cutting edge technologies in conducting research, collaboration and partnership with other companies, vendors, recruitment of contingent workers. These adaptive methodologies have shown proven results in speeding up research and thereby reducing drug development and approval time. As a result of this it was evident that pharmaceutical organizations were able to launch Covid-19 vaccines in a very short span of time which was indeed needed for mankind. Thus, it's time for us to build robust platform and implement it across organizations with help of technologies (R, Python etc.), it eventually helps employees such as Statistical Programmers who can efficiently leverage their skills to support various end-to-end activities in achieving key goals for their respective organizations.

#### **2. Adapting new Technologies**

As we know that in recent years adapting and implementing new technologies have been successful across pharmaceutical organizations such as cloud computing platforms Azure, AWS, SCE, R, Python etc. and providing opportunities to our employees should help them to be future ready. This will eventually provide us opportunities to explore more with these emerging technologies and also be competitive.

### **AI Era: ChatGPT**

In recent years, artificial intelligence (AI) has made significant strides in understanding and generating human-like text. One remarkable breakthrough in this field is ChatGPT, a conversational AI model based on the GPT-3.5 architecture developed by OpenAI. ChatGPT has gained widespread attention for its ability to engage in coherent and contextually relevant conversations with users, revolutionizing the way we interact with AI systems.

# ChatGPT Technology & its features



ChatGPT simulates human interaction, and it is probably the most powerful language model currently available. It uses a deep learning technique called a 'transformer architecture' that is trained on massive text datasets obtained from the internet.

Its applications may include customer service bots, creation of content for social media or blogs, and translation of text from one language to another and healthcare industry.

AI models may help in clinical decision support, clinical trial recruitment, clinical data management, research support, patient education and other fields. In some cases, AI models can automate certain tasks performed by humans, such as data analysis, image acquisition and interpretation. This may increase the efficiency and reduce the workload of healthcare professionals, allowing them to focus on higher level tasks that require their expertise and clinical judgement. On the other hand, the results must originate from the extraordinary human mind and its ability to process information and communicate through them.

Therefore, it is important to use AI models such as ChatGPT as a tool to support, rather than replace, healthcare professionals in their decision-making process.

At present, ChatGPT cannot update the training data in a real-time manner. In addition, it can only give general and vague answers in some existing medical-related conversations. Clinical studies are based on real time data and most often endpoints are newly derived results. Hence, human intervention such as new endpoints-based programming logic will need skilled SAS Programmers hence there would be no replacement. Therefore, researchers and healthcare providers must consider these factors when implementing ChatGPT-based interventions.

Overall, ChatGPT has the potential to revolutionize clinical and translational medicine, but we need to develop appropriate strategies to mitigate potential risks and negative outcomes. ChatGPT technology, the development of AI is unstoppable. The best course of action is to embrace it, use its capabilities to improve our lives, and foster mutually beneficial relationships by evolving it in clinical medicine.



### 3. Open-Source Technologies (R, Python etc.)

#### a. What are open-source technologies?

Open-source is a licensing agreement that allows users to freely inspect, copy, modify and redistribute a software, and have full access to the underlying source code. In simple words, it's nothing but the source code of software created by a developer is freely available to all users. On the other hand, "closed source" is a proprietary software that has source code only the creator can legally copy, inspect, and modify.

There are different types of open-source software's however in this paper we talk about R and Python and its widespread usage in statistical programming.

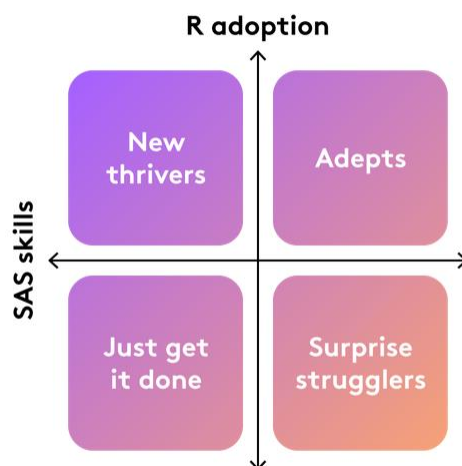
**R** - It is statistical software which is widely used in data analysis and visualization. R generally comes with the Command-line interface. R is available across widely used platforms like Windows, Linux, and macOS. Also, the R programming language is the latest cutting-edge tool which can easily be integrated with SAS.

**Python** - One of the most popular programming and scripting languages used by software developers across globe. It is powerful and fast, easy to learn and use, and runs on various platforms, which is why developers choose Python as one of the top open-source technologies available.

## b. Who can implement it?

Statistical Programmers with expertise on SAS and R programming skills can contribute to development of R libraries. A group of experienced statistical programmers can be assigned to build SAS → R libraries as a pilot project. After successful testing and implementation of R-Libraries, a study level initiative can be taken further.

**Figure:** A SAS Programming team would typically fit anywhere along either axis (shown below). These groupings are for illustrative purposes and offer practical guidance.



**‘Just get it done’** – This group uses SAS to achieve what they need to do on a daily basis, see it purely as a tool and don’t engage with complex/custom functionality. They will probably not be particularly excited about the prospect of having to learn something new. In addition, they may struggle with some training content, however once shown the practicalities of how to continue achieving what they currently do (often in tutorial setting), they will be OK with the process. These individuals will likely rely heavily on others in the community when faced with new problems. Encourage them to engage.

**New Thrivers** – This user group may not be SAS experts but see this as an opportunity to shine. They may put themselves forward to take on additional responsibility after the transition and can act as champions within your new R community, helping others along the way. Encourage them to present new ideas in community forums and share what they have achieved.

**Adepts** – These are the SAS wizards who will pick up R easily and run with it. They will want to stretch their understanding and explore what possibilities the new language unlocks. This group needs to be engaged to establish internal best practices moving forwards and will help shape the latter half of the migration effort.

**Surprise Strugglers** – SAS experts who could handle any task you throw at them may not be entirely sure why they are being made to learn something new, and not necessarily happy about it. They might struggle to step away from the SAS way of approaching a problem (e.g. sequential processing of data vs vectorised approach). It would benefit this group from being shown 'the art of the possible' early on to try and drive excitement in the new technology and encourage engagement with the process. Additional support such as code-pairing these people with Adepts who have gone through the same process can also help. If required, roles may evolve away from day-to-day code ownership tasks, but it is incredibly important to retain their knowledge of processes involved with relevant products/projects.

Going into the process with an idea of how your team may experience the migration process is critical - put a plan in place for how to sufficiently challenge those who thrive and manage any difficulties for those who struggle – and will help ensure a successful transition.

#### c. Advantages of using Open source – R

Over the years Statistical Programming has evolved and so are SAS Programmers. Endless efforts have been made by both Pharma and CROs to develop and maintain Global SAS Macro programs. No one would want to take the risk of losing out all those precious programs developed and would like to take leverage out of new programming platform such as R and Python. We would like to focus here on R programming and showcase that how SAS Programmers can easily learn R and still be the vital resource for the growth of company.

#### d. Challenges of using Open Source in real time situation?

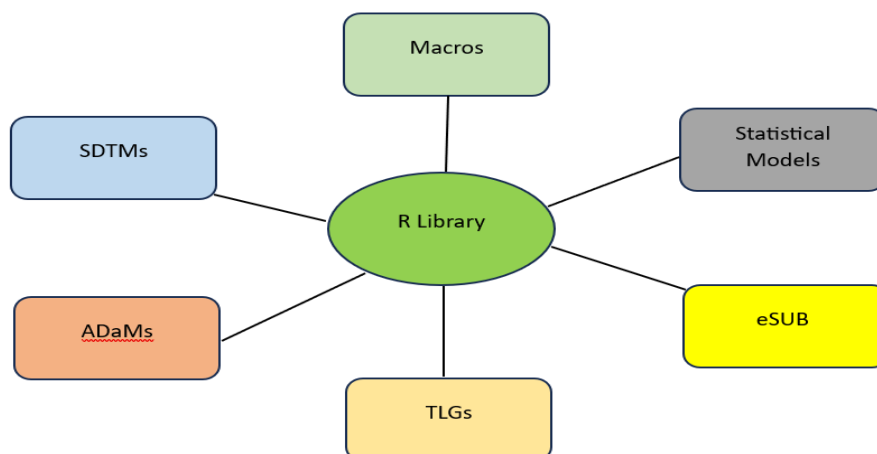
**Lack of Documentation:** Since R has started gaining momentum in Clinical Report Submission, there seems to be few documentations available on complex logic.

**Speed and Efficiency:** R is little slower than other programming languages when dealing with large datasets because it takes up a lot of memory and R's data must be stored in physical memory. However, increasing use of cloud-based memory can make R usage more productive.

**Missing Value:** R handles missing values differently than SAS. All missing values in a R dataset is represented by 'NA' which means 'not available' however SAS displays blank values (' ') for character value and period (.) for numeric value. Hence, an explanation needs to be given in documentation for the same.

### 4. Building SAS to R Libraries (SAS Programs ----> R Programs)

A typical flow diagram, how statistical programmers can start building R programs library.



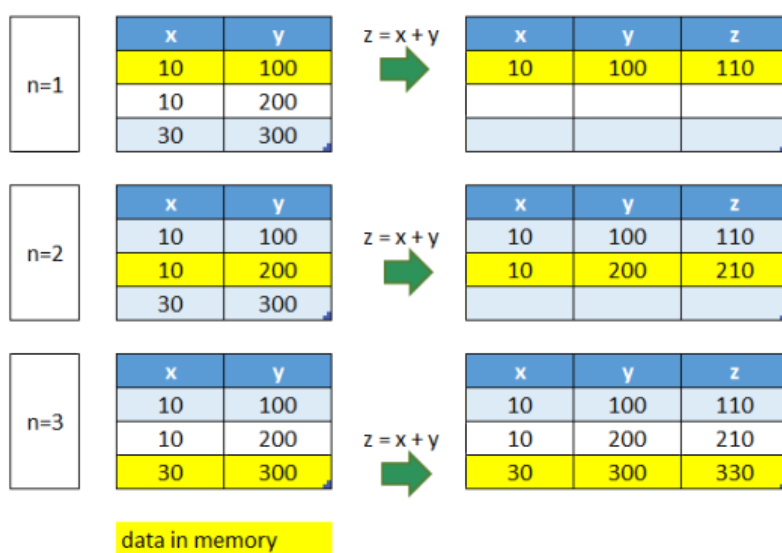
## Figure: SAS Programs to R Programs Library

A few examples, how existing SAS Programs can be converted into R programs. This provides an easy access to R programs libraries for future use.

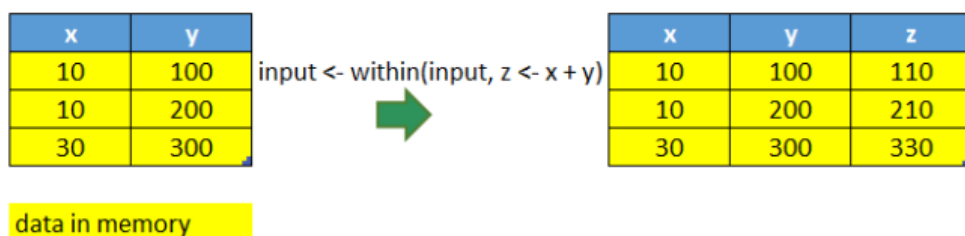
An example of SAS and R program is mentioned below:

| SAS                                                                                     | R                                                                            |
|-----------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| <pre>* add two variables in SAS ; data output ;   set input ;   z = x + y ; run ;</pre> | <pre># add two variables in R input &lt;- within(input, z &lt;- x + y)</pre> |

**Step #1.** SAS processes data row by row, by using an implied loop in Data Step.



**Step #2.** R, on the other hand, applies functions at the column level by processing all rows at the same time, as shown in the following graph.



## 5. Final Goal – Planning effective Change Management

It's widely known that growth of an organization depends on its people and utilizing resources effectively. The leadership skills in bringing latest trends and technologies, motivating and guiding people eventually helps team members to be future ready. Opportunities for skills enhancement and continuous improvement helps employee's satisfaction level and engagement which in turn helps us achieve a successful transition between latest technologies.

## Conclusion

This paper presents a practical approach on change management and how to develop skills that are needed to effectively manage latest technologies and smooth transition. Open-source technologies that involves R programming, Python etc. can play a pivotal role for the growth of organization, and it also provides end to end array of possibilities to employees. Building libraries in R can be useful in many aspects such as cost optimization, be competitive in the market and capability building. This may provide an opportunity to explore many new custom-made libraries, that would have more advantages to other users within R community.

## Acknowledgments

Authors would like to extend their sincere thanks to Efficacy Lifescience Analytics for giving them an opportunity to write and present this paper. Any brand and product names are trademarks of their respective companies.

## Recommended Reading

ChatGPT website: [https://www.itesolution.com/chatgpt.shtml#google\\_vignette](https://www.itesolution.com/chatgpt.shtml#google_vignette)

The potential impact of ChatGPT in clinical and translational medicine:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9976604/>

## Contact Information

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

Shashikant Kumar, Director - Biostatistics & Programming  
Efficacy Lifescience Analytics  
Bengaluru, India  
E-mail: [shashikant.kumar@efficacy.com](mailto:shashikant.kumar@efficacy.com)  
[www.efficacy.com](http://www.efficacy.com)

Mrityunjay Kumar, Associate Director - Biostatistics & Programming  
Efficacy Lifescience Analytics  
Bengaluru, India  
E-mail: [mrityunjay.kumar@efficacy.com](mailto:mrityunjay.kumar@efficacy.com)  
[www.efficacy.com](http://www.efficacy.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.