

# Bridging the Gap: Role Transition from SAS Programmer to Data Scientist

**Authors:** Shashikant Kumar & Mrityunjay Kumar

Presentation No. - PD07

**Stream:** Professional Development (PD)

PHUSE EU Connect 2024, Strasbourg, France 10–13 November

Disclaimer: The opinions expressed in this talk are solely those of the authors and do not reflect the views of their employers

# Agenda

1. Introduction
2. Understanding the Skills Gap
3. Role Transition: SAS Programmers to Data Scientist
4. Adapting Skills for Data Science
5. Relevance of SAS Programming Skills in Data Science
6. Some Examples of Real-world Applications as Data Scientists
7. Future Opportunities
8. Challenges in Role Transition
9. Conclusion

# Introduction

- SAS programmer needs in-depth statistical knowledge, familiarity with clinical trial data, and a structured approach to programming
- Data science is an interdisciplinary field
  - deals with structured and unstructured data
  - data mining techniques, machine-learning (ML) algorithms, and big data to simplify decision making process
- The adoption of data science in clinical trials is seen across the globe
- Diverse data sources may lead to more accurate and timely insights
  - Electronic Health Records (EHRs),
  - Wearables, and
  - Real-World Evidence (RWE) etc.
- Some examples in clinical applications, identify suitable patient populations, forecast dropout rates, and streamline trial workflows

# Understanding the Skills Gap

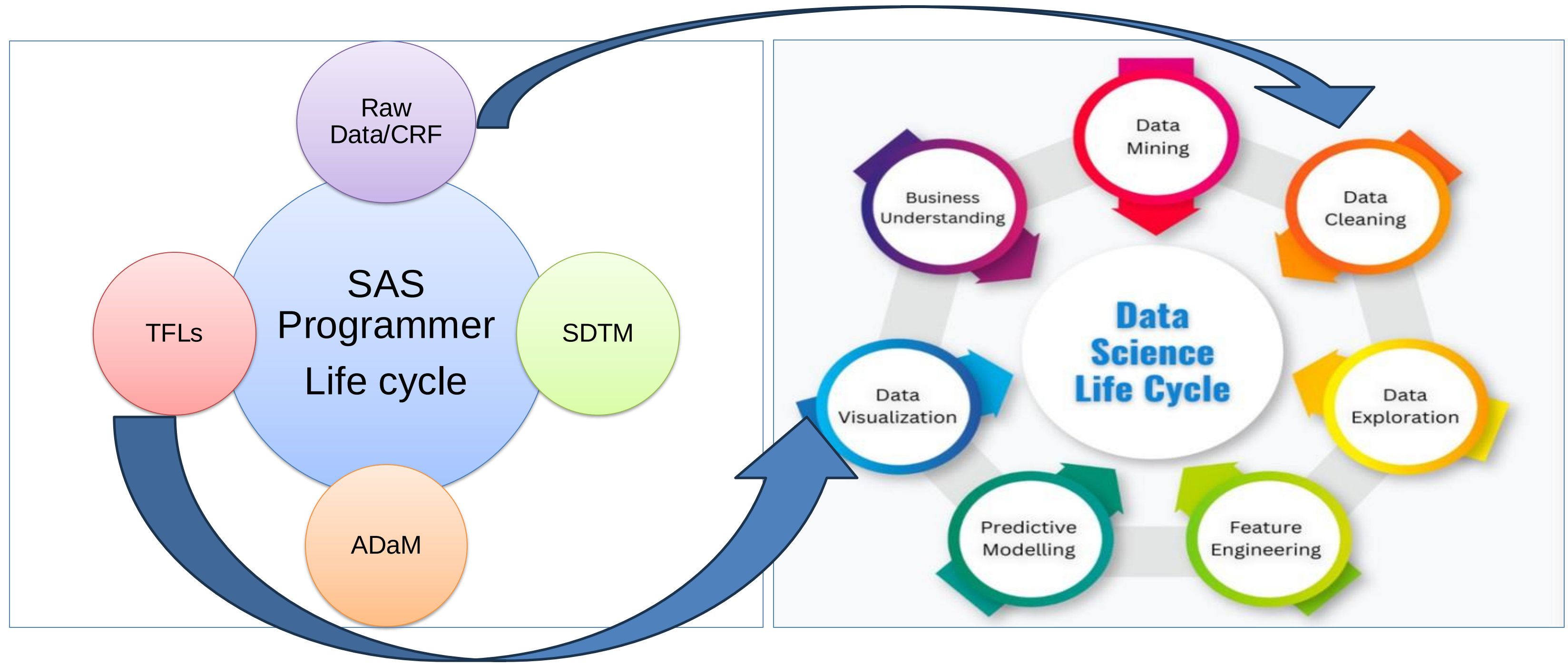
The skills gap is a multifaceted issue, encompassing a lack of technical expertise, soft skills, and industry-specific knowledge:

- Technological advancements
- Skill gaps
- Changing job roles



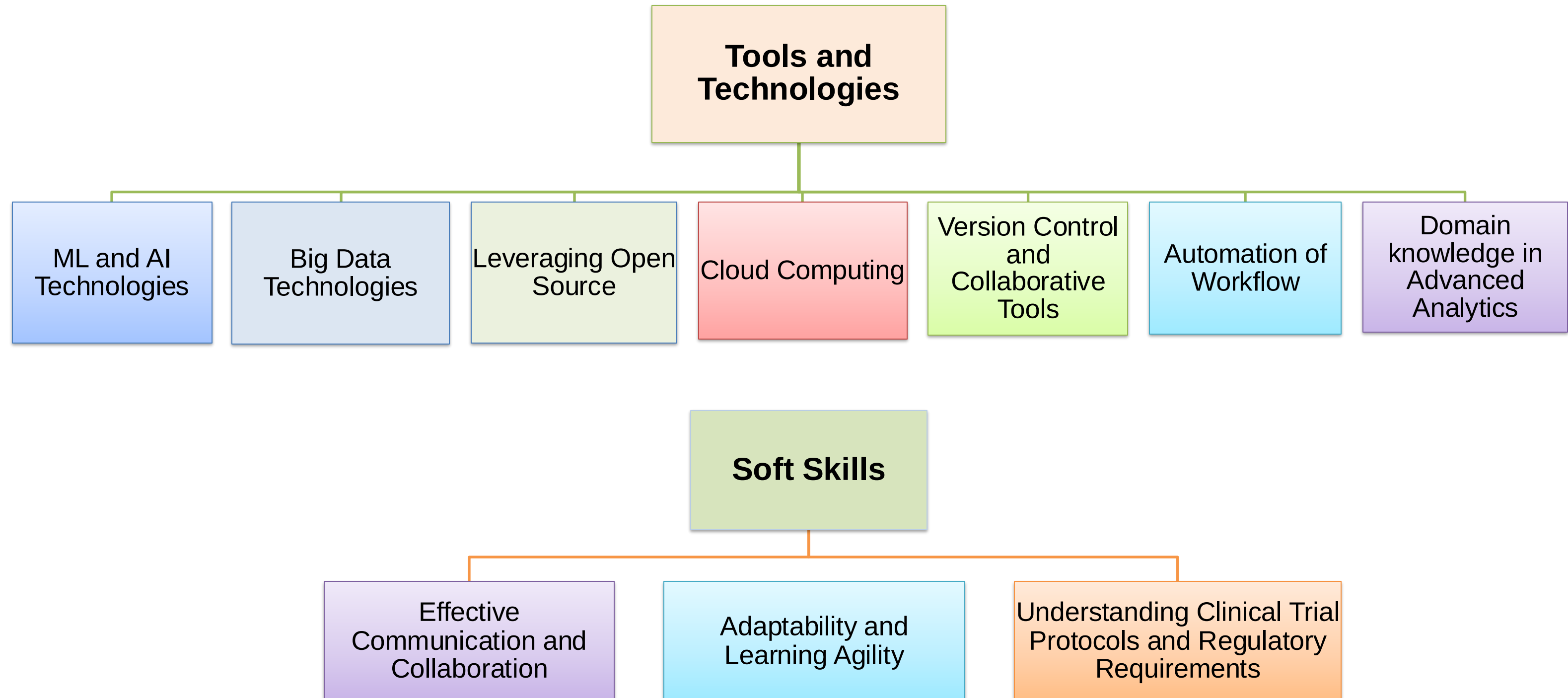
Source: [analyticsweek.com](https://analyticsweek.com)

# Role Transition: SAS Programmers to Data Scientist



Flow Diagram: SAS Programmer vs Data Scientist Life Cycle

# Adapting Skills for Data Science



# Adapting Skills for Data Science contd.

Data scientists typically do the following:

Determine which data are available and useful for the project

Collect, categorize, and analyze data

Create, validate, test, and update algorithms and models

Use data visualization software to present findings

Make business recommendations to stakeholders based on data analysis

# Relevance of SAS Programming Skills in Data Science

Below table demonstrates the transferable skills and applications of SAS programming in data science projects that can be leveraged

| Leveraging SAS Skills                             | Description                                                                                                     | Application in Data Science                                                                    |
|---------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|
| Data Management and Manipulation                  | SAS programmers excel in handling large datasets and applying complex data transformations                      | Data wrangling and preprocessing are essential steps in building ML models                     |
| Statistical Analysis and Model Building           | SAS procedures like PROC REG, PROC LOGISTIC, and PROC MIXED offer a strong foundation in statistical techniques | Help in transitioning to ML models such as regression, decision trees, and clustering          |
| Handling Complex Data Structures and Missing Data | Expertise in dealing with hierarchical data, repeated measures, and missing data imputation                     | Applicable in hierarchical clustering, time-series forecasting, & predictive modelling         |
| Automation and Reporting                          | Proficiency in automating workflows and generating reports using macros and the ODS system                      | Automation of data science workflows, reproducible reporting using Jupyter Notebook, RMarkdown |
| Compliance and Data Privacy                       | Familiarity with industry standards e.g., CDISC, HIPAA, and GDPR for clinical data                              | Ensures data science projects adhere to compliance and privacy guidelines                      |

# Examples of Real-world Applications as Data Scientists

Data science has played a transformative role in the clinical research industry, enhancing drug development, patient care, and regulatory submissions

Predictive Analytics  
for Patient  
Recruitment and  
Retention

ML Models for  
Predicting Disease  
Progression

RWD Analysis for  
Drug Repurposing

Digital Biomarkers  
for Neurological  
Disorders

AI-Driven Clinical  
Trial Monitoring  
and Risk-Based  
Monitoring (RBM)

NLP for Adverse  
Event (AE)  
Detection

Image Analysis  
Using Deep  
Learning in  
Oncology Trials

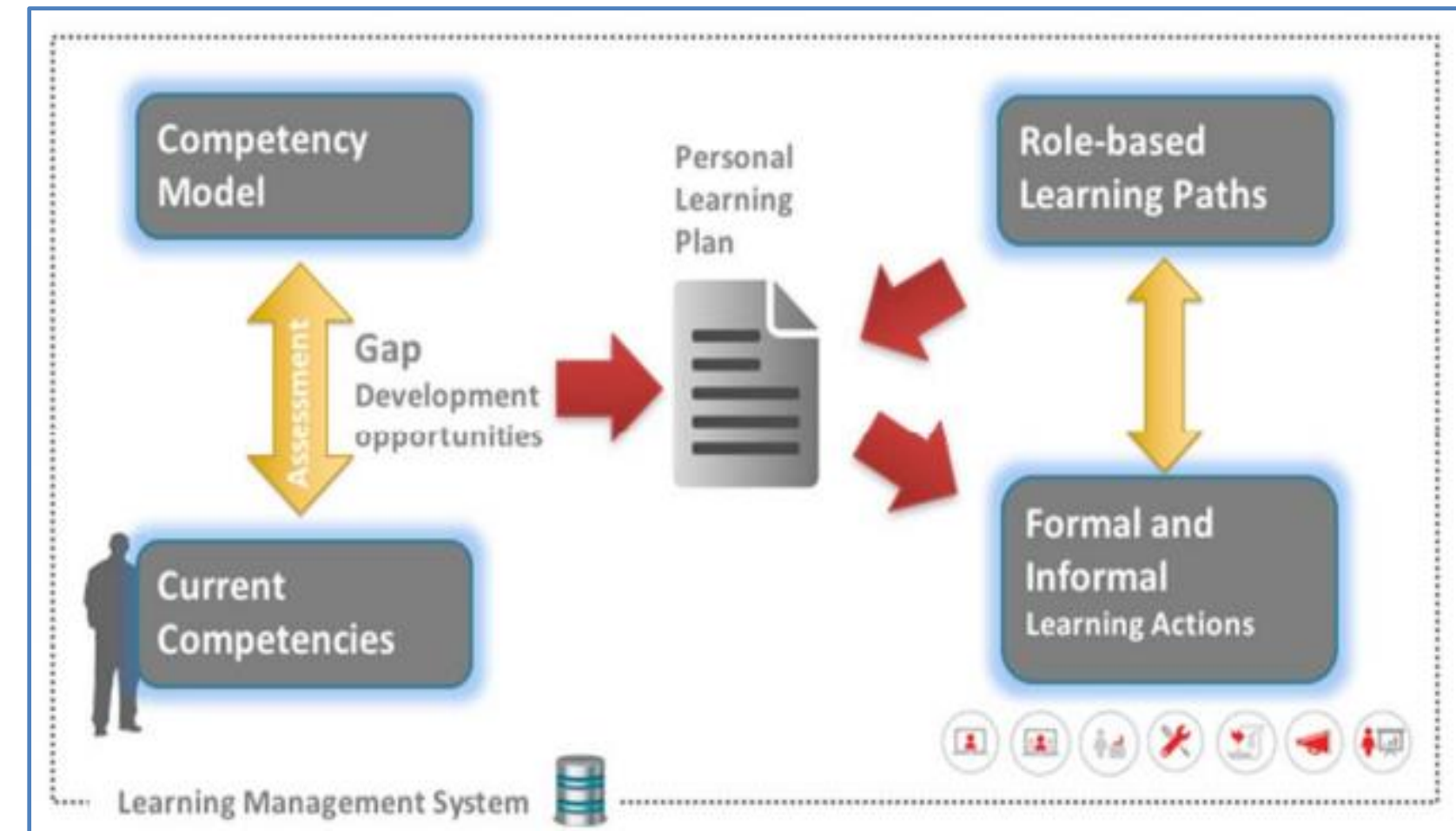
AI-Powered Drug  
Discovery and  
Development

# Future Opportunities

## Professional Development

- Stages of Professional Growth
  - Entry → Intermediate → Advanced
- Continuing Education and Certification Programs
- Networking and Mentorship Opportunities in the Clinical Research Community

## Competency-Based Learning



# Challenges in Role Transition

## Steep Learning Curve

Rule-based SAS programming environment to the dynamic landscape of data science

The continuous evolution of new technologies to advance learning.

## Lack of Real-World Experience

SAS programmers have extensive experience in structured data but limited exposure to unstructured data.

Data sources like text or images, which are common in the field of data science.

## Integration with Current Role

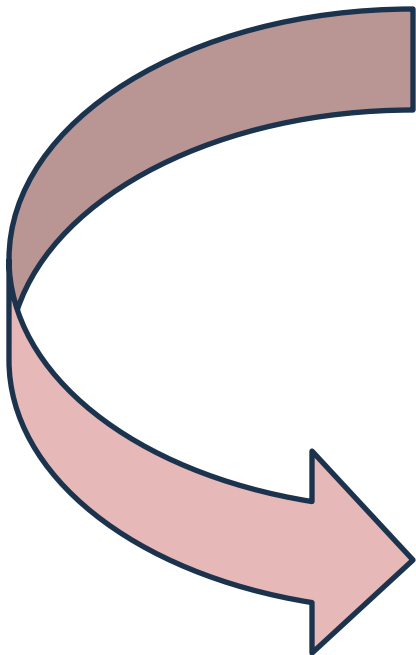
Balancing existing responsibilities while upskilling can be challenging.

Managing time effectively and prioritizing learning goals is crucial.

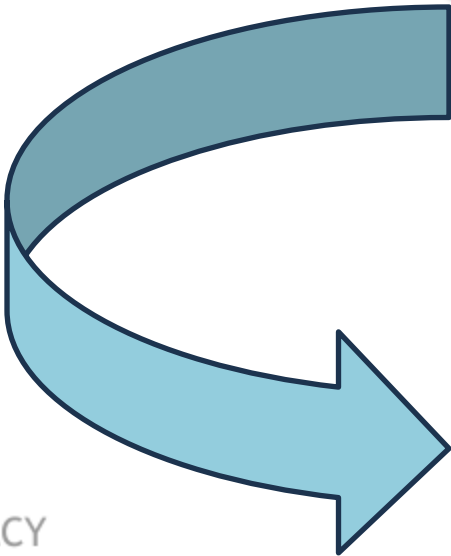
# Conclusion

Data science is rapidly transforming the landscape of clinical trials by:

- improving study design, enhancing patient recruitment & retention
- enabling real-time monitoring
- optimizing data management thus increasing the need of data scientists



As the pharmaceutical industry continues to evolve, data science will play an increasingly critical role in shaping the future of clinical research where SAS programmers can effectively use their clinical and programming knowledge by aforesaid means in role transition to Data Scientist.



With professional training, organization can develop an expert data science team and level up existing employees to suggest data science projects and facilitate collaboration and communication.



THANK YOU