



*Next evolution step(s) for statistical
programmers:
“Clinical Data Scientist”
or
“Clinical Data Analytics Engineer”?*



Amar Mahidadia & Sascha Ahrweiler

Bayer LLC

June 2021



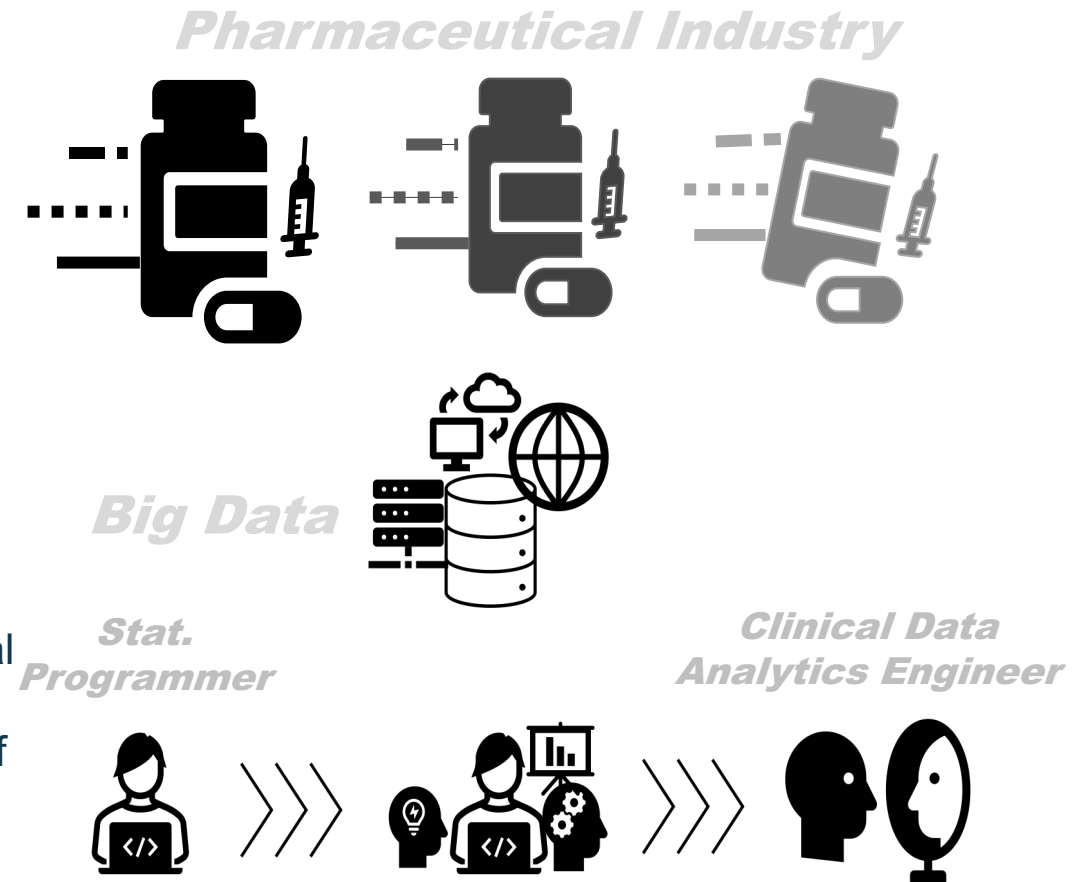


Abstract

The pharmaceutical industry is changing at a rapid pace, as is the role of a statistical programmer. With the evolution of digital capacity, cloud computing, and handheld devices, much more clinical data is produced and stored in the digital space. The traditional statistical programmer role may not be enough to handle these needs.

We need to harness advanced analytical techniques including sophisticated statistical methods, machine learning, and artificial intelligence to realize the full potential of the opportunity and unlock the insights from this data. For these reasons, the role of the statistical programmer needs an upgrade.

Should this upgraded role be defined as “Clinical Data Scientist” (i.e., defining the best advanced analytical methods) or “Clinical Data Analytics Engineer” (i.e., implementing advanced analytical methods)? This paper will illustrate that the upgraded role, which is application-based, should be defined as “Clinical Data Analytics Engineer”.





Agenda

- // Big Data Impact on Pharma & Other Industries
- // Do We Need Data Science Approach in Clinical Development?
- // Current Roles & Responsibilities
- // Future Roles & Responsibilities
- // Conclusion



Big Data Impact on Pharma & Other Industries

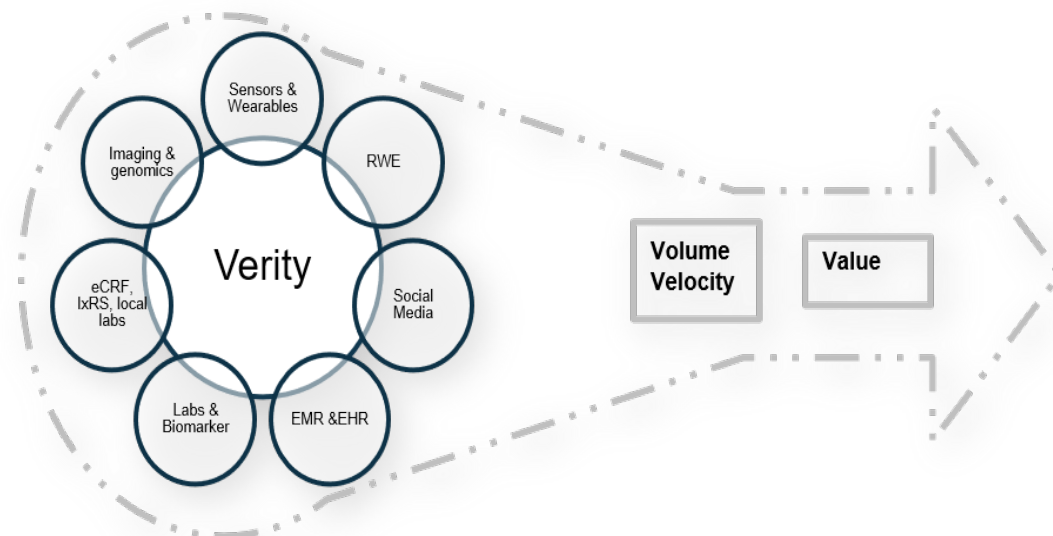
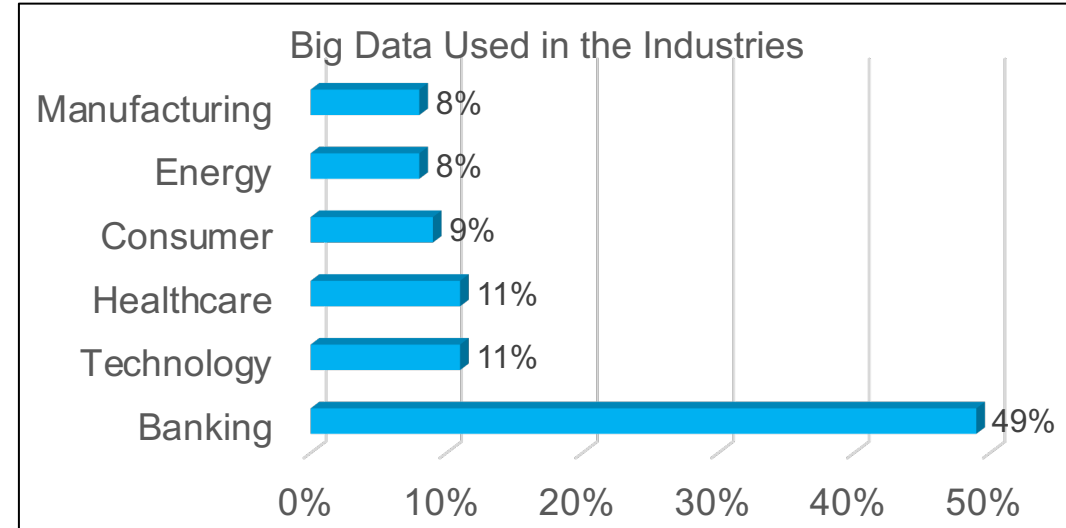
What are the global trends seen in the industry?

// Big Data

How are industries dealing with this challenge?

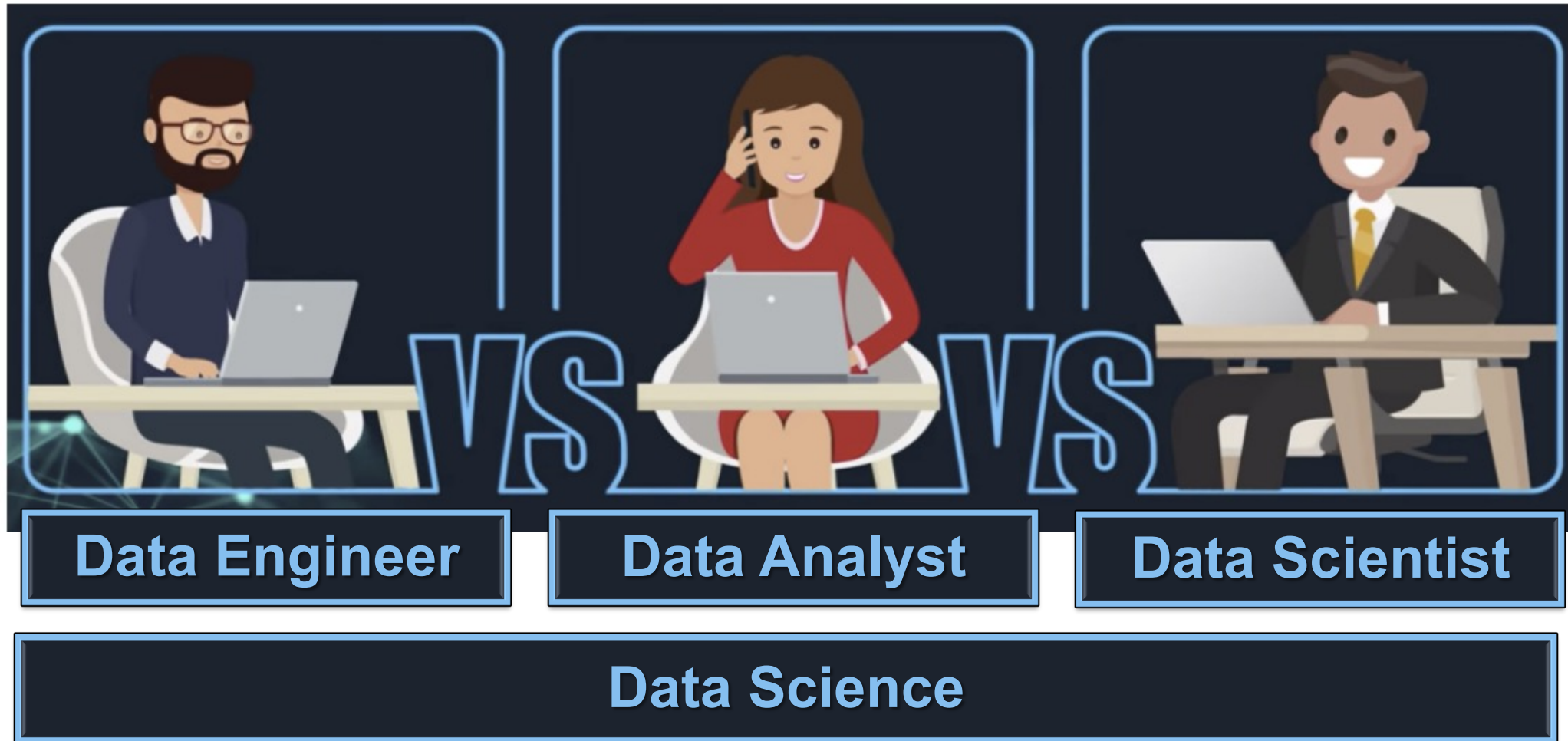
// In other industries, the data science approach has been applied by:

- Data Engineer
- Data Analyst
- Data Scientist





“Big Data” Solution Using A Data Science Approach (Other Industries)





What do these three functions do? (other industries)

Key Skills



Data Engineer

- Data architecture & pipelining
- Data cleaning, anomaly detection
- Data storage(structured/unstructured)
- Sensors, device, external data collection
- Data Warehousing & ETL



Data Analyst

- Reporting and visualization
- Statistical and analytics skills
- Advanced programming knowledge
- Analyze, model and interpret data



Data Scientist

- Statistics
- Data mining
- Machine learning concept
- Data Visualization
- Data Manipulation and Analysis

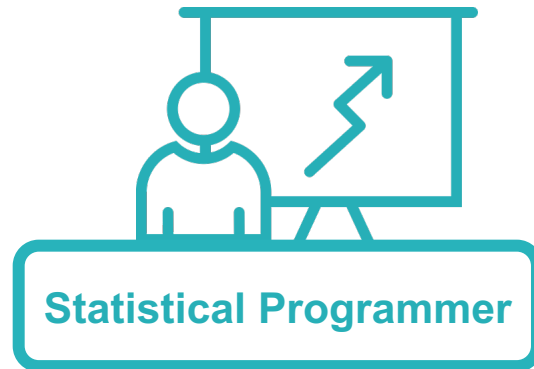


Big Data Impact in Clinical Development

Should clinical development jump on the data science bandwagon?

// Every function thought so (*few years ago*)!

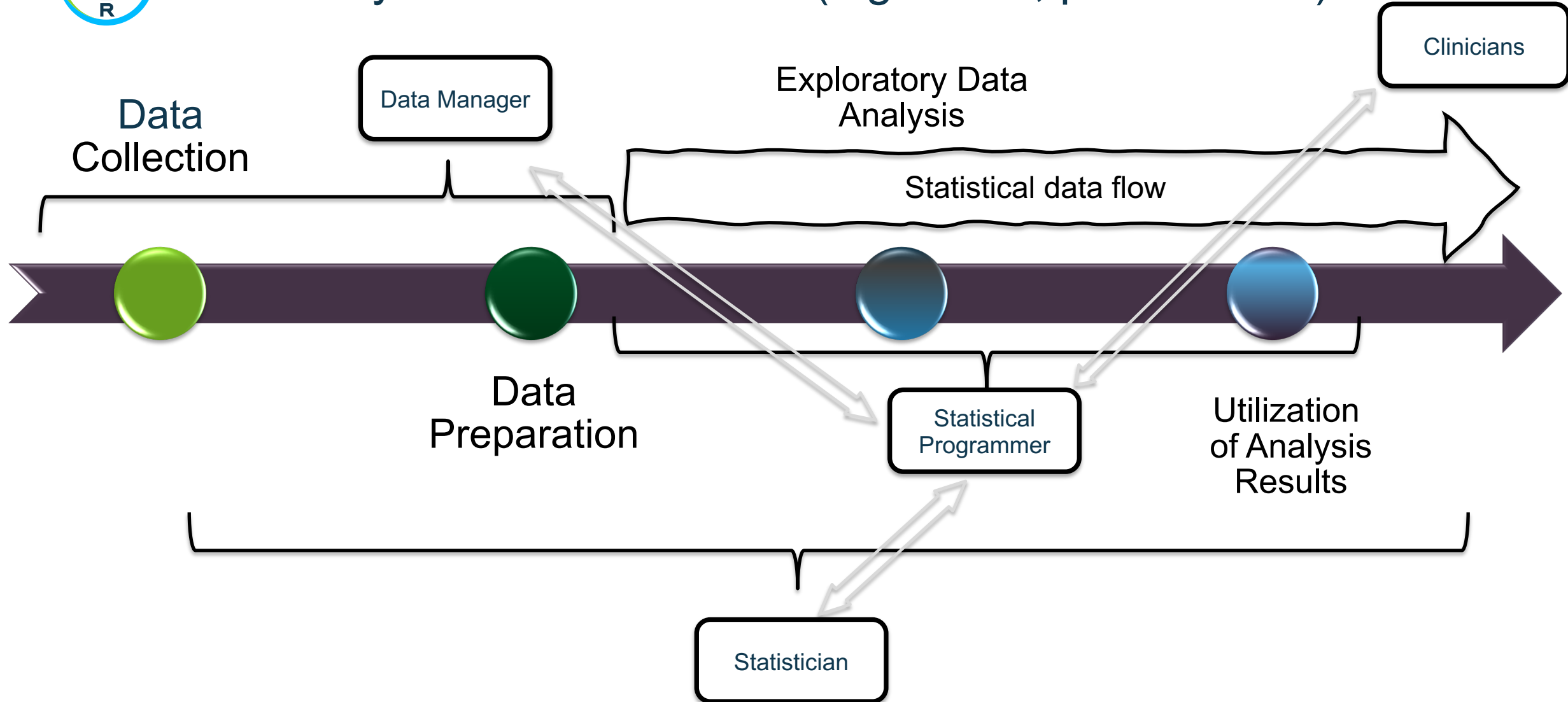
// “Data Science” has become a buzz word. 



- Data managers/engineers thought they were in the best position to advance to the Clinical Data Scientist role. Michaela Jahn touched upon this: [Paper Phuse PD08 Data Scientist Final 20151030](#)
- Statistical programmers were not far behind. Sascha Ahrweiler presented the idea in 2014 ([Evolving from Statistical Programmers to Clinical Data Scientists](#))
- Karl Pearson, a British mathematician and arguably the father of modern statistics, is credited with the quote, “**Statistics is the grammar of (Data) science**”.



Data Lifecycle in Clinical Trials (high-level, past-current)



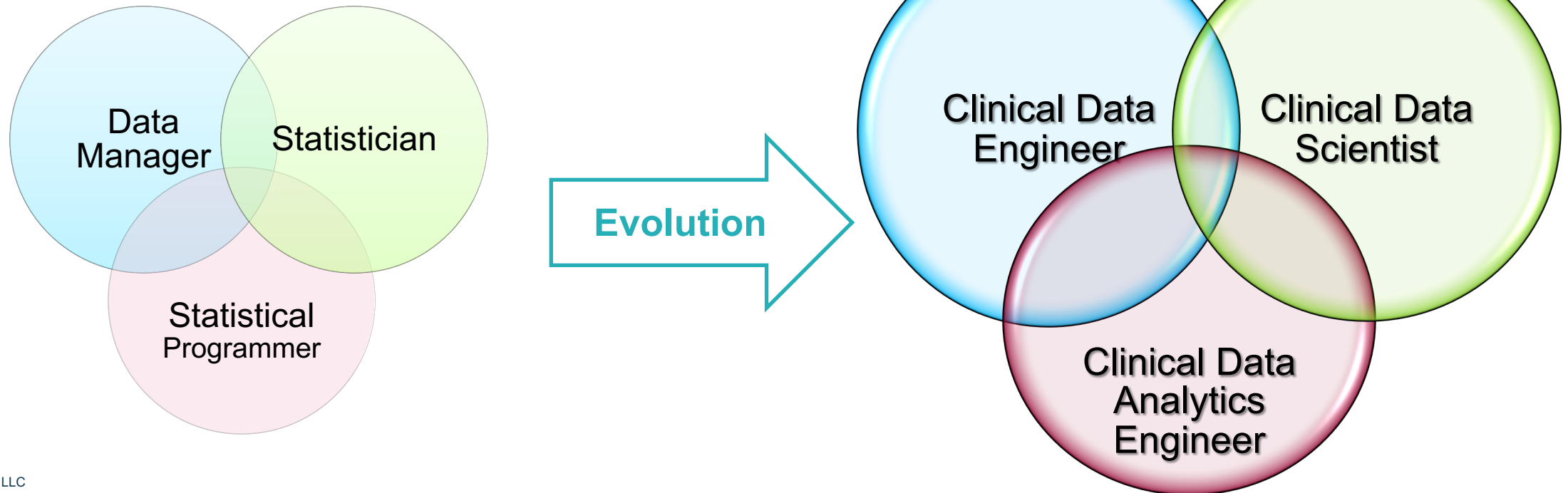


Big Data Impact in Clinical Development (Regulated Industry)

In order to realize the full potential of the opportunity and unlock the insights from this data, we need to harness advanced analytical techniques including sophisticated statistical methods, machine learning, and artificial intelligence.

Who is going to do that?

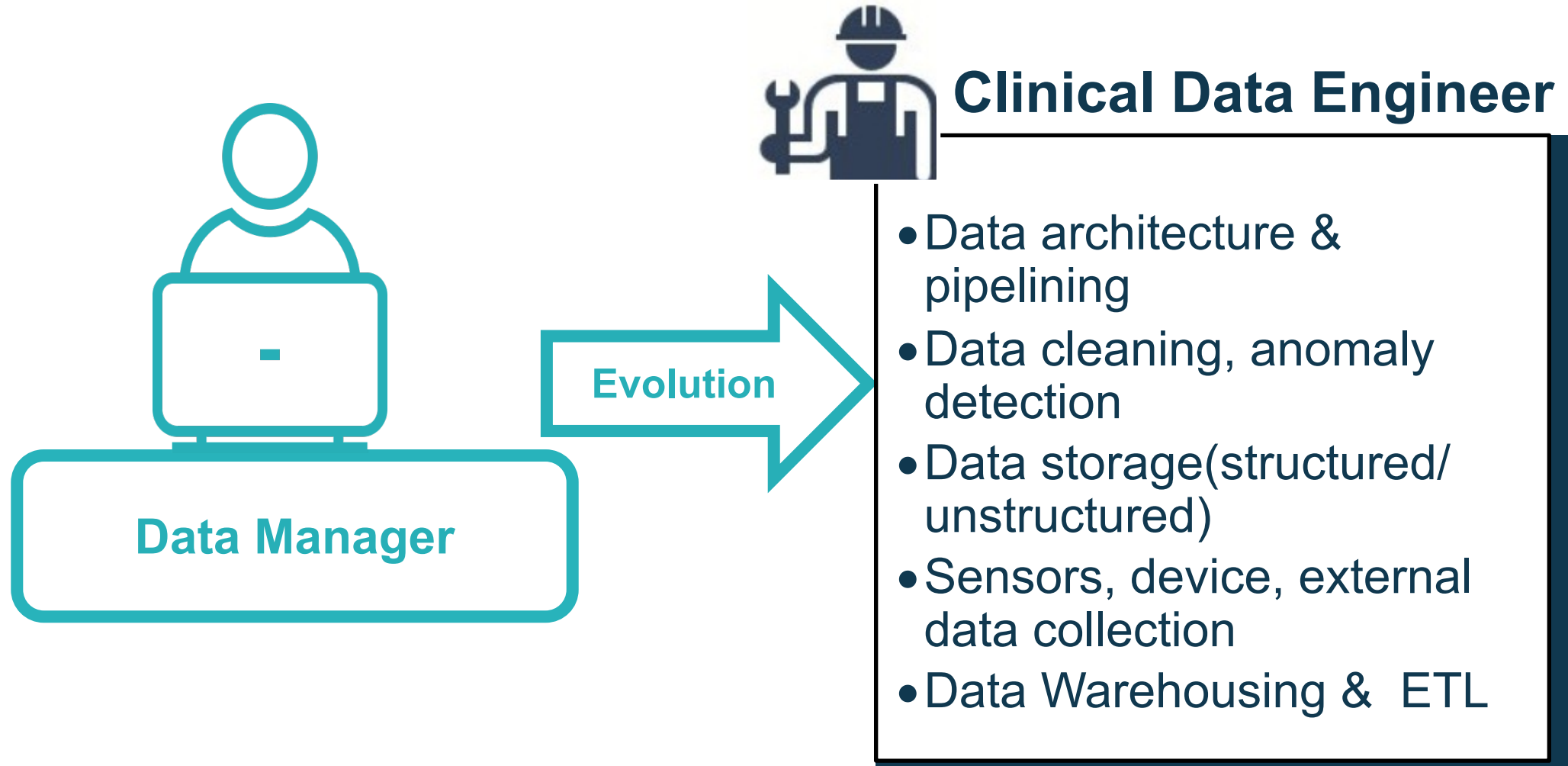
- Every function needs to evolve to next stage to embark upon new responsibilities.





Big Data Impact in Clinical Development (Regulated Industry)

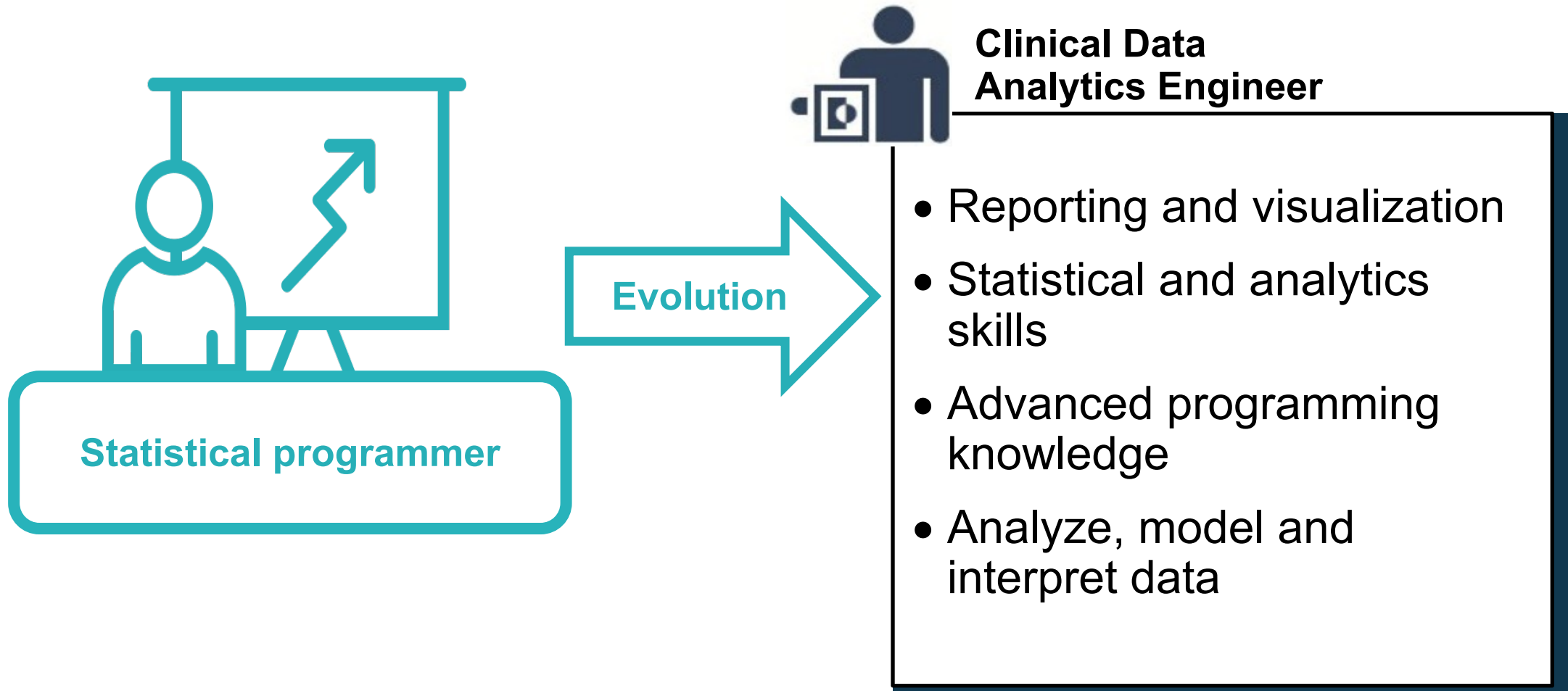
Clinical Trials





Big Data Impact in Clinical Development (Regulated Industry)

Clinical Trials





Big Data Impact in Clinical Development (Regulated Industry)

Clinical Trials



Evolution



Clinical Data Scientist

- Statistics
- Data mining
- Machine learning concepts
- Data Visualization
- Data Manipulation and Analysis



Big Data Impact in Clinical Development (Regulated Industry)

How is pharmaceutical industry different from other industries?

Regulations



- Requires strict compliance
- Patient data results in privacy issues
- Strong government oversight

Data Transformations



Statistical programmers are best suited for

- Transforming
- Massaging
- Handling sensitive data

Collaborations



Statistical programmers should work with:

- Statisticians
- Data managers
- Other Stakeholders

to engineer the statistical data flow with full compliance



Future Role / Skills for **Clinical Data Analytics Engineer**

Key skills



Data Engineer

- Data architecture & pipelining
- Data cleaning, anomaly detection
- Data storage(structured/unstructured)
- Data Optimization
- Sensors, device, external data collection
- Data Warehousing & ETL



Data Analyst

- Reporting and visualization
- Statistical and analytics skills
- Advanced programming knowledge
- Analyze, model and interpret data

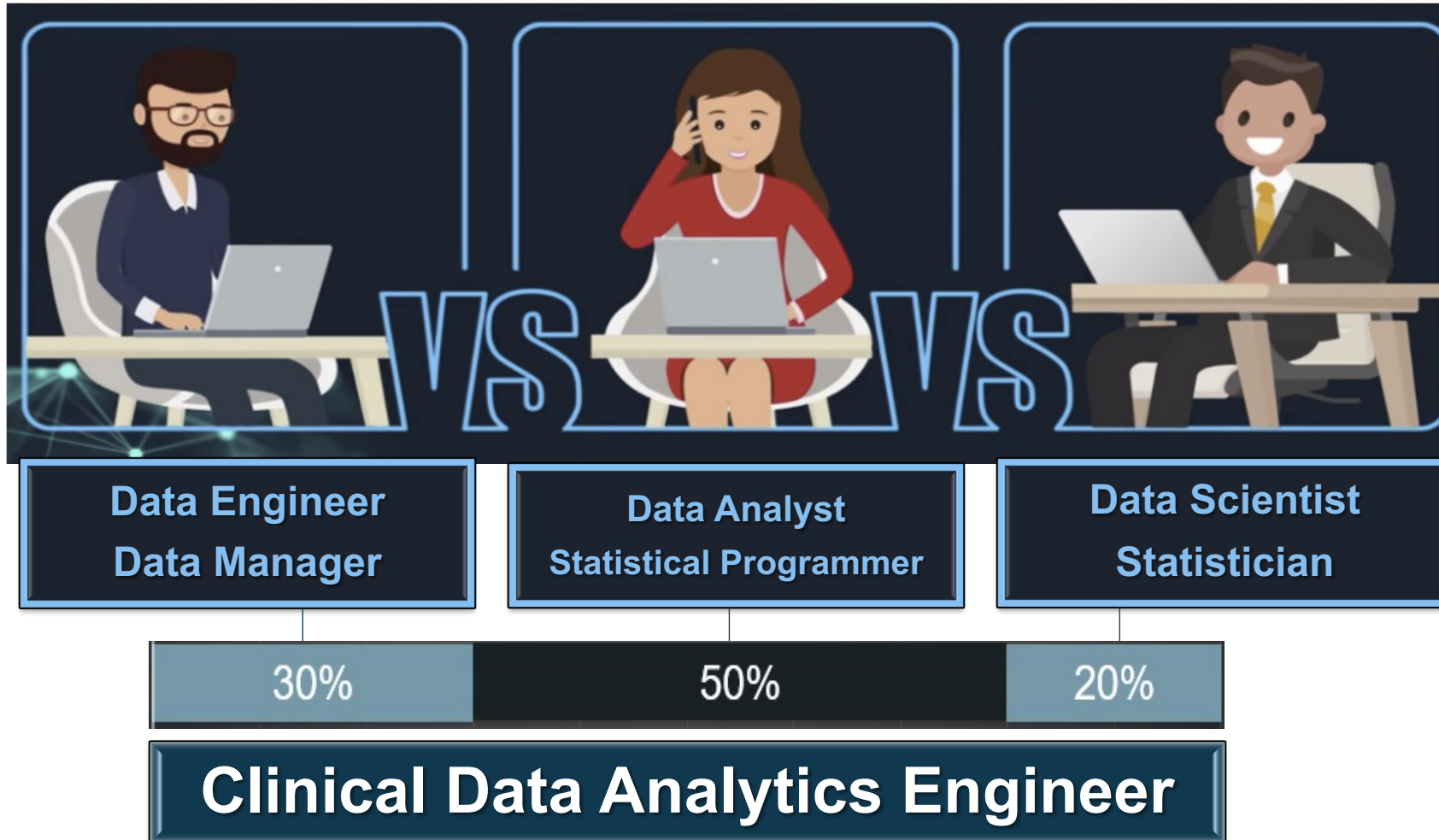


Data Scientist

- Statistics
- Apply data mining
- Apply machine learning concepts
- Apply data manipulation and analysis
- Communications & Soft Skills



“Big Data” Solution Using A Data Science Approach (Pharmaceutical: New Role)





Conclusion

The landscape of the clinical development (data and analytics) is shifting rapidly. In many companies, the roles and responsibilities of data managers, statistical programmers, and statistician are changing.

This change has created the need for an updated new role, “Clinical Data Analytics Engineer”. We think that statistical programmers will fit in this role as they are more equipped to add the data engineering part of data lifecycle into clinical trials.