



PhUSE SDE London - 20th May 2019

Structuring Your Data Landscape for Real World Evidence

Abstract



Guy Garrett CITP FBCS
Managing Director
Achieve Intelligence

Guy is Managing Director of Achieve Intelligence, a consultancy offering Enterprise Architecture and Information Strategy. He is a guest lecturer, teaches the SAS EMEA: Data Warehouse Architecture & Design Principles course, and co-leads the PhUSE Data Engineering Project.

Clinical development has made huge efficiencies over recent decades, (eCRF and CDSIC standards to name a few).

However, just as we're reaching the summit of traditional database management techniques we're facing another step-change in our journey – the big data mountain.

The methodologies we're currently using cannot cater for these vast and varied data sources, without a huge increase in resource - which is something most BioPharma companies haven't provisioned for. So how can we bridge the gap between current database management techniques and faster more innovative digital technologies?

This presentation discusses a practical step-by-step approach for Data Analyst to begin to transform themselves into Data Scientists; bringing with them cost-saving efficiencies for clinical development and career-enhancing skills in the realm of digital innovation.



James McDermott MSc CStat FIoD CDir
Executive Chairman
Achieve Intelligence

James McDermott is an entrepreneur and independent consultant with over 25 years' experience in the life sciences industry. As a Chartered Statistician and a Chartered Director, he plays a leading role in shaping the clinical data analytics sector of the pharmaceutical industry.

RWE Data Landscape Roadmap



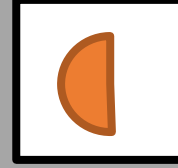
[10 mins]

Past



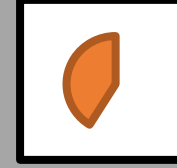
[10 mins]

Transition



[10 mins]

Future



[10 mins]

Q&A

RWE Data Landscape Roadmap



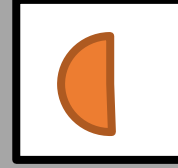
[10 mins]

Past



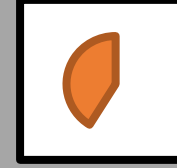
[10 mins]

Transition



[10 mins]

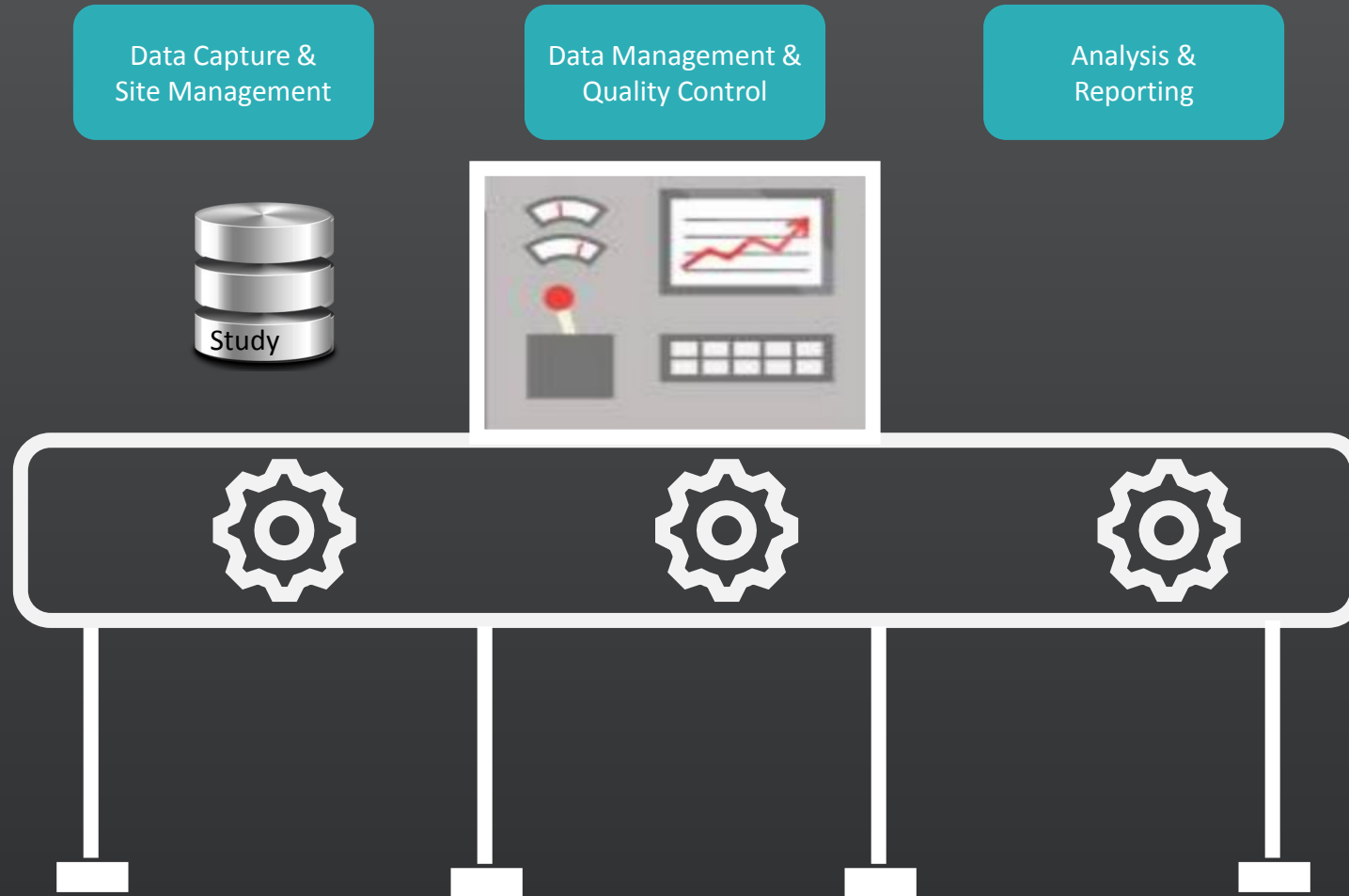
Future



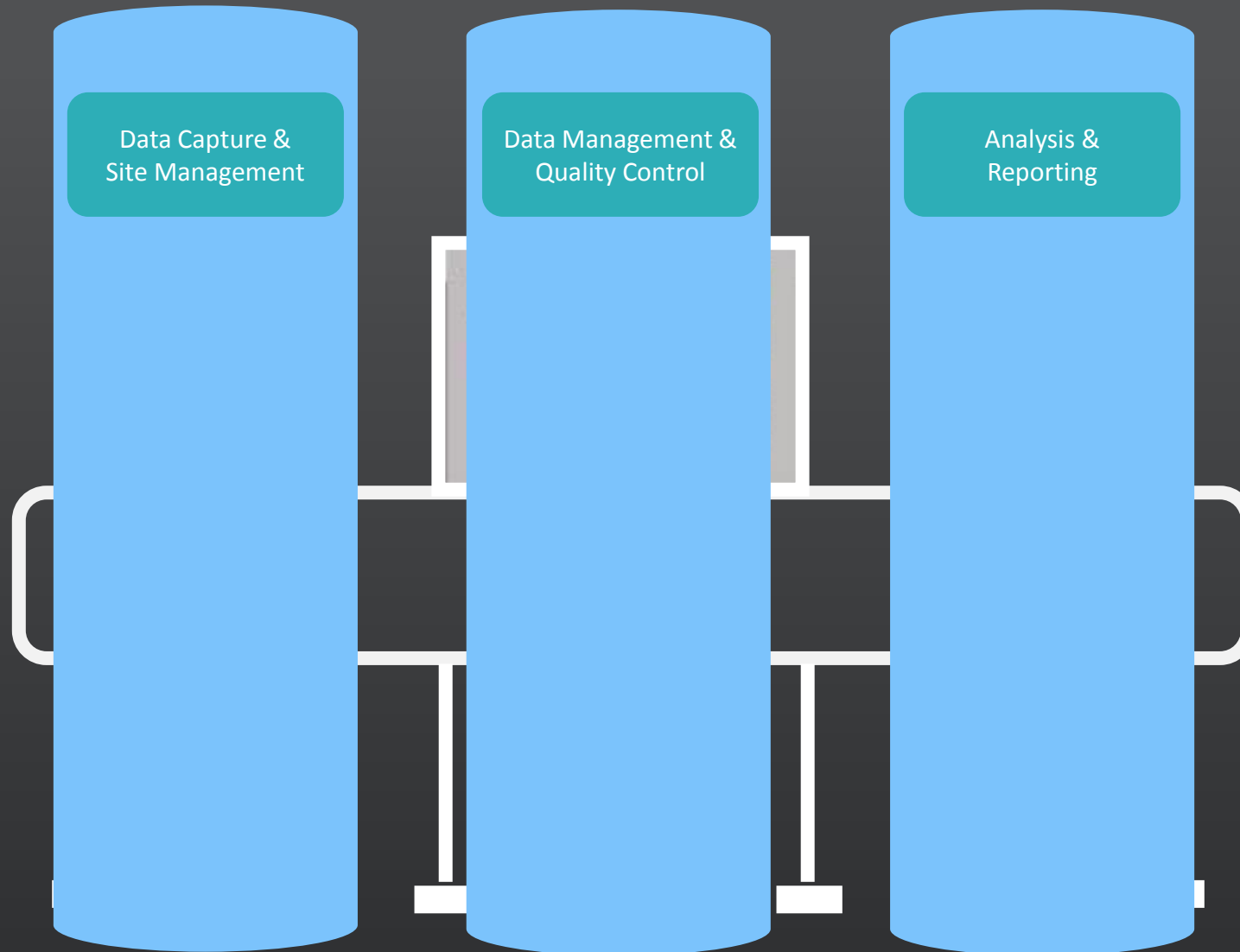
[10 mins]

Q&A

Assembly line / functional silo



Assembly line / functional silo



Assembly line / functional silo



Multiple Studies = Multiple Databases



Multiple Studies = Multiple Databases



Multiple Studies = Multiple Databases



RWE Data Landscape Roadmap



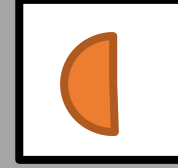
[10 mins]

Past



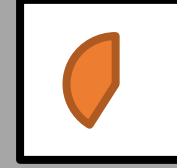
[10 mins]

Transition



[10 mins]

Future



[10 mins]

Q&A

A New Approach



Data Engineering in other industries



Finance



Retail



Telecoms



Insurance

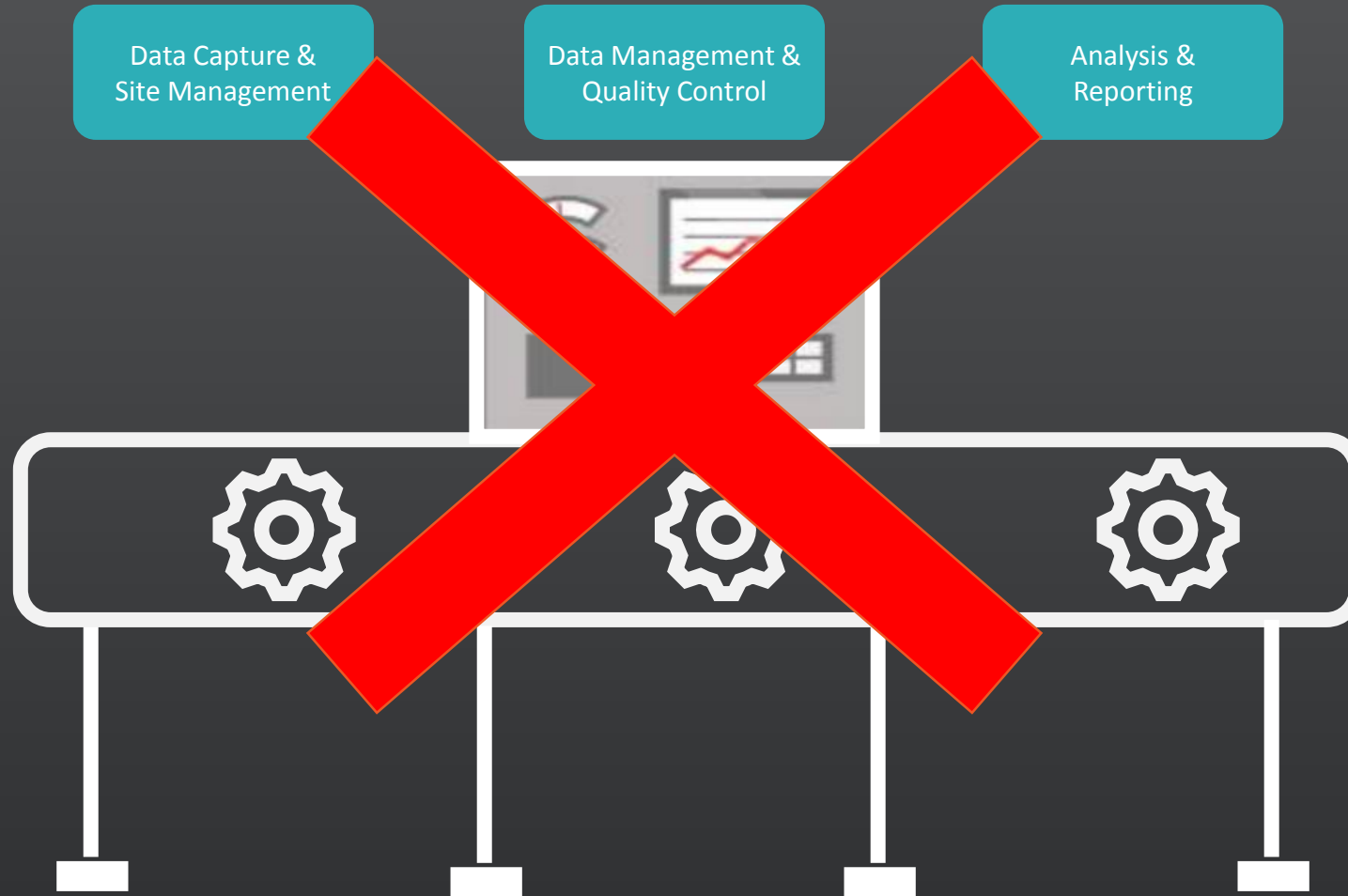
Why isn't BioPharma utilizing Data Engineering?

1. We have focused on other improvements (CDISC and communication with health authorities)
2. Fragmentation prevents holistic decisions
3. Budget pressures have been lower than other industries (until recently)
4. Resistance to change / Inertia of perceived cost of change

or

5. Data Engineering **just cannot work** for the BioPharma industry

REMOVE Assembly line



REMOVE functional silos



Multiple Studies = Multiple Databases



Conformed Data Model



Conformed Data Model

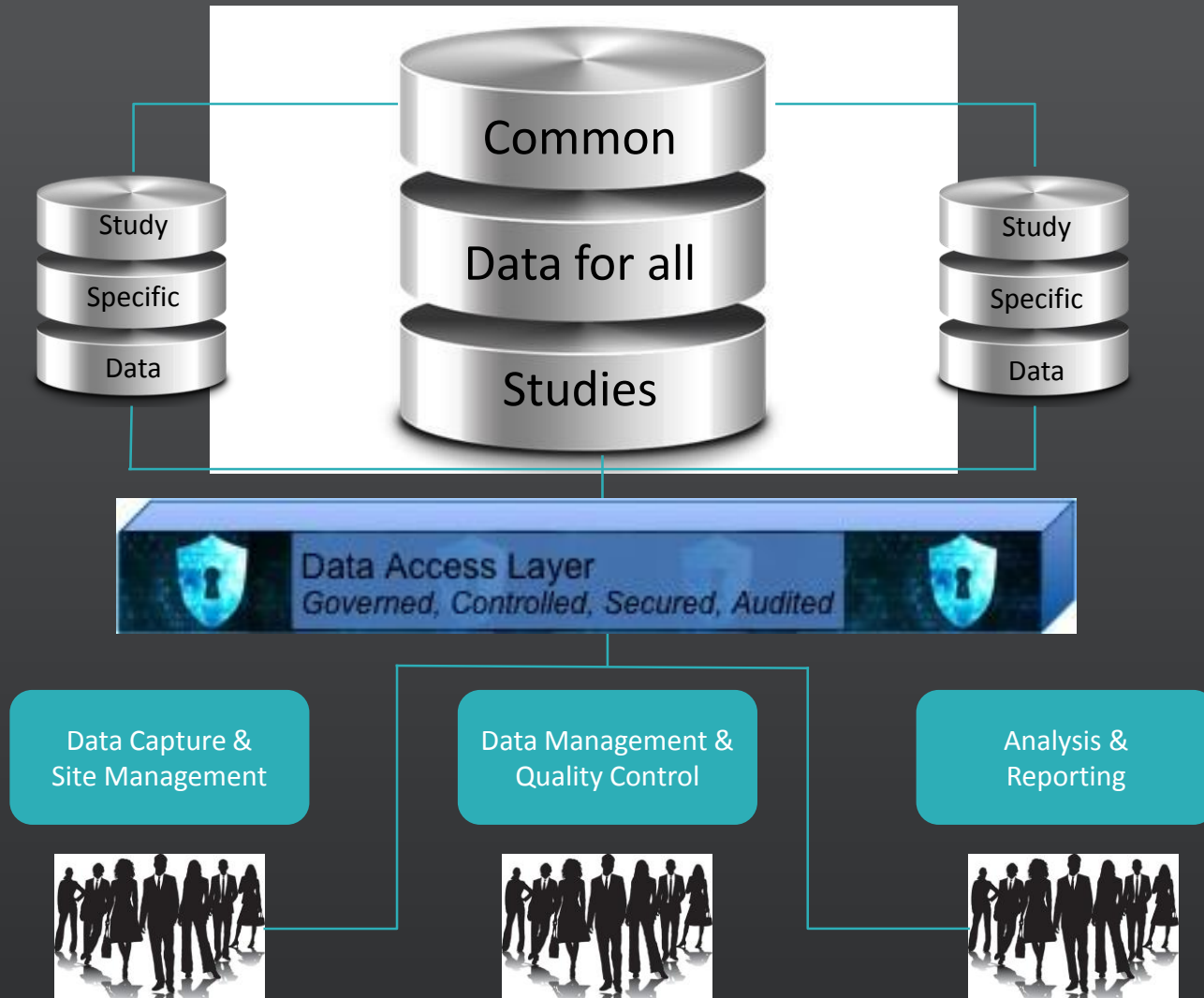


One Central database
build once what you use repeatedly

Conformed Data Model



Conformed Data Model



RWE Data Landscape Roadmap



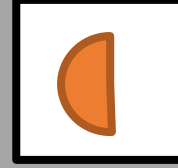
[10 mins]

Past



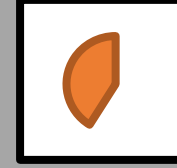
[10 mins]

Transition



[10 mins]

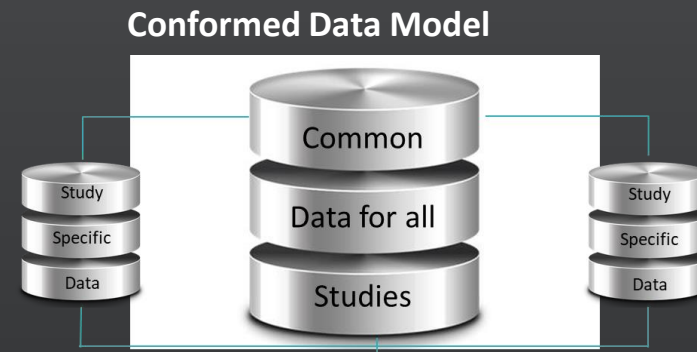
Future



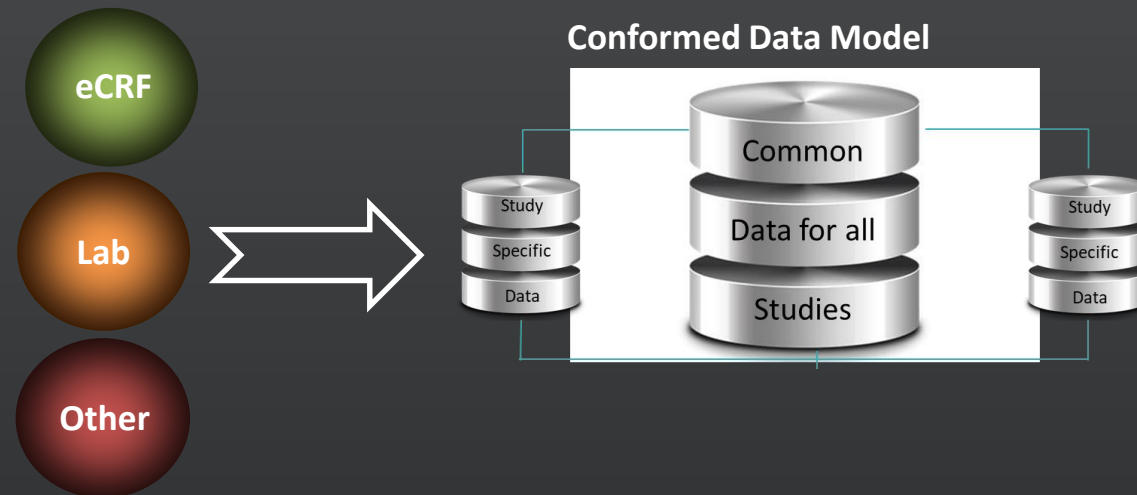
[10 mins]

Q&A

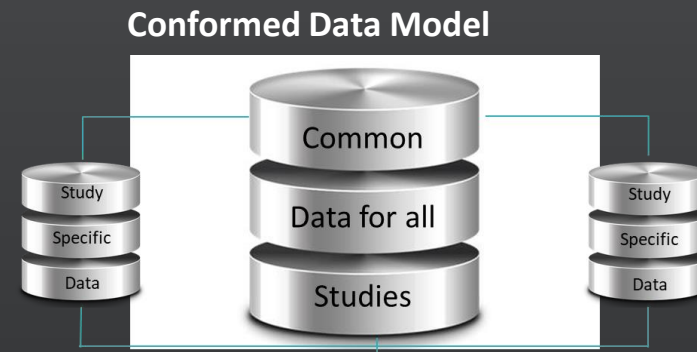
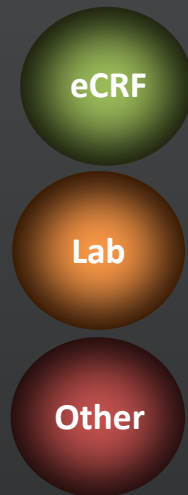
Data Science Pipelines



Data Science Pipelines



Data Science Pipelines



Data Science Pipelines

Real World Data



eCRF

Lab

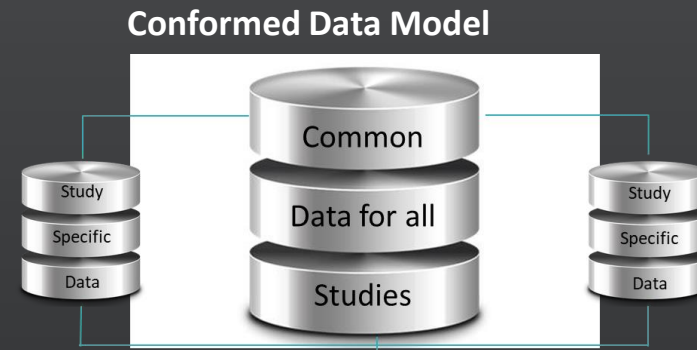
Other



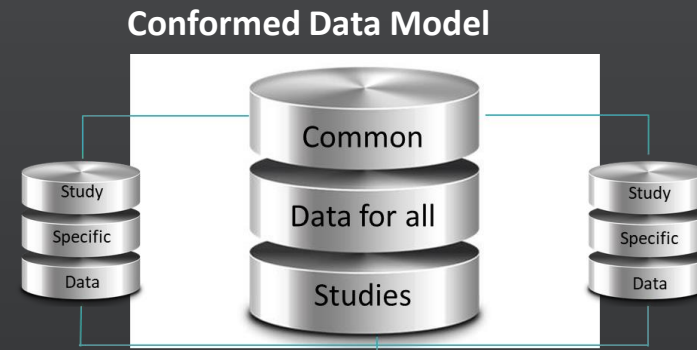
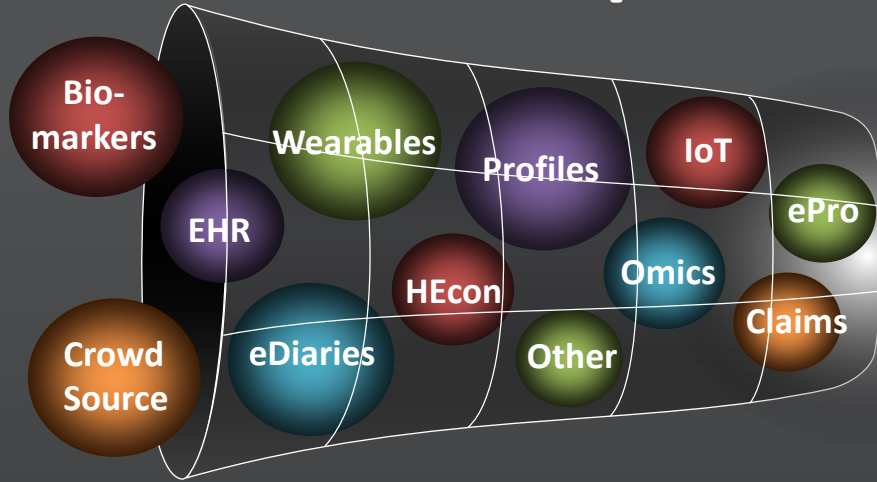
Conformed Data Model



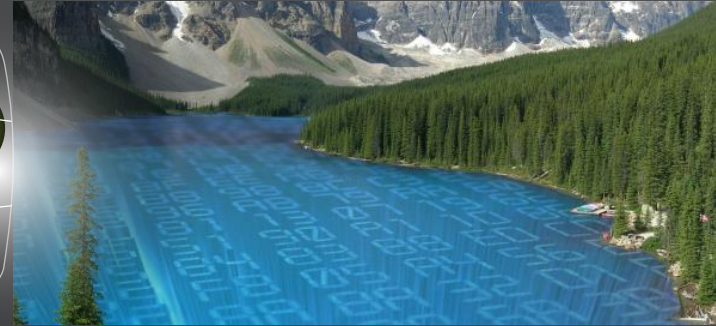
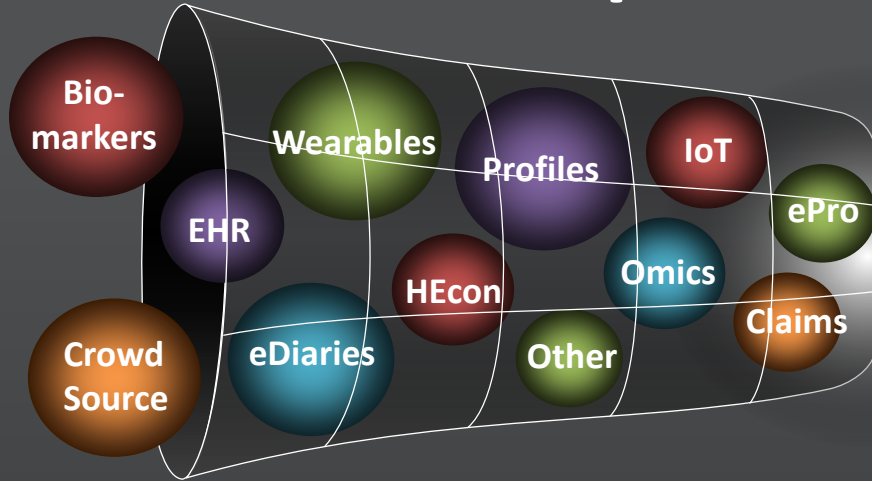
Data Science Pipelines



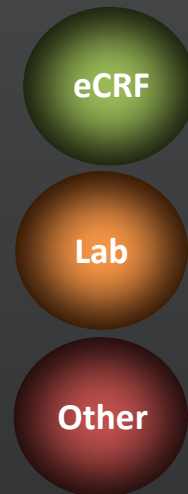
Data Science Pipelines



Data Science Pipelines



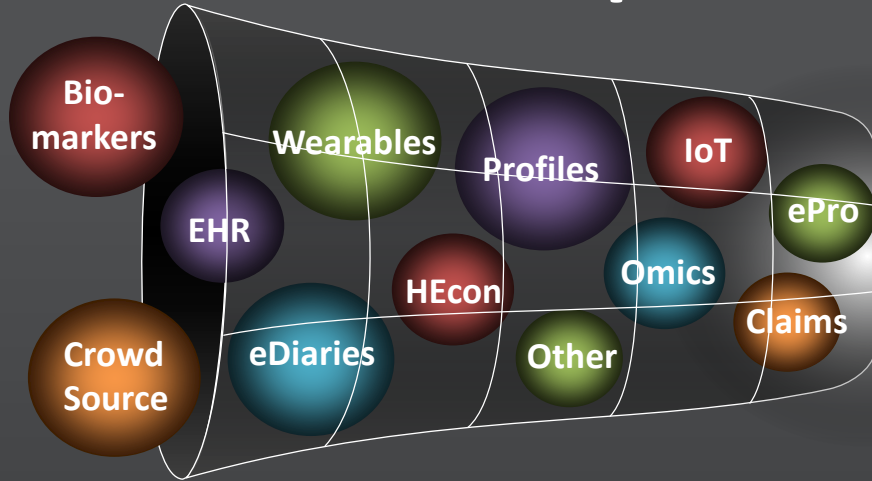
Unstructured Data Lake



Conformed Data Model



Data Science Pipelines



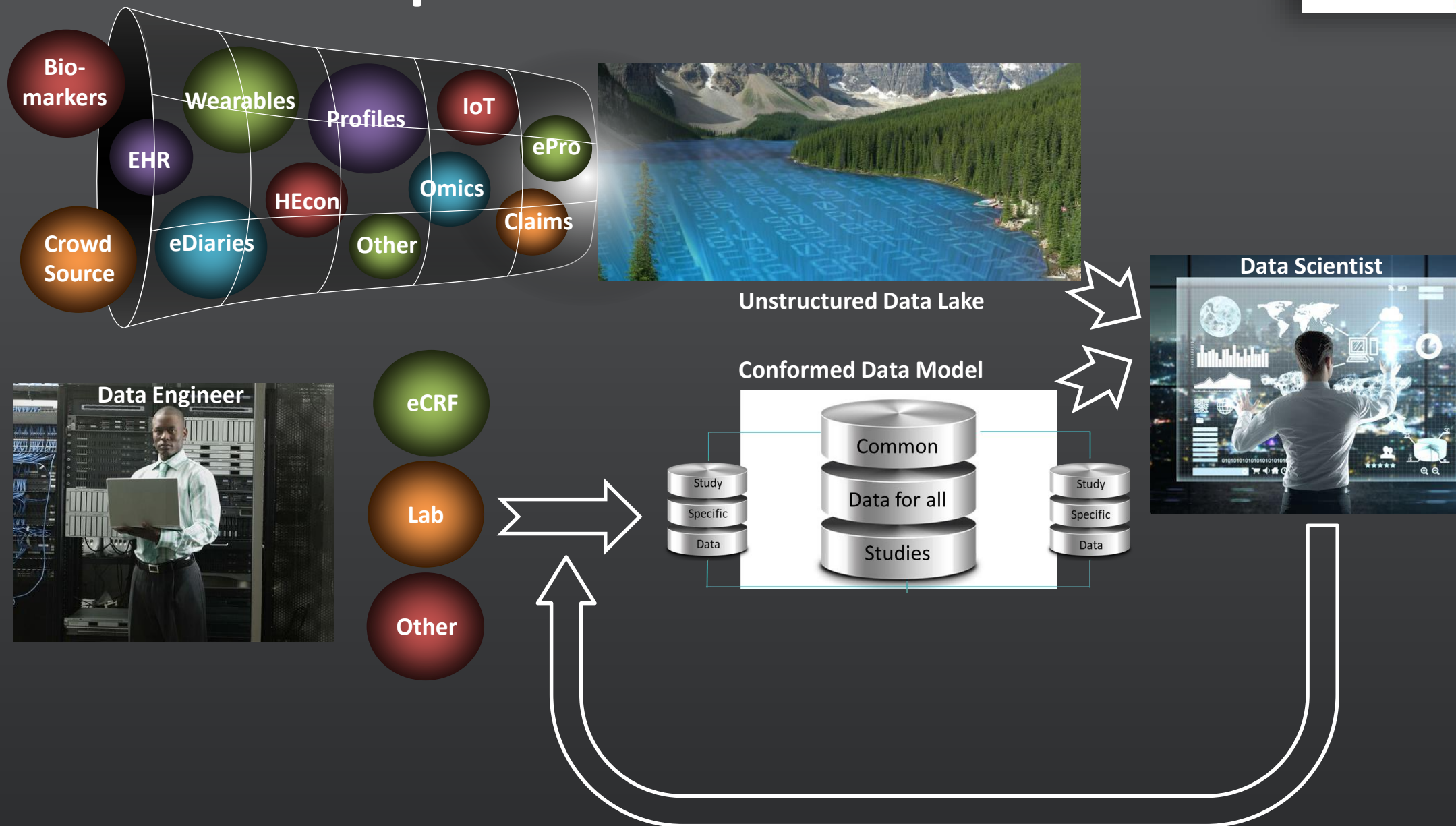
Unstructured Data Lake



Conformed Data Model



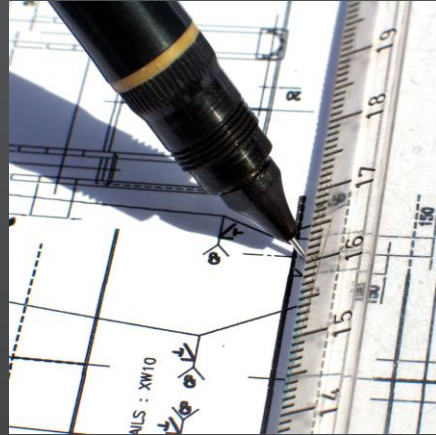
Data Science Pipelines



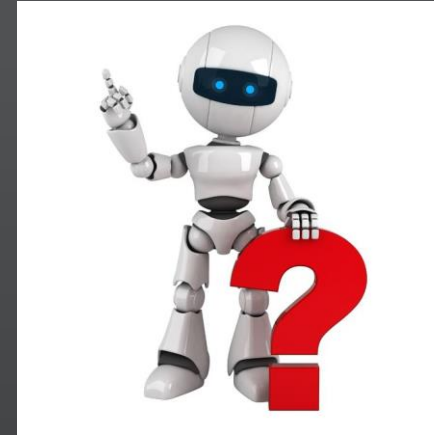
Benefits



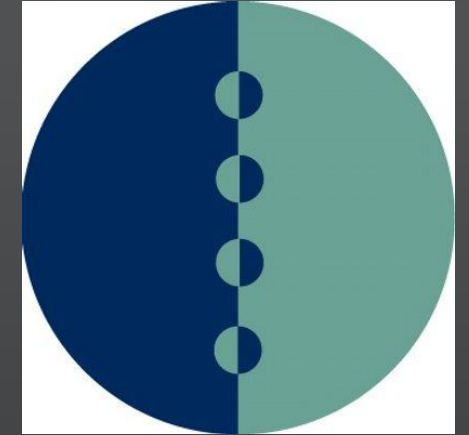
Statisticians & Programmers
-> Data Scientists



Data Managers & Programmers
-> Data Engineers



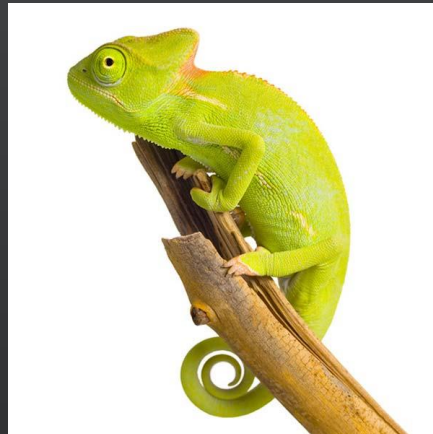
Automatic Processing of
Data and Information



Better Standards
Management



KPI Capability



Better Adaptive Trial
Capability



Serendipitous
Results



Risk-Based
Data Quality

RWE Data Landscape Roadmap



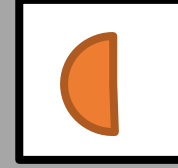
[10 mins]

Past



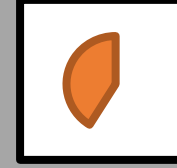
[10 mins]

Transition



[10 mins]

Future



[10 mins]

Q&A