



SP05 – Data Engineering Project

(Educating for the Future PhUSE Working Group)

Presentation:

Guy Garrett, *Achieve Intelligence*

Bev Hayes, *Janssen Research & Development*

Contribution:

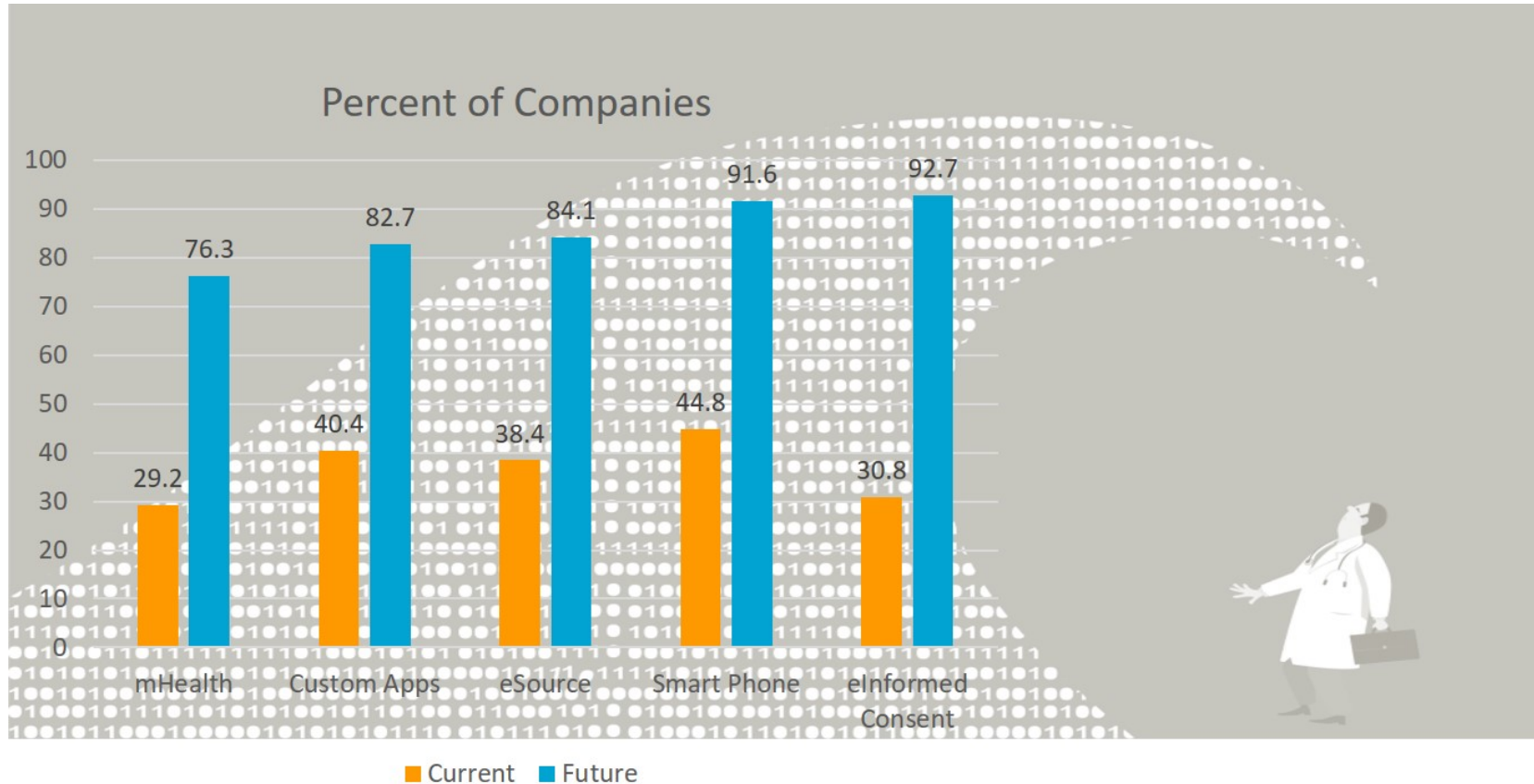
Amy Gillespie, Beate Hientzsch, Jagdev Bhogal, Jatin Patel, Karnika Dalal, Mark Bynens, Mike Carniello, Paul Slagle, Sascha Ahrweiler, Shaaz Ansari, Vijay Pasapula, Vince Marinelli

Special thanks:

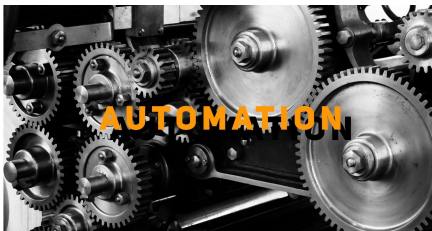
Ian Fleming, James McDermott, Wendy Dobson

<https://education.phuse.eu/eftf/data-engineering/>

Data Sources Used Currently and Projected Use in 3 Years



Data Engineering Project Overview



Data Engineering Project Overview



What is Data Engineering?



What is Data Engineering?

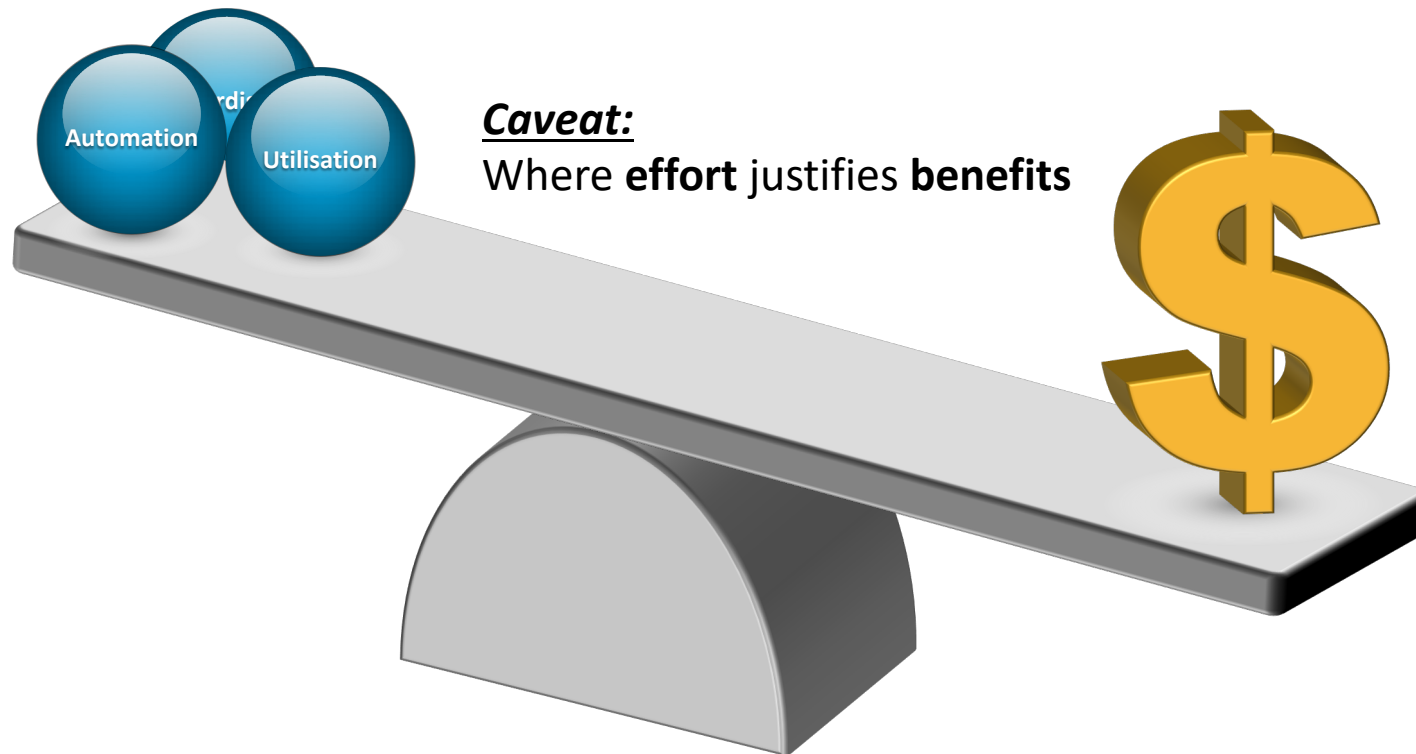
Standardisation



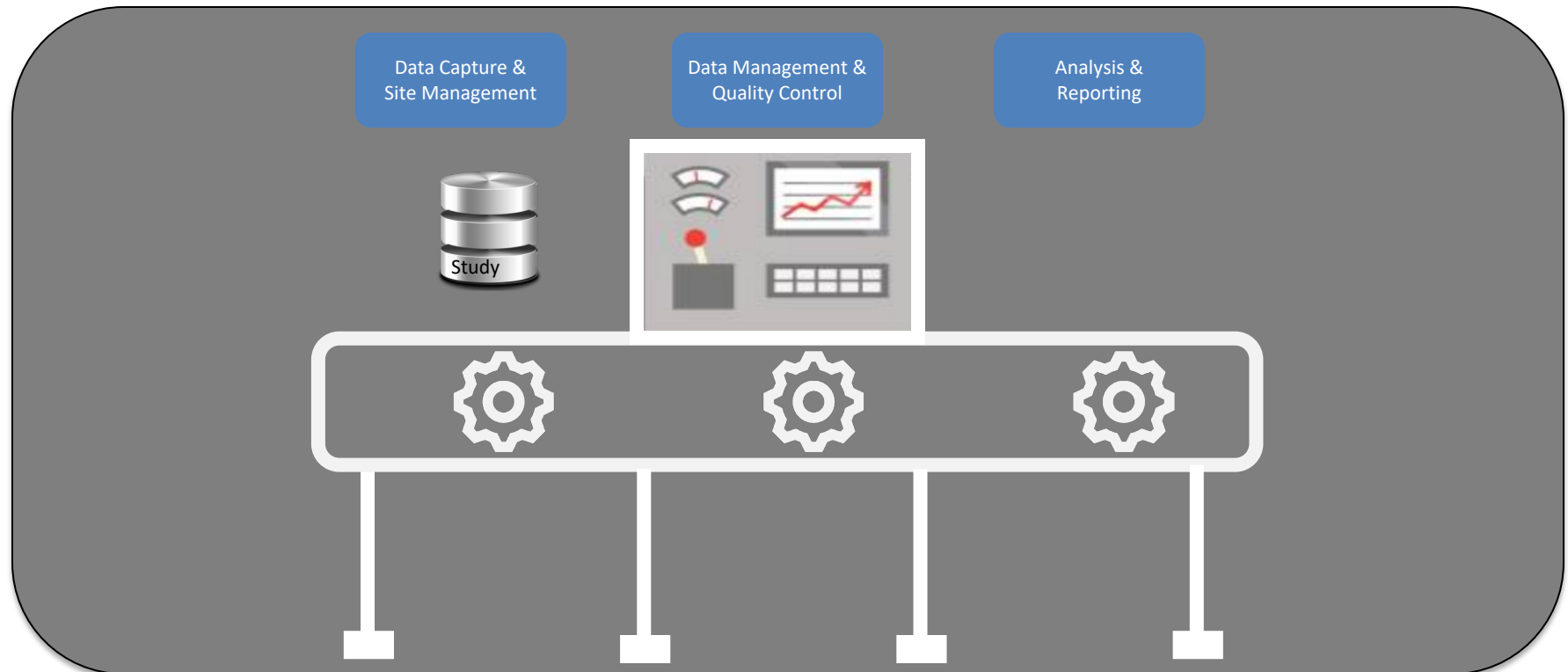
Utilisation

Automation

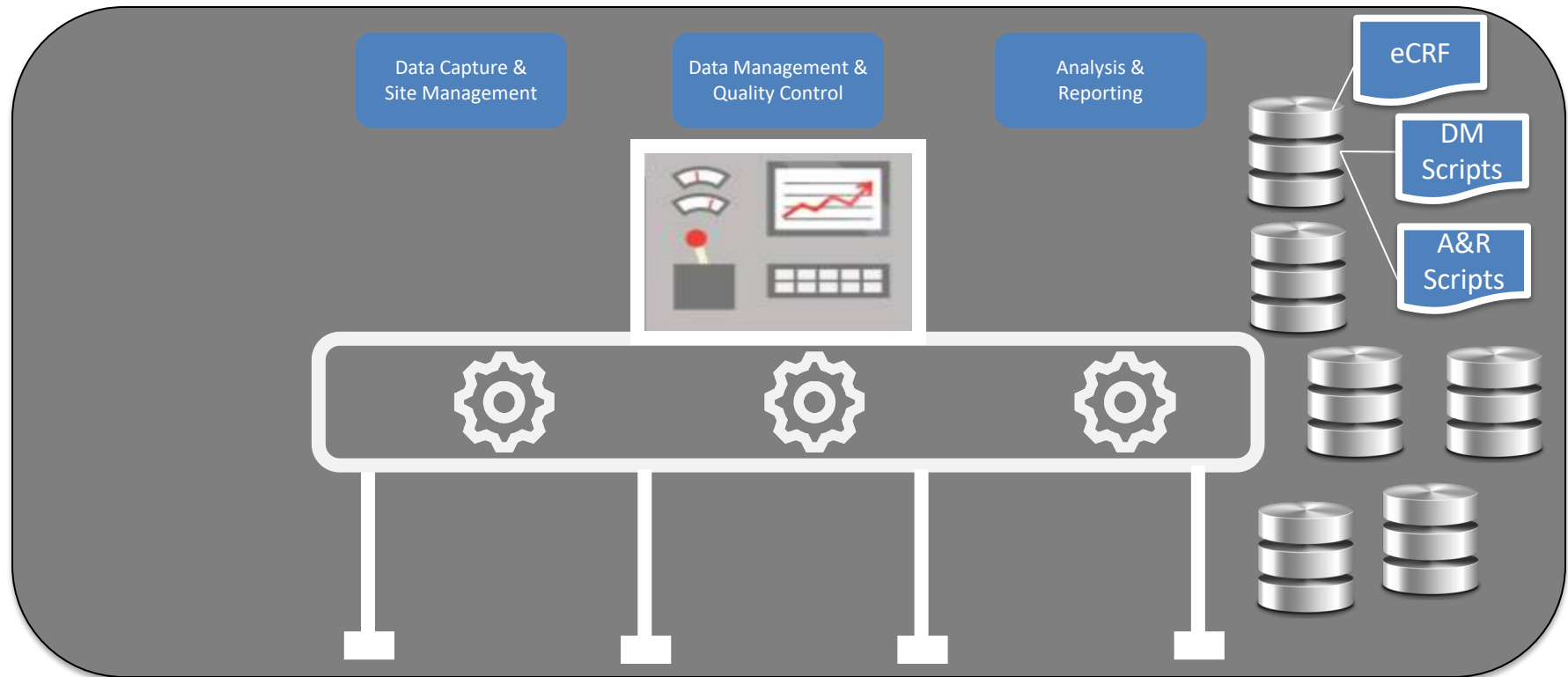
What is Data Engineering?



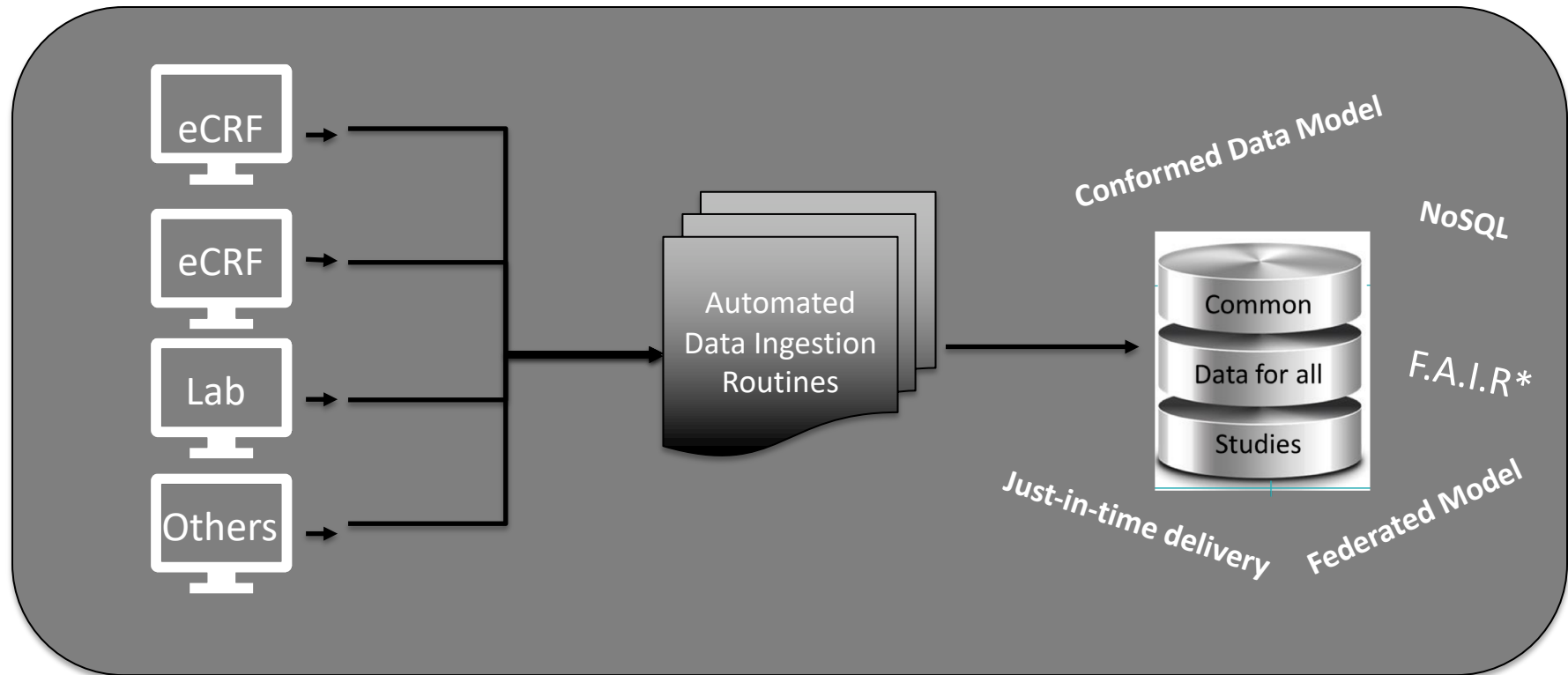
What is Data Engineering?



What is Data Engineering?



What is Data Engineering?



* F.A.I.R. = Findability, Accessibility, Interoperability, Reusability

Data Engineering Project Overview



Data Engineering vs Data Science



Data Engineer Role

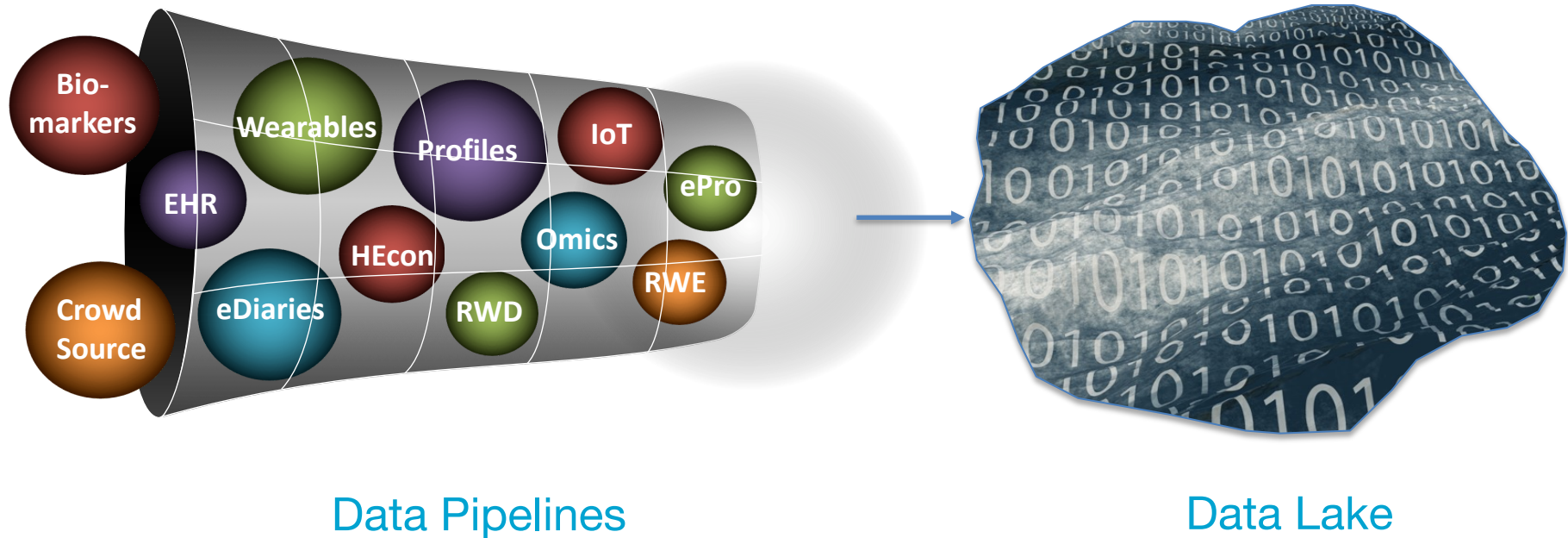
Identification
Automation
Curation



Data Scientist Role

Analysis
Predictions
Serendipity

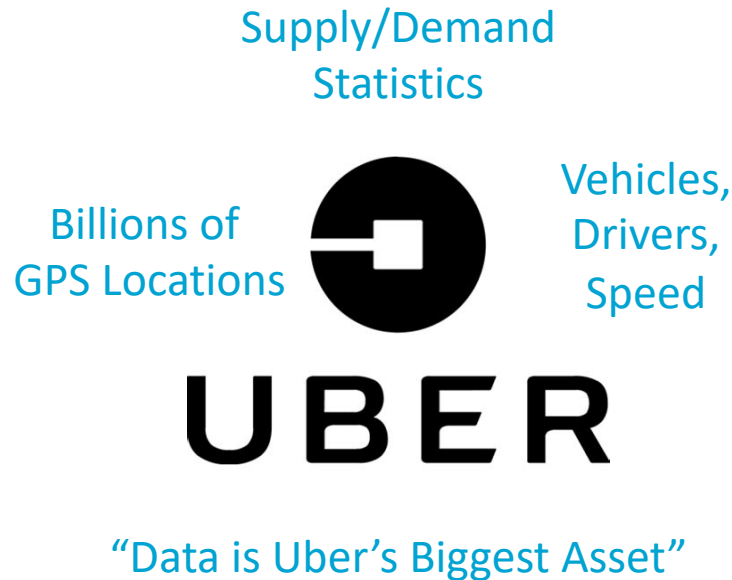
Data Engineering vs Data Science



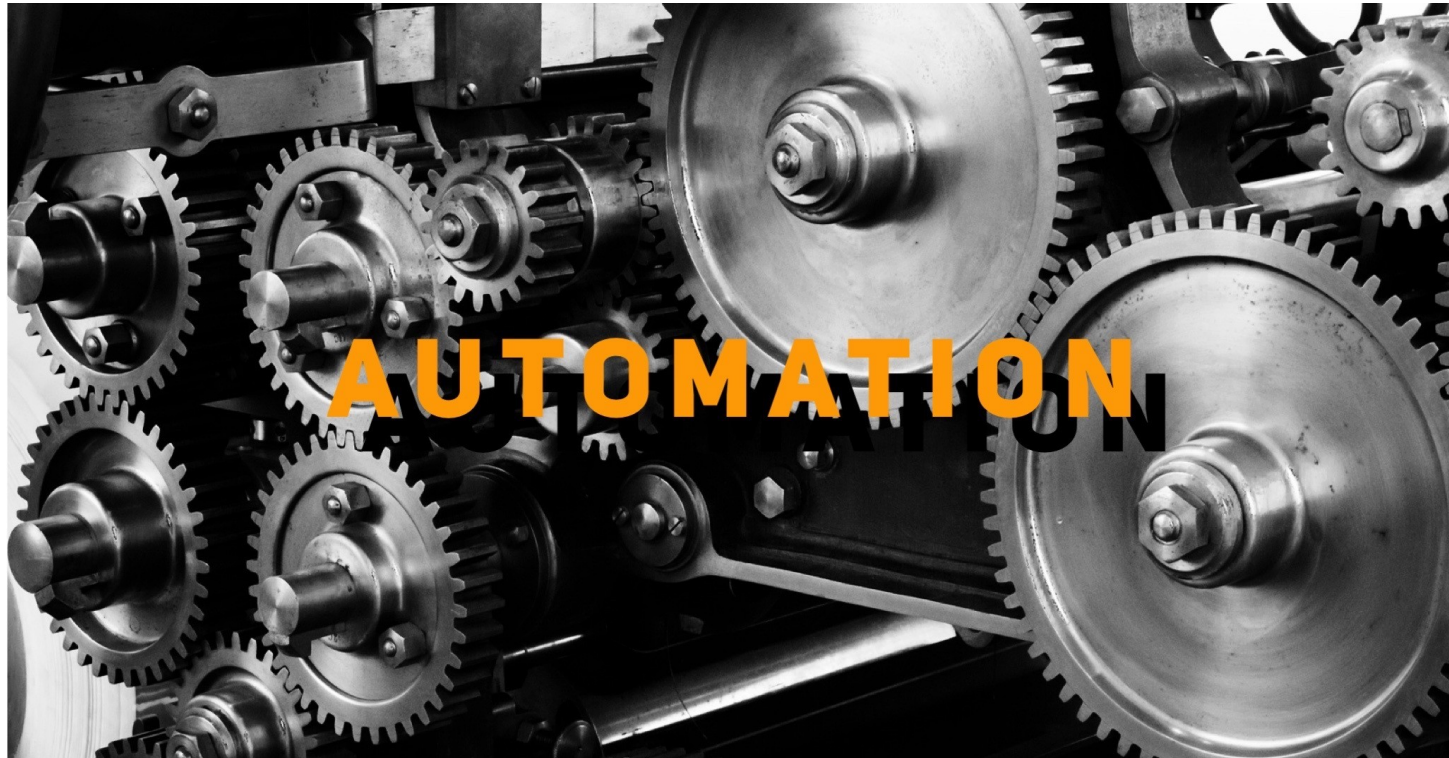
Data Engineering Project Overview



Other Industry Use Cases



Data Engineering Project Overview



Developing an Automation Strategy



Why Automate?

What Should Be Automated?

How Might Automation Impact our Business?

Developing an Automation Strategy

Consider...

- Manual tasks, repetitive or standardized
- Enabling users to be better at their jobs
- Impact to the organization
 - embracing change
 - value creation



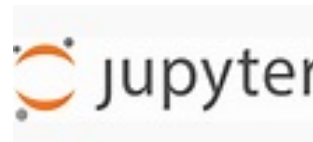
A selection of Automation Software...



Apache Arrow™



SQL/NoSQL



Data Engineering Project Overview



Culture



Patient Safety
Patient Privacy
Effective treatments
High Integrity



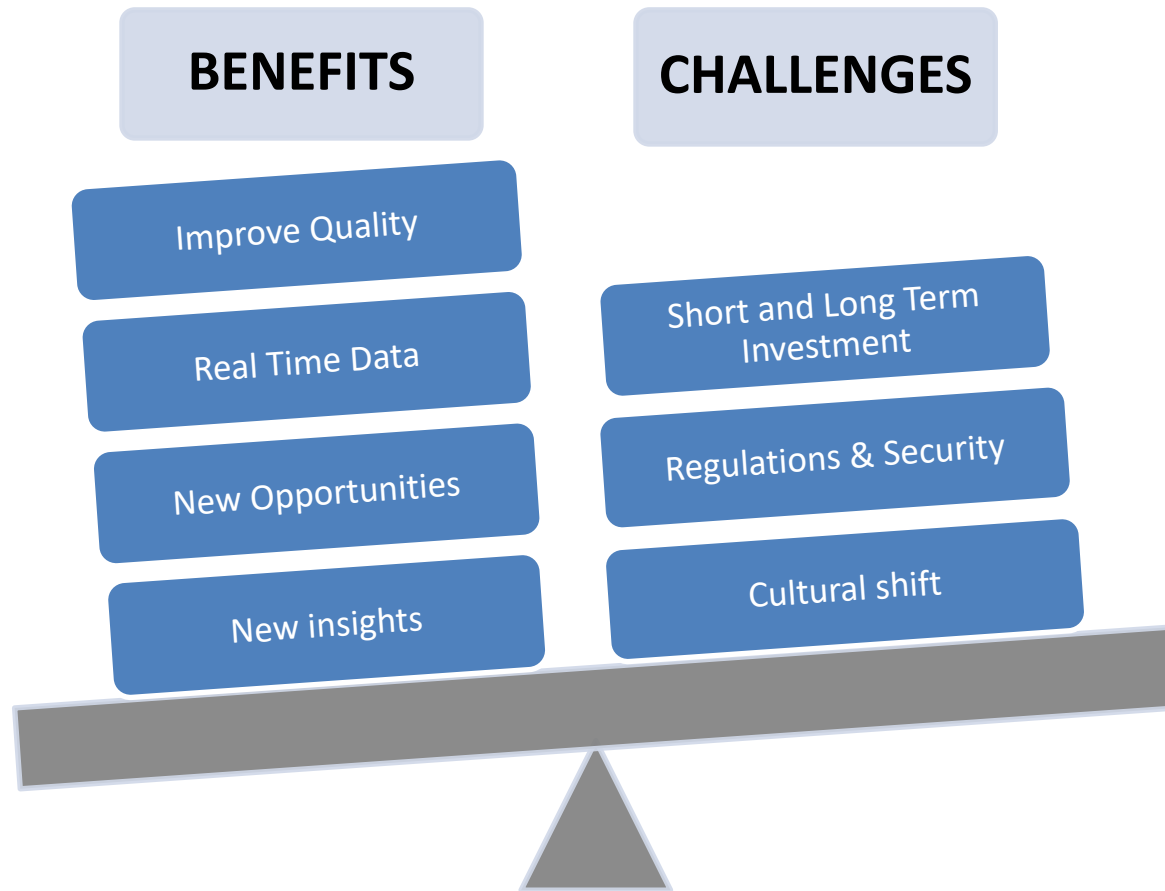
Internal drug development
Licensing & Acquisitions
Distributed data utilization
via Data Engineering



Investment
Legacy Systems
Reduce costs
Shorten Timelines



Benefits & Challenges



Data Engineering Project Overview



Education Methods

- Videos
- Textbooks or articles on selected topics
- Blogs
- Podcasts
- Social Media Channels
- Interactive Components
- Professional trainings
- Pages of/Links to other websites





[Home](#) [Educating for the Future](#) [Industry Starters](#) [Standard Scripts](#) [Events](#)

This section has its focus on the works of the PhUSE **Education for the Future: Data Engineering** work group. With the formation of the work group in early 2018 , the team has taken on the mission to explore how data engineering techniques may affect the industry culture. You will learn about [use cases of data engineering in other industries](#) and how these advances in digital capability has affected their business model. We will also share some of the many software packages and tools available to build further [automation](#) in data engineering and data science.

If you are interested in our up to date news and information, check out our Data Engineering [Blog](#). Here we review books or other resources, introduce new software versions or write about interesting events around the topic Data Engineering.

<https://education.phuse.eu/eftf/data-engineering/>

Data Engineering Project Overview



Data Engineering Project (Educating for the Future PhUSE Working Group)



BHayes2@its.jnj.com

Guy.Garrett@AchieveIntelligence.com

<https://education.phuse.eu/eftf/data-engineering/>

No Copyright infringement intended.