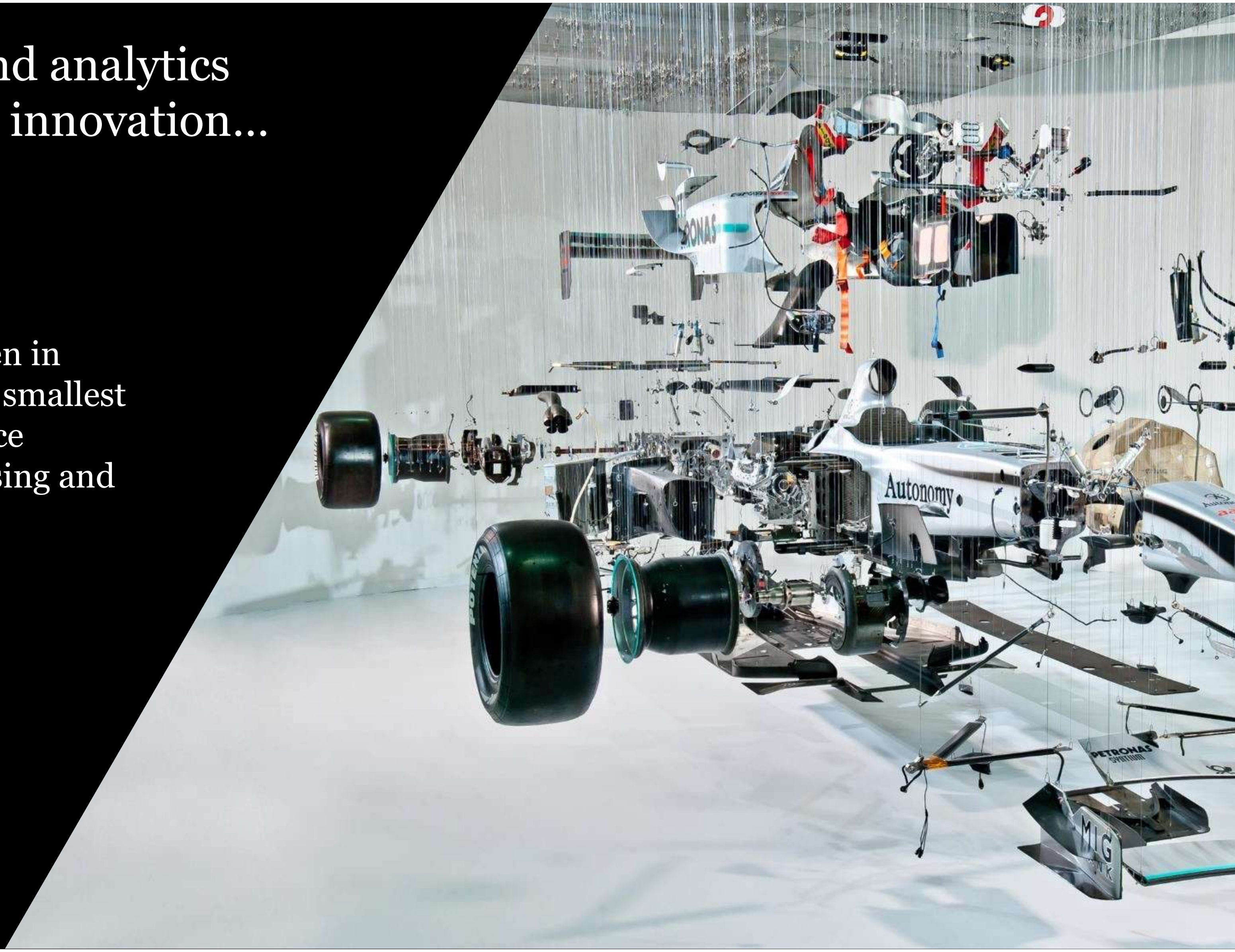
Rethinking organizations in the AI era

Dr Jacomo Corbo
Chief Scientist, QuantumBlack

QUANTUMBLACK
A MCKINSEY COMPANY

QB exploit data and analytics in an arms race in innovation…

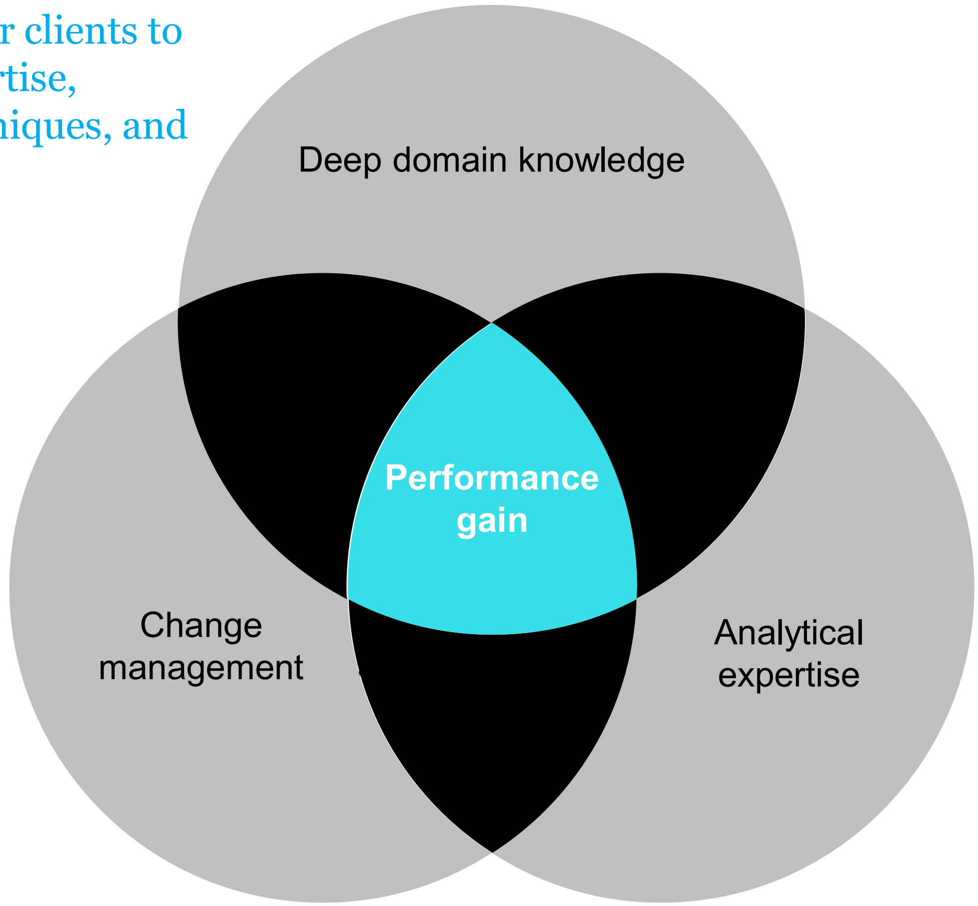
We were born and proven in Formula One, where the smallest margins are the difference between winning and losing and data has emerged as a fundamental element of competitive advantage



QuantumBlack is McKinsey's global centre of excellence and innovation for Advanced Analytics

Our partnership allows our clients to get the best industry expertise, change management techniques, and analytical horsepower

McKinsey & Company



QUANTUMBLACK
A MCKINSEY COMPANY

"The McKinsey of AI may yet turn out to be McKinsey itself"

The Economist

June 22, 2017

QB has grown from 35 to 430 people in the last 4 years, focusing on scaling our impact by combining talent, technology and protocols – and continuous learning!

We have dedicated people …

From 35 people in 2015 to **430** in **2020**

including **>100** PhD Data Scientists

… with distinct capabilities for delivery of advanced analytics…

- Data engineering
- Data science
- User design
- Delivery

… who create high impact solutions across industries…

Healthcare & Pharma


Financial Services


Insurance


Advanced Industries


Telecoms


Natural Resources


Sport


Public & Social Sector


Infrastructure


… and continually learn through proprietary tools and assets…

QBProtocols
Best practice for delivering analytics engagements

Studio
Data management & collaboration application

Kedro
Data and analytics pipeline development framework

CausalNex
Python library to help establish causality

QBBlueprints
Tech blueprints for 1-click deployments of Lab platforms

Performance AI
Tech tooling and playbooks for operational AI

Brix
Capture and share reusable analytics notebooks

… while deploying complimentary McKinsey Capabilities

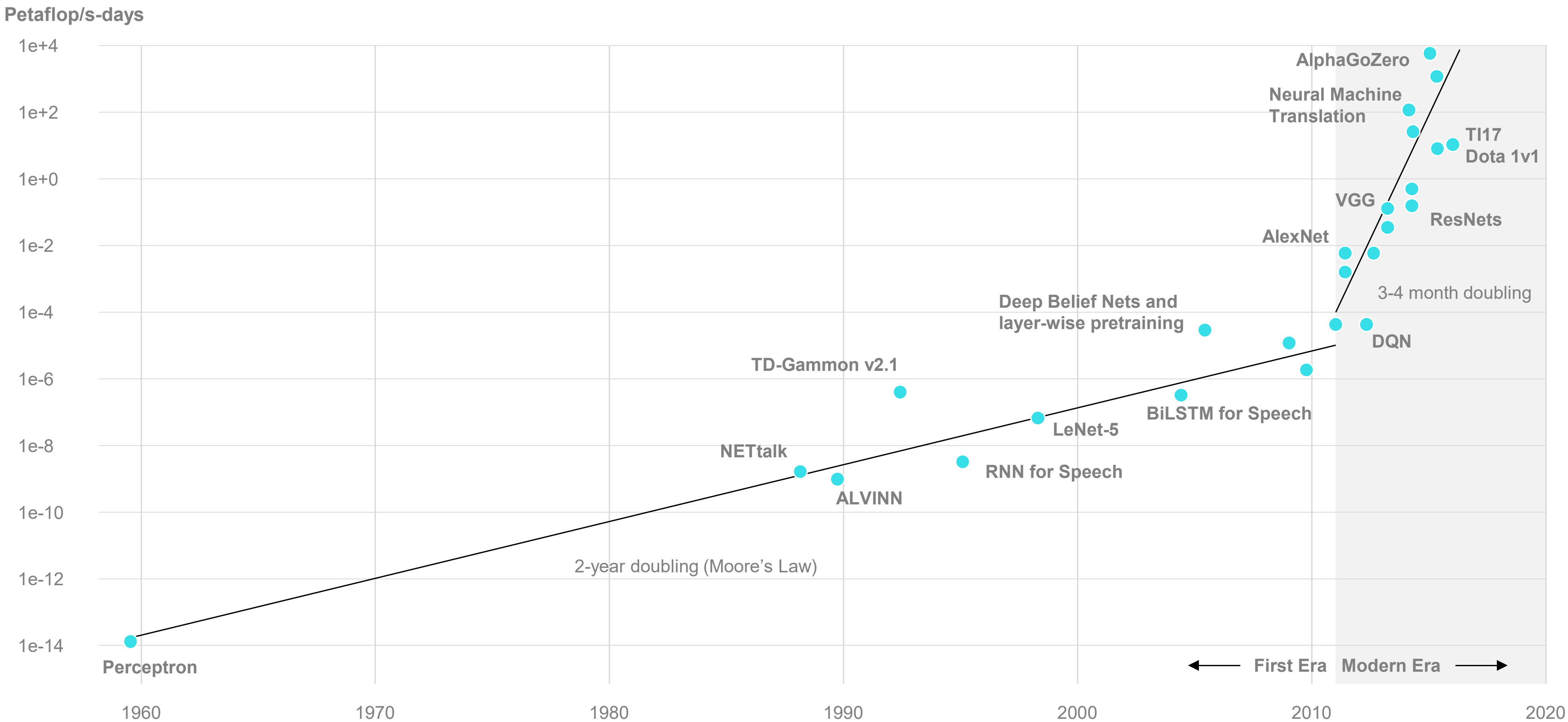
- Domain expertise
- Digital
- Design
- Analytics academy
- Change management and implementation

SOURCE: QuantumBlack + McKinsey


MAJOR 3 TRENDS


1
MORE DATA
2
UNCON-STRAINED COMPUTING POWER
3
FAST-MATURING SOFTWARE STACK

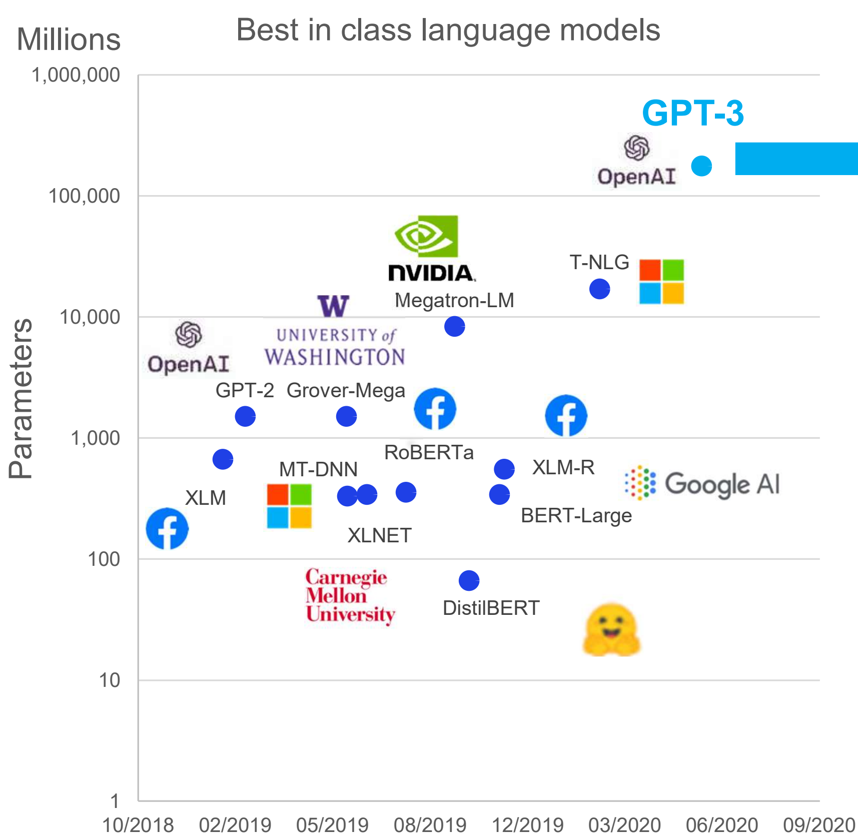
Two Distinct Eras of Compute Usage in Training AI Systems

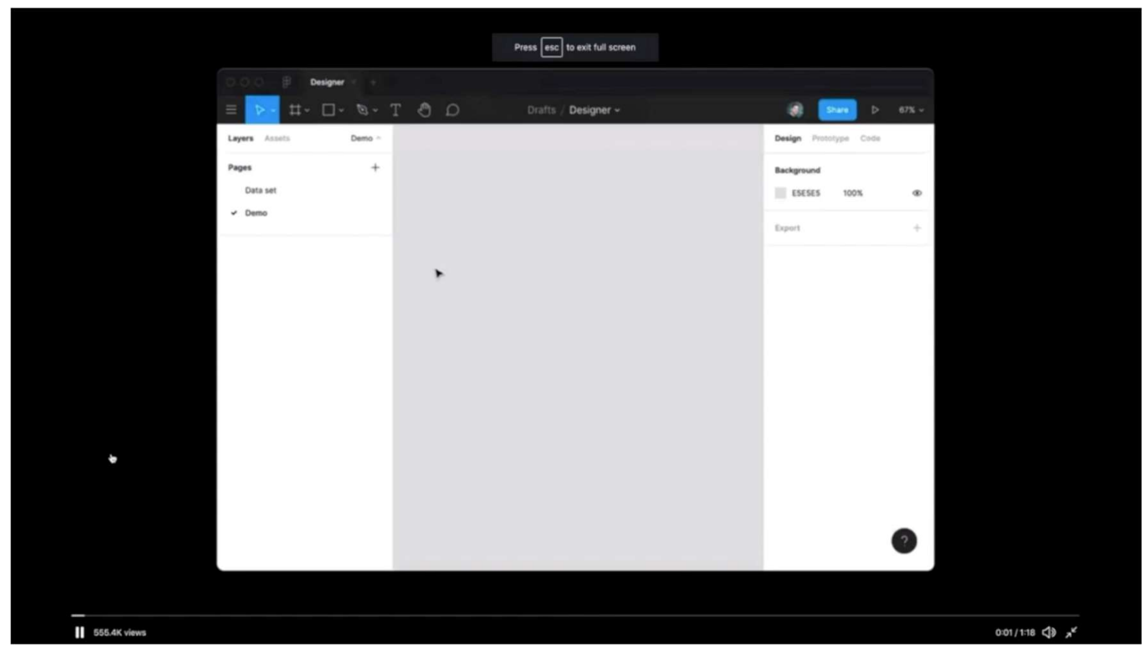


OPENAI

SOURCE: MIT Technology Review October 2019: The computing power needed to train AI is now rising seven times faster than ever before

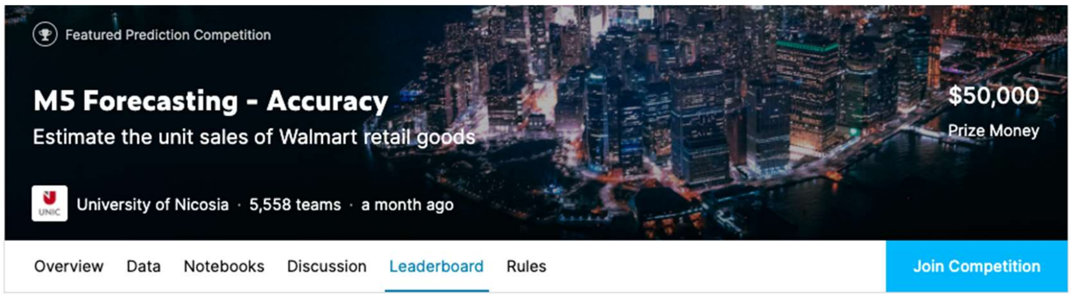
Which improve accessibility of increasingly powerful capabilities





Prestigious competitions are accelerating best-in-class approaches to solving real world problems

Challenge and Approach



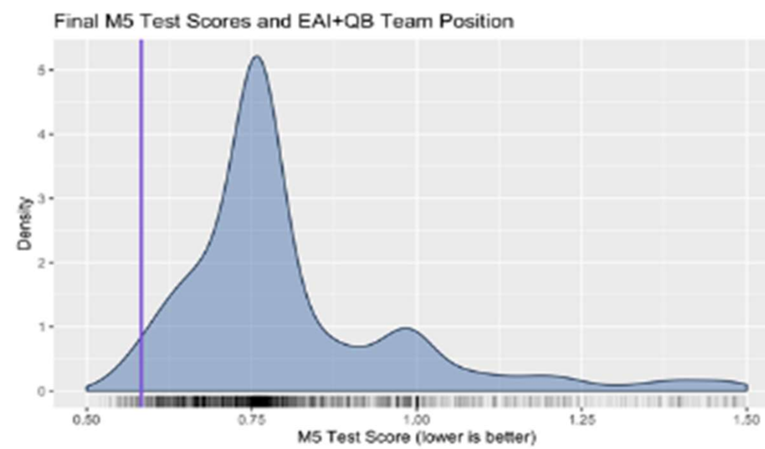
The team created Fusenets, a **neural network based approach** to produce forecasts.

It is applicable to problems where the target prediction is directly driven by the covariates in some multiplicative / additive / functionally defined way. It first estimates **the effect of covariates** on the signal, then **removes it,** and then **forecasts using the residual pattern** left in the historical sales.

The method was tested using **ElementAI's GPU cluster**, which enabled the team to carry out **rigorous backtesting** in order to avoid overfitting to the public leaderboard.

Impact

Team **ranked 65th from more than 5500 entries**, including the winner of the M4 competition (Uber).



And industries are taking advantage of the change

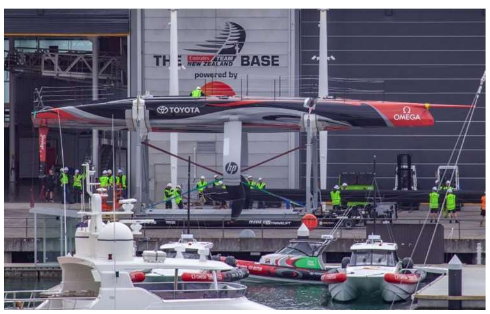
Designing an America’s Cup boat using Reinforcement Learning and Software 2.0 principles

Background

- Since 1851 the Americas Cup has been won by the teams that innovated and advanced technology
- The 36th Americas Cup will be raced in the new AC75 class
- The variety in the four AC75s launched to date highlights the complex choices the designers must make

How to design an Americas Cup winning boat

- Digital simulation drove Emirates Team New Zealand to victory in the 35th Americas Cup
- Boat designs are loaded into the simulator and then sailed by the sailors. However the sailors have limited availability and produce inconsistent results
- By using Deep Reinforcement Learning we are helping ETNZ increase the number of designs they can test and the precision with which they can be evaluated



Emirates Team New Zealand

Prada Luna Rosa Italy

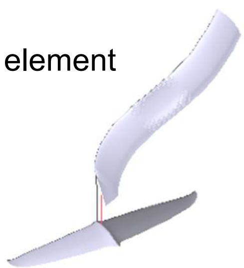
American Magic


INEOS GBR

Foil design is a critical driver of boat performance

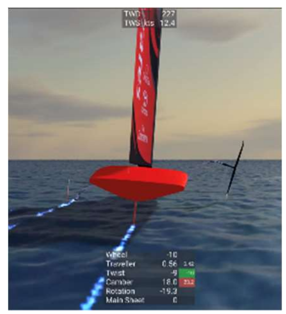
Highly iterative, slow and noisy process

Key design element



Evaluate design performance (sailors in simulator)

Noisy and slow



Specify design (hydrodynamicist)

Fairly ad-hoc

Build 3 pairs of foils for on-water testing and racing

- Current foil design evaluation is slow and noisy: it involves sailors doing multiple laps in the **sailing simulator** for each design
- Can we automate the design evaluation to reduce variability and improve the likelihood of finding an optimal solution?
- In the 35th AC, the performance margin between the boats was 5 seconds per race.
- A 10x increase in the speed of foil development could equate to a 10 second advantage per race

We built an AWS tech stack for large scale Reinforcement Learning

Support distributed computing across 1000s of workers

Frameworks, libraries and tools

Deployment

Scalable Reinforcement Learning

rllib
PyTorch
Gym
RAY
tune
Hydra
python

docker

Infrastructure and Orchestration

Amazon SageMaker
PostgreSQL
Amazon S3
Amazon ECR
Amazon ECS
AWS Fargate

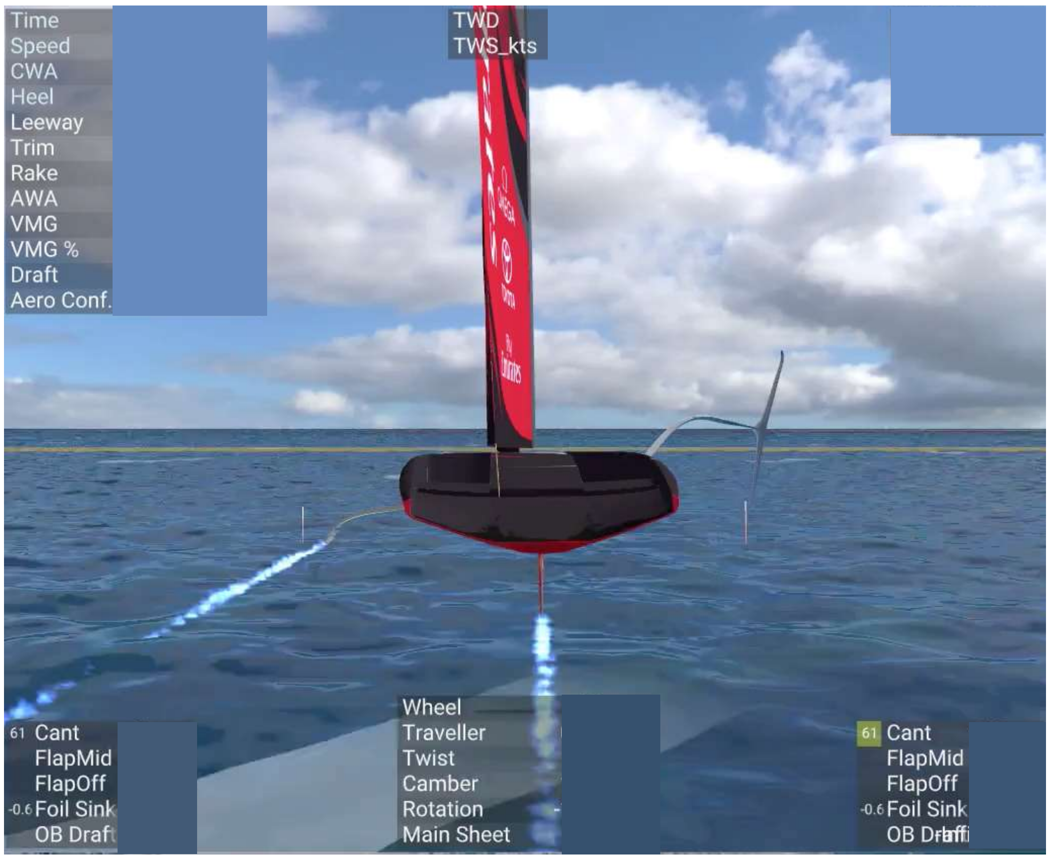
aws Cloud Development Kit

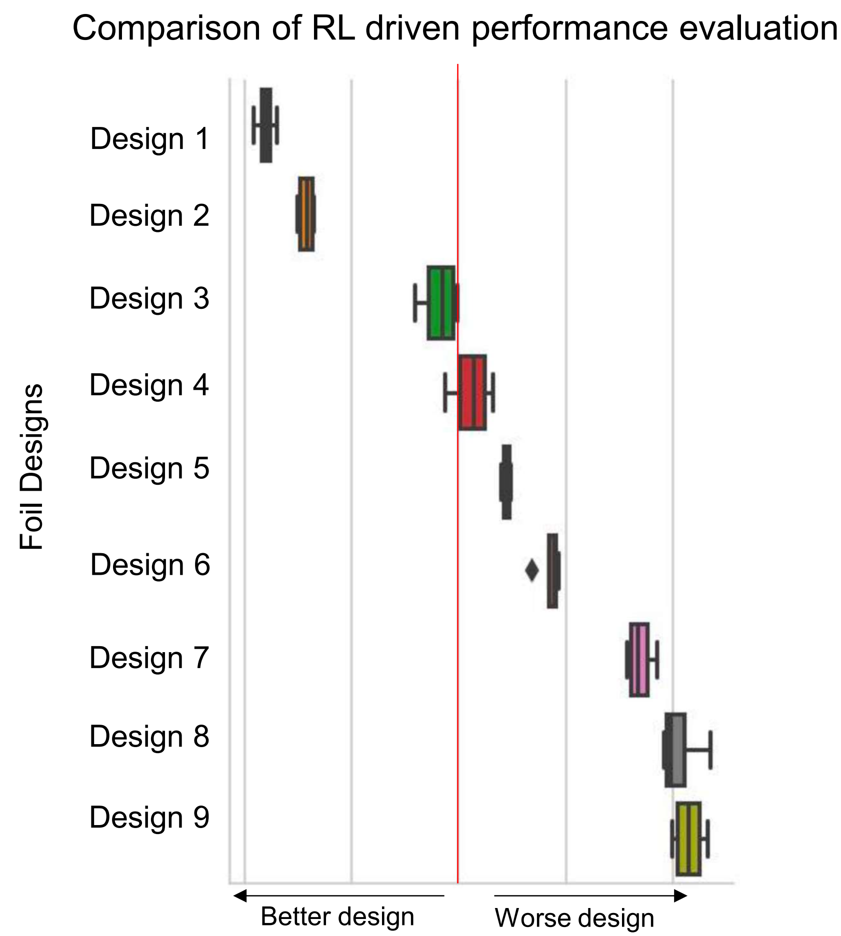
Software 2.0 approach

- Use high level APIs where possible
- Manage large compute costs through emphasizing sample efficiencies and focusing effort on value adding models not yet commoditized (scalable Soft Actor Critic)

  Fully scripted client infrastructure deployment
- End deliverable is a set of pre-tuned weights for 2 neural networks

RL allows to rank foil designs with high statistical significance

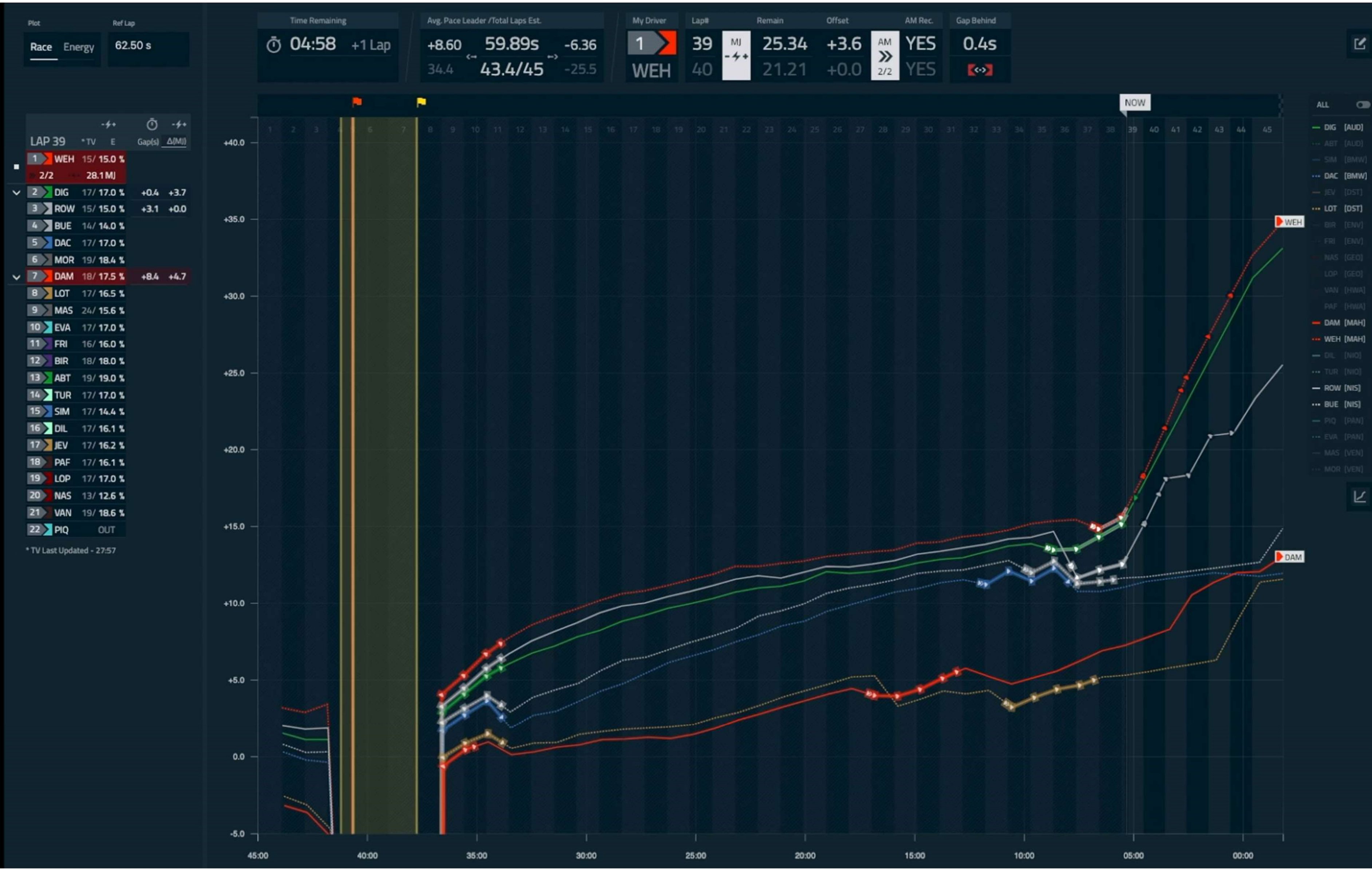




How can we help a Formula E team make decisions about race strategy?

QUANTUMBLACK
A MCKINSEY COMPANY

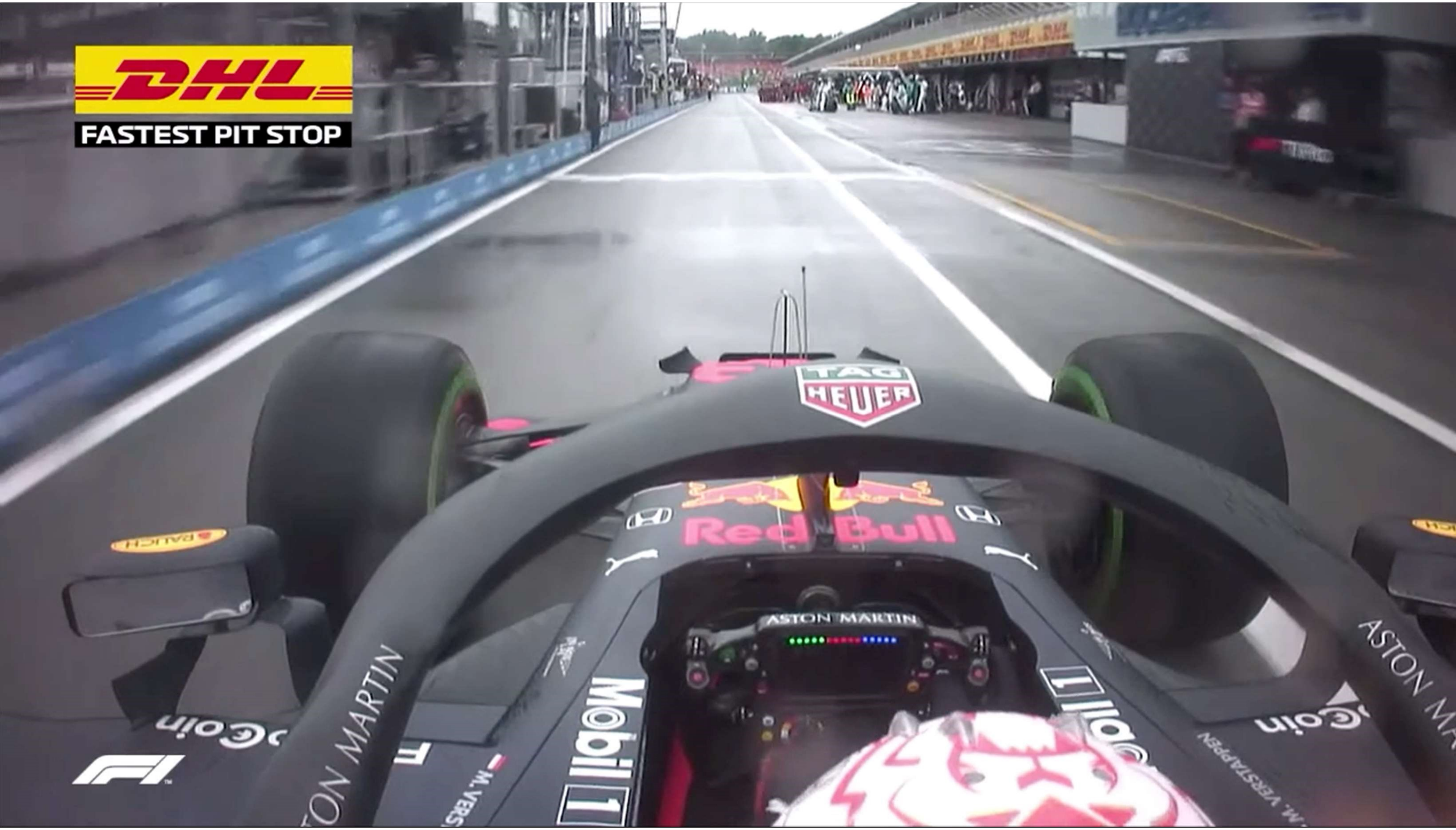

45:00 +1 LAP
ABB FORMULA-E
FIA FORMULA-E CHAMPIONSHIP
CBMM NIOBIUM
MEXICO CITY
E-PRIX 2019
FORMULA-E
FIA
CBMM
Niobium
TAG Heuer
DHL
#MEXICO



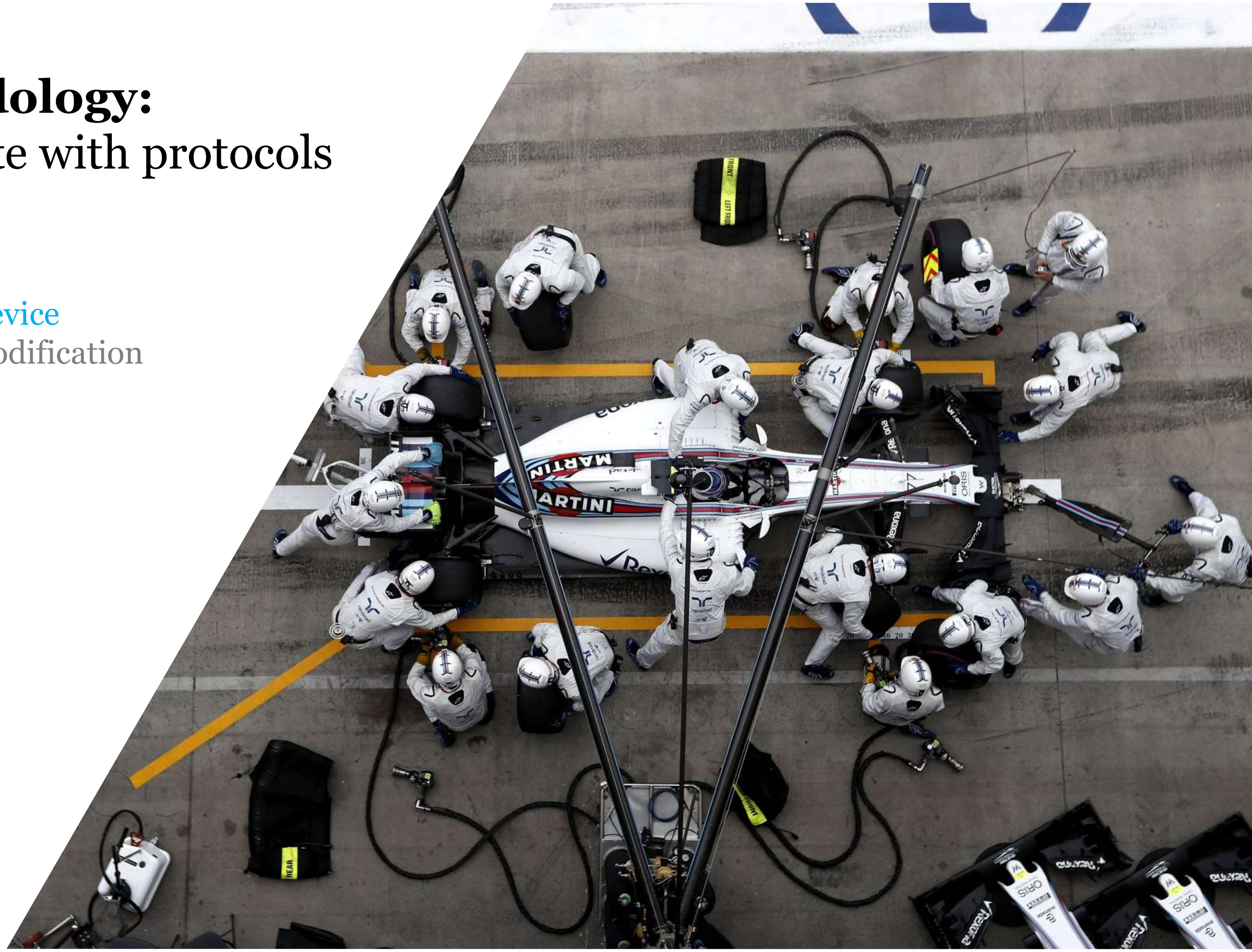
Plot
Race Energy
Ref Lap
62.50 s
Time Remaining
04:58 +1 Lap
Avg. Pace Leader /Total Laps Est.
+8.60 59.89s -6.36
34.4 43.4/45 -25.5
My Driver
1 WEH
Lap#
39
40
MJ
Remain
25.34
21.21
Offset
+3.6
+0.0
AM Rec.
AM 2/2
YES
YES
Gap Behind
0.4s
NOW
LAP 39 *TV E Gap(s) Δ(MJ)
1 WEH 15/ 15.0 %
2/2 28.1 MJ
2 DIG 17/ 17.0 % +0.4 +3.7
3 ROW 15/ 15.0 % +3.1 +0.0
4 BUE 14/ 14.0 %
5 DAC 17/ 17.0 %
6 MOR 19/ 18.4 %
7 DAM 18/ 17.5 % +8.4 +4.7
8 LOT 17/ 16.5 %
9 MAS 24/ 15.6 %
10 EVA 17/ 17.0 %
11 FRI 16/ 16.0 %
12 BIR 18/ 18.0 %
13 ABT 19/ 19.0 %
14 TUR 17/ 17.0 %
15 SIM 17/ 14.4 %
16 DIL 17/ 16.1 %
17 JEV 17/ 16.2 %
18 PAF 17/ 16.1 %
19 LOP 17/ 17.0 %
20 NAS 13/ 12.6 %
21 VAN 19/ 18.6 %
22 PIQ OUT
* TV Last Updated - 27:57
+40.0
+35.0
+30.0
+25.0
+20.0
+15.0
+10.0
+5.0
0.0
-5.0
45:00
40:00
35:00
30:00
25:00
20:00
15:00
10:00
05:00
00:00
WEH
DAM
ALL
DIG [AUD]
ABT [AUD]
SIM [BMW]
DAC [BMW]
JEV [DST]
LOT [DST]
BIR [ENV]
FRI [ENV]
NAS [GEO]
LOP [GEO]
VAN [HWA]
PAF [HWA]
DAM [MAH]
WEH [MAH]
DIL [NIO]
TUR [NIO]
ROW [NIS]
BUE [NIS]
PIQ [PAN]
EVA [PAN]
MAS [VEN]
MOR [VEN]

Protocol as accelerators for learning & assetization

FORMULA 1 PIT STOPS

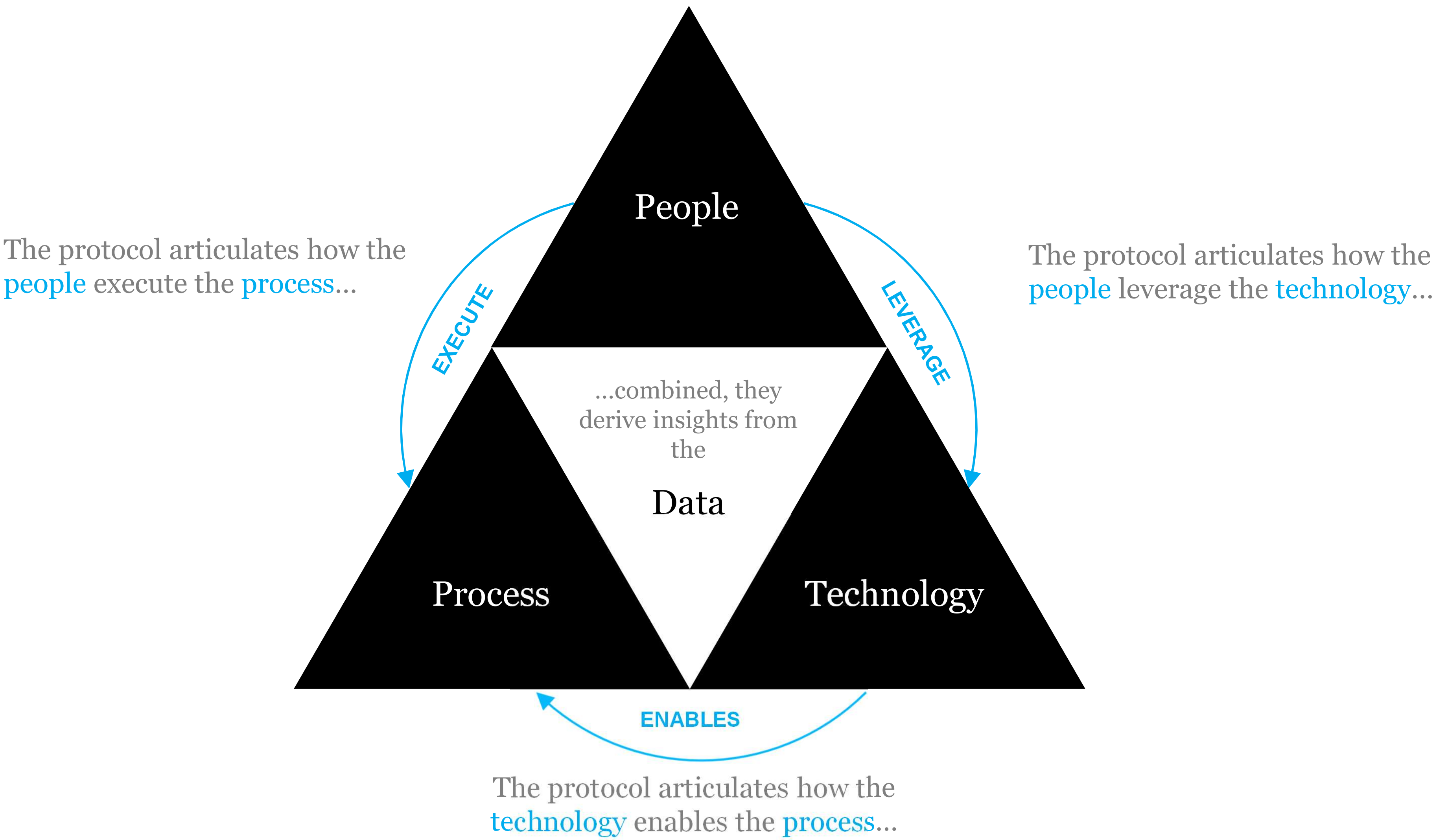

DHL
FASTEST PIT STOP
TAG HEUER
Red Bull
ASTON MARTIN
Mobil 1
M. VERSTAPPEN

Delivery methodology:
Elite teams compete with protocols



Repeatability is a scaling device

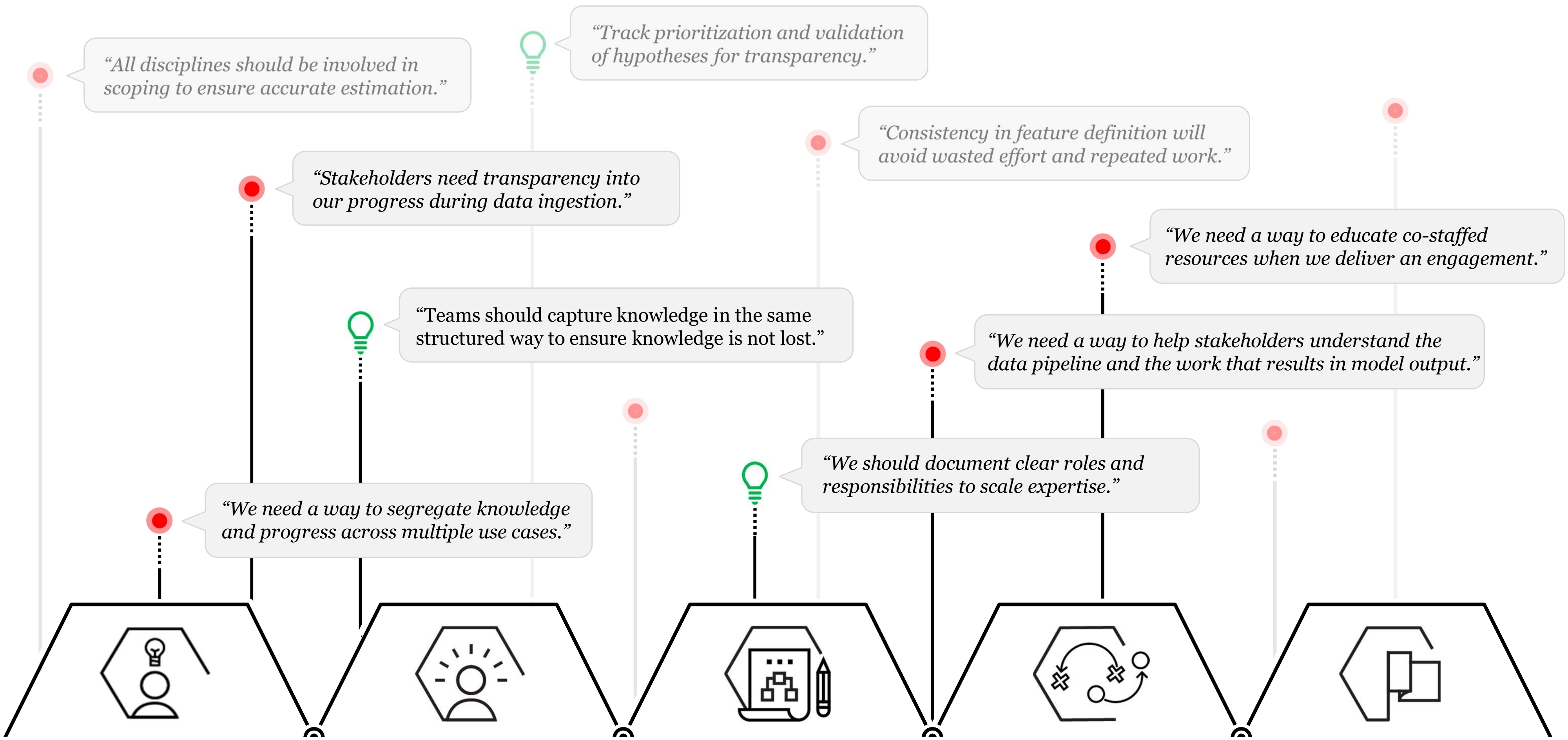
Practice, consistency and codification


The protocol articulates how the people execute the process…
The protocol articulates how the people leverage the technology…
People
EXECUTE
LEVERAGE
…combined, they derive insights from the
Data
Process
Technology
ENABLES
The protocol articulates how the technology enables the process…

How do protocols accelerate learning?[2]

The protocol provides a scaffold for capturing historical learnings and reported pain points – frequency informs prioritization for codification into the protocol.

LEARNING PAIN POINT



How do protocols accelerate learning?[1]

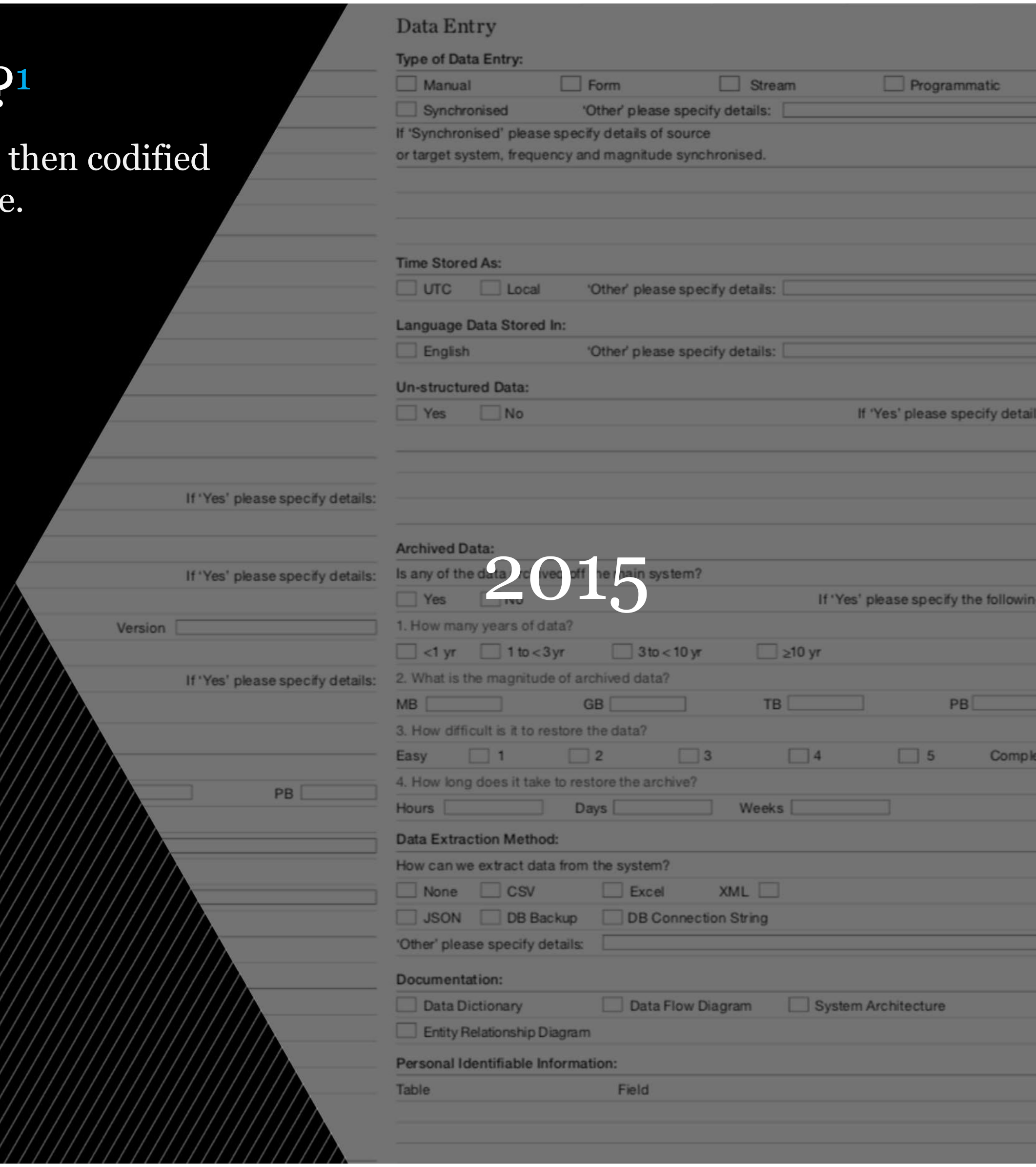
Learnings are first codified into best practice. Best practice is then codified into technology. This unlocks adoption of the protocol at scale.

PAIN POINT

"Project knowledge is unstructured, in multiple locations and lost at the end of an engagement."

SOLUTION - 2015

Meta information for data dictionaries captured in PDF and XLS templates.



REQUIRED GOVERNANCE 1 2 3 4

PROCESS ADOPTION 1

How do protocols accelerate learning?[2]

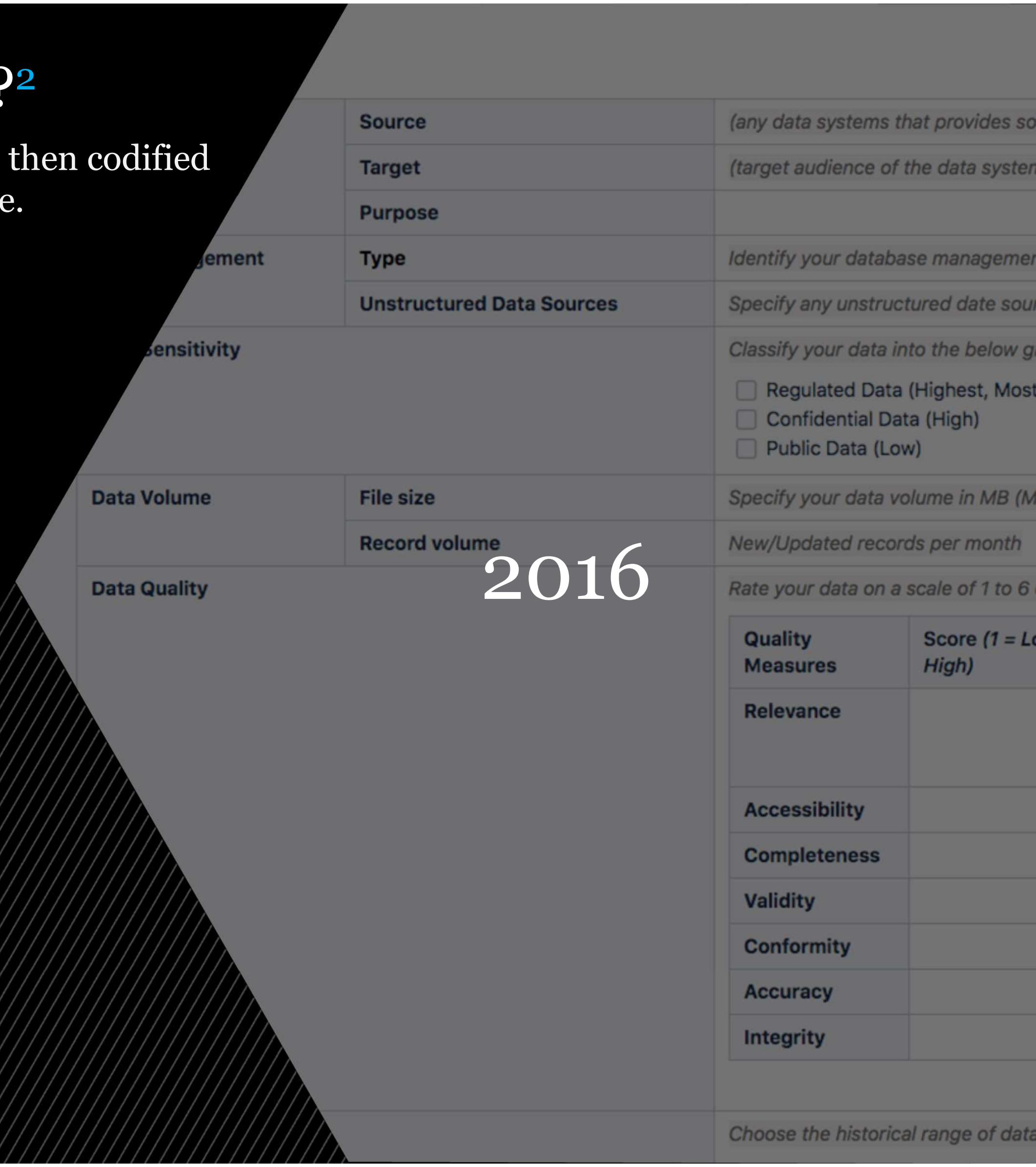
Learnings are first codified into best practice. Best practice is then codified into technology. This unlocks adoption of the protocol at scale.

PAIN POINT

"Project knowledge is unstructured, in multiple locations and lost at the end of an engagement."

SOLUTION - 2016

Data due diligence questions built into confluence so the whole team can access the information.

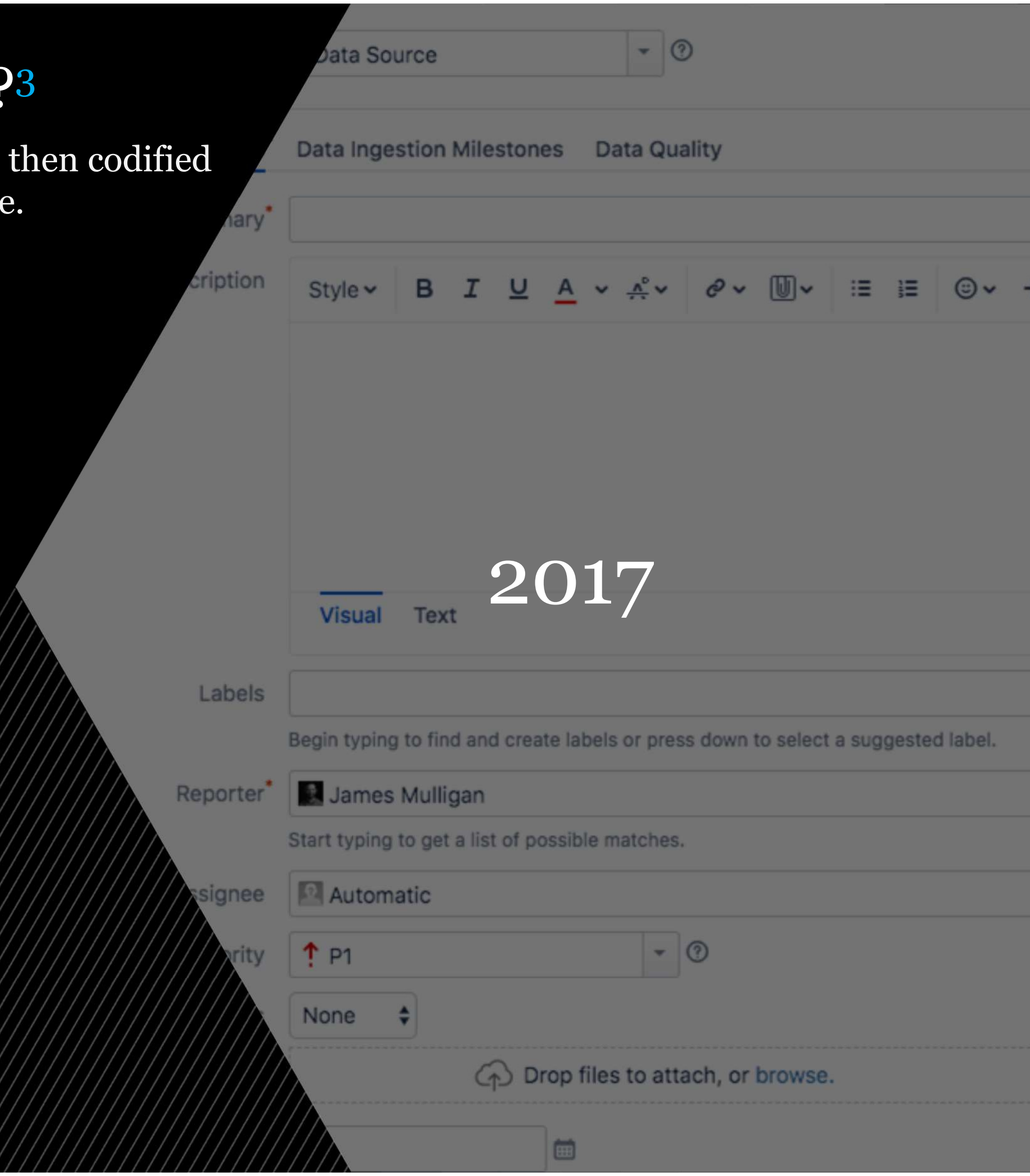


REQUIRED GOVERNANCE 1 2 3

PROCESS ADOPTION 1 2

How do protocols accelerate learning?[3]

Learnings are first codified into best practice. Best practice is then codified into technology. This unlocks adoption of the protocol at scale.



PAIN POINT

"Project knowledge is unstructured, in multiple locations and lost at the end of an engagement."

SOLUTION - 2017

Data due diligence questions built into JIRA tickets, so ingestion progress can also be tracked.

REQUIRED GOVERNANCE 1 2

PROCESS ADOPTION 1 2 3

How do protocols accelerate learning?[4]

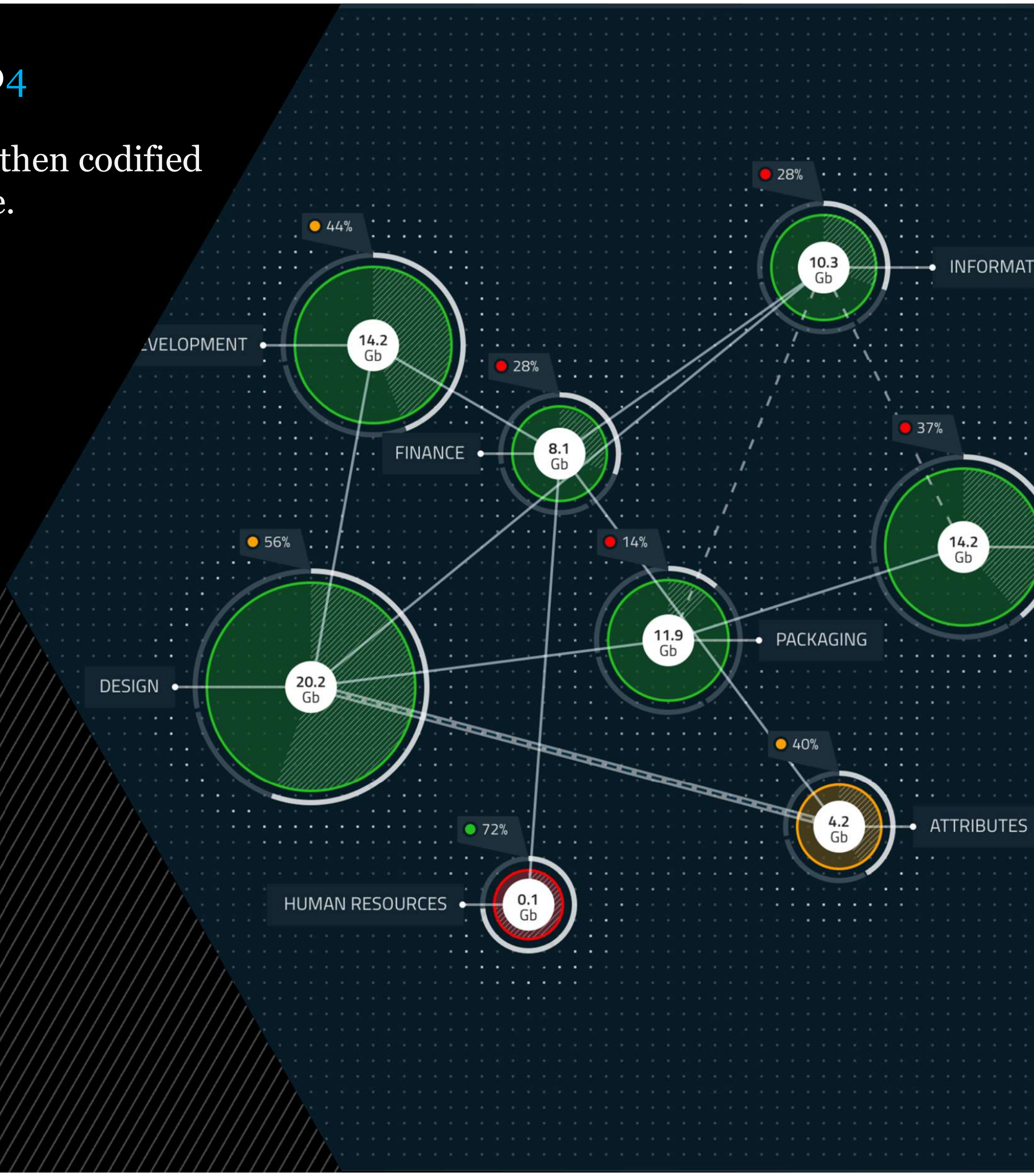
Learnings are first codified into best practice. Best practice is then codified into technology. This unlocks adoption of the protocol at scale.

PAIN POINT

"Project knowledge is unstructured, in multiple locations and lost at the end of an engagement."

SOLUTION - 2018

Learnings consolidated into Studio, to provide a single source of truth for teams and clients – visualizing the overall data landscape and ingestion progress.



REQUIRED GOVERNANCE 1

PROCESS ADOPTION 1 2 3 4

A horizontal assets ecosystem to improve end-to-end model lifecycle development and management

QB/Protocols

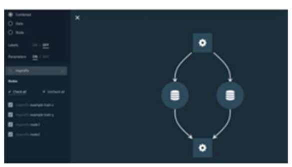
Protocols for delivering Analytics transformations

A codification of QB's best practices for delivering AI @ scale provided by technical practitioners that continues to evolve based on new learnings and experiences.

Kedro

Data and analytics pipeline development framework

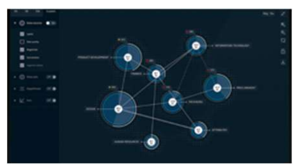
Kedro is a workflow development tool that helps you build data pipelines that are robust, scalable, deployable, reproducible and versioned.



Studio

Data management application

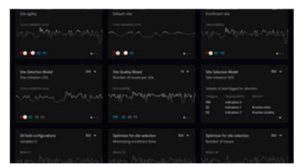
An application for capturing project knowledge, hypotheses, data source metadata, data landscapes, data ingestion status, and data quality metrics.



PerformanceAI

Tech tooling and playbooks for operational AI

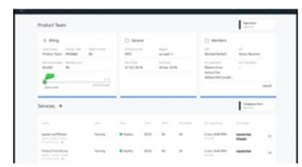
A product for sustaining model performance in a live environment and frameworks to accurately scope and consistently deliver models into production.



Platform McKinsey

Self-service cloud infra for analytics engagements

Automated provisioning of templated analytics environments with a single management layer for all security policies and monitoring, logging, audit.



CausalNex

Python library to help establish causality

A Python library that leverages Bayesian Networks to help data scientists to infer causation rather than observing correlation.

Note: We recently open sourced Kedro & CausalNex – video available here: https://youtu.be/KEdmJ2ADy_M

Technology

Brix

Brix aims to scale firm knowledge and improve the way McKinsey data scientists, data and machine learning engineers collaborate across projects, in order to:

- Accelerate project delivery, thereby shortening time to impact

- Reduce project costs

- Lower the risk of untested, outdated and fragile code

- Enable the distribution and use of novel capabilities on the back of R&D initiatives.

Brix aspires to achieve that by becoming the single source of truth for technical analytic assets ranging from small code snippets through component pipelines to entire end-to-end ML cookbooks together with any associated data schemas and business hypotheses. Brix is excited to leverage the internal McKinsey community of nearly 2000 ML practitioners in order to create a dynamic content ecosystem, where users can contribute, improve and reuse analytics assets. It endeavours to become the preferred way to share and consume technical analytics assets across the Firm

Technology

OptimusAI

OptimusAI accelerates the application of advanced analytics in processing operations (e.g. manufacturing, mining, refining), by leveraging prebuilt code, algorithms, automated processes, and visualisation tools.

Processing can be highly complex, with variability in feed conditions, interactions across thousands of variables, and trade-offs across multiple performance objectives.

While it is common to have process controls systems in place and access to enormous amounts of data, most plants are not fully optimised to maximise performance across operation conditions.

OptimusAI is a proprietary solution for developing control room advisor systems to improve

Example case: Freeport-McMoRan

Case study: Freeport

2nd
Largest copper producer

1st
Largest molybdenum producer

3.8 bn
pounds of copper sales in 2018

12
Operating mines worldwide

Global Leader
in mining and mineral processing

50k
Employees worldwide

Copper

Molybdenum (MO)

Gold

Freeport-McMoran Inc

Freeport turns to artificial intelligence to raise copper output by 90,000 tonnes

US company to roll out machine learning technology across mines in the Americas

Neil Hume, Natural Resource Editor 6 HOURS AGO

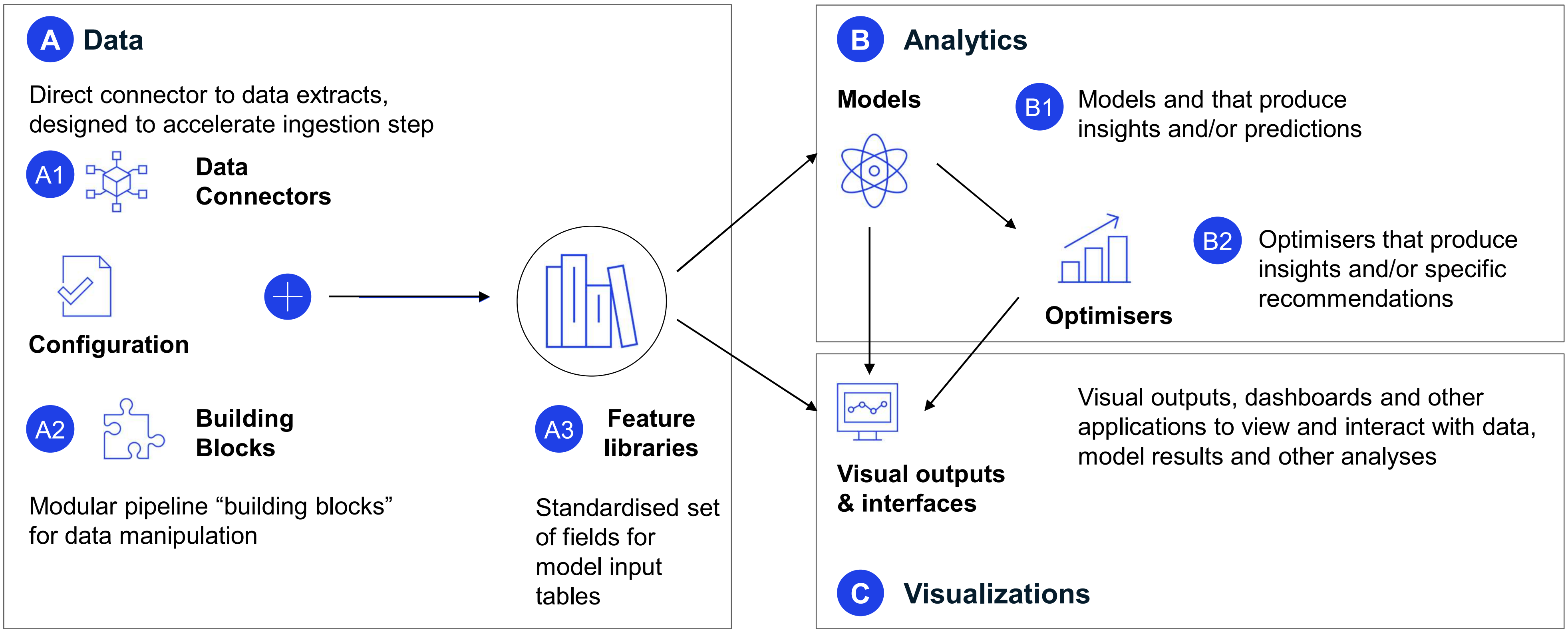
Freeport-McMoran, one of the world's biggest copper producers, is targeting a 90,000-tonne increase in annual output from the introduction of machine learning technology at its mines.

The US-listed company has been testing an artificial intelligence model at its Bagdad mine in Arizona, which it is now planning to roll out across all of its operations in the Americas. The move will lift its yearly copper production by about 5 per cent.

"We have set an aspirational goal of adding 200m pounds [90,000 tonnes] of copper from these initiatives with very little capital investment," Richard Adkerson, chief executive, said in an interview in London.

Copper is critical to the switch away from fossil fuels to renewable energy, with the red metal used in wind turbines, batteries that power electric vehicles and charging points.

PMPX deployable analytics assets allow a full pipeline to be quickly configured and deployed, then iteratively improved



Overview of relevant case examples

	Context	Market	Impact
A	Understand drivers of physician prescribing, improve actions and develop a vision for Advanced Analytics	US	+1mn physicians 15 – 20% sales impact
B	Aiming to accelerate growth of biologic drug – 8-10 yrs after launch for a top 10 Biopharma	Germany	30 – 40% of calls reallocated 7 – 15% sales impact
C	Spanish affiliate of top 10 global pharmaco with a new specialty drug 18 months post launch	Spain	9% estimated sales impact 7% of opportunity captured
D	Create multi-channel engagement engine to provide physician level recommendations in EU countries	EU	50 mEUR revenue increase over 3 year window 100% sales uplift
E	Transforming pharma customer experience	Canada and 4 large European markets	New operating model New Agile way of working

The AI era requires organizations to operate differently in 7 fundamental ways…

	Yesterday's	AI-era
1	**Data collected** but not actioned consistently	**Data-driven decision-making** driving competitive insights
2	**Piece-meal investments** in capabilities targeted areas	**All-in approach** to invest in capabilities in multiple domains across the organization
3	**Technology is the "back office"**, supports the rest of the organization	**AI/ Technology is the "front office",** embedded throughout the organization
4	**Siloed execution** of largely routine tasks	**Cross-functional and agile collaboration** to solve dynamic problems
5	**Business case** focused execution	**Pilot focused execution** at incredible speed as cost of experimentation goes to zero
6	**Deep experts** know one specific area	**Lifelong learning** to develop expertise across and within domains
7	**Stress-testing and certainty** based decision making	**Speed and agility** based decision making

"Companies that inject big data and analytics into their operations show productivity rates and profitability that are 5% to 6% higher than those of their peers"

Andrew McAfee & Erik Brynjolfsson
HBR: Big Data: The Management Revolution

"While cutting-edge technology and talent are certainly needed, it's equally important to align a company's culture, structure, and ways of working to support broad AI adoption…"

Tim Fountaine, Brian McCarty & Tamim Saleh
HBR: Building the AI-Powered Organization

QUANTUMBLACK
A MCKINSEY COMPANY