# DevOps Enterprise Summit 2022
# Data Operations

Simone Steel
Chief Data and Analytics Officer &
CIO for Enterprise Data Platforms

# A bit about Nationwide…

We're a **building society**, or **mutual**, **owned by our members**. That's anyone who banks, saves or has a mortgage with us.

We're run for **their benefit** and to **help the communities** around us.

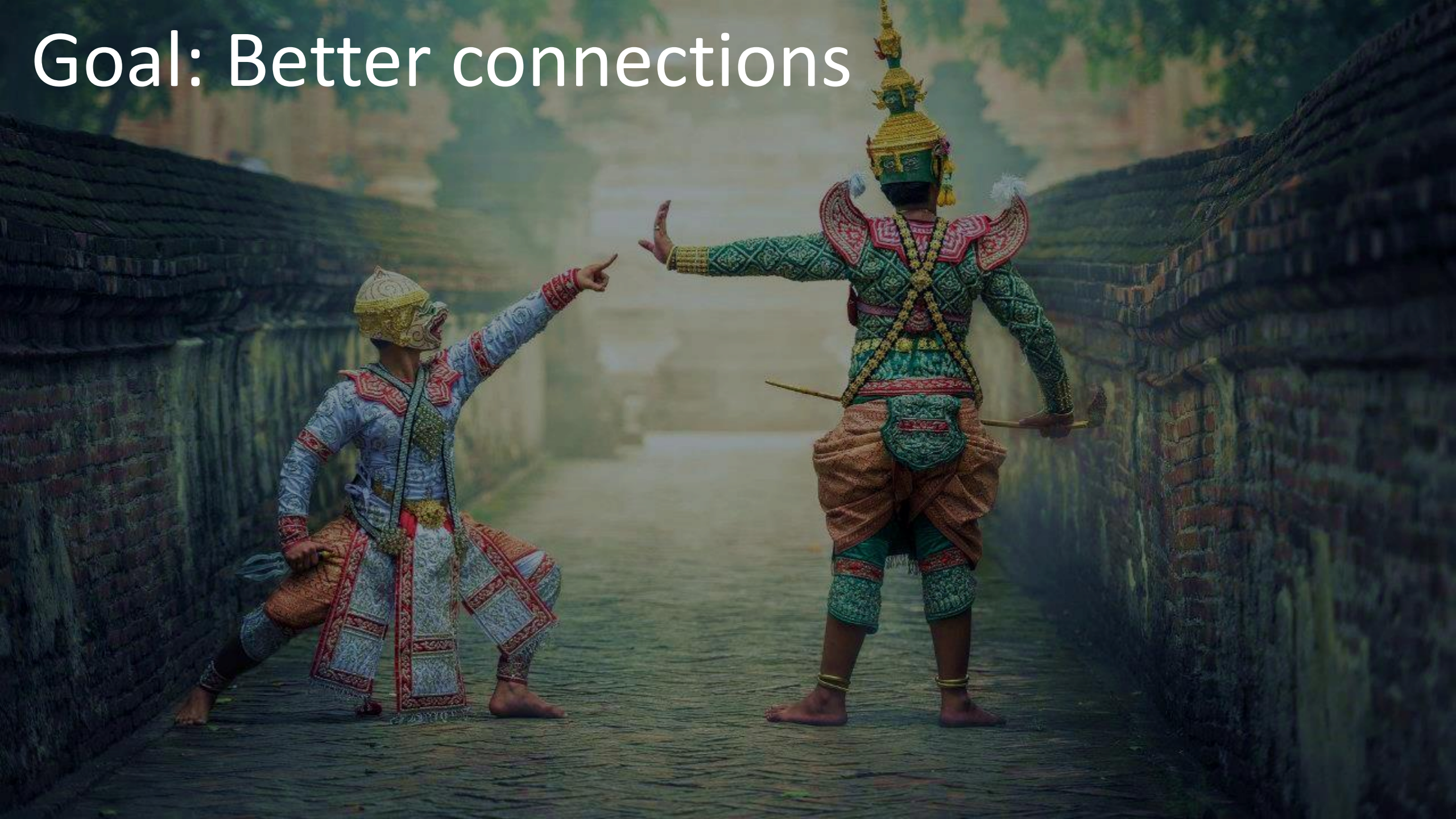
- ✓ Founded in 1884
- ✓ UK Financial Institution
- ✓ Size of a FTSE 250 company
- ✓ c.17 million members
- ✓ +15,000 colleagues in branches and offices

# How did I get here?

Technology and Data Services
Academia – Computing, Economics, Data Science
Financial Services

… 80’s 90’s 2000’s 2010’s 2020’s…

# Goal: Better connections



# Data

Business or Technology?
Business and Technology?

# My job in a nutshell…

to ensure that the **use and management of data** in the Society is **ethical, valuable, safe and compliant**

Claire Tracey

Chief Strategy and Sustainability Officer

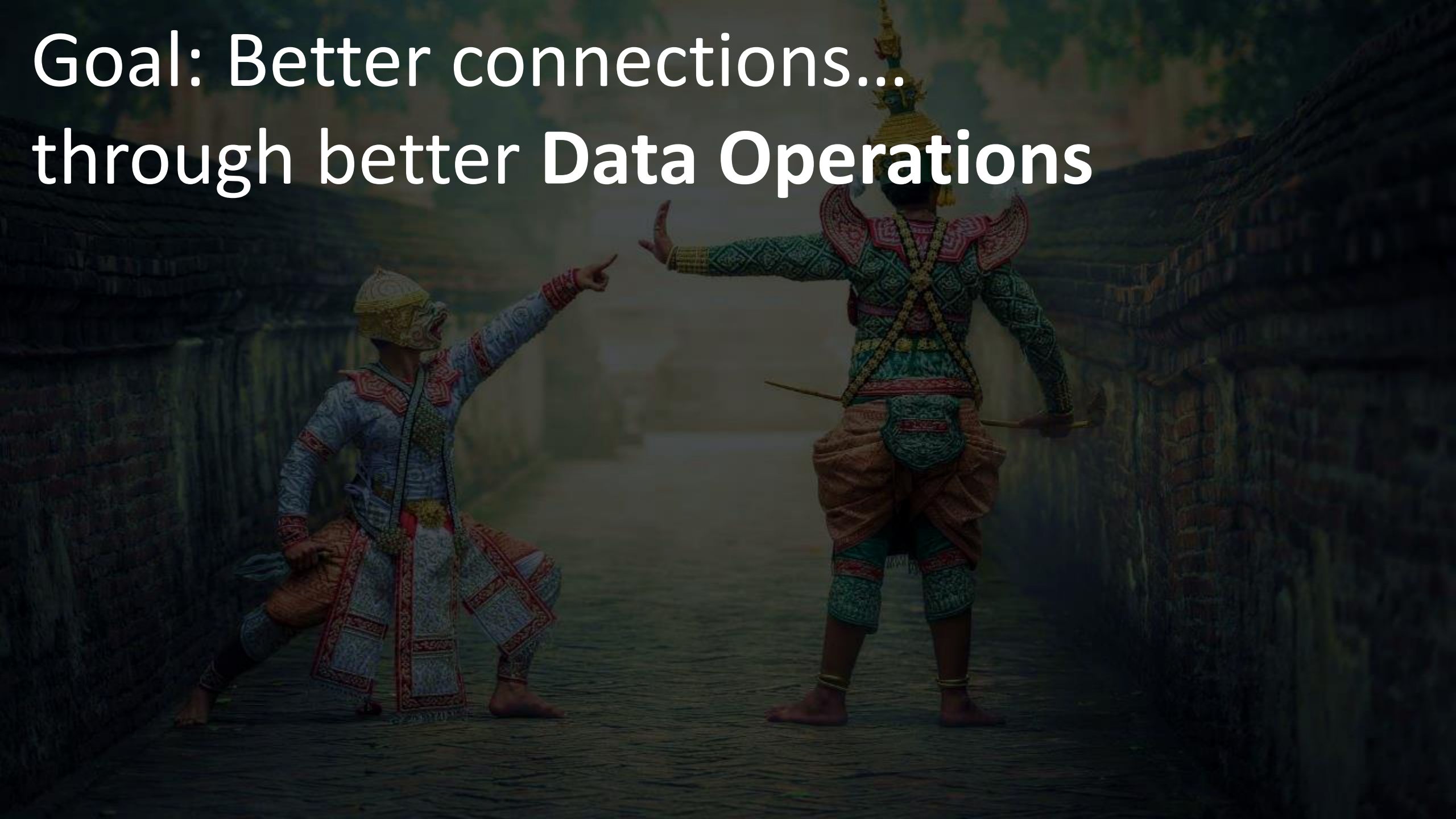
*"The way we've been able to use advanced analytics, including machine learning, to* ***model the carbon emissions from our mortgage book*** *has been a game-changer, as we think about how to help our members* ***make their homes greener****."*

Sara Bennison

Chief Proposition and Marketing Officer

*"Working with Data & Analytics allowed us to get* ***accurate understanding of our membership****, which enabled the timely and successful launch of our* ***Member Prize Draw (£1 million/month)****, returning value directly to our members."*

# Goal: Better connections…
# through better **Data Operations**



# A bit of data analysis history…

Earliest data records of periodic motions of celestial bodies

*Assyro-Babylonians, 104-101 BCE*

…c.200 BCE …

…

*© Trustees of the British Museum*

*Source: https://www3.astronomicalheritage.net/images/astronomicalheritage.net/gallery/ahp_themes/12_mesopotamia-and-the-middle-east/002.jpg*

# A bit of modern data analysis history…

…90’s 2000’s 2010’s 2020’s…

Dawn of databases
Analysis through Visicalc, Lotus, Excel

MIS and Report-based business management
BI tools emerge

Big Data technologies, accessible computing power and specialist languages and tools enables the emergence of enterprise data scientists*

*Arguably data scientists are just a recent expression of earlier professions assisted by modern tools (statisticians, economists, geophysicists, geneticists, etc.)

# Context (Financial Sector)

Software-enabled business => Data-enabled business

…90’s | 2000’s | 2010’s | 2020’s…

Dodd-Frank, Sarbannes-Oxley:
Control software => DevOps

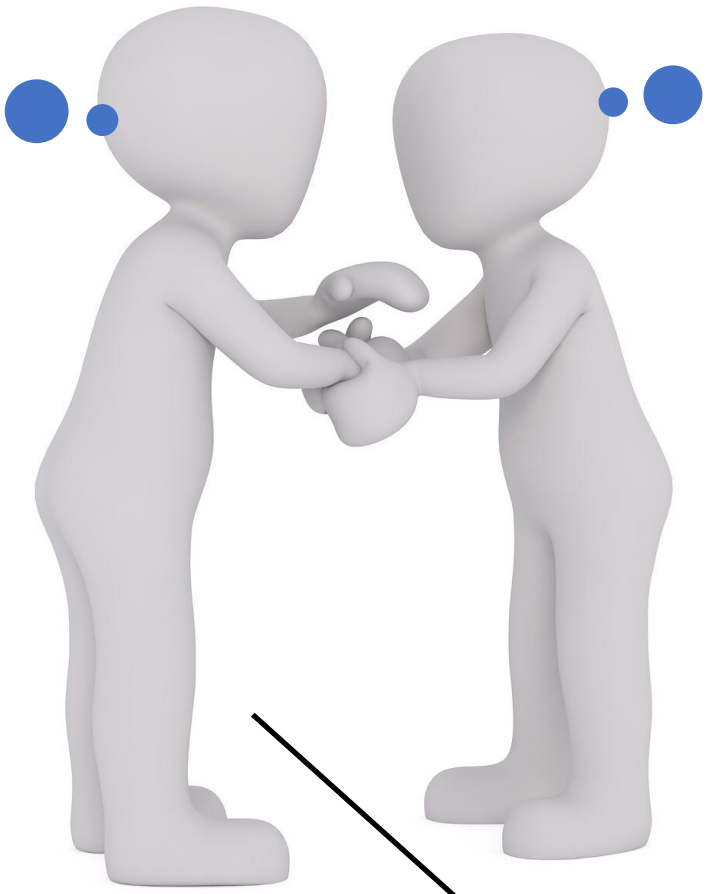
Financial Crisis and the dawn of CDO:
Govern data

Data professionals are citizen developers, not ***End Users***
Data scientists, data engineers and data managers straddle the Tech and End User labels

# Misunderstanding and mistrust?

Your primitive handling of the code lifecycle is shocking and will get us into trouble!

Why do you have to automate everything before understanding what the data insights are?



Tech camp

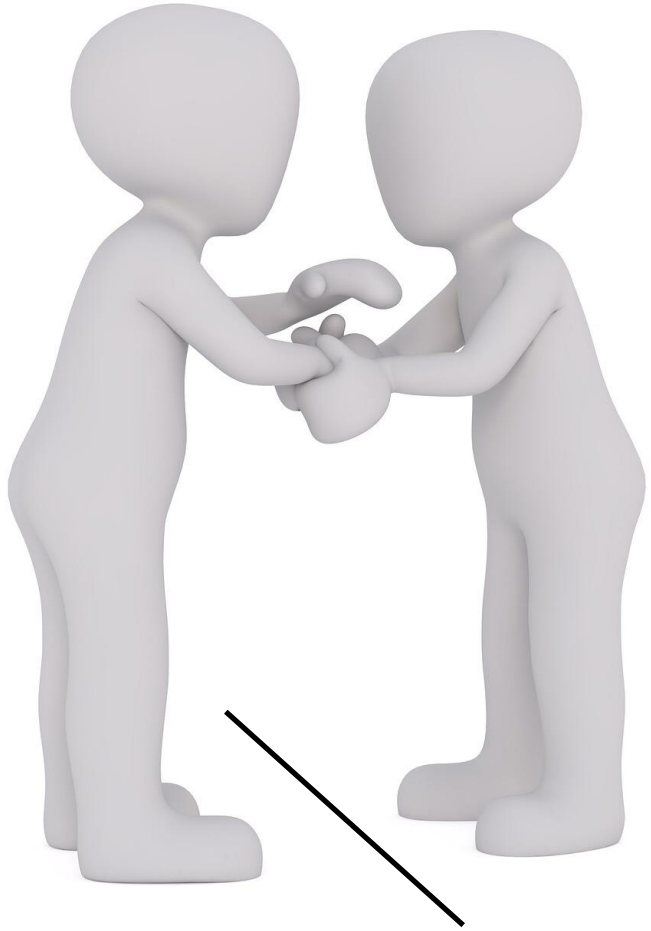
Data camp

# Common goal

To offer a financial product at real time, based on the customer’s circumstances and behaviours

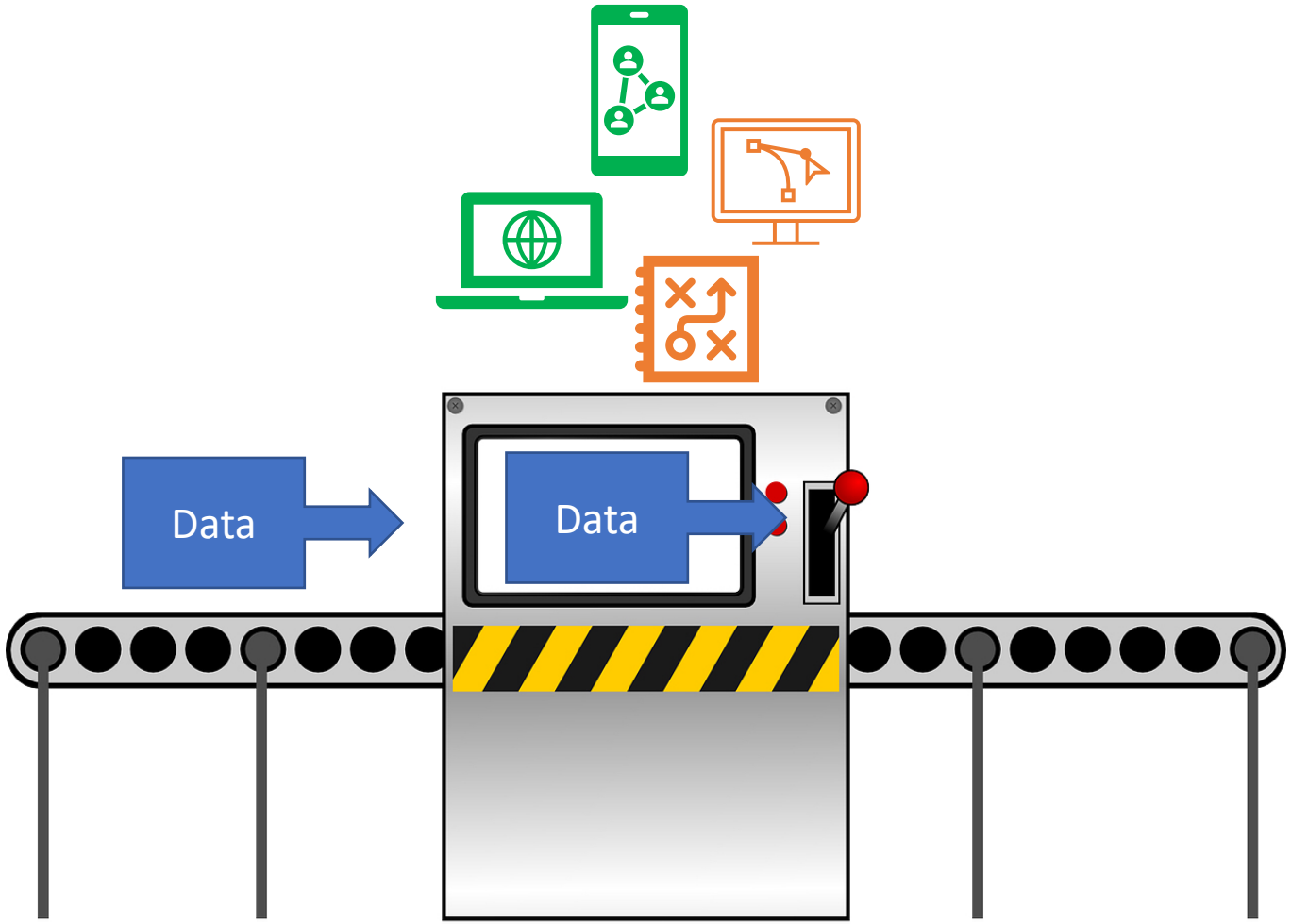
# Software and Data needs

Software products



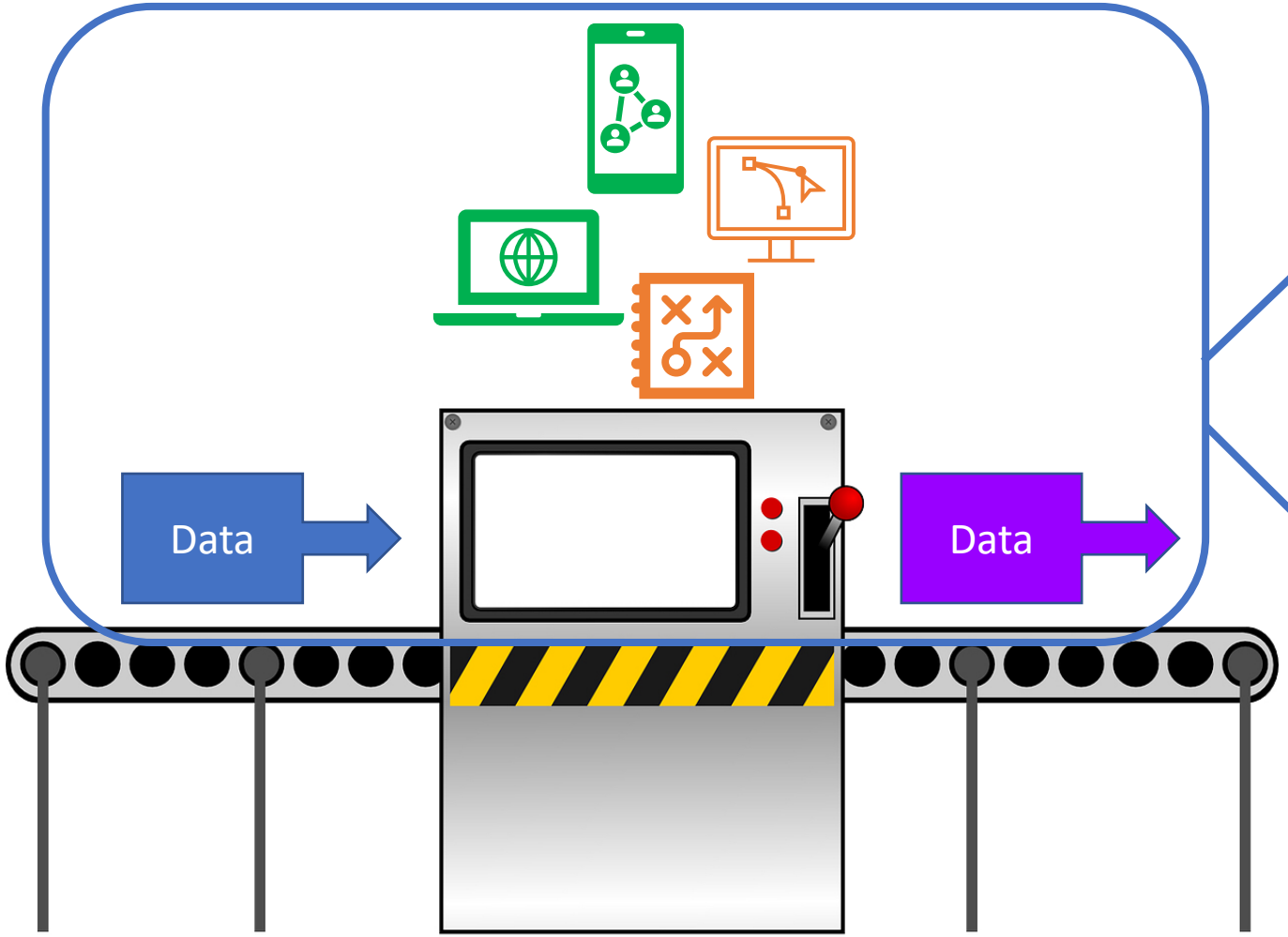
Predictive models, insights

# Business virtual production line




We need better team **interactions** and appreciation for their contribution to business results

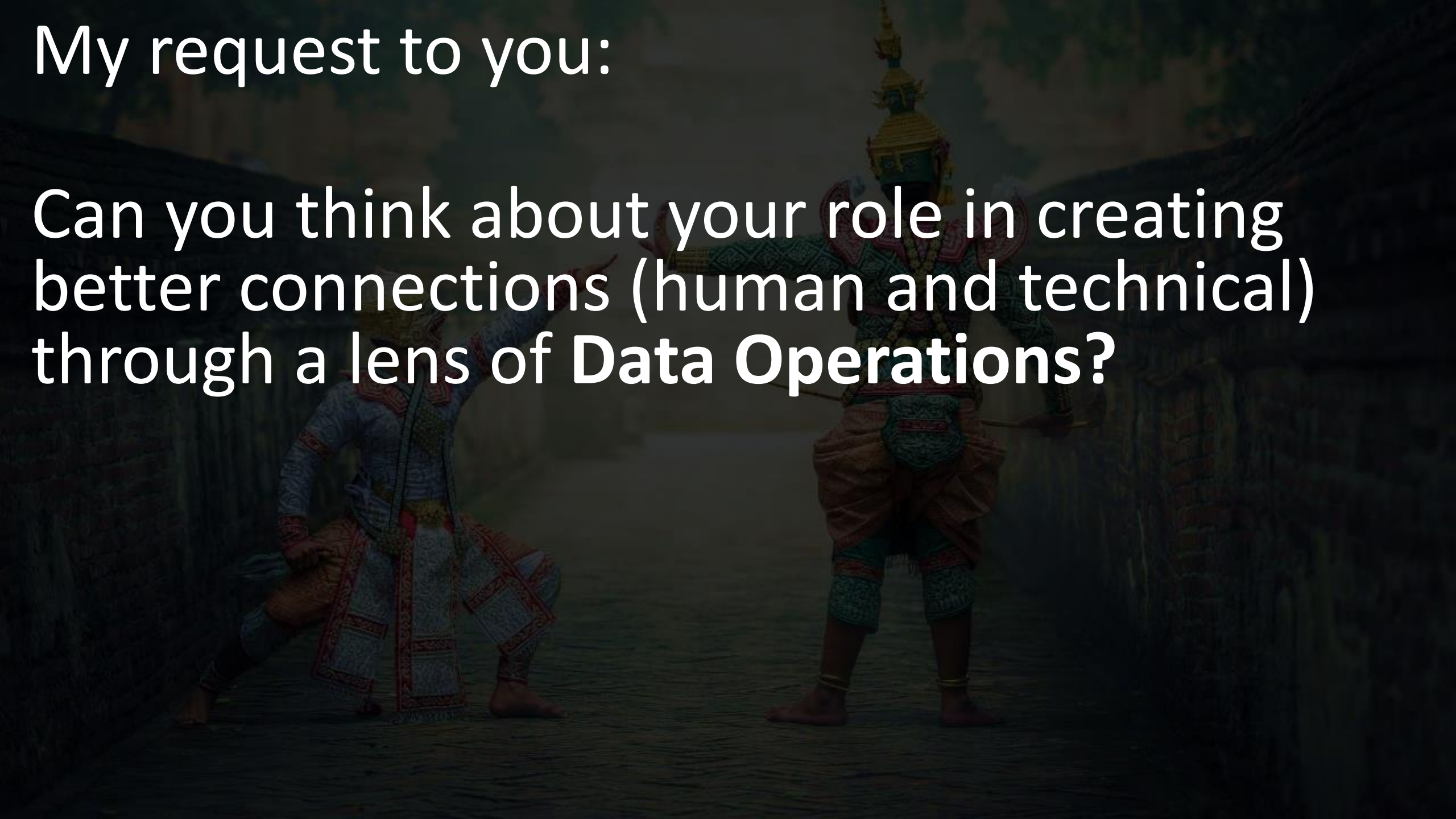
# Data Operations - DataOps




Automate and monitor the business (data) production line

Data observability is different from system observability

# Reality is a bit more complex



# My request to you:

Can you think about your role in creating better connections (human and technical) through a lens of **Data Operations?**

# DataOps – how did we get here?

HBR article defining the problem space

…2013 2015 2017 2019 2021…

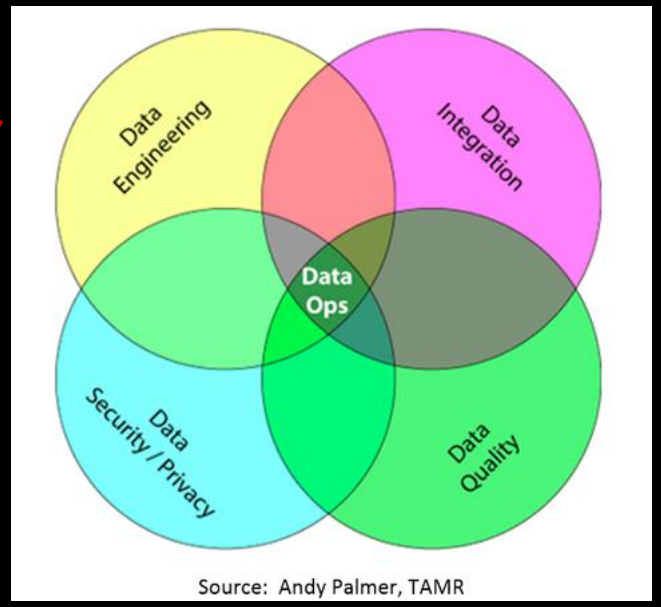
*"[…] to extract insights from the massive amounts of data now available from internal and external sources, many companies are **spending heavily on IT tools and hiring data scientists**.*

*Yet most are **struggling to achieve a worthwhile return**. That's because they treat their big **data and analytics** projects the same way they treat all **IT projects**, not realizing that the two are completely different animals."*

# DataOps

TAMR's Andy Palmer's working definition of practices

…2013 2015 2017 2019 2021…



# DataOps

Definition from a Data Science perspective
Bill Vorhies, contributing editor for Data Science Central

…2013 2015 2017 2019 2021…

*Recognising "**conflicting goals of the different data tribes** in the organization, data science, BI, line of business, operations, and IT"*

*… and proposing "basis for a new category of data access, blending, and deployment platforms that may **solve data conflicts in your organization**"*

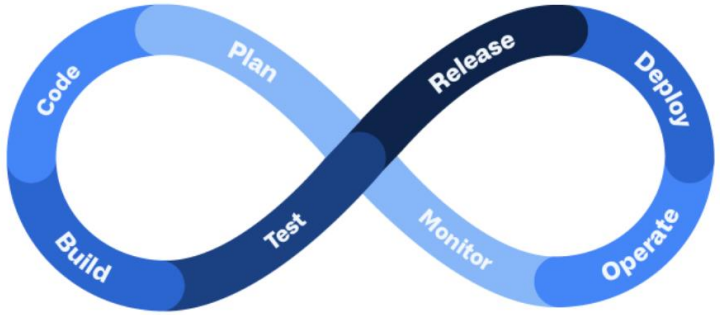
# DataOps

Definition from a Technology perspective

…2013 2015 2017 2019 2021…

"Inspired by the DevOps movement, the DataOps strategy strives to **speed the production of applications** running on big data processing frameworks.

[…] including data development, data transformation, data extraction, data quality, data governance, data access control, data center capacity planning and system operations"

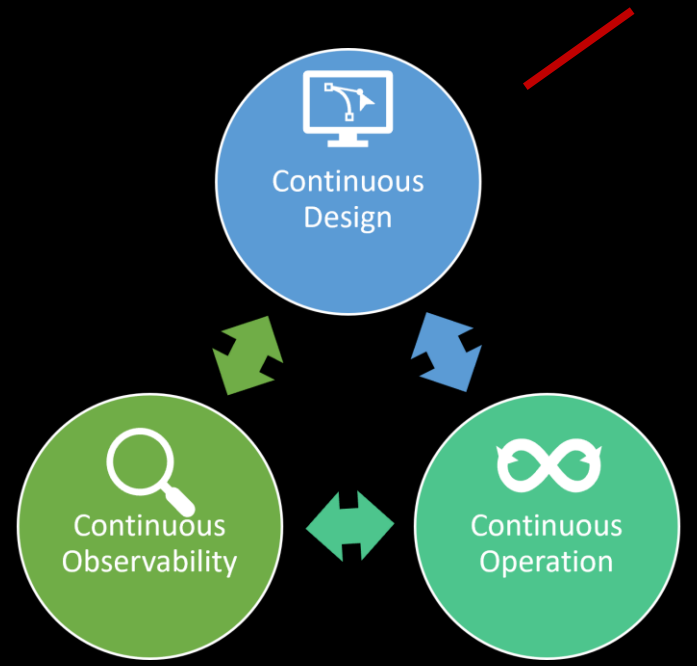



Agile approach to designing, implementing and maintaining a distributed data architecture

# DataOps

Community is maturing

…2013 2015 2017 2019 2022…



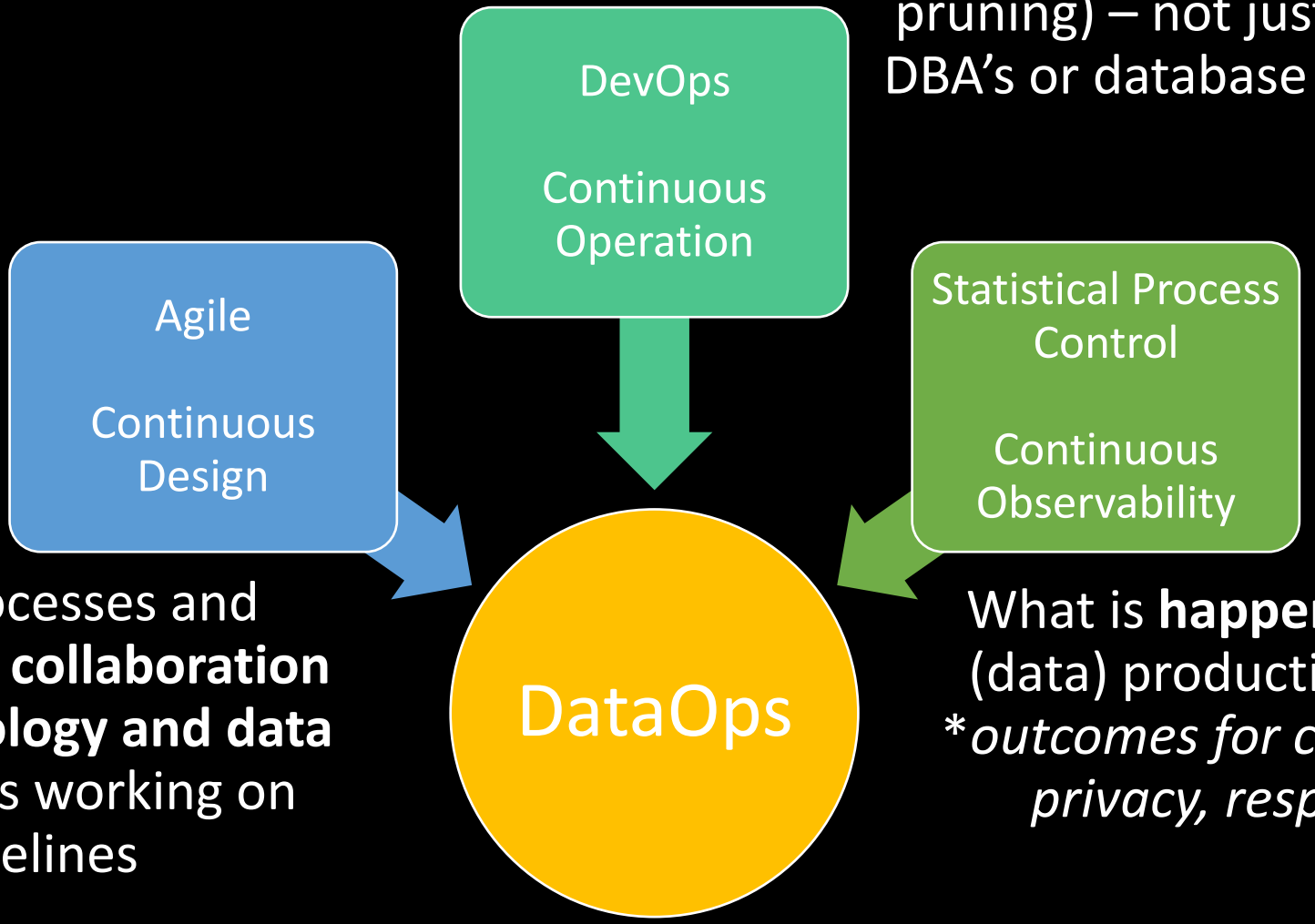
# DataOps – do we share a North Star?

## Integrity of data and timeliness of decisions, with…

- Better team interactions and technology interoperability for reduced data science applications' cycle time, acknowledging that…
- Data analytics has moved away from static MIS (heavy ETL, brute force wrangling) towards dynamic insights needing rigorous metadata management.

# The ingredients…

**Extend mature practices into large-scale data management** (wrangling, pruning) – not just your standard DBA's or database developer's job



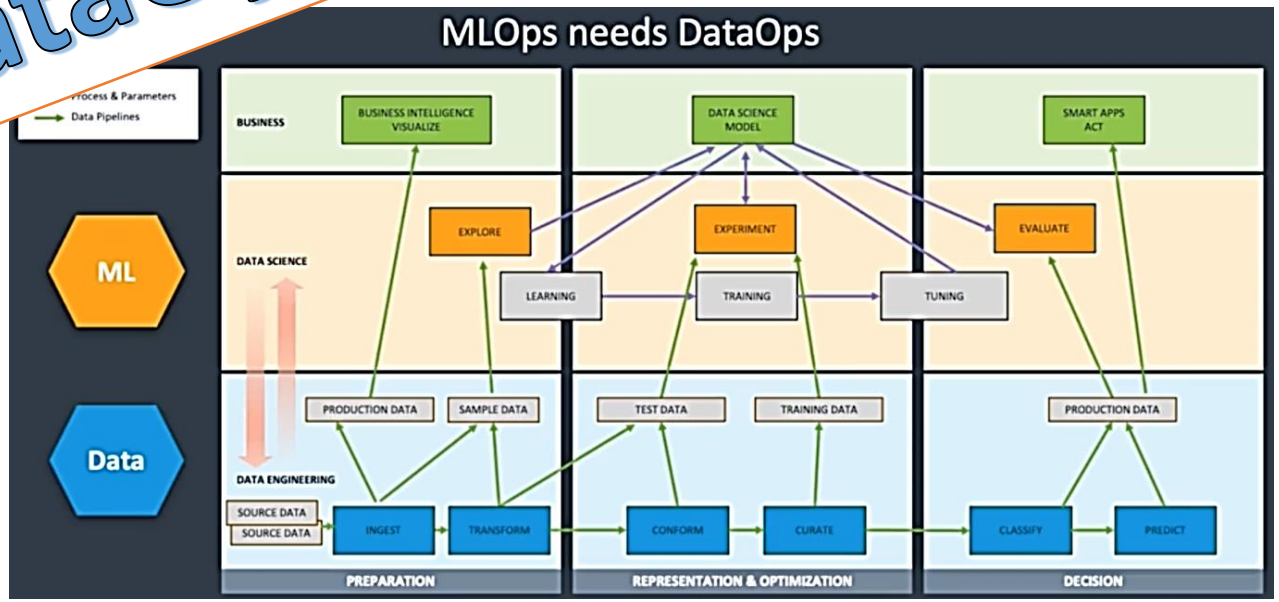

Practices, processes and technologies for **collaboration between technology and data analytics** teams working on data pipelines

What is **happening*** in the virtual (data) production line **right now**?
**outcomes for customers: accuracy, privacy, responsiveness, etc.*

# Not covered today… MLOps and AnalyticsOps

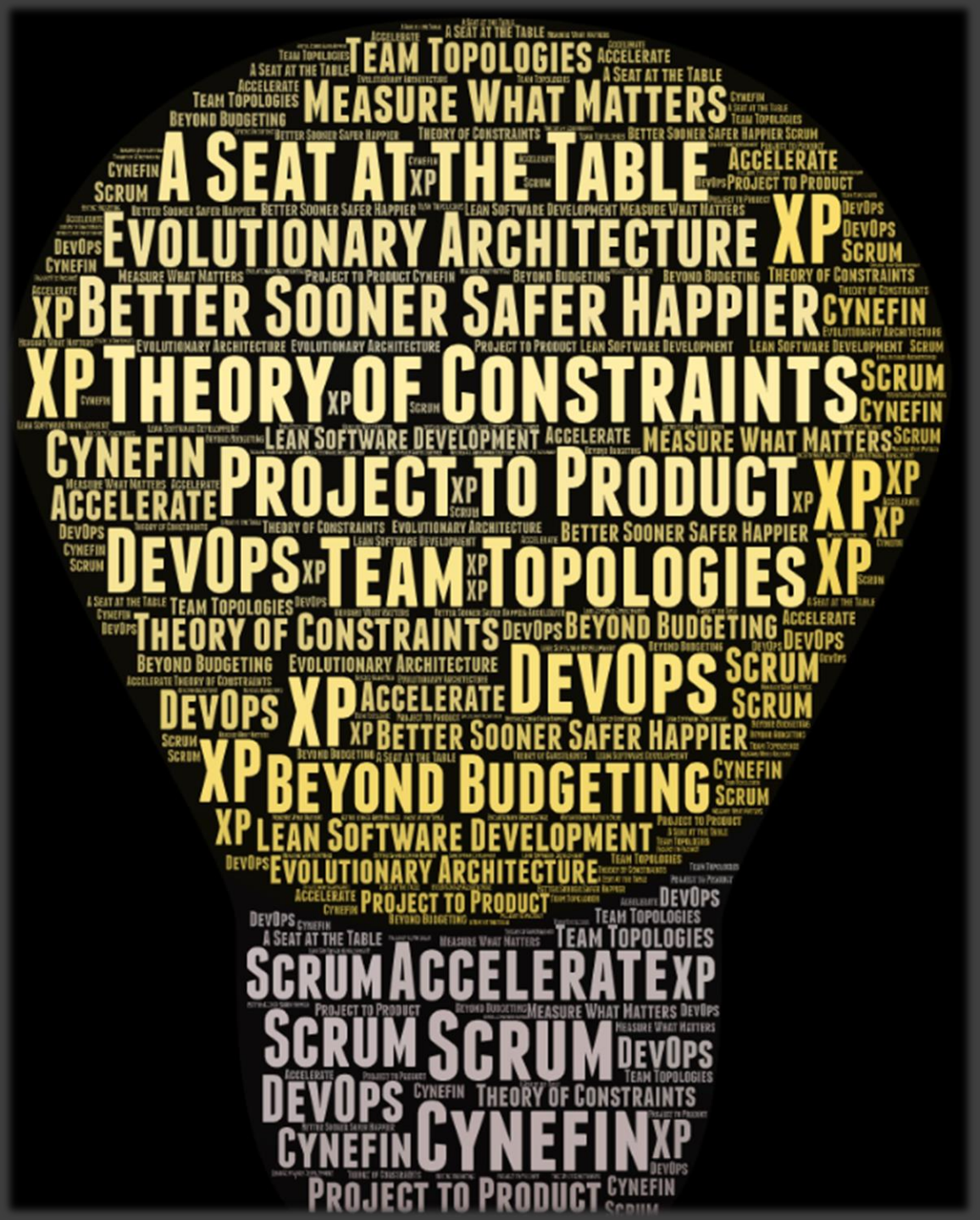
Emerging practices of automation and monitoring of deployment, delivery and performance of AI/ML models

DataOps Summit

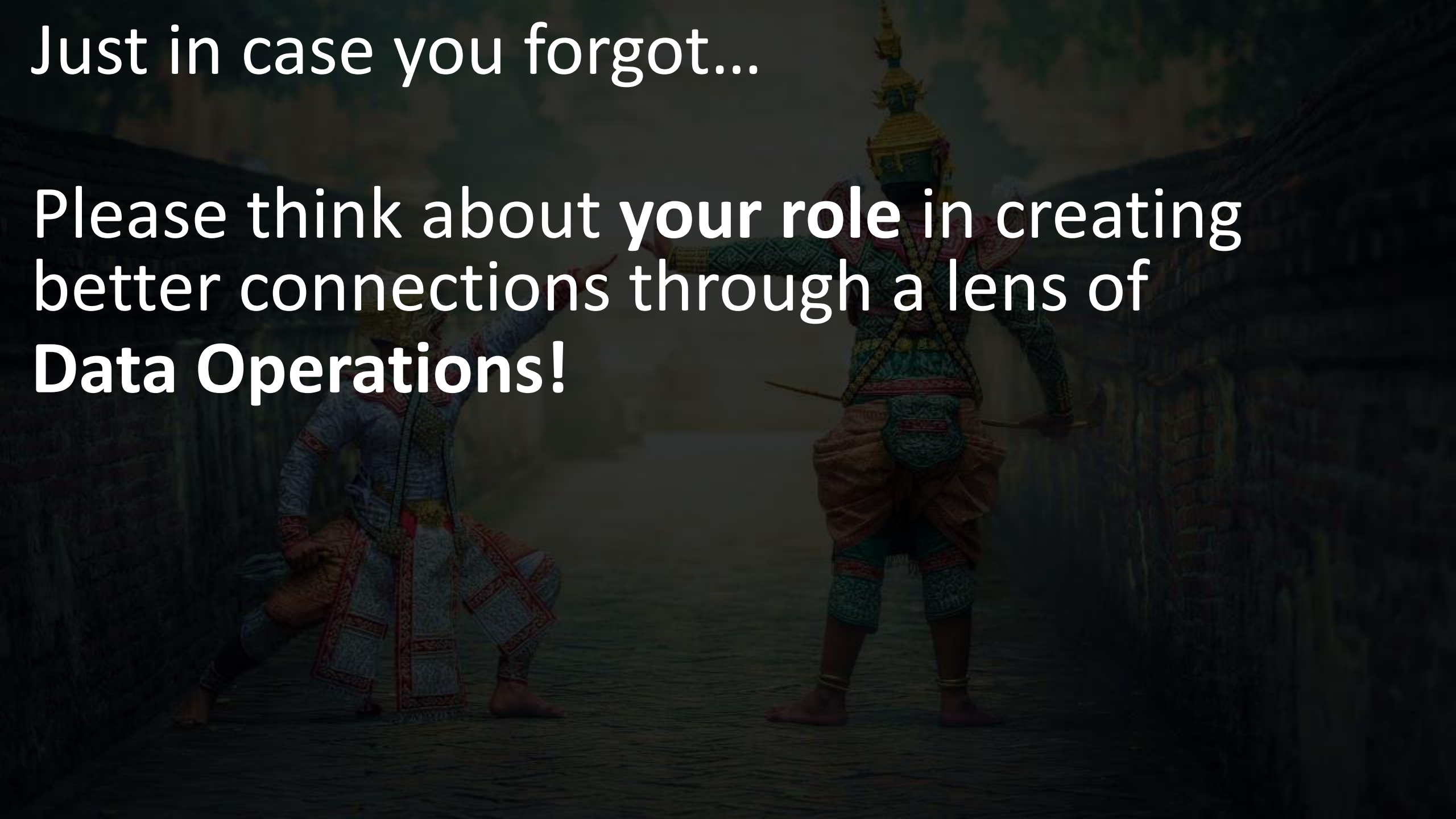


# References


- Astrometry
  https://sci.esa.int/web/gaia/-/58212-astrometry-through-the-ages
- A lunar ephemeris from Babylon calculated using ‘System B’ in cuneiform astronomical tablets. BM 34580+42690.
  © Trustees of the British Museum
  https://www3.astronomicalheritage.net/images/astronomicalheritage.net/gallery/ahp_themes/12_mesopotamia-and-the-middle-east/002.jpg
- February 2013, Harvard Business Review - Why IT Fumbles Analytics, by Donald A. Marchand and Joe Peppard
  Why IT Fumbles Analytics (hbr.org)
- May 2015, Tamr – From DevOps to DataOps, by Andy Palmer
  https://www.tamr.com/blog/from-devops-to-dataops-by-andy-palmer/
- April 2017, Data Science Central – DataOps, it’s a secret, by Bill Vorhies, contributing editor
  https://www.datasciencecentral.com/author/williamvorhies/
  https://www.datasciencecentral.com/dataops-it-s-a-secret/
- May 2019, TechTarget – DataOps definition, by Jack Vaughan, Senior News Writer
  https://www.techtarget.com/searchdatamanagement/definition/DataOps

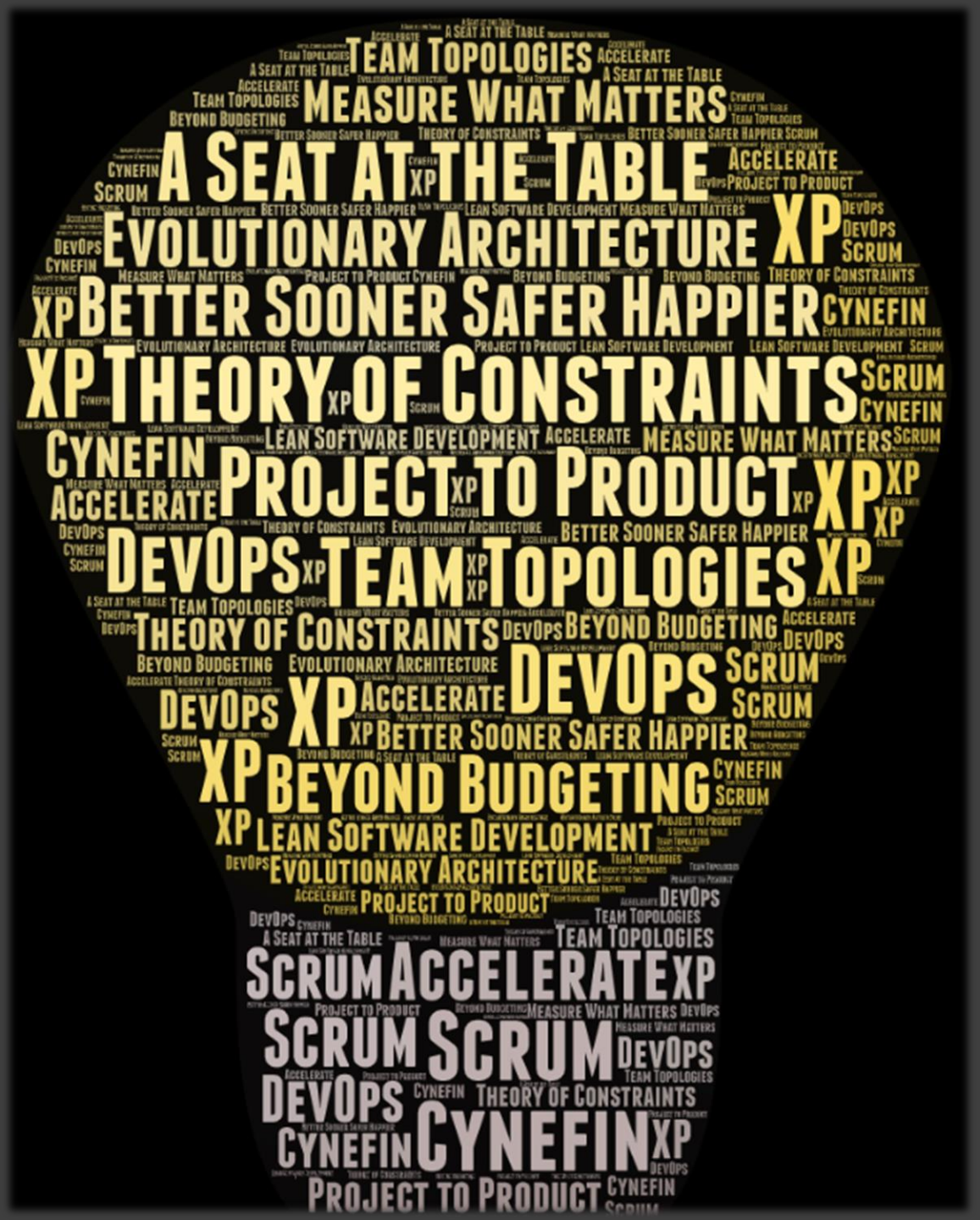



# Standing on the shoulders of giants



Just in case you forgot…

Please think about **your role** in creating better connections through a lens of **Data Operations!**


TEAM TOPOLOGIES
MEASURE WHAT MATTERS
A SEAT AT THE TABLE
EVOLUTIONARY ARCHITECTURE
XP
BETTER SOONER SAFER HAPPIER
THEORY OF CONSTRAINTS
CYNEFIN
LEAN SOFTWARE DEVELOPMENT
PROJECT TO PRODUCT
DEVOPS
BEYOND BUDGETING
SCRUM
ACCELERATE

# Thank you!