Welcome & Introductions

- **Welcome & Introductions**

- **The Data Fit Organisation (DFO) - What is it & why it matters?**

- **Applied Research Outcomes - Why & how we mapped a Data Workflow?**

- *Break (5 mins)*

- **The Data Workflow Method - How it works?**

- **Exercise: Mapping your own Data Workflow**

- **Discussion & Insights**

Zane Prickett

- Co-founder Unearthed, Co-founder & Director CORE Innovation Hub, Founding Director of StartupWA
- Leading contributor to the resources innovation community. Unearthed leads the world's largest community of innovators in the resources sector in crowdsourcing solutions to industry challenges. Founded two startups focused on the use of large geoscience and operation datasets in Oil & Gas applications. Previously worked in a variety of roles around the world with Schlumberger.
- Founder of the Data Fit Organisation concept and method.
- **zane@corehub.com.au**

Dr Eden (Keyao) Li

- Research Fellow, Future of Work Institute, Curtin University
- Research Fellow, Australian Research Council Industrial Transformation Training Centre for Transforming Maintenance through Data Science.
- Lead researcher for the Data Fit Organisation concept and method.
- **keyao.li@curtin.edu.au**

Understand concept of a Data Fit Organisation

Map Data Roles across a Data Workflow

The Data Fit Organisation (DFO)
What is it & why it matters?

- **Data science is hard.** Most data science projects (>75%) in industry fail and most AI initiatives (85%) do not deliver value to the business (Deloitte, McKinsey). Issues exist with data quality, data capability, business understanding and deployment.

- **However, predictive models enabled by ML/AI bring massive opportunity and potential.**

- **As they change the way we work, we need to rethink how roles and capabilities are organised and supported across an organisation.**

What simple approach can help us align People, Process and Technology to bring new value and continuous improvement through data? →→

→→**We call this the Data Fit Organisation.**

- **Where <u>all roles</u> in your organisation deliver value through data**.

- Every role has a data capability need and value to an organisation.

- **Data Fitness** is the capability of an organisation to use data to inform decision-making processes.

- **A Data Workflow** captures a standardised process to realise value through leveraging data.

- Across the whole organisation, each individual has an important yet different **Data Role** to play.

- A **Data Capability Framework** describes different requirements for different roles to ensure value delivery.

- **Fitness-To-Operate is only achieved when safety is delivered by all roles across an organisation; with aligned organisational capability and behaviours.**

- Similarly for data, an organisation's Data Fitness reflects the capability and behaviours of all roles to deliver value through data.

- The '3 Capitals' provides a comprehensive view of capability across people, process and technology.
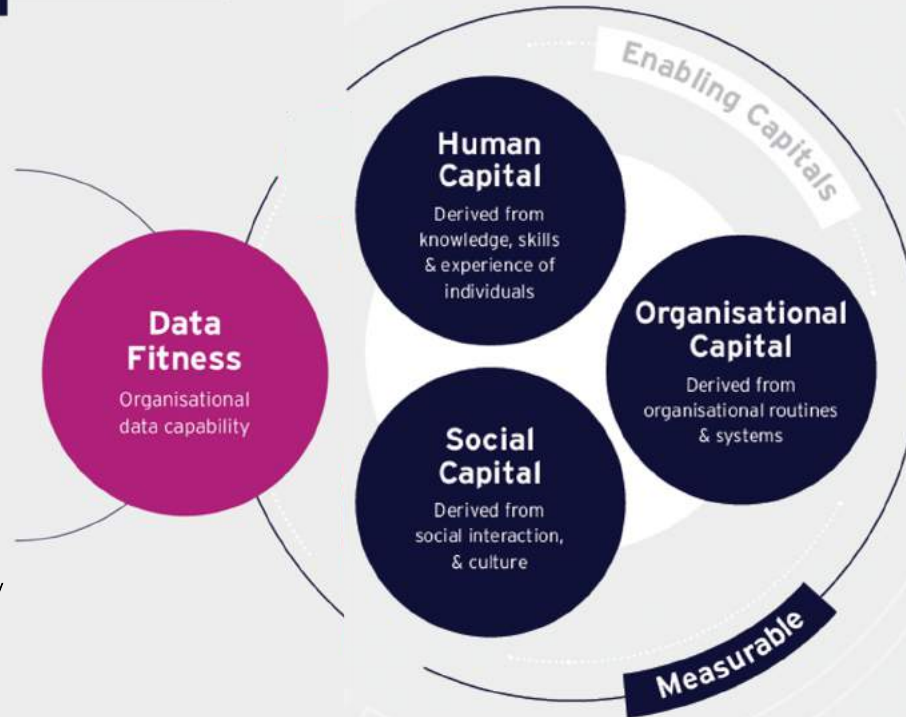
The three capitals, each with verified dimensions, have been informed by the conceptual framework developed by Griffin MA, Hodkiewicz MR, Dunster J, Kanse L, Parkes KR, Finnerty D, Cordery JL, Unsworth KL. A conceptual framework and practical guide for assessing fitness-to-operate in the offshore oil and gas industry. Accident Analysis & Prevention. 2014 Jul 1;69:156-71.
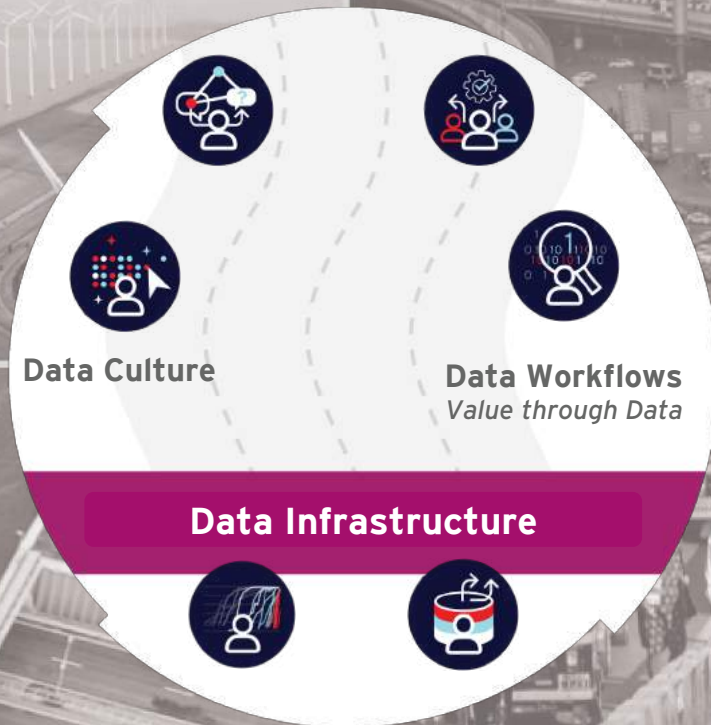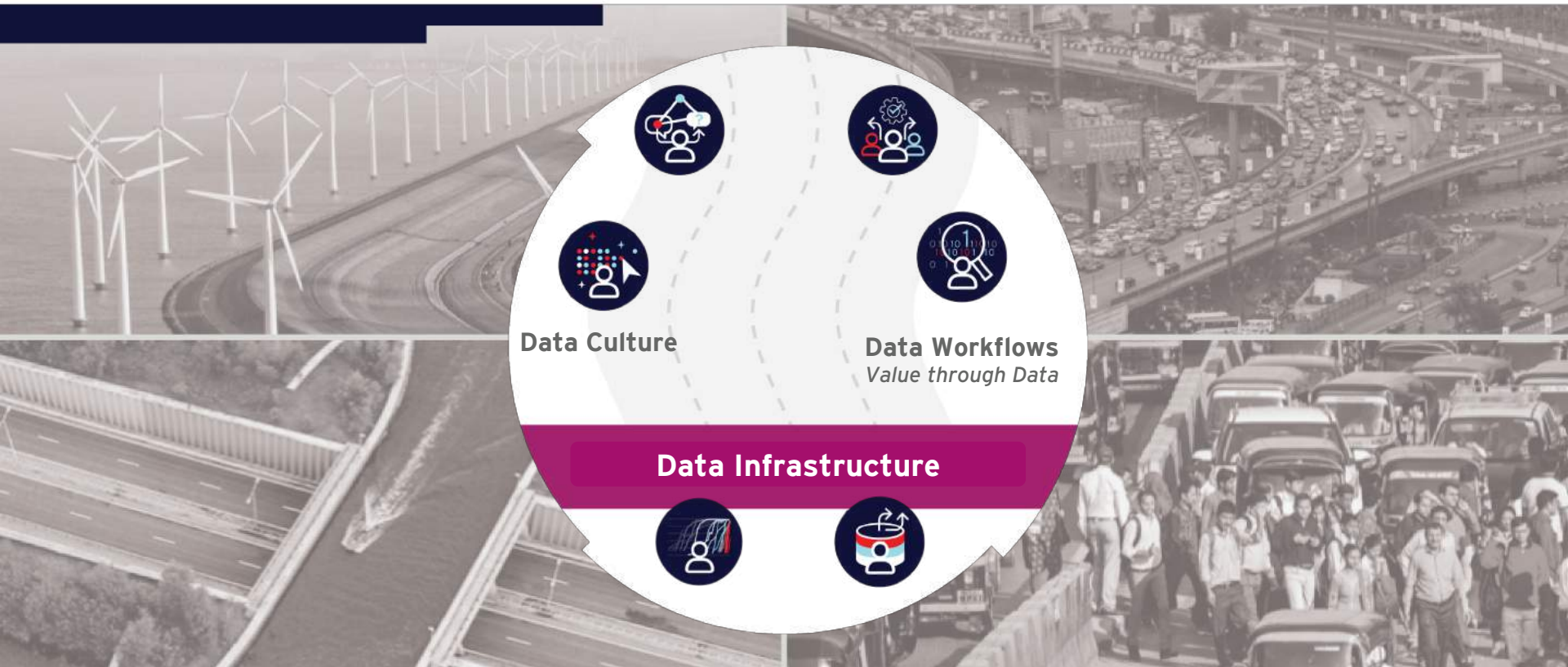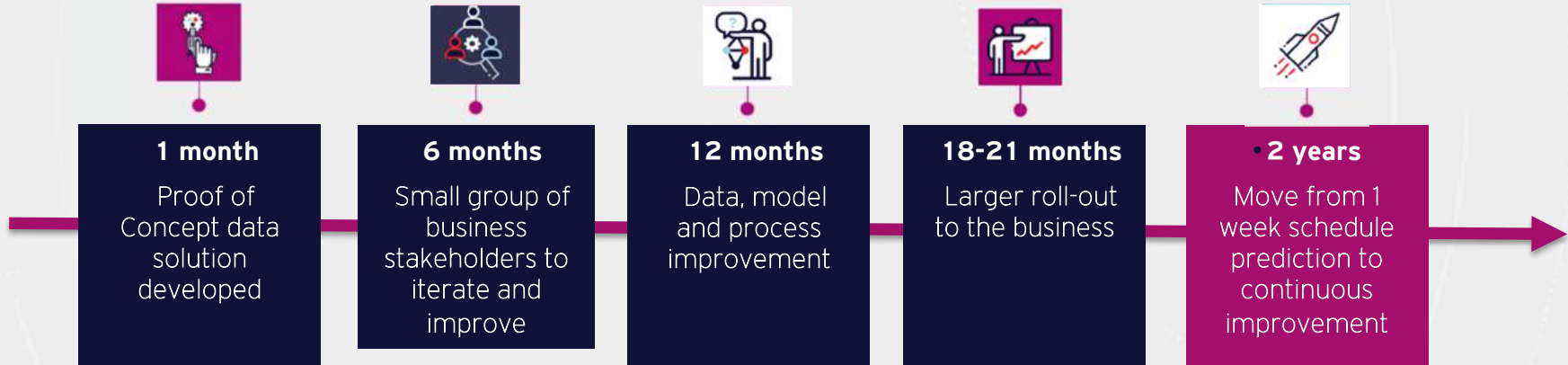
Applied Research Outcomes
Why & how we mapped a Data Workflow?

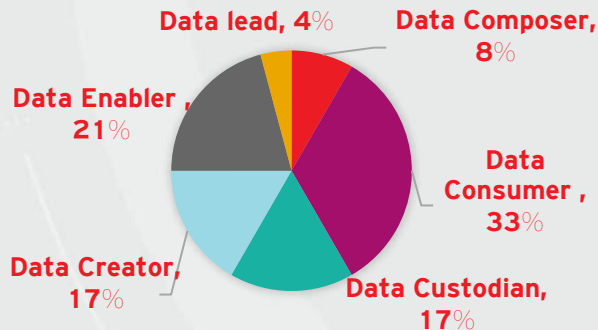**Through this research, we want to understand:**

- What are the **core features of an embedded operation-specific data workflow?**

- What are **the roles, and the associated key capabilities** that could enable efficient collaboration among the roles?

To deepen the understanding of a Data Fit Organisation (DFO), we investigated a now-embedded Data Workflow (DWF) in operation that has been use geological and production data with a machine learning model to predict 24 hours in advance the percent <1mm fines

1 month
Proof of Concept data solution developed

6 months
Small group of business stakeholders to iterate and improve

12 months
Data, model and process improvement

18-21 months
Larger roll-out to the business

2 years
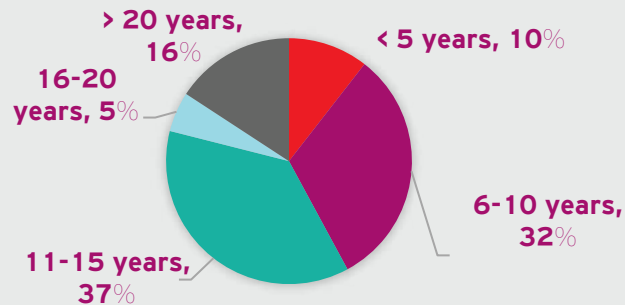Move from 1 week schedule prediction to continuous improvement

- Online survey + in-depth interviews (60mins) with 19 key stakeholders
- Survey & Interview time: October 2021
- 19 survey respondents, 17 interview participants
- Multiple workgroups: Geologist, Data Scientist, General Manager, Scheduler, Superintendent Scheduling & Execution

Data Role

- Data lead, 4%
- Data Composer, 8%
- Data Enabler, 21%
- Data Consumer, 33%
- Data Creator, 17%
- Data Custodian, 17%

Working Experience

- > 20 years, 16%
- < 5 years, 10%
- 16-20 years, 5%
- 6-10 years, 32%
- 11-15 years, 37%

Experience in Current Position

- > 20 years, 5%
- < 5 years, 95%
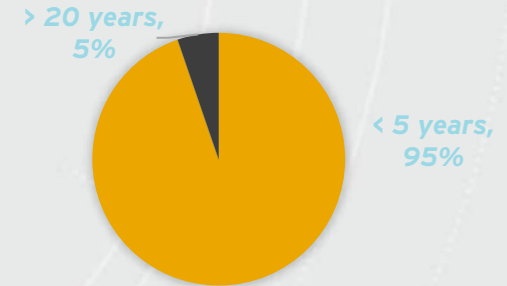
A whole-of-organisation approach to data capability, with a common language and way of working to consistently deliver value through data ('Data Workflows').

Example of quotes	Task Description (First-order themes)	Task Dimension (Second-order themes)	Stages of Workflow (Aggregated concept)
"In terms of our process plant, at the heart of it, is to separate material into sizes... the best thing we can do for the business to get the most repeatable, predictable product outcome, as well as maximize revenue is to have a very low level of variation in the percentage of minus one millimetre in what we feed plant."	Predicting the yield outcome is key to bring business benefits	Understand business priorities and potentials	Data Opportunities
"Around one year or two years ago, we found a correlation between a natural lump percentage and the percentage of minus one millimetre of the product outcome... and we ran some trials of machine learning and statistical algorithms to predict the yield outcome"	Machine learning and statistical algorithms could help predict the yield outcome	Understand the potential of data analytics	
"Understanding the potential risk of the data and limitation of the results are critically important because the decisions are made on a weekly horizon to driving the mine schedule...if that data has a level of inaccuracy, then that's going to develop potential constraints within the organization and, we must instigate particular controls and responses."	Understand the possibilities of things going wrong and the limitations of the predictive approach, so as to make corresponding contingency plans	Understand the uncertainty and risks	
"Before building a model to analyse the data, we really need to understand the story of what are the problems, challenges and pain points for operation and how this could affect business outcome, more specifically, what are the practical meanings and importance of different predictors, where and how they were collected"	Have sufficient domain knowledge to understand the practical meanings and impact of model input and output	Understand the operation problem	Data Solution
"I'm a firm believer that no data is better than wrong data".	Ensure data collected are accurate and representative	Have fit-for-use data	
"We'll use machine learning to build a predictive model to tell what's coming tomorrow and make sure this thing is running, believed and embedded and it doesn't fail"	Build predictive model to inform operation decisions	Build prediction model	

© CORE Hub 2022

Example of quotes	Task Description (First-order themes)	Task Dimension (Second-order themes)	Stages of Workflow (Aggregated concept)
"Being an engineer, I need to understand why. It is important that I understand the data that I'm using, where were they from and the purpose and the outcome of the analyses."	Understand the why, how and so what of the process	Understand the entire data pipeline	Deploy Solutions
"I feel like that goes down to this whole trust element, it is about the moral support. When some models predictions went wrong, we feed that back to the technical team and trust it will improve the model and predictions."	Trust and recognise the potential of using data analytics	Trust the model and the process	
"When we are seeing those differences (between actuals and predicted results), we update the spreadsheet to reflect and capture it. We got the crusher delivery team trained, that they could actually take that number and feed it back into system, and this will flow back to the data science team."	Provide real-world observations to build back into the model	Review and provide feedback	
"We just have to use it, and allow it to get better and in positively promoting it...the positivity is not being negative about it – it is the absence of negativity, it's the absence of 'oh this is garbage. I'm not going to use it. Don't worry about that'."	Promote the positivity/acceptance of the model across business	Perpetuate a positive data culture	Embed Solutions
"When there is a failure of the model accuracy, some stakeholders lose faith in the credibility of the data. And you have to reiterate and communicate the value in the long term, and the value if everything works according to the plan, how much that can impact the business bottom line."	Communicate the long-term benefits of using data analytics	Communicate long-term business value	
"My role for the last couple of months has been trying to educate the supervisors within the mobile plant team, on how to respond and what the triggers are and the actions they can take and who they can escalate to... the biggest part is people management and make them feel more comfortable about the new process and new standard"	Support people to adapt to changes and new ways of working and thinking	Facilitate the transition to new ways of working	

Thematic Mapping
Data Roles

	Capabilities	Behaviour requirement
Data creator Collect data from different data sources	1. I know the data sources that can support this data workflow.	· Know where to collect data. · Be able to explain the data source.
	2. That I know how to enter the data accurately in the system for this data workflow.	· Manage the data inputs to be accurate and representative.
	3. I know how to check the quality of the data in this data workflow.	· Conduct weekly and monthly data validations. · Conduct both people-driven & automated checks.
Data custodian Own, clean and maintain good quality dataset	1. I know how to check the quality of the data in this data workflow.	· Conduct weekly and monthly data validations. · Conduct both people-driven & automated checks. · Check the discrepancies between actual and predicted results.
	2. That I know how to clean the data for this data workflow.	· Identify and remove outliers that might bias the model
	3. That I understand the potential risks of managing the data for this data workflow.	· Explore alternatives and make adjustments. · Be able to predict the potential risks. · Have contingency plans.

© CORE Hub 2022

Data composer **Analyse data to generate actionable insights**	1. That I build models to generate predictions for this data workflow.	· Build the minimal viable product.
	2. That I have a problem-solving mindset for this data workflow.	· Figure out the problem, focus on the big picture and explore different ways of problem solving. · Look at data continually with a more holistic approach and ensure it performs and serves the longer-range plan.
	3. That I demonstrate the efficiency of the data analysis to a non-technical audience in this data workflow.	· Understand people with different profession backgrounds might have different perspectives. · Be able to explain data solutions with small and simple presentations.
Data consumer **Utilise data to perform their daily tasks and make decisions**	1. That I embrace emerging technologies and techniques in this data workflow.	· Open to new methods and techniques.
	2. That I use the insights from this data workflow for new ways of working.	· Use prediction results to inform new plans. · Use the insights to point to future improvement directions
	3. That I provide constructive feedback on data practices in this data workflow.	· Communicate real world situations and observations to the geology team and the data scientist team.

Data enabler **Promote the use** **of data for** **decision making** **within the team**	1. That I enable others to use the insights from this data workflow.	· Ensure the accessibility of up-to-date data to all the stakeholders. · Communicate data vision within business.
	2. That I promote collaboration on data usage across different teams in this data workflow.	· Provide training and upskilling to the champions in the team. · Organise regular meetings with different departments.
	3. That I build and sustain trust in this data workflow across the business.	· Explain the "why", the process and emphasises the potential. · Reiterate and communicate the value in the long-term, rather than focusing on the shortcomings.
Data lead **Actively seek** **potentials, align** **and adjust the** **data workflow with** **business KPIs**	1. That I focus on keeping this data workflow aligned to business KPIs.	· Adjust the data workflow to get the intended output (business KPIs).
	2. That I have a positive influence on the organisational data culture in this data workflow.	· Live a working approach in using and reporting data to support decision making. · Keep a positive attitude to the data workflow and avoid negativity.
	3. That I promote continuous improvement of this data workflow.	· Provide feedback based on frontline inspections to improve the model.

Priority Data Capabilities
Top 3 capabilities across data roles

Foundational Skills

1. Deep understanding of the business problem in the data workflow.

2. Knowledge of the data sources that can support the data workflow.

3. Understanding of the potential risks of managing the data for the data workflow.

Transformational Skills

1. Promotes continuous improvement of the data workflow.

2. Has a problem-solving mindset for the data workflow.

3. Embraces emerging technologies and techniques in the data workflow.

Networked Skills

1. Can build and sustain trust in the data workflow across the business.

2. Can provide constructive feedback on data practices in the data workflow.

3. Promotes collaboration on data usage across different stakeholders in the data workflow.

Integration Skills

1. Understanding of the entire data workflow.

2. Positively influences the organisational data culture in the data workflow.

3. Commits organisational resources under own control to data initiatives in the data workflow.

- **Networked skills were ranked as most important** by Data Lead, Data Composer, Data Consumer and Data Enabler.

- **Transformational skills were ranked as most important** by Data Creator and Data Custodian, 2nd most important by Data Consumer, Data Composer, Data Lead and Data Enabler.

Data Capability

Data Composer · Data Consumer · Data Custodian · Data Creator · Data Enabler · Data lead

Networked skills · Transformational skills · Foundational skills · Integrational skills

© CORE Hub 2022
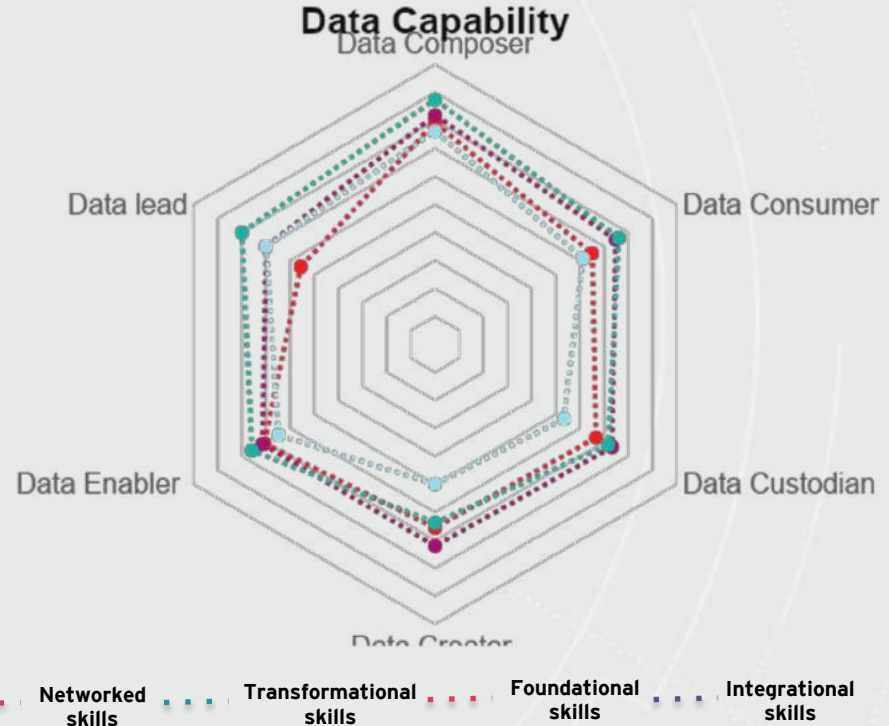
- **How is data capability understood and organised in your organisation?**

- **How do you drive value through data?**

- **Do you define data role-based capabilities?**

- **Does the Data Fit Organisation concept and purpose resonant with you and your business goals?**

Data Workflow
How it works?

Our organisations are investing in capability to deliver business outcomes through standardised processes that leverage data.

We call these data workflows, in which everyone in the organisation will have a data role.

The Data Workflow is a way of thinking and working to realise value consistently through data

- Iterative framework for organisational alignment and continuous improvement.

- Role-led, everyone in the organisation has a Data Role.

- All roles need to be defined within the data workflow and with understanding of requisite capabilities.

- Successful if embedded in the business, is a continuous source of value and is continually improved.

- As there are many and ongoing data opportunities in a business, there will be many data workflows.

- It establishes a common method and language by:

  - Having clear phases
  - Mapping roles/key tasks
  - Supporting continuous improvement
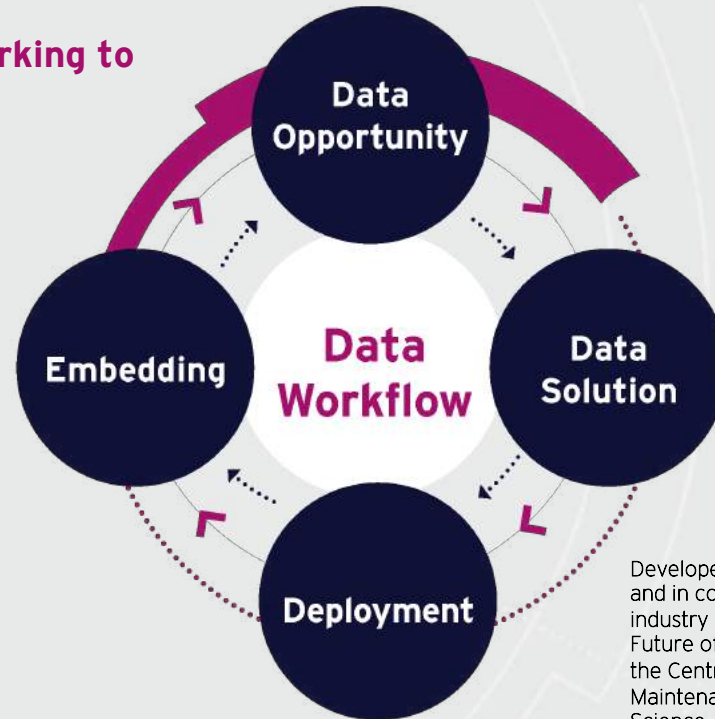
Developed from experience and in collaboration with industry consultation, the Future of Work Institute and the Centre for Transforming Maintenance Through Data Science.

© CORE Hub 2022

Data Roles

Data Lead

Actively uses data to make decisions of value to the business as a whole during role.

1. Identify the opportunities of using data.
2. Understand the potential of data usage in improving business.
3. Build data awareness.
4. Communicate and promote data usage across the business.

Data Composer

Actively works with data (i.e. transforming, analysing, producing) to inform work processes during role.

1. Build model and ensure it is running, believed and embedded.
2. Educate other roles.

It is possible to fulfill more than 1 data role at a time.

Data Enabler

Oversees and actively uses data to make decisions of value to the business within the team during role; and mobilises resources (i.e. finance, time, facilitator).

1. Communicate data findings to the business.
2. Ensure accessibility of up-to-date data.
3. Support and enable people (in the team) to implement changes in the new work process.
4. Facilitate others to use model prediction results.

Data Custodian

Owns, consumes, and generates outputs using data.

1. Ensure data input accuracy and representativeness.
2. Provide context and subject matter knowledge to data composers.
3. Interpret and reconcile model predictions.

Data Consumer

Utilises data as a primary function of their daily tasks to perform their role and make decisions.

1. Monitor and review prediction results.
2. Use data prediction results to daily work.
3. Validate and provide feedback back into the model.

Data Creator

Creates data assets during role or data use automatically creates data assets.

1. Collect data (e.g. sample data from the blast cores).
2. Ensure data input accuracy and representativeness.
3. Understand and use the data collected.

DATA OPPORTUNITY

An opportunity to drive business value through data:

- Can come from anywhere in the organisation.
- Utilise data and the potential of data (ie predictions, visualisations,...) to realise value.
- When embedded, have clear value for the business.

Examples:

- Water use reduction through tailing slurry density prediction.
- Increase core sample information availability through automated image cropping and data ingestion.
- Increase conveyor uptime through failure estimation.
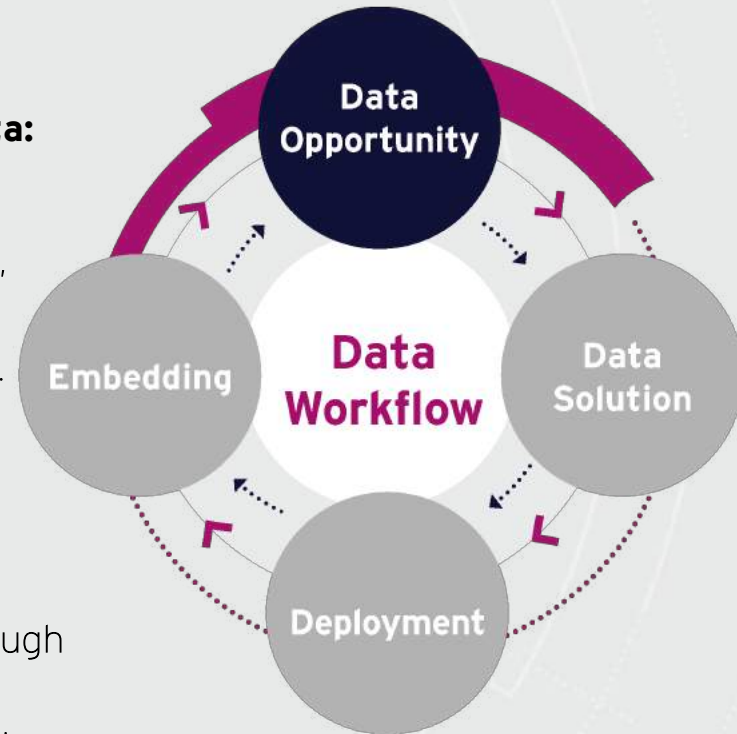
DATA SOLUTION

A solution utilising data that delivers outcomes:

• Can be simple or complex, data display to predictive machine learning model(s)

• Meets the requirements of the business need

Examples:

• An optimisation

• A statistical prediction

• A machine learning classifier

• A user interface to display information
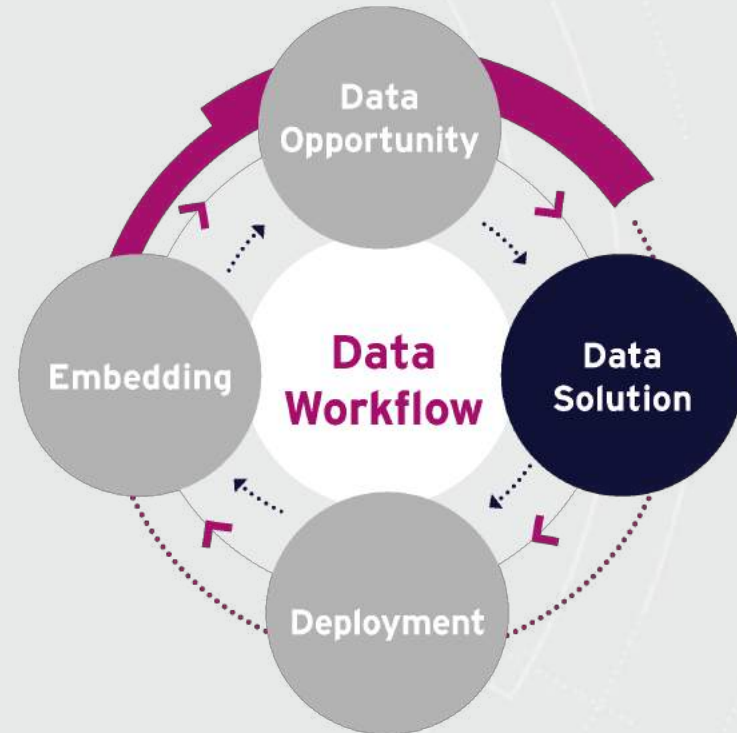
DEPLOYMENT

Ensure the data solution gets to the data consumer to embed and realise value:

- Deploys the data solution to the data consumer.
- Ensures the validation, reproducibility, and monitoring of the solution.
- Meets the requirements of the business need.

Examples:

- Dashboards, Integrations
- APIs, Third party provided
- Edge deployment
- Model monitoring/management
- Github
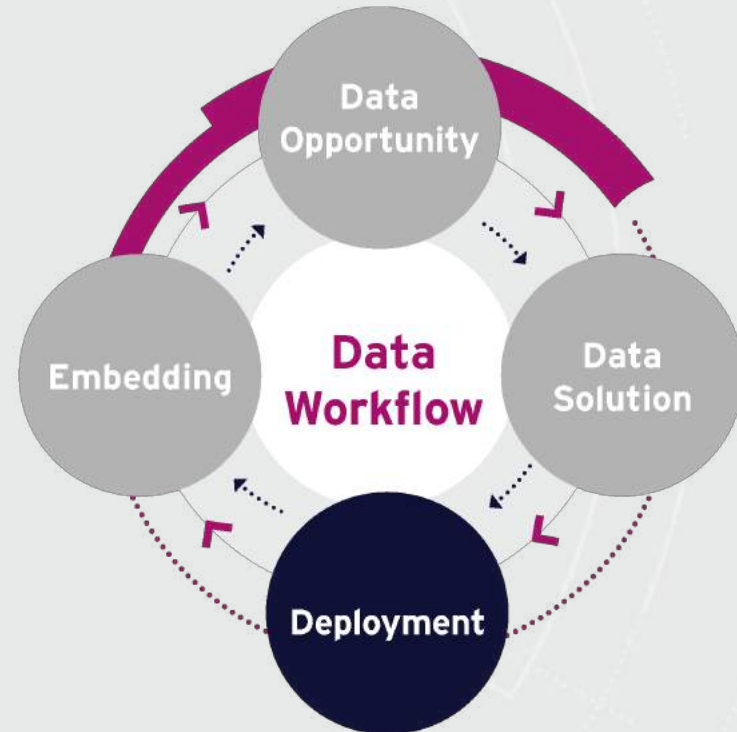
EMBEDDING

Ensure a deployed data solution is integrated into the business and final value is realised:

- Ensuring the solution is understood and trusted across the data roles.
- Building a path to continuous improvement.

Questions:

- Are the outputs understood and trusted?
- Are the outputs used and the business value realised?
- Are feedback and continuous improvements pathways established?
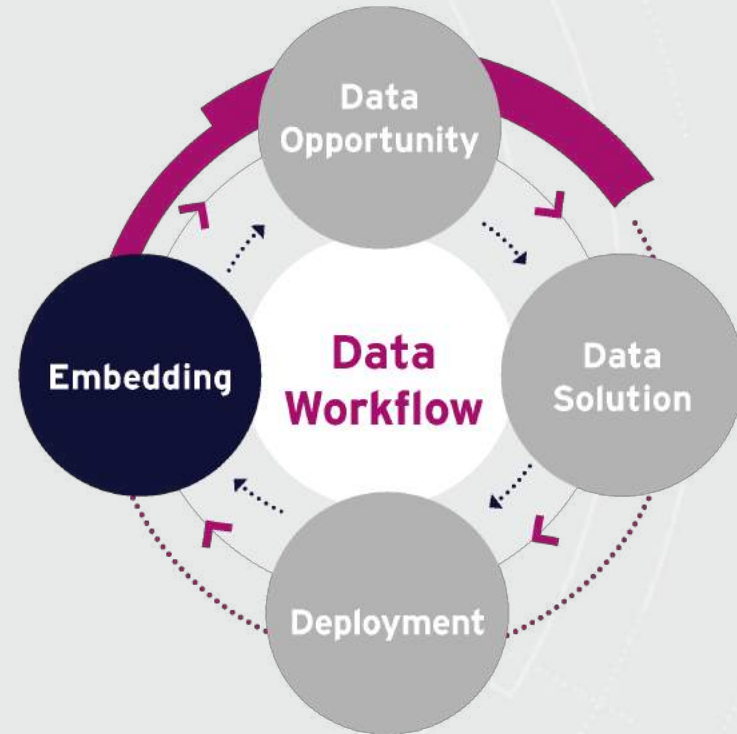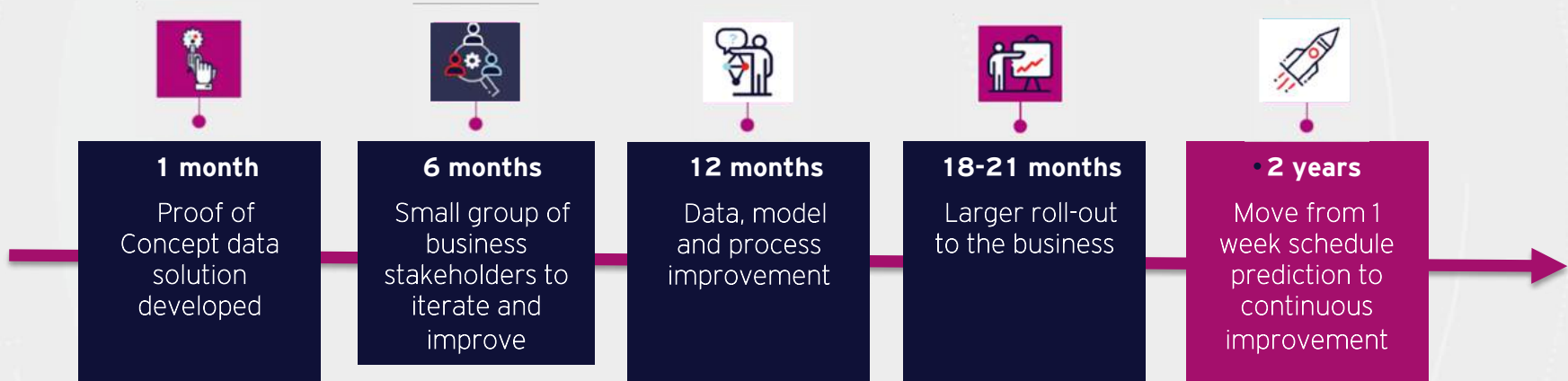
Ore fragmentation size prediction

- Ore fragmentation size is highly correlated to final recovery

- Business Opportunity: Can we accurately predict fragmentation size from data inputs to:

  - Improve recovery variability

  - Improve forward scheduling

  - Proactively manage recovery challenges

- **Value of solution was transformative → moved the operations from reactive to proactive.**
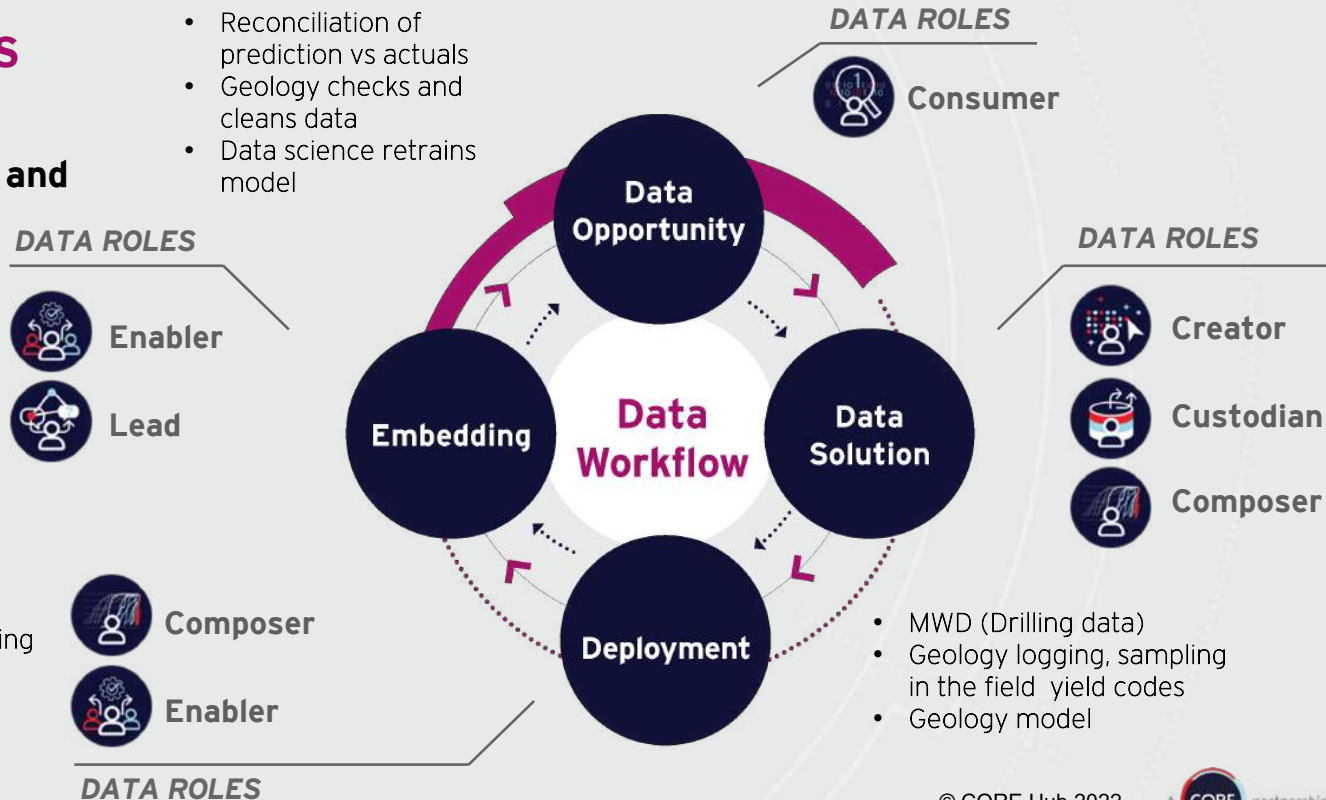
DEVELOPMENT TIMELINE

1 month	6 months	12 months	18-21 months	• 2 years
Proof of Concept data solution developed	Small group of business stakeholders to iterate and improve	Data, model and process improvement	Larger roll-out to the business	Move from 1 week schedule prediction to continuous improvement

MAPPING DATA ROLES ACROSS A DWF

All phases have data roles and capability to deliver an effective data workflow.

On-premise:
- CSV
- AMPLA
- Weekly, daily scheduling

Cloud Processing:
- Machine Learning model
- Prediction

- Reconciliation of prediction vs actuals
- Geology checks and cleans data
- Data science retrains model

DATA ROLES

Enabler

Lead

Composer

Enabler

DATA ROLES

DATA ROLES

Consumer

DATA ROLES

Creator

Custodian

Composer

- MWD (Drilling data)
- Geology logging, sampling in the field yield codes
- Geology model

Data Opportunity

Data Workflow

Embedding

Data Solution

Deployment

PEOPLE, PROCESS, TECHNOLOGY

- **It takes all Data Roles to embed a successful Data Workflow.**
- **In most Data Science solutions, the technology is easy.**
- **Let the business value drive the Data Workflow requirements.**
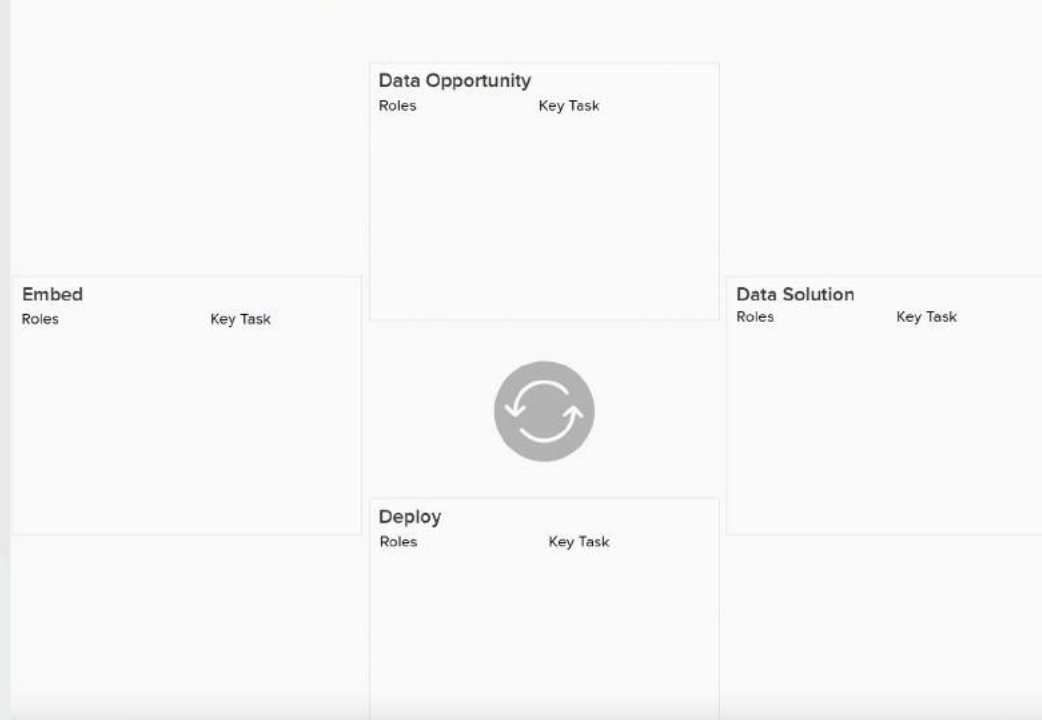- **Most first-time models don't realise value → continuous improvement must be baked in.**

- **Drive value & sustainability:** A method to help embed a predictive solution. Shifting the way we work, from a project to a continuous improvement data workflow.

- **Improve role clarity, engagement & impact:** Increase understanding of role expectations and responsibilities; enhance engagement and strengthen feedback loops; more readily identify support needs.

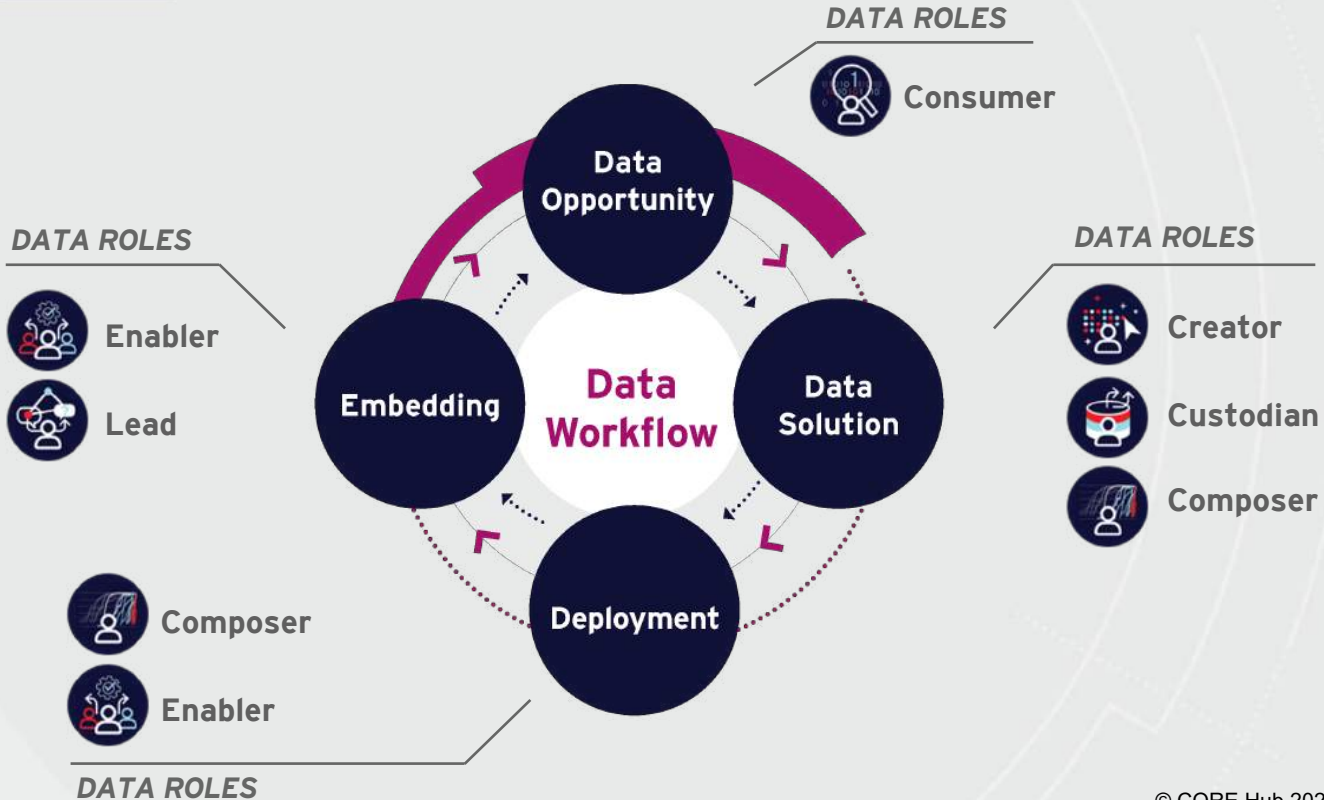- **Shared strategic approach:** A shared approach to data workflows; common language and mindset across the organisation and vendors.

Exercise: Mapping your own Data Workflow

Data Workflow on a page

A single sheet to map the opportunity, key task and data roles to realize value through data

Data Opportunity
Roles   Key Task

Embed
Roles   Key Task

Data Solution
Roles   Key Task

Deploy
Roles   Key Task

Discussion & Insights

Research & Industry Partners

- **The Future of Work Institute (FOWI) at Curtin University is CORE Skills' research partner for the Data Fit Organisation concept and method**, providing the research design and analysis.

- **CORE is part of the Centre for Transforming Maintenance Through Data Science (CTMDS), a multi-research and industry collaboration**. The Data Fit Organisation research and capability framework development contributes to the Centre's work and focus areas with industry partners.

- **CORE Skills, CORE Innovation Hub's industry skills initiative** is committed to delivering a unique model for industry, education and research expertise to accelerate professional learning and business outcomes through accessible state-of-the-art knowledge and insights.