

# THE JOURNEY FROM DATA FABRIC TO DATAOPS TO DATA PRODUCTS

*"Data has no value unless the business trusts and uses it"*

FORRESTER

- 1 **Data Productization** creates trusted data that speaks human
- 2 **DataOps** is how you get there
- 3 **Breaking Down** monolithic enterprise architecture thinking is where you start

## DROWNING IN DATA, STARVING FOR CONTEXT

It is not the data alone that holds the keys to value, but our ability to understand and operationalize the data. The key to creating value from data lies in data context and interpretability by data consumers in business operations.

*"Data needs to become self-explanatory to data consumers without needing subject matter expert support. A company is digitally mature when they can enable citizen data scientist and citizen developers to do more with advanced data and analytics."*

FORRESTER

**FROM** *Do you speak data?* → **TO** *Does your data speak human?*

For data to be operationally used at scale, especially for critical operations, it needs to be trusted. For data to be trusted, it needs to be productized.

| Common Data Product Challenges                      | Cost to Business                       |
|---|--|
| No easy way to find which data assets are available | Poor data utilization                  |
| No ownership model for existing data assets         | Slow innovation                        |
| No documentation of existing data                   | Delayed and poor decision making       |
| Many current data sets are not directly available   | Loss of business confidence towards IT |
| Difficult to compile data across                    | Information management risks           |
| Data is exported to external parties ad-hoc         |  |

## THINK BEYOND DATA LAKES

*"DataOps is the ability to enable solutions, develop data products, and activate data for business value across all technology tiers from infrastructure to experience"*

FORRESTER

Adopting a data product-centric mindset, along with a DataOps practice to create and manage data products, is needed. To successfully implement Industrial DataOps, it is essential to move from a conventional centralized data architecture into a domain data architecture (also referred to as data mesh).

For domain data architecture to work, the data product owner teams need to ensure their data is discoverable, trustworthy, selfdescribing, interoperable, secure, and governed by global access control. In other words, they need to manage their data products as a service, not as data.

## USE DATAOPS TO DELIVER CONTEXTUALIZED DATA TO BOTH SMEs AND PROFESSIONAL DATA SCIENTISTS

Industrial DataOps is about breaking down silos and optimizing the broad availability and usability of industrial data generated in asset-heavy industries. Automating the data process is the only way to make sure that live data triumphs over static documentation and reports in the decision-making process.

**FROM** *How does this solution retain and control data?* → **TO** *How does this solution enable people to access and use data?*

Here is a convenient guide to understanding the key characteristics of horizontal DataOps, as well as which features to look out for that will significantly catalyze success with DataOps adoption for heavy-asset organisations confronted with a somewhat different data source, data type, data quality as well as data consumer landscape.

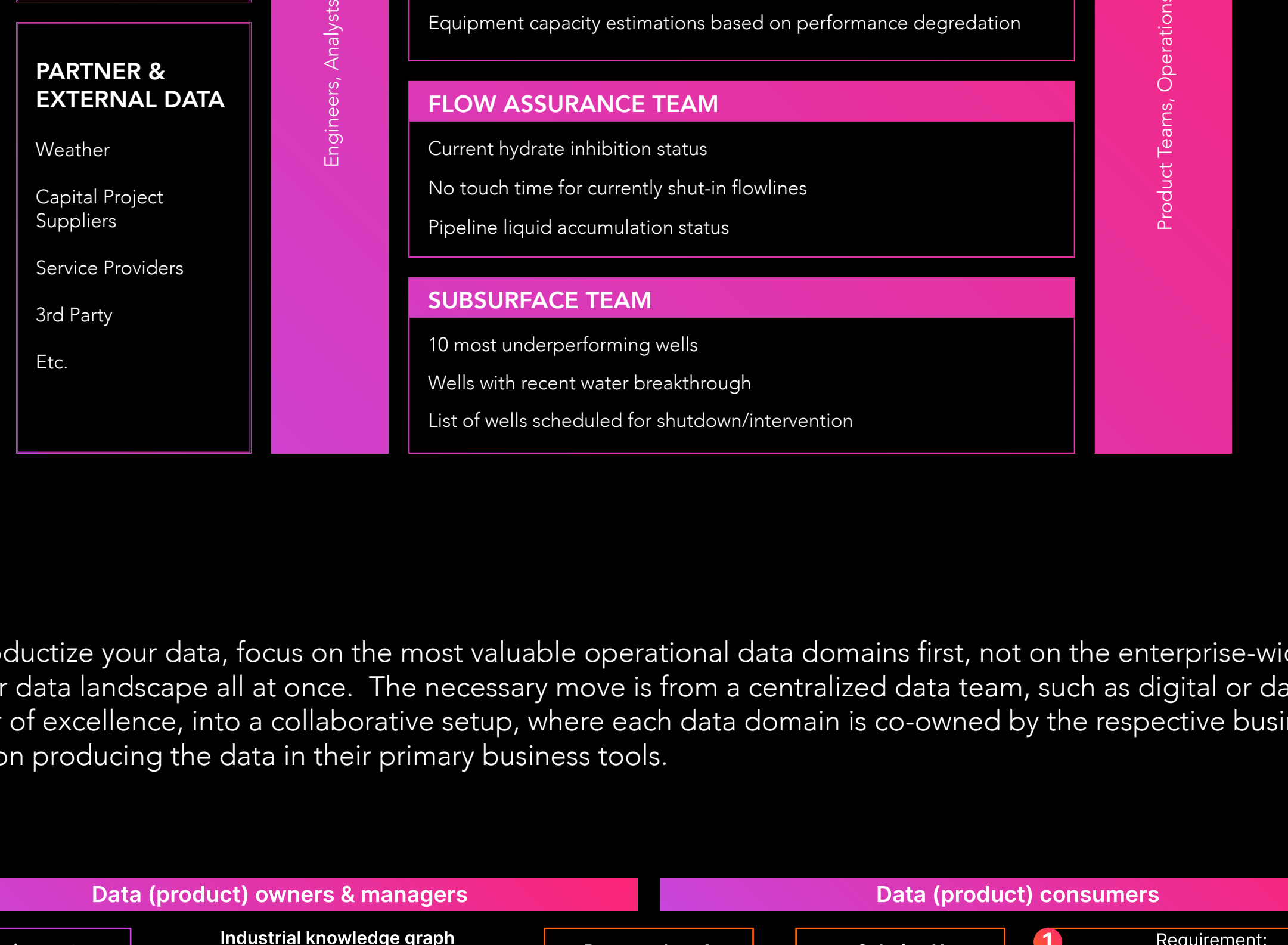
| HORIZONTAL DATAOPS CAPABILITIES   |  | SPECIALIZED INDUSTRIAL DATAOPS CAPABILITIES                                      |  |
|---|--|--|--|
| <b>Development environments support</b><br>PowerBI, data, ODBC, SPKs, Spark, Jupyter plugins, Low-code frameworks | <b>Access control and sharing</b><br>The possibility to restrict and provide access to data sets and data kits | <b>Data contextualization</b><br>Transforming abstract/OT/ET data into knowledge | <b>Data type support</b><br>Native handling of time-series data, unstructured and semi-structured data (e.g. PIDs)                       |
| <b>Versioning</b><br>Code, models, data and pipelines need to be version controlled                               | <b>Pipeline orchestration</b><br>The ability to build pipelines where one can reuse components                 |  | <b>Live data access</b><br>The ability to work with live operational (OT) data   |
| <b>Data lineage</b><br>The ability to track dependency of data  | <b>Observability</b><br>The ability to inspect, monitor and debug solutions in production                      |  | <b>Industrial equipment and process data models and templates</b><br>Talking domain language. Scaling models from one to many            |
|   |  |  | <b>Model governance and time series data quality monitoring</b><br>Ensuring data completeness and use case specific quality requirements |
|   |  |  | <b>Data discovery</b><br>The ability for SMEs to self-service explore data in full operational systems context                           |
|   |  |  | <b>Incorporating physics</b><br>The possibility to use physical simulations as part of model pipelines                                   |

DataOps enables efficient development, operationalization and scaling of digital solutions through supporting both data owners and consumers with a common toolset.

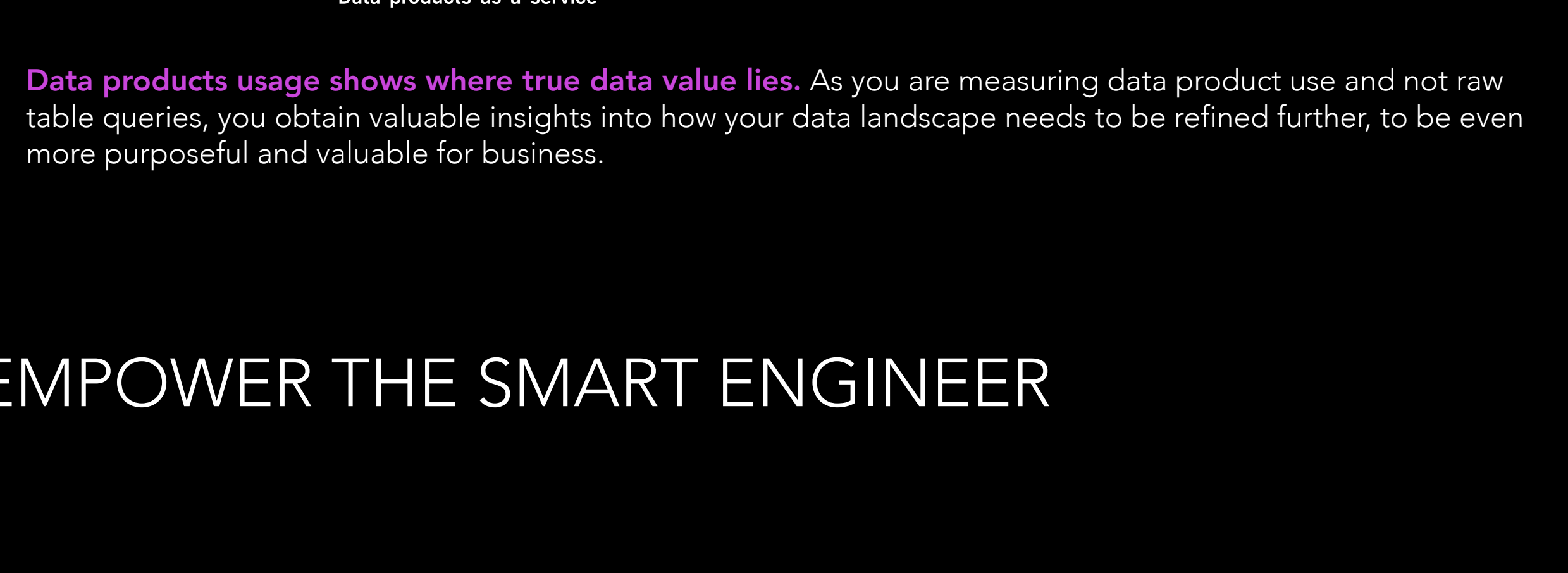
## THINK IN DATA PRODUCTS, EXECUTE IN DATA DOMAINS

The shift from data availability to data products as a service is what will allow us to transform our data swamps into operational data architectures of real business value.

### DATA PRODUCTS-AS-A-SERVICE MODEL



To productize your data, focus on the most valuable operational data domains first, not on the enterprise-wide master data landscape all at once. The necessary move is from a centralized data team, such as digital or data center of excellence, into a collaborative setup, where each data domain is co-owned by the respective business function producing the data in their primary business tools.



Data products usage shows where true data value lies. As you are measuring data product use and not raw table queries, you obtain valuable insights into how your data landscape needs to be refined further, to be even more purposeful and valuable for business.

## EMPOWER THE SMART ENGINEER

As industrial organizations will increasingly rely on the work of smart engineers, individuals and interactions (far more than processes and tools) are essential to make data valuable and useful for data consumers across an organization.

**Not traditional data professionals, but subject matter experts empowered with the right capabilities and practices to harness data effectively:**

RELIANCE ON SCARCE, HIGH-COST DATA SCIENTISTS AND OTHER DATA PROFESSIONALS

DEMOCRATIZING OF DATA UNDERESTIMATING FOR IMPROVED USE CASE SOLVING