Maximising the effectiveness of data professionals

Much of our consulting is focused on implementing platforms and tools that provide data teams (data analysts, data engineers and data scientists) the ability to uplift their productivity and solve more complex business problems or create insights into the operations of their organisation that reveal new strategic opportunities.

Technology enablement, however, will only take an organisation so far in the quest to improve the way data is managed and data 'products' are created, and to coordinate these improvements with the goals of the business. This is where methodology or 'ways of working' comes to the fore.

It is quite common for data teams to be comfortable with and follow some form of agile methodology. The way the team deals with changing requirements, the way they work together with businesspeople, and the way motivated individuals in the team are supported to deliver their best work are challenges for which the Agile Manifesto provides useful and relevant guidance. However, the agile principles of "deliver working software frequently" and "working software is the primary measure of progress" are not so relevant for data teams as they are for agile software development teams.

Another software development approach that is increasingly referenced by data teams is DevOps. DevOps combines methods followed to develop software with methods followed to put the software into production. The objective here is to slash the time taken to get a new feature into production by automating the testing and deployment and developing in such a way that facilities continuous integration of changes into the delivery pipeline. DevOps concepts are generally applicable to teams delivering data products, but there are gaps around managing the data resources, the more critical role of the data consumers in quality testing, and the monitoring of ongoing usage of the data.

In response to these gaps, the concept of DataOps emerged through the back half of last decade. The scope and definition of DataOps is still somewhat inconsistent, but generally it takes key concepts from Agile and DevOps (and arguably Lean Manufacturing) and places a data lens over the top. The DataOps leadership principles authored by Toph Whitmore are a very common reference:

- Establish progress and performance measurements at every stage of the data flow. Where possible, benchmark data-flow cycle times.
- Define rules for an abstracted semantic layer. Ensure everyone is "speaking the same language" and agrees upon what the data (and metadata) is and is not.
- Validate with the "eyeball test": Include continuous -improvement -oriented human feedback loops. Consumers must be able to trust the data, and that can only come with incremental validation.
- Automate as many stages—even BI, data science, and analytics, when possible—of the data flow.
- Using benchmarked performance information, identify bottlenecks and then optimise for them. This may require investment in commodity hardware, or automation of a formerly-human-delivered data-science step in the process.
- Establish governance discipline, with a particular focus on two-way data control, data ownership, transparency, and comprehensive data-lineage tracking through the entire workflow.
- Design process for growth and extensibility. The data flow model must be designed to accommodate volume and variety of data. Ensure enabling technologies are priced affordably to scale with that enterprise data growth.

For a data team, the take on DataOps principles espoused by consultancy DataKitchen in their [DataOps Manifesto](#) is a useful framework for assessing maturity. We find this list provides a great tool for facilitating improvement opportunities for data teams that aspire to best practice. For example, one of the most common principles that teams identify as an area for improvement is 'Make it reproducible'. Investigating versioning practices across data sets and software configurations often exposes risks that the data team can address before those risks manifest themselves as hard to analyse data quality issues.

In summary, if you are responsible for the effectiveness of a data team, keep an eye on the DataOps movement and look to leverage thought leadership around getting the most out of your increasingly valuable data professionals.

If your business needs help maximising the effectivessness of their DataOps, please contact a Fujitsu Data & AI specialist now.

Contact

Fujitsu Data & AI

+61 3 9924 3000