



# THE ULTIMATE GUIDE TO MODERN DATAOPS



**PEOPLE, PROCESS & TECHNOLOGY -  
YOUR KEYS TO TRUSTED DATA**

---

# Table of Contents

Introduction	_____	<b>03</b>
The DataOps imperative	_____	<b>04</b>
Is data your top priority yet?	_____	<b>05</b>
The Modern DataOps Way	_____	<b>08</b>
How did you do?	_____	<b>16</b>
Seize the DataOps Advantage	_____	<b>17</b>
Get in touch	_____	<b>18</b>

# Introduction

Innovative organisations across the world are mastering the management and utilisation of data to create superior products, better experiences, and higher margins.

Data Operations (DataOps) has risen as a pivotal methodology, redefining the way organisations handle their data pipelines, processes, and teams.

Our guide begins with an overview of the multifaceted benefits of DataOps - how it drives trusted data, accelerates delivery aligned with business velocity, fosters scalable and efficient teams, simplifies operations, minimises issues, and ultimately reduces costs.

The second half of our guide is dedicated to technology. Our team at Matatika are data experts who understand that strategy needs to be made practical. In our technical section entitled "*The Modern DataOps Way*", we share practical steps for your teams to implement DataOps starting now.



# The DataOps imperative

Data is the lifeblood of modern organisations - steering strategic decisions, nurturing customer relationships, and enabling AI innovation.

The exponential growth of data, coupled with the increasing complexity of data ecosystems, has posed significant challenges in managing, processing, and deriving meaningful insights from vast troves of information. As businesses strive to grow in this data-driven environment, conventional approaches to data management and operations are no longer good enough. The systems and processes are failing to deliver the agility, accuracy, and efficiency modern data-driven organisations demand.

Enter DataOps, an innovative data operations methodology inspired by DevOps, Agile, and Lean methodologies. DataOps strives to create a collaborative and streamlined environment that bridges the gap between data engineers, data scientists, analysts, and business stakeholders. By fostering a culture of collaboration, automation, and continuous improvement, DataOps supports a data lifecycle that delivers huge benefits across the entire spectrum of business operations.



...conventional approaches to data management and operations are no longer good enough.

# Is data your top priority yet?

## Transform your results with these 6 key objectives.

Data is the most valuable asset held by all companies in the world. Yet 56% of CEOs don't trust their data. Organisations embracing DataOps position themselves at the forefront of innovation and competitiveness, ready to harness the full potential of their data assets.



### TRUSTED INFORMATION

#### **The Foundation of Informed Decision-Making**

Business decisions require trusted information, from a trusted source. DataOps first and foremost addresses these key issues of trust. Whereas traditional data management practices often led to siloed and inconsistent data sources - creating mistrust among stakeholders - DataOps addresses these demands by establishing robust data pipelines, automated quality checks, and data lineage tracking. This paves the way for a single source of truth, instilling confidence in data consumers and empowering them to make well-informed decisions.

---



### LOWER COSTS

#### **Optimising Resource Utilisation**

Sustainable business growth requires a constant focus on costs. DataOps aligns with this objective by optimising resources allocated to systems, efficient collaboration, and reducing human based operational overhead. By minimising waste, streamlining processes, and focusing on optimal resource utilisation, DataOps helps organisations achieve their goals while maintaining a competitive cost structure.

---



## SCALABLE TEAMS

### **Empowering Collaboration and Cross-Functionality**

Traditional data management has suffered from silos that hinder collaboration and knowledge sharing between data-related functions. DataOps shifts the focus to cross-functional teams that are able seamlessly collaborate across the data lifecycle. By promoting knowledge sharing, skill diversification, and continuous learning, DataOps not only ensures streamlined requirements to delivery but also empowers teams to tackle complex challenges with collective expertise.

---



## RELIABLE OPERATIONS

### **Increased Automation and Simplification**

Automation, data as code, and continuous integration & deployment enables DataOps teams to reduce manual work, minimise errors, and maximise their operational efficiency. This automated and code promotion approach not only reduces effort but also allows data professionals to focus on value-add activities rather than repetitive donkey work.

---



## REDUCED OPERATIONAL RISKS

### **Fewer Issues and Enhanced Reliability**

The modern complexity of data systems can breed issues and impede progress through constant operational issues. DataOps combats this challenge by introducing data centric automated testing, monitoring, and error detection mechanisms. This proactive approach allows teams to identify and address potential issues before they impact business processes. As a result, teams deliver more changes with reduced downtime, improved data quality, and higher confidence.

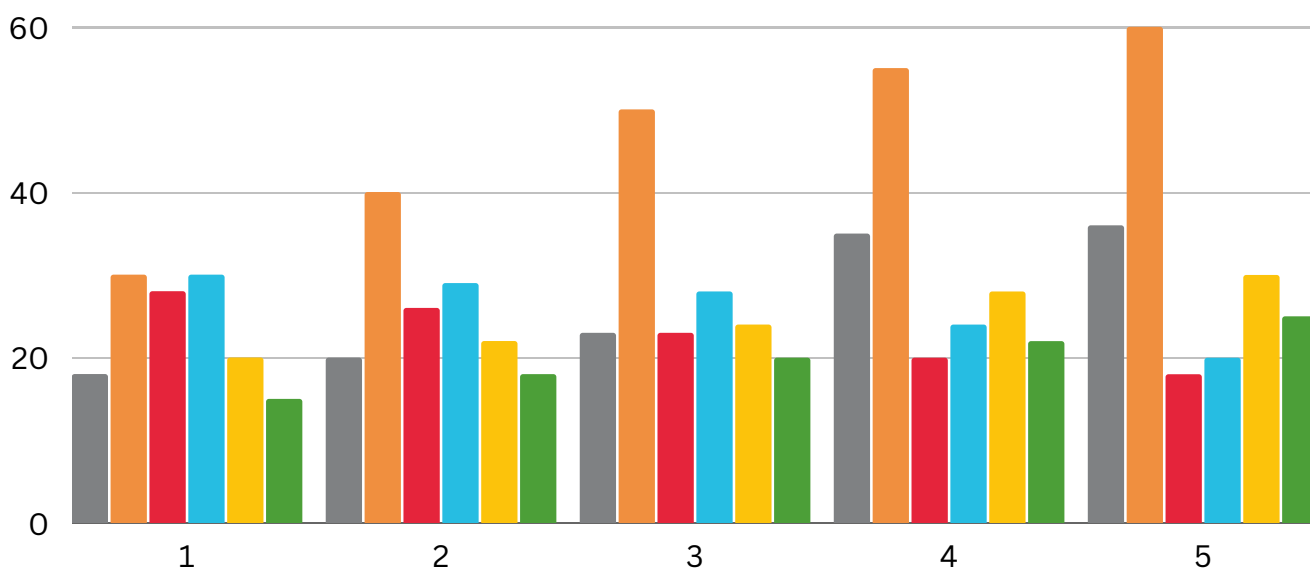
---



## BUSINESS AGILITY

### Accelerated Insights and Innovation

To keep pace with rapidly changing marketplaces, businesses must swiftly transform raw data into actionable insights. A DataOps environment aims for data workflows that are automated, all code is versioned, and collaboration is seamless. As a result, teams can rapidly iterate on data processing tasks, experiment with new ideas, and introduce innovations to market faster. By aligning data delivery with the speed of business requirements, DataOps greatly reduces innovation cycles, giving organisations a competitive edge.



Increase trust, reduce costs, be more agile!

# The Modern DataOps Way

This section highlights the strategies and tactics your data team can use to implement DataOps today. Over and above technology, DataOps is a culture shift - best practices that need to be encoded into your organisational ways of working. Using our guide we aim to help teams accelerate that shift. These technologies and techniques are applicable even if you are not yet a Matatika customer.

## Expect changes

“Change is the only constant.” - Heraclitus.

Data teams must frequently modify systems, connections, and logic sometimes with very little advance notice. The process of delivering these changes can end up more time consuming than the functional change itself and each change carries an inherent risk. Counter intuitively, the remedy lies in increasing the frequency of change delivery. Striving for this objective compels the adoption of automation, testing and other practices that mitigate the delivery overheads and potential hazards of human oriented mistakes.

A repeatable and automated release process is an essential modern data management practice, driving efficiency and reliability in your data teams' workflows. A successful release and change promotion process relies on data teams adopting source code management of configuration, versioning of artefacts, and testing of correct behaviour. By treating all components as code, teams ensure consistency and traceability in their operations. In addition to configuration, data transformation logic must also be managed as source code, allowing for trusted data and rapid iteration. Fundamentally, we are aiming to automatically propagate all our changes in deployment pipelines - removing human errors.



As data teams gain experience, their focus extends beyond the basics towards optimising the entire value chain. From developers to production, each phase has optimisation opportunities. This growing sophistication demands a repeatable change promotion process, spanning connectors, models, reports, and visualisations. Leveraging automated release systems, teams guarantee that every modification is consistent and reliable. Change promotion should become a 'boring' task that never goes wrong, freeing time to navigate all the complexities of modern data management.

## Level

## Activity

---

### Foundational



Data teams use source code to manage configuration, versioning and testing of connectors and transformations



Data teams are able to automatically promote all changes for connectors and transformations



Data teams are able to automatically rollback all changes for connectors and transformations

---

### Advanced



Data teams are optimising the delivery of the entire value chain; during development the file you change is the file you commit



Data teams promote all changes to connectors, models, reports, and visualisations with automated release systems



Data teams are releasing multiple times per day with no down time

Expect changes continued...

---

#### Links

- [Introduction to change promotion Learn more...](#)
- [Setup change promotion in the Matatika platform Read our documentation...](#)

## Think about the future

“If a tree falls in the forest, and there's no one around to hear it, does it make a sound?” - Dr. George Berkeley

Embracing an automated testing philosophy guarantees seamless and accurate operation of an organisation's entire data pipeline. Crucially, by validating data model correctness, organisations drastically reduce the risk of flawed outputs and subsequent downstream consequences, bolstering data quality and fostering confidence in the business insights. Data teams adopting this philosophy also expedite their development by pre-emptively identifying coding issues. Up to date documentation and tests go hand in hand to create a maintainable codebase - one that is robust and trusted even in the face of complex data and highly changeable workflows. This small additional effort reduces development time and increases responsiveness to business needs.

Establishing a solid foundation for efficient data processing often begins with the adoption of self-documenting artifacts, where data teams are seamlessly integrating documentation within the code itself. This approach not only enhances understanding but also expedites collaboration as the code becomes more comprehensible to fellow team members. Data teams that are embracing maintainability with documentation, will naturally seek out further opportunities such as incorporating unit tests to validate the accuracy of their transformation code. With minimal effort the codebase can begin to evolve, ensuring the established functionality remains intact, minimising the risk of unintended errors and facilitating smoother iterations.

## Think about the future continued...

Moving beyond the basics, advanced data strategies such as linting and modularisation seek to optimise the code quality even further. Linting tools scrutinise the style and formatting of the codebase, ensuring adherence to established coding conventions. Whereas creation of modular and reusable components aims for consistency and correctness in common functionality. Modules adopted across the entire organisation not only save effort, but are a strong indication of a move away from siloed thinking toward a culture of collaboration. By prioritising maintainability through self-documentation, testing, style checks, and modular design, data teams are actively driving efficiency, reliability, and scalability in their endeavours.

Level		Activity
Foundational	<input type="checkbox"/>	Data teams are creating and publishing documentation with self-documenting artefacts - documentation in code.
	<input type="checkbox"/>	Data teams are writing maintainable transformation code using unit tests to assert the correct behaviour
	<input type="checkbox"/>	Data teams are writing maintainable transformation code with code reviews
Advanced	<input type="checkbox"/>	Data teams are checking the style and formatting of their transformation code with linting tools
	<input type="checkbox"/>	Data teams are writing reusable and testable modules, adopted by the whole organisation
Links		<ul style="list-style-type: none"><li>• Introduction to developing and testing transformations <a href="#">Learn more...</a></li></ul>

Think about the future continued...

Links

- Publishing modules with an Enterprise Data Hub [Learn more...](#)
- Setup automated deployment and testing [Read our documentation](#)

## Manage expectations

“Beep, beep... Ring, ring...” - Iam, Data Team nightmare

Data teams are accustomed to the intricacies of managing business expectations and providing swift answers. However, the management of the data pipeline expectations is equally vital but often overlooked. The production of insightful, business-ready information hinges on data teams precisely aligning their data requirements to the systems of record. The arrival of subpar data can have damaging consequences for the entire business when decisions are made on the wrong basis or loss of trust leaves leaders no choice but to ignore the data. Teams need streamlined processes, seamless collaboration, and codified expectations to promptly identify and address potential data discrepancies.

At the foundational level, asserting the availability and correctness of requisite data and its corresponding format is imperative. Collaborative processes facilitate the seamless flow of feedback, allowing discrepancies in data expectations to be effectively communicated back to the data producers. This foundational step ensures that the groundwork for data alignment is laid, preventing misunderstandings and maintaining trust.

As the level of data dependency and sophistication in an organisation grows, it becomes crucial to gain early access to any modifications made by data producers and to coordinate changes proactively. This proactive coordination prevents middle-of-the-night surprises and supports the alignment of data with evolving business needs. As your organisation matures, gradually shifting the responsibility of testing data expectations to the data producers becomes

## Manage expectations continued...

viable. Empowering data producers to participate in expectations testing fosters a culture of mutual accountability and facilitates the early identification of potential data issues, resulting in more robust and trusted business-ready data.

Level		Activity
Foundational	<input type="checkbox"/>	Data teams assert required data availability and type
	<input type="checkbox"/>	Data teams and data producers have established collaborative processes to communicate broken data expectations
Advanced	<input type="checkbox"/>	Data teams are testing data producers changes early and coordinate changes
	<input type="checkbox"/>	Data producers are testing expectations within their development and change management



# Reduce noise

“Wolf! Wolf!” - The Boy Who Cried Wolf, Aesops Fables

The availability and reliability of information is the benchmark for all data teams. As organisations increasingly derive business critical decisions from their data, the operational demands on the data teams grow correspondingly. At a minimum, data teams are required to promptly address incidents and need to move towards anticipating issues. In more sophisticated operations, it becomes essential to operate services with a focus on reliability such as alerting, proactive monitoring and dedicated teams. Recognising that issues, challenges, and errors are a part of every service, operational services that generate excessing noise risk being ignored.

Level	Activity
Foundational	<input type="checkbox"/> Data teams have monitoring and alerting for business critical functions
Advanced	<input type="checkbox"/> Data teams are aiming for zero errors
Links	<ul style="list-style-type: none"><li>• Setup notifications and alerts in the Matatika platform <a href="#">Learn more...</a></li></ul>

# Be agile

“Quickly come and quickly go” - Chinese sage, circa 400 B.C.

Recognising that working directly on production is the root cause of many avoidable incidents and outages, these systems are justifiably locked down. To facilitate innovation, experimentation, testing and problem resolution, data teams need a production-like systems without the production risk. By

## Be agile continued...

establishing a seamless process for replicating environments, the pathway to effective resolution is made considerably easier, and potentially all projects can be delivered faster. Environments need to become throw away items - easy come, easy go.

In today's data-driven landscape, data teams play a pivotal role in not only delivering insights but also shaping them into valuable products. The concept of data as a product has gained significant traction, wherein data teams meticulously curate, refine, and package raw data into consumable offerings for both internal stakeholders and external customers. This approach involves understanding the unique needs of various users and tailoring the data product to cater to their specific requirements, be it visualisations, reports, or interactive dashboards. Simultaneously, data teams are also at the forefront of enabling self-service analytics, empowering users across the organization to explore data independently. By developing user-friendly tools, interfaces, and platforms, data teams democratise access to data-driven insights, fostering a culture of informed decision-making. This dual role of data teams in delivering data as a product and providing self-service analytics not only enhances efficiency but also drives innovation and strategic growth within the organisation.

Level		Activity
Foundational	<input type="checkbox"/>	Data teams are able to develop in an environment near identical to production
	<input type="checkbox"/>	Data teams are able to create environments on demand
Advanced	<input type="checkbox"/>	Data teams delivery data as a product
	<input type="checkbox"/>	Data teams provide self-service analytics
Links		<ul style="list-style-type: none"><li>• Setup of development workspaces in the Matatika platform <a href="#">Read our documentation...</a></li></ul>

# How did you do?

Using our guide, consider each foundational level activity and any advanced practices you've mastered.



## Apprentice

Applying some foundational practices.



## Craftsman

Applying all foundational and some advanced practices



## Master

Consistently applying all practices.



What actions will you take now?



# Seize the DataOps Advantage

Embracing DataOps enables organisations to not only survive, but thrive - driving trusted data, facilitating swift and agile delivery, empowering scalable teams, enhancing efficiency, simplifying operations, mitigating issues, and lowering costs.

At Matatika, we are committed to building the next-generation DataOps platform for business. We are continually expanding our integrations with a wider variety of data sources, apps, insights, and plugins that can be immediately utilised by businesses of all sectors and sizes.

Our team at Matatika loves to talk about data; if any of these DataOps topics resonate with you, please [get in touch](#).

## DataOps empowers organisations to harness the full potential of their data.

# Get in touch

[www.matatika.com](http://www.matatika.com)

[info@matatika.com](mailto:info@matatika.com)

---