



# DATAOPS in Practice

Vincent Goris & Niels Naglé

# ▲ Vincent Goris

**Analyst**

**Chapter lead Data & AI - Healthcare**

Vincent.Goris@infosupport.com

*[www.linkedin.com/in/vincent-goris](http://www.linkedin.com/in/vincent-goris)*



# ▲ Niels Naglé

**Data Architect / Trainer / Speaker**  
**Chapter lead Data & AI - Agri, food and retail**  
**Niels.Nagle@infosupport.com**

*[www.linkedin.com/in/nielsnagle](http://www.linkedin.com/in/nielsnagle)*





# Inhoud

01

DataOps != DevOps

02

Impact of DataOps

03

Roadmap to DataOps

04

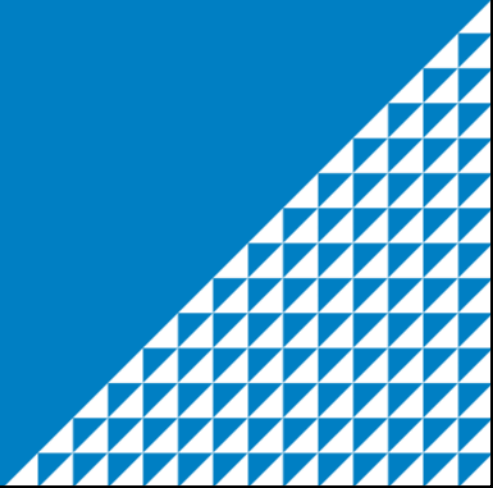
Challenges of DataOps

05

Wrap up & Questions



DataOps  $\neq$  DevOps



“

DevOps is the ‘father’ of DataOps

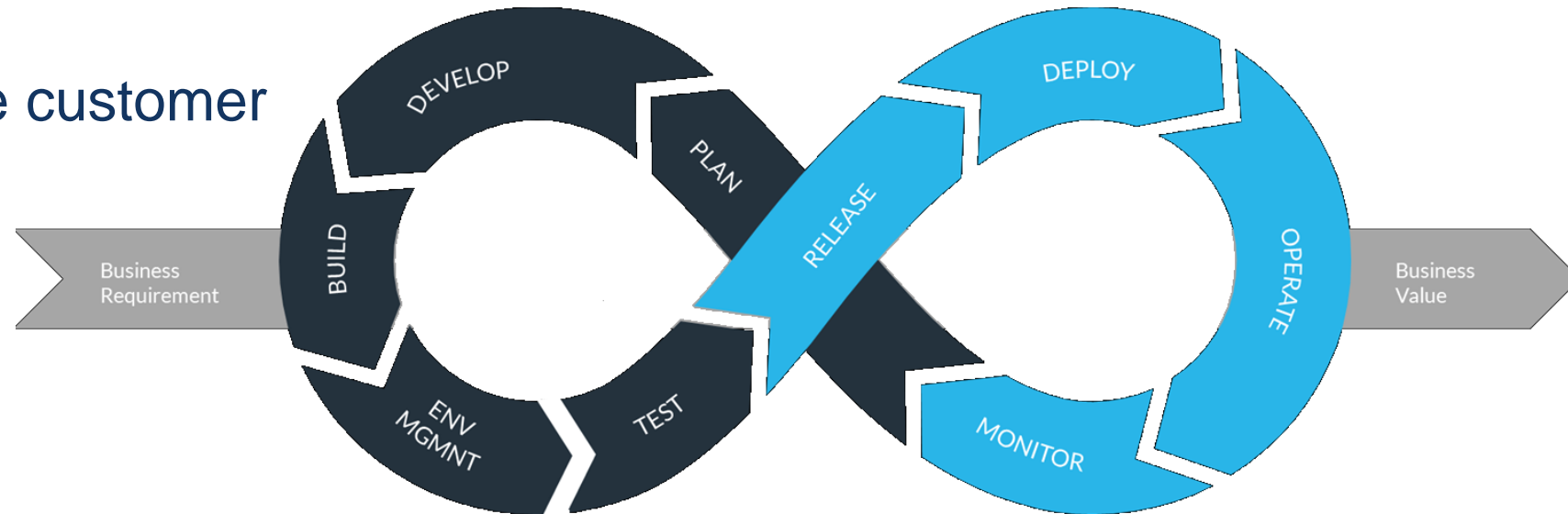
**Peter Travkin**

*DataOps vs DevOps, what's the difference?*



# ▲ What is DevOps

- Shared Responsibility
- Automating, Monitoring and Testing
- Shortened release cycle
- Feedback cycle
- More value for the customer



# What is DataOps?

- Little agreement
- Many definitions





"spans the entire analytic process, from data acquisition to insight delivery"

"addresses the needs of data professionals on the modern internet and inside the modern enterprise"

"all activities between the data and operation teams"

"an approach to eliminate data silos by connecting different data pipelines"

"an integrated perspective over the entire data lifecycle"

Gartner: "hub for collecting and distributing data, with a mandate to provide controlled access to systems of record for customer and marketing performance data, while protecting privacy, usage restrictions and data integrity"



**Table 2.** A list of DataOps components and the emphasize of the interviewed experts

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>	<b>C5</b>	<b>C6</b>	<b>C7</b>	<b>C8</b>
<b>Goals:</b>								
Continuous improvement	X	X	X	X	X		X	
Orchestration		X				X	X	X
Empowerment of citizen users		X			X		X	
Agility & speed	X	X	X	X	X	X	X	X
Collaboration & trust		X		X	X		X	
<b>Principles:</b>								
Reuse of artifacts	X	X						X
Automation	X	X	X	X	X	X	X	X
Integrated end-to-end thinking		X	X			X	X	X
Short cycles & incremental change	X	X	X			X		
Analytics as code	X	X	X			X		
Testing	X	X		X		X	X	X
Monitoring	X	X		X	X	X	X	X
Data-driven improvement	X	X				X		
Process-oriented Data pipelines		X	X	X			X	X

Source: DataOps – Towards a Definition, J. Ereth



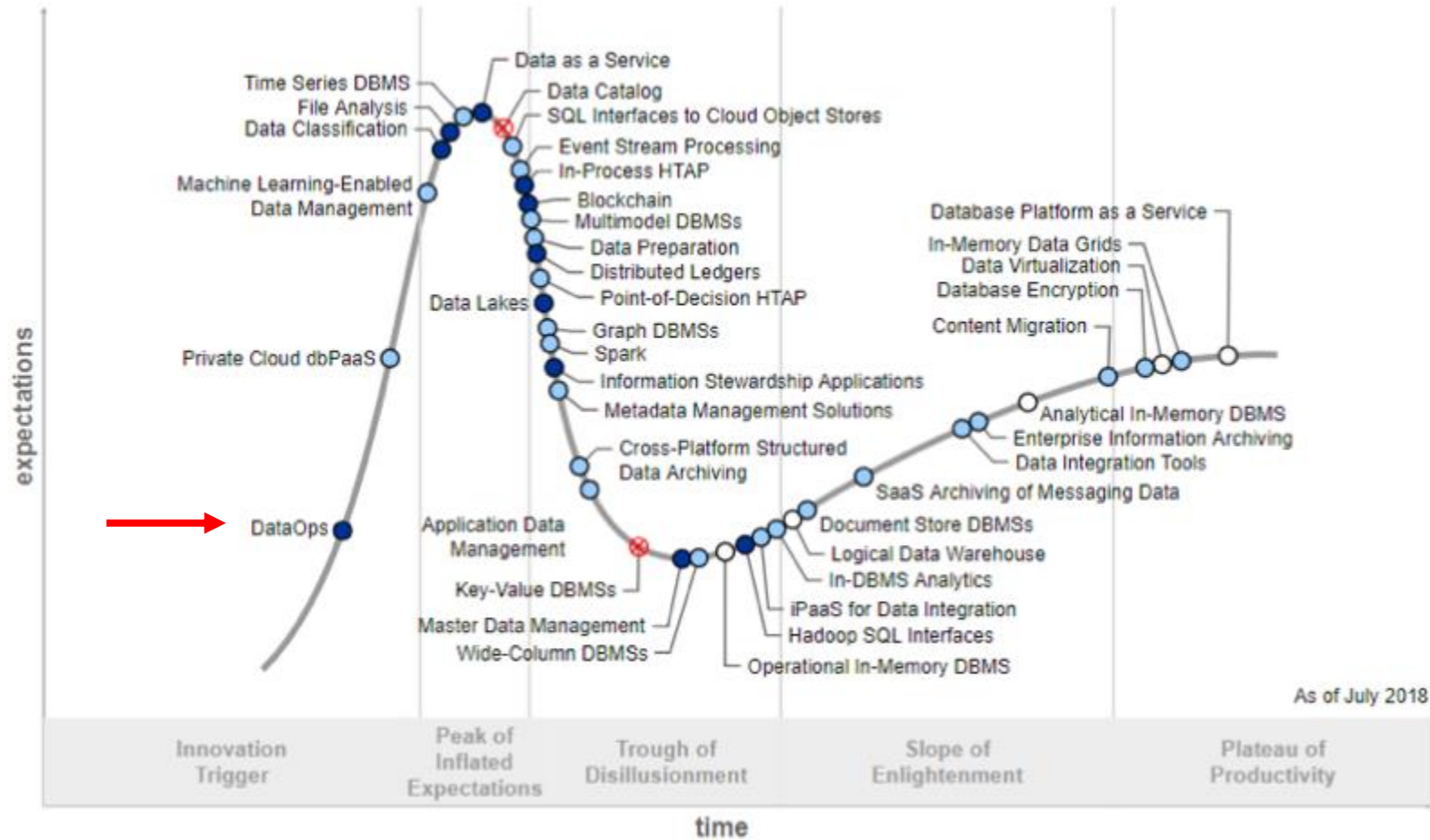
# DataOps

## DEFINITION

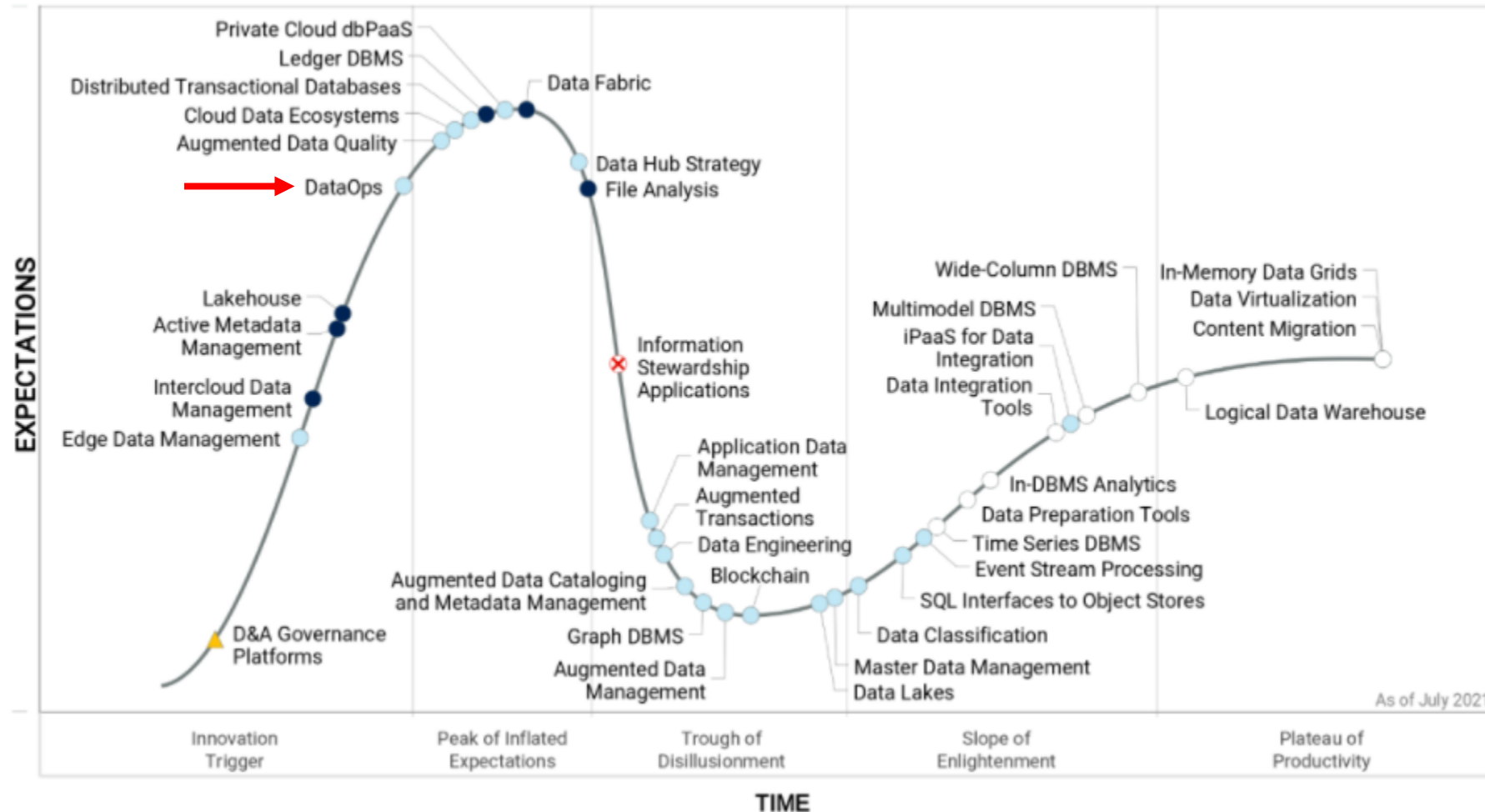
"DataOps can be defined as an approach that accelerates the delivery of high-quality results by **automation** and orchestration of **data life cycle stages**. DataOps adopts the best practices, processes, tools and technologies from **Agile software engineering** and **DevOps** for governing analytics development, optimizing code verification, building and delivering new analytics thereby promoting the **culture of collaboration and continuous improvement**.

*From Ad-Hoc Data Analytics to DataOps, Munappy, A., Issa Mattos, D., Bosch, J. et al (2020)*

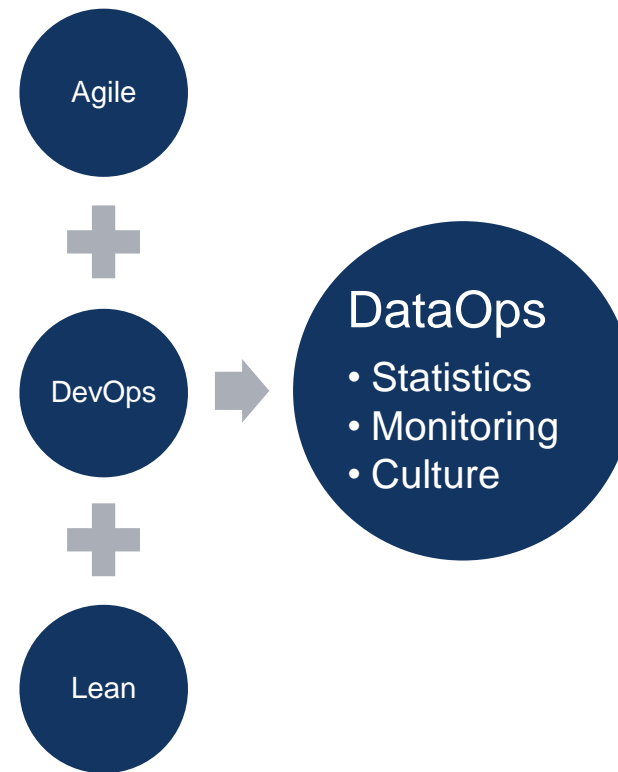
# ▲ Gartner Hype Cycle for Data Management



# ▲ Gartner Hype Cycle for Data Management



# Relation DevOps - DataOps





# Relation DevOps - DataOps

- **Differences: product delivered**
  - Static product vs mutating object
  - Fluid data and business needs
  - Metrics
  - Difference in pipelines









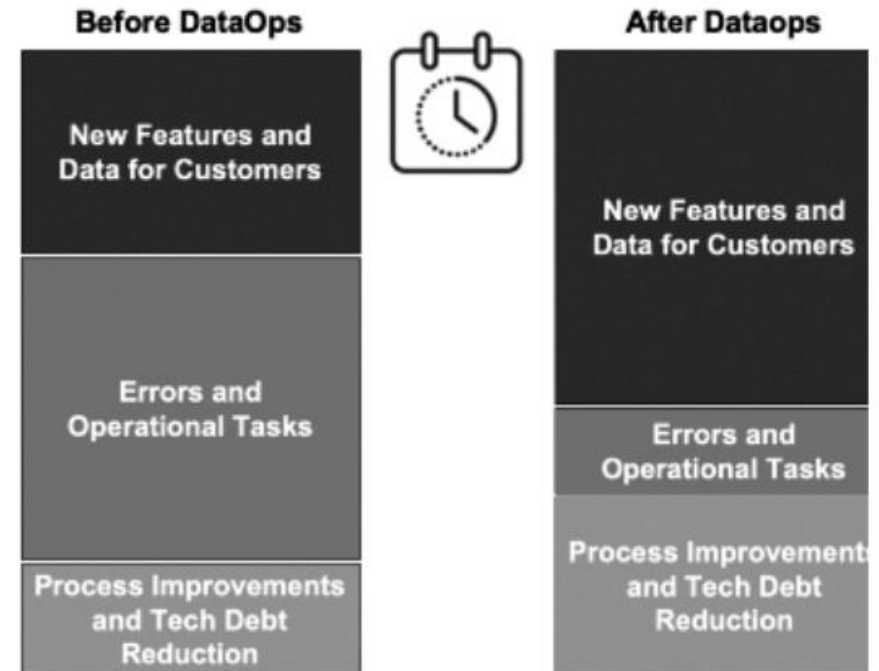
# Impact of DataOps





# ▲ Impact of DataOps

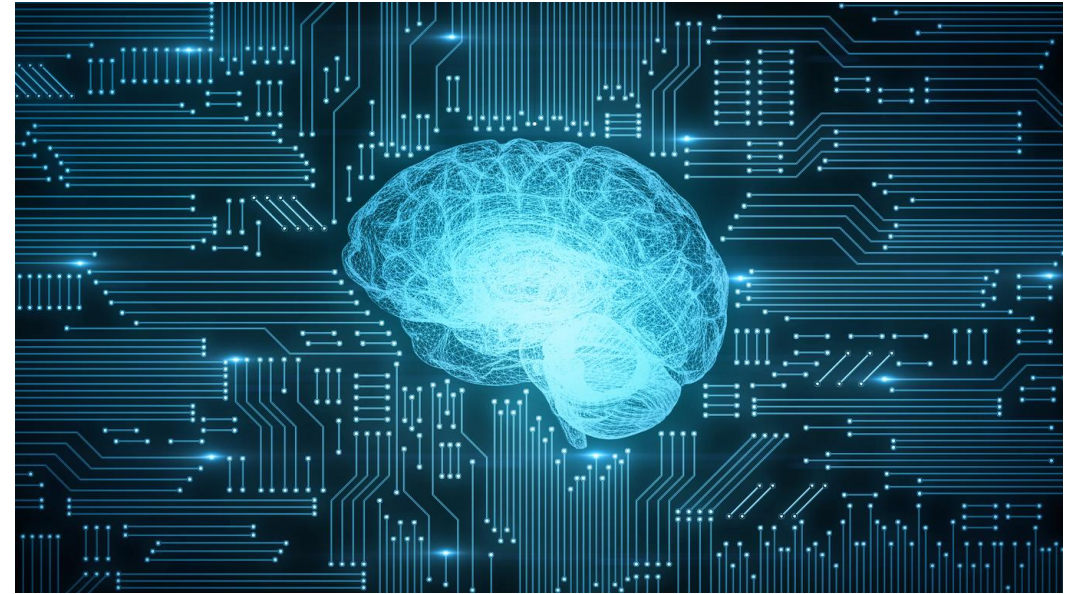
- Faster time to release
- More effective
- More value
- Better quality assurance



*Recipes for DataOps Success - DataKitchen*

# ▲ Knowledge is power

- Quality assurance
- Monitoring data
- Automated pipelines
- Mature Data Governance
- User collaboration
- Feedback loops
- In it for the long run



“

Building a factory is 100 times as hard as building a car

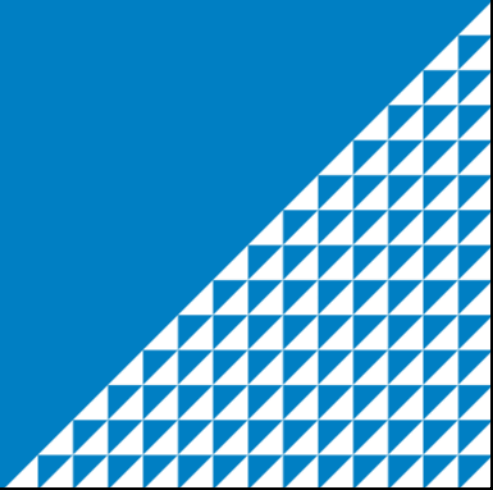
**Elon Musk**

*Techno King*





# Roadmap to DataOps



“

Culture eats strategy for breakfast

**Peter Drucker**

*Management Consultant*






# Roadmap to implement DataOps

## 1. Scan

 Data Driven Maturity

 DataOps Maturity

## 2. Educate

 People, Process & Technology

 Lean manufacturing


## 3. Pilot

 Business Use-cases

 CI/CD

 Showcase

## 4. Productize

 Metadata, catalog and governance

 Automate & Smart



## 5. Grow



# ▲ 1. Scan

- First step of the road to implement DataOps is to understand the gap between your current state and your desired state.



# Data-driven maturity model



# Data-driven maturity model

## Collect & Store

Focus on orchestrating efforts to **set up** analytics .

- ✓ Make data available for analyzing,
- ✓ Basic reporting.
- ✓ Analytics and a data-driven way of working must become official.

## Explore

Analytics is getting more **formal** and **organized**.

- ✓ Insight shifts from the manager to the analytics department
- ✓ Many of the prerequisites have been taken care of the next stage.
- ✓ Management views predetermined reports and KPI 's for more informed decisions.

## Stream & Operationalize

Focus on **developing** and **operationalizing** prescriptive analytics .

- ✓ Now what to do in the future to optimize our profitability
- ✓ Operationalizing insight without manual actions
- ✓ Recommendation services prescribe products to customers they will more likely respond to.

## Predict

Focus on **empowering** everyone in the organization.

- ✓ Allowing rapid integration of analytical features into products or services.
- ✓ Everyone is warmed up for change during the previous stages.
- ✓ Continually showing the success of an analytical approach.

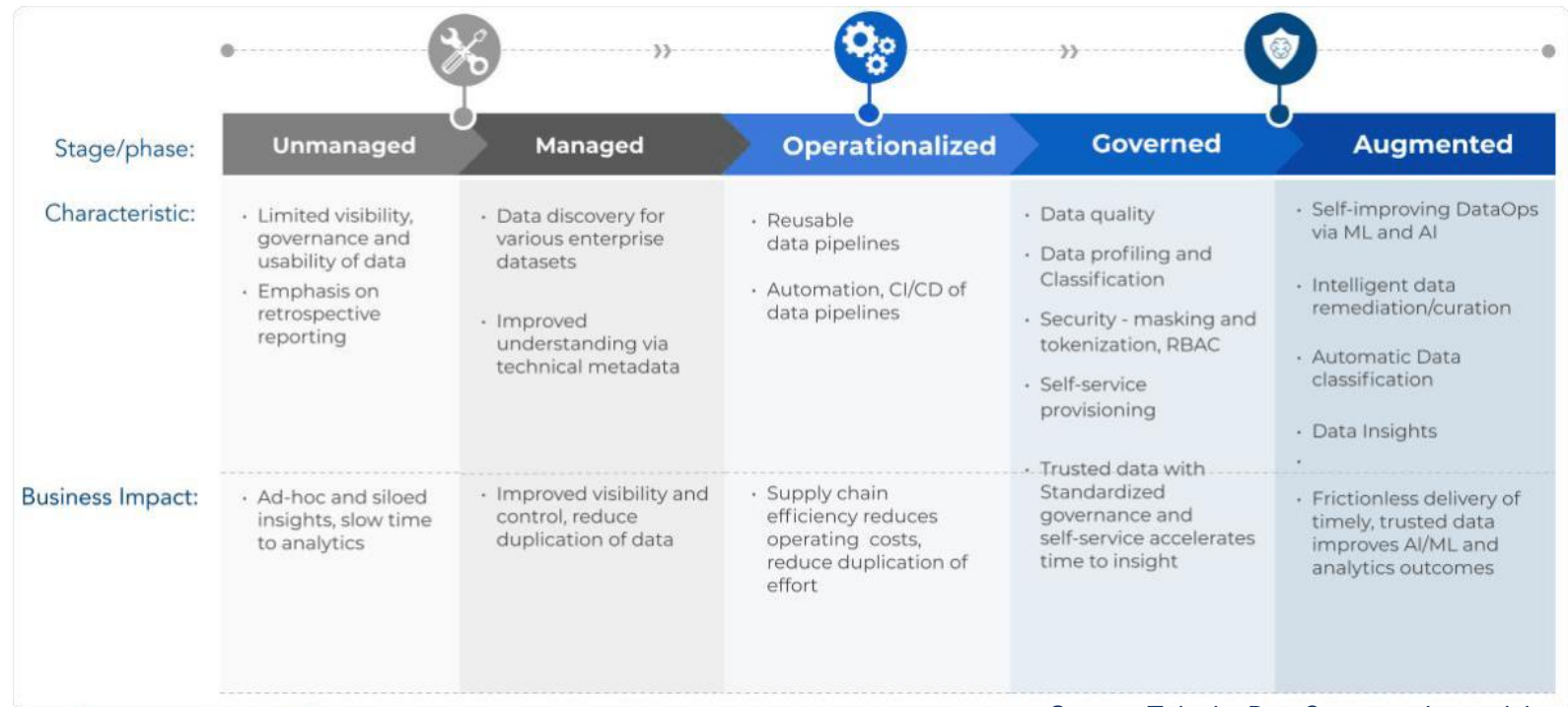
## Prescribe

Main source of **competitive advantage**

- ✓ The way the business is done is defined
- ✓ Innovative products and services can redefine the business model
- ✓ Analytical transformation is a journey that never ends. The organization keeps progressing and keeps redefining itself in everything it does.
- ✓ Continuous improvement is so deeply ingrained in the organization, it becomes second nature.

# ▲ DataOps maturity

- Unmanaged
- Managed
- Operationalized
- Governed
- Augmented



Source: Zaloni – DataOps maturity model

# Roadmap to implement DataOps

## 1. Scan

 Data Driven Maturity

 DataOps Maturity

## 2. Educate

 People, Process & Technology

 Lean manufacturing


## 3. Pilot

 Business Use-cases

 CI/CD

 Showcase

## 4. Productize

 Metadata, catalog and governance

 Automate & Smart

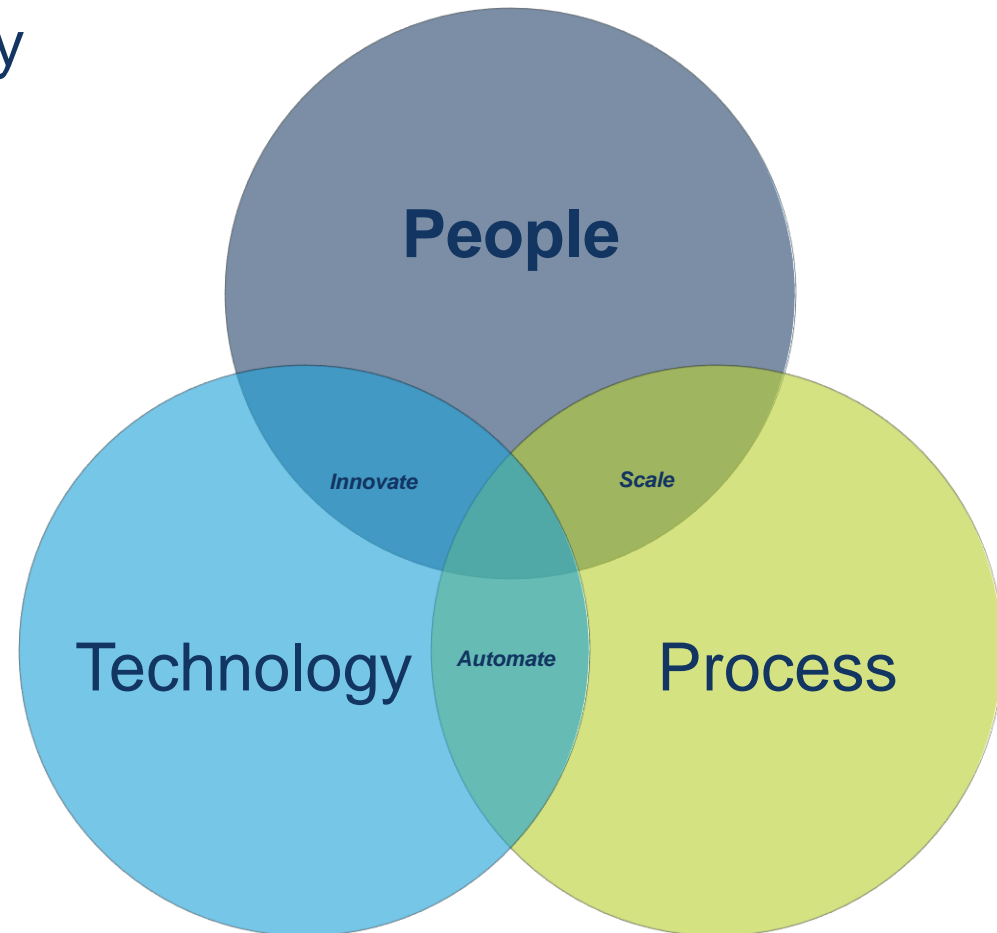


## 5. Grow



## ▲ 2. Educate

- **People, Process & Technology**



## ▲ 2. Educate

- Technologies and process (might be logical and straightforward)
  - Technology guidance
  - Continuous Integration / Continuous Deployment
  - Data Quality processes
- Make people a priority, because it is **a cultural change** that needs to be lead
  - Roles and responsibilities
  - Demonstrate
  - Build trust
  - Build on this in small steps
  - Continuous delivery
  - Feedback loop
  - Communicate




# Roadmap to implement DataOps

## 1. Scan

 Data Driven Maturity

 DataOps Maturity

## 2. Educate

 People, Process & Technology

 Lean manufacturing


## 3. Pilot

 Business Use-cases

 CI/CD

 Showcase

## 4. Productize

 Metadata, catalog and governance

 Automate & Smart



## 5. Grow



## ▲ 3. Pilot

- Business use-cases
  - Start small with high impact
- CI/CD
  - Use version control system
  - Branch and merge
  - Use multiple environments (one for each developer)
  - Reuse (data pipeline reuse)
  - Monitoring
  - Automated Testing

## ▲ 3. Pilot

- Showcase
  - Impact of DataOps in the Business use case
  - Data Insights – Quality, size
  - Amount of rows (new, updated, deleted)
  - Throughput
  - Error rate
  - Test performance
  - Burn down charts
- Don't forget the people side
  - Happiness factor of users


# Roadmap to implement DataOps

## 1. Scan

 Data Driven Maturity

 DataOps Maturity

## 2. Educate

 People, Process & Technology

 Lean manufacturing


## 3. Pilot

 Business Use-cases

 CI/CD

 Showcase

## 4. Productize

 Metadata, catalog and governance

 Automate & Smart

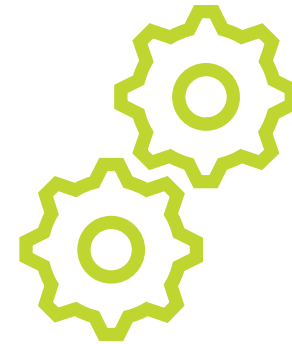


## 5. Grow



## ▲ 4. Productize

- Continually focus on lowering
  - Error rates
  - Cycle times
  - External dependency's
- Metadata, Catalog and Governance
  - Is at the core of DataOps
  - Without Metadata no automation
  - Traceability and ownership



## 4. Productize



### Automate

Adaptable architecture in which data flows continuously, designers must automate everything

- Profile and tag data
- Map data to existing data sets (metadata injection)
- Detect changes and identify the impact on downstream objects and applications
- Detect anomalies and notify.



## ▲ 4. Productize



### Smart

The ideal data architecture is more than just automated. It uses machine learning and artificial intelligence to learn, adjust, alert, and recommend.

- Build data objects
- Identifying data types, common keys and relations
- Identify and fix data quality errors
- Recommend related data sets and analytics.




# Roadmap to implement DataOps

## 1. Scan

 Data Driven Maturity

 DataOps Maturity

## 2. Educate

 People, Process & Technology

 Lean manufacturing


## 3. Pilot

 Business Use-cases

 CI/CD

 Showcase

## 4. Productize

 Metadata, catalog and governance

 Automate & Smart



## 5. Grow



“

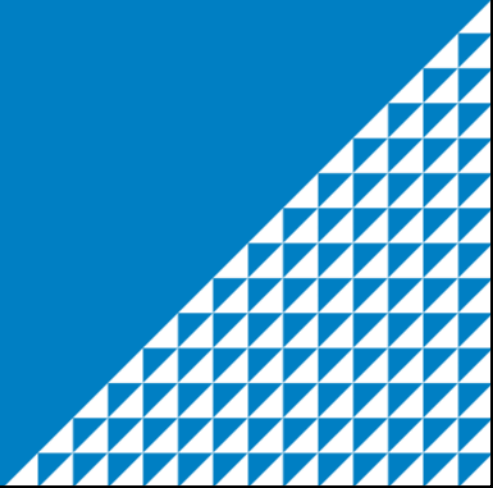
Life's a marathon, not a sprint

Phillip C. McGraw





# Challenges of DataOps





# Data-driven culture

- One of the key roadblocks to developing such a data-driven environment can be an organization's culture.
- Teams typically need to transition from more traditional hierarchical chains of command to a data-driven culture, an environment in which data and evidence are transparently used to make business decisions as opposed to more opaque top-down management.





# Common Mistakes

- Spend too little time on education (roles & responsibility)
- Focus only on the Technology/Tool
- Stopping DataOps at the border of the central team
- Only collecting the basic metrics and quality aspects
- Misalignment of the data team and business users



# ▲ Critical success factors

- Acknowledge the cultural change
- Communication, collaboration and governance
- Iterate and gather feedback from different aspects and learn
- Invest in the platform and technology
- Embracing AI and machine learning as part of DataOps

“

Dream big. Start small.  
But most of all, start!

Simon Sinek



# Info Support

PASSION LED US HERE



# ▲ Questions





*.infoSupport*  
*Solid Innovator*

# ▲ Data Driven Maturity scan

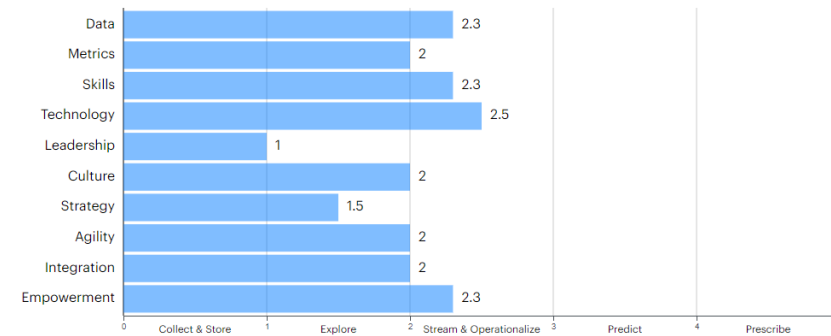
Uw organisatie ziet de waarde van data.  
Benieuwd hoe uw organisatie er voor staat  
op de data driven maturity.

Binnen 10 minuten een beeld van je data  
maturity en tips om nog verder te groeien.  
**Data-Driven Maturity Scan!**



[Data-driven maturity scan \(data-maturity-scan.azurewebsites.net\)](https://data-maturity-scan.azurewebsites.net)

Data maturity detailed per dimension



## Improve your scores

Based on your scores, we recommend taking a look at guidance about the following dimensions and how to improve them:

- Leadership
- Strategy
- Metrics

# ▲ Data-Driven Business, Make IT Happen!

Uw organisatie bezit veel data.  
Dat is een goudmijn die niet altijd  
optimaal wordt benut.

Meer weten over Data-Driven business?

Download gratis onze compacte gids  
**Data-Driven: Make IT Happen!**

