

# Substring - The Data Company

---

**Komponenten und Charakteristiken  
moderner Datenarchitekturen**

Dataholics Webinar-Reihe



**Nino Müller**  
Partner & Lead Technology  
@ Substring

Teilnehmen unter  
**attend.sl/egev**

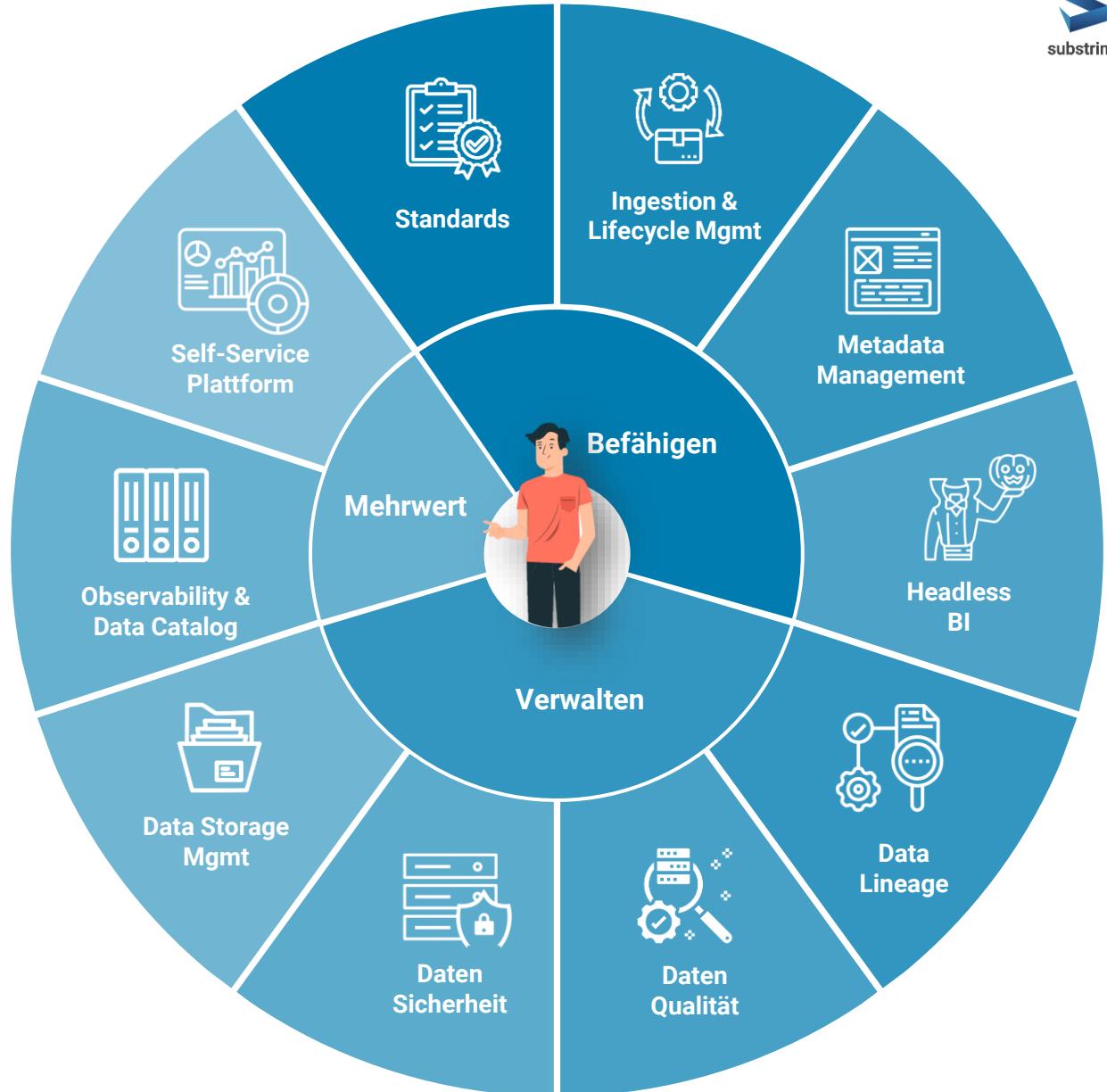




Eine  
Organisation  
generiert heute  
Daten. Überall  
und konstant



**Eine Datenarchitektur beinhaltet Technologie, Standards & Komponenten, um Daten effizient und effektiv bewirtschaften und nutzen zu können.**





**Vom Data Lake...**

A photograph of a swampy area. The foreground is covered in numerous green lily pads floating on the water. In the background, there is a dense line of green vegetation, including tall grasses and some palm trees. The overall scene is a natural, overgrown wetland.

**... zum Data Swamp.**

# Das Versprechen einer Datenarchitektur



Effektive und effizientes Verarbeiten von wachsenden Datenmengen



Business Objekte, Use Case-basierte Datasets, Historische Daten



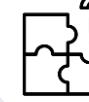
Standards wie Once-Only setzen



Zugang zu Daten für Mensch & Maschine um Datensilos aufzubrechen

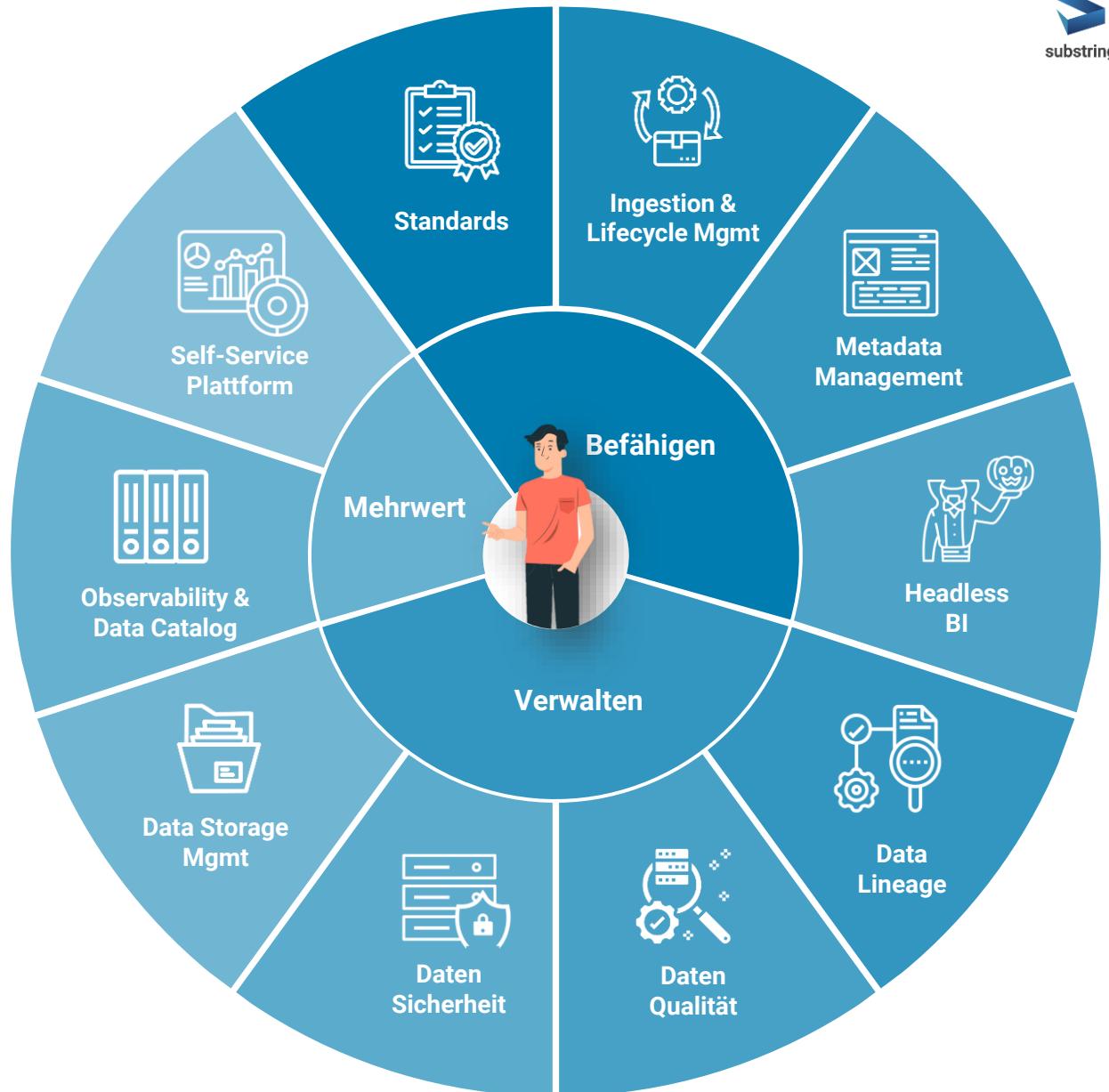


Kürzere Time-to-Insight



Technologisch flexibel und adaptierfähig

# Unsere 10 Komponenten im Überblick



Befähigen

# Datenstandards



Befähigen

# Data Ingestion Management

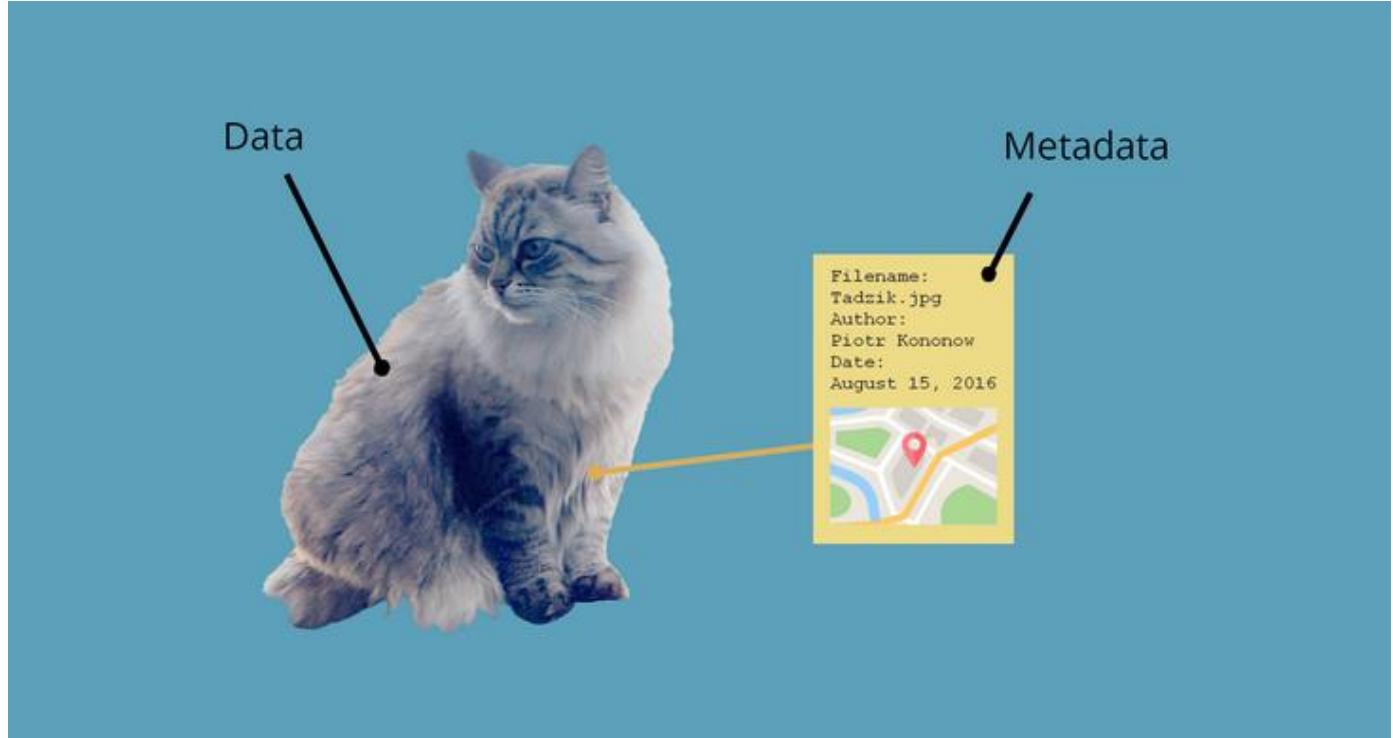


# The Data Lifecycle



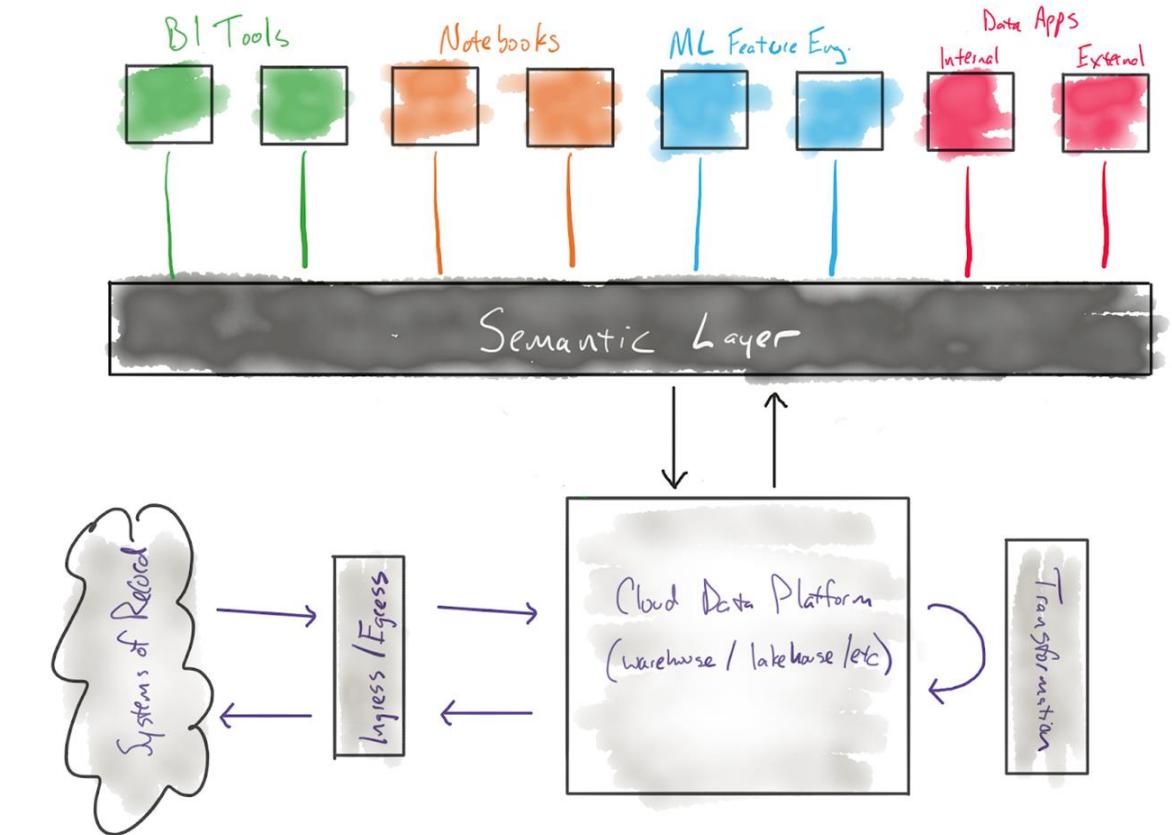
Befähigen

# Metadaten Management

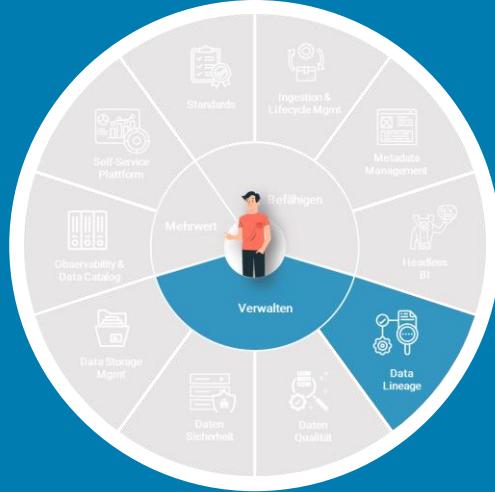


Befähigen

# Headless BI / Semantic Layer

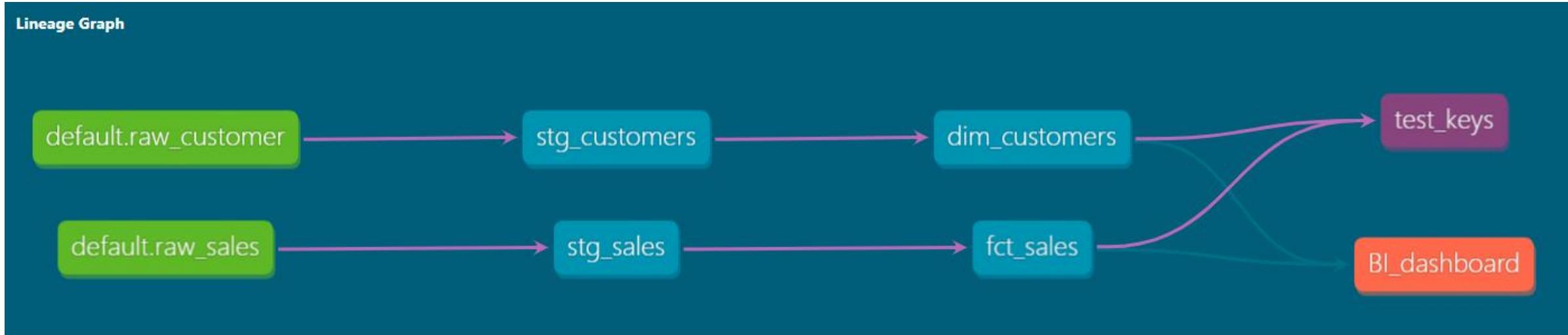


<https://www.getdbt.com/blog/next-layer-of-the-modern-data-stack/>



Verwaltung

# Data Lineage



Verwaltung

# Datenqualität



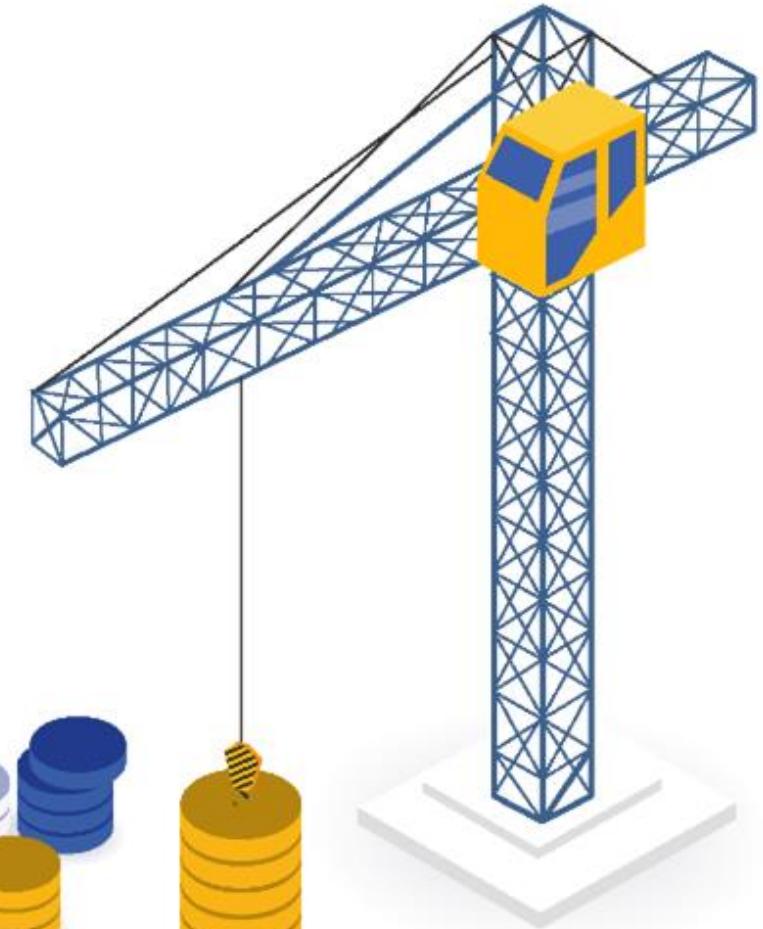
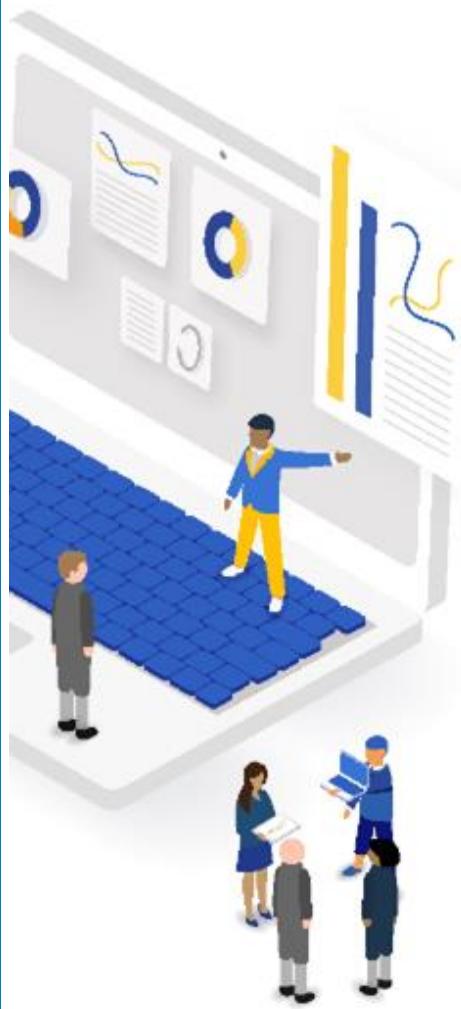
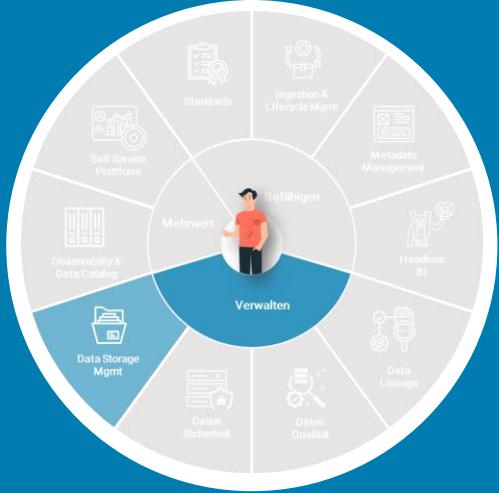
Verwaltung

# Datensicherheit & Datenschutz



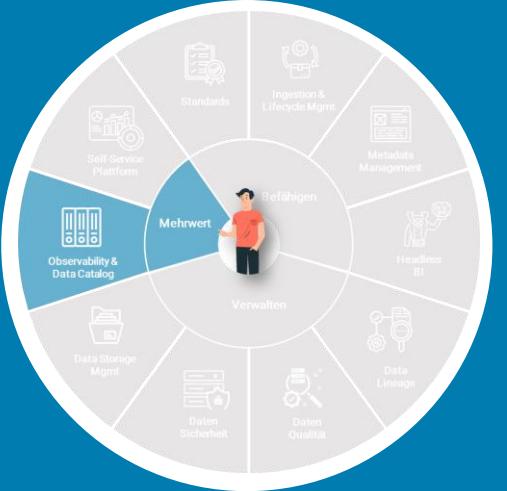
Verwaltung

# Data Storage Management



Business Mehrwert

# Data Observability & Data Catalog



AMUNDSEN

test\_schema.test\_table1 ★

Datasets • Hive • gold Beta

Description

1st test table

Last Updated

Oct 04, 2019 7pm EDT

Date Range

From: Apr 22, 2017 To: Sep 30, 2019

Owners

roald.amundsen@example.org, chrisc@example.org

Frequent Users

B, C, H, B, G

Tags

tag1, tag2, expensive, low\_quality

Quality Service

Quality Report:

Ipsus enim. Ipsus enim ipsum loremum.

build unknown

S3 Crawler

Size: 50T

Columns (5)

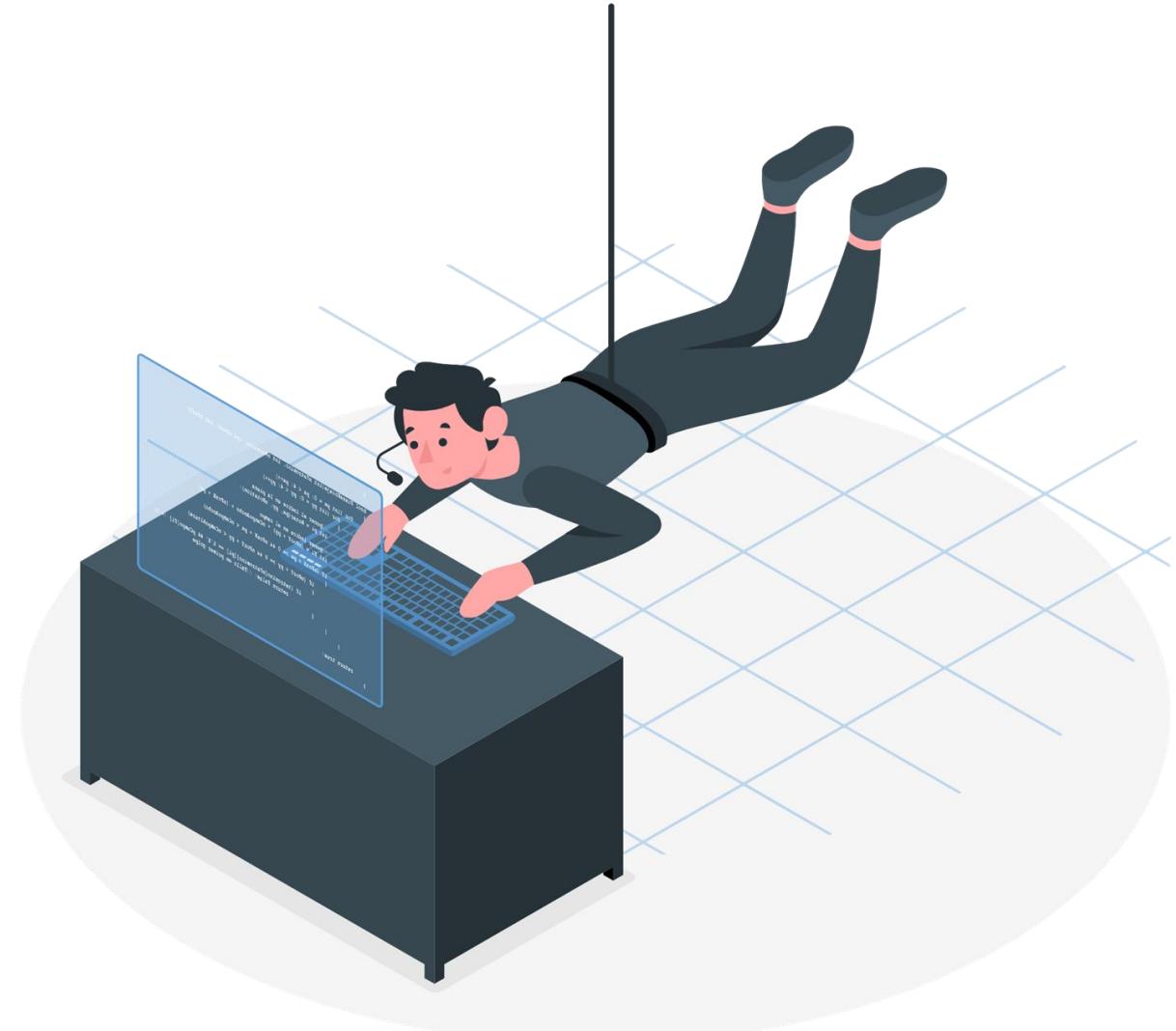
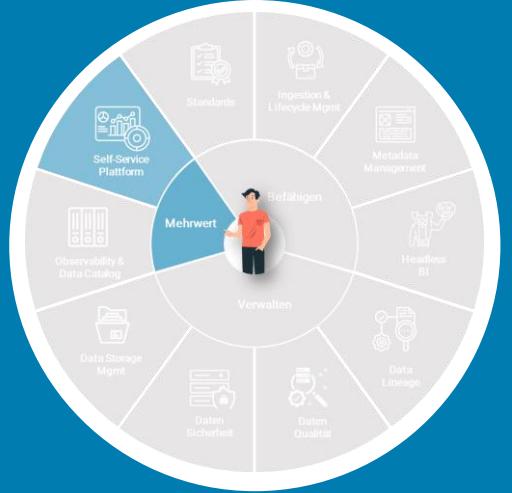
NAME	TYPE	BADGES
col1	string	Pk
col2	string	PII
col3	string	PII, Fk
col4	string	
col5	float	

Sort by ▾

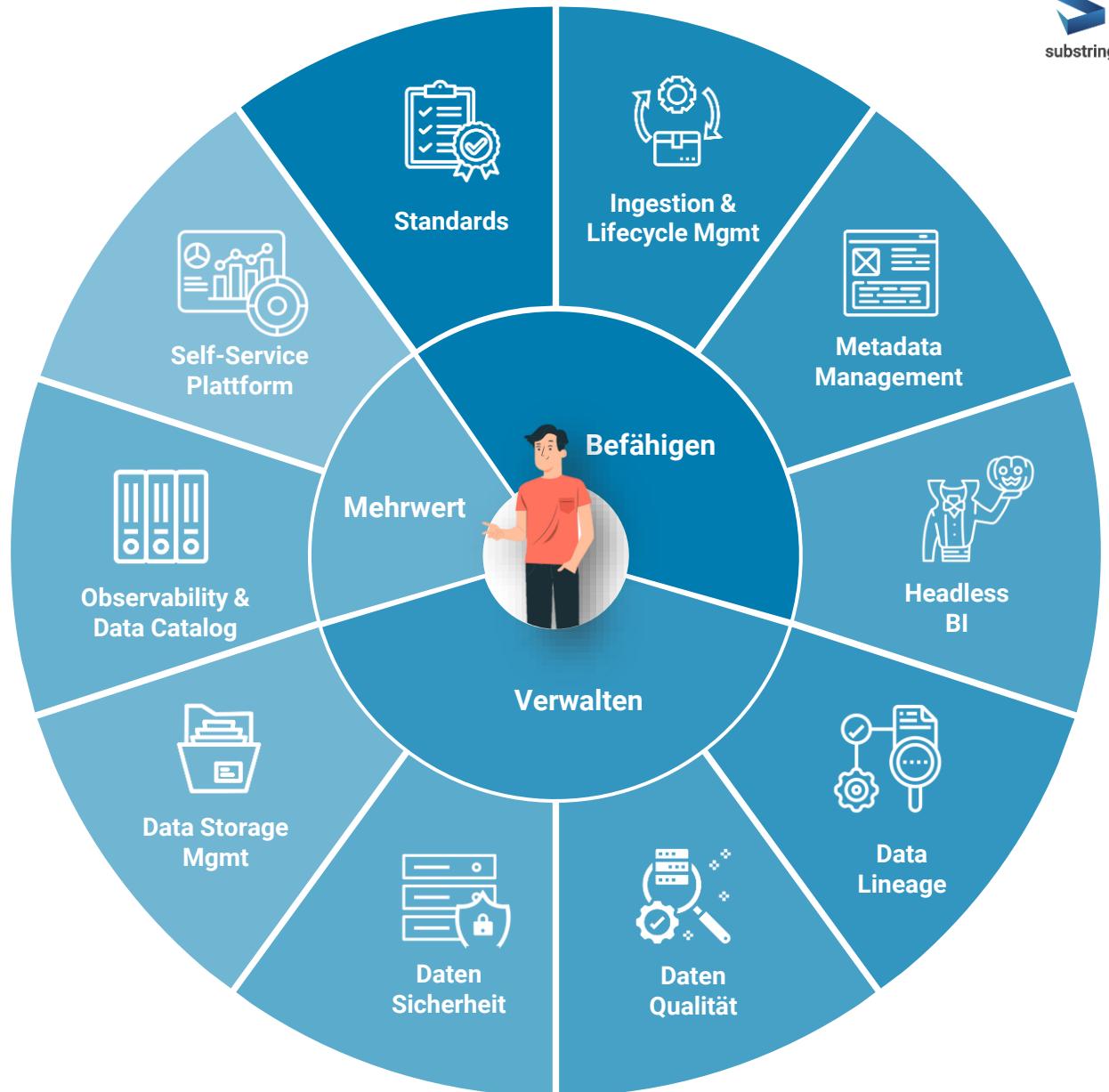
Amundsen was last indexed on May 20th 2021 at 3:40:45 pm

Business Mehrwert

# Self-Service Plattform



# Unsere 10 Komponenten im Überblick



# The Data Company.

## Services Technical

- Data Science & Machine Learning
- Business Intelligence & Data Discovery
- Data Lake & Data Warehouse
- DataOps & DevOps
- APIs & Microservices für Datenappl.

## Services Business

- Datenlandkarte
- Daten Strategien
- Business Analysen
- Datenarchitektur Design
- Schulungen zu Datenkompetenzen

## Aus-/Weiterbildung

- CAS Digitale Organisation (BFH)
- CAS Data-Driven Organization (BFH)
- CAS New Quality Management (SAQ/HWZ)

## Team

- 10 Mitarbeitende mit Leidenschaft für Daten
- Interdisziplinäres Team (Didaktik, Ökonomie, Statistik, SW-Entwicklung)
- + 20 Jahre Erfahrung in der Entwicklung von digitalen Data Lösungen

## Zufriedene Kunden



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra



HAUTE ÉCOLE DE GESTION  
HOCHSCHULE FÜR WIRTSCHAFT  
SCHOOL OF MANAGEMENT

Fribourg  
Freiburg



# Nächster Dataholics Termin

**Die Datenlandkarte als nützliches Werkzeug zur Datenorientierung.**

Wann? 28.03.2023

Uhrzeit? 11:15 Uhr - 12:00 Uhr

Anmeldung: <https://share.hsforms.com/19fr9ebzwSEeY2REOpXSbeQ5mdeq>



# Substring - The Data Company

---

*This wasn't very data-driven of you!*