

# Substring - The Data Company

---

## Erfolgsfaktoren und Kosten beim Aufbau eines MVP DWHs

Dataholics #12 Webinar-Reihe



**Sebastian Lienert**  
Data Expert  
@ Substring

Teilnehmen unter  
**`attend.sl/bfq9`**



**Ziel:**



***Was sind die Erfolgsfaktoren bei der Umsetzung eines MVP DWH für ein KMU?***



***Welche einmaligen und wiederkehrenden Kosten sind zu erwarten?***

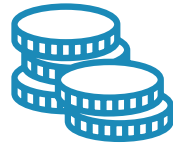
# Agenda



**Einleitung: Von Datensilos zum DWH**



**ELT, Datenarchitektur & BI:  
Konzept & Erfolgsfaktoren**



**Kosten: Aufbau und Betrieb**



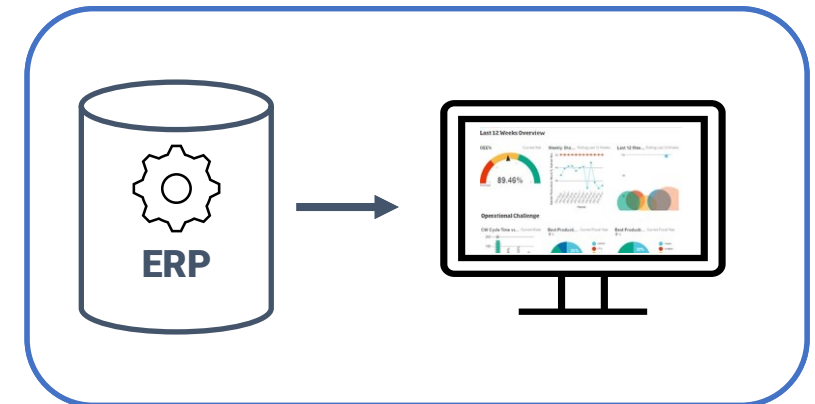
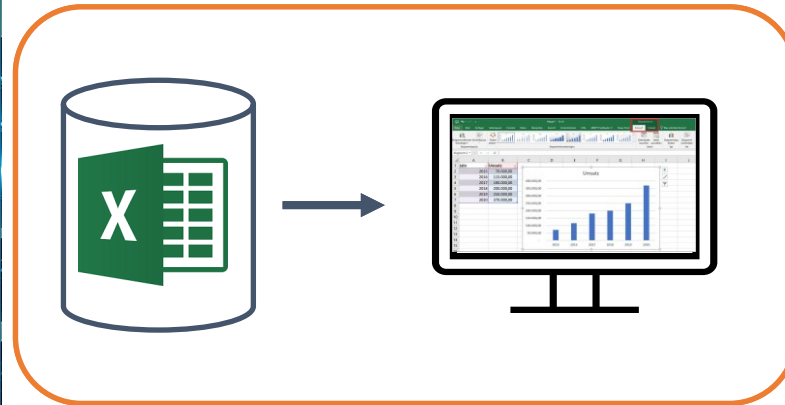
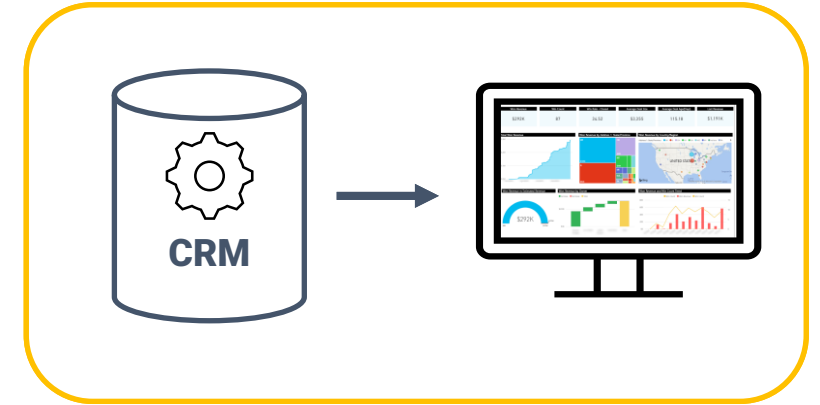
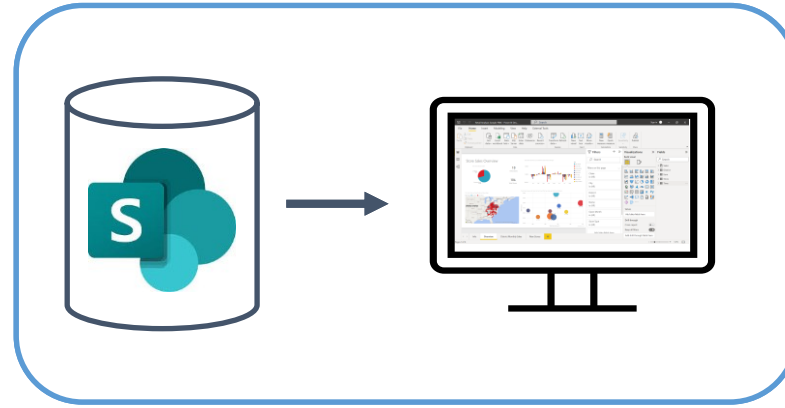
**Konkretes Kundenprojekt:  
Datenarchitektur & Betriebskosten  
SIPRO**



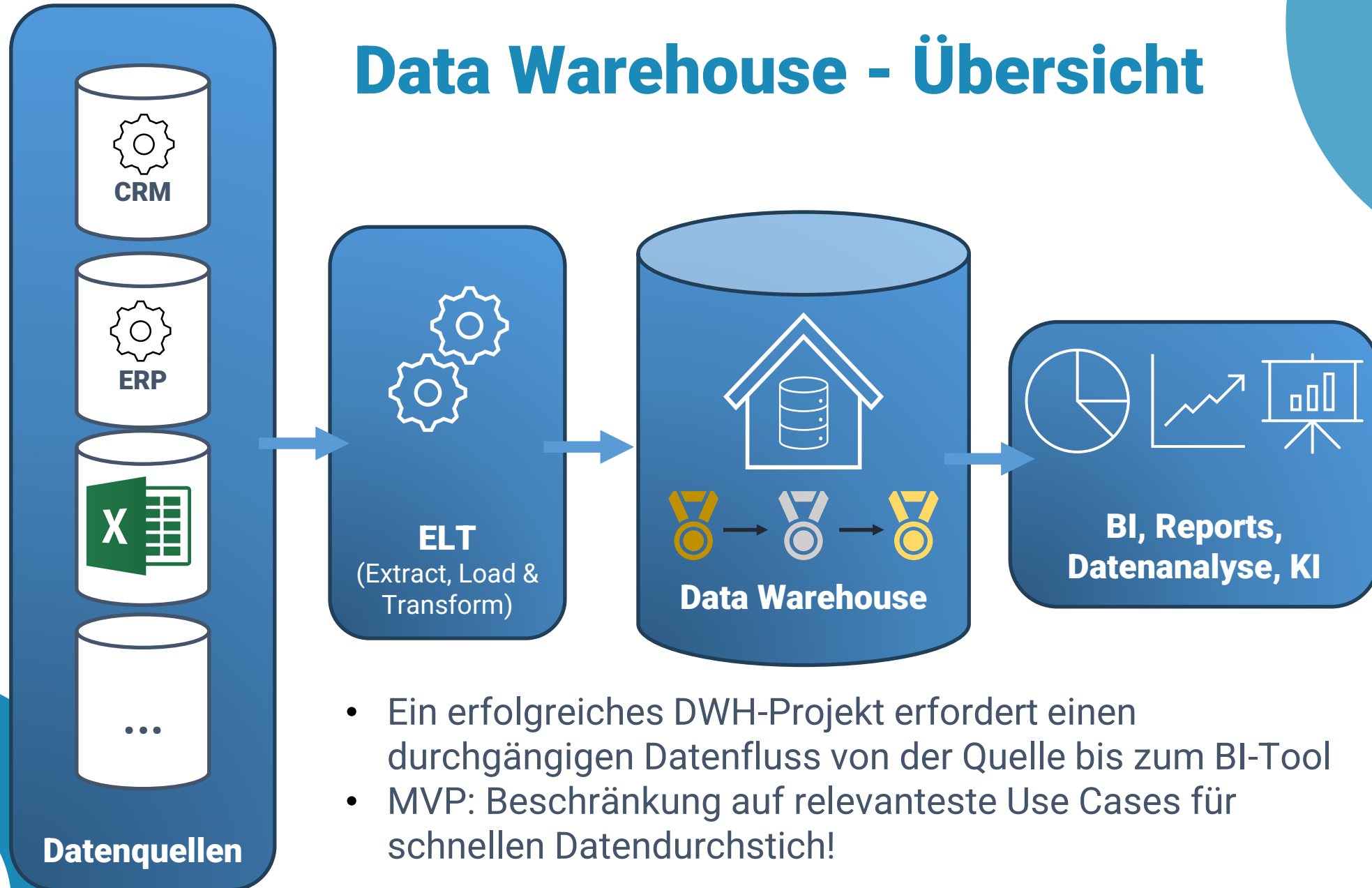
## Problematik

# Datensilos

- Eingeschränkter Informationsfluss
- Inkonsistente Daten
- Ineffizienz und Doppelarbeit
- Verlangsamte Entscheidungsfindung
- Lizenz- & Betriebskosten



# Data Warehouse - Übersicht

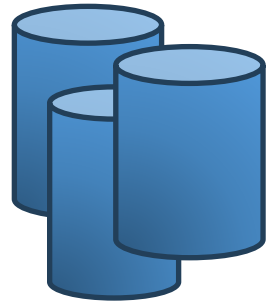


- Ein erfolgreiches DWH-Projekt erfordert einen durchgängigen Datenfluss von der Quelle bis zum BI-Tool
- MVP: Beschränkung auf relevanteste Use Cases für schnellen Datendurchstich!

# ELT Pipeline

## Extraktion

Abruf von Daten aus verschiedenen Quellen



## Laden

Laden von Rohdaten ins DWH



## Transformation

Bereinigung, Harmonisierung und Transformation von Daten



### Erfolgsfaktoren

- Skalierbarkeit
- Qualitätskontrolle & Fehlerhandling (Testing)
- Automatisierung & Reproduzierbarkeit
- **Datenarchitektur**



### Tools (Beispiele)



... und viele mehr



# Datenarchitektur



## Medaillon Schema



**Bronze**

*Rohdaten*



**Silber**

*Bereinigte Daten*



**Gold**

*Daten Bereit für BI*



**Gold Layer**

- Daten in Star/Snowflake Schema (Fakten & Dimensionen)
- Datamarts für spezifische Use Cases (optional)



**Erfolgsfaktoren**

- Architektur und Granularität stimmt mit vorgesehenen Anwendungen überein.
- Struktur der Daten ist verständlich für Endnutzer
- Datentransformationen mit Versionskontrolle, Tests, wohl definierter Data Lineage und Dokumentation





# BI Umsetzung



## Tools



... und viele mehr

+ Benutzerfreundlich

+ Open Source

+ Komplexe Visualisierungen



## Datenquelle

- Aufbereitete Daten aus Gold Layer
- Direkter Import von Star/Snowflake Schema
- ... oder massgeschneiderte Data Marts



## Erfolgsfaktoren

- **Self-Service fördern:** Endbenutzer können am besten Beurteilen welche Reports/Dashboards nützlich sind
- **Benutzerfreundlichkeit:** Sowohl bei Konsum als auch Erstellen von Inhalten hilft sicherzustellen das die BI Tools genutzt werden

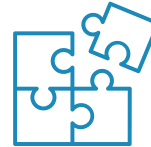


Generelle

# Erfolgsfaktoren



MVP: Reduktion der Komplexität und Fokussierung auf die wichtigsten Use Cases – Synergie: Reevaluation



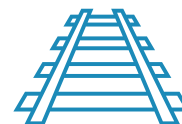
Architektur passend zu Geschäftsanforderungen.



Datenarchitektur und BI-Tool sollen Self-Service fördern



Schneller Datendurchstich zum BI-Tool (Allenfalls Begrenzung auf wichtigste Use Cases & Datenquellen)



Agile Weiterentwicklung nach MVP



# Kosten



## Substring: Kostenrechner für Data Warehouse MVP

WIP: Dieses Tool hilft Ihnen, die Entwicklungskosten und monatlichen Cloud-Kosten für ein Data Warehouse Minimum Viable Product (MVP) zu schätzen.

### Konfigurationsoptionen

Anzahl der Schnittstellen zu Datenquellen



Komplexität der Datentransformationen

Mittel

Ungefähres Datenvolumen in GB

5

Anzahl der Dashboards für den MVP

0

Datenaktualisierungsfrequenz

Täglich

Kosten berechnen

### Kostenschätzungen

Geschätzte MVP-Entwicklungskosten

# Preview: Kostenrechner für MVP DWH



Erkenntnis: Viele KMUs haben keine klare Vorstellung über die Kosten eines DWHs.



Unsere Lösung: Ein interaktiver Kostenrechner für eine schnelle und grobe Kostenschätzung.









Derzeit in Entwicklung: Wir informieren über den Release!



# Kosten für den Aufbau und Betrieb eines MVP DWH

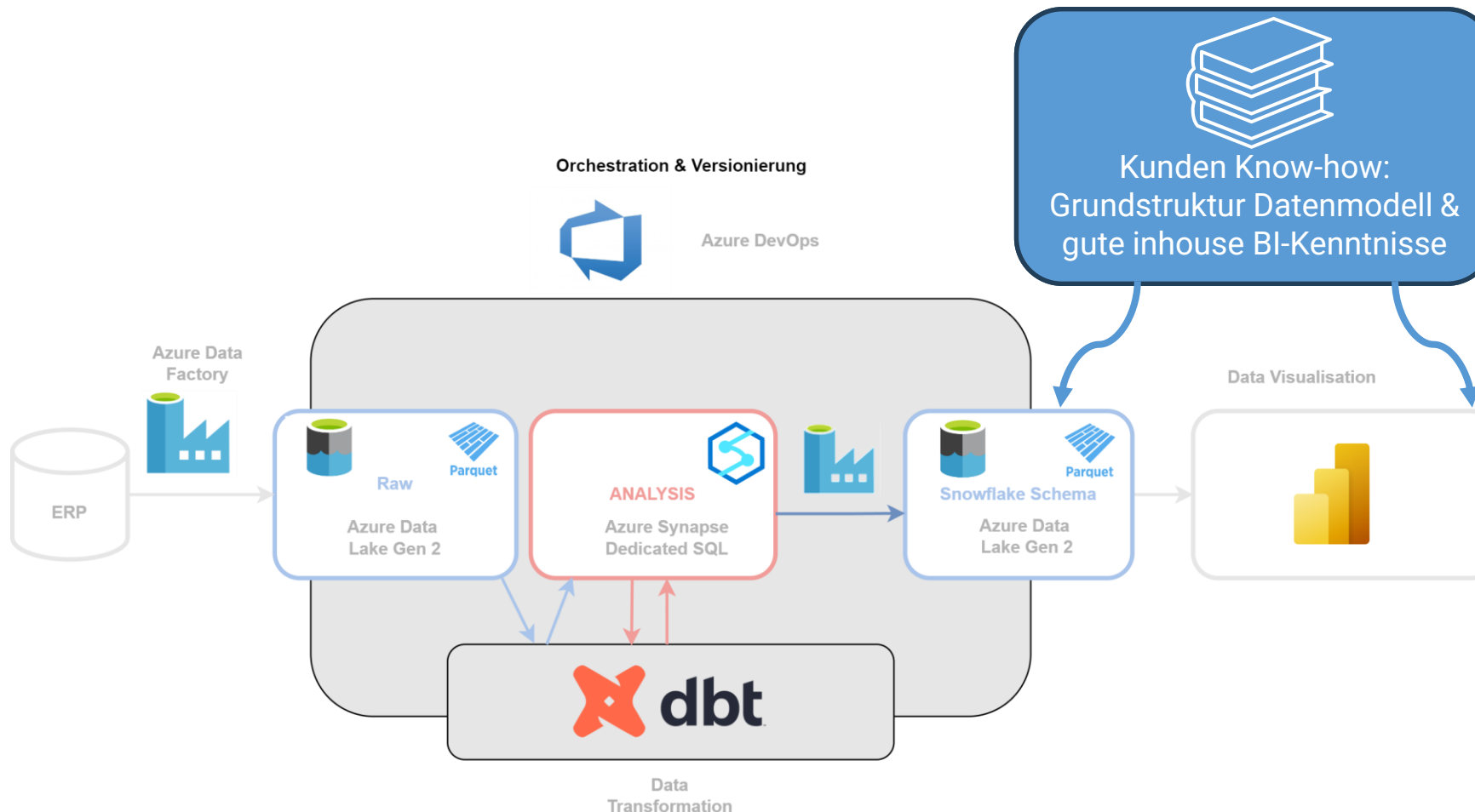
- Unverbindliche Angaben
- Orientierungshilfe
- Fall KMU

		 <b>Aufgaben</b>	 <b>Kostenfaktoren</b>	 <b>~ Kosten</b>
	<b>Aufbau MVP</b>	Konzept, Konfiguration Infrastruktur, Implementierung (ELT, Anbindung BI, erste Berichte, Anbindung erste Schnittstellen, ...)	Datenmenge & Komplexität, ELT Toolchain, Arbeitsaufteilung und Zusammenarbeit Kunde, ...	40'000 – 160'000 CHF
	<b>Betriebskosten</b>	(Cloud) Infrastruktur, Wartungsaufwand	Architektur & Provider, Aktualisierungsfrequenz, Datenmenge, Wartung, ...	200 – 2000 CHF / Monat
	<b>Lizenzkosten (BI)</b>	Kommerzielle Plattform: Lizenzkosten	BI Tool, # Nutzer, Existierende Lizenzen (z.B. Microsoft E5 -> PowerBI)	0-50 CHF / User/ Monat
		OpenSource Plattform: Infrastrukturkosten	Performance Bedürfnisse	100 – 1000 CHF / Monat

# Kundenprojekt: Datenarchitektur & Betriebskosten



**SIPRO**<sup>®</sup>  
STEEL SOLUTIONS



## Optimierte Betriebskosten:

- ADF: Pay-Per-Use
- Synapse Dedicated SQL: (~2 CHF/hr), aber limitierte Uptime!

→ ~ 200 CHF / Monat

## Überschaubare Lizenzkosten

- PowerBI Pro:

10 CHF / Benutzer / Monat

# Zusammenfassung



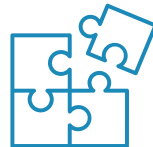
## Schnelle Wertschöpfung durch MVP

*mit einem MVP schnell Mehrwert generieren und aufzeigen, dass die gewählte Architektur alle Use Cases unterstützt*



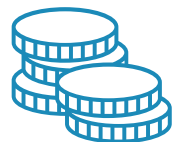
## Kunden-Know-how nutzen

*Integration des Kunden-Know-hows in das Projekt, z.B. für Self-Service BI oder Infrastrukturwartung*



## Massgeschneiderte Lösung

*Kein «One-Size fits all»*



## Kosten im Rahmen halten

*Durch gute Architekturentscheide können Kosten optimiert werden*

Substring AG

# The Data Company.

## Services Technical

- Data Science & Machine Learning
- Business Intelligence & Data Discovery
- Data Lake & Data Warehouse
- IoT, Smart Factory & Industrie 4.0
- DataOps & DevOps
- APIs & Microservices für Datenapplikationen

## Services Business

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

## Lehraufträge & Weiterbildung

- CAS Digitale Organisation (BFH)
- CAS Data-Driven Organization (BFH)
- CAS Customer Experience & Service Design (BFH)
- CAS Supply Chain & Logistik (ab 2024 BFH)

## Team

- 14 Mitarbeitende mit Leidenschaft für Daten & KI
- Interdisziplinäres Team (Didaktik, Maschinenbau, Physik, Ökonomie, Statistik, SW-Entwicklung)
- + 20 Jahre Erfahrung in der Entwicklung von Daten Lösungen

## Auszug unserer Kunden



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Direktion für Entwicklung  
und Zusammenarbeit DEZA



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Eidgenössisches Departement für  
Wirtschaft, Bildung und Forschung WBF  
Staatssekretariat für Bildung,  
Forschung und Innovation SBF



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra



Schweizerisches Rotes Kreuz  
Croix-Rouge suisse  
Kanton Bern - Canton de Berne



BLUTSPENDE SRK SCHWEIZ



**Stöcklin**  
Home of Intralogistics



**GEBERIT**



die Mobiliar



*u<sup>b</sup>*

UNIVERSITÄT  
BERN



HAUTE ÉCOLE DE GESTION  
HOCHSCHULE FÜR WIRTSCHAFT  
SCHOOL OF MANAGEMENT

Fribourg  
Freiburg



Berner  
Fachhochschule



biomondo

**INSELSPITAL**

UNIVERSITÄTSSPITAL BERN  
HÔPITAL UNIVERSITAIRE DE BERNE



Kanton Bern  
Canton de Berne



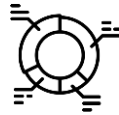


Substring

# Ihr Data Partner



**Langjährige  
Erfahrung und  
breites Know-How**



**Ganzheitliche,  
langfristige und  
partnerschaftliche  
Zusammenarbeit**



**Vielseitige  
Kompetenzen**



**Offenheit,  
Transparenz und  
Ehrlichkeit**



**Nachhaltige  
Verbesserung von in  
Richtung Data-Driven**



**Agil, kundenorientiert  
& flexibel**

Substring

# Nächster Dataholics Termin

**Erfahrungsbericht: Warum man sich für den «Data-Driven Weg» entscheidet und was Learnings daraus sind?**

Wann: Dienstag, 22. Oktober 2024

Uhrzeit: 11:15 Uhr - 12:00 Uhr

Anmeldung: [https://share.hsforms.com/1NVSKJK6STMW2D\\_RvOCVew5mdeq](https://share.hsforms.com/1NVSKJK6STMW2D_RvOCVew5mdeq)







# substring - the data company

---

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