



CubeServ®

# BAUSTEINE FÜR IHRE ANALYTICS ARCHITEKTUR

18. CUBESERV CONGRESS

ZÜRICH, 10. MAI 2022





# **DATA INSIGHTS GEWINNEN UND INTEGRIEREN: "R FÜR FREESTYLE-USER" ARCHITEKTURPRINZIPIEN IN DER AWS**

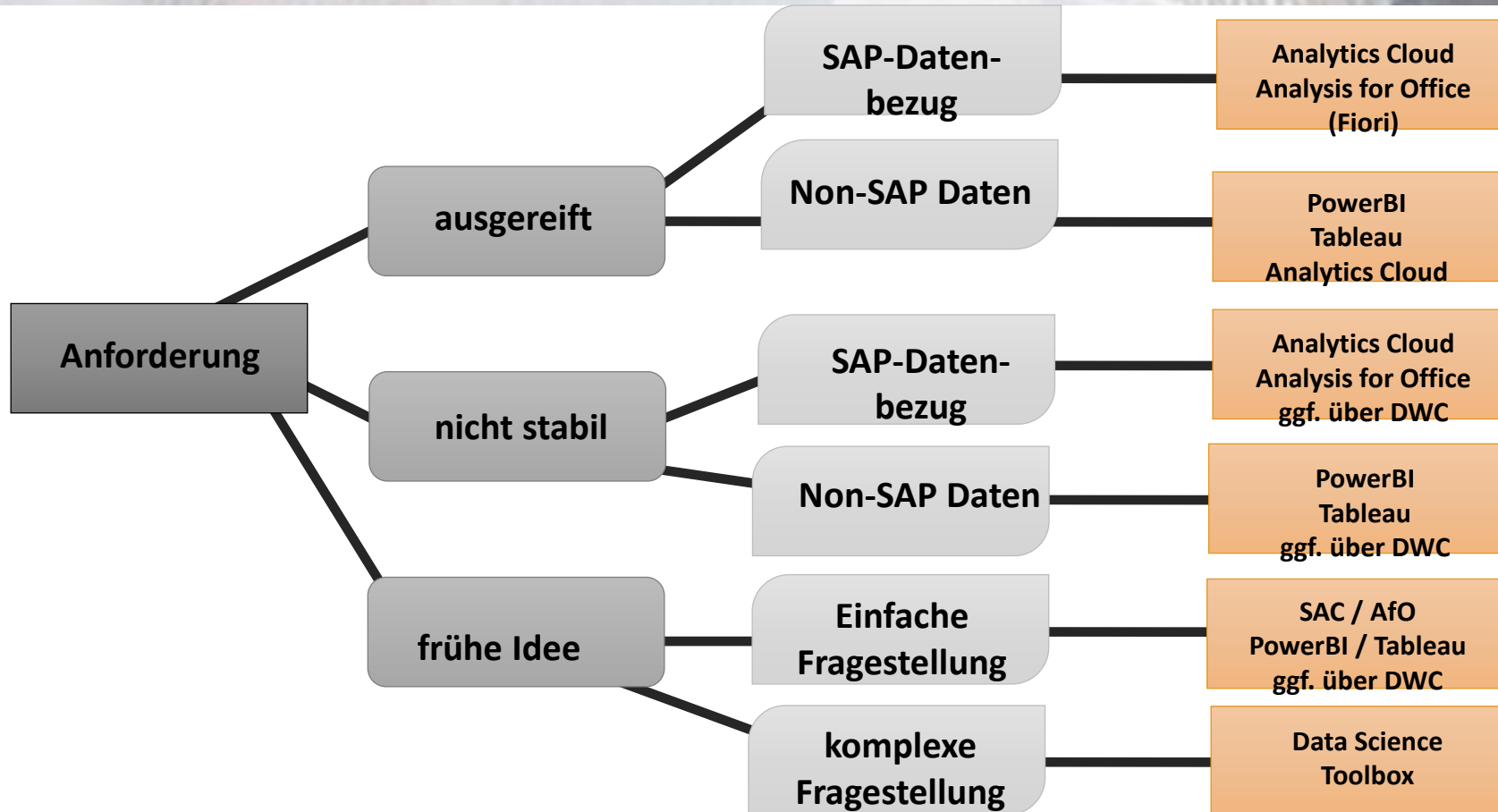
**Thorsten Stoßmeister | Adrian Bourcevet**

# Data Insights

## Statt APD neue Herausforderungen

- Das BW hatte zur Überraschung aller, auch zur Überraschung des BWs selbst, mit dem APD ein einfaches Tool, das simplen Self-Service erlaubte.
- Der APD ist in BW/4HANA nicht mehr vorhanden. Gleichzeitig ergeben sich neue Herausforderungen:
  - Wie kann man grosse Datenmengen analysieren, ohne auf die zentrale IT zu warten und ohne das System teuer zu vergrössern?
  - Wie können Services wie AWS S3 (Simple Storage Service für theoretisch beliebig grosse Datenmengen) und Athena (SQL-Abfrageservice) genutzt werden?
  - Wie können R-Algorithmen der weltweiten Community genutzt und an der Community teilgenommen werden?
  - Wie kann Self-Service angeboten werden?
  - Wie können die Ideen von Mitarbeitern genutzt werden, die auch mal was ausprobieren möchten?

## Entscheidungsbaum nach Anforderung



# Data Insights

## Ein Business Case

- Folgendes Business Szenario angenommen:
  - Aufgrund des Ukraine-Krieges sind Standorte ausgefallen. Ein neuer Standort wird gesucht.
  - Wegen immer noch niedriger Kosten kommt die grösste Demokratie der Welt, Indien, in Betracht.
  - Bei der Suche nach Daten über Standorte fällt ein Datensatz auf:

<https://www.kaggle.com/code/prattusha/r-code-protests-in-india-from-2016-2022>

Es sind mehr als 100k Daten über Proteste und Unruhen in Indien, die für die Standortsuche von Relevanz sein könnten.

# Data Insights

## 3 Use Cases

- Use Case 1:

Der Indien-Datensatz soll ad-hoc in einer Data Science-Toolbox ausgewertet werden. Es soll analysiert werden, welcher Art die Daten sind welche Analysen bereits von der Community vorgenommen wurden.

Zunächst ist aber nicht sicher, ob Daten, erstelltes Coding und benötigte Software später weiterverwendet oder weggeworfen werden. Weiterverwendung soll aber potenziell möglich sein.

Die Analyse soll daher zunächst in einer separaten Umgebung erfolgen, vorinstallierte benötigte Software wäre aber hilfreich.

# Data Insights

## 3 Use Cases

- Use Case 2:

Die Daten erscheinen nützlich, eine weitere Verarbeitung dieser oder abgeleiteter Daten in der DWC soll möglich gemacht werden (z.B. für eine Geo-Visualisierung der Daten).

- Use Case 3:

Die Daten werden mit SAP-Daten verknüpft (wie z.B. Standorte von Kunden oder Lieferanten).

# Data Insights

## Demo Use Case 1

- Eine Instanz mit vorinstallierter Software (AirLiveDrive für S3, R-Studio, gitHub).
- User erhält eine IP-Adresse für die RDP-Anmeldung.
- Weitere Software-Installationen lokal möglich.
- Daten im S3-Bucket gespeichert, als Laufwerk D: verfügbar (man kann auch ganz normal eigene Dateien anlegen und editieren).
- R-Analysen können gefahren werden.
- Coding kann per gitBash genutzt oder zur Verfügung gestellt werden.



# Data Insights

## Demo Use Case 2

- Das AWS Glue kann automatisch die Metadaten von Dateien ermitteln, sozusagen die Tabellenstruktur.
- Der Athena Service kann dann SQL-Abfragen gegen diese Tabelle (=File) ausführen.
- Genauso kann diese Tabelle in der DWC verfügbar gemacht werden, z.B. für einen SQL-View.

# Data Insights

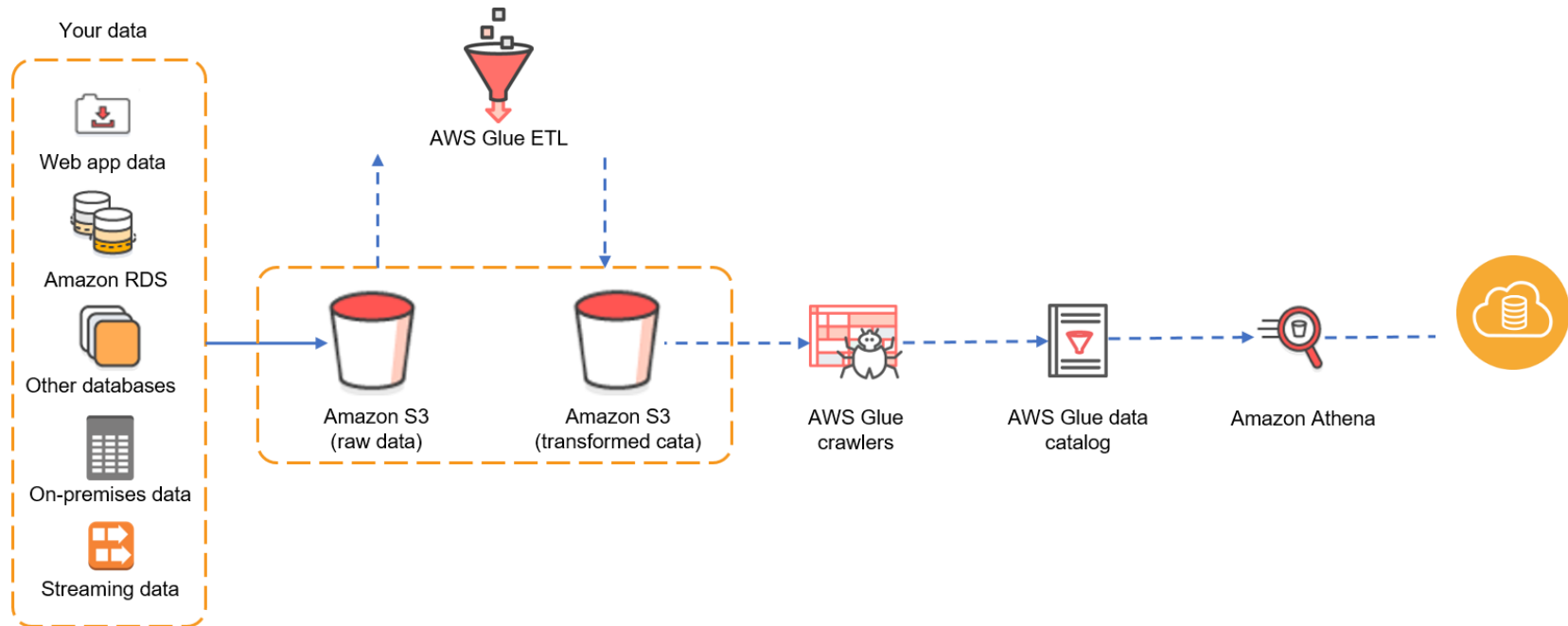
## Use Case 3 (Ohne Demo)

- SAP Daten können im Datalake strukturiert per SAP Data Intelligence zur Verfügung gestellt werden.
- SAP plant eine stärkere Zusammenführung von DWC und DI.

# Data Insights

## Fazit

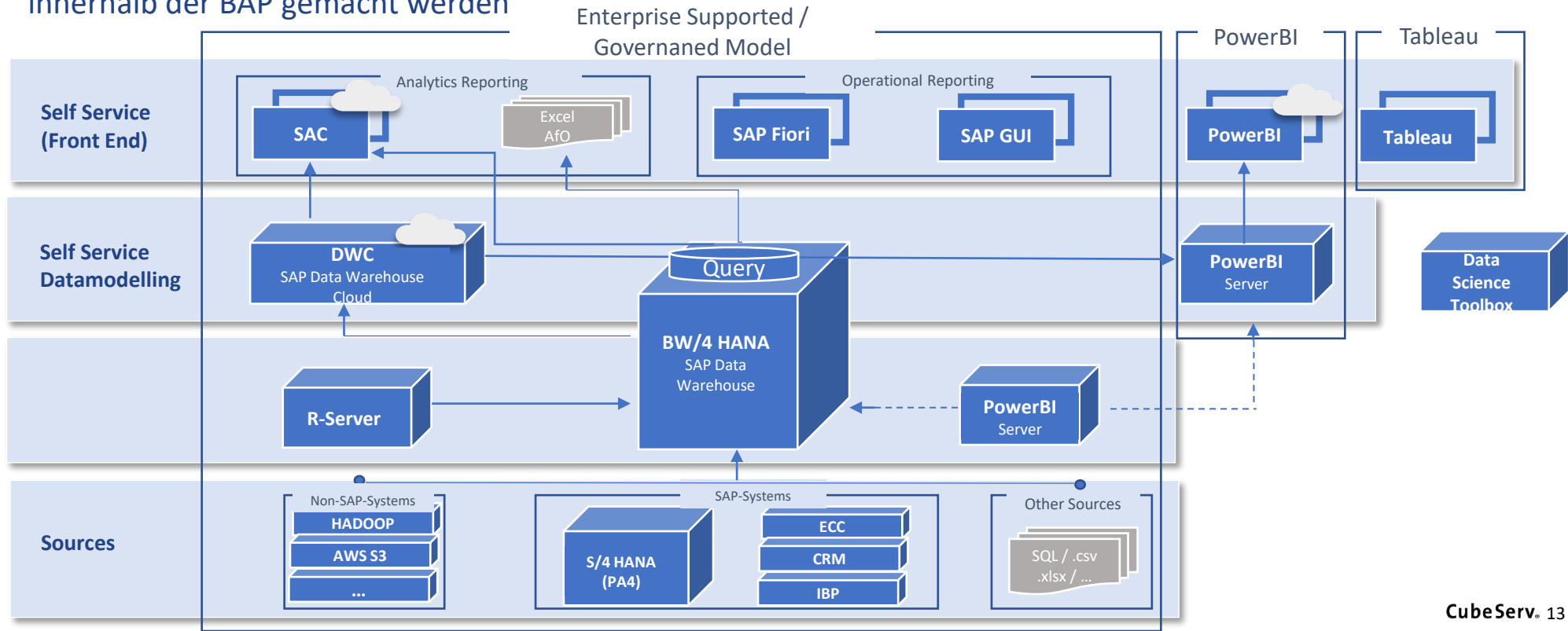
- Die Integration von AWS Services in die analytische Plattform ist mit geringem Aufwand möglich.



# Data Insights

## Fazit

- Mit der Data Science Toolbox kann ein flexibles, unabhängiges, aber auch wiederverwendbares Angebot innerhalb der BAP gemacht werden



Vielen Dank

Für Ihre Aufmerksamkeit



Ihr Ansprechpartner

**Adrian Bourcevet**

Managing Director / Chairman

T +41 79 691 97 33

adrian.bourcevet@cubeserv.com



Ihr Ansprechpartner

**Dr. Thorsten Stoßmeister**

Competency Lead EDW / HANA

T +41 79 175 62 72

t.stossmeister@cubeserv.com

# CubeServ Group

## Kontakte

### **CubeServ AG (Head Office)**

Neue Jonastrasse 52  
CH-8640 Rapperswil  
T +41 55 224 30 00

### **CubeServ GmbH**

Am Prime Parc 4  
D-65479 Raunheim  
T +49 6142 210 170

### **CubeServ GmbH**

Bösendorferstraße 3/Top 10  
A-1010 Wien  
T +43 1 28 68 800

### **CubeServ d.o.o.**

II. Vrandučka br. 8  
HR-10000 Zagreb  
T +385 1 2851 049





**©COPYRIGHT 2022 CUBESERV GROUP**

**ALL RIGHTS RESERVED.**

The information in this document is confidential and proprietary to CubeServ Group (hereinafter referred to as CubeServ) and may not be disclosed without the permission of CubeServ. No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of CubeServ. The information contained herein may be changed without prior notice.

Some software products marketed by CubeServ and its distributors contain proprietary software components of other software vendors. All other product and service names mentioned are the trademarks of their respective companies. National product specifications may vary.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. Please see <https://www.sap.com/corporate/en/legal/trademark.sap-trademark-list.html#sap-trademark-list> for additional trademark information and notices.

These materials are provided by CubeServ for informational purposes only and may not be incorporated into a contract.

The materials are provided without representation or warranty of any kind, and CubeServ shall not be liable for errors or omissions with respect to the materials. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. The only warranties for CubeServ products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, CubeServ has no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and CubeServ's strategy and possible future developments, products and/or platform directions and functionality are all subject to change and may be changed at any time for any reason without notice. The information in this document is not a commitment, promise or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements which speak only as of their dates, and they should not be relied upon in making purchasing decisions.