ELCA We make it work.

17.10.2024

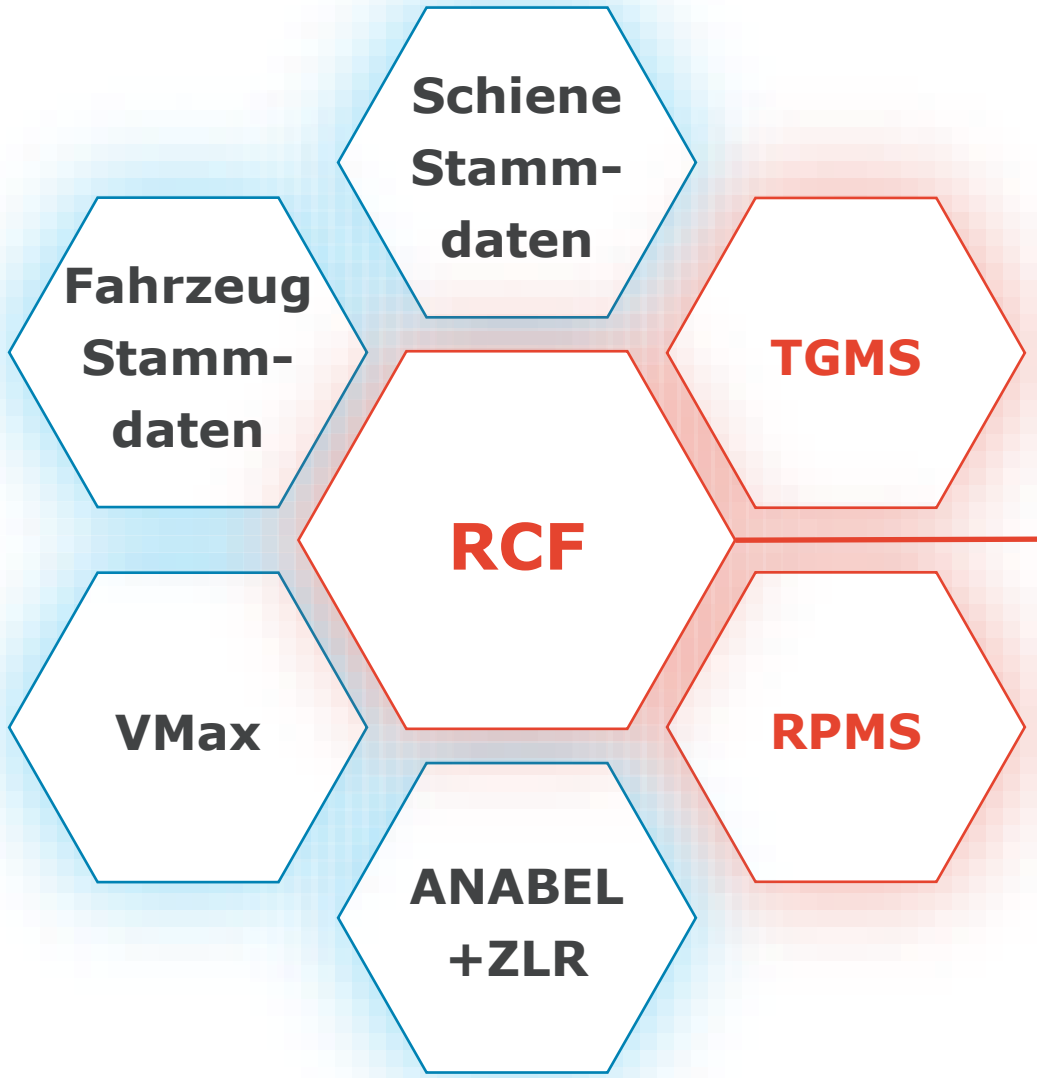
TU Graz

Komplexität: wie man eine Daten-Schlacht gewinnt

Von steter Innovation zu stabiler Produktion am Beispiel von Rollkontaktermüdung

ELCA We make it work.

Challenge: RCF-Prediction



complexity ⚠

Performance in memory:

20 Milliarden Datensätze/Tag

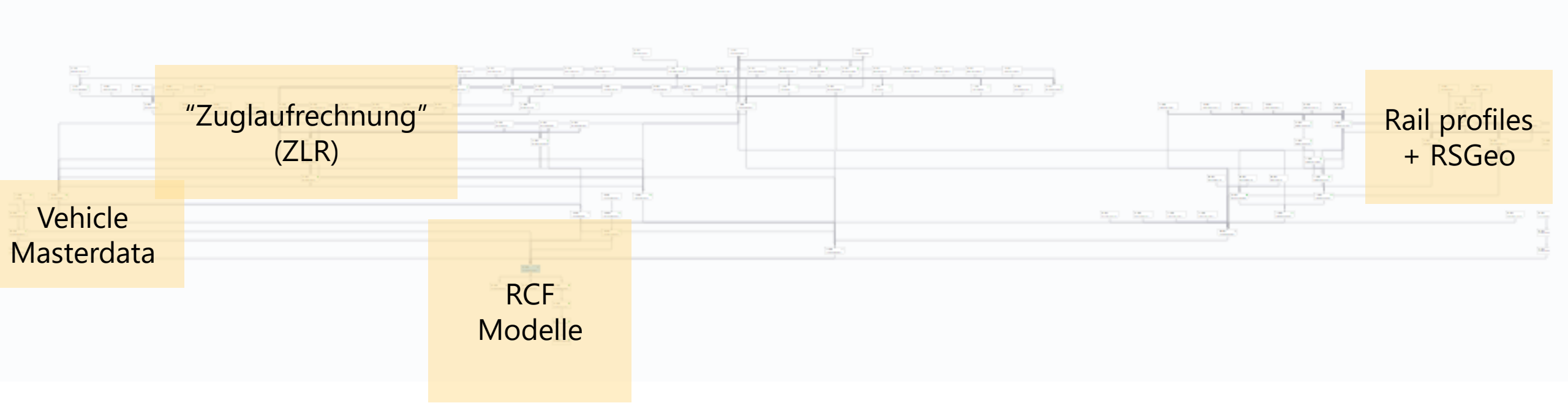
1.4TB Daten/Tag

120 Cores -> 4h Rechenzeit

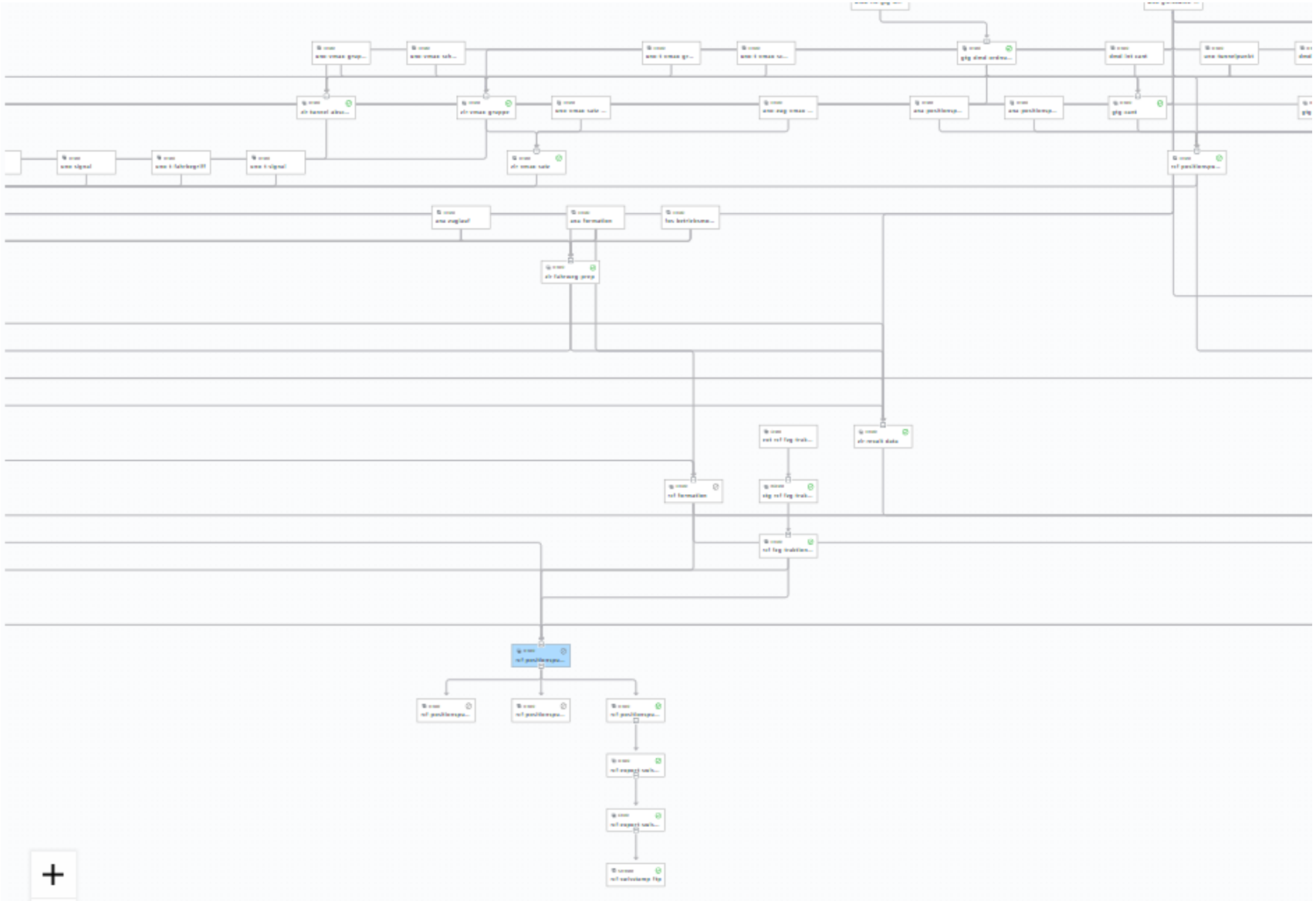
✓

data quality ⚠

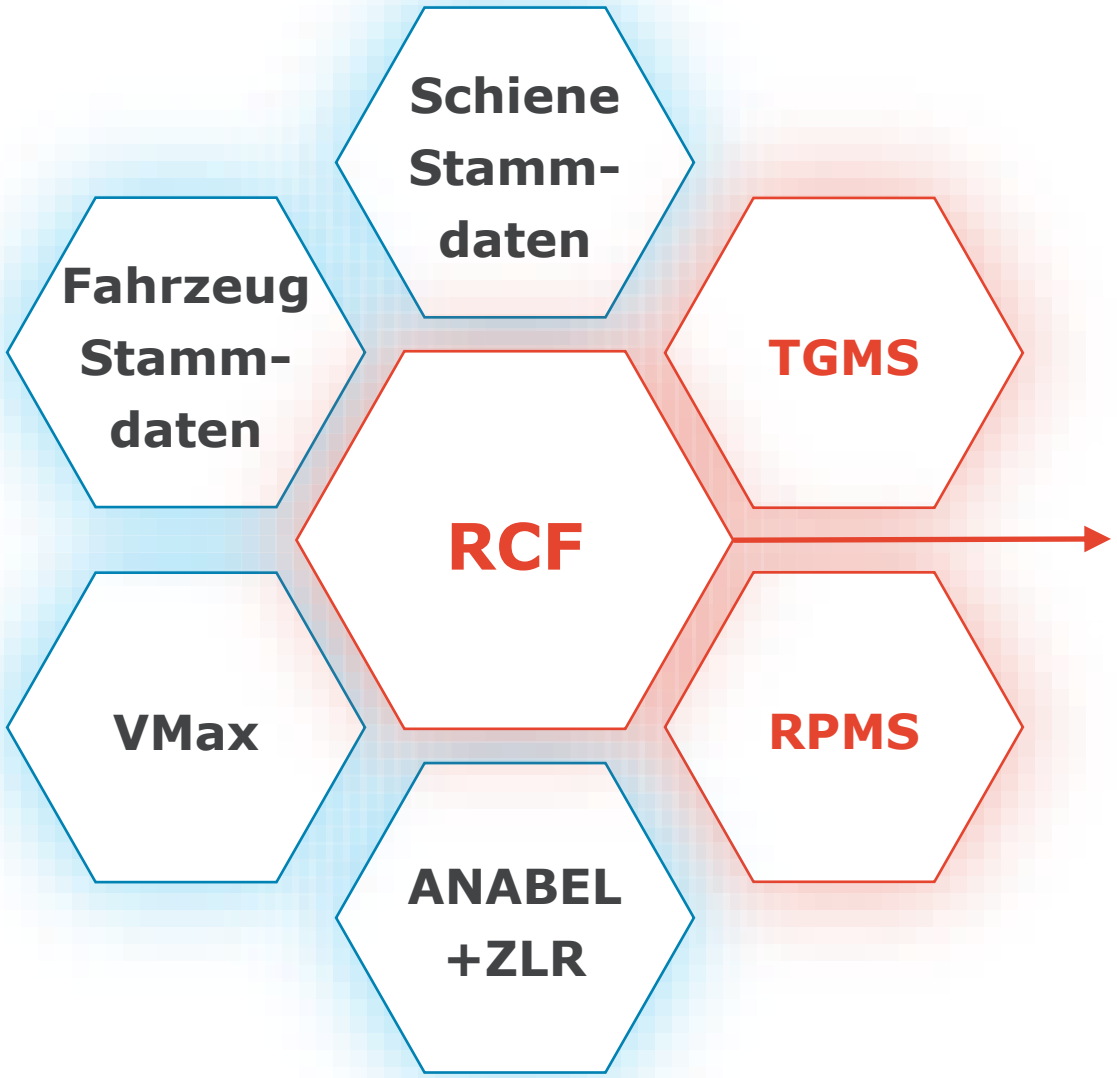
RCF-Prediction data flow -> complex



RCF-Prediction data flow -> complex



Challenge: RCF-Prediction «mit Durchdringung»



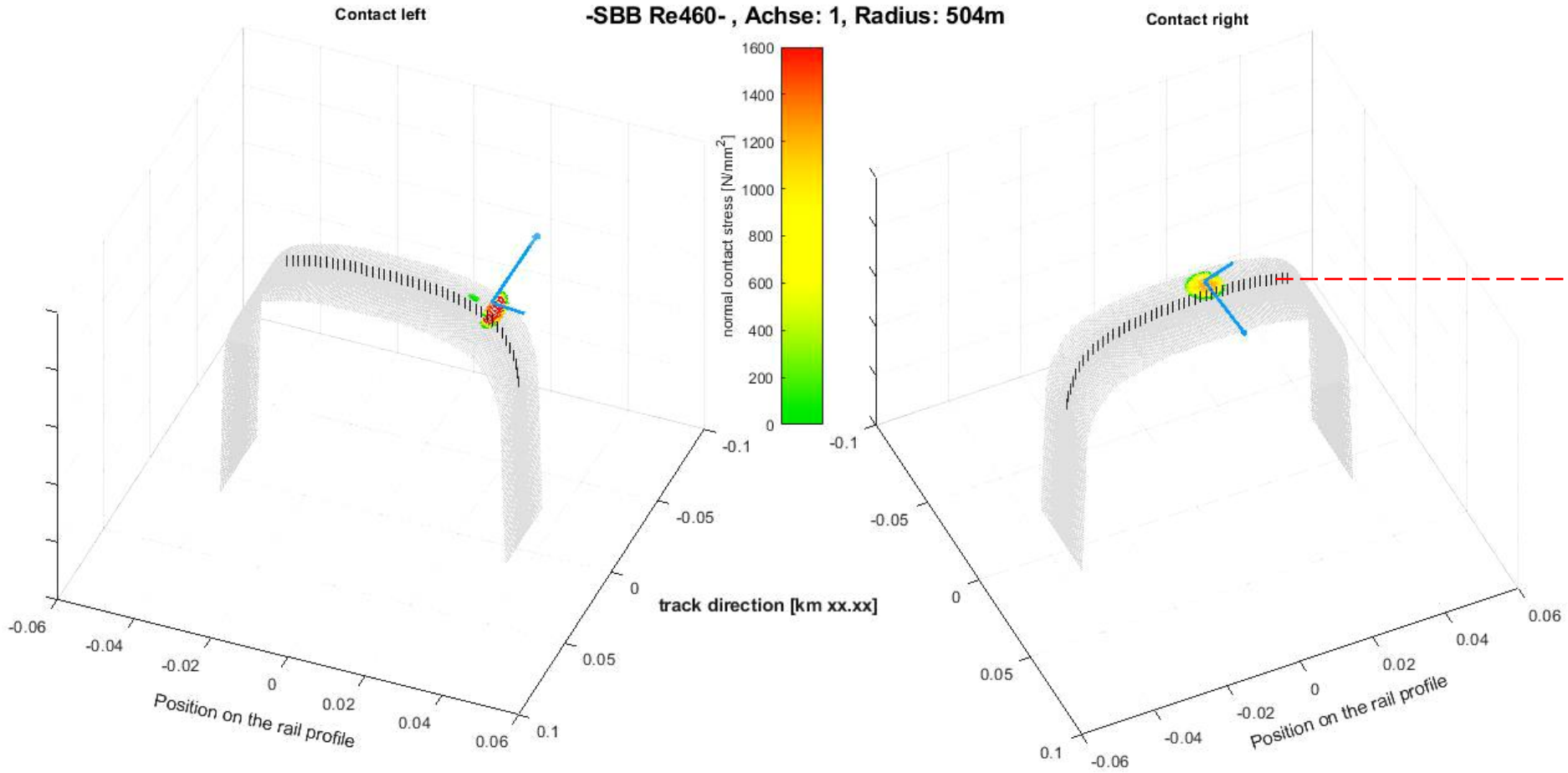
Performance on Disk:

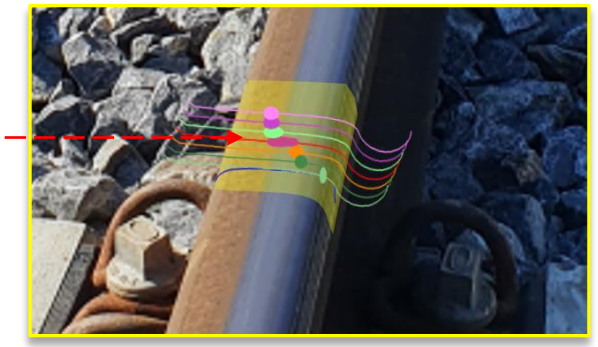
11Mio -> 1.4Mia Datensätze/Tag

40GB -> 215GB Daten/Tag

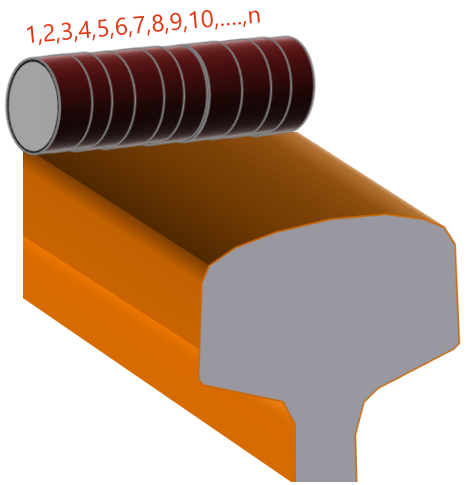
4h -> 13h Rechenzeit

6

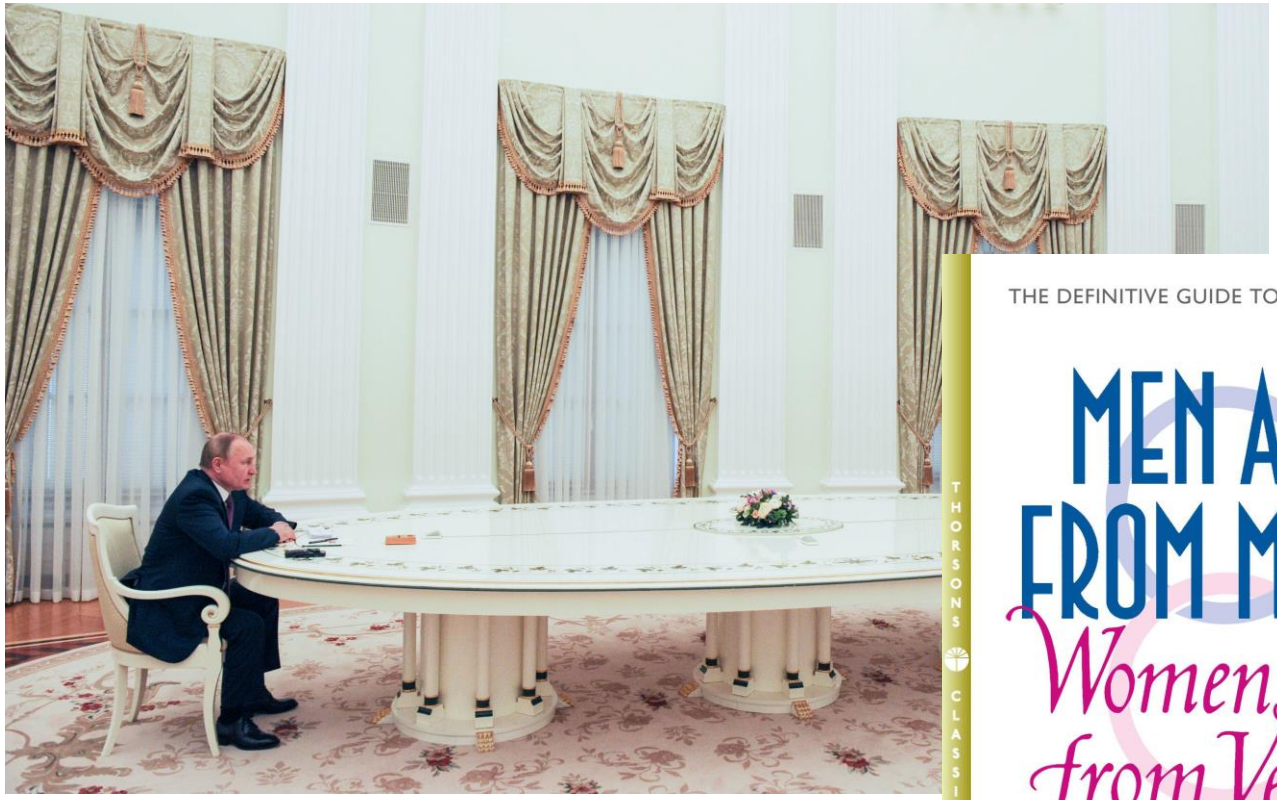


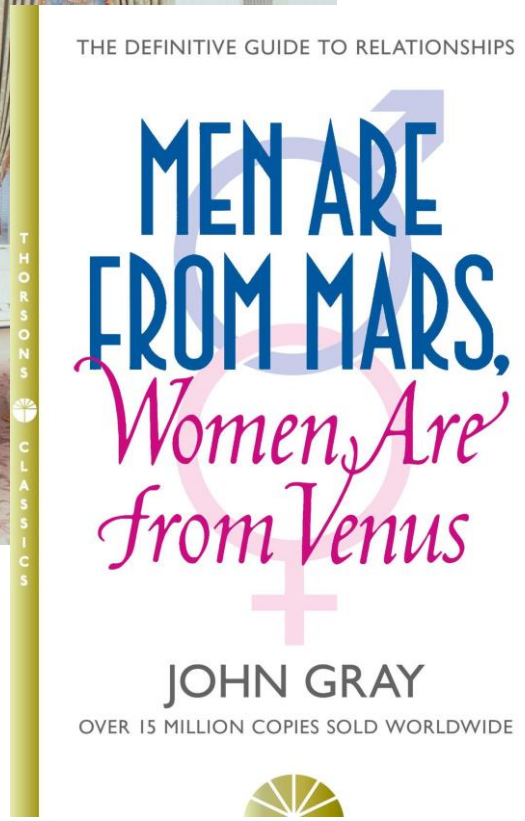


«Streifisierung» z.B. 1mm
= Multiwalzenkontakt



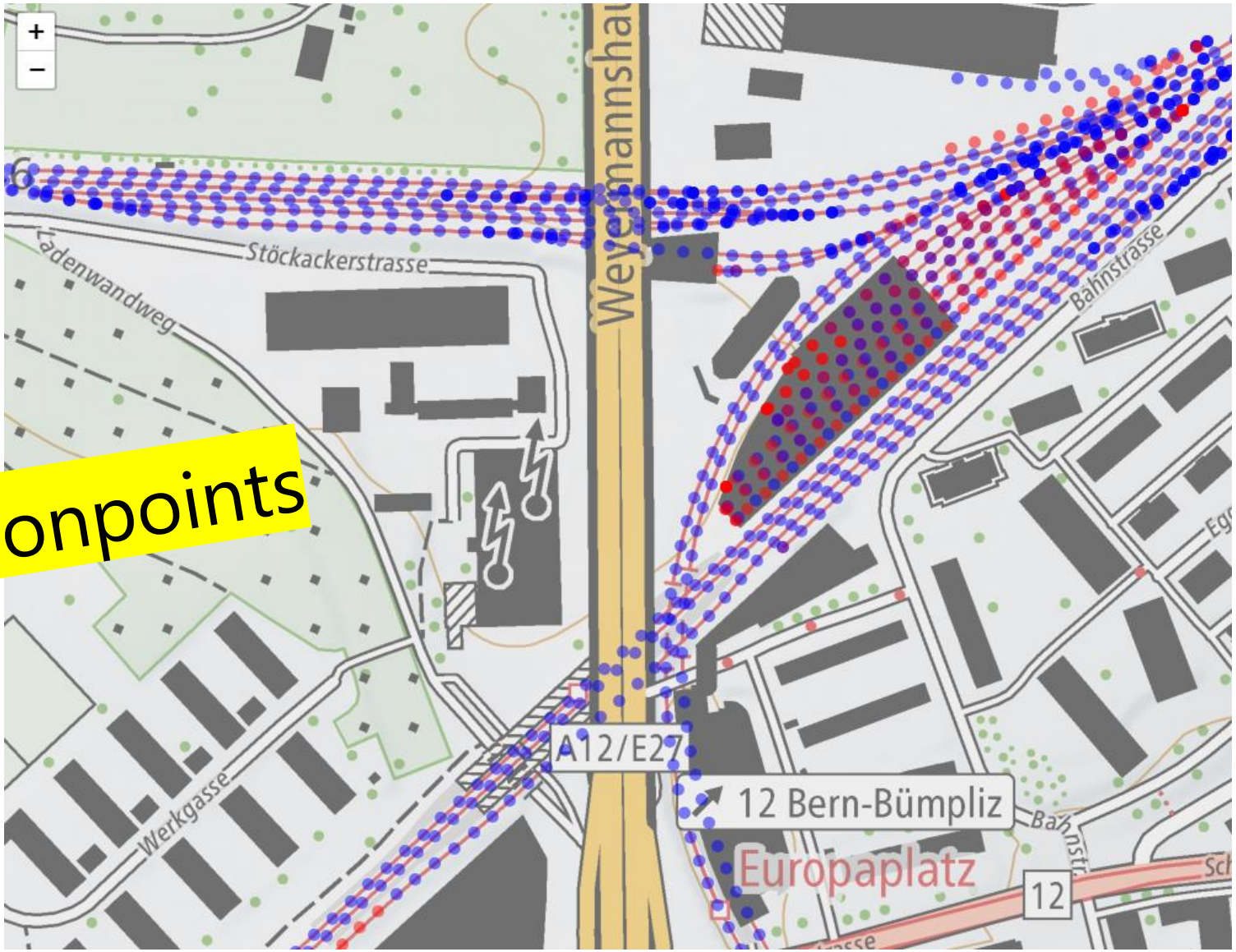
Business and IT...





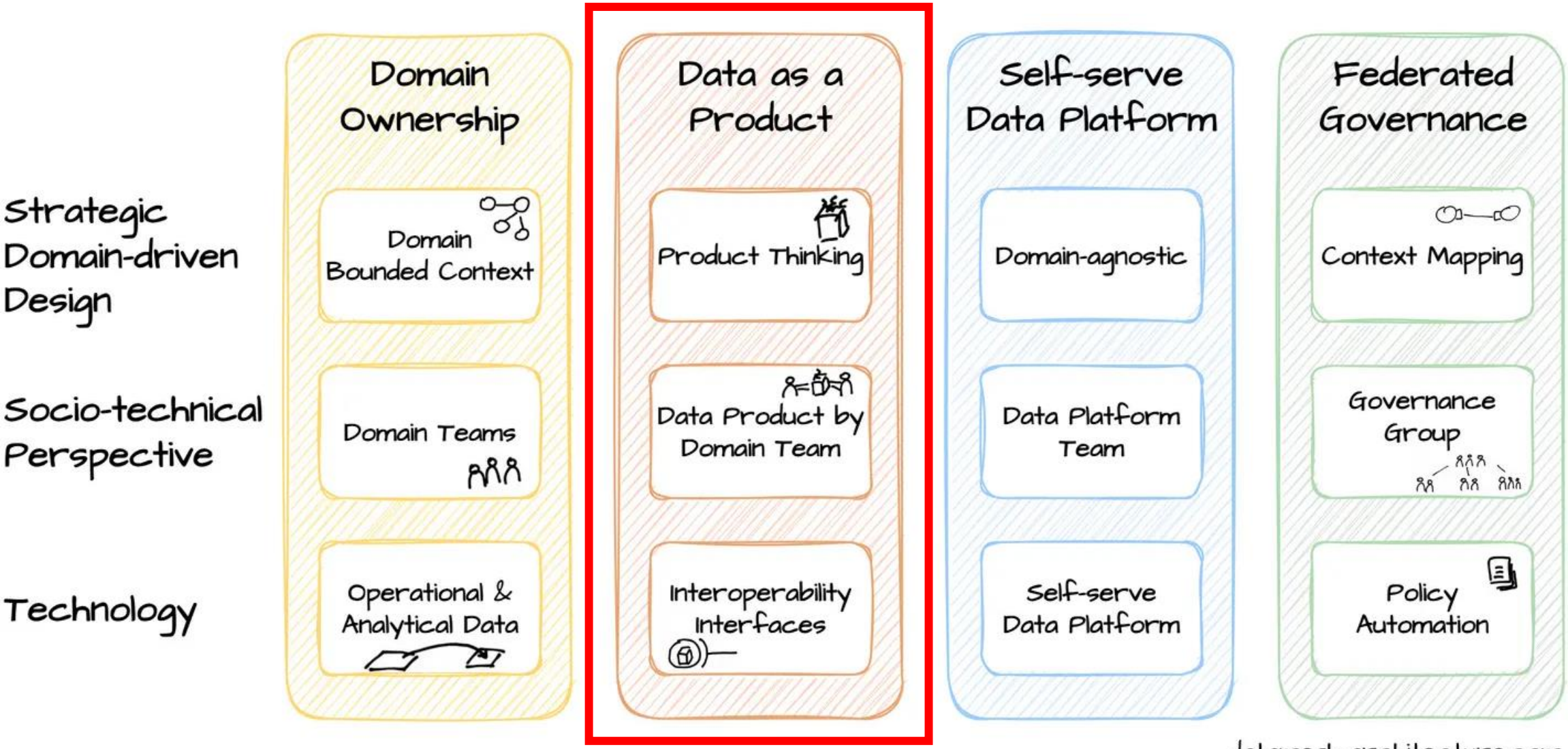
ELCA We make it work.

Architecture Principle: Keep it simple and stupid (KISS)



ELCA We make it work.

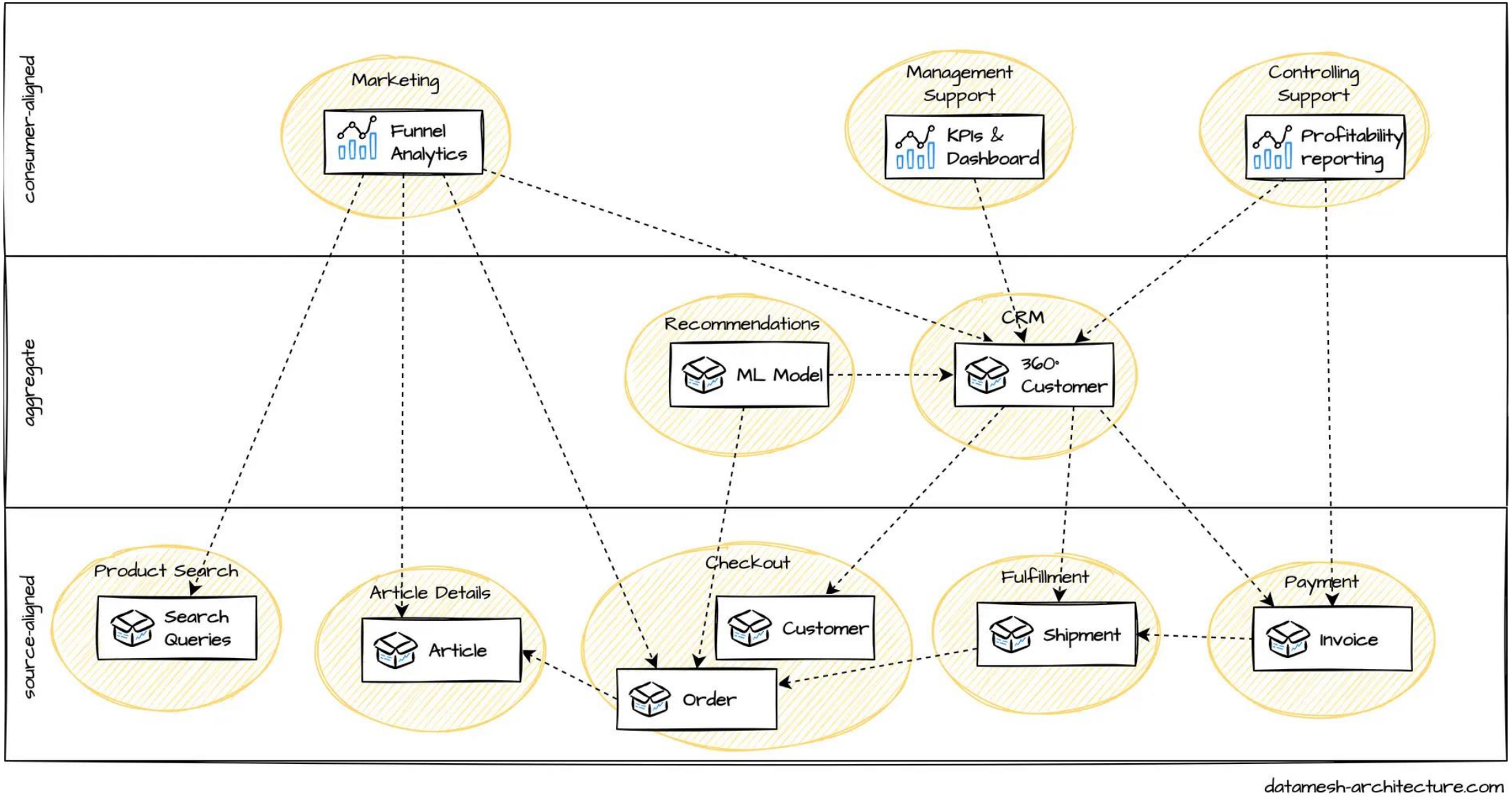
Data Mesh: Organizational Scaling



datamesh-architecture.com

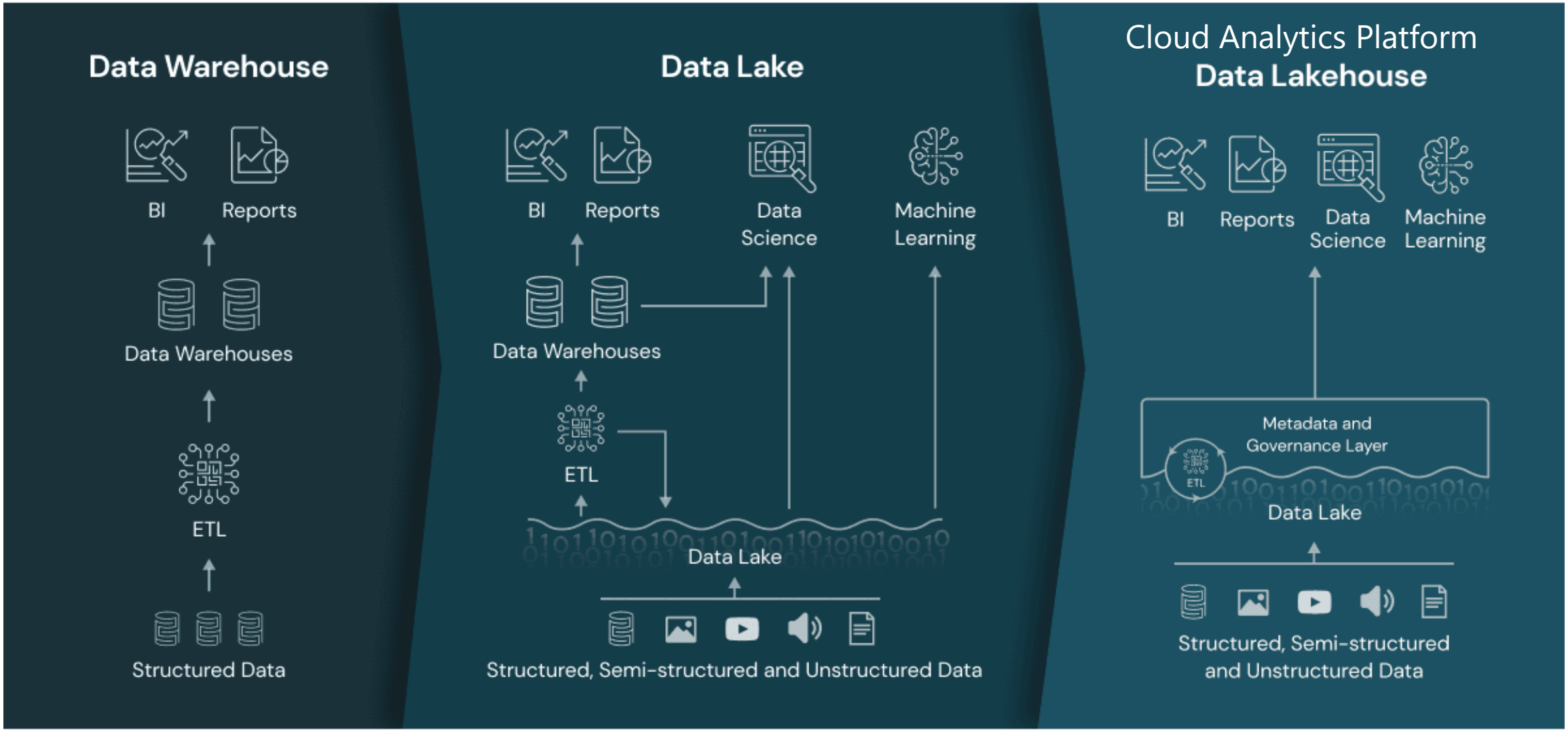
ELCA We make it work.

Data Mesh: Organizational Scaling



ELCA We make it work.

Cloud Analytics Plattform: State-of-the-art Technology



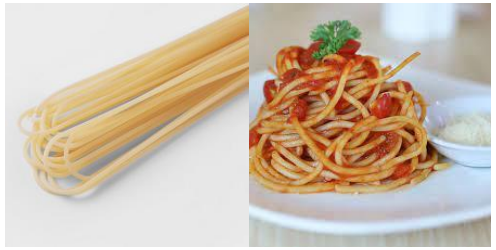
Databricks

ELCA We make it work.

Data Pipeline Automation Tools: State-of-the-art Technology

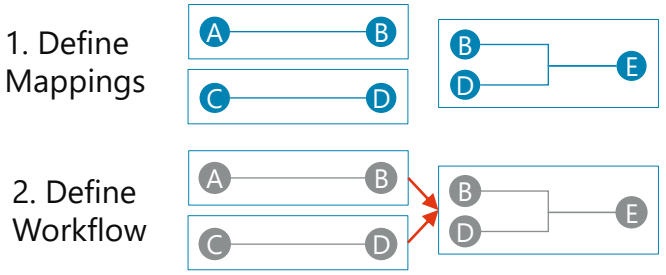
Scripting only
> 20 years

- SAS or SQL Scripts
- PL/SQL Stored Procedures
- Snowflake Stored Procedures
- …



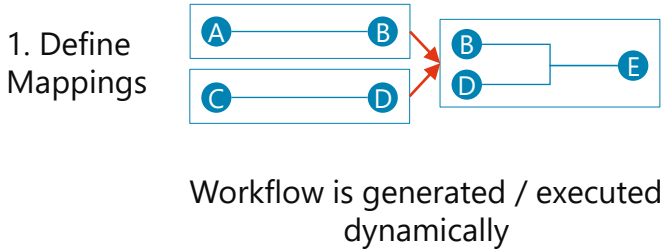
Traditional ETL Tool
> 15 years

- Informatica
- Oracle Warehouse Builder
- SSIS
- Talend
- Azure Data Factory
- …



Automation ETL Tool
> 10 years

- Wherescape
- Datavault Builder
- DBT, SQLMesh
- SmartDataLakeBuilder
- …



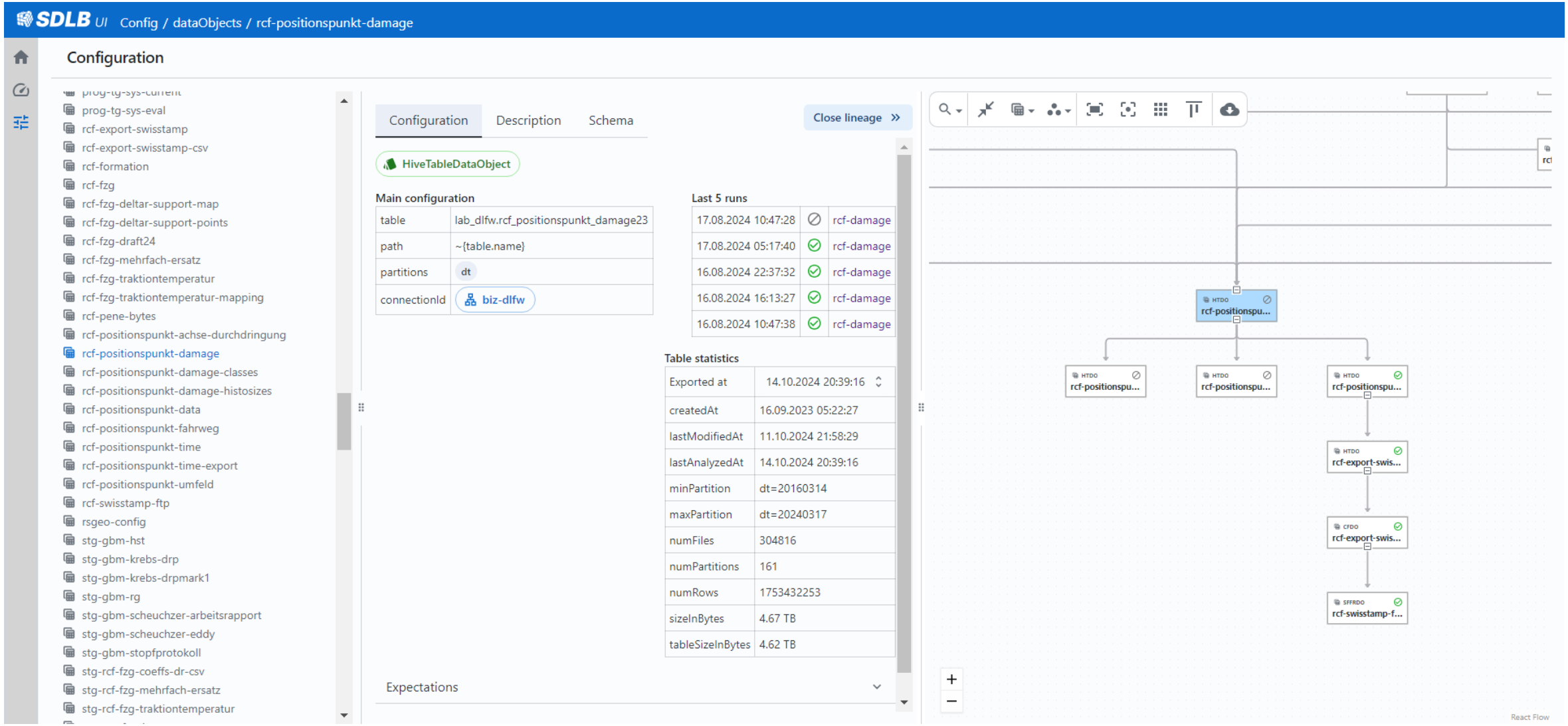
Workflow is generated / executed dynamically

SmartDataLakeBuilder

smartdatalake.io
github.com/smart-data-lake

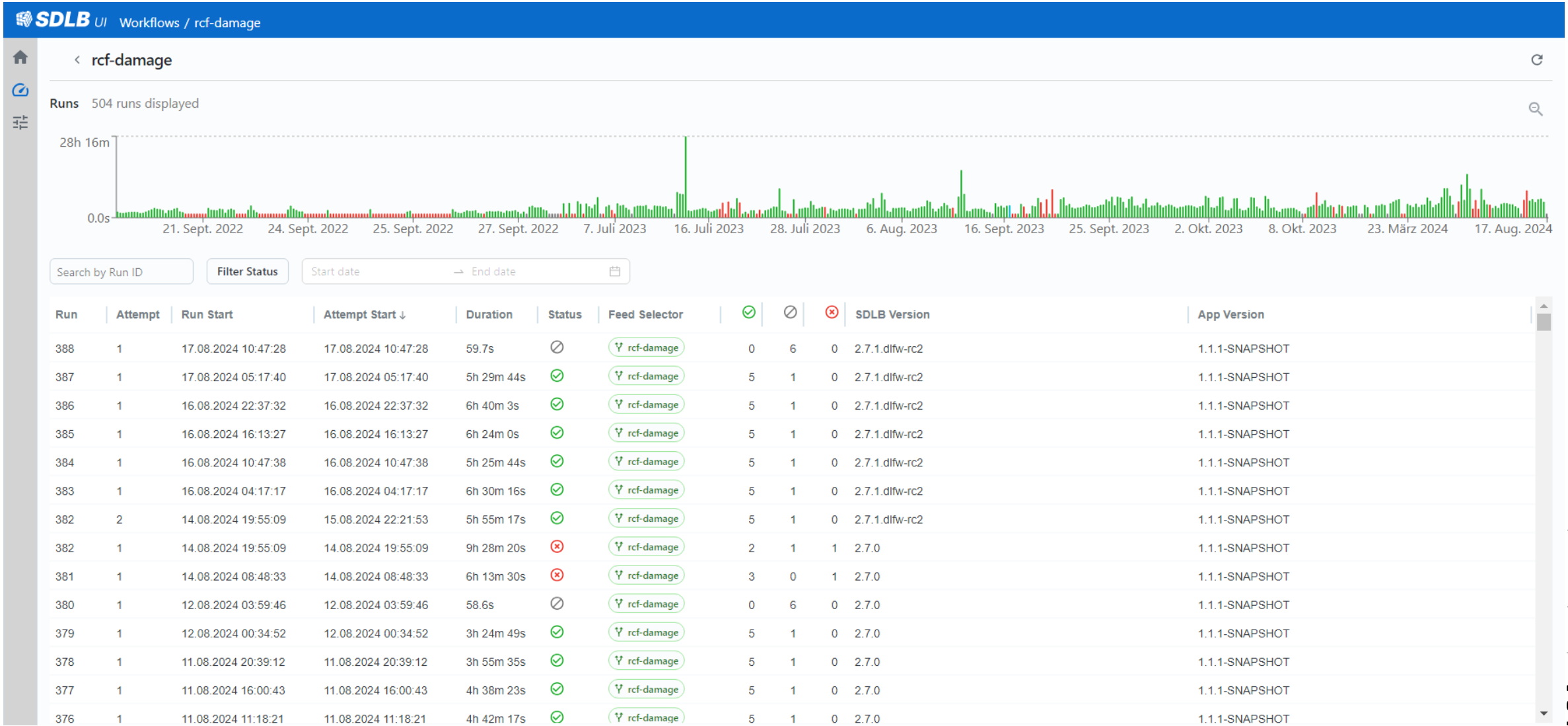
Making data pipelines explorable!

Catalog, Lineage, Documentation



ELCA We make it work.

Monitoring



Run	Attempt	Run Start	Attempt Start ↓	Duration	Status	Feed Selector				SDLB Version	App Version
388	1	17.08.2024 10:47:28	17.08.2024 10:47:28	59.7s	⊘	rcf-damage	0	6	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
387	1	17.08.2024 05:17:40	17.08.2024 05:17:40	5h 29m 44s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
386	1	16.08.2024 22:37:32	16.08.2024 22:37:32	6h 40m 3s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
385	1	16.08.2024 16:13:27	16.08.2024 16:13:27	6h 24m 0s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
384	1	16.08.2024 10:47:38	16.08.2024 10:47:38	5h 25m 44s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
383	1	16.08.2024 04:17:17	16.08.2024 04:17:17	6h 30m 16s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
382	2	14.08.2024 19:55:09	15.08.2024 22:21:53	5h 55m 17s	✓	rcf-damage	5	1	0	2.7.1.dlfw-rc2	1.1.1-SNAPSHOT
382	1	14.08.2024 19:55:09	14.08.2024 19:55:09	9h 28m 20s	✗	rcf-damage	2	1	1	2.7.0	1.1.1-SNAPSHOT
381	1	14.08.2024 08:48:33	14.08.2024 08:48:33	6h 13m 30s	✗	rcf-damage	3	0	1	2.7.0	1.1.1-SNAPSHOT
380	1	12.08.2024 03:59:46	12.08.2024 03:59:46	58.6s	⊘	rcf-damage	0	6	0	2.7.0	1.1.1-SNAPSHOT
379	1	12.08.2024 00:34:52	12.08.2024 00:34:52	3h 24m 49s	✓	rcf-damage	5	1	0	2.7.0	1.1.1-SNAPSHOT
378	1	11.08.2024 20:39:12	11.08.2024 20:39:12	3h 55m 35s	✓	rcf-damage	5	1	0	2.7.0	1.1.1-SNAPSHOT
377	1	11.08.2024 16:00:43	11.08.2024 16:00:43	4h 38m 23s	✓	rcf-damage	5	1	0	2.7.0	1.1.1-SNAPSHOT
376	1	11.08.2024 11:18:21	11.08.2024 11:18:21	4h 42m 17s	✓	rcf-damage	5	1	0	2.7.0	1.1.1-SNAPSHOT

ELCA We make it work.

Data Products and Modules

Gleisgeometrie

- Verifizierung
- Gleisarbeiten
- Sys.Prognose
- Einzel.Prognose

Schiene

- Verifizierung
- RSGeo
- Wear
- RCF
- Squatana

Trassenpreis

- Fzg Modelle
- Streckenklass.
- Kalibration
- Preissimulation

Common

- Fahrwegcheck
- Positionpoints
- Feinverortung

GBM Datahub

- Perimeternetz
- Open Data Spec.
- Datenspez.

ELCA We make it work.

Together?!

ELCA We make it work.

Data Products and Modules

Gleisgeometrie

- Verifizierung
- Gleisarbeiten
- Sys.Prognose
- Einzel.Prognose

Schiene

- Verifizierung
- RSGeo
- Wear
- RCF
- Squatana

Trassenpreis

- Ezg M…
- …
- …
- …

Common

- Fahrwegcheck
- Positionpoints
- Feinverortung

GB…

- …z
- Open Data Spec.
- Datenspez.

Our aim…

- Share more than **Concepts** and scientific **Datasets** -> **Share Code**
- Create an **open-source ecosystem** with multiple parties

ELCA We make it work.

ELCA We make it work.

Thank **you**

Zacharias Kull
Data Architect / Senior Manager

zacharias.kull@elca.ch
+41 79 438 42 54

ELCA We make it work.