



**Simplex** 4 Data

# Unsere Daten neu denken – Twin Information System (TwIS)

Digitaler Zwilling –  
organisiert als Wissensgraph

# Datenquellen, Fach-, Geographische Informationssysteme

Die Daten werden in Systemen mit ganz speziellen fachlichen Aufgaben erzeugt und verwaltet. Die Daten sind so strukturiert, dass sie die Fachaufgaben bestmöglich erfüllen. Für themenübergreifende Auswertungen sind die Daten erstmal nicht geschaffen, sie existieren auf strukturellen Inseln.



## Ein Beispiel zum Datenmanagement (Lärmkartierung)

Heute ist es immer noch sehr kompliziert, Daten für konkrete neue Aufgaben zusammenzuspielen. Die Praxis ist durch komplexe Datenmanipulationen geprägt, und für jede neue Aufgabe sind die Prozesse neu zu erstellen. – Ein riesiger Programmieraufwand entsteht, die neu zusammengestellten Daten liegen oftmals erst nach mehreren Wochen vor, ungeeignet für schnelle und aktuelle Entscheidungen.



„Kannst Du mir mal schnell das Straßennetz mit Topologien ausspielen?“



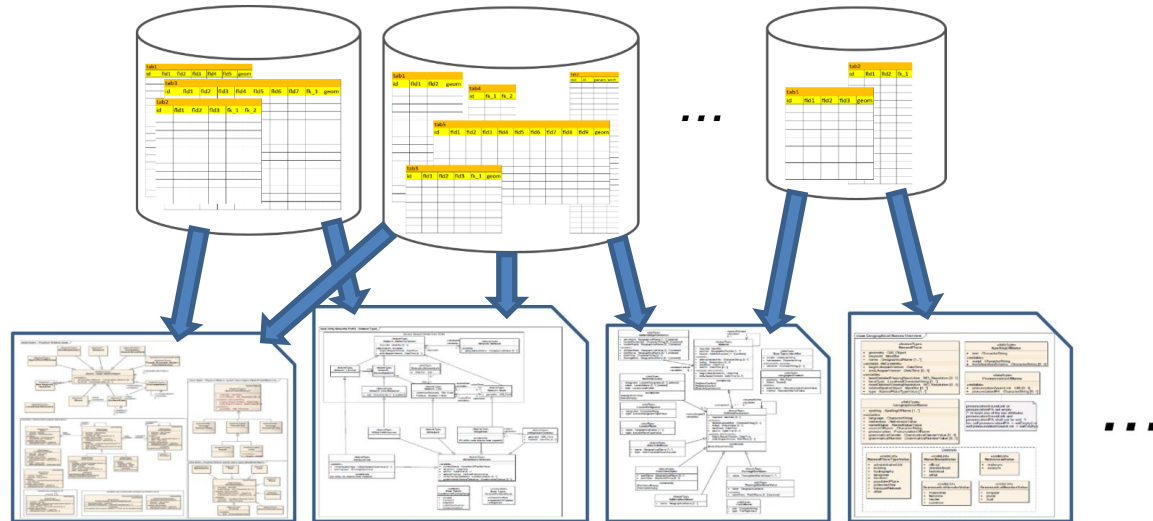
„Wir haben doch Gebäude, Adressen, Einwohnerzahlen und öffentliche Einrichtungen zu den Straßen?“



„Schön wären jetzt noch die Verkehrsmengen, -regeln und Geschwindigkeiten.“

## So werden Daten heute üblicherweise bereitgestellt.

Für das Zusammenspiel von Daten aus Fachsystemen (als Zylinder symbolisiert) werden für jede neue Sichtweise (als Fünfecke dargestellt) sogenannte ETL-Prozesse (Extract Transform Load) programmiert. Diese werden durch die dicken Pfeile charakterisiert (da es oftmals aufwendig zu erstellende Prozesse sind). Auch die Bereitstellung von standardisierten Open Data erfolgt heute auf diese Art und Weise. Dabei wirkt sich noch zusätzlich negativ aus, dass für spezielle Zielstellungen immer wieder Standards mit neuen Fachstrukturen kreiert werden. – Wir sehen heute ein „Wuchern“ an fachlichen Standards und ETL-Prozessen.



## Das MUSS auch anders möglich sein!

Schnell ist ein zentraler Datenpool hingemalt. Es geht aber nicht darum, die Daten einfach in einen Topf zu schütten, sondern sie dort themenübergreifend gemeinsam in einer Struktur abzulegen. Dann werden auch die ETL-Prozesse viel einfacher, denn es gibt eine einzige Zielstruktur. Notwendig ist dafür ein themenunabhängiges Datenmodell – die Lösung ist der Wissensgraph, der alle Objekte in einem Knoten und alle Objektverknüpfungen in einer Verbindungstabelle ablegt.



# Digitaler Zwilling – organisiert in einem einzigen Wissensgraphen

Die Fachsysteme werden nicht angefasst, die DV-Systeme bleiben so wie sie sind. Es wird ein eigenständiges Tool ergänzt, das die Daten in einem Digitalen Zwilling in einem einzigen Wissensgraphen ablegt. Wir trennen die Interoperabilität von den Fachsystemen.



# Zwillinge – mehr als oberflächliche Abbilder?



**August der Starke:  
ein analoger Zwilling**

- **dreidimensional**
- **optisch ansprechend**
- **ausmessbar**

**Was macht den Zwilling zu mehr  
als einem oberflächlichen  
Abbild?**

# Wissensgraph mit Objekten und Verknüpfungen



© OpenStreetMap Mitwirkende

Eine Datensammlung ist dann ein digitaler Zwilling,  
wenn sie neben den  
Eigenschaften von Objekten auch  
deren Beziehungen und  
Prozesse im Kontext abbildet.

# Beispiel für einen Wissensgraphen im TwIS

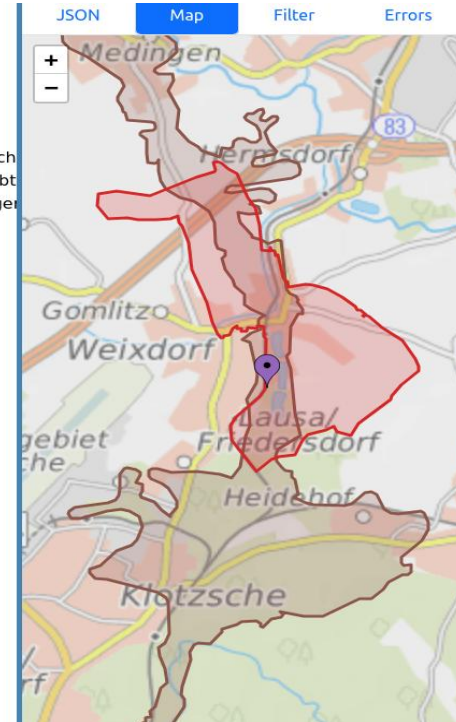
Links unten sehen wir die Objekte und ihre Beziehungen untereinander, rechts sind die Objekte grafisch dargestellt:

- Bäder als öffentliche Einrichtungen
- Betreiber
- Grundwasserleiter, die weiter mit Messstellen verknüpft sind
- Adressen
- Administrative Einheiten.

Die Daten sind aus mehreren Quellen im TwIS zusammengespield.



Statistischer Bezirk  
Adressen  
öffentliche Einrichtungen  
Städtische Einrichtungen und Unternehmen  
Grundwasserleiter  
öffentliche Einrichtung -> Adresse  
öffentliche Einrichtung - (liegt in)-> statistisch  
Städtische Einrichtung/Unternehmen -betreibt  
Grundwasserleiter -> öffentliche Einrichtungen



## Auswertungen im Wissensgraphen

Auswertungen und damit beliebige Datenzusammenstellungen können jetzt ganz einfach aus dem Wissensgraphen „herausgeschnitten“ werden, da alles in einem einzigen themenübergreifenden Graphen in einem Knoten und einer Kante organisiert ist – im Beispiel: Überschreitungen von Nitratwerten im Grundwasser mit den zugehörigen Bädern.

Auch standardisierte Datenzusammenstellungen werden einfach als Sichten auf den Wissensgraphen kreiert.

|  | Grundwasserleiter Name<br>gwl_name | Grundwasserleiter Geometrie<br>gwl_geom                                            | Messung<br>messung                    | Messwert<br>messwert | Einrichtung<br>einrichtung | Träger<br>traeger   | Träger Webseite<br>traeger_webseite | s  | c   | o |
|--|------------------------------------|------------------------------------------------------------------------------------|---------------------------------------|----------------------|----------------------------|---------------------|-------------------------------------|----|-----|---|
|  | Berliner Elbelauf (E2n-S1v (Ho))   |   | Nitratwert in mg/l, Merkmal 210, 2006 | 56.6                 | Badestelle Weixdorf        | Dresdner Bäder GmbH | www.dresdner-baeder.de              | 10 | 109 | 1 |
|  | Elbtal-GWL                         |  | Nitratwert in mg/l, Merkmal 210, 2006 | 211                  | FKK-Strandbad Wostra       | Dresdner Bäder GmbH | www.dresdner-baeder.de              | 10 | 109 | 2 |

## Datenaufbereitung aus den Fachsystemen

Die Prozesse vereinfachen sich um alles rechts grün Geschriebene, sodass für das Datenmanagement Einsparungen von bis zu 80% auftreten.



- Vorbereitung der Daten für interoperable Bereitstellungen im Fachsystem
- Export der Daten
- Sichtung der Daten, Objektdefinition und –strukturierung für Zielsystem
- Modellierung der Zielstruktur
- Aufsetzen des Systems mit den modellierten Strukturen



- Das Fachsystem wird nicht angefasst.
- Export der Daten
- Sichtung der Daten, Objektdefinition für Zielsystem
- Modellierung entfällt
- System ist bereits verfügbar.

## „Mapping“ und Konvertierung

Die Prozesse vereinfachen sich um alles rechts grün Geschriebene, sodass für das Datenmanagement Einsparungen von bis zu 80% auftreten.



- „Mapping“ der Fachdaten auf die Zielstrukturen
- Erstellen der ETL-Prozesse  
sowohl für INSERT als auch UPDATE
- Konvertierung
- Metadaten sind gesondert einzuspielen.



- „Mapping“ wird deutlich einfacher, da die Zielstrukturen immer gleich sind.  
(Machbarkeitsstudie zum Einsatz von KI)
- QS zu den mit KI erstellten Zuordnungen  
Art des UPDATE wird einfach angekreuzt.
- Konvertierung
- Metadaten werden automatisch mit angelegt.
- zusätzliche QS durch (teil-) automatische Rückkonvertierung der Daten in die Fachstrukturen

# Datenbereitstellung

Die Prozesse vereinfachen sich um alles rechts grün Geschriebene, sodass für das Datenmanagement Einsparungen von bis zu 80% auftreten.



- Die Dienste sind aufzusetzen.
- Neue Datenzusammenstellungen sind wieder komplett neu umzusetzen.
- Metadaten sind gesondert bereitzustellen.
- Für die Nutzung der Daten sind die phänomenalen Datenstrukturen auszuwerten.  
→ riesiger Programmieraufwand

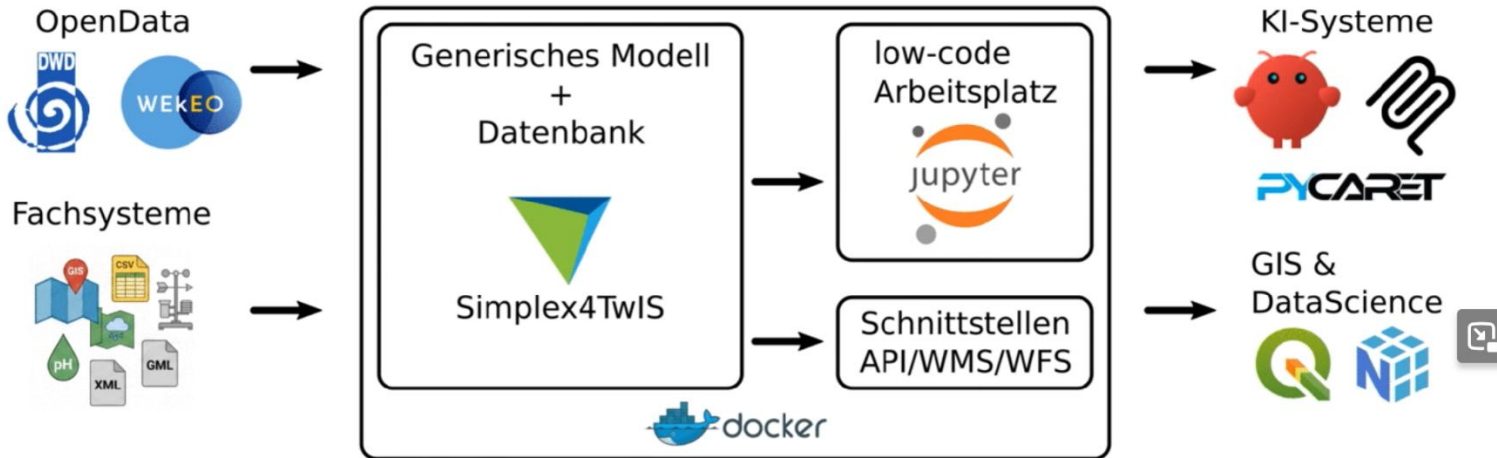
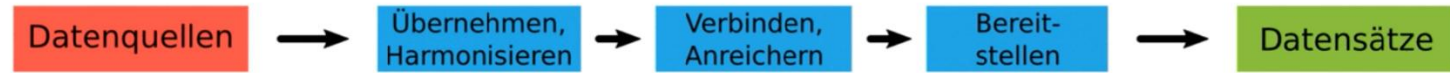


- Aufsetzen von Diensten entfällt. Die Dienste sind sofort verfügbar.
- Einfachstes Anlegen von Szenarien für spezielle Datenzusammenstellungen
- Metadaten werden automatisch bereitgestellt.
- Die Daten werden immer wieder in einheitlicher Struktur geliefert und sind damit einfach weiterverarbeitbar.  
→ einmaliger Programmieraufwand

# TwIS und KI: Mit dem TwIS „Ready for Artificial Intelligence“

Durch die Organisation der Daten in einem Wissensgraphen sind diese bereits für Anwendungen mit KI vorkonfektioniert. Wir arbeiten aktuell an mehreren KI-Ansätzen:

- Auswertungen von Messdaten mit Machine Learning Prozessen
- Bereitstellung von Agenten zum Einsammeln und Mappen der Daten
- Automatische Erstellung von Auswertungen durch Anreicherung der Large Language Modelle um den Wissensgraphen.



# Twin Information System (TwIS)

## Alles – Einmalig – Einfach



**Simplex**4Data

- Nutzbarmachung komplexer Datenlandschaften
- Objekte in einem Wissensgraphen
- Digitaler Zwilling der Realität und der Fachinformationssysteme
- Datenbereitstellungen: Interoperabilität – Nachhaltigkeit – Weiterverarbeitbarkeit
- Daten sind für KI vorbereitet
- Agilität und Open Source: In bestehende Systemlandschaften integrierbar und sofort einsetzbar



[www.simplex4data.de](http://www.simplex4data.de)

[www.hrd-consulting.eu](http://www.hrd-consulting.eu)

<https://db01.simplex4data.de/projekte/sandbox>