



# Data Governance Einführung bei der Deutschen Bahn Regio

Erfahrungen und Best Practices zum Rollenmodell  
und Datenqualitätsmanagement

---

12.04.2024 | Wien



**Tanja Schlesinger**



Tanja Schlesinger ist Data Officer und verantwortet die Data & Analytics Initiative bei der DB Regio AG. Ihr Fokus liegt auf der digitalen Transformation unter Bereitstellung eines verlässlichen Datenökosystems, das Entscheidungsempfehlungen, Prozessautomatisierungen und ein digitales Kundenerlebnis ermöglicht. Als kreative, innovative und welt-offene Vordenkerin blickt Tanja Schlesinger auf mehr als drei Jahrzehnte Berufspraxis in verschiedenen Top-Management-, Finanz- und Strategiefunktionen zurück.

**Dr. Carsten Dittmar**



Dr. Carsten Dittmar ist Executive Partner und Area Director West bei der Alexander Thamm GmbH. Er beschäftigt sich seit über 20 Jahren intensiv mit den Themenfeldern Business Analytics, Data Science und Artificial Intelligence mit dem Fokus auf strategische und organisatorische Beratung von datengetriebenen Vorhaben. Carsten Dittmar ist europäischer TDWI Fellow und Autor diverser Fachpublikationen und Referent bei zahlreichen Fachveranstaltungen. Dr. Carsten Dittmar leitet den TDWI Themenzirkel Data Strategie & Data Governance

**Dr. Christian Fürber**



Dr. Christian Fürber ist Autor, Dozent und Experte für Themen des Datenmanagements und der Digitalisierung. Mit seinem Unternehmen, dem Information Quality Institute, unterstützt er Organisationen und Unternehmen aller Branchen dabei, spürbaren Mehrwert aus Ihren Daten zu generieren. Christian hat seit 2005 mehr als 20 große Daten- und Digitalisierungsprojekte erfolgreich geleitet. Dr. Christian Fürber leitet den TDWI Themenzirkel Data Strategie & Data Governance

1. **Grüne Mobilität erfordert gesunden Daten**
2. Data Governance Initiative bei der DB und der DB Regio
3. Prozessorientierte Data Governance und Framework für Datenqualitätsanalysen
4. Lessons Learned und Best Practices

DIE BAHN ALS  
RÜCKGRAT DER  
TRANSFORMATION  
ZUR GRÜNEN  
MOBILITÄT IN  
DEUTSCHLAND



Unser Anliegen: 1 Mrd. mehr Fahrgäste!



Einsatz von alternativen Antrieben, Grünstrom und energieeffizientem Fahren entlasten die Umwelt.



Bis 2038 sollen alle Zugleistungen mit erneuerbarer Energie erbracht werden.



Bis 2040 wollen wir klimaneutral sein.

# DEUTSCHLAND BRAUCHT EINE STARKE GRÜNE SCHIENE



- Reduzierung der Umweltbelastung

- Absenkung der Abhängigkeit von fossilen Brennstoffen

- Erhalt natürlicher Rohstoffe

- Förderung gesünderer und lebenswerterer Gemeinschaften



 **ROBUST**

 **EFFEKTIV**

 **MODERN**

 **GRÜN**

# DB Regio 2023 auf einen Blick



Pünktlichkeit  
Schiene 91 %



39.587 Mitarbeiter

Jeden Tag werden Millionen von  
Daten generiert, eine  
vertrauenswürdige Datenbasis  
ist ein Must Have

9,7 Mrd. € Umsatz

22.708 Züge pro Tag



4,65 Mio. Reisende pro Tag

Jährlich 1,7 Mrd. Reisende



Ziel: Klimaneutrales Beheizen  
der Werke



Betrieb Elektrobuse mit 100%  
Grünstrom



GRÜNE  
MOBILITÄT

Betrieb und Instandhaltung der  
weltgrößten Flotte mit  
Wasserstoffantrieb

Fahrassistenz Dieseltriebzüge

Steigender Einsatz von  
ökologischem Strom



Intelligente Kombination  
Fahrrad & Zug



# ZUVERLÄSSIGE DATEN. GRUNDLAGE FÜR KI.



## DATENVERFÜGBARKEIT UND -VERTRAUENSWÜRDIGKEIT ALS ERFOLGSFAKTOR



**CHALLENGE** Fehleranfällige manuelle Befundung möglicher Schäden

**LÖSUNG**  
Kamera-basierte Befundung der Züge zur Feststellung von Auffälligkeiten (Sichtprüfung)



**CHALLENGE**  
Stetig wachsende Menge an Sensordaten kann manuell nicht mehr verarbeitet werden

**LÖSUNG**  
Automatisierte Datenverarbeitung und -plausibilitätsprüfung



**CHALLENGE**  
Komplexität der Fahrpläne stoßen mit mathematischen Lösungsansätzen ihre Grenzen

**LÖSUNG**  
KI-basierte Unterstützung der Fahrplanoptimierung unter Berücksichtigung von Kapazitätsvorgaben



**CHALLENGE**  
Informationsüberflutung der Zugdisponenten, die kurzfristig komplexe Entscheidungen treffen

**LÖSUNG**  
Echtzeit Datenzugriff auf KI-gestützte Analysen für optimierte Dispositionsentscheidungen



# DIE ERREICHUNG DER NACHHALTIGKEITSZIELE ERFORDERT...



... EFFIZIENZ IM RESSOURCENEINSATZ



... ÜBERZEUGUNG DER FAHRGÄSTE STATT AUTO, MIT DEM ZUG ZU FAHREN



... EINE QUELLE DER WAHRHEIT FÜR DIE VERSORGUNG DER DB REGIO MIT ERFORDERLICHEN GESCHÄFTSINFORMATIONEN UND –ANALYSEN

- ONESOURCE

... EIN KOMPETENZCENTER FÜR DATEN-TRANSPARENZ, SICHERHEIT IN DER TEILUNG UND DIE NOTWENDIGE QUALITÄT ALS GESUNDE BASIS FÜR DIE DIGITALE TRANSFORMATION

- DAS DATA OFFICE

# JEDE ERFOLGREICHE DIGITALE TRANSFORMATION BASIERT AUF GESUNDEN DATEN.

## WERT DER DATEN

Wir heben unseren Datenschatz und ermöglichen **fundierte Entscheidungen**, das Treiben von **Innovation** und Verbessern der **wirtschaftlichen Leistungsfähigkeit**

## DATENKULTUR

**Befähigung** von Mitarbeitenden und **Förderung von Datenkenntnissen**, -vertrauen zur Zusammenarbeit für eine datenbasierten Entscheidungsfindung, für Innovationen und kontinuierlicher Verbesserung der betrieblichen Performance

## DATA OPERATING MODEL

Etablieren von klaren Verantwortlichkeiten, **Prozessen** und **Governance** als Wegbereiter für ein effektives **Datenmanagement**, Ermöglichen von Datenteilung und für relevante Insights

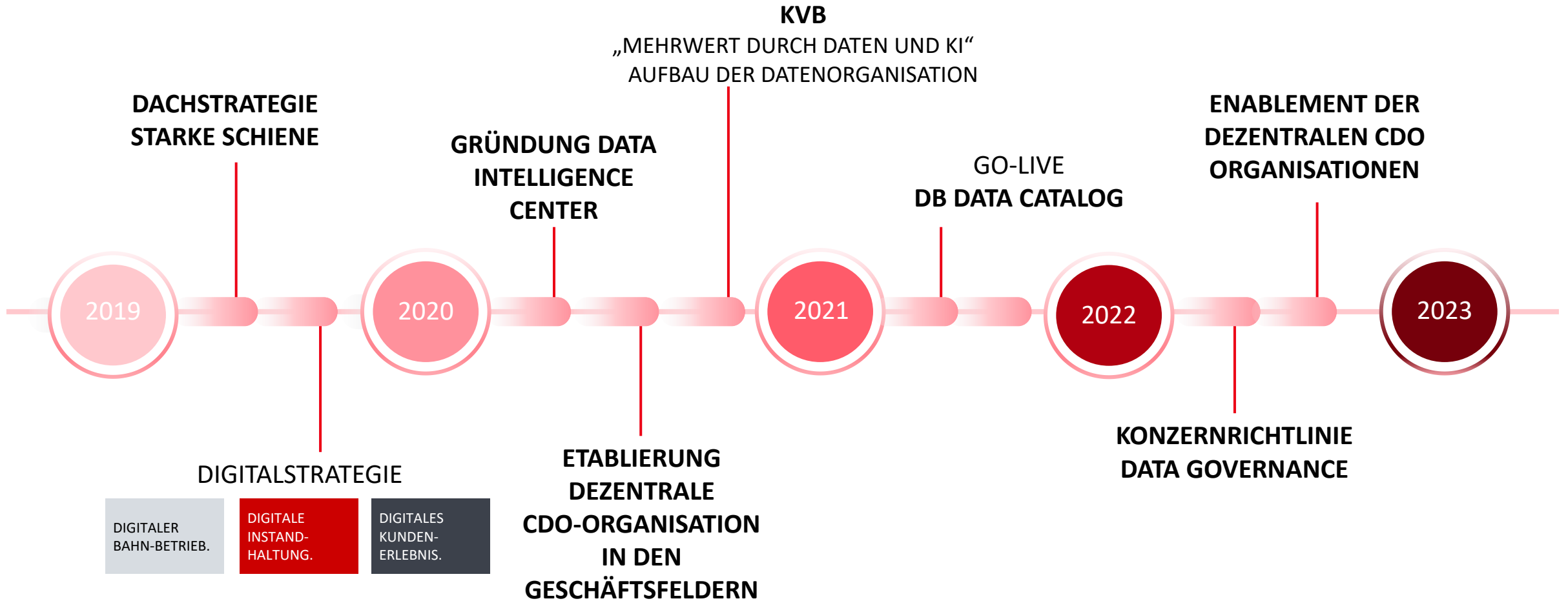
## DATENGRUNDLAGEN

Sicherstellung einer **hohen Datenqualität, Zuverlässigkeit und Nutzbarkeit** als gesunde Basis für aussagekräftige und richtige Analyseergebnisse



1. Grüne Mobilität erfordert gesunden Daten
2. **Data Governance Initiative bei DB und DB Regio**
3. Prozessorientierte Data Governance und Framework für Datenqualitätsanalysen
4. Lessons Learned und Best Practices

# DIE DATENJOURNEY DER DEUTSCHEN BAHN SEIT 2019



## DATA ACCESS - UMKEHR DER DATENNUTZUNGSBEGRÜNDUNGSPFLICHT

Daten sind im DB-Konzern grundsätzlich frei zugänglich und über alle Geschäftsfelder hinweg nutzbar  
Ausnahmen sind zu begründen. Nicht der Datenzugang, sondern eine etwaige Ausnahme vom freien Datenzugang ist begründungspflichtig



### DB DATA CATALOG

- Verpflichtende Datentransparenz durch einen zentralen, konzernweiten Datenkatalog
- Einschränkungen werden durch betroffene Datenfelder gekennzeichnet



### DATA GOVERNANCE-STRUKTUR

- Vorgaben & Vorgehensweisen zu allen datenrelevanten Fragen wie Datenqualität, -schutz und -sicherheit
- Handlungssicherheit in der Umsetzung des Data Access



### DATA MESH

- Sukzessive Bereitstellung eines konzernweit geregelten Zugangssystems auf Daten
- Daten und Lakes bleiben in den GF/SE, der Data Mesh ermöglicht gemeinsames Arbeiten in einem föderalen System



### KOSTENFREIHEIT

- Die DB-interne Datenweitergabe erfolgt kostenfrei
- Aufwand durch die Bereitstellung von notwendigen Datenschnittstellen sowie Datenaufbereitung ist davon ausgenommen



### DATEN- ÖKOSYSTEM

- Datenökosystem mit Geschäftspartnern ermöglicht vertraglich geregelten zentralen Zugriff auf benötigte, externe Daten

# RAHMENRICHTLINIE FÜR DATA GOVERNANCE



**Rahmenrichtlinie**  
111.0001 **Rahmenrichtlinie Data Governance**  
Version 1.0 - Gültig ab 01.06.2022

**Inhalt**

**Zusammenfassung**

Diese Rahmenrichtlinie setzt die Zielvorgaben des Konzerns zu Transparenz, Verfügbarkeit, Qualität sowie von Verantwortlichkeiten im Datenmanagement im DB-Konzern. Diese Richtlinie definiert den Zielzustand der Datenorgane und Metadatenmanagement im DB-Konzern. Dazu gehören die Beschreibung, Bewertung, Pflege sowie von im DB-Konzern erzeugten und gespeicherten Daten. Der DB Datenkatalog wird damit zur zentralen, laufend und über Daten und Datennutzung im DB Konzern. Transparenz im Sinne dieser Rahmenrichtlinie bedeutet, dass Mitarbeiter der Deutschen Bahn AG und ihrer verbundenen Unternehmen, die im Rahmen über Daten benötigen, Zugriff auf Beschreibungen von Daten und Beispieldaten der jeweiligen Quelle mittels des zentralen Datenkatalogs erhalten.

**Änderungshinweise**

**Regelwerkseigenschaften**

<b>Hauptgruppe</b>	Organisation und Managementsysteme
<b>Untergruppe</b>	Datenmanagement
<b>Herausgeber</b>	Deutsche Bahn AG
<b>Geltungsbereich</b>	Die Regelung gilt Konzernweit mit Ausnahme von DB Schenker, DB Arriva, Clever Shuttle
<b>Rechtsraum</b>	Geltung ist eingeschränkt auf Deutschland

Die aktuelle Version dieses Regelwerks finden Sie in der Konzern-Lokal gespeicherte Dokumente und Ausdrücke könnten veraltet sein.  
Regelwerk in der KRWD aufrufen  
Nur für den internen Gebrauch!

**Rahmenrichtlinie** 111.0001 **Data Governance**  
Seite 1 von 8

**Inhaltsverzeichnis**

01	Zielstellung und Aufbau der Rahmenrichtlinie	1
02	Geltungsbereich und Zielgruppe	2
03	Verantwortung und Zuständigkeiten	3
04	Datenqualität	4
05	Datentransparenz	5
06	Datenverfügbarkeit	7
07	Verweise	8
08	Anlagen	8

**01 Zielstellung und Aufbau der Rahmenrichtlinie**

(1) Mit stetig wachsenden Anforderungen an Technologien und schnelle Prozesse wachsen der Bedarf und die Anforderungen an eine konzernweit geregelte Data Governance. Am 8. September 2020 hat der Konzernvorstand die Konzernvorstandsbeschlussvorlage („KVB“) „Meinwert durch Daten und KI“ beschlossen, um diesem Bedarf nachzukommen. Diese Rahmenrichtlinie („Richtlinie“) setzt die Zielvorgaben des KVB zu Transparenz, Verfügbarkeit und Qualität sowie von Verantwortlichkeit für Daten um.

Über eine verpflichtende Nutzung des DB Datenkatalogs werden konzernweite Datentransparenz erreicht und Standards für ein einheitlich geregeltes Zugangssystem auf Daten gesetzt. Der DB Datenkatalog wird damit zur zentralen, laufend und verpflichtend gepflegten Wissensinstanz über Daten und Datennutzung im DB Konzern. Transparenz im Sinne dieser Richtlinie bedeutet, dass alle internen und externen Mitarbeitenden der Deutschen Bahn AG und ihrer verbundenen, zum Geltungsbereich dieser Richtlinie zählenden Unternehmen, die im Rahmen ihrer Tätigkeit Informationen über Daten benötigen, Zugriff auf Beschreibungen von Daten und Beispieldaten der jeweiligen Quelle mittels des zentralen Datenkatalogs erhalten.

Fachautorin: Anna Einhorn | TDD | Tel.: (+49) 0160 97404044 | Version 1.0 | Gültig ab: 01.06.2022

## WESENTLICHE INHALTE DER RICHTLINIE

- Zuständigkeiten des Chief Digital Officer (“CDO”)
- Zuständigkeiten des Data Intelligence Center (“DICE”)
- Zuständigkeit des zentralen Führungsgremiums “Data Council”
- Rollenprofil Data Officer
- Rollenprofil Data Owner
- Zielsetzung der zentralen Ziele Steigerung Datenqualität, Datentransparenz, Datenverfügbarkeit

# INHALTE DER RICHTLINIE DATA GOVERNANCE



## ZIELSETZUNG & GELTUNGSBEREICH

- Aufbau einer konzernweiten Data Governance Struktur und Organisation
- Eine zentrale Vorgabe ist die **Umkehr der Datennutzungsbegründungspflicht**
- Ziel der Richtlinie ist es, die Mitarbeitenden des DB-Konzerns bei dem effizienten, (rechts-)sicheren und wertstiftenden Umgang mit Daten zu unterstützen
- **Konzernweiter Geltungs-bereich** innerhalb der DB, ihren GF/SE & Konzernleitung innerhalb der Deutschlands. Umfasst alle Geschäfts-prozesse, Verfahren, Organisationsbereiche und Mitarbeitende

## RELEVANTE ROLLEN

- CIO/CDO oder **äquivalente** Rollen in den GF/SE sind für die Umsetzung verantwortlich
- **DICe** gibt **Standards/Anforderungen** für das Datenmanagement vor
- **Data Council entscheidet bei Uneinigkeit:** Qualität, Standards, Transparenz, etc. Besteht aus Data Officern der GF und (wahlweise) weiteren Experten
- Der **Data Officer** wird von CIO/CDO oder äquivalente Rolle ernannt. Er wirkt bei der Benennung des **Data Owners** mit. Der **Data Owner** verantwortet Qualität, Verfügbarkeit, und Management der Meta-(Daten) seines Zuständigkeitsbereichs

## ZIEL 1: DATENQUALITÄT

- Der **Data Owner** verantwortet das **Datenqualitäts-management** unter Beachtung der **Konzernvorgaben** des Data Councils hinsichtlich Standards, Policies und Richtlinien
- Definition und Erhebung **KPIs** zur Pflege und Überwachung der Datenqualität

## ZIEL 2: DATENTRASPARENZ

- **Vorrangiger Anschluss** von priorisierten Datenquellen. Metadaten sind in **systematischer Art** und Weise im zentralen DB Datenkatalog nach zentral vorgegebener **Metadatenstruktur** des **DICe** zu erfassen
- **Klassifizierung von Datenobjekten** anhand der Datengruppen: **Vertraulichkeit, Personenbezug, Lizenz und regulierte Daten.**

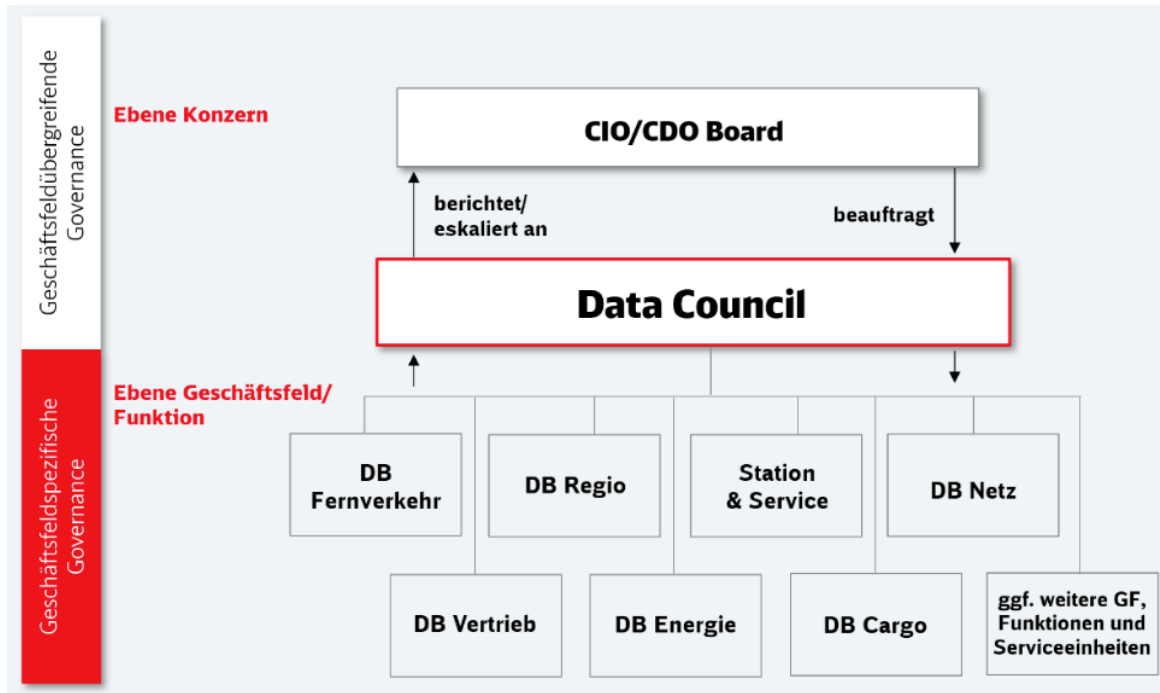
## ZIEL 3: DATENVERFÜGBARKEIT

- Jedes GF/SE verantwortet die Verfügbarkeit und Teilung seiner Daten entsprechend den **Konzernvorgaben.**
- **Antragsprozess** erfolgt über den DB Data Catalog mit einem **standardisierten Antragsprozess**
- **Dezentrale Systeme und Vorgehensweisen** für die Beantragung und Vereinbarung der Datennutzung können **eingeschränkt beibehalten werden**, wenn sie für angrenzende Funktionen eine Arbeitserleichterung bieten und der Datennutzungsvereinbarungsprozess beibehalten werden kann

# ZENTRALE GREMIEN FÜR DATA GOVERNANCE BEI DER DB



## EINORDNUNG DES DATA COUNCIL IM DB KONZERN







## DAS DATA COUNCIL

### BESCHREIBUNG

Das Data Council ist das höchste Gremium, welches sich **exklusiv** mit dem Thema **Daten** beschäftigt.

### LEITUNG

 Vorsitz des Boards:	 Geschäftsführung:	 Protokoll & Einladungsmanagement:	 Sitzungsintervall und Dauer:
Claudia Pohlank, Leiterin Data Intelligence Center Arlene Bühler, CIO/CDO DB Cargo AG	Alexander Stavinski, Data Intelligence Center	Data Intelligence Center	monatlich – 90 Minuten

### TERMINE UND ORGANISATION

Das Data Council tagt **monatlich** im **engeren** bzw. zur **Quartalsmitte** im **erweiterten** Kreis, mit einer Sitzungslänge von 90 Minuten

### THEMEN UND ZIELE

**Geschäftsfeldübergreifende Organisation**, die der strategischen Diskussion, Steuerung und Wirkungsüberwachung der Daten und KI-Themen im Konzern dient



# DB DATA CATALOG - DATENTRANSparenZ UND -TEILUNG

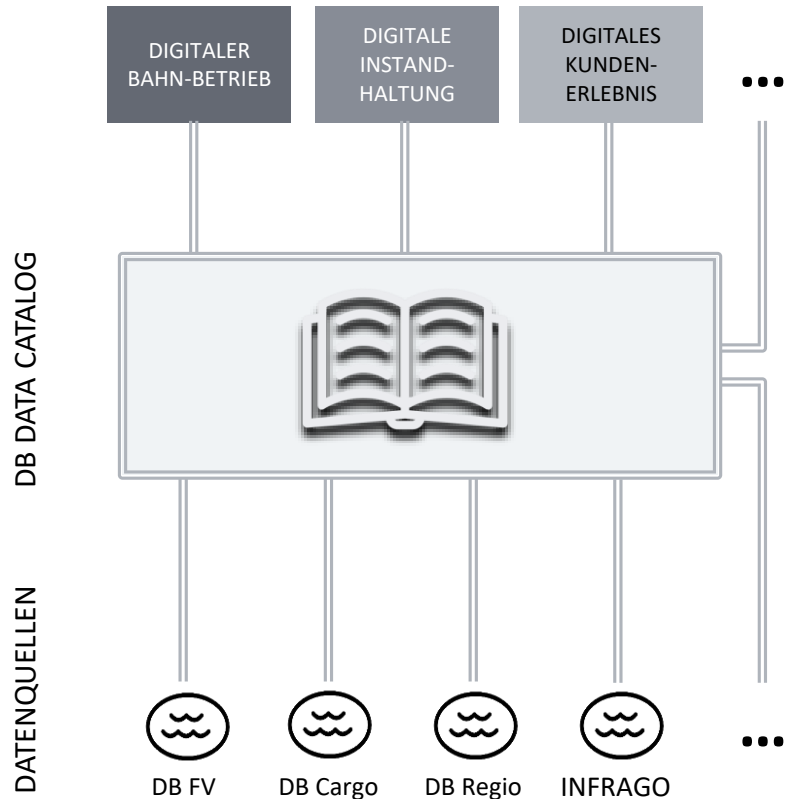


The screenshot shows the DB Data Catalog interface. A search bar at the top contains 'Suchbegriff eingeben'. The left sidebar lists navigation options like 'Home', 'Schnelleinstieg', and 'Publikation'. The main content area features a 'Home' section with a welcome message and a 'Meine laufenden Aktivitäten' table. Red callouts highlight: 'Suche und finde Daten' (pointing to the search bar), 'Deine To Do's' (pointing to the activities table), 'Hauptmenü' (pointing to the sidebar), and 'Hilfe' (pointing to a help icon). Below the callouts, there are four informational cards: 'Zur Orientierung' (explaining the main menu), 'Suche und finde Daten' (explaining the search function), 'Deine To Do's' (explaining the activities list), and 'Hilfe benötigt?' (explaining how to find help).

Icon	Entitätstyp	Entität	Status	Letzte Änderung	Abgelehnt von / Kommentar
	Data Set Proposal	<a href="#">Kontaktstammdaten</a>	approval pending	02.05.2023	-

- ## DB DATA CATALOG
- Zentrales Datenverzeichnis des DB Konzern
  - 360 Grad-Blick auf Daten (fachlich, technisch, Governance)
  - Lesezugriff für alle DB Mitarbeiter über Single-Sign-On

# STUFENWEISER ROLL-OUT DES DB DATA CATALOG



## ■ PHASE 1

ist der DB Data Catalog eines der essentiellen Enabler-Tools zur Digitalisierung der Bahn und ist **konzernweit** das **Datentransparenzwerkzeug**

## ■ PHASE 2

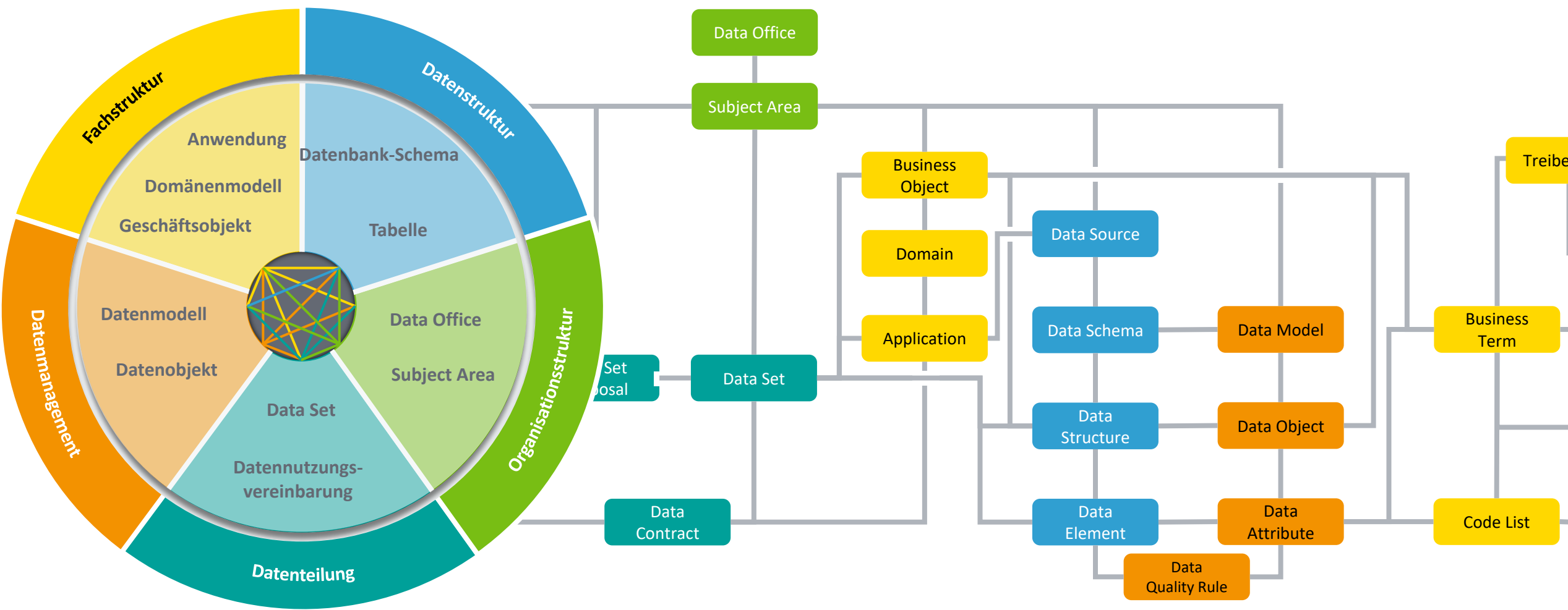
wird der DB Data Catalog das **Cockpit der Datenteilung**. Neben Informationen über Daten werden dort auch die Spielregeln zur Datenteilung transparent und intuitiv nutzbar gemacht sowie die Datenteilung gemanaged

ZIELSETZUNG - IN 3-5 JAHREN WERDEN DATEN KONZERNWEIT INNERHALB VON 25 H VERFÜGBAR SEIN.  
DER DB DATA CATALOG ALS WESENTLICHER ENABLER

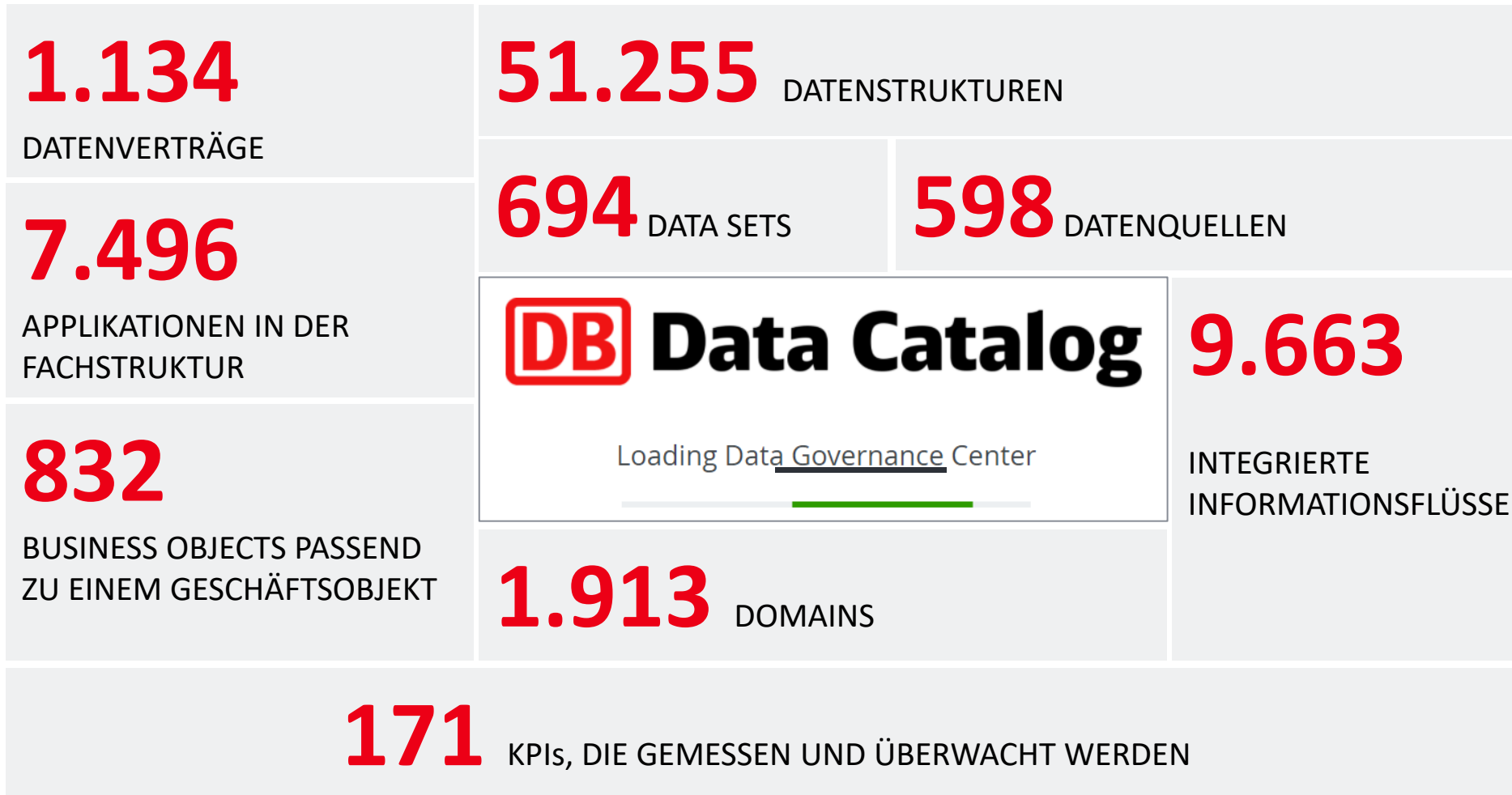
# AUFBAU DES DB DATA CATALOG



TECHNISCHE, FACHLICHE & DATA GOVERNANCE SICHTEN IN RELATION ZUEINANDER



# AKTUELLER STAND DER UMSETZUNG DER DATA GOVERNANCE



Stand Februar 2024

# MEHRWERT DER DATA GOVERNANCE FÜR DB REGIO



Herzlich Willkommen bei  
Data Excellence – DB Regio

- In der **DB Regio** werden täglich Millionen von Daten erzeugt, die das Potenzial haben, Geschäftsprozesse durch ein stärker datengetriebenes Arbeiten effizienter und effektiver zu gestalten.
- Als Hindernis stand allerdings im Raum, dass es keinen vollständigen Überblick über die verfügbaren Daten gibt und die jeweiligen **Ansprechpartner zu den Daten nicht klar definiert** waren. Weiterhin **fehlte ein prozessübergreifendes Monitoring der Datenqualität**, um eine breite Nutzung von qualitativ hochwertigen Daten sicherzustellen und einer **breiten Datenteilung gemäß definierter rechtlicher und ethischer Rahmenbedingungen** zu erlauben.



Effektives **Data Governance-System** sorgt für schlanke Prozesse und erwirkt prozessübergreifendes effizientes Arbeiten mit Daten



Klare Regelungen zum effizienten Umgang mit Daten und breites Bewusstsein für **prozessübergreifende Datenqualität**



**Enablement** des **Data Officers** und **Data Owner** sorgt für Qualität, Transparenz, Verfügbarkeit und Compliance ihrer Datenobjekte



Klare aus der Sachkompetenz abgeleitete Ansprechpartner und Verantwortlichkeiten zur Transparenz und zur Bereitstellung von **qualitätsgesicherten Daten**

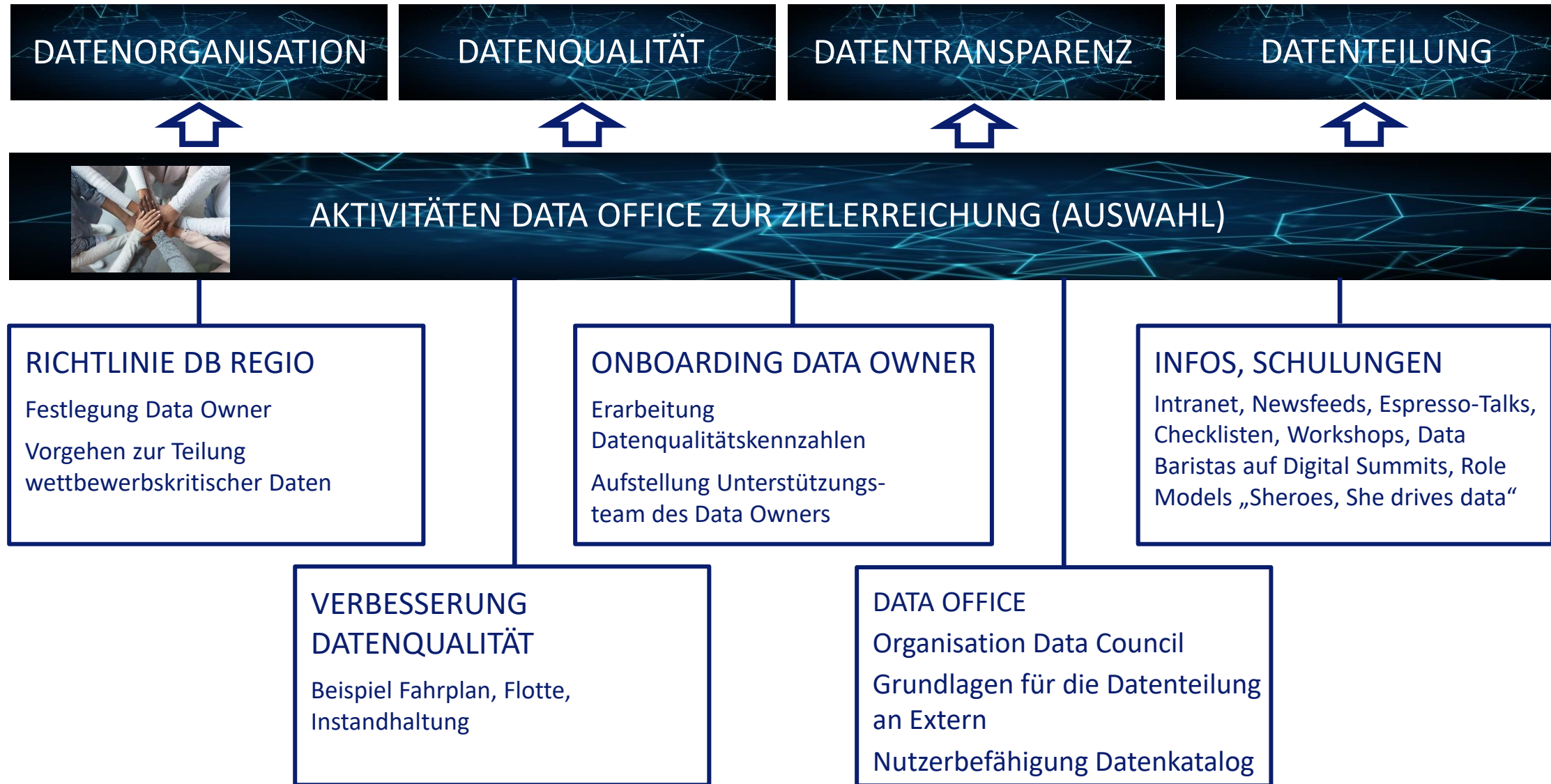


Vertrauensvolle und agile Zusammenarbeit von **Data Governance Rollen und ihrer Gremien** liefert schnelle und sachkompetente Entscheidungen



Sicherung des **rechtssicheren Umgang mit personenbezogenen, vertraulichen oder regulierten Daten** auf Basis abgestimmter Vorgaben

# AKTIVITÄTEN ZUR OPERATIONALISIERUNG BEI DB REGIO



# WORKSHOPBASIERTES VORGEHEN ZU HERAUSFORDERUNGEN DER DATENQUALITÄT



## AKTIVITÄTEN DATA OFFICE ZUR ZIELERREICHUNG – ONBOARDING DATA OWNER



PROZESSBEZOGEN

FOKUSSIERT

INDIVIDUELL



### Prozessorientierte Datenfluss

### Steckbriefe Datenqualität

Beschreibung	Einheit
Messung der Nutzung der Kennzeichnung von Stammdatenbeziehungen durch DB Regio an DB Netz. Die Nutzung der Kennzeichnung sichert eine frühzeitige Erkennung und einen dokumentierten und nachvollziehbaren Kennzeichnungsbedarf.	Prozent Anteil der ausgenutzten Stellungsnummern
<b>Ermittlungslage / Prämissen</b> <b>Berechnung</b> F. abgezogen/Anzahl Stellungsnummern DB Regio/8 maximal verfügbare Stellungsnummern DB Regio/2 z.B. 1 Stellungsnummern pro Maßnahme/Modul <b>Weitere Prämissen</b> Spezifische Stellungsnummern werden elektronisch erfasst und mit Stammdaten (Ereignis, Datum, Stellungsnummer,...) identifiziert	<b>Zielwert</b> 100 % im Erhebungszeitraum
<b>Datenquelle</b> Plato Portal	

### DQ-Konzept

OPN	Erklärung	DQ Dimension
<b>ZyF Rate</b>	Anzahl von Abweichungen der ZyF mit identischer Kennung pro Bezugsgebiet (Anzahl/Anzahl)	Zweitesignale Nachvollziehbarkeit
<b>Falsch Rate</b>	Anzahl von Abweichungen der Falsch mit identischer Kennung pro Bezugsgebiet	
<b>Quote KQ/Bus</b>	Anzahl KQ/Bus Abweichungen je Eisenbahn/Anzahl KQ/Bus Eisenbahn	
<b>Quote ZyF</b>	Anzahl aktiver ZyF Nachlieferungen/Anzahl aktiver Datensätze/100	
<b>Fehlerquote KQ/Bus ZyF</b>	Anzahl ZyF Einträge mit nicht belassenen Modulen/Anzahl ZyF Einträge pro Eisenbahn	
<b>Fehlerquote KQ/Bus ZyF</b>	Anzahl ZyF mit Verletzungen bei der Gültigkeitsregel zur Korrektheit von Zugnummern/Anzahl ZyF in Summe pro Bezugsgebiet/100	
<b>Fehlerquote Vergleichswerte ZyF</b>	Anzahl ZyF mit Verletzungen bei der Gültigkeitsregel zur Korrektheit der Werte/Anzahl ZyF in Summe pro Bezugsgebiet/100	
<b>Fehlerquote Informationsstruktur Falsch</b>	Anzahl Falsch mit Verletzungen bei der Gültigkeitsregel zur Nutzung der Falsch Informationsstruktur/Anzahl Falsch in Summe pro Bezugsgebiet/100	
<b>Auslastungs KQ/Bus</b>	Anzahl von zurückgenommenen Maßnahmen/Anzahl KQ/Bus Eisenbahn je Eisenbahn	
<b>Inkonsistenzquote ZyF</b>	Anzahl ZyF mit Inkonsistenzen zur letzten Version der ZyF/Anzahl ZyF mit gültigen Verletzungen bei der Gültigkeitsregel/Anzahl ZyF mit gültigen Verletzungen bei der Gültigkeitsregel/Anzahl ZyF mit gültigen Verletzungen bei der Gültigkeitsregel/100	
<b>Unvollständigkeit</b>		

### Output des Analyseprozesses

#### DG Organisation

**Data Quality Gruppe Nordost (Regelabstimmung)**

DQ Issues	DQ ZyF	Einschränkungen	Umsetzung
<b>Data Owner</b> Lutz Wille DB Regio Region Nordost (Bauplanung)	<b>(Delegated) Data Owner</b> Thomas Lohr DB Regio Region Nordost (Verkehrsvorgang)	<b>Data Custodian</b> Dimitri Stelzer DB Regio Region Nordost (System Master)	<b>Data Custodian</b> Alex Bensch DB Regio Region Nordost (System Master)
<b>Data Steward</b> Bergt Stefan DB Regio Region Nordost (Bauplanung)	<b>Data Steward</b> Ralf Kuhn DB Regio Region Nordost (Verkehrsvorgang)	<b>Data Steward</b> Bergt Stefan DB Regio Region Nordost (Angebotplaner)	<b>Data Steward</b> Jens Hammer DB Regio Region Nordost (Angebotplaner)
<b>Key User</b> DB Regio Region Nordost (z.B. Rolle Abrechnung)	<b>Key User</b> DB Regio Region Nordost (z.B. Rolle Abrechnung)		

#### Meine Data Story

1. Grüne Mobilität erfordert gesunden Daten
2. Data Governance Initiative bei der DB und der DB Regio
3. **Prozessorientierte Data Governance und Framework für Datenqualitätsanalysen**
4. Lessons Learned und Best Practices



Der **Data Owner (DO)** ist eine **Management-Rolle** in der Data Governance Struktur der DB Regio. Bei dieser Rolle liegt die **Verantwortung** für die **Transparenz, Qualität, Verfügbarkeit und Compliance (Datenteilung) der Datenobjekte**, die der Rolle zugeordnet sind.



Der Data Owner ist der **Leiter eines virtuellen Teams** aus Data Stewards und Data Custodians und **Eskalationsinstanz** bei Themen, die seine Datenobjektive betreffen.

Um die **Verantwortung** für den Fluss der Daten entlang der **zentralen Fachprozesse** sicherzustellen, erfolgt die Festlegung der **Data Owner** im **Geschäftsfeld DB Regio i.d.R. prozessorientiert**.

Die identifizierten Data Owner decken sich i.d.R. mit den zuständigen **Leitern der OrgEinheiten**, die den **jeweiligen Prozess bei DB Regio** verantworten.

Der Data Owner kann innerhalb der DB Regio die Durchführung von Regelaufgaben innerhalb seiner **Data Ownership** an einen **Deputy Data Owner** (responsible, nicht accountable) übertragen.

Bei Freigaben von **wettbewerbskritischen Daten** (z.B. Relevanz für Ausschreibungs-kalkulationen) muss auch eine regionale Freigabe durch den Vorsitz der **Regionalleitung** eingeholt werden.

Der **Data Steward (DS)** ist die **fachliche Umsetzungsrolle** in der Data Governance Struktur der DB Regio und ist Teil des (virtuellen) Teams des Data Owners. Ein Data Owner kann mehrere Data Stewards für die ihm zugeordneten Datenobjekte ernennen.



Der Data Steward ist Teil des (virtuellen) Teams des Data Owners und diesem fachlich (und ggf. auch disziplinarisch) unterstellt. Er ist zuständig für die **Erreichung der vom Data Owner ausgegebenen Ziele** und die **Umsetzung der mit ihm entwickelten (Datenqualitäts-)Maßnahmen**.

Aufgrund der dezentralen Organisation von DB Regio sind die Data Stewards in den Regionen **Führungskräfte** (speziell Vertreter von Finanz & Controlling- und Betriebs-Funktionen) bzw. Führungskräfte, die entlang des **jeweiligen Prozesses** beteiligt sind.

Bei der DB Regio sind die Data Stewards von der **unmittelbaren Pflege des Datenkatalogs** entbunden.

Der Data Steward führt in seinem **Verantwortungsbereich Mitarbeiter**, die diverse datenrelevante Anwendungen nutzen. Seine Mitarbeiter sind im **operativen Prozess für Dateneingaben/-pflege/-prüfung** zuständig und nehmen häufig die Rolle eines **Key-Users** ein.

Der **Data Custodian (DC)** ist die **technische Umsetzungsrolle** in der Data Governance Struktur der DB Regio und ist Teil des (virtuellen) Teams des Data Owners. Ein Data Owner kann mehrere Data Custodians für die ihm zugeordneten Datenobjekte ernennen.



Der Data Custodian ist Teil des (virtuellen) Teams des Data Owners und diesem fachlich unterstellt. Er ist **Hauptansprechpartner** für den Data Owner im **CIO-Bereich** oder unter den Verfahrensverantwortlichen in der DB Regio.

Data Custodian bei der DB Regio sind in der Regel die **Verfahrensverantwortlichen** der **IT-Verfahren** der DB Regio.

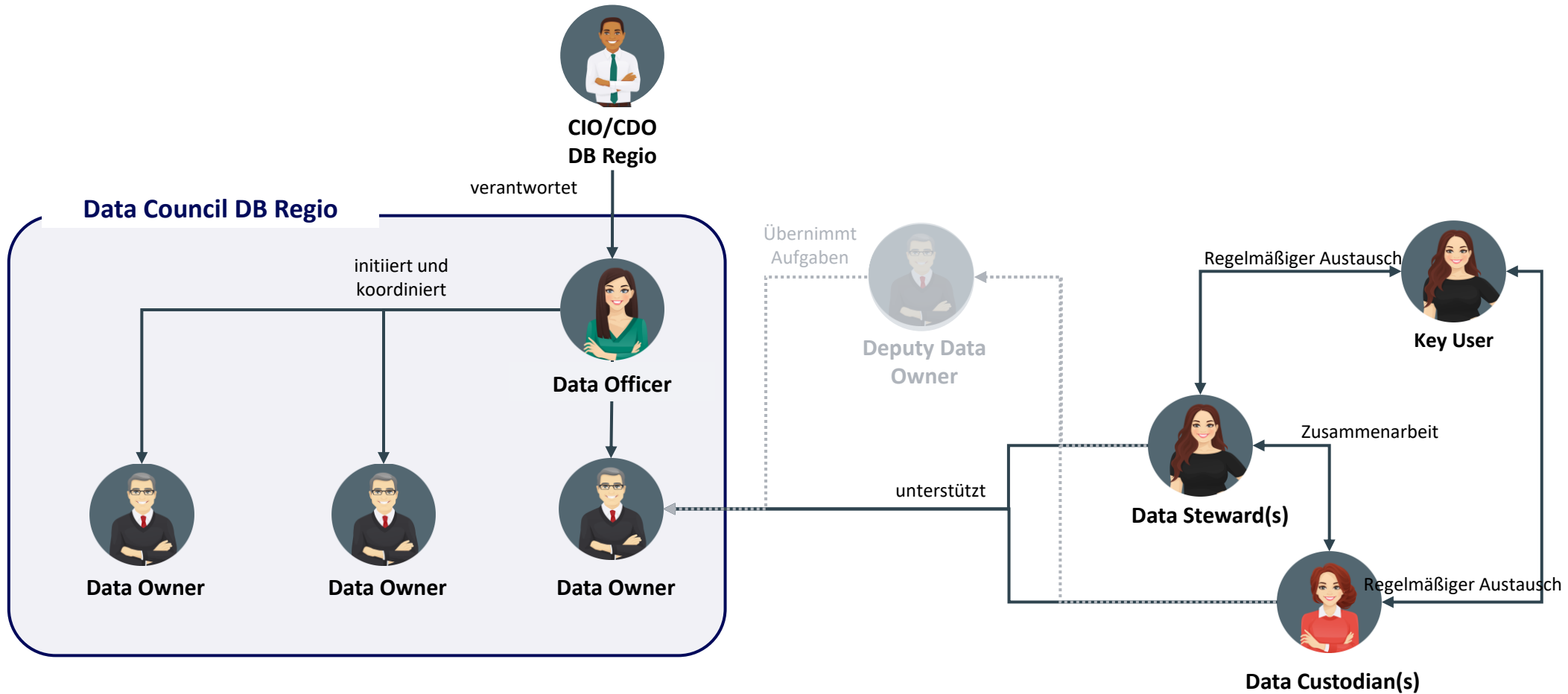
Data Custodian bei der DB Regio sind für technische **Implementierung eines Datenqualitätsmonitoring** zuständig, um Datenqualitätsmaßnahmen zu evaluieren.

Data Custodian bei der DB Regio sind für die **Integration der Datenquellen in den DB Datenkatalog** verantwortlich.

Data Custodian bei der DB Regio sind in der Rolle als Anwendungsverantwortliche für die **Implementierung und Nachhalten eines spezifischen Berechtigungs-konzepts** für die Datennutzung der Daten in den Anwendungen verantwortlich.

# Das Rollenkonzept und Ihre Zusammenarbeit im Data Council DB Regio

## Data Owner und das Unterstützungsteam



Weitere Arbeitsgruppen z.B. zu den Themen Datenqualität oder Datenteilung entstehen je nach Datendomäne unter Führung der beteiligten Data Owner mit den Data Stewards, Data Custodian, Key Users etc.

Auf Basis einer Fachprozessanalyse wird der Datenfluss entlang der beteiligten Systeme und Prozesse erhoben und entsprechende Datenqualitätsprobleme identifiziert. Dazu wird ein passendes Data Quality Konzept erarbeitet und dazu wird wiederum die passende Data Governance aufgesetzt



- Einschätzung **Ausgangssituation Data Quality und Data Governance** aufstellen
- **Fachprozessanalyse** mit Aufnahme der zentralen Prozessbeteiligten und Prozessschnittstellen
- **Verständnis** für Zusammenhang von **Datenqualitäts- und operativen Problemen** aufbauen
- Erhebung **DQ Konsequenzen für DQ Bus. Case**
- Erhebung **IT-Systemeinsatz** im Fachprozess
- Ableitung **Ist-Datenfluss** aus Fachprozess mit den wesentlichen Datenobjekten (Grobdarstellung und Detail)
- **Identifikation** von Datenqualitätsproblemen im Fachprozess
- **Verständnis** für Zusammenhang von **Datenqualitäts- und operativen Problemen** aufbauen
- Erhebung **DQ Konsequenzen für DQ Bus. Case**
- Festlegung eines **Datenqualitätsziels für geschäftskritische Daten**
- **Prozessspezifische Quality Performance Indicators (QPIs)** für die Datenqualität definieren
- **Messpunkte** zur Datenqualität in **Fachprozesse** integrieren
- **Bewertungs- und Entscheidungsprozesse** zur Datenqualität definieren
- **Regelmäßige und automatisiertes DQ-Monitoring** aufbauen
- **DG Gestaltungsgrundsätze** aufstellen
- **DG Aufgaben und Verantwortlichkeiten** der Datenrollen auf **GF spezifische Belange anpassen**
- **Ablauforganisation mit Gremien & Abstimmungsprozessen** konzipieren
- Zusammenarbeit **Data Owner mit den Data Managern vor Ort** festlegen und verproben
- Liste **potenzieller Rolleninhaber** erstellen und **Besetzung abstimmen**
- **Befähigungsmaßnahmen** definieren und einleiten

# Exemplarische Ergebnisse aus dem Vorgehens Framework zur Analyse von DQ Herausforderungen in den Fachprozessen



## Fachprozessanalyse: Identifikation & Analyse Datenqualitätsprobleme

**Vielzahl von unterschiedlichen Datenqualitätsproblemen wurden identifiziert**

Im Rahmen der Begleitung der WS zur Prozessanalyse sowie verschiedener Experteninterviews konnten zahlreiche Daten der jeweiligen Prozessphasen identifiziert und beschrieben werden. Einige von ihnen lassen sich direkt in eine IT-Anforderung übersetzen.

Datenqualitätsdimension	(1) Bewerte Auswirkungen aus Baumaßnahmen aufbauend	(2) Phase Fahrplan für kommende Fahrplanperiode	(3) Phase unterjähriger Fahrplan	(4) Nachbereitung	(5) U
Aktualität/Zeitraum			1		1
Genauigkeit				1	1
Vollständigkeit	2		1		
Konsistenz	2	3	2		
Korrektheit/Fehlerfreiheit			4	1	
Redundanzfreiheit			4		
Relevanz	1	1	1		
Verständlichkeit			1		
Verfügbarkeit	1	2	2	1	
Verlässlichkeit				3	1
<b>Gesamtergebnis</b>	<b>7</b>	<b>9</b>	<b>18</b>	<b>6</b>	<b>6</b>

DB Regio AG | Bau & Regio | Teilprojekt Datenqualität

---

**Prozess „Auswirkungen aus Integrierter Bündelung und Baukapazitätsmanagement bewerten“ [x-26 bis x-17]**

Der Prozess erfährt seitens DB Regio keine systemische Unterstützung und ist gekennzeichnet von vielen manuellen Arbeitsschritten. Allerdings sehen in der frühen Phase (FPL-Wechsel – 26 Monate bis – 17 Monate) die Regionen keinen unmittelbaren, technischen Handlungsbedarf.

Prozess (VORBEREITUNG)	Sicht Baumaßnahmen aus integrierter Bündelung	Sicht Auswirkungen in Abteilungsgruppen (in einem)	Informationen zu Großveranstaltungen	Bewerte Baumaßnahmen aus integrierter Bündelung
R-Systeme	MS Portal der DB Netz AG	MS Excel-Templates der DB Netz AG	MS Excel-Templates der DB Netz AG	MS Portal der DB Netz AG
Information / Formate	<ul style="list-style-type: none"> <li>Standardisierungsrichtlinien (u.a. VOP, Servicekennzeichen, BOP, Durchlaufzeitplan, Betriebsstellen)</li> <li>erweiterte Auswertungen auf Mehrfachnutzung</li> <li>Werkzeuge (z.B. Excel, Word)</li> </ul>	<ul style="list-style-type: none"> <li>Sichtungen der betroffenen DB Netz Fahrpläne</li> <li>Erweiterte Besuchszeit</li> <li>Sonderverkehrs</li> </ul>	<ul style="list-style-type: none"> <li>u.a.</li> <li>bestmögliche</li> <li>Erweiterte Besuchszeit</li> <li>Sonderverkehrs</li> </ul>	<ul style="list-style-type: none"> <li>Individuelle Sichtungen zur integrierten Bündelung (DB des EV/Busbereichs, Einreisereiseleistungen und Umleitungen)</li> </ul>
DQ Probleme	<ul style="list-style-type: none"> <li>Kein Regio-System der DB Netz, um relevant integrierte Informationen in PPT in relevanten Dimensionen, wenig standardisiert und nicht auswertbar</li> <li>Unstrukturiertes Dokument zum Download einer EV-spezifische Datenlieferung</li> <li>Detailtiefe nicht gegeben für präzisere Bewertung / Planung</li> <li>Informationsverlust bei MFI-Budget und -Planung im Bereich SEV-Budget</li> </ul>	<ul style="list-style-type: none"> <li>Informationen in Fileformat (Dateien, wenig standardisiert und nicht auswertbar)</li> </ul>	<ul style="list-style-type: none"> <li>Mehrereinstufige (Pilot) absehbare Datenlieferung</li> </ul>	<ul style="list-style-type: none"> <li>Eingabe als Freitext z.B. die Bitte, Maßnahme eintragung, zu behaltene Informationen, wenig standardisiert und nicht auswertbar</li> <li>Manuelle Übertrag von Sichtungen (Informationen aus vorherigen Prozessschritten)</li> <li>Nicht die Informationen werden nachgelagert neue Bündelung</li> <li>Keine Informationen zur Bewertung / Berücksichtigung, sondern DB Netz</li> </ul>

DB Regio AG | Bau & Regio | Teilprojekt Datenqualität

## Datenfluss: Systembeteiligung & Schnittstellen und relevante Dateninhalte

**Übersicht wesentliche Systeme im Gesamtprozess Baumanagement**

**Detaildatenfluss: Auswirkungen aus Integrierter Bündelung und Baukapazitätsmanagement bewerten**

DB Regio | Referenzmodell Baumanagement v0.11 | P.1008 | S.17 | 03.11.2021

DB Masterfirma | Vorname Nachname | Abteilung | 01.11.2021

## Datenorganisation: Data Governance Blaupause für Aufbau- & Ablauforga

**Der Data Owner hat eine hohe Affinität zu Daten und ist Ansprechpartner für Datenweitergabe, Optimierung der Datenqualität und Datenstruktur**

**Data Owner**  
Der Data Owner ist Ansprechpartner für die Datenweitergabe, übernimmt eine Management-Rolle zur Optimierung der Datenqualität und verantwortet die Datenstruktur seiner Daten zur Steigerung der Datenqualität und -verfügbarkeit.

**Grundlegende Aufgaben**

- Steuerung von Anfragen zur Beachtung des Schutzbedarfs der Daten
- Abstimmung und Entwicklung von Datenqualitätsmaßnahmen
- Verantwortung der Datenstruktur der Daten und der Datenqualität

**Typische Fragen an einen Data Owner**

- Ist es möglich, Zugriff auf die Daten zu bekommen?
- Wie sollen Daten nach Schutzbedarf, Vertraulichkeit und Personenbezug klassifiziert werden?
- Wer kann bzgl. fachlichen oder technischen Fragen zu den Datenquelle angesprochen werden?
- Wie sehen die Qualitätskennzahlen für die Daten aus?
- Wie kann die konkrete Datennutzung und Datenverfügbarkeit gesteigert werden?

**Wichtige Eigenschaften**

- Entscheidungskompetenz zu Prinzipal, Datenhaltung, Datenqualität, Datenstruktur
- Affinität zu den Daten mit hohem Business-Nutzen der Daten sowie
- Berennung und Führung von fachlichen Ansprechpartnern mit Verantwortung der Daten

**Deep Dive: Zuordnung operativer Tätigkeiten zu grundlegenden Aufgaben eines Data Owners**

- Aufgabe** Steuerung von Anfragen zur Datenhaltung unter Beachtung des Schutzbedarfs der Daten
- Aufgabe** Abstimmung und Entwicklung von Maßnahmen zur Steigerung der Datenqualität
- Aufgabe** Verantwortung der Datenstruktur zur Erleichterung der Datennutzung und der Datenstruktur

**Operative Tätigkeiten**

- Dokumentation von Anfragen zur Datenweitergabe
- Zuordnung von Daten zu Datenqualitätsmaßnahmen
- Prüfung und Umsetzung technischer Rahmenbedingungen und Vorgaben im Zusammenhang mit Daten, deren Schutz und ihrer Weitergabe
- Etablierung von Policies und Richtlinien zur Datenethik und -compliance

- Durchführung von Data Profiling der Datenkatalog angebotenen Data Source
- Entwicklung und Umsetzung von Maßnahmen zur Steigerung der Qualität der Metadaten im Datenkatalog
- Messung und Berichterstattung von KPIs zum Status der Anbindung an den Datenkatalog

- Aktualisieren von Kataloginhalten einer Datenquelle, dazu gehören das Anlegen, die Pflege und ggf. die Löschung von Data Objects sowie Data Sets
- Leitung Virtuellet Teams bestehend aus Data Stewards und Custodians

DB Regio | Tatja Schlesinger/Carsten Dittmar | Baumanagement@regio | 13.01.2022

(1) Datenqualitäts sind gruppiert nach Schutzbedarf, u.a. Open Data, DB Intern, DB vertraulich, Pers.-bez. Daten, etc.

DB Masterfirma | Vorname Nachname | Abteilung | 01.11.2021

## DQ Konzept: QPI zur Messung von Datenqualität

**Quality Performance Indikatoren erlauben ein Monitoring von identifizierten DQ Issues (1/2)**

Die vorgeschlagenen QPIs sollten sich multidimensional mit Bezug zu den Dimensionen Zeit (Fahrplanperiode, Monat), Strecken, Bauvorhaben, Baumaßnahmetyp gruppieren. Die Zuordnung der QPI an Messpunkte sowie die Präzisierung der Gültigkeitsregeln sind abhängig von der Gestaltung der zukünftigen IT Unterstützung.

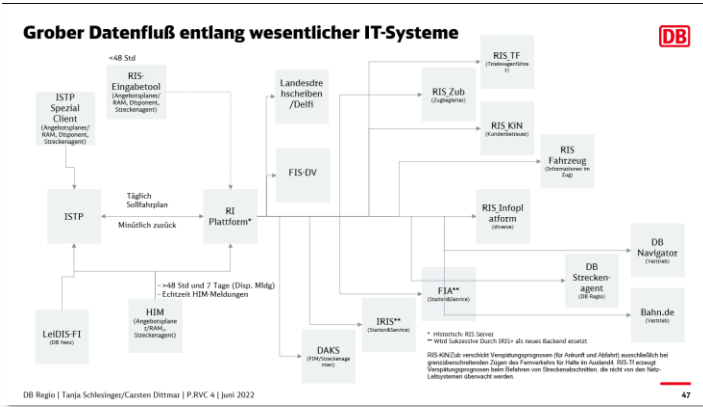
QPI	Erklärung
<b>Nachfragequalität</b>	Anzahl der eingereichten Subanfragen pro Prozessschritt (z.B. Baugewerbliche Subanfragen pro Prozessschritt)
<b>Beurteilungsgeschwindigkeit</b>	Anzahl der eingereichten Subanfragen pro Prozessschritt (z.B. Baugewerbliche Subanfragen pro Prozessschritt)
<b>Gleichung</b>	Anzahl der eingereichten Subanfragen pro Prozessschritt mit 'Warten' Status pro Baugewerbliche (z.B. Baugewerbliche Subanfragen)
<b>Anzahl aktive Zf ohne korrekte Historie</b>	Anzahl aktive Zf mit Historie/Status/Anzahl aktive Zf pro Baugewerbliche (z.B. Baugewerbliche Subanfragen)
<b>Anzahl aktive Zf ohne korrekte Historie</b>	Anzahl aktive Zf mit Historie/Status/Anzahl aktive Zf pro Baugewerbliche (z.B. Baugewerbliche Subanfragen)
<b>Frühzeitige Schließung</b>	Anzahl zu später Umstellungen gemäß Tarifvertrag und Betriebsvereinbarung bzw. RL Vorgaben für Schließung/Anzahl der Umstellungen (z.B. Baugewerbliche Subanfragen)
<b>Frühzeitige Fpl</b>	Anzahl zu später (unvollständiger) Zf gemäß RL Vorgaben für Umstellung/Anzahl der Zf für Umstellung pro Baugewerbliche (z.B. Baugewerbliche Subanfragen)
<b>Frühzeitige Zf</b>	Anzahl zu später (unvollständiger) Zf gemäß RL Vorgaben für Umstellung/Anzahl der Zf für Umstellung pro Baugewerbliche (z.B. Baugewerbliche Subanfragen)
<b>Abschreibungen</b>	Anzahl der Abschied mit unterschiedlichen Baumaßnahmen Zuordnung in zwei aufeinanderfolgenden Prozessschritten/Anzahl der Abschied pro Baugewerbliche

**Kriteriensteckbrief – Datenqualität Bauinformationen Fristverzug Zf-Entwurf & Endstück**

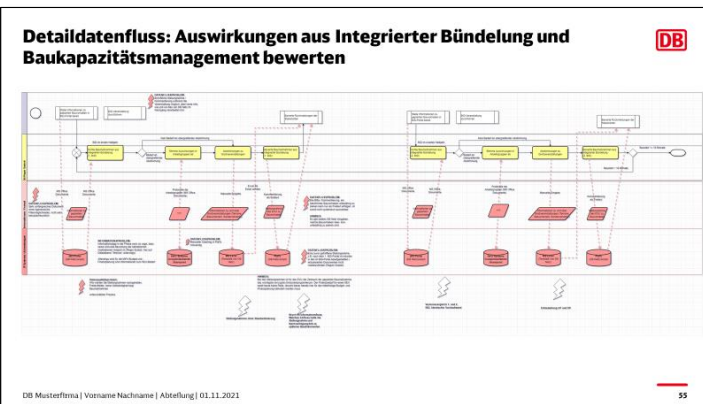
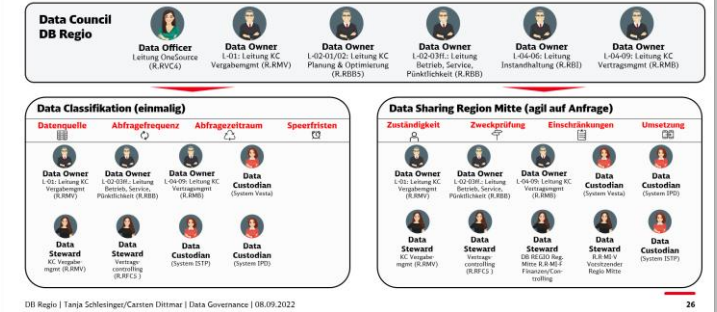
Beschreibung	Einheit	Drilldown
Monitoring der fristgerechten Zustellung der Zf durch DB Netz an DB Regio in den Erhebungsgraden Entwurf und Endstück. Die Fristenhaltung stellt die planmäßige Prüfung der Zf entlang der regio-internen Kommunikation sicher.	Prozent	<ul style="list-style-type: none"> <li>Fahrplanperiode, Monat</li> <li>Prozessschritt Baugewerbliche</li> <li>Region, Strecken</li> <li>Bauvorhaben</li> <li>Baumaßnahme (A-Maßnahmen (Entwurf/Endstück), B-Maßnahmen, F-Maßnahme)</li> </ul>
<b>Ermittlungslage / Prämissen</b>	<b>Zielwert</b>	
<b>Berechnung</b> # Zf mit verspätetem Eingang (Datum Beginn Baumaßnahme - Datum Eingang Zf Entwurf o. Endstück) > 1 Tag / # Zf im Erhebungszeitraum	0 % im Erhebungszeitraum Grenzwerte Fristen: – 24 Wochen vor Baubeginn – Zf-Entwurf (A) – 13 Wochen vor Baubeginn – Zf-Endstück (A) – 10 Wochen vor Baubeginn – Zf (B)	
<b>Weitere Prämissen</b> – Berechnung der Fristabweichung auf Basis der Datumswerte der jeweiligen Version 1		
<b>Datenquelle</b>	<b>Erhebungsfrequenz</b>	
Max/Bau Datenbank	– täglich	

DB Regio AG | Bau & Regio | Teilprojekt Datenqualität

# Der Data Owner wird in seinen ersten Aufgaben aktiv begleitet und bekommt SEINE jeweilige Data Story als Ergebnis der DQ Analyse



## Zweistufige Entscheidungsebenen zur Klassifikation und Freigabe von wettbewerbskritischen Daten



## Data Story für Verbundprozess RI

- (1) Ziele der Data Ownerschaft
- (2) Relevante Datenobjekte und Datenflüsse
- (3) Data Ownerschaft Team
- (4) Gremien und Abstimmungen
- (5) Initiale Datenqualitätskriterien
- (6) Initiale Datenklassifikation
- (7) Nächste Maßnahmen

10/22 | R.RMP 3

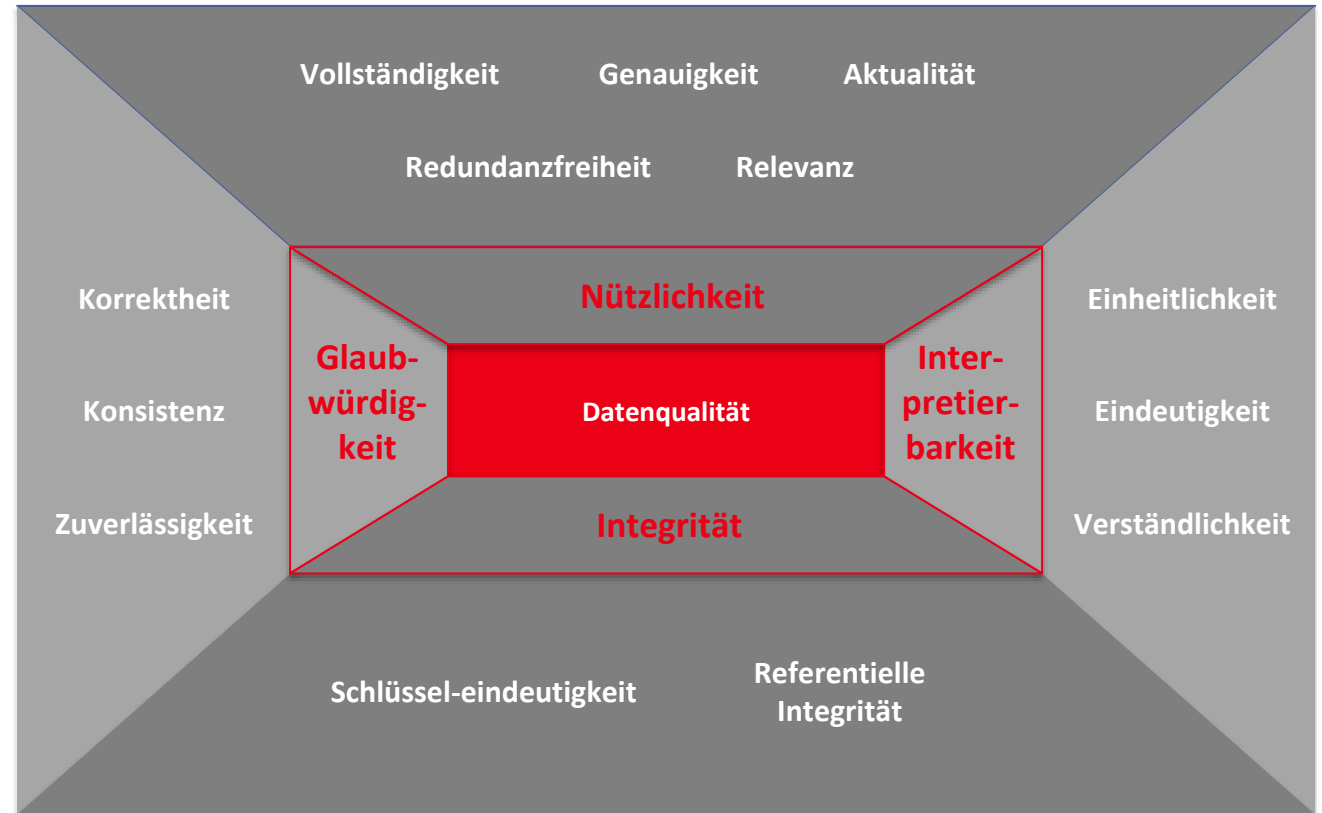
Kriteriensteckbrief – Datenqualität Bauinformationen Fristverzug ZvF-Entwurf & Endstück		
<b>Beschreibung</b>	<b>Einheit</b>	<b>Drilldown</b>
Monitoring der fristgerechten Zustellung der ZvF durch DB Netz an DB Regio in den Erarbeitungsgraden Entwurf und Endstück. Die Fristeinholung stellt die planmäßige Prüfung der ZvF entlang der regio-internen Terminkette sicher	Prozent	<ul style="list-style-type: none"> <li>Fahrplanperiode, Monat</li> <li>Prozessschritt: Baufahrplan</li> <li>Region, Strecken</li> <li>Bauvorhaben</li> <li>Baumaßnahme (A-Maßnahmen (Entwurf/Endstück), B-Maßnahmen, F-Maßnahmen)</li> </ul>
<b>Ermittlungslogik / Prämissen</b>	<b>Zielwert</b>	
<p><b>Berechnung</b></p> $\# ZvF \text{ mit verspätetem Eingang} / (\text{Datum Beginn Baumaßnahme} - \text{Datum Eingang ZvF Entwurf o. Endstück}) \times \text{Tag} / \# ZvF \text{ im Erhebungszeitraum}$ <p><b>Weitere Prämissen</b></p> <ul style="list-style-type: none"> <li>Berechnung der Fristabweichung auf Basis der Datumswerte der jeweiligen Version 1</li> </ul>	<ul style="list-style-type: none"> <li>0 % im Erhebungszeitraum</li> <li>Grenzwerte Fristen</li> <li>24 Wochen vor Baubeginn – ZvF-Entwurf (A)</li> <li>19 Wochen vor Baubeginn – ZvF-Endstück (A)</li> <li>10 Wochen vor Baubeginn – ZvF (B)</li> </ul>	
<b>Datenquelle</b>	<b>Erhebungsfrequenz</b>	
MaxBau Datenbank	Täglich	

Es gibt nicht DIE Datenqualität für jeden Prozessbeteiligten, sondern je nach DQ-Dimension sind Anforderungen unterschiedlich

> **Qualität (DIN/ ISO 9904):**  
 Qualität ist die Gesamtheit von **Eigenschaften und Merkmalen** eines Produktes oder einer Dienstleistung, die sich auf deren **Eignung zur Erfüllung** festgelegter oder vorausgesetzter **Erfordernisse** bezieht.

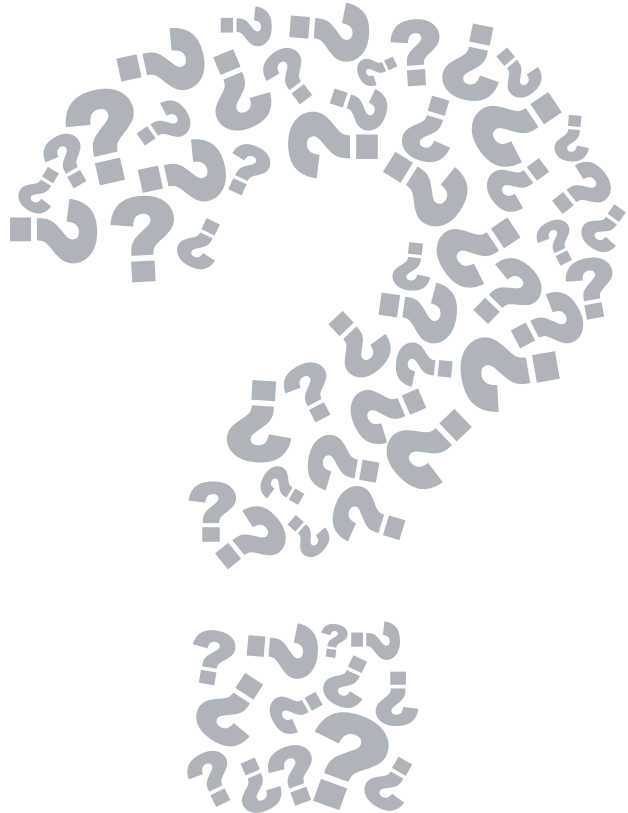


> **Datenqualität:**  
 Die Gesamtheit aller Eigenschaften von Daten hinsichtlich der Fähigkeit, den an ihre Erfassung/Generierung gebundenen Zweck zu erfüllen bzw. die Anforderungen des Anwenders zu erfüllen („meeting user requirements“ / „fit for purpose“).





DQ Kriterium	Definition
<b>Korrektheit/ Fehlerfreiheit</b>	Die Attributwerte eines Datensatzes entsprechen denen der modellierten Objekten der realen Welt, d.h. die Daten stimmen mit der Realität überein.
<b>Konsistenz</b>	Die Attributwerte eines Datensatzes weisen keine logischen Widersprüche untereinander oder zu anderen Datensätzen (in anderen Systemen) auf.
<b>Zuverlässigkeit/ Nachvollziehbarkeit</b>	Die Attributwerte eines Datensatzes sind vertrauenswürdig, d.h. die Entstehung der Daten in ihren Versionen und Zwischenschritten ist nachvollziehbar und zuverlässig.
<b>Vollständigkeit</b>	Vollständigkeit beschreibt die generelle Verfügbarkeit von Daten, die der Anwender benötigt, um seine Arbeit überhaupt durchführen zu können. Alle wichtigen Daten sind in einem System enthalten.
<b>Genauigkeit</b>	Abhängig vom jeweiligen Kontext liegen die Daten in der geforderten Genauigkeit (z.B. Anzahl Nachkommastellen) vor.
<b>Aktualität/ Zeitnähe</b>	Alle Datensätze entsprechen jeweils dem aktuellen Zustand der modellierten Welt und sind damit nicht veraltet. Die Daten bilden die tatsächlichen Eigenschaften des Objekts zeitnah ab.
<b>Redundanzfreiheit</b>	Innerhalb eines Systems dürfen keine Duplikate vorkommen. Als Duplikate werden Datensätze verstanden, die das selbe Objekte der realen Welt beschreiben, die aber nicht unbedingt in allen Attributwerten übereinstimmen.
<b>Relevanz</b>	Das Informationsangebot der Daten zu einem realem Objekt bezüglich eines definierten Anwendungskontextes deckt sich mit dem Informationsbedarf des Anwenders.
<b>Verständlichkeit</b>	Die Datensätze stimmen in ihrer Begrifflichkeit und Struktur mit den Vorstellungen der Anwender überein. Die abgebildete Fachlichkeit ist für den Anwender verständlich.
<b>Integrität</b>	Die Datensätze können über Schlüsselfelder eindeutig mit anderen Datenquellen integriert werden, da sie auf der gleichen Systematik von Stammdaten basieren.



### Schritt 1: Identifikation & Beschreibung

- 1 Welches **DQ-Problem** unterliegt der Prozessherausforderung?
- 2 Welche **DQ-Dimension** ist betroffen?
- 3 Wie lässt sich das DQ-Problem **feststellen**?
- 4 Wann im Prozess oder in nachgelagerten Prozessen **fällt das DQ-Problem auf bzw. wird es geprüft**?

### Schritt 2: Bewertung / Business Case

- 1 Welche **negative Fehlerkonsequenz** resultiert aus dem DQ-Problem?
- 2 Welches **Ausmaß** hat die negative Fehlerkonsequenz?
- 3 In welcher **Häufigkeit** tritt das DQ-Problem auf?
- 4 Welcher **Aufwand** fällt zur **Fehleranalyse** an?
- 5 Welcher **Aufwand** fällt zur **Fehlerkorrektur** an?

# Schritt 1: Identifikation & Beschreibung



Prozess: L-02-04-01 Bedarf an Fahrzeuginstandhaltung planen



## Prozessherausforderung: Unvollständige Zeitangabe Haltepunkterreichung (BUS)

1. Beschreibung	<ul style="list-style-type: none"><li>▪ Unterschiedliche Systeme im Bus zur Protokollierung von Fahrereignissen sowie Unterschiedlich geschulte Fahrer</li><li>▪ Kein Systematisches Stammdaten und Datenqualitätsmanagement und viele unterschiedliche dezentrale Systeme</li><li>▪ Haltepunkterreichung nicht ausgefüllt</li></ul>
2. DQ-Dimension	<ul style="list-style-type: none"><li>▪ Vollständigkeit</li></ul>
3. DQ-Regeln zur Erkennung	<ul style="list-style-type: none"><li>▪ Feld Ankunft leer oder keine chronologische Folge entlang der Haltepunkte</li></ul>
4. Bewusstwerden bzw. Existenz von definierten Prüfschritten	<ul style="list-style-type: none"><li>▪ Erst bei Verkehrsleistungsabrechnung gegenüber Aufgabenträger erfolgt Prüfung</li><li>▪ Höchst individuell je nach Region</li></ul>
Prozessschritt	<ul style="list-style-type: none"><li>▪ 40</li></ul>
Bezug Datenobjekt	<ul style="list-style-type: none"><li>▪ Fahrzeugdaten</li></ul>
Bezug IT-Systeme	<ul style="list-style-type: none"><li>▪ Fahrzeugdatenbank</li></ul>

# Schritt 2:

## Bewertung / Business Case



Prozess: L-02-04-01 Bedarf an Fahrzeuginstandhaltung planen

### Prozessherausforderung: Unvollständige Zeitangabe Haltepunkterreichung (BUS)

1. Negative Fehlerkonsequenz (Wo)	<ul style="list-style-type: none"> <li>Zu Unrecht bezahlte Pönalzahlungen</li> </ul>
2. Ausmaß der negativen Fehlerkonsequenz	<ul style="list-style-type: none"> <li>Ca. 10.000 EUR / Woche</li> </ul>
3. Häufigkeit des DQ-Problems	<ul style="list-style-type: none"> <li>Sehr regelmäßig in einigen Regionen</li> <li>Kein prozessuales Problem, sondern eher Unkenntnis bei Busfahrer über die Relevanz der Daten</li> </ul>
4. (Regelmäßiger) Aufwand zur Fehleranalyse (bei Wem)	<ul style="list-style-type: none"> <li>Quartalsweise bei der Verkehrsleistungsabrechnung gegenüber Aufgabenträger</li> <li>Eintägiger Prüfaufwand</li> </ul>
5. Aufwand zur Fehlerkorrektur (bei Wem)	<ul style="list-style-type: none"> <li>Rückrechnung aus Endpunkt Erreichung</li> <li>Quartalsweise eintägiger Korrekturaufwand</li> </ul>
Gesamtaufwand	<p>Im Quartal:</p> <p>Pönalzahlungen: 13x10.000 EUR=<b>130.000 EUR</b></p> <p>Suchaufwand &amp; Fehlerkorrektur: 1+1 PT=<b>2 PT</b></p>

1. Grüne Mobilität erfordert gesunden Daten
2. Data Governance Initiative bei der DB und der DB Regio
3. Prozessorientierte Data Governance und Framework für Datenqualitätsanalysen
4. **Lessons Learned und Best Practices**

