

# DATA ANALYTICS 2018

19. & 20. März 2018 im Phantasialand Brühl

EINE GEMEINSAME KONFERENZ VON DOAG UND ORACLE

**DOAG ORACLE®**

## Datenqualitäts-Cockpit (Data Quality Result Mart) zur Analyse und Steuerung der Datenqualität auf Basis eines DWH nach Data Vault 2.0

Christiane Breuer  
Consulting Manager  
DATA MART Consulting GmbH



## I N H A L T E

- **Warum Datenqualitätsmanagement**
- **Beispiel einer Implementierung auf einem DWH (Oracle 12c / Data Vault 2.0)**
- **Beispiel einer Visualisierung mit Oracle BI-EE**

- Unabhängiges Beratungs- und Implementierungshaus
- Spezialisiert auf Data Warehouse- / BI- Systeme und Internationale Konzernrechnungslegung
- Verschiedenste Branchenerfahrungen und Technologien
- Über 300 Mannjahre Projekterfahrung
- 2017: Seit 20 Jahren erfolgreich am Markt!



Niederlassung Hamburg  
Behringstraße 16 b  
22765 Hamburg  
Tel: 040 - 359 661120  
Fax: 040 - 359 661199

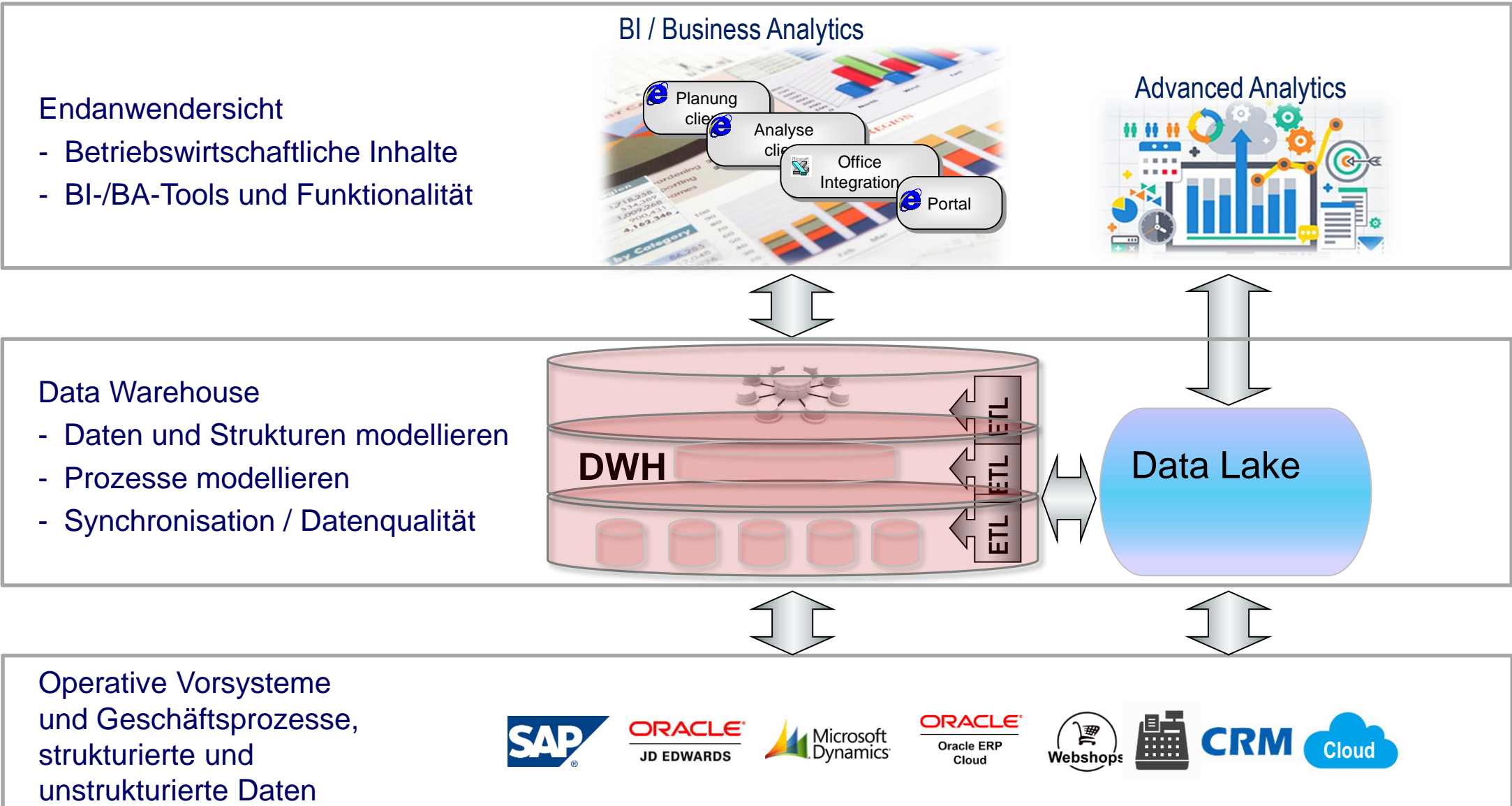
Niederlassung Frankfurt  
Martin-Behaim-Straße 12  
63263 Neu-Isenburg  
Tel: 06102 - 73160  
Fax: 06102 – 731622

Email: [kontakt@datamart.de](mailto:kontakt@datamart.de)  
Web: [www.datamart.de](http://www.datamart.de)



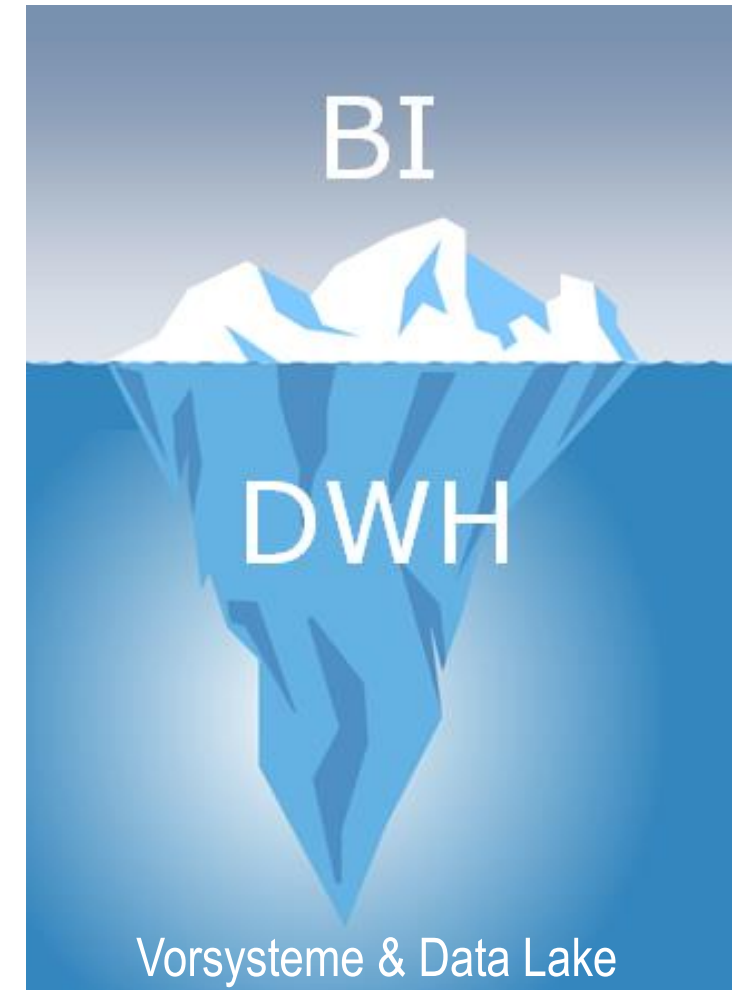
# Aufgabenstellung / Erwartungshaltung „BI“ und Data Warehouse ?

Datenqualitätsmanagement im Data Warehouse



# Aufwandverteilung in einem typischen Data Warehouse- / BI-Projekt

- Auswertung, Berichtswesen, Analyse (BI-/BA-Frontends)  
10-25%
- **Datenbewirtschaftung und Modellierung:**  
Datenmodellierung, Extraktion, Transformation, Laden, Compliance, Verarbeitung,  
**Datenqualitätsmanagement**  
75-90%



- Datenqualität muss messbar sein
- Dokumentierte Datenqualität schafft Sicherheit und Vertrauen in Informationen
- Legale Anforderungen zum Sign-off von Finanz-Informationen, z.B.
  - Sarbanes-Oxley Act (SOX 302 / SOX 404)
  - Basel Committee on Banking Supervision (BCBS 239)



## Wo sind die Datenqualitätsmanagementaufgaben (in Projektierung und Betrieb)

- Technische Prozesse
  - Teil der Datenbewirtschaftungsprozesse
  - Automatisiert, protokolliert, messbar, idealerweise „selbstheilend“
- Fachliche Prozesse
  - DQ-Prüfungen (initial, laufend) nach fachlicher Verantwortlichkeit
  - Spezifikationsqualität incl. DQ-Prüfungen, Testverfahren und Business-Logik
- Organisatorische Prozesse und Verantwortung
  - Verantwortung definieren (fachlich, technisch, Betrieb, BICC ....)
  - Prozesse der Spezifikation / Umsetzung / Testverfahren / laufende Kontrolle
  - Change-Management, Release Management, Dokumentation, DQ-Pflichtkriterien

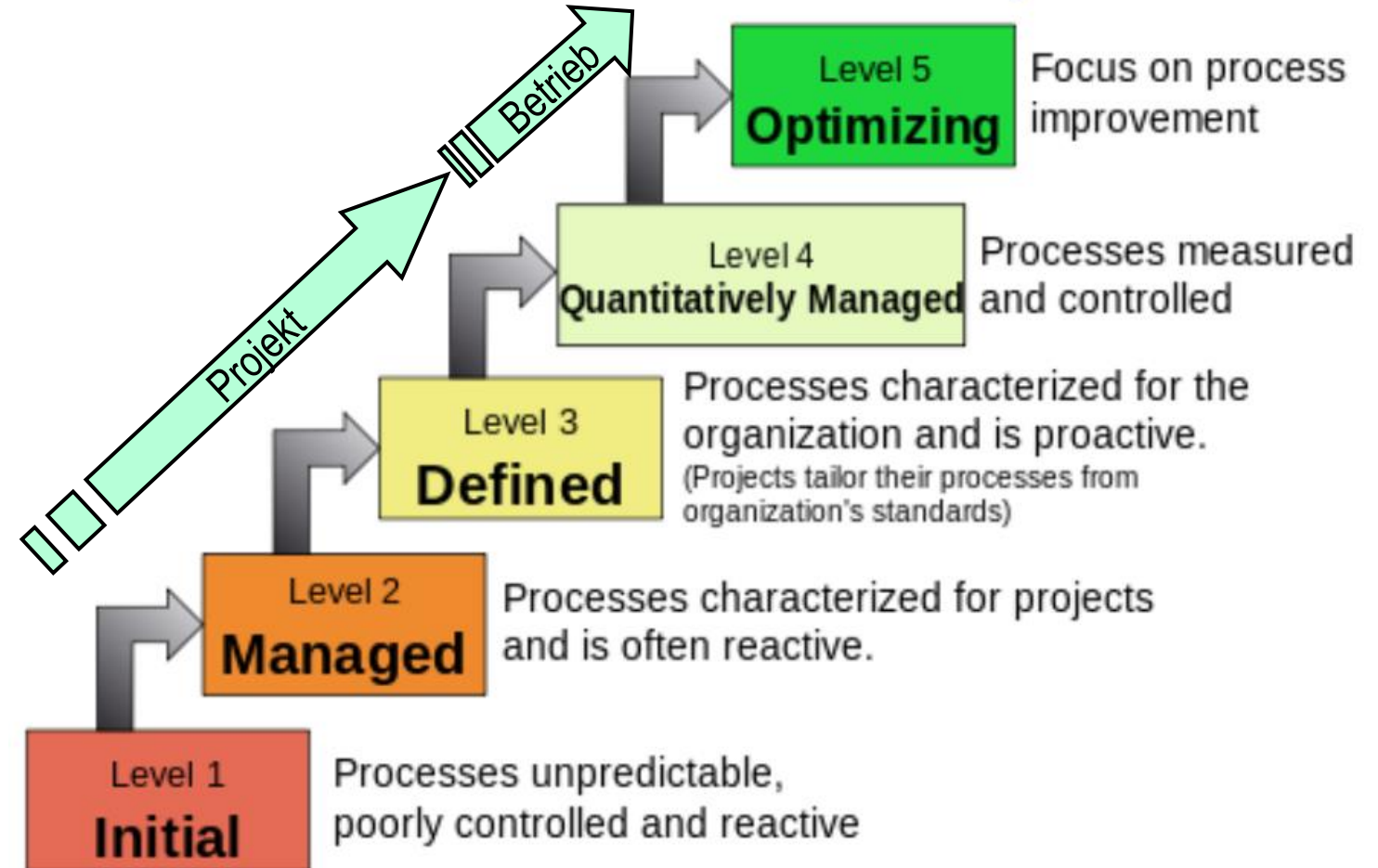


## Reifegrade der QM-Prozesse

CMMI-Modell  
(Capability Maturity Model Integration)

Basis 1979 von Philip B. Crosby :  
Quality Management Maturity Grid (CMM),  
CMMI 1.0: US Ministry of Defense (2000–2002) ,  
CMMI 1.3: SEA 2010 (Software Engineering Institute)

## Characteristics of the Maturity levels





## Was ist jetzt „Qualität“ in diesem Kontext?

- Qualität der Daten und Informationen
  - ⇒ Qualitätsattribute für Daten
- Qualität der Prozesse
  - ⇒ Zuverlässigkeit, Protokollierung, Dokumentation, Audit der Prozesse
  - ⇒ Verfügbarkeit, Wartbarkeit, Nachvollziehbarkeit
- Qualität der Definitionen / Vorgaben
  - ⇒ Spezifikationsqualität
  - ⇒ Definition der DQ-Kriterien
- Qualität der Umsetzung
  - ⇒ Messbarkeit, Kontrolle/Maßnahmen, unterstützende Organisation, Reviews
  - ⇒ Zieldefinitionen, Messung der Zielerreichung, Verantwortung



Als „Qualitätsdimensionen“ für DQ-Kennzahlen werden wir konkreter, um Qualitätsmängel fassbar / analysierbar zu machen und die Ableitung von Maßnahmen zu erleichtern

Qualitätsattribute / -dimensionen	Qualitätsprüfungen und Bildung von Qualitätskennzahlen (Beispiele)
Vollständigkeit der Daten	Prüfung Quellsysteme, Abweichung z. Durchschnitt, Datensätze Quellvergleich ....
Fachliche Korrektheit der Zahlen	Fehleranzahl Validierung, Fehlende Regeln, Fehlende Anreicherungen, .....
Technische Korrektheit der Zahlen	Abweichungen Prüfsummen (pro Ebene), Datentypfehler, Feldfehler, NULL-Prüfungen, ...
Konsistenz	Anzahl Stammdatenfehler, Anzahl Historisierungsfehler, Anzahl Zuordnungsdopplungen, .... (jeweils pro Fact / fachlicher Dimension)
Strukturintegrität	Anzahl Transformationsfehler, Anzahl unlogische Veränderungen, .....
Aktualität	Aktualisierung Quellsystem, Aktualität Ladeprüfung, Umfang Delta-Load

# Datenqualität in Informations-Systemen

## DQ-Matrix (Beispiel)

Letzte Beladung

Woche

Monat

Verlaufsgrafik

DQ-Dimensionen \ DQ-Kennzahlen	Quell-system-fehler	Ladelos fehlend	Ladelos unvollständig	Ladelos Abweichungen $\phi$	Staging-Abweichung	Hist-Abweichung	Analyse-Abweichung	Stamm-daten-fehler	Plausibi-litäts-fehler	Fehler Business-Logik	Fehler Anreicher-ungen	Prüf-summen-fehler	Fehler Cube-Verarbei-tung
<b>Stammdatenverarbeitung</b>													
DQ-Kunde/Debitor	0	0	0	0	0	0	0	5	0	0	0	0	0
DQ-Kunde/CRM	0	0	0	0	0	0	0	12	0	0	0	0	0
DQ-Organisation	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Mandant	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Produkte	0	0	0	0	0	0	0	22	0	0	0	0	0
DQ-Lieferanten/Kreditor	0	0	0	0	0	0	0	0	0	0	32	0	0
DQ-Kostenstellen	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Kostenarten	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Sachkonten	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Verträge	0	0	1	0	0	0	0	3	56	0	41	0	0
<b>DQ-Stamm-Summe</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>42</b>	<b>56</b>	<b>0</b>	<b>73</b>	<b>0</b>	<b>0</b>
<b>Faktenverarbeitung</b>													
DQ-Fakturadaten	0	0	0	0	0	0	0	28	0	22	0	0	0
DQ-CRM-Daten	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-Mailing-Daten	0	0	0	1	0	0	0	0	0	0	0	0	0
DQ-WEB-Tracking	2	2	0	0	0	0	0	0	0	0	0	0	0
DQ-Vertragsdaten	0	0	0	0	0	0	0	46	0	0	0	0	0
DQ-Konditionen	0	0	0	0	0	0	0	12	0	0	0	0	0
DQ-Kostenrechnung	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-FiBu Salden	0	0	0	0	0	0	0	0	0	0	0	0	0
DQ-FiBu Anlagen	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>DQ-Fakten-Summe</b>	<b>2</b>	<b>2</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>86</b>	<b>0</b>	<b>22</b>	<b>0</b>	<b>0</b>	<b>0</b>

# Datenqualität in Informations-Systemen

## ETL-DQ-LOG (Beispiel)

Protokollierung der Prozesse und Prüfungen auf der Ebene der technischen Bewirtschaftungsprozesse (hier per ODI auf ORACLE), angereichert um die Ergebnisse jeder Qualitätsprüfung in den Einzelschritten:

Auswahl Betrachtungstag

Auswahl Ebene

Auswahl Job

nur Fehlerbericht

ETL-Prozess	Datum	Server	JobID	Start	Ende	Laufzeit	Lade-Modus	Event	result	#records source	#records target	Target / message	System-fehler	System-warnings	Stamm-daten-fehler	Plausibi-litäts-fehler	Fehler Business-Logik	Fehler Anreicher-ungen
HIS_ANA_TA_AGENTEN	2017-04-25	DWDBPROD	1,87E+08	07:28:13.000	07:38:39.000	10:26	DELTA	End	success	657388	657388	'ANA_TA_AGENTEN' (21250)' schrieb 657388 Zeilen.	0	0	0			
SRC_STG_TA_TRANSAKTIONSCODE_MASTER	2017-04-25	DWDBPROD	1,87E+08	06:14:31.000	06:14:42.000	00:11	FULL	End	success	8	8	'STG_TA_TRANSAKTIONSCODE_MASTER' (17442)' schrieb 8 Zeilen.	0	0	0			
SRC_STG_TA_TRANSAKTIONSCODE	2017-04-25	DWDBPROD	1,87E+08	06:13:07.000	06:14:42.000	01:35	DELTA	End	success	16973	16973	'STG_TA_TRANSAKTIONSCODE' (17433)' schrieb 16973 Zeilen.	0	0	0			
SRC_STG_TA_STATISTIK_PILOTEN	2017-04-25	DWDBPROD	1,87E+08	06:14:31.000	06:14:37.000	00:06	DELTA	End	success	4515	4515	'STG_TA_STATISTIK_PILOTEN' (17433)' schrieb 4515 Zeilen.	0	0	3			
SRC_STG_TA_PILOTEN_SPLIT_PILOTEN	2017-04-25	DWDBPROD	1,87E+08	06:13:47.000	06:14:31.000	00:44	FULL	End	success	47444	47444	'STG_TA_PILOTEN_SPLIT_PILOTEN' (17580)' schrieb 47444 Zeilen.	0	0	0			
SRC_STG_TA_PILOT_MASTER	2017-04-25	DWDBPROD	1,87E+08	06:13:44.000	06:13:46.000	00:02	FULL	End	success	28	28	'STG_TA_PILOTEN_MASTER' (17276)' schrieb 28 Zeilen.	0	0	0			
SRC_STG_TA_PILOTEN	2017-04-25	DWDBPROD	1,87E+08	06:02:32.000	06:13:36.000	11:02	DELTA	End	success	1081111	1081111	'STG_TA_PILOTEN' (17079)' schrieb 1081111 Zeilen.	0	0	0			
SRC_STG_TA_BEARBEITUNGSGRUPPE_PRO_PILO	2017-04-25	DWDBPROD	1,87E+08	06:00:30.000	06:02:27.000	01:57	DELTA	End	success	86320	86320	'STG_TA_BEARBEITUNGSGRUPPE_PRO_PILOT' (17961)' schrieb 86320 Zeilen.	0	0	0			
SRC_STG_TA_BEARBEITUNGSGRUPPE_MASTER	2017-04-25	DWDBPROD	1,87E+08	05:59:01.000	05:59:45.000	00:44	DELTA	End	success	7	7	'STG_TA_BEARBEITUNGSGRUPPE_MASTER' (18084)' schrieb 7 Zeilen.	0	0	0			
SRC_STG_TA_BEARBEITUNGSGRUPPE	2017-04-25	DWDBPROD	1,87E+08	05:31:45.000	05:57:48.000	26:03	DELTA	End	success	1323673	1323673	'STG_TA_BEARBEITUNGSGRUPPE' (17654)' schrieb 1323673 Zeilen.	0	0	0			
SRC_STG_TA_AGENTENPILOT	2017-04-25	DWDBPROD	1,87E+08	05:10:08.000	05:10:33.000	00:25	FULL	End	success	10	10	'STG_TA_AGENTENPILOT' (16884)' schrieb 10 Zeilen.	0	0	0			
SRC_STG_TA_AGENTEN_MASTER	2017-04-25	DWDBPROD	1,87E+08	05:10:01.000	05:10:06.000	00:05	FULL	End	success	18	18	'STG_TA_AGENTEN_MASTER' (16582)' schrieb 18 Zeilen.	0	0	0			
SRC_STG_TA_AGENTEN	2017-04-25	DWDBPROD	1,87E+08	05:09:55.000	05:10:06.000	00:07	FULL	End	success	10	10	'STG_TA_AGENTEN' (17241)' schrieb 10 Zeilen.	0	0	0			
ANA_ANZAHL_RECHNUNGEN	2017-04-25	DWDBPROD	1,87E+08	04:49:45.000	04:52:11.000	02:26	DELTA	End	success	226312	226312	'OLE DB-Ziel' (8764)' schrieb 226312 Zeilen.	0	0	0			
ANA_ANZAHL_RECHNUNGEN	2017-04-25	DWDBPROD	1,87E+08	04:49:45.000	04:52:11.000	02:26	DELTA	End	success	226312	226312	'Flatfileziel ExportRechnung' (8642)' schrieb 226312 Zeilen.	0	0	0			
HIS_ANA_LM_REICHWEITEN_TAEGLICH	2017-04-25	DWDBPROD	1,87E+08	03:06:39.000	03:09:12.000	02:33	DELTA	End	success	216794	216794	'ANA_KD_REICHWEITEN_ALLE_WOCHEN1' (770)' schrieb 216794 Zeilen.	0	0	0	0	0	0

Für eine zuverlässige Datenbewirtschaftung UND das zugehörige Datenqualitätsmanagement sind sowohl ein Datenmodell als auch ein Prozessmodell zu definieren

- **Prozessmodell:** Definierte Prozesse in standardisierten Prozessebenen (ETL)
- **Datenmodell:** Definierte Fehlertabellen und Datenqualitätstabellen, die ebenso wie die Finanz- oder Controlling-Daten für die Verantwortlichen aufzubereiten sind!

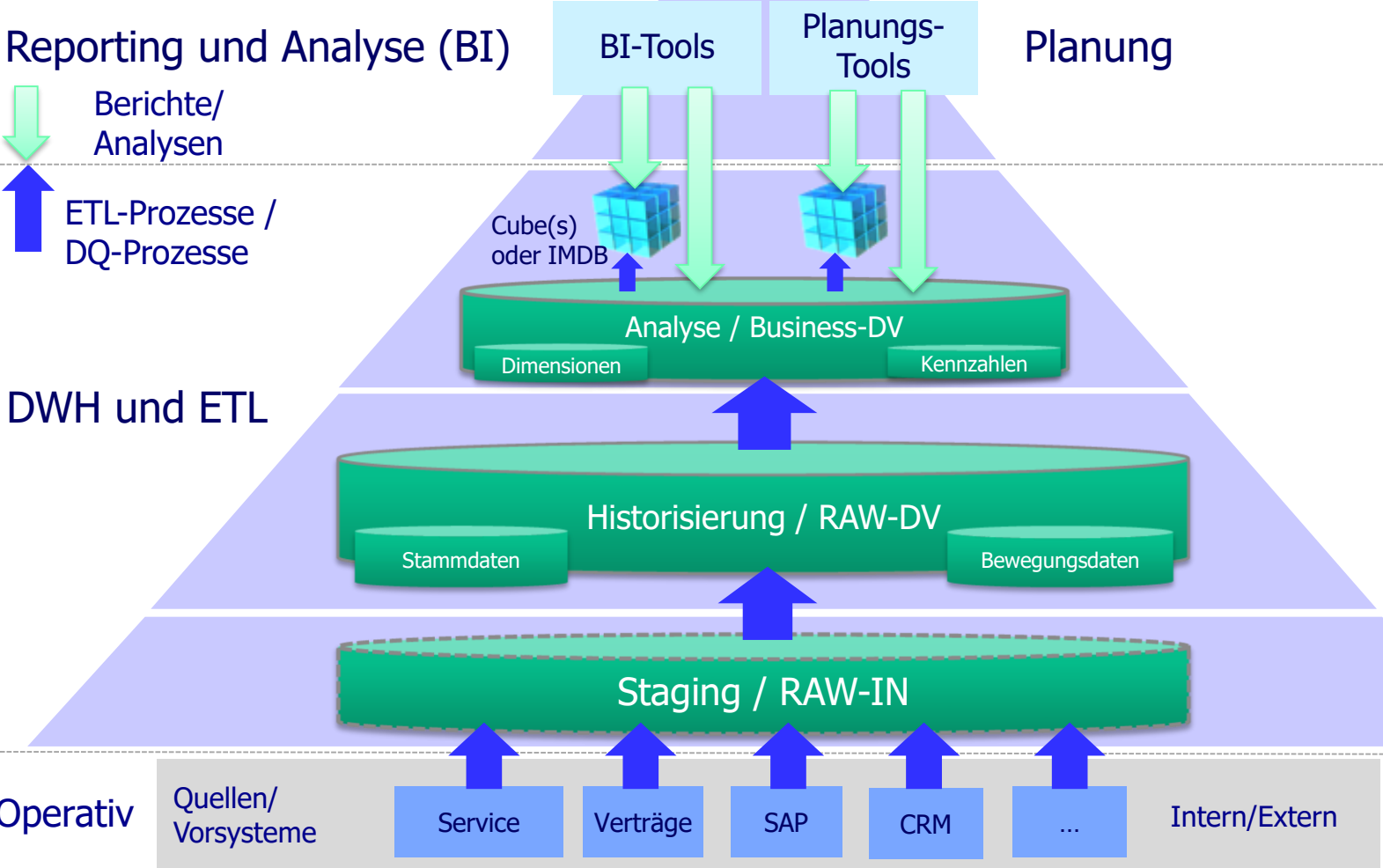
Die Monitoring der Datenqualität hat ebenfalls zwei Perspektiven:

- **Prozessqualität:** wie sind die Prozesse gelaufen
- **Datenqualität:** Stimmen die Daten inhaltlich

# Beispiel Architekturmodell und Prozessmodell

Datenqualitätsmanagement im Data Warehouse

## Data Warehouse und BI - Aufbau



## DQ-Modell

## DQ-Prozesse

- BI-Schicht**
- DQ-Kennzahlen
  - Visualisierung in den Daten
  - Qualitätscockpit

- Informieren/Messen**
- Prüfberichte
  - Exceptions
  - Abgleich
  - Statistik

- Analyseschicht**
- DQ-Kennzahlen
  - Stammdatenfehler
  - Ersetzungen
  - Zuordnungen
  - Prüfschritte Validierung

- Semantische Prüfung**
- Integrität
  - Stammdaten
  - Business-rules
  - Validität

- Historisierung**
- Fehlerdaten
  - Prüfsummen
  - Protokollierung
  - Ladeloskontrolle

- Ladung/Historisierung**
- Vollständigkeit
  - Konsistenz
  - Feldprüfungen
  - Datentypen

- Staging**
- Fehlertabellen
  - Warnings
  - Rejections

- Schnittstellen**
- Ladelosprüfung
  - Vollständigkeit
  - Aktualität
  - Technische Fehler

## Vollständigkeit

- Prüfung der Ladelose pro Quelle im ETL-Prozess (date-range)
- Prüfung der Systemfehler Connect-to-source (System nicht abfragbar / vorhanden)
- Abweichungen zu Messgrößen (threshold) oder Abweichungen zum Durchschnitt
- Zählungen (z.B. Anzahl Filialmeldungen, Anzahl Datensätze pro Quelle, Anzahl Konditionen..)
- Vergleich (pro Ebene): Anzahl Datensätze Source => Target

## Fachliche Korrektheit der Daten

- Prüfung in Validierungsschritten mit Protokollierung und Fehlertabelle (zusätzlich, appended)
- Prüfung aller Anreicherungen (z.B. Konditionen, Bezugspreise, Bewertungen...)
- Protokollierung der Business-Rules (welche Fälle trafen zu / nicht zu)
- Plausibilisierungen (z.B. versus Durchschnitt, versus Plan etc., Abweichung Mittelwerte, Range von Datenwerten)



## Technische Korrektheit der Daten

- Prüfsummenvergleiche pro Ebene/Tabelle
- Datentypfehler pro Ebene/Tabelle
- Feldfehler (z.B. truncation) und NULL-Fehler
- Dubletten-Prüfungen
- Prüfung gegen externe Quellen (z.B. Adressprüfungen)
- Parallelprüfung („Audit-Load“) über alle Ebenen auf separatem Informationsfluss  
=> Prüfung Quelldaten gegen Ziel (Analyseschicht)

## Konsistenz der Daten und Strukturintegrität der Daten

- Vollständigkeit von Stammdatenhistorisierungen (SCD2)
- Konsistenz von hierarchischen Zuordnungen und Attributen der Stammdaten
- Prüfung zu führenden Stammdatensystemen
- Prüfung der Zuordnung zu Analysestrukturen (Clustering, ABC, Segmentierungen ...)
- Prüfung der Transformationen und Synchronisationen (no-match cases)
- Prüfung der Stammdatenanreicherung aller Bewegungsdaten, orphan-value-replacement
- Prüfung der logischen Schlüsseldefinitionen (Dublekken)

# Datenqualität im Einzelnen (ETL-Prozesse)

## Aktualität der Daten

- Aufgabe des technischen Monitoring
- Visualisierung der Aktualität bis zum Fachanwender
- Prüfung der Aktualität im ETL-Prozess (siehe Vollständigkeit)

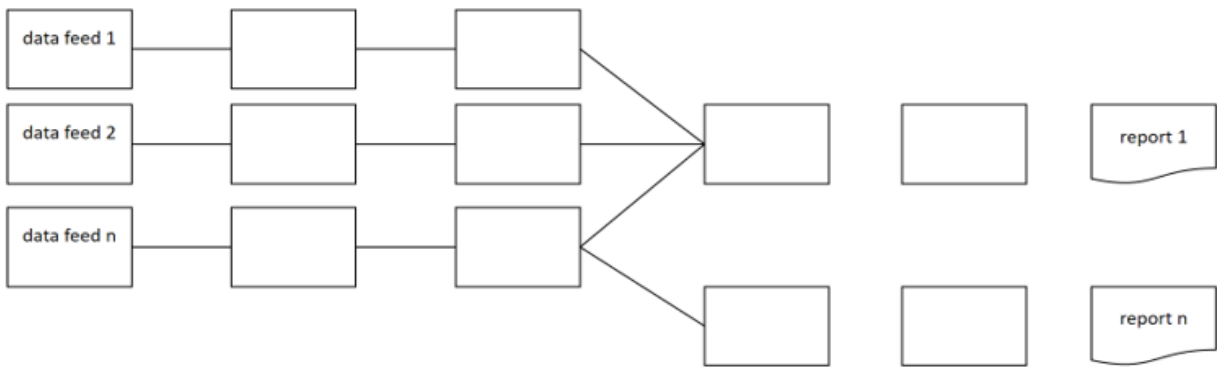
## Prozessqualität

- Automation
- Monitoring und Alerting
- Protokollierung
- Nachvollziehbarkeit (Daten: data lineage, Prozesse: Protokollierung in den Daten)
- Fehlertoleranz / Abbruchsicherheit
- „Selbstheilung“ von Stammdatenfehlern und Anreicherungsfehlern
- .....

# Data Quality Control Framework

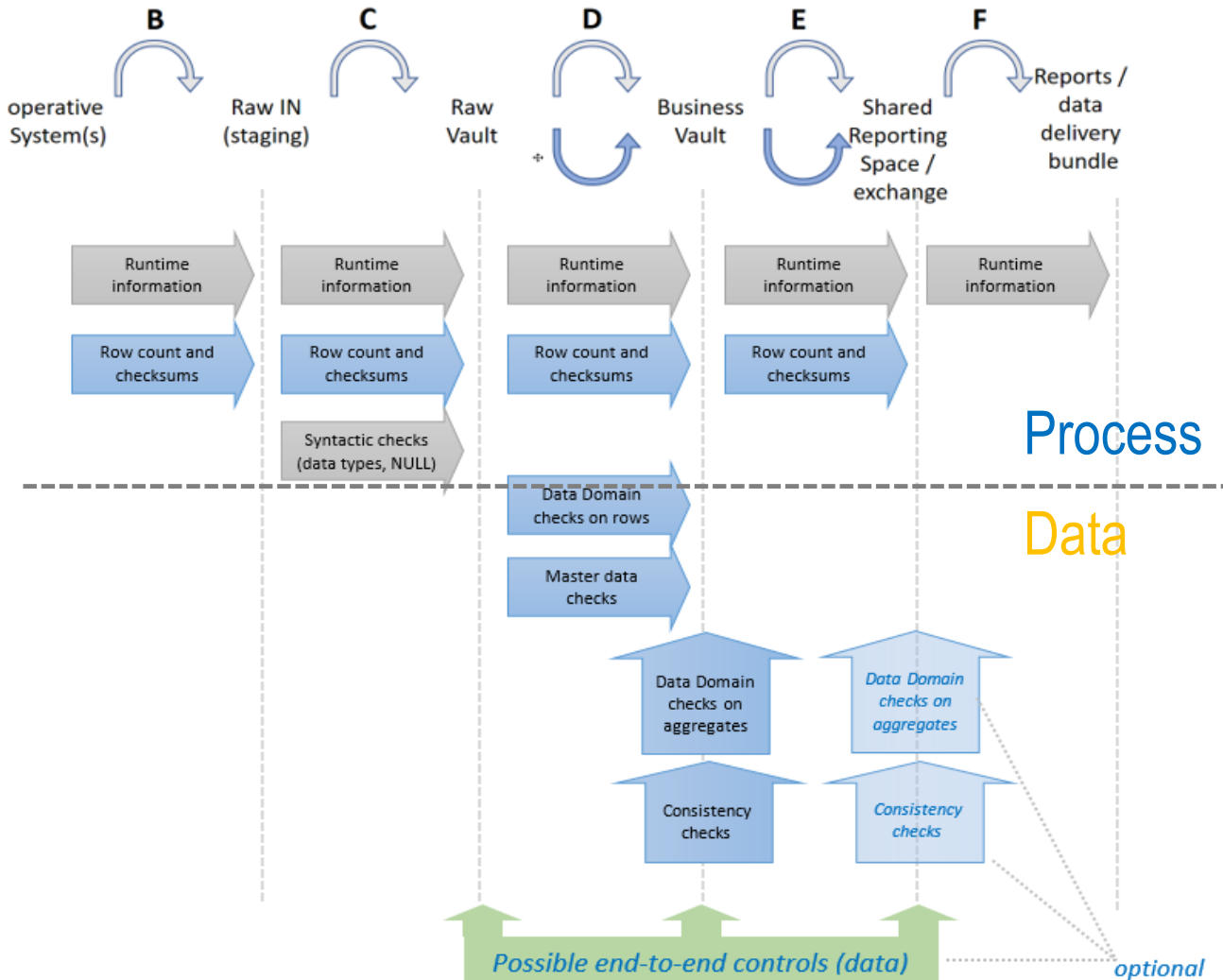
## Dimensions of controls:

<b>Daten</b>	Datenwertprüfung	Datenwert liegt innerhalb eines definierten Wertebereiches
	Konsistenz	Daten sind konsistent mit anderen Daten
	Aktualität	Die aktuellsten Daten sind verarbeitet
	Eindeutigkeit	Eineindeutige Relation von einem Datensatz im Reporting zur Quelle
<b>Prozess</b>	Trend/Entwicklung	Messung aller Prozesse u. Vergleich aktuell vs. gleitender Ø (Volatilität der Werte)
	Integrität	Prozess hat den Inhalt von der Quelle bis zum Ziel unverfälscht transportiert
	Zeiteinhaltung	Prozess ist innerhalb der gesetzten Zeit beendet, Reporting kann beginnen



## Prozessbezogene und Datenbezogene Kontrollen

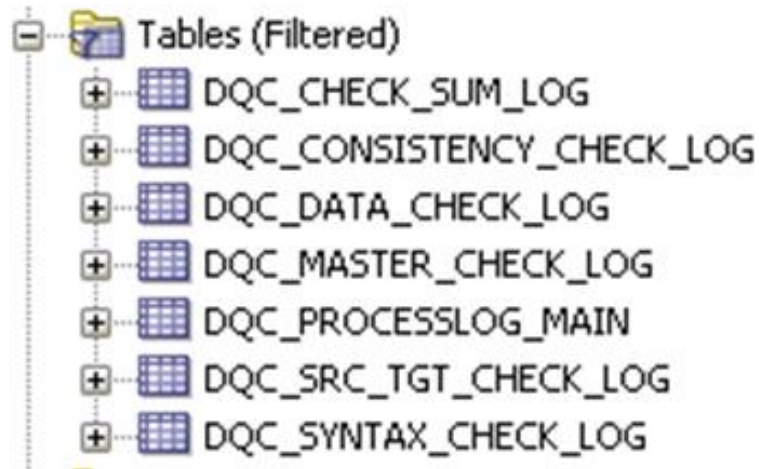
- Technische Prüfungen vorwiegend bis Raw Data Vault
- End-to-end Kontrollen und semantische Prüfungen im Business Data Vault



# Data Quality Control Framework

## Vereinfachtes DQCF Datenmodell

⇒ Ziel ist es, alle DQ-Prozesse in ein DQ-Modell zu protokollieren



DQC_SRC_TGT_CHECK_LOG
PROC_ID
LOG_DATETIME
SRC_OBJ
TGT_OBJ
SRC_ROW_CT
TGT_ROW_CT
MUST_MATCH
TGT_UPDATE_CT
TGT_INSERT_CT
MIN_BUSINESS_DATE
MAX_BUSINESS_DATE
MESSAGE

DQC_MASTER_CHECK_LOG
PROC_ID
LOG_DATETIME
FIELDNAME
REFERENCE
NO_MATCH_CT
MESSAGE

DQC_CONSISTENCY_CHECK_LOG
PROC_ID
LOG_DATETIME
RULE_NAME
RULE_DESCRIPTION
ERROR_CT
MESSAGE

DQC_PROCESSLOG_MAIN
PROC_ID
PROC_NAME
LOG_DATETIME
MAIN_SOURCE_NAME
EXPECTED_ROWS
ROWTOTAL_SRC
START_DATETIME
END_DATETIME
LOAD_TYPE
DOMAIN
STATUS
RESULT
MESSAGE

DQC_SYNTAX_CHECK_LOG
PROC_ID
LOG_DATETIME
FIELDNAME
CHECK_TYPE
ERROR_CT
MESSAGE

DQC_CHECKSUM_LOG
PROC_ID
LOG_DATETIME
NAME_SRC
VALUE_SRC
NAME_TGT
VALUE_TGT
MESSAGE

DQC_DATA_CHECK_LOG
PROC_ID
LOG_DATETIME
MEASURE_NAME
RULE_DESCRIPTION
ERROR_CT
MESSAGE

- Oracle 12c (12.1)
- Oracle Business Intelligence Enterprise Edition 12c (12.2)
- Data Vault 2.0  
(Schichten: Sources, Raw-In, Raw Data Vault, Business Data Vault, Information Mart)
  - Herausforderungen:
    - große Anzahl Tabellen von RDV nach BDV
    - nicht persistentes BDV



# Oracle BIEE Implementation Highlights

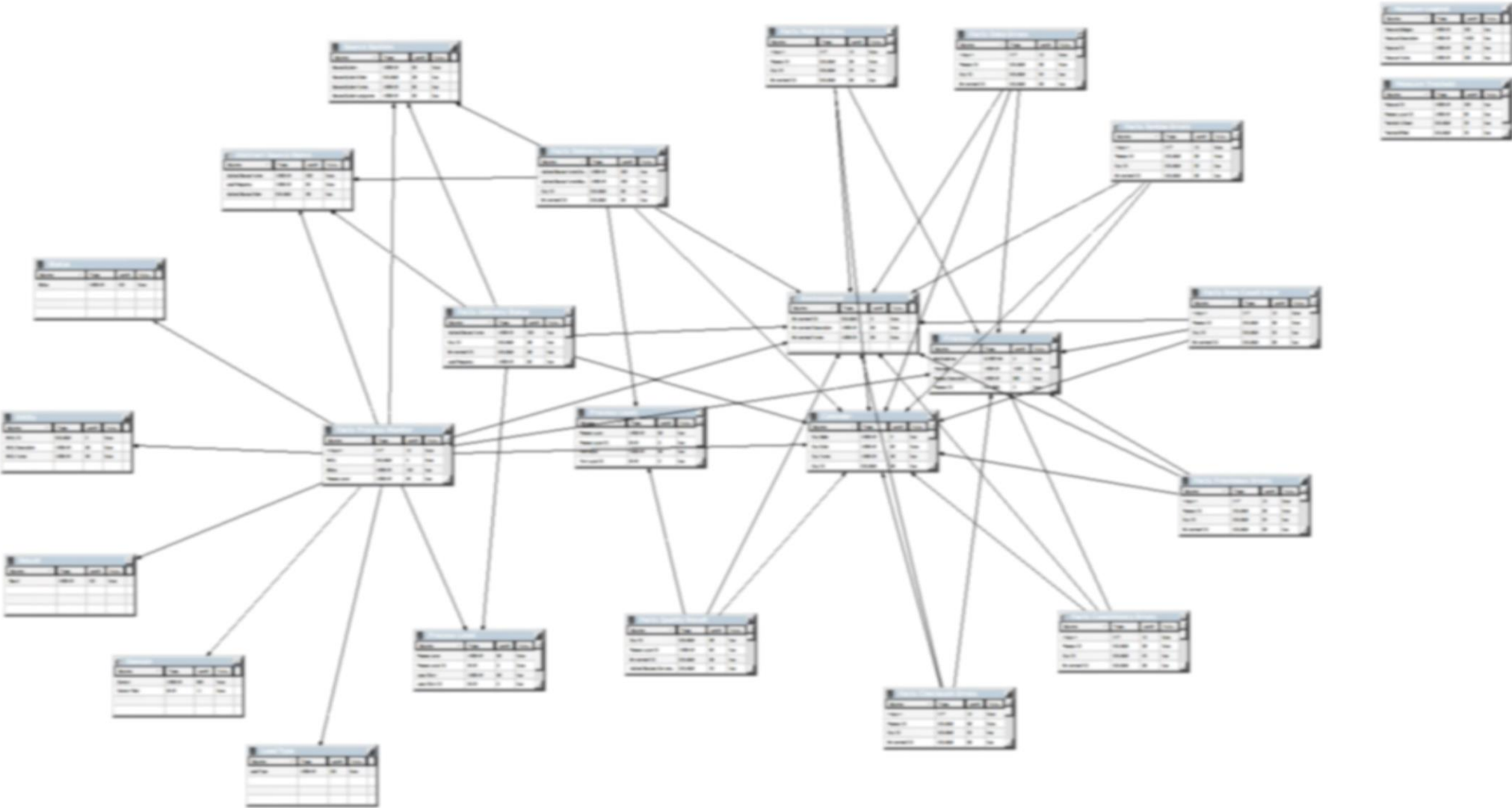
DOAG ORACLE®



- Repository Highlights:
  - Dynamic row-wise Initialization Variables used for Measure Tooltips
  - Writeback for Tooltip Maintenance
  - Dynamic Connection Pool (Different Databases: DEV,UAT,PRD)
- Visualization Highlights:
  - Master-Detail Reports
  - Conditional Drill-Downs
  - Report Pop-ups
  - Conditional Formatting based on complex calculations

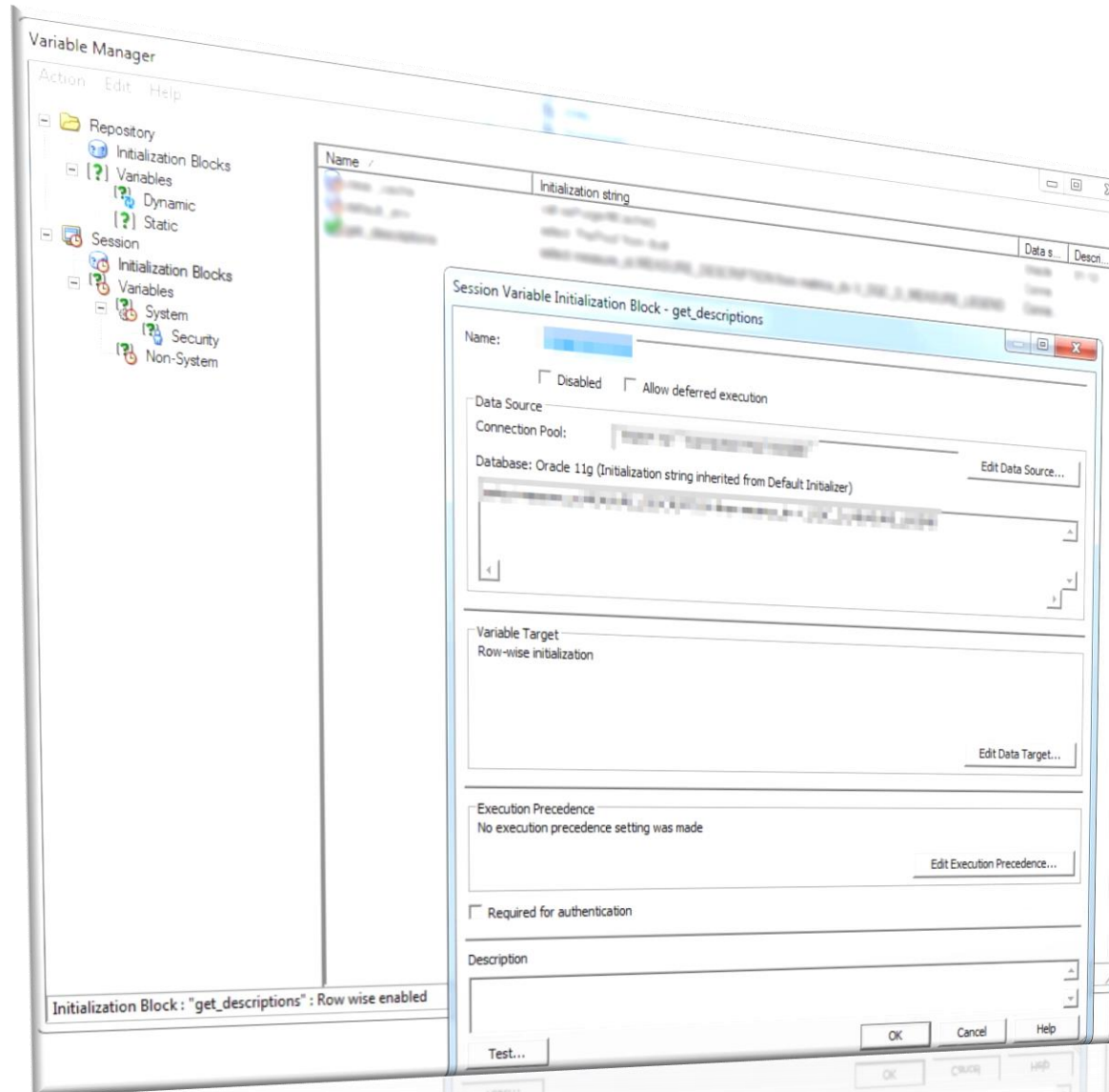


# OBIEE Repository – Business Model



Datenqualitätsmanagement im Data Warehouse

# OBIEE Repository – Dynamic Variables



- Abfrage für dynamische Tooltips
- Row-wise initialization
- HTML Code in Analysis Spalten Überschriften:

```
<b title="{biServer.variables['NQ_SESSION.Variable1']}">Measure 1<b>
```



# OBIEE Repository – Dynamic Connection Pool

Datenqualitätsmanagement im Data Warehouse

- Prompt mit DSN Selektion
- Repository Session Variable
- Dynamische Connection Pool Selektion: VALUEOF(NQ\_SESSION.DSN)

**Set DSN Value**

(DESCRIPTION=(ADDRESS\_LIST=(ADDRESS=...))  
 (DESCRIPTION=(ADDRESS\_LIST=(ADDRESS=...))

---

**Edit Prompt: Set DSN Value**

Prompt For Column:  fx

Label:

Custom Label

Description:

Operator:

User Input:

---

**Options**

Radio Buttons Values:  ++

(DESCRIPTION=(ADDRESS\_LIST=(ADDRESS=...))  
(DESCRIPTION=(ADDRESS\_LIST=(ADDRESS=...))

Require user input  
 Limit values by:

Auto Fill:  None  First Available Value

Default selection:

Radio Buttons Layout:  Vertical  Horizontal

Set a variable:

**Sessionvariable – DSN**

Name:

Festlegen des Wertes durch jeden Benutzer zulassen  
 Sicherheitssensitiv

Initialisierungsblock:

Standardinitialisierungsblock:

Beschreibung:

**Verbindungspool – Sample Relational Connection**

Name:  Berechtigungen...

Aufrufsschnittstelle:

Maximale Anzahl:

Vollständig angegebene Tabellenamen erforderlich

Datenquellenname:  ←

Gemeinsame Anmeldung

Benutzername:  Kennwort:

Verbindungspooling aktivieren

Timeout:  Endlos

Multi-Threaded-Verbindungen verwenden

Unterstützte Parameter

Isolationsebene:

Beschreibung:

# OBIEE Visualization – Master-Detail Reports

The screenshot displays a Master-Detail report in OBIEE. On the left, a table lists data for various regions, with one row highlighted in blue. On the right, a line chart shows 'Sales' (orange line) and 'Profit' (blue line) over time. The chart is filtered to show data only for the selected region in the table, demonstrating context-filtering.

- Drill-Down Bericht
- Context-Filtering auf Grafik
- Master-Detail Bericht



# OBIEE Visualization – bedingter Drill-Down

Datenqualitätsmanagement im Data Warehouse

Item ID	Item Name	Item Desc	Item Type	Item Status	Item Count	Item Error Rate	Item Error Rate	Item Error Rate	Item Error Rate	Item Error Rate	Item Error Rate
103					103	4.85%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	100.00%	0.00%				
1					1	100.00%	0.00%				
1					1	100.00%	0.00%				
1					1	100.00%	0.00%				
1					1	100.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				
1					1	0.00%	0.00%				

- Drill-Down nur bei Fehlern
- Drill-Down nur wenn bestimmte Ebene



# OBIEE Visualization – Kontext Mini-Report

The screenshot displays a data table in an OBIEE interface. The table has several columns, and one row is highlighted in red. A pop-up mini-report is visible over the table, showing a detailed view of the selected row's data. The table data includes a total count of 103 and several percentage values, with some cells highlighted in red or green.

Count	Percentage	Percentage	Percentage
103	4.85%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	0.00%	0.00%	
1	100.00%	0.00%	
1	100.00%	0.00%	
1	100.00%	0.00%	

- Kontext-Drill-Down
- Mini-Report als Pop-Up im Daten-Kontext

# OBIEE Dashboards – Beispiel DQ-Matrix

Data loaded on 28-02-2018 (FRI)  
@ PRD

The high Level DQ-Matrix provides an aggregated management view of the DQ data by Data Domains (Types) across Dimensions of Control: grouped by PQ Process Quality (Trend/Timeliness and Integrity) and DQ Data Quality (Freshness, Uniqueness, Consistency, Domain of Value)

Day Date: 28-02-2018 (WED)

Domain: --Select Value--

Entity: SALES-COMMERCIAL

Target Load Time: 10:00:00 AM

Environment: PRD

## DQ-Matrix

## DQ-Matrix-Rates

Day Date	Domain	PQ Trend/Timeliness	PQ Integrity	DQ Freshness	DQ Uniqueness	DQ Consistency	DQ Domain of Value
28-02-2018 (WED)	Domain Total	Warning	Success	Warning	Warning	Success	Success
	ASSET	Success	Success	Success	Success	Success	Success
	BUSINESS PARTNER	Success	Success	Success	Success	Success	Success
	BUSSINESS PARTNER						
	FACILITY	Success	Success	Success	Success	Success	Success
	GENERAL LEDGER	Success	Success	Success	Success	Success	Success
	TRANSACTION	Success	Success	Success	Success	Success	Success
	UNDEFINED	Warning	Success	Warning	Warning	Success	Success
	domain_A						
	domain_B	Warning	Success	Success	Success	Success	Success
domain_D							

Day Date	Domain	Row Error Rate %	Freshness Error Rate %	Syntax Error Rate %	Match Error Rate %	Consistency Rate %	Data Error Rate %	Sum Error Rate %
28-02-2018 (WED)	Domain Total	0.28%	0.00%					
	ASSET	0.00%	0.00%					
	BUSINESS PARTNER	0.00%	0.00%					
	BUSSINESS PARTNER							
	FACILITY	0.00%	0.00%					
	GENERAL LEDGER	0.00%	0.00%					
	TRANSACTION	0.00%	0.00%					
	UNDEFINED	0.30%	0.00%					
	domain_A							
	domain_B	0.00%	0.00%					
domain_D								



# OBIEE Dashboards – Beispiel Duration

Data loaded on 03-03-2018 (FRI)  
@ PRD

Dashboard shows load duration by load process along BI Platform architecture.  
Significant slow-down of performance helps to identify issues.  
Dashboard information is base for IT Operations Team to:  
o find solution for slow-down,  
o share issue with stakeholders.

← Back to Overview

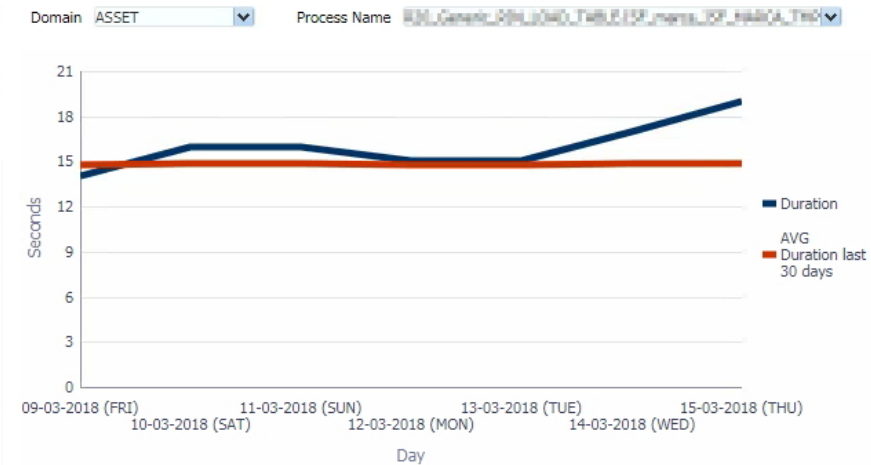
Day Between 02/28/2018 - 02/28/2018  
 Domain --Select Value--  
 Process Level File-loading to view  
 Process Name --Select Value--  
 Entity --Select Value--  
 Environment PRD

## Duration Result Table

(Expand to Process Detail or drill to get detail information report)

Day Date	Process Level	Process	Main Source Name	Domain	Number of Executions	Duration	AVG Duration	AVG Duration last 30 days
28-02-2018 (WED)	File-loading to view	Process Total		ASSET	1.00	12	12.0	15
				BUSINESS PARTNER	4.00	69	17.3	79
				FACILITY	3.00	53	17.7	73
				GENERAL LEDGER	5.00	72	14.4	83
				TRANSACTION	4.00	175	43.8	186
				UNDEFINED	106.00	38,432	362.6	52,989
				ASSET	1.00	43	43.0	30
				UNDEFINED	1.00	30	30.0	39
				UNDEFINED	1.00	2,116	2,116.0	2,101
				UNDEFINED	1.00	303	303.0	219
				UNDEFINED	1.00	50	50.0	42
				UNDEFINED	1.00	319	319.0	92
				UNDEFINED	1.00	584	584.0	445
				UNDEFINED	1.00	29	29.0	38
				UNDEFINED	1.00	32	32.0	44
				UNDEFINED	1.00	44	44.0	54
				UNDEFINED	1.00	768	768.0	770
				UNDEFINED	1.00	384	384.0	233
				UNDEFINED	1.00	314	314.0	354
				UNDEFINED	1.00	2,640	2,640.0	4,404
				UNDEFINED	1.00	2,028	2,028.0	3,131
				UNDEFINED	1.00	536	536.0	784
				UNDEFINED	1.00	4,147	4,147.0	6,738
				UNDEFINED	1.00	5,078	5,078.0	8,111
				UNDEFINED	1.00	11	11.0	47
				UNDEFINED	1.00	22	22.0	22
				UNDEFINED	1.00	21	21.0	23

## Duration Result Graph



# OBIEE Dashboards – Beispiel Process Overview

Data loaded on 28-02-2018 (FRI)  
@ PRD

Dashboard shows status on expected Delivery Sources, as degree of fulfillment for each Process Layer. It only measures the **completely successful results**. We look at the Source in its entirety as a process. This means, even if a part was OK, we rate the Source as not loaded. Success rate in this case would be 0%.

← Back to Overview



Day Date 28-02-2018 (WED) ▼

Apply Reset ▼

Environment PRD ▼

Apply Reset ▼

## Process Overview

Day Date	Main Layer	Process Layer	ASAP (ASAP-Germany)	IMP	SAP	ASHT (ASHT)	SPC (SAP/PEL)	IN (ASHT)	SAP (ASHT)	ASHT (ASHT)
			Process Status %	Process Status %	Process Status %	Process Status %	Process Status %	Process Status %	Process Status %	Process Status %
28-02-2018 (WED)	Data Transfer / Sources	File Delivery	97% ▼	100%	100%	100%	100%			
		File-loading to 	97% ▼	100%	100%	100%	100%			
	Data Preparation / Processing	ASAP to RDV	100%		100%	88%	100%			
		RDV to BDV								
	Provision of Information	BDV to Information Mart								
	Information Delivery	Information Mart to 								
	other	other								

Conditional Formatting  
based on complex calculations  
with different rules for each layer

# OBIEE Dashboards – Beispiel Delivery Status

Data loaded on 28-02-2018 (FRI)  
@ PRD

Dashboard shows Status on expected Delivery Sources.  
It contrasts the number of EXPECTED runs and the number of effectively performed runs (DELIVERED independent of the Result), and thereof the successful runs (PROCESSED)

← Back to Overview

Day Date: 28-02-2018 (WED) ▾

Apply Reset ▾

Environment: PRD ▾

Apply Reset ▾

Source System: (All Column Values) ▾

Apply Reset ▾

## Delivery Status

(filtered on Process Level: File-loading to martain)

Day Date	Abstract Source Name	Load Frequency	BASE (WOLFGANG)			BAP			BAP			BENT (WALBERT)			BPCO (Lorenz/WILHELM)			BR (WOLFGANG)			BAPS (WOLFGANG)			LWENTFELD (WOLFGANG)		
			Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed	Expected	Delivered	Processed
28-02-2018 (WED)	WICHA_DOSBERG	WorkingDays	1	1	1																					
	[A]BAC	WorkingDays	1	1	1																					
	DAUFHA	WorkingDays	1	1	1																					
	DEUCON	WorkingDays	1	1	1																					
	[D]GCB	WorkingDays	1	1	1																					
	DCUPOT	WorkingDays	1	1	1																					
	DCLIDE	WorkingDays	1	1	1																					
	[D]LPR	WorkingDays	1	1	1																					
	DODABL	WorkingDays	1	1	1																					
	DODCUB	WorkingDays	1	1	1																					
	[D]ODU	WorkingDays	1	1	1																					
	DODORO	WorkingDays	1	1	1																					
	DODODH	WorkingDays	1	1	1																					
	[D]ODPL	WorkingDays	1	1	1																					
	DODCFE	WorkingDays	1	1	1																					
	DODDMA	WorkingDays	1	1	1																					
	[D]ODPE	WorkingDays	1	1	1																					
	DODHEZ	WorkingDays	1	1	1																					
	DODHEP	WorkingDays	1	1	1																					
	[D]ODHP	WorkingDays	1	1	1																					
	DODDPE	WorkingDays	1	1	1																					
	DODHAT	WorkingDays	1	1	1																					
	[D]ODCT	WorkingDays	1	1	1																					
	DODPME	WorkingDays	1	1	1																					
	DODPME	WorkingDays	1	1	1																					
	[D]ODRH	WorkingDays	1	1	1																					
	DODDCO	WorkingDays	1	1	1																					
	DEMOE	WorkingDays	1	1	1																					
	[D]ODCL	WorkingDays	1	1	1																					
	DREDOC	WorkingDays	1	1	1																					
	DVBIHO	WorkingDays	1	1	1																					
	[D]VBIQE	WorkingDays	1	1	1																					
	DVBIOR	WorkingDays	1	1	1																					

## Rolle der Datenqualität in der Entscheidungsunterstützung

- Datenqualität ist das Fundament der Systeme zur Entscheidungsunterstützung
- Payback von Investition in Datenqualität wird durch Wartbarkeit, niedrigere Betriebskosten und sichere Entscheidungsgrundlagen gebildet
- Qualität in der Entscheidungsunterstützung  
= gesicherte Datenqualität der Geschäftsdaten  
+ intelligente Strukturierung / Anreicherung der Daten  
+ Qualität der Datenpräsentation und -analyse
- Datenqualität ist unsichtbar – wenn sie vorhanden ist ...





# Vielen Dank