

# Informations-Dimension der DG

---

*Data Governance*

*Thilo Riegel, München*

## Zusammenfassung

Data-Governance-Serie, Teil 2: Im Teil 1 haben wir die klassischen Probleme der Data Governance beschrieben. In den folgenden Teilen beschreiben wir konzeptionell einen effektiven und Praxiserprobten Lösungsansatz. Entgegen häufigen Vermutungen basiert dieser nicht auf Tools bekannter Hersteller, die i.W. die Buchhaltung der technischen Felder der Überleitungen zwischen ihnen vereinfachen, sondern er basiert auf Aggregation und fachlicher Abstraktion.

In diesem Artikel gehen wir auf die sog. Informations-Dimension der Data Governance ein.

## Begriffsklärung

Im BI-Umfeld haben wir es nicht „einfach nur mit Daten“ zu tun, die wir in einem DWH aus verschiedenen Quellen zusammentragen. Es geht vielmehr um die Steuerung eines Unternehmens auf Basis von *Informationen*, die aus diesen Daten gewonnen werden.

Diese Informationen werden auf fachlicher und teilw. auch auf Management-Ebene klar definiert und zu einem Report kompiliert. Damit sie entsprechend kompiliert werden können, müssen sie wiederum nach bestimmten Dimensionen aufgefächert vorliegen. Die klassischen BI-Informations-Dimensionen sind hierbei:

## Direkt auswertbare Daten

Das sind die Daten für den Kunden, also letztlich das Management:

- Kern-Geschäftsdaten, low-level:
  - Ist-Daten (actual data)
  - Plandaten (plan data)
- Kern-Geschäftsdaten, aggregiert:
  - KPIs Ist
  - KPIs Plan
  - KPIs Benchmark
- Zeit
  - historisch (versch. Stände)
  - Aktuell
  - Forecast
- Statistische Auswertungen:
  - Min

- Max
- Mittelwert
- Median
- usw.
- Fachliche Drill-Down-Dimensionen der Berichts- und Steuergrößen (KPIs).

Beispiele:

- Geographischer Markt
- Marktsegment
- Produkt(kategorie)
- Legal Entity
- Organisationseinheit

Pro KPI sind unterschiedliche Dimensions-Kombinationen relevant.

## Organisatorische Daten

Das sind ergänzende Daten für den das Management, die üblicherweise nicht unbedingt Data Governor verantwortet werden, deren Entstehung und Verarbeitung er jedoch unterstützt.

- Versionsstände:

### Technische Versionierung

- Aktuelle Version
- Ältere (interne) Versionsstände

### Fachliche Versionierung<sup>1</sup>

- Aktuelle Version
- Ältere (interne) Versionsstände

- Korrekturstände<sup>2</sup>

- Rohdaten unkorrigiert
- Korrigierte Daten

- Kommentare

## Organisatorische Metadaten

Das sind die Daten für den Data Governor. Er gibt diese, sofern überhaupt, nur als entspr. Meta-Information nach außen, und nicht in den normalen Reports.

- „Reife“:<sup>3</sup>
  - vorläufig / indikativ

1 Hier nur der Vollständigkeit halber aufgeführt. Es gibt Experten, die generell davon abraten, innerhalb eines DWH / BI-Systems fachlich zu versionieren. Dem schließe ich mich generell an. Allerdings gibt es von Zeit zu Zeit Sonderfälle, in denen man legitimerweise von dieser Regel abweichen kann/muss.

2 Hier gemeint: Korrekturen, die in den dispositiven Systemen gemanagt werden. Und es sind auch nicht Stände im Rahmen der üblichen Versionierung gemeint, die oben erwähnt sind, sondern vielmehr manuelle Einzelkorrekturen aus ganz bestimmten Gründen, die innerhalb des BI-Systems gepflegt werden. Auch hierzu gibt es Experten-Meinungen, die sagen, so etwas sollte es grundsätzlich nicht geben. Auch hier schließe ich mich generell der Meinung an, und auch hier erkenne ich aber auch legitime Ausnahmen und Sonderfälle an.

3 Ein Grenzfall – diese Information kann durchaus auch direkt in einen Report einfließen.

- intern gereviewt / verlässlich
- extern auditiert / juristisch verbindlich
- DQ-Daten
  - DQ-Regeln
  - DQ-KPIs
- Strukturelle Meta-Daten / Planungs- und Umsetzungs-Status:
  - Datenanforderungen (Datenanforderungen, noch nicht implementiert)
  - Implementierter Stand (implementierte Datenanforderungen)

## Auswirkungen

Sie sehen: Es kommt ganz schön was zusammen. Vermutlich habe ich noch ein paar Sonderfälle vergessen. Es ist unter den genannten Dimensionen nicht alles mit allem frei kombinierbar, aber die potenziell mögliche Anzahl der Kombinationen kommt ist von der theoretisch maximalen Menge manchmal gar nicht so weit entfernt, je nach Beschaffenheit aus Ausbaustufe des Systems.

In jedem Fall haben wir es nicht nur mit riesigen Datenmengen zu tun (das ist ja ohnehin jedem klar), sondern auch mit riesigen Mengen an Informations-Kategorien, wenn man es einmal genauer betrachtet. Damit erklärt sich natürlich auch guter Teil der hohen Kosten von BI-Systemen, nicht nur in der Anschaffung, sondern auch im Betrieb.

Eine der Kernfragen in den nächsten Teilen wird nun sein, wie man denn dieser großen Menge an Informations-Ausprägungen Herr wird. Folgende Aussagen dazu können wir bereits jetzt machen:

- Die Informations-Ausprägungen und -strukturen sind sehr komplex und vielschichtig. Ihre Verwaltung ist technisch wie organisatorisch eine Herausforderung.
- Ähnlich, wie das Management eine geeignete Aggregation der Daten in einem Report braucht, so braucht auch der Data Governor eine geeignete Aggregation dieser (Meta-)Daten bzw. Datenausprägungen. Dazu werden wir ein paar Objekte auf höherer Abstraktionsebene als „fachliche Organisationseinheit“ einführen.
- Das ist aufgrund der schier Menge nicht mehr mit einfachen Mitteln wie Listen in Word-Dokumenten, Excel-Dateien o.ä. zu bewerkstelligen. Es braucht eine professionelle, spezialisierte IT-Lösung.

Lesen Sie mehr im nächsten Teil.