

RDA Deutschland Tagung, Potsdam, 25.–27.02.2020

## DATA GOVERNANCE

### ALS VORAUSSETZUNG FÜR DATENEXZELLENZ AN FORSCHUNGSORGANISATIONEN

Johannes Sautter<sup>1</sup>, Andrea Wuchner<sup>2</sup>

<sup>1</sup>Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, johannes.sautter@iao.fraunhofer.de

<sup>2</sup>Fraunhofer-Informationszentrum Raum und Bau IRB, andrea.wuchner@irb.fraunhofer.de

#### DATENEXZELLENZ

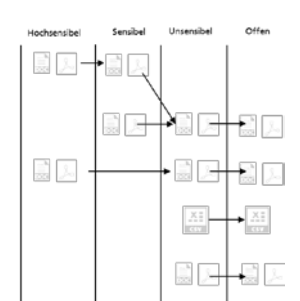
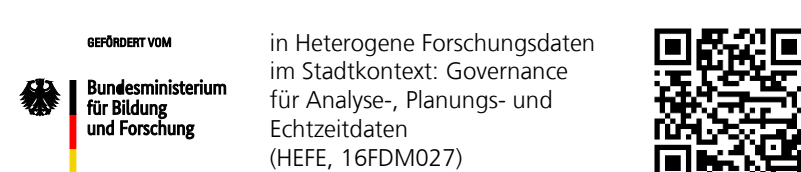
Jeder Wissenschaftler und jede Wissenschaftlerin steht vor der immer komplexer werdenden Herausforderung, mit Daten angemessen umzugehen. Anforderungen an Daten ergeben sich einerseits aus dem Transparenzanspruch von Open Science und der Interoperabilität sensibler und geheimer Daten andererseits. Wichtig ist jedoch zu verhindern dass die dafür notwendige Standardisierung der Methodendurchführung nicht zur Entstehung getrennter Datensilos führt.

Datenexzellenz bezieht sich nicht auf ein Datenobjekt, sondern auf die Organisation, in der es erstellt wurde (Sautter et al. 2018). Wissenschaftsorganisationen stehen vor der Herausforderung, Maßnahmen zu ergreifen, um ihre Datenexzellenz sicherzustellen.

#### DATA GOVERNANCE

Data Governance ist ein Rahmenkonzept in einer Organisation, welches festlegt welche Entscheidungen im Kontext von Daten zu treffen sind und wer sie trifft (Otto Österle 2016). Die Data Governance-Funktion dient dem Data Management als Führungsfunktion. Sie stellt durch Organisation und Moderation Datenprodukte her, die den FAIR<sup>1</sup>-Kriterien genügen. Dazu konzipiert und pflegt ein Data Governance-Team unter anderem ein Glossar sowie ein Daten- und Prozessmodell nach dem Grundsatz »so universell und zentral wie möglich«.

<http://s.fhg.de/datenexzellenz-forschung>  
Hinterlassen Sie uns Ihre Kontaktdaten, um über Neuigkeiten und Publikationen informiert zu werden.

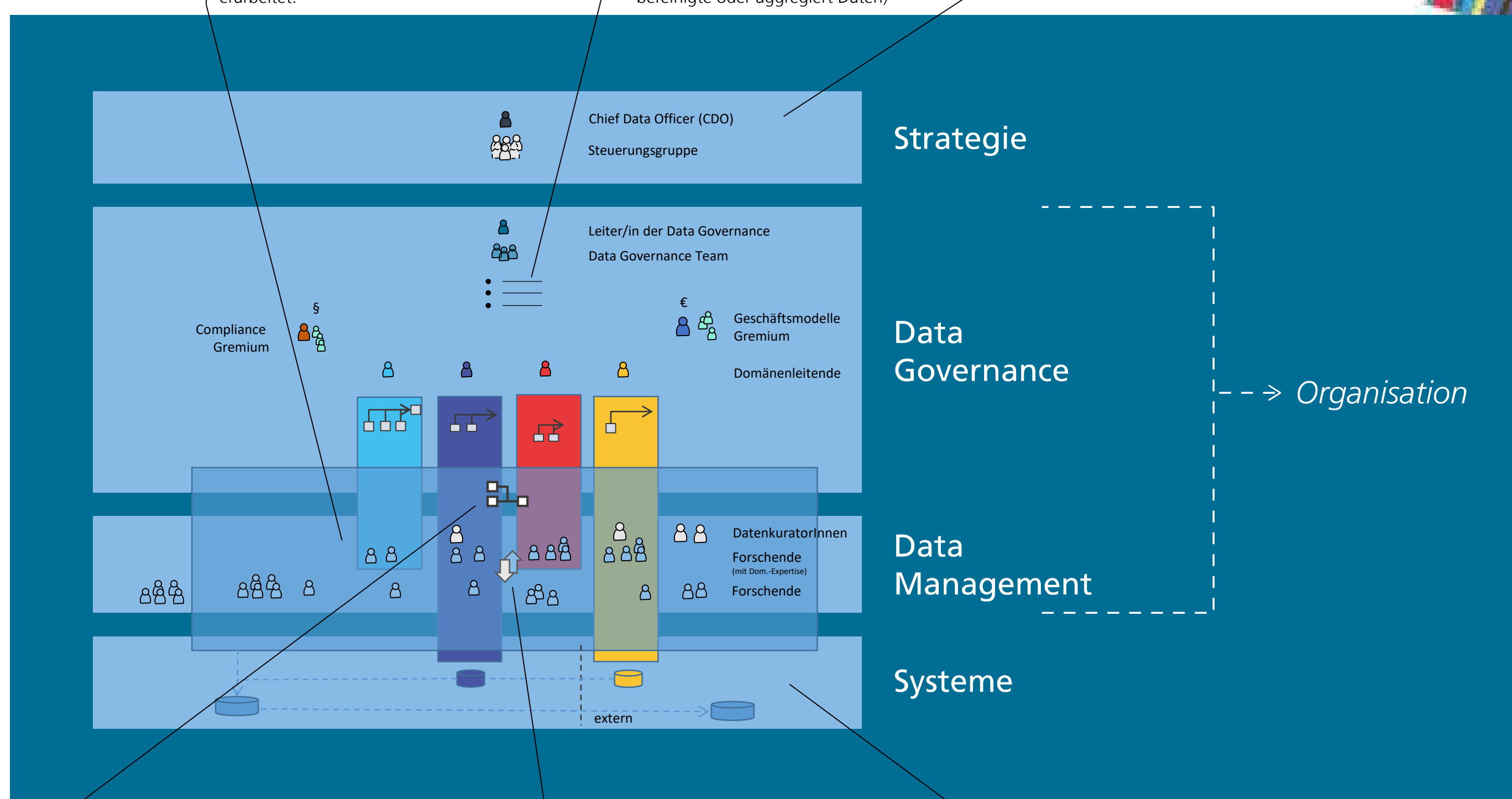


Forschende sowie Fachinformationsmanager und Fachinformationsmanagerinnen werden durch definierte **Prozesse zum Transfer von Dateien zwischen Sensibilitätsstufen** unterstützt. Die zugrundeliegenden Checklisten werden vom Data Governance Team mit Unterstützung des Compliance Gremiums sowie der Domänenleitenden erarbeitet.

**Geschäftsregeln** und ein **Glossar** erläutern zentrale Begriffe, sowie deren Zusammenhänge. Am relevantesten sind die Begriffe:

- Domäne
- Datensatz
- Metadaten (Bibliogr., Inhaltsbeschr.)
- Dateien (Roh- oder Metadaten, bereinigte oder aggregiert Daten)

Um eine strategische Richtung vorzugeben, ohne die Wissenschaftsfreiheit einzuschränken, benötigt die Forschungsorganisation eine möglichst konkrete **Strategie** für Forschungsdaten. Eine Forschungsdatenpolicy wird hierbei als Vorstufe einer Strategie verstanden.



Im Mittelpunkt des **domänenübergreifenden Meta-Datenmodells** steht der Datensatz. Ein Datensatz entsteht aus einem Projekt und wird mit bestimmten Metadatenfeldern beschrieben, welche im Datenmodell dargestellt sind. Jeder Datensatz hat einen Status, welcher sich auf den Forschungsprozess bezieht. Je nach Status besteht er aus Dateien und weiteren Entitäten wie bspw. Software. Weitere Entitäten beschreiben Methoden und Forschungssoftware.

Ein iterativer »**Forschungsprozess der Zukunft**« steuert das phasenorientierte Zusammenspiel der Rollen Forschungsprojektleitende, Forschende, Domänenleitende sowie Bibliothekarinnen und Bibliothekare. Phasenorientierte Regelwerke realisieren zweckorientierte Datenqualität. Mittels intern frei zugänglicher Metadaten wird Wissensmanagement zu geplanten, stattfindenden oder vergangene Erhebungen möglich.

Mittels einer **Systemarchitektur** und entsprechender Prozessautomatisierung wird angestrebt, dass Daten für Projektpartnerinnen und Projektpartner sowie Forschende komfortabel an einer Stelle in bestmöglicher Datenqualität nichtredundant abgelegt werden als sog. »Single Version of Truth« (Dobrokhotova et al 2015). Die jederzeitige volle Datensouveränität des Dateneigentümers bzw. der Dateneigentümerin ist dabei essentiell.

#### DATA MANAGEMENT

Das Data Management führt die Prozesse der Data Governance operativ aus und ermöglicht somit Datenexzellenz. Eine frühe Metadaten-Dokumentation und Standards für Datenschutzerklärungen sowie Erhebungsmethoden vereinfachen den Forschenden die Arbeit.

#### ZUSAMMENFASSUNG & AUSBLICK

Als Vision für die Zukunft der Wissenschaftslandschaft muss sowohl eine Standardisierung jeder Wissenschaftsdisziplin erfolgen, als auch eine interdisziplinäre Integration. Data Governance leistet dazu einen wesentlichen Beitrag und realisiert eine Entlastung der Wissenschaftlerinnen und Wissenschaftler gegenüber den komplexer werdenden Anforderungen unserer Zeit.

#### Kernthemen eines Pilotprojekts am Fraunhofer IAO

- Werteformulierende Policy für Datenexzellenz als Verantwortung des Gesamtinstituts
- Identifikation der Datendomänen, qualitative Leitfadengespräche, quantitative Befragungen sowie Bauwerks- und Mobilitätsgeodaten
- Ernennung von Domänenleitenden und Gründung von Domänengremien
- Strategisches Steuerungsgremium auf Leitungsebene mit Verantwortlichen aus Bibliothek, IT und Compliance sowie Domänenleitenden
- Zentrales Data Governance-Team
- Compliance-Gremium
- Operative Hilfsmittel wie Methodenchecklisten für Forschende
- »Fraunhofer Datenraum« für die interne Datenablage avisiert
- Datenveröffentlichung in Fordatis, dem Forschungsdaten-Repository der Fraunhofer-Gesellschaft

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz. Das Key Visual rechts (© bei Tartila AdobeStock) ist hiervon ausgenommen. <http://creativecommons.org/licenses/by/4.0/>

#### LITERATUR

Sautter, J., Litauer, R., Fischer, R., Klages, T., Wuchner, A., Müller, E., Schaj, G., Dobrokhotova, E., Drews, P. and Riess, S.: Beyond Data Quality: Data Excellence Challenges from an Enterprise, Research and City Perspective. In Proceedings of the 7th International Conference on Data Science, Technology and Applications (DATA 2018), pages 245-252, ISBN: 978-989-758-318-6, <http://dx.doi.org/10.5220/0006912902450252>  
Otto, Boris, Österle, Hubert (2016): Corporate Data Quality. Voraussetzung erfolgreicher Geschäftsmodelle. 1. Aufl. 2016. Berlin: Gabler; Springer Gabler. <http://cdq-buch.de/>  
Dobrokhotova, Ekaterina; Engelbach, Wolf; Sautter, Johannes (2015): Marktstudie 2015 Multidomänen-Stammdatenmanagementsysteme. Stuttgart: Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO). <http://publica.fraunhofer.de/documents/N-364091.html>

© Tartila AdobeStock