

# Data-Steward(-ship) in der Digitalisierung des Gesundheitswesens

Im Rahmen der fortschreitenden Digitalisierung im Gesundheitswesen diversifiziert sich die benötigte Expertise im Berufsfeld des Medizinischen Informationsmanagements und der Medizinischen Informatik. Dies gilt im Besonderen für Tätigkeitsbereiche in der Datenbereitstellung und -aufbereitung für die medizinische Forschung und die Gesundheitsversorgung, für die sich in den letzten Jahren zunehmend die Bezeichnung Data Stewardship bzw. Data Steward als Rollenbezeichnung etabliert hat. Mit der Frage einer definitorischen Beschreibung dieses Tätigkeitsbereichs bzw. dieser Rolle insbesondere im Hinblick auf die für diesen Tätigkeitsbereich erforderlichen Kompetenzen haben sich Vertreter:innen des DVMD und der GMDS, namentlich der Arbeitsgruppen (AG) Datenmanagement (DM) und Curricula der Medizininformatik (CMI) sowie Vertreter:innen der NFDI4Health in mehreren Sitzungen und Workshops auseinandergesetzt. Hierauf aufbauend haben die GMDS-AGs DM und CMI in Abstimmung mit dem DVMD eine Begriffsdefinition erarbeitet, die im Folgenden vorgestellt wird.

## Gemeinsame Begriffsdefinition Data Steward (-ship) der AG Datenmanagement (DM) und Curricula der Medizininformatik (CMI) der GMDS

Aus Sicht des Datenmanagements eignet sich als Ausgangspunkt der Definition der Rolle bzw. des Tätigkeitspektrums eines Data Stewards die Betrachtung des Lebenszyklus der Daten, den Data Stewards in ihrer Tätigkeit partiell begleiten. Dieser umfasst Tätigkeiten der Modellierung, Erhebung/Erfassung, Speicherung, Manipulation, Transformation/Konvertierung nach national und international gültigen Vorschriften und Guidelines, Übertragung, Publikation, Aggregation, Archivierung, Abfrage und Löschung von Daten (s. Abb. 1).

Hierauf aufbauend kann in Anlehnung an [1] verallgemeinert auf Tätigkeiten im Kontext der Gesundheitsversorgung oder der medizinischen Forschung die Rolle des Data Stewards wie folgt definiert werden:

### Definition Data Steward

Eine Person, die dafür verantwortlich ist, die Qualität, Integrität und Zugriffsregelungen von Daten und Metadaten in Übereinstimmung mit geltendem Recht, institutionellen Richtlinien und individuellen Berechtigungen zu wahren. Data Stewardship beinhaltet den professionellen und sorgfältigen Umgang mit Daten in allen Phasen eines Datenverarbeitungsprozesses. Ein Data Steward soll gewährleisten, dass Daten in

allen Phasen des Datenlebenszyklus (d.h. Modellierung, Erfassung, Verarbeitung, Analyse, Aufbewahrung, Datenfreigabe und Wiederverwendung) adäquat verarbeitet werden.

Aus Sicht des Datenmanagements werden die Tätigkeiten eines Data Stewards in zwei unterschiedliche Perspektiven eingeordnet, sodass der Begriff Data Steward(-ship) sowohl aus fachlicher als auch aus technischer oder organisatorischer Perspektive betrachtet werden kann (vgl. [2], Seite 19).

Während fachliches Data Steward(-ship) insbesondere inhaltliche Expertise und Beratung aus dem eigentlichen Anwendungs-/Einsatz- und Erhebungsbereich der Daten bietet, ist technisches Data Stewardship vor allem verantwortlich für das Design, die Implementierung und den Einsatz relevanter Hard- und Software-Infrastruktur und für die Prozessierung und Verwaltung der Daten.

Fachliches und technisches Data Stewardship ist abhängig voneinander und wird in der Praxis partiell oder auch vollständig in Personalunion wahrgenommen.

**Fachliche Data Stewards** (z.B. klinische Data Stewards) sind für die inhaltliche Korrektheit von Datenmodellen verantwortlich und besitzen Kenntnis über die Herkunft und den Fachbereich der erhobenen Daten, welche kontrolliert und strukturiert erfasst werden. Sie arbeiten eng mit den Expert:innen des jeweiligen Fachbereichs zusammen (z.B. Onkologie, Kardiologie oder Pathologie) und arbeiten an der Schnittstelle als Vermittlungs- und Kontrollinstanz zwischen der rein technischen Ebene und Anwendungsebene (s. Abb. 2). Des Weiteren haben sie eine unterstützende Funktion bei der Durchführung klinischer Studien (Studienunterlagen, Einwilligung, Rekrutierung, Erfassung der Daten und Überwachung der Datenqualität), wobei hier eine Überschneidung zum Begriff der Study Nurse gesehen werden kann.

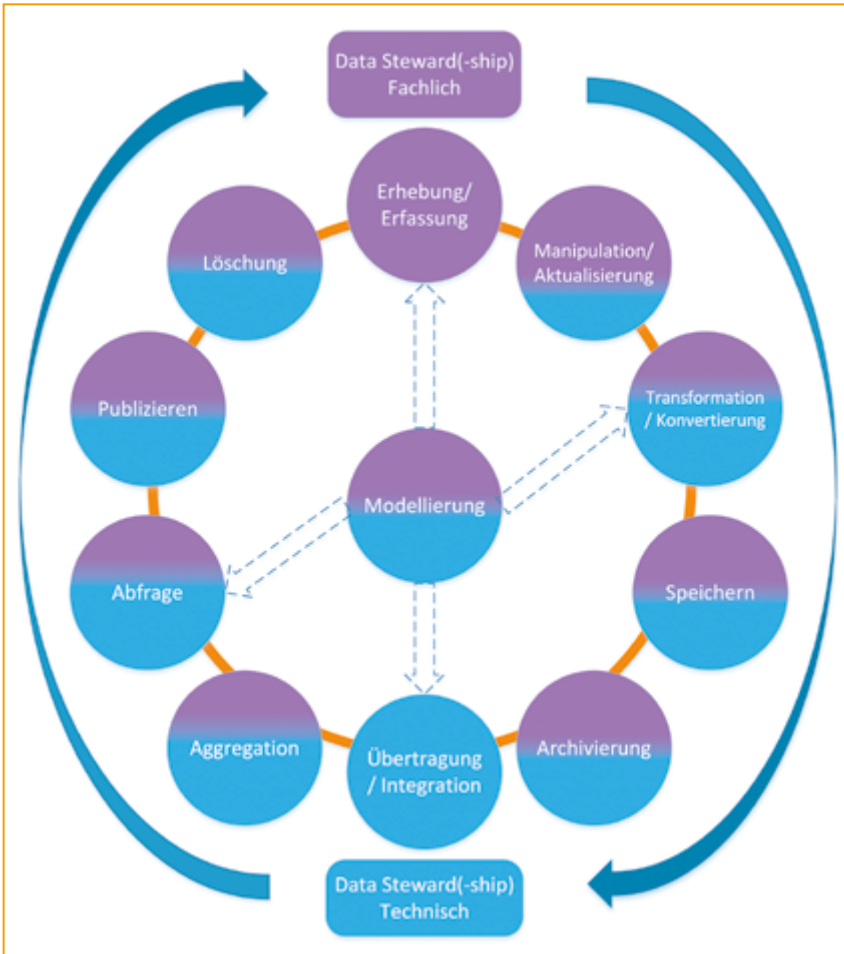
**Technische Data Stewards** (zur Abgrenzung ggü. den fachlichen Data Stewards gelegentlich auch als **Data Custodians** bezeichnet ([2], S. 19)) sind für die technische Anwendung und Umsetzung von Datenmodellen verantwortlich und besitzen Kenntnis über den korrekten Einsatz bestimmter Modellierungsmethoden (bspw. Mappingbasierte Modellierung, Modellierung von Templates und Eingabeformularen) und Technologien (z.B. openEHR oder FHIR auf Ebene der relevanten Interoperabilitäts- bzw. Datenmodellierungsstandards sowie z.B. Data Warehouses auf Ebene der relevan-



**Matthias Katzensteiner, M.A.**  
WiMi HS Hannover  
Leitung GMDS-AG Datenmanagement (DM)  
Matthias.Katzensteiner@hs-hannover.de

Als Autoren gelten die GMDS-Arbeitsgruppen »Datenmanagement in klinischen und epidemiologischen Studien – gemeinsame Arbeitsgruppe der Fachbereiche Med. Informatik, Biometrie und Epidemiologie« (AG DM) und »Curricula der Medizinischen Informatik« (AG CMI). Die AGs werden vertreten von den jeweiligen Leiter:innen.

**Annett Müller**  
Vorsitzende des DVMD  
annett.mueller@dm.de  
**Marie-Louise Witte, M.A.**  
WiMi HS Hannover  
Stv. Leitung GMDS-AG Datenmanagement (DM)  
Marie-Louise.Witte@hs-hannover.de  
**Nina Schewe, M.A.**  
WiMi im PLRI für Medizinische Informatik der MHH und der TU Braunschweig  
Stv. Leitung GMDS-AG Datenmanagement (DM)  
Schewe.Nina@mh-hannover.de  
**Prof. Dr.-Ing. Oliver J. Bott, HS Hannover**  
Leitung GMDS-AG Curricula der Medizinischen Informatik (CMI)  
Oliver.Bott@hs-hannover.de



**Abb. 1: Lebenszyklus von Daten in Abhängigkeit zur Rolle und Aufgaben**

ten Datenbanktechnologien). Einheitliche Modelle dienen letztlich dazu, die Schnittstelle zu zum Beispiel anderen medizinischen Datenintegrationszentren (DIZ) einzurichten. Ziel dabei ist es, die ETL-Prozesse aus den Systemen zu begleiten. Sie können spezifische Fragen zum korrekten Einsatz von Kombinationen verschiedener Methoden beantworten und sind für die Umsetzung verantwortlich. Technische Data Stewards arbeiten eng mit Entwickler:innen (z.B. Software- oder Datenbankentwickler:innen) zusammen und fungieren an der Schnittstelle als Vermittlungs- und Kontrollinstanz zwischen der rein technischen Ebene und der Anwendungsebene (s. Abb. 2).

Data Stewardship ist damit an der Schnittstelle zwischen End-Anwendung und Software- sowie Datenbankentwicklung verortet, wobei die tatsächliche Ausrichtung (technisch oder fachlich) vom konkreten Tätigkeits- und Aufgabenfeld abhängt. Hier

ran ist die Schwierigkeit erkennbar, Aufgaben des Data-Stewardships eindeutig der fachlichen oder technischen Ebene zuzuordnen.

Für die Ausübung des Data Stewardships ist eine fundierte Ausbildung durch ein i.d.R. grundständiges einschlägiges Studium notwendig. Je nach Ausbildung ist es durchaus möglich, dass nur eine der beiden Perspektiven von Data Stewards bedient werden kann, da entweder der fachliche Hintergrund oder das technische Wissen nicht in der nötigen Tiefe gegeben ist.

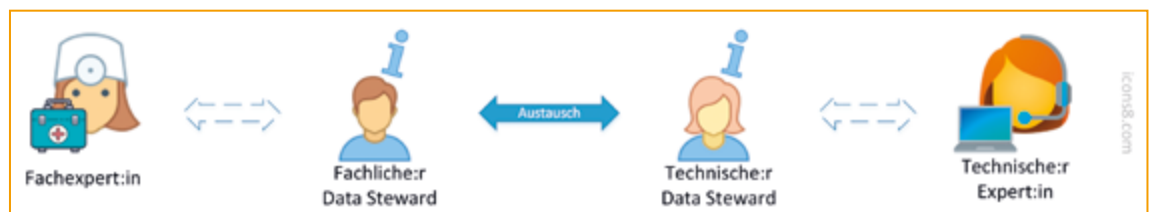
Als optimale Ausbildung für den Einsatz als Data Steward an der Schnittstelle zwischen fachlicher und technischer Perspektive eignet sich für den Gesundheitsbereich insbesondere das Studium »Medizinisches Informationsmanagement«. Hier werden neben dem Anwendungswissen aus Medizin mit den Ausprägungen Gesundheitsversorgung und medizinische Forschung dokumentarische, informatische und medizininformatische Kenntnisse zu Datenbanktechnologien, Programmierung und Datenmodellierung vermittelt.

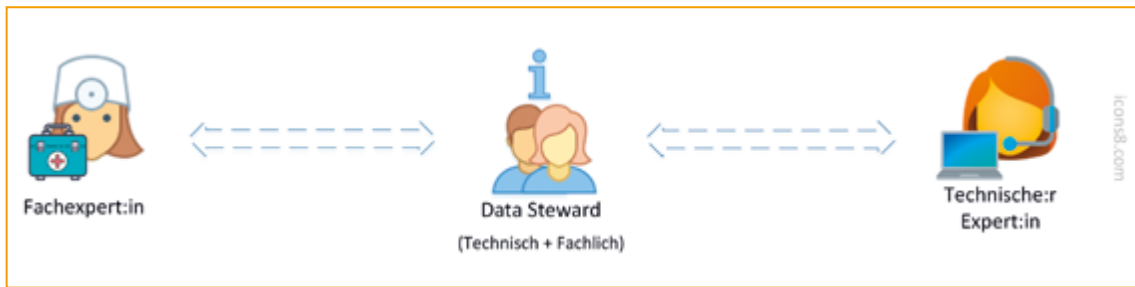
Ausgehend von einer fachbezogenen Ausbildung (z.B. Studium der Human-/Tiermedizin oder Biologie) kann durch einschlägige Weiterqualifikation (z.B. Masterstudiengang des Medizinischen Informationsmanagements) oder eine entsprechende Zusatzausbildung ein grundlegendes technisches Verständnis erlernt werden, dass die Tätigkeit als fachliche:r Data Steward ermöglicht. Für den Einsatz als technische:r Data Steward sind dagegen umfangreichere informatische Kenntnisse notwendig, die in der Regel eine grundlegende, einschlägig medizininformatische Ausbildung bedingen.

Umgekehrt kann durch technische Ausbildung (z.B. Studium der angewandten Informatik) ein tiefes technisches Verständnis aufgebaut und eine fachliche Expertise geboten werden. Jedoch ist es einer fachfremden Person nicht möglich, inhaltliche Aussagen über die zu verwaltenden Daten zu treffen. Reine technische Expertise (z.B. in der Software- oder Datenbankentwicklung) kann ohne inhaltliches Hintergrundwissen nicht nutzengenerierend eingesetzt werden.

Durch eine fundierte Ausbildung in einem praktischen Anwendungsgebiet, die sowohl technische als auch anwendungsfachbezogene Kenntnisse vermittelt, kann ein Data Steward sowohl fachliche als auch technische Expertise und somit beide Ausprägungen

**Abb. 2: Zusammenarbeit zwischen Data Stewards fachlicher und technischer Ausprägung und Expert:innen**





**Abb. 3: Zusammenarbeit von Data Steward in Personalunion und Expert:innen**

gungen der Rolle Data Steward vereinen. In diesem Fall kann eine Rollendifferenzierung ausgelassen werden, da ein solcher Data Steward beide Rollen bekleiden und zwischen Fachexpert:innen und technischen Expert:innen vermitteln kann (s. Abb. 3).

Eine spezifische Facette der Rolle des fachlichen Data Stewards, die sich auch auf Aufgaben im Kontext der Gesundheitsversorgung beziehen kann, ist die des **forschungsbezogenen Data Stewards** (engl. Research Data Steward). Diese sind im Kontext des Forschungsdatenmanagements (FDM) zuständig für die langfristige und nachhaltige Pflege von Forschungsdaten, vom Studiendesign bis zur Datenerhebung, -analyse, -speicherung und -weitergabe. Es umfasst alle Aktivitäten, die erforderlich sind, um sicherzustellen, dass digitale Forschungsdaten langfristig auffindbar, zugänglich, interoperabel und nachnutzbar (FAIR) sind, einschließlich Datenmanagement, Archivierung und Nachnutzung durch Dritte [5]. Hervorzuheben ist die Aufgabe der disziplinspezifischen Beratung und Schulung im FDM von Forschenden [6].

Das Data Stewardship ist damit Teil der **Data Governance** eines Unternehmens. Data Governance wiederum umfasst alle Organisationseinheiten, Rollen, Entscheidungsregeln und Verantwortlichkeiten von Personen und Informationssystemen in dem Umfang, in dem diese an informationsverarbeitenden Prozessen des Unternehmens beteiligt sind [3]. Das Data Stewardship eines Unternehmens bzw. die mit der Rolle eines Data Stewards verknüpften Aufgaben orientieren sich an den im Rahmen der Data Governance auf strategischer und taktischer Ebene festgelegten Leitlinien bzw. Policies [3]. Insofern sind Rahmenseetzungen des Data Stewardships auf Managementebene des Unternehmens angesiedelt, Data Stewards arbeiten dieser Rahmenseetzung bzw. der Definition entsprechender Policies zu (vgl. [4], [6]).

Neben der Bezeichnung Data Steward werden in der Literatur häufig synonym, partiell aber auch voneinander abgegrenzt, die Bezeichnungen **Data Manager**, **Data Curator** oder **Data Librarian** verwendet [6]. Eindeutiger abgrenzbar sind die Rollenbezeichnungen **Data Analyst**, da hier die Kompetenz der aufgabenbezogenen Anwendung von insbesondere statistischen Analysemethoden zur Vorbereitung von Entscheidungen hinzu kommt

sowie die Rollenbezeichnung **der Datenwissenschaftlerin/des Datenwissenschaftlers** (engl. **Data Scientist**) mit darüber hinausgehenden, tiefgehenden analytischen Fähigkeiten einschließlich vertieften Kenntnissen mathematischer Algebra und der Fähigkeit zum Modellieren von Algorithmen ([7], S. 10) mit dem Ziel des Gewinnens von Information und Wissen aus Daten ([6], S. 3). Der Komplexität des Aufgabenfeldes eines Data Scientists ist es geschuldet, dass die Vermittlung entsprechender Kompetenzen in der Regel Gegenstand eines wissenschaftlichen Masterstudiums ist [7]. In der beruflichen Praxis wird die Rolle des Data Stewards oft auch mit der Rolle des Data Analysts in Personalunion verknüpft bzw. sind beide Kompetenzgebiete Gegenstand der einschlägigen Ausbildung z.B. im Rahmen eines Bachelorstudiums des Medizinischen Informationsmanagements. ■

#### Quellen

- [1] Jetten, M., Grootveld, M., Mordant, A., Jansen, M., Bloemers, M., Miedema, M., & van Gelder, C. W. G. (2021). Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship. Zenodo.
- [2] Gluchowski, P. (Hrsg.) (2020): Data Governance – Grundlagen, Konzepte, Anwendungen. dpunkt.verlag.
- [3] The Data Governance Institute (DGI) (2020). The DGI Data Governance Framework, 2020; [https://datagovernance.com/wp-content/uploads/2020/07/dgi\\_data\\_governance\\_framework.pdf](https://datagovernance.com/wp-content/uploads/2020/07/dgi_data_governance_framework.pdf). Zuletzt zugegriffen am 09.12.2022.
- [4] Martateperek, W. by. (2019, Januar 25). Data Stewardship at TU Delft – 2018 report. Open Working. <https://openworking.wordpress.com/2019/01/25/data-stewardship-at-tu-delft-2018-report/>
- [5] Jansen, P., van den Berg, L., van Overveld, P., & Boiten, J.-W. (2019). Research data stewardship for healthcare professionals. In *Fundamentals of Clinical Data Science* (S. 37–53). Springer International Publishing.
- [6] Steinke, B., Hausen, D., Kuberek, M., Hora, M., Kessler, K., Kramer, C., Fuhrmans, M., Grunwald-Eckhardt, L., Hermann, S., Müller-Pfefferkorn, R. et al. (2022). Data Stewards an den TU9-Universitäten – Bestandsaufnahme, Handlungsfelder und Kooperationspotenzial. Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managerinnen
- [7] Lübcke M., Wannemacher, K. (2018). : Vermittlung von Datenkompetenzen an den Hochschulen: Studienangebote im Bereich Data Science. HIS-HE, Hannover 2018. URL: <https://medien.his-he.de/publikationen/detail/vermittlung-von-datenkompetenzen-an-den-hochschulen-studienangebote-im-bereich-data-science> Zuletzt zugegriffen am 09.12.2022.