

VEREIN
DEUTSCHER
INGENIEURE

Datenmanagement in den Life Sciences
Data management for life sciences

VDI 6320

Ausg. deutsch/englisch
Issue German/English

Die deutsche Version dieser Richtlinie ist verbindlich.

The German version of this standard shall be taken as authoritative. No guarantee can be given with respect to the English translation.

Inhalt	Seite	Contents	Page
Vorbemerkung	2	Preliminary note.....	2
1 Anwendungsbereich	2	1 Scope	2
2 Begriffe	3	2 Terms and definitions	3
3 Abkürzungen	9	3 Abbreviations	9
4 Grundsätze	9	4 Principles	9
5 Anleitungen	15	5 Instructions	15
5.1 Organisatorische Datenverfügbarkeit.....	16	5.1 Organisational data availability	16
5.2 Maschinelle Verarbeitbarkeit.....	34	5.2 Machine processability	34
5.3 Nutzbarkeit (Integrität, Aussagekraft, Nachvollziehbarkeit).....	39	5.3 Usability (integrity, significance, traceability).....	39
5.4 Qualitätssicherung.....	46	5.4 Quality assurance.....	46
Anhang A Risikomatrix	50	Annex A Risk matrix.....	50
Anhang B Beispiel – Datenmanagementplan	51	Annex B Example – Data management plan.....	57
Anhang C Beispiel – Arbeitsanweisung ELN	63	Annex C Example – Working instruction ELN ...	66
Anhang D Beispiel – Datenformate	69	Annex D Example – Data formats.....	70
Schrifttum	71	Bibliography	71

VDI-Gesellschaft Technologies of Life Sciences (TLS)
Fachbereich Biotechnologie

VDI-Handbuch Biotechnologie
VDI-Handbuch Informationstechnik, Band 1: Angewandte Informationstechnik
VDI-Handbuch Medizintechnik
VDI-Handbuch Verfahrenstechnik und Chemieingenieurwesen, Band 2: Planung/Projektierung

Vorbemerkung

Der Inhalt dieser Richtlinie ist entstanden unter Beachtung der Vorgaben und Empfehlungen der Richtlinie VDI 1000.

Alle Rechte, insbesondere die des Nachdrucks, der Fotokopie, der elektronischen Verwendung und der Übersetzung, jeweils auszugsweise oder vollständig, sind vorbehalten.

Die Nutzung dieser Richtlinie ist unter Wahrung des Urheberrechts und unter Beachtung der Lizenzbedingungen (www.vdi.de/richtlinien), die in den VDI-Merkblättern geregelt sind, möglich.

Allen, die ehrenamtlich an der Erarbeitung dieser Richtlinie mitgewirkt haben, sei gedankt.

Eine Liste der aktuell verfügbaren und in Bearbeitung befindlichen Blätter dieser Richtlinienreihe sowie gegebenenfalls zusätzliche Informationen sind im Internet abrufbar unter www.vdi.de/6320.

1 Anwendungsbereich

Diese Richtlinie betrifft die Grundregeln des guten Datenmanagements im Bereich der Life-Science-Forschung in nicht regulierten Umgebungen. Sie richtet sich an organisatorisch Verantwortliche für die Erhebung und Verarbeitung von digitalen Daten im Zusammenhang mit forschungsorientierten Aktivitäten (z.B. Projekt- oder Laborleiter, deren Vorgesetzte, Beauftragte für die IT oder das Qualitätsmanagement).

Die Richtlinie beschreibt Aktivitäten zur Umsetzung guten Datenmanagements vornehmlich innerhalb einer Organisation oder Organisationseinheit und grenzt sich damit zu den sogenannten FAIR-Prinzipien ab (FAIR steht für „*to be findable, accessible, interoperable, re-usable*“ – Daten sollen also auffindbar, zugänglich, interoperabel und wiederverwendbar sein (vgl. [1])). Während die FAIR-Prinzipien in erster Linie Forschende adressieren und Anforderungen an die Annotierung von Daten in öffentliche Repositorien stellen, spricht diese Richtlinie darüber hinaus auch das Management einer Organisation und andere verantwortliche Bereiche an, z.B. die IT. Die Kategorien der hier aufgestellten Regeln sind in Bezug auf den Mehrwert für eine Organisation formuliert. Im Gegensatz dazu sind die FAIR-Prinzipien in Bezug auf die weitere Nutzbarkeit von Daten und Metadaten strukturiert. Da die FAIR-Prinzipien für das Forschungsdatenmanagement immer mehr an Bedeutung gewinnen, erklärt diese Richtlinie im Abschnitt 4, welche der hier formulierten Regeln zur Erfüllung der FAIR-Prinzipien angewendet werden müssen.

Preliminary note

The content of this standard has been developed in strict accordance with the requirements and recommendations of the standard VDI 1000.

All rights are reserved, including those of reprinting, reproduction (photocopying, micro copying), storage in data processing systems and translation, either of the full text or of extracts.

The use of this standard without infringement of copyright is permitted subject to the licensing conditions (www.vdi.de/richtlinien) specified in the VDI Notices.

We wish to express our gratitude to all honorary contributors to this standard.

A catalogue of all available parts of this series of standards and those in preparation as well as further information, if applicable, can be accessed on the Internet at www.vdi.de/6320.

1 Scope

This standard concerns the basic rules of good data management in the field of life sciences research in non-regulated environments. It is aimed at those organisationally responsible for the collection and processing of digital data in connection with research-oriented activities (e.g. project or laboratory managers, their superiors, IT, or quality management representatives).

The standard describes activities for the implementation of good data management primarily within an organisation or organisational unit and thus differentiates itself from the so-called FAIR principles (FAIR stands for “*to be findable, accessible, interoperable, re-usable*” – data should therefore be findable, accessible, interoperable and reusable (cf. [1])). While the FAIR principles primarily address the researcher and place requirements on the annotation of data in public repositories, this standard also addresses the management of an organisation and other responsible areas, e.g., IT. The categories of the rules set out here are formulated in terms of the added value for an organisation. In contrast, the FAIR principles are structured in terms of the further usability of data and metadata. As the FAIR principles are becoming increasingly important for research data management, this standard explains in Section 4 which of the rules formulated here must be applied to fulfil the FAIR principles.

Die Richtlinie dient als Leitfaden für

- die Sicherung der Nutzbarkeit und Nachvollziehbarkeit der Daten über die konkrete Form der Datenerhebung hinaus,
- den Austausch der Daten mit Projektpartnern und Kunden sowie in der wissenschaftlichen Gemeinschaft und
- die Erstellung organisationspezifischer Datenmanagement-Arbeitsanweisungen.

Die Anwendung der Richtlinie

- erleichtert die Weiterverwertung wissenschaftlicher Ergebnisse,
- schafft Handlungssicherheit für die Verantwortlichen (Ausführende *und* Auftraggeber),
- unterstützt dabei, Konformität mit Richtlinien von Fördergebern und wissenschaftlichen Verlagen zu erreichen und
- kann auch als erste Übersicht über relevante Standards in diesem Themengebiet genutzt werden.

Die Richtlinie gilt im forschungsorientierten Umfeld und beschränkt sich ausdrücklich auf nicht GMP-regulierte Bereiche. Darüber hinaus werden in dieser Richtlinie keine technischen Einzelfestlegungen (Datenformate, Protokolle, Schnittstellen) getroffen und keine dateninhaltlichen Aspekte (Vokabulare, Ontologien, minimale Attributmengen) vorgegeben oder entwickelt.

The standard serves as a guide for

- ensuring the usability and traceability of the data beyond the concrete form of data collection
- the exchange of data with project partners and customers and in the scientific community, and
- the creation of organisation-specific data management work instructions.

The application of the standard

- facilitates the further exploitation of scientific results,
- creates security of action for those responsible (contractors *and* clients),
- helps to achieve conformity with the guidelines of funding bodies and scientific publishers and
- can also be used as a first overview of relevant standards in this subject area.

The standard applies in a research-oriented environment and is explicitly limited to non-GMP areas. In addition, this standard does not contain any individual technical specifications (data formats, protocols, interfaces) and no data content aspects (vocabularies, ontologies, minimum attribute sets) are specified or developed.