

DIN ISO 26162-1



ICS 01.020; 35.240.30

Entwurf

Einsprüche bis 2019-02-07
Vorgesehen als
teilweiser Ersatz für
DIN ISO 26162:2016-12

**Management von Terminologieressourcen –
Terminologiedatenbanken –
Teil 1: Design (ISO/DIS 26162-1:2018);
Text Deutsch und Englisch**

Management of terminology resources –
Terminology databases –
Part 1: Design (ISO/DIS 26162-1:2018);
Text in German and English

Gestion des ressources terminologiques –
Bases de données terminologiques –
Partie 1: Conception (ISO/DIS 26162-1:2018);
Texte en allemand et anglais

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2018-12-07 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfs besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal von DIN unter www.din.de/go/entwuerfe bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an nat@din.de möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/go/stellungnahmen-norm-entwuerfe oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an den DIN-Normenausschuss Terminologie (NAT), 10772 Berlin, Burggrafenstr. 6, 10787 Berlin.

Die Empfänger dieses Norm-Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 46 Seiten

DIN-Normenausschuss Terminologie (NAT)

Inhalt

	Seite
Nationales Vorwort.....	3
Vorwort	4
Einleitung	5
1 Anwendungsbereich.....	6
2 Normative Verweisungen	6
3 Begriffe	6
3.1 Begriffe	7
3.2 Terminologiedatenbanken	8
4 Erstellung einer Terminologiedatenbank.....	10
4.1 Allgemeine Grundsätze	10
4.2 Terminologisches Metamodell	11
4.3 Datenkategorien	12
4.3.1 Allgemeine Überlegungen	12
4.3.2 Datenkategorietypen	13
4.3.3 Gemeinsame Ressourcen.....	15
4.3.4 Begriffsbeziehungen.....	15
4.4 Begriffseinträge.....	17
4.4.1 Begriffsorientierung.....	17
4.4.2 Sprache.....	17
4.4.3 Abhängigkeit und Wiederholbarkeit von Datenkategorien	18
4.4.4 Datengranularität.....	19
4.4.5 Datenelementarität	19
4.4.6 Datenmodellierungsvarianz.....	20
4.5 Rollen.....	20
Anhang A (informativ) Auszug aus einer TMF-Metamodell-konformen Terminologiedatenbank.....	21
Literaturhinweise.....	24

Nationales Vorwort

Das internationale Dokument ISO/DIS 26162-1:2018 ist unverändert in diesem Norm-Entwurf übernommen worden.

Das zuständige deutsche Normungsgremium ist der NA 150-00-05 AA „Systeme für die Verwaltung von Terminologie, Wissen und Content“ im DIN-Normenausschuss Terminologie (NAT).

Änderungen

Gegenüber DIN ISO 26162:2016-12 wurden folgende Änderungen vorgenommen:

- a) komplette inhaltliche und redaktionelle Überarbeitung;
- b) Aufteilung der Norm in mehrere Teile;
- c) Übernahme von ISO/DIS 26162-1:2018 als DIN ISO 26162-1 und von ISO/DIS 26162-2:2018 als DIN ISO 26162-2;
- d) in Teil 1 geht es um die Gestaltung von Terminologiedatenbanken;
- e) in Teil 2 geht es um die wesentlichen Funktionen des Terminologieverwaltungssystems;
- f) die Zukunft von Teil 3 ist auf ISO-Ebene derzeit unklar, in ihm geht es ggf. um den Inhalt von Terminologiedatenbanken.

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung nationaler Normungsorganisationen (ISO-Mitgliedsorganisationen). Die Erstellung von Internationalen Normen wird üblicherweise von Technischen Komitees von ISO durchgeführt. Jede Mitgliedsorganisation, die Interesse an einem Thema hat, für welches ein Technisches Komitee gegründet wurde, hat das Recht, in diesem Komitee vertreten zu sein. Internationale staatliche und nichtstaatliche Organisationen, die in engem Kontakt mit ISO stehen, nehmen ebenfalls an der Arbeit teil. ISO arbeitet bei allen elektrotechnischen Themen eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Direktiven, Teil 1 beschrieben. Es sollten insbesondere die unterschiedlichen Annahmekriterien für die verschiedenen ISO-Dokumentenarten beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Direktiven, Teil 2 erarbeitet (siehe www.iso.org/directives).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der erhaltenen Patenterklärungen (siehe www.iso.org/patents).

Jeder in diesem Dokument verwendete Handelsname dient nur zur Unterrichtung der Anwender und bedeutet keine Anerkennung.

Eine Erläuterung zum freiwilligen Charakter von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO) hinsichtlich technischer Handelshemmnisse (TBT) berücksichtigt, enthält der folgende Link: www.iso.org/iso/foreword.html.

Dieses Dokument wurde vom Technischen Komitee ISO/TC 37, *Language and terminology*, Subkomitee SC 3, *Management of terminology resources*, erarbeitet.

Diese erste Ausgabe von ISO 26162-1 ersetzt zusammen mit ISO 26162-2 die ISO 26162:2012, die technisch überarbeitet wurde.

Eine Auflistung aller Teile der Normenreihe ISO 26162 ist auf der ISO-Internetseite abrufbar.

Einleitung

Terminologien sind die Gesamtheit von Begriffen in bestimmten Fachgebieten, die durch Bezeichnungen dargestellt und durch Verwendung zusätzlicher terminologischer Daten beschrieben werden. Im Allgemeinen sind diese Daten in strukturierten Terminologiedatenbanken organisiert und werden in der Regel in spezifischen Softwareanwendungen wie Terminologieverwaltungssystemen bearbeitet. Terminologiedatenbanken unterscheiden sich in der Regel in Bezug auf die ihnen zugrunde liegenden Datenmodelle und bestehen aus unterschiedlichen Zusammenstellungen von Datenkategorien, während Terminologieverwaltungssysteme auf unterschiedlichen Plattformen von mobilen Geräten bis hin zu Desktop-Computern, Client-Server- oder webbasierten Architekturen von Einzelpersonen, Unternehmen und staatlichen Behörden eingesetzt werden.

Diese Normenreihe enthält Anleitungen zur Gestaltung von Terminologiedatenbanken und zu wesentlichen Funktionen von Terminologieverwaltungssystemen. Die Reihe kann auch verwendet werden, um die Eignung von bestehenden Terminologiedatenbanken und Terminologieverwaltungssystemen zu bewerten.

1 Anwendungsbereich

Die vorliegende Internationale Norm legt allgemeine, d. h. implementierungs- und anwendungsunabhängige, Prinzipien für die Erstellung von Terminologiedatenbanken fest, um ein Maximum an Effizienz und Qualität im Bereich der Terminologearbeit und damit bei der Erstellung, Verarbeitung und Verwendung hochwertiger Terminologie zu ermöglichen. Die Zielgruppe dieses Dokuments sind Terminologen, Übersetzer, Technische Redakteure, Sprachplaner, Fachexperten und Entwickler von Terminologieverwaltungssystemen. Die Norm kann jedoch auch zur Bewertung der Konformität und Eignung von Terminologiedatenbanken herangezogen werden.

Die vorliegende Internationale Norm beschreibt einen maximalen Ansatz für die Erstellung von Terminologiedatenbanken in dezentralen, mehrsprachigen Arbeitsumgebungen. Sie kann aber auch für kleinere Lösungen verwendet werden, für die nur eine Teilmenge des vorgeschlagenen Datenmodells erforderlich ist.

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 704, *Terminology work — Principles and methods*

ISO 1087, *Terminology work — Vocabulary*

ISO 12620, *Management of terminology resources — Data category specifications*

ISO 16642, *Computer applications in terminology — Terminological markup framework*

ISO 23185, *Assessment and benchmarking of terminological resources — General concepts, principles and requirements*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- ISO Online Browsing Platform: verfügbar unter <http://www.iso.org/obp>
- IEC Electropedia: verfügbar unter <http://www.electropedia.org/>

3.1 Begriffe

3.1.1

Objekt

alles Wahrnehmbare oder Vorstellbare

Anmerkung 1 zum Begriff: Objekte können materiell (z. B. eine Maschine, ein Blatt Papier, ein Diamant), immateriell (z. B. ein Konversionsverhältnis, ein Projektplan) oder erdacht (z. B. ein Einhorn, eine wissenschaftliche Hypothese) sein.

Anmerkung 2 zum Begriff: Objekte können Veränderungen unterliegen, die eine Änderung des Begriffs oder eine Änderung der Bezeichnung nach sich ziehen.

[QUELLE: ISO/DIS 1087:2018, 3.1.1, modifiziert — In Anmerkung 1 zum Begriff wurde „darf“ zu „kann“ geändert und eine Anmerkung 2 zum Begriff wurde ergänzt.]

3.1.2

Begriff

Wissenseinheit, die durch eine einmalige Kombination von Merkmalen gebildet wird

Anmerkung 1 zum Begriff: Begriffe sind nicht zwangsläufig an bestimmte Sprachen gebunden. Sie werden jedoch durch den sozialen oder kulturellen Hintergrund beeinflusst, der häufig zu verschiedenen Kategorisierungen führt.

Anmerkung 2 zum Begriff: Aufgrund ihres dynamischen Charakters werden Begriffe auch als Denkeinheiten definiert (siehe ISO 704:2009, 5.1 und DIN 2342:2011-08, 4.1).

[QUELLE: ISO/DIS 1087:2018, 3.2.3, modifiziert — Anmerkung 2 zum Begriff wurde entfernt und durch eine neue Anmerkung 2 zum Begriff ersetzt.]

3.1.3

Bezeichnung

Bezeichner

Darstellung eines *Begriffs* (3.1.2) durch ein Zeichen, das ihn bezeichnet

Anmerkung 1 zum Begriff: In der Terminologearbeit werden vier Bezeichnungsarten unterschieden: *Benennungen* (3.1.4), Eigennamen, Namen und Symbole.

Anmerkung 2 zum Begriff: Bezeichnungen können verbale oder nonverbale bzw. eine Kombination aus beidem sein.

[QUELLE: ISO/DIS 1087:2018, 3.4.1]

3.1.4

Benennung

sprachliche *Bezeichnung* (3.1.3) die einen allgemeinen Begriff in einem bestimmten Fachgebiet darstellt

BEISPIEL „Planet“, „Turm“, „Stift“, „Ziffer“, „Anzahl“, „Quadratwurzel“, „Logarithmus“, „Maßeinheit“, „Logarithmusbasis“, „chemisches Element“, „chemische Verbindung“, „HP Laserjet 1100“, „Nobelpreis für Physik“.

Anmerkung 1 zum Begriff: Benennungen können Buchstaben und Kurzzeichen, Ziffern, mathematische Symbole sowie Wort- und Satzzeichen beinhalten (z. B. Interpunktionszeichen wie Bindestriche, Klammern, eckige Klammern und sonstige Verbindungs- und Trennzeichen), u. a. auch in Zeichenformaten (d. h. bestimmte Schriftarten und Fett, Kursiv, Fettkursiv oder sonstige Darstellungsweisen), die fachgebiets- oder sprachspezifischen Konventionen unterliegen.

[QUELLE: ISO/DIS 1087:2018, 3.4.2, modifiziert — „Bezeichnung“ wurde durch „verbale Bezeichnung“ ersetzt und Anmerkung 1 zum Begriff wurde entfernt]

3.2 Terminologiedatenbanken

3.2.1

Terminologiedatenbank

Termbank

Datenbank, die eine *Terminologiedatensammlung* (3.2.4) enthält

[QUELLE: ISO/FDIS 30042:2018, 3.28]

3.2.2

Datenmodell

graphische und/oder lexikalische Darstellung von Daten, in der die Eigenschaften, die Struktur und die Beziehungen der Daten zueinander festgelegt werden

[QUELLE: ISO/IEC 11179-1:2015, 3.2.7]

3.2.3

terminologisches Metamodell

Datenmodell (3.2.2), das die Grundlage für die Erstellung von *Terminologiedatensammlungen* (3.2.4) beschreibt

3.2.4

Terminologiedatensammlung

TDS

Datensammlung, die aus *Begriffseinträgen* (3.2.7) sowie den zugehörigen Metadaten und dokumentarischen Informationen besteht

[QUELLE: ISO/FDIS 30042:2018, 3.29]

3.2.5

globale Informationen

GI

technische und verwaltungstechnische Informationen, die für die gesamte *Terminologiedatensammlung* (3.2.4) gelten

BEISPIEL Beispiele für globale Informationen sind Titel von Terminologiedatensammlungen, Änderungshistorien oder Eigentümer- und Copyright-Informationen.

[QUELLE: ISO 16642:2017, 3.11, modifiziert — Anmerkung 1 zum Begriff wurde durch „BEISPIEL“ ersetzt und als vollständiger Satz formuliert.]

3.2.6

Zusatzinformationen

ZI

Informationen, die den Inhalt von *Begriffseinträgen* (3.2.7) ergänzen und für die gesamte *Terminologiedatensammlung* (3.2.4) gelten

BEISPIEL Typische Beispiele für Zusatzinformationen sind hierarchische Fachgebietsklassifikationen, komplette Beschreibungen der Institution, bibliographische Angaben und Verweise auf Textkorpora.

[QUELLE: ISO 16642: 2017, 3.2, modifiziert — „terminologische Einträge“ wurde in der Definition durch „Begriffseinträge“ und „Anmerkung 1 zum Begriff“ durch „BEISPIEL“ ersetzt.]

3.2.7

Begriffseintrag

BE

terminologischer Eintrag

Teil einer *Terminologiedatensammlung* (3.2.4), der terminologische Daten zu einem *Begriff* (3.1.2) enthält

Anmerkung 1 zum Begriff: Ein Begriffseintrag kann Informationen zu einer oder mehreren Sprachen enthalten.

[QUELLE: ISO/FDIS 30042:2018, 3.3, modifiziert — Abkürzung wurde hinzugefügt, „Anmerkung 1 zum Begriff“ wurde geändert.]

3.2.8

Sprachebene

S

Teil eines *Begriffseintrags* (3.2.7), der Informationen zu einer bestimmten Sprache enthält

[QUELLE: ISO 16642:2017, 3.13, modifiziert – „Begriffseintrag“ wird anstelle von „terminologischer Eintrag“ verwendet.]

3.2.9

Benennungsebene

B

Teil der *Sprachebene* (3.2.8), der Informationen zu einer *Benennung* (3.1.4) enthält

[QUELLE: ISO 16642:2017, 3.20]

3.2.10

Benennungskomponentenebene

BK

Teil der *Benennungsebene* (3.2.9), der linguistische Informationen zu den Bestandteilen einer *Benennung* (3.1.4) enthält

[QUELLE: ISO 16642:2017, 3.19]

3.2.11

Datenkategorie

Verallgemeinerung eines Datenfeldes in einer Datenbank

BEISPIEL /definition/

Anmerkung 1 zum Begriff: In diesem Dokument werden Datenkategorienamen im Fließtext zwischen Schrägstrichen (/) angegeben.

[QUELLE: ISO/DIS 12620:2018, 3.2]

3.2.12

Begriffsorientierung

Grundsatz, demzufolge ein *Begriffseintrag* (3.2.7) einen einzigen *Begriff* (3.1.2) beschreibt

Anmerkung 1 zum Begriff: Ein Begriffseintrag kann mehrere Benennungen enthalten, die somit als semantisch äquivalent gelten.

3.2.13

Benennungsautonomie

Grundsatz, demzufolge die *Benennungen* (3.1.4) in einem *Begriffseintrag* (3.2.7) als eigenständige Teileinheiten betrachtet werden, die sich mit einem identischen Satz von *Datenkategorien* (3.2.11) beschreiben lassen

3.2.14

Datengranularität

Grad der Datengenauigkeit

BEISPIEL Der Datenkategoriesatz /Wortklasse /, /Genus / und /Numerus / verfügt z. B. über eine höhere Datengranularität als die Datenkategorie /Grammatik/.

3.2.15

Datenelementarität

Grundsatz, demzufolge ein Datenfeld nur ein einziges Datenelement enthält

BEISPIEL Werden z. B. sowohl Vollform als auch Abkürzung in ein und demselben Datenfeld aufgeführt, so stellt dies einen Verstoß gegen die Datenelementarität dar.

3.2.16

Datenmodellierungsvarianz

Varianz im Hinblick auf die Zuordnung von *Datenkategorien* (3.2.11) zu *Datenmodellen* (3.2.2) infolge unterschiedlicher Philosophien bezüglich der Informationsanordnung in einem *Begriffseintrag* (3.2.7)

4 Erstellung einer Terminologiedatenbank

4.1 Allgemeine Grundsätze

Die Erstellung einer Terminologiedatenbank erfordert grundlegende terminologiewissenschaftliche und terminologiarbeitstechnische Kenntnisse. Aus diesem Grund, und um qualitativ hochwertige Ergebnisse zu erzielen, sind folgende Ressourcen heranzuziehen:

- festgelegte Benennungen und Definitionen nach ISO 1087;
- terminologische Grundsätze und Verfahren nach ISO 704;
- Datenmodellierungskriterien nach ISO 16642 und ISO 12620;
- Metriken der Gebrauchstauglichkeit nach ISO 23185.

Terminologiedatenbanken verfügen über eine logische Struktur, die in einem allgemeinen hierarchischen Datenmodell beschrieben wird (siehe 4.2). Dieses Datenmodell enthält verschiedene Ebenen, denen konkrete Datenkategorien (siehe 4.3) zugeordnet werden können. Dieser Datenmodellierungsansatz bietet die notwendige Flexibilität, um den unterschiedlichen Anforderungen bei der Erstellung von Terminologiedatenbanken in Bezug auf spezifische Arbeitsprofile (Erstellen von Texten, Übersetzung, fachliche Prüfung usw.) und organisationsspezifische Erfordernisse (Freiberufler, Übersetzungsagenturen, interne Abteilungen von Unternehmen oder Organisationen, usw.) gerecht werden zu können. Daher muss bereits in der Frühphase der Konzeption einer Terminologiedatenbank ein langfristiger und detaillierter Managementplan aufgestellt werden, der alle absehbaren Benutzergruppen sowie organisatorische und technische Fragen berücksichtigt, um umfangreiche zeit- und kostenintensive Änderungen nach Abschluss des Designprozesses zu vermeiden.

4.2 Terminologisches Metamodell

Terminologiedatenbanken müssen dem in ISO 16642:2017 festgelegten terminologischen Metamodell (oder einem Teil davon) entsprechen (siehe Bild 1 und Bild 2). Den Kern des Metamodells stellt das Prinzip der Begriffsorientierung dar (siehe 4.4.1), d. h.:

- in einer Terminologiedatenbank beschreibt ein Eintrag (Begriffseintrag, BE) einen Begriff, der weiter beschrieben werden kann über:
- zusätzliche Unterebenen zur Instanziierung von Sprachen und sprachbezogenen Informationen (Sprachebene, S-Ebene), die weitere Informationen enthalten zu:
- Benennungen und benennungsbezogenen Informationen (Benennungsebene, B-Ebene);
- einzelnen Wörtern einer Mehrwortbenennung oder einem der Bestandteile einer Einzelwortbenennung, wie etwa einem Morphem (Benennungskomponentenebene, BK-Ebene).

Darüber hinaus enthält das Metamodell den Begriffseinträgen übergeordnete Ebenen, auf denen weitere Informationen abgelegt werden können:

- globale Informationen (GI), die für die gesamte Terminologiedatenbank gelten (Name der Terminologiedatenbank, Institution oder Person, die die Datei erzeugt, Copyright-Informationen, Historie usw.);
- Zusatzinformationen (ZI), wie vollständige bibliographische oder verwaltungstechnische Informationen, binäre Daten oder Picklistenwerte, die für alle Begriffseinträge verfügbar sind und nicht für verschiedene Begriffseinträge wiederholt werden.

Eine schematische Darstellung der oben genannten Ebenen des terminologischen Metamodells zeigt Bild 1.

Bild 2 stellt die Metamodellebenen, ihre Beziehungen und Kardinalitäten in Form eines UML-Diagramms dar.

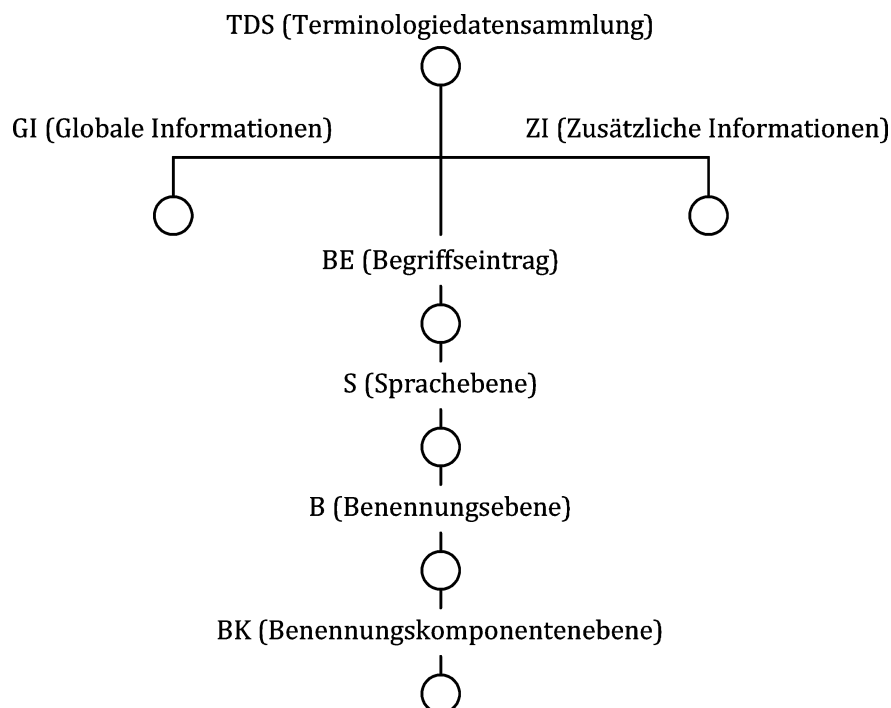
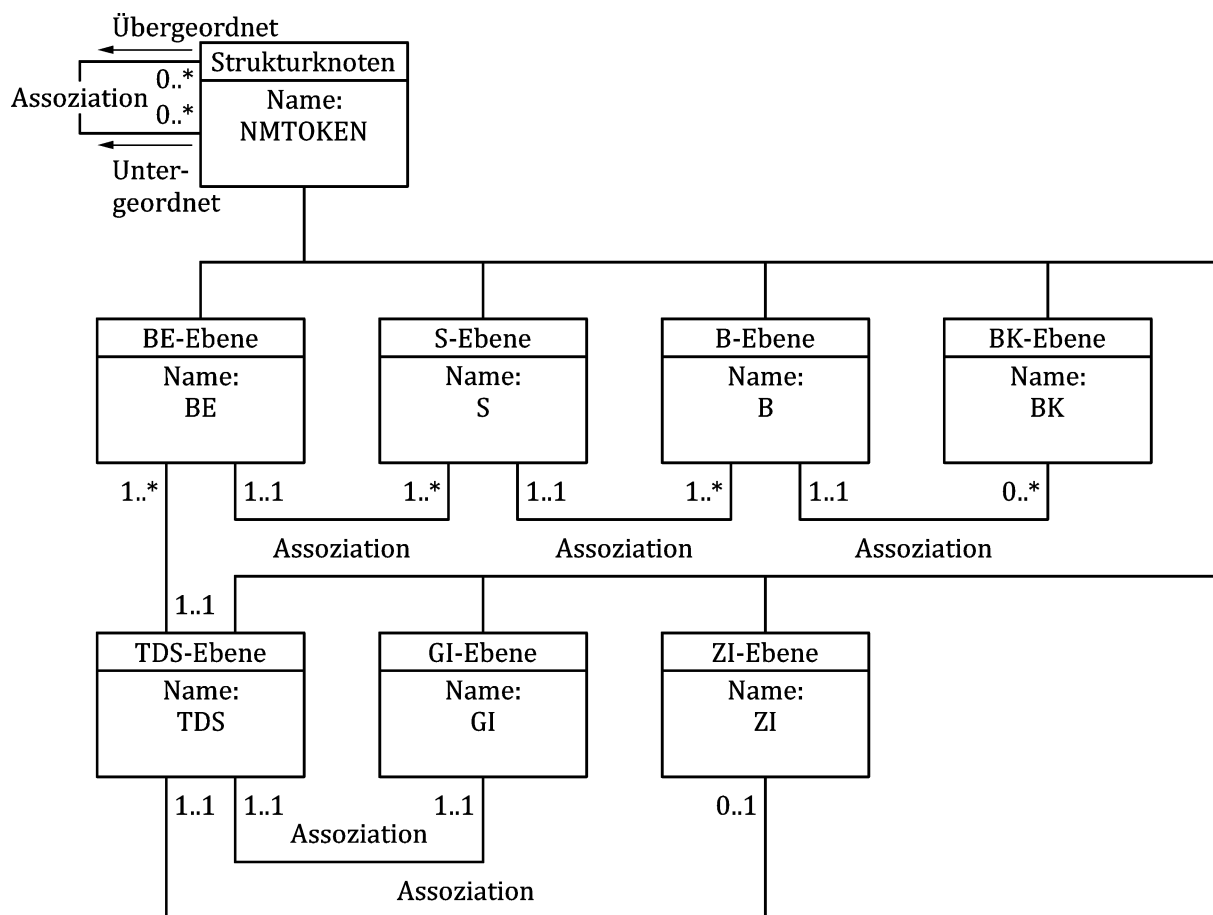


Bild 1 — Terminologisches Metamodell — Schematische Darstellung (abgeänderte Fassung aus ISO 16642:2017)



Kardinalitäten:

1..1 = muss genau einmal vorkommen

1.* = muss mindestens einmal, kann aber auch mehrmals vorkommen

0..1 = kann fehlen oder darf einmal vorkommen

0..* = kann fehlen und darf mehrmals vorkommen

Bild 2 — Terminologisches Metamodell — UML-Diagramm (abgeänderte Fassung aus ISO 16642:2017)

Ein Beispiel eines Terminologiedatenbankauszugs, der dem terminologischen Metamodell entspricht, ist in Anhang A enthalten.

4.3 Datenkategorien

4.3.1 Allgemeine Überlegungen

Begriffseinträge enthalten bestimmte Informationseinheiten, wie z. B. Benennungen, Definitionen und Kontexte. Jede Klasse dieser Informationseinheiten wird als Datenkategorie bezeichnet. Bei der Erstellung von Terminologiedatenbanken können Datenkategorien natürlich nach spezifischen Nutzungsanforderungen ausgewählt werden. In den letzten Jahrzehnten haben Terminologieexperten jedoch relevante Datenkategorien gesammelt, standardisiert, klassifiziert und in anerkannten Datenkategorieregistern wie DatCatInfo veröffentlicht (siehe www.datcatinfo.net). So ist z. B. /grammatical gender/ eine Datenkategorie, die normalerweise als Feldname verwendet wird, um typischerweise die Werte „feminine“, „masculine“, „neuter“ usw. zu instanziiieren. Es kann natürlich vorkommen, dass die von anerkannten Datenkategorieregistern vorgeschlagen Datenkategorienamen nicht den Nutzungsanforderungen entsprechen. Falls Anpassungen an Datenkategorienamen vorgenommen werden müssen, ist dafür Sorge zu tragen, inhaltliche Überschneidungen oder Fehlinterpretationen zu vermeiden.

ISO 10241-1 und ISO 12616 enthalten eine Liste gängiger Datenkategorien zur Instanziierung von:

- Benennungen (in jeder beliebigen Sprache);
- benennungsbezogenen Informationen (z. B. grammatischen Angaben, Benennungstypen, Sprachregistern, Status, Kontexten usw.);
- sprachbezogenen Informationen (z. B. Definitionen, Anmerkungen usw.);
- begriffsbezogenen Informationen (z. B. Fachgebieten, Definitionen, Beispielen, Anmerkungen, Abbildungen usw.);
- verwaltungstechnischen und bibliographischen Informationen:
 - Bezeichnern unterschiedlicher Art (etwa zur Identifikation von Produkten oder Projekten, auf die sich die Benennungen beziehen);
 - Erstellungs- und Änderungsdatum sowie Namen der Autoren, die den Begriffseintrag bzw. Teile davon erstellt oder geändert haben;
 - Eintragsstatus (z. B. „vorgeschlagen“, „in Bearbeitung“, „freigegeben“);
 - Quellen von Benennungen, Definitionen, Kontexten, Anmerkungen usw.

Werden Datenkategorien benötigt, die nicht in anerkannten Datenkategorieregistern vorgehalten werden, sind die Anforderungen zum Erstellen, Dokumentieren, Harmonisieren und Pflegen von Datenkategorie-spezifikationen nach ISO 12620 zu berücksichtigen.

4.3.2 Datenkategorietypen

Terminologiedatenbanken enthalten in der Regel unterschiedliche Datenkategorietypen, abhängig von der Art der darin enthaltenen Informationen. Eine sorgfältige Auswahl der Datenkategorietypen ist die Voraussetzung für hochqualitative Inhalte einer Terminologiedatenbank.

4.3.2.1 Offene und geschlossene Datenkategorien

Offene Datenkategorien ermöglichen die Instanziierung freien Texts, der ihrer Datenkategoriedefinition entspricht. Beispielsweise gehört /term/ zu den offenen Datenkategorien, da die Benennung, die im entsprechenden Datenfeld erfasst werden kann, nicht vorhersehbar ist. Weitere Beispiele für offene Datenkategorien sind /definition/ oder /context/.

Im Gegensatz dazu gibt es andere Datenkategorien, die lediglich einen oder mehrere Werte aus einer begrenzten Anzahl zulässiger Werte umfassen können. Sie werden als geschlossene Datenkategorien bezeichnet, da ihr Inhalt aus einer geschlossenen Wertemenge besteht. Werden z. B. deutschsprachige Benennungen dokumentiert, so kann die Datenkategorie /grammatical gender/ nur die Werte „masculine“, „feminine“ oder „neuter“ annehmen. Neben /grammatical gender/ sind /part of speech/, /term type/, /geographical usage/, /administrative status/ oder /subject field/ typische Vertreter geschlossener Datenkategorien.

Auf der anderen Seite können Werte aber auch selbst als Datenkategorie fungieren. In manchen Terminologiedatenbanken könnten Nutzer z. B. für /masculine/ die Booleschen Werte „ja|nein“ oder „richtig|falsch“ verwenden wollen.

Die Verwendung geschlossener Datenkategorien stellt sicher, dass vorgegebene Wert ausgewählt werden können oder müssen, wodurch die Eingabe falscher oder abweichender Schreibweisen vermieden wird. Steht es den Benutzern dagegen frei, wie etwa das Genus eines Substantivs beschrieben werden soll, könnte

„masculine“, „masc.“ oder einfach „m.“ für ein maskulines Substantiv eingetragen werden. Eine einheitliche Darstellung dieser Werte ist die Grundlage für konsistente Terminologiedatenbanken und stellt sicher, dass Suchanfragen, Datenfilterungen und sonstige Datenverwaltungsaufgaben zu optimalen Ergebnissen führen.

4.3.2.2 Obligatorische, fakultative, systemgenerierte und voreingestellte Datenkategorien

Datenkategorien können obligatorisch oder fakultativ sein. Werden obligatorische Datenfelder nicht ausgefüllt, lassen sich Begriffseinträge in Terminologieverwaltungssystemen in der Regel nicht speichern. Zu den Mindestanforderungen einer Terminologiedatenbank gehört, dass für jeden Begriffseintrag mindestens eine Sprache und eine Benennung in dieser Sprache vorhanden sind.

Andere typische obligatorische Datenfelder sind etwa solche, die Informationen zum Fachgebiet oder zur Wortart enthalten, da diese Informationen für die Unterscheidung von Homographen relevant sein können. Die Festlegung obligatorischer Datenkategorien kann u. U. aber auch problematisch sein wie etwa im Fall der Datenkategorien /definition/ oder /context/, denn es in der Regel relativ aufwendig, geeignete Definitionen und Kontexte zu finden und sie zu besprechen. Es lassen sich aber noch komplexere Bedingungen implementieren, wie etwa die obligatorische Angabe der Quelle einer Definition oder des Genus bei Substantiven in bestimmten Sprachen. Die Produktivität beim Befüllen einer Terminologiedatenbank lässt sich steigern, wenn das Terminologieverwaltungssystem das Speichern von unvollständigen Begriffseinträgen im Entwurfsmodus ermöglicht, sodass obligatorische Informationen zu einem späteren Zeitpunkt hinzugefügt werden können.

Bei systemgenerierten Datenfeldern werden die Informationen nicht von den Benutzern (per Hand) eingegeben, sondern vom Terminologieverwaltungssystem generiert. Viele Systeme weisen den Begriffseinträgen automatisch Eintragsnummern, Erstellungs- und Änderungsdatum sowie die Namen der Benutzer zu, die den Begriffseintrag erstellt oder geändert haben.

Bei der Erstellung von Terminologiedatenbanken muss auch entschieden werden, auf welchen Ebenen eines Begriffseintrags solche verwaltungstechnischen Informationen dokumentiert werden sollen. Gewöhnlich reicht es nicht aus, verwaltungstechnische Informationen nur auf der Begriffsebene zu erfassen, da verschiedene Mitarbeiter für unterschiedliche Teile eines Begriffseintrags – insbesondere für verschiedene Sprachen – zuständig sein können.

Zudem wird empfohlen, für bestimmte Datenkategorien Standardwerte festzusetzen. Möchte ein Benutzer etwa während einer Arbeitssitzung ausschließlich Benennungen dokumentieren, die mit einem bestimmten Projekt oder Ausgangstext zusammenhängen, so ist es vorteilhaft, die Werte für die Datenkategorien /project subset/ oder /source/ voreinstellen zu können. Sämtliche in der entsprechenden Arbeitssitzung erstellten Begriffseinträge erhalten dann automatisch diese Werte. Wenn feststeht, dass die meisten Benennungen in einer Terminologiedatenbank Substantive sein werden, spart es dem Benutzer Zeit, wenn /part of speech/ auf „Substantiv“ voreingestellt ist und der Benutzer diesen Wert für andere Wortarten ändern kann.

4.3.2.3 Datenkategorien mit Lese- und Schreibzugriff, Nur-Lesezugriff sowie ausgeblendete Datenkategorien

In einer Terminologiedatenbank muss es möglich sein, unterschiedliche auf die Anforderungen und Rollen der Benutzer zugeschnittene Zugriffsrechte festzulegen. Sichtbare Datenfelder, deren Wert geändert werden kann, sind Datenfelder mit „Lese- und Schreibzugriff“. Sichtbare Datenfelder, deren Wert nicht geändert werden kann, sind Datenfelder mit „Nur-Lesezugriff“. Ausgeblendete Datenfelder können verwendet werden, wenn die darin enthaltenen Informationen für manche Benutzer nicht relevant sind.

So kann es z. B. empfehlenswert sein, einem leitenden Terminologen Lese- und Schreibzugriff auf sämtliche Felder zu gewähren, insbesondere auf das Statusfeld des Begriffseintrags, um den entsprechenden Wert einzustellen („freigegeben“, „in Bearbeitung“, „abgelehnt“ usw.). Diese Person kann auch dafür verantwortlich sein, anderen Benutzern Berechtigungen für bestimmte Ebenen der Begriffseinträge zuzuteilen. Benutzer, die in den Begriffseinträgen nur bestimmte Sprachen bearbeiten, sollten nur für eben diese