

DIN ISO 26162-3



ICS 01.020; 35.240.30

**Management von Terminologieressourcen –
Terminologiedatenbanken –
Teil 3: Inhalt (ISO 26162-3:2023)**

Management of terminology resources –
Terminology databases –
Part 3: Content (ISO 26162-3:2023)

Gestion des ressources terminologiques –
Bases de données terminologiques –
Partie 3: Contenu (ISO 26162-3:2023)

Gesamtumfang 29 Seiten

DIN-Normenausschuss Terminologie (NAT)

Inhalt

	Seite
Nationales Vorwort	3
Nationaler Anhang NA (informativ) Literaturhinweise	4
Vorwort	5
Einleitung	6
1 Anwendungsbereich	7
2 Normative Verweisungen	7
3 Begriffe	7
4 Identifizierung von Benennungen	10
4.1 Anforderungen an die Auswahl von Benennungen	10
4.2 Unithood und Termhood	11
4.3 Korpora und Terminologieextraktion	12
5 Sammlung von terminologischen Daten	12
5.1 Anforderungen an Daten	12
5.1.1 Bewertungsverfahren	12
5.1.2 Qualitativ hochwertige Daten	13
5.1.3 Zweck der Termbank	13
5.1.4 Datenkorrektheit	13
5.1.5 Gebrauchstauglichkeit	14
5.2 Datenmodell	15
5.3 Datenkategorien	15
5.4 Datenportabilität	15
6 Validierung der Begriffseintragsqualität	15
6.1 Allgemeine Validierungskriterien	15
6.2 Fehlertypologie und Systemdesign	15
6.3 Fehlertypen	16
6.3.1 Termbank-Spezifikation und -Pflege	16
6.3.2 Fehlertypen, die der Aufnahme von Begriffseinträgen zugeordnet sind	16
6.3.3 Fehlertypen, die automatisch erzeugten Inhalten zugeordnet sind	16
6.3.4 Fehlertypen, die offenen Datenkategorien zugeordnet sind	17
6.3.5 Fehlertypen, die geschlossenen Datenkategorien zugeordnet sind	18
Anhang A (informativ) Bewertungsmodelle	19
A.1 Scoring-Modelle	19
A.2 Prüfen und Bearbeiten von Begriffseinträgen	19
A.3 Gewichtete Bewertungsskalen	19
Anhang B (informativ) Beispiel für einen Begriffseintrag	21
Anhang C (informativ) Beispiel für eine Fehlertypologie	24
Literaturhinweise	29

Bilder

Bild 1 — Termhood	11
-----------------------------	----

Tabellen

Tabelle A.1 — Beispiel eines Bewertungsbogens für Sprachblöcke	20
Tabelle C.1 — Fehlertypen	24

Nationales Vorwort

Diese Deutsche Norm ist die Übersetzung der Internationalen Norm ISO 26162-3:2023, die im Technischen Komitee ISO/TC 37 „*Language and terminology*“ erarbeitet wurde.

Das zuständige nationale Normungsgremium ist der Arbeitsausschuss NA 105-00-05 AA „Management von Terminologieressourcen“ im DIN-Normenausschuss Terminologie (NAT).

Wegen einer sprachlichen Ungenauigkeit in ISO 26162-3:2023 wurde in 5.1.4.3 eine Nationale Fußnote ergänzt. Zwei weitere Nationale Fußnoten wurden in 6.3.4 und 6.3.5 ergänzt.

Für die in diesem Dokument zitierten Dokumente wird im Folgenden auf die entsprechenden deutschen Dokumente hingewiesen:

ISO 9001	siehe	DIN EN ISO 9001
ISO 26162-1	siehe	DIN ISO 26162-1
ISO 26162-2:2019	siehe	DIN ISO 26162-2:2020-05

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN EN ISO 9001, *Qualitätsmanagementsysteme — Anforderungen*

DIN ISO 26162-1, *Management von Terminologieressourcen — Terminologiedatenbanken — Teil 1: Design*

DIN ISO 26162-2:2020-05, *Management von Terminologieressourcen — Terminologiedatenbanken — Teil 2: Software (ISO 26162-2:2019)*

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung nationaler Normungsinstitute (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 Normungsthemen 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 Directives, 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 Directives, 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.

Für eine Erläuterung des freiwilligen Charakters von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO, en: World Trade Organization) hinsichtlich technischer Handelshemmnisse (TBT, en: Technical Barriers to Trade) berücksichtigt, siehe www.iso.org/iso/foreword.html.

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

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

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Auflistung dieser Institute ist unter www.iso.org/members.html zu finden.

Einleitung

Manager, Lehrende und Zuständige für die Unterhaltung von Terminologiedatenbanken führen sowohl in regelmäßigen Abständen wie auch kontinuierlich Bewertungen von Terminologiedatenbanken mit Begriffseinträgen zu unterschiedlichen Zwecken durch:

- qualitätssichernde Validierung von terminologischen Datensammlungen in Wirtschafts-, Regierungs- und Nichtregierungsbereichen, und/oder
- formative Beurteilung und summative Evaluierung und Rückmeldung in Ausbildungs- und Schulumgebungen.

ISO 26162-1 und ISO 26162-2 spezifizieren Gestaltungsprinzipien und Überlegungen hinsichtlich Software zur Modellierung von Terminologiedatenbanken (Termbanken). ISO 26162-1 legt die allgemeinen Prinzipien für die Gestaltung von Termbanken fest, wie diese in grundlegenden ISO/TC 37-Normen aufgeführt sind, wie etwa in ISO 704, in der neben anderen Themen die allgemeinen Prinzipien für Inhalte und Strukturen von Begriffseinträgen, Identifizierung von Benennungen, Grundprinzipien für die Modellierung von Begriffssystemen und eine Reihe anderer Bereiche im Zusammenhang mit Terminologiarbeit behandelt werden. ISO 26162-1 fördert zudem die Übereinstimmung mit dem terminologischen Metamodell, das in ISO 16642 aufgeführt ist. Das Dokument beschreibt die Rolle, die Datenkategorien bei der Modellierung von terminologischen Daten spielen, und legt Grundprinzipien fest für die Sicherung und Bewertung der Qualität von Daten, die in Termbanken gespeichert werden, wie Granularität, Elementarität und Modellierungsvarianz. Diese Kriterien umfassen wesentliche Bezugsgrößen für die Messung der Qualität und Zuverlässigkeit von terminologischen Daten. ISO 26162-2 setzt die Prinzipien, die in ISO 26162-1 aufgeführt sind, zur Umsetzung von Datenbankgestaltung im Hinblick auf Überlegungen zu Software und Benutzerschnittstellen in Beziehung. Daneben gibt sie Arbeitsabläufe zur pragmatischen Umsetzung in Terminologieverwaltungsumgebungen vor.

Dieses Dokument gibt Empfehlungen für die Festlegung von Verfahren zur Sicherstellung von qualitativ hochwertigen Inhalten in Terminologiedatensammlungen, die dazu entwickelt wurden, Dokumentationsbedürfnisse in unterschiedlichen Bereichen wie beispielsweise Übersetzen, Dolmetschen und Technische Kommunikation zu erfüllen. Die Konformität mit diesem Dokument kann Prozesse stärken, die dazu entwickelt wurden, ein Qualitätsmanagementsystem wie beispielsweise ISO 9001 und die damit zusammenhängenden Auditverfahren in den Bereichen Übersetzen, Dolmetschen oder Technische Kommunikation zu unterstützen. Eine Fehlertypologie wird im Rahmen eines Gesamtbewertungsmodells mit generischen (nicht genormten) Optionen für die Erstellung eines Modells zur Evaluierung von Begriffseinträgen in Abhängigkeit von den Bedürfnissen der Nutzer und Trägerorganisationen vorgelegt.

Anhang A, Anhang B und Anhang C enthalten praxisrelevante Hinweise zu Fehlerbewertungsmethoden. Anhang A beschreibt die Erstellung von Scoring-Modellen, die die in Abschnitt 6 beschriebene Fehlertypologie wiedergeben, wobei Abweichungen bei der Gestaltung in Abhängigkeit von den Bedürfnissen der jeweiligen Organisation möglich sind. Beispielsweise kann ein vorgegebenes Scoring-Modell die Grundlage für einen Bewertungsbogen bilden, der für Studierende und Auszubildende verwendet wird und der sich wahrscheinlich von Bewertungsbögen, wie sie für Großunternehmen oder nationale Termbanken verwendet werden, unterscheidet.

Anhang B enthält ein Beispiel für einen Begriffseintrag. Anhang C zeigt ein Beispiel für ein Bewertungsmodell, das übernommen oder angepasst werden kann, um die Bedürfnisse von Terminologen, Einzelpersonen, die als Freiberufler oder in Unternehmen arbeiten, von Regierungsorganisationen und NGOs, zu erfüllen. Die Werte in diesem Bewertungsmodell können verwendet werden, um ein Scoring-Verfahren zu erstellen, wobei sich versteht, dass die tatsächliche Scoring-Praxis entsprechend Spezifikationen und Anwenderbedürfnissen voraussichtlich abweicht.

1 Anwendungsbereich

Dieses Dokument legt inhaltsbezogene Aspekte der Verwaltung von Terminologiedatenbanken fest. Es gibt Leitlinien zum Inhalt von terminologischen Datensammlungen vor, wobei der Fokus auf der Bewertung der Datenqualität liegt.

Dieses Dokument enthält Leitlinien für Personen, die Begriffseinträge modellieren, um die Interoperabilität und qualitativ hochwertige Inhalte sicherzustellen. Sein Ziel ist es, sicherzustellen, dass terminologische Datensammlungen ihrerseits hohe Standards bei der Gestaltungskonformität mit Normen wie ISO 12620-1 und ISO 16642, bei der Datengenauigkeit und bei der Performanz erfüllen. Es beschreibt Prinzipien zur Sicherstellung der Datenqualität (siehe ISO 9001) und zur Bewertung von terminologischen Datensammlungen mit dem Ziel der kontinuierlichen Verbesserung. Dieser Ansatz unterscheidet sich von dem in ISO 23185:2009, der sich auf die Nutzbarkeit von bestehenden Terminologieressourcen konzentriert.

Dieses Dokument ist nicht anwendbar für die Verwaltung von Textkorpora oder für Terminologieextraktionswerkzeuge.

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 1087, *Terminology work and terminology science — Vocabulary*

ISO 26162-1, *Management of terminology resources — Terminology databases — Part 1: Design*

ISO 26162-2:2019, *Management of terminology resources — Terminology databases — Part 2: Software*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO 1087 und 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 <https://www.iso.org/obp>

— IEC Electropedia: verfügbar unter <https://www.electropedia.org/>

3.1

Begriffseintrag

Eintrag

terminologischer Eintrag

Teil einer *Terminologiedatensammlung* (3.14), der die terminologischen Daten enthält, die sich auf einen Begriff beziehen

Anmerkung 1 zum Begriff: Ein Begriffseintrag kann Informationen enthalten, die sich auf eine oder mehrere Sprachen beziehen.

Anmerkung 2 zum Begriff: In diesem Dokument wird die Benennung *Eintrag* als Kurzform von *Begriffseintrag* verwendet.

[QUELLE: ISO 30042:2019, 3.5, modifiziert — „Eintrag“ als zweite Vorzugsbenennung aufgeführt, Anmerkung 1 zum Begriff und Anmerkung 2 zum Begriff hinzugefügt.]

3.2

Bewertungsmodell

Modell zur Analyse von Daten innerhalb eines *Begriffseintrags* (3.1) nach terminologischen Prinzipien und festgelegten Datenanforderungen, die dem Zweck der *Termbank* (3.11) entsprechen

3.3

Kernstruktur

gemeinsame Struktur und *Datenkategorien* (3.6), die für alle *TBX-Dialekte* (3.5) verwendet werden

Anmerkung 1 zum Begriff: Die Kernstruktur entspricht ISO 16642 (TMF).

[QUELLE: ISO 30042:2019, 3.6]

3.4

Fehlertypologie

systematische Auflistung von Fehlertypen

3.5

TBX-Dialekt

eXtensible-Markup-Language-Dokument (XML-Dokument), das nach der *Kernstruktur* (3.3) von TermBase eXchange (TBX) validiert werden kann und das genau die *Datenkategorien* (3.6) auf den Ebenen erlaubt, die durch eine präzise definierte Konfiguration von Datenkategorien vorgegeben sind

Anmerkung 1 zum Begriff: Siehe ISO 30042 hinsichtlich weiterer Einzelheiten.

[QUELLE: ISO 30042:2019, 3.12, modifiziert — Für dieses Dokument vereinfacht, Anmerkung 1 zum Begriff ausgetauscht.]

3.6

Datenkategorie

Klasse formal oder semantisch ähnlicher Datenelemente

BEISPIEL /Wortklasse/, /Fachgebiet/, /Definition/.

Anmerkung 1 zum Begriff: Eine Datenkategorie kann als Verallgemeinerung der Bedeutung eines Datenfeldes in einer Datenbank betrachtet werden.

Anmerkung 2 zum Begriff: Datenkategorienamen werden im Fließtext, wie in diesem Dokument, zwischen Schrägstrichen angegeben, z. B. /Wortklasse/.

[QUELLE: ISO 30042:2019, 3.8]

3.7

Datenintegrität

Übereinstimmung von Datenwerten mit einem vorgegebenen Regelwerk

[QUELLE: ISO/IEC TR 10032:2003, 2.23]

3.8

lexikalische Einheit

Spracheinheit, die dem Wortschatz einer bestimmten Sprache angehört

[QUELLE: ISO 1951:2007, 3.8, modifiziert — Verweisung auf ein Wörterbuch entfernt.]

3.9**Qualitätssicherung****QS**

Menge von geplanten und systematischen Aktivitäten, die notwendig sind, um sicherzustellen, dass ein *Begriffseintrag* (3.1) die Akzeptanzkriterien auf Grundlage von terminologischen Prinzipien und festgelegten Datenanforderungen erfüllt

3.10**Spezifikation**

Dokument, das die Anforderungen, die von einer *Terminologiedatensammlung* (3.14) zu erfüllen sind, ausführlich darlegt

Anmerkung 1 zum Begriff: Spezifikationen können Verfahren für das Prüfen der Übereinstimmung mit diesen Anforderungen einschließen.

3.11**Termbank****Terminologiedatenbank**

Datenbank, die eine *Terminologiedatensammlung* (3.14) enthält

[QUELLE: ISO 30042:2019, 3.28, modifiziert — „Terminologiedatenbank“ ist nicht mehr nur zugelassene Benennung, sondern zweite Vorzugsbenennung.]

3.12**Termbank-Qualitätsbewerter**

Person, die als Terminologe oder Experte des jeweiligen Fachgebiets qualifiziert ist und eine Qualitätsbewertung einer *Terminologiedatensammlung* (3.14) vornimmt

3.13**Termhood, f.**

Fachwortlichkeit

Grad der terminologischen Relevanz einer *lexikalischen Einheit* (3.8)

BEISPIEL „Massengutfrachter“ weist eine stärkere Termhood als „Schiff“ alleine auf. „Maus“ verfügt im Bereich Computeranwendungen über eine gewisse Termhood, jedoch nicht in der Allgemeinsprache.

Anmerkung 1 zum Begriff: Termhood ist sowohl auf Einwortbenennungen (die aus einem einzelnen Wort bestehen) und Mehrwortbenennungen (die aus mehr als einem Wort oder einer lexikalischen Einheit gebildet sind) als auch auf weitere Bezeichnungen wie Eigennamen, Titel, Marken- und Gattungsbezeichnungen sowie Formeln und Symbole anwendbar.

3.14**Terminologiedatensammlung**

Ressource, die aus *Begriffseinträgen* (3.1) mit zugehörigen Metadaten und dokumentarischen Informationen besteht

BEISPIEL Eine TBX-Dokumentinstanz, ISO 1087.

[QUELLE: ISO 30042:2019, 3.29, modifiziert — Zweite Vorzugsbenennung „TDC“ entfernt.]

3.15**Unithood, f.**

Kollokationsstärke

Grad, in dem eine bestimmte Abfolge von Wörtern über ausreichende inhaltliche Kombinierbarkeit verfügt, um eine stabile *lexikalische Einheit* (3.8) zu bilden

BEISPIEL „absolute Mehrheit“ weist eine höhere Unithood als „stille Mehrheit“ auf.