

DIN 5007-1

ICS 01.140.20

Ersatz für
DIN 5007:1991-04**Ordnen von Schriftzeichenfolgen –
Teil 1: Allgemeine Regeln für die Aufbereitung (ABC-Regeln)**Filing of character strings –
Part 1: General rules for processing (ABC rules)Classement de chaînes de caractère –
Partie 1: Méthodes générales de pré-traitement (Règles ABC)

Gesamtumfang 18 Seiten

Normenausschuss Bürowesen (NBü) im DIN
Normenausschuss Bibliotheks- und Dokumentationswesen (NABD) im DIN
Normenausschuss Informationstechnik (NI) im DIN
Normenausschuss Terminologie (NAT) im DIN

Inhalt

	Seite
Vorwort	3
Einleitung.....	4
1 Anwendungsbereich	4
2 Normative Verweisungen.....	5
3 Begriffe	5
4 Allgemeine Regeln.....	6
5 Schriftzeichen	6
5.1 Schriftgrundzeichen	6
5.1.1 Buchstaben	6
5.1.2 Ziffern und Zahlzeichen	7
5.2 Schriftsonderzeichen	7
5.2.1 Leerzeichen	7
5.2.2 Andere Schriftsonderzeichen.....	7
6 Ordnungswert der Schriftzeichen und Regeln für die Aufbereitung.....	8
6.1 Schriftgrundzeichen	8
6.1.1 Buchstaben	8
6.1.2 Ziffern bzw. Zahlzeichen und Zahlen	10
6.2 Schriftsonderzeichen	11
6.2.1 Leerzeichen	11
6.2.2 Andere Schriftsonderzeichen.....	11
7 Einordnung von gleichgesetzten Schriftzeichenfolgen.....	11
7.1 Einordnung von nicht beeinflussten und durch Gleichsetzung beeinflussten Schriftzeichenfolgen	11
7.2 Einordnung von durch Gleichsetzung beeinflussten Schriftzeichenfolgen untereinander.....	12
7.2.1 Ligaturen.....	12
7.2.2 Umlaute.....	13
7.2.3 Sonstige Buchstaben mit diakritischen Zeichen.....	13
7.3 Einordnung von gleichgesetzten Großbuchstaben	13
Anhang A (normativ) Zusammenfassung der Regeln aus den Abschnitten 6 und 7.....	14
Anhang B (normativ) Profil in der in ISO/IEC 14651 spezifizierten Syntax	15
B.1 Erläuterungen.....	15
B.1.1 Syntax	15
B.1.2 Codierung der Schriftzeichen.....	15
B.1.3 Schlüsselwörter	15
B.1.4 Verarbeitungsvorschrift	15
B.1.5 Kommentare.....	15
B.2 ISO/IEC 14651-konformes Profil für DIN 5007-1	16
Literaturhinweise	18

Vorwort

Diese Norm wurde vom Gemeinschaftsausschuss GA „DIN 5007“ unter Federführung des Normenausschusses Bürowesen (NBü) erarbeitet.

DIN 5007 *Ordnen von Schriftzeichenfolgen* besteht aus:

- *Teil 1: Allgemeine Regeln für die Aufbereitung (ABC-Regeln)*
- *Teil 2: Ansetzungsregeln für die alphabetische Ordnung von Namen*

Änderungen

Gegenüber DIN 5007:1991-04 wurden folgende Änderungen vorgenommen:

- a) Angleichung an DIN 5007-2;
- b) Überarbeitung von 5.1.1.3 der Ausgabe von 1991, siehe 6.1.1.4 dieser Ausgabe;
- c) vollständige redaktionelle Überarbeitung auch im Hinblick auf die modalen Hilfsverben;
- d) Ergänzung eines normativen Anhangs B Profil in der in ISO/IEC 14651 spezifizierten Syntax.

Frühere Ausgaben

DIN 5007: 1948-12, 1962-11, 1991-04

Einleitung

Als Voraussetzung für das Ordnen von Schriftzeichenfolgen ist es nötig, eine Art Richtschnur oder Messlatte festzulegen, die zum einen die Grobeinteilung von Schriftzeichen in bestimmte Gruppen bestimmt (siehe Aufzählung a)) und zum anderen eine Feineinteilung der Schriftzeichen innerhalb dieser Gruppen ermöglicht (siehe Aufzählung b)).

a) Ordnung nach Gruppen

Siehe Anhang A.

b) Ordnung innerhalb der Gruppen

Für manche der in dieser Norm angeführten Gruppen legt diese Norm keine Ordnung der Schriftzeichen untereinander innerhalb der Gruppe fest. Ein Anwender kann für die jeweiligen Schriftzeichen einer solchen Gruppe eine Ordnungsreihenfolge selbst bestimmen. Das wäre z. B. für die Gruppe nach A.2 möglich, ebenso für die Gruppe nach A.4, dort allerdings nur mit Ausnahme der in DIN 31638 angegebenen Schriftzeichen.

Während diese Norm die Ordnung von Buchstaben aus nichtlateinischen Alphabeten nach den Regeln des entsprechenden Ursprungsalphabets festlegt, ist die Ordnungsreihenfolge von mehreren nicht-lateinischen Alphabeten untereinander nicht vorgegeben; sie kann nach den Wünschen des Anwenders erfolgen.

Die Ordnung von Ziffern bzw. Zahlenzeichen und Zahlen bedarf keiner Erläuterung, da der Ordnungswert sich aus dem jeweiligen Zahlenwert, und zwar in aufsteigender Reihenfolge ergibt.

Die Gruppe nach A.3, nämlich die Gruppe der Grundbuchstaben, sieht vor, dass es nur die Reihenfolge der Buchstaben von a bis z, also 26 Buchstaben, gibt. Es ist wichtig, zu beachten, dass es hier keine Ordnungswerte für andere lateinische Buchstaben wie Ligaturen oder mit diakritischem Zeichen versehene Buchstaben gibt. Ebenso sind keine getrennten Ordnungswerte für Groß- und Kleinbuchstaben vorgesehen.

Das bedeutet aber, dass alle nicht in der Reihe von a bis z enthaltenen Buchstaben einem oder mehreren Grundbuchstaben gleichgesetzt werden müssen, damit sie einerseits überhaupt und andererseits auch richtig geordnet werden können.

Die Gleichsetzung erfolgt im Rahmen der Aufbereitung zum Ordnen der zu ordnenden Schriftzeichenfolgen. Die Aufbereitung (und damit die Gleichsetzung) wird in dieser Norm festgelegt.

Anhang B ermöglicht die automatisierte Sortierung durch ein Profil in der in ISO/IEC 14651 spezifizierten Syntax.

1 Anwendungsbereich

Diese Norm legt allgemeine Regeln für das Ordnen von Schriftzeichen und Schriftzeichenfolgen (Wörtern, Wortgruppen und sonstigen Zeichen) für alle Bereiche fest, in denen alphabetisch geordnet werden soll, z. B. in Lexika, Registraturen, Katalogen oder Namensverzeichnissen. Für spezielle Anwendungen können in anderen Normen darüber hinausgehende Festlegungen getroffen werden.

Festlegungen für die Ansetzung (Bilden von Ordnungswörtern, Ordnungsgruppen usw.) sind nicht Gegenstand dieser Norm. Zu Ansetzungsregeln für das Ordnen von Namen siehe DIN 5007-2.

2 Normative Verweisungen

Die folgenden zitierten Dokumente sind für die Anwendung dieses Dokuments erforderlich. 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).

DIN 31638:1994-08, *Bibliographische Ordnungsregeln*

ISO/IEC 14651, *Information technology – International string ordering and comparison – Method for comparing character strings and description of the common template tailorable ordering*

ENV 13710, *European Ordering Rules – Ordering of characters from the Latin, Greek and Cyrillic scripts*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

3.1

Aufbereitung

Gleichsetzung von Schriftzeichen, für die diese Norm keinen eigenen Ordnungswert vorsieht, mit anderen Schriftzeichen, für die in dieser Norm ein Ordnungswert festgelegt ist

3.2

Schriftzeichen

Schriftgrundzeichen und Schriftsonderzeichen, die zur Textdarstellung dienen

3.3

Schriftgrundzeichen

Buchstaben und Ziffern bzw. Zahlzeichen

3.4

Buchstaben

Grundbuchstaben sowie Buchstaben mit diakritischen Zeichen einschließlich Umlauten und das lateinische Alphabet ergänzender Buchstaben

3.5

Ziffern und Zahlzeichen

Zeichen, die dazu dienen, einen Zahlenwert darzustellen

3.6

Schriftsonderzeichen

Schriftzeichen, die neben Schriftgrundzeichen zur Textdarstellung dienen

BEISPIELE Satzzeichen, Zeichen für Wörter, Einheitenzeichen (siehe DIN 1301-1), Zahlensymbole, Rechenzeichen und das Leerzeichen.

3.7

Leerzeichen

Schriftsonderzeichen zur Unterteilung von Schriftzeichenfolgen, im Allgemeinen durch die Abwesenheit eines graphischen Symbols dargestellt

3.8

diakritische Zeichen

Zusatzzeichen zu Buchstaben

4 Allgemeine Regeln

Der Ordnungsvorgang muss durch eine Aufbereitung (siehe 3.1) der zu ordnenden Schriftzeichenfolgen vorbereitet werden.

ANMERKUNG 1 Die Aufbereitung geht dem Ordnungsvorgang voraus und kann unter anderem vom jeweiligen Anwendungszweck abhängig sein. Der Aufbereitungsvorgang kann zeitlich oder räumlich getrennt vom Ordnungsvorgang ablaufen, aber auch mit dem Ordnungsvorgang so verbunden sein, dass für den Benutzer der Eindruck getrennt ablaufender Vorgänge nicht entsteht. Das Ergebnis der Aufbereitung, die Aufbereitungsform, kann schriftlich niedergelegt oder gedanklich beim Ordnen berücksichtigt werden. Die aufbereitete Form einer Schriftzeichenfolge kann mit der originären Schriftzeichenfolge identisch sein, kann aber auch von dieser abweichen. Die durch die Gleichsetzung entstehende gezielte Veränderung einer Schriftzeichenfolge dient ausschließlich zum Zwecke des Ordnen. Die originäre Form der zu ordnenden Schriftzeichenfolgen wird dadurch nicht ersetzt, sondern bleibt unverändert bestehen.

Buchstaben, die weder Grundbuchstaben nach 5.1.1.1 sind noch einer anderen Gruppe in der Hierarchie nach Anhang A zugeordnet werden können, müssen durch Aufbereitung Grundbuchstaben gleichgesetzt werden, damit sie geordnet werden können.

Die Aufbereitung ist nötig, weil diese Norm nur für einen Teil der Schriftzeichen (z. B. Grundbuchstaben) eigene Ordnungswerte festlegt, für andere (z. B. Buchstaben mit diakritischen Zeichen) jedoch nicht. Diese anderen Schriftzeichen müssen zum Zwecke des Ordnen Schriftzeichen mit Ordnungswert gleichgesetzt werden.

Nur die durch die Gleichsetzung im Rahmen der Aufbereitung entstandene Aufbereitungsform bestimmt den Ordnungswert der betroffenen Schriftzeichen oder Schriftzeichenfolgen und nicht die originäre Darstellung der Schriftzeichen oder Schriftzeichenfolgen.

Die Ordnungsregeln dieser Norm beziehen sich immer auf die aufbereitete Form der zu ordnenden Schriftzeichenfolgen, d. h., maßgebend ist der Ordnungswert der Schriftzeichen in der Aufbereitungsform, nicht der in der originären Form, weil in dieser manche Schriftzeichen keinen Ordnungswert haben.

Für Schriftzeichen, für die diese Norm nicht ausdrücklich eine andere Ordnung festlegt, ist die in ENV 13710 vorgeschriebene Anordnung zu verwenden. Die Ordnung von Buchstaben aus nichtlateinischen Alphabeten ergibt sich aus den in ENV 13710 getroffenen Regelungen.

Bei der Anordnung von Ziffern und Zahlenzeichen sollte man sich, wo immer sinnvoll, an ihrem Zahlwert orientieren.

5 Schriftzeichen

5.1 Schriftgrundzeichen

5.1.1 Buchstaben

5.1.1.1 Grundbuchstaben des lateinischen Alphabets

Kleinbuchstaben:

a b c d e f g h i j k l m n o p q r s t u v w x y z

Großbuchstaben:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z