

DIN ISO/IEC 6937

ICS 35.040

Ersatz für
DIN ISO/IEC 6937:1996-11**Informationstechnik –
Zeichensätze für die Textkommunikation –
Lateinisches Alphabet (ISO/IEC 6937:2001)**

Information technology –
Coded graphic character set for text communication –
Latin alphabet (ISO/IEC 6937:2001)

Technologies de l'information –
Jeu de caractères graphiques codés pour la transmission de texte –
Alphabet latin (ISO/CEI 6937:2001)

Gesamtumfang 47 Seiten

Normenausschuss Informationstechnik und Anwendungen (NIA) im DIN

Inhalt

	Seite
Nationales Vorwort	3
Nationaler Anhang NA (informativ) Literaturhinweise.....	3
Einleitung.....	5
1 Anwendungsbereich	6
2 Normkonformität und Implementierung.....	6
2.1 Normkonformität.....	6
2.1.1 Normkonformität des Informationsaustauschs.....	6
2.1.2 Gerätekonformität.....	6
2.2 Implementierung	7
3 Normative Verweisungen.....	7
4 Begriffe	7
5 Kennzeichnung, Code-Tabelle und Namen.....	10
5.1 Kennzeichnung	10
5.2 Code-Tabelle	11
5.3 Namen	11
6 Festlegung von LEERZEICHEN, NICHT-TRENNBARES LEERZEICHEN und SILBENTRENNUNGSSTRICH	11
6.1 LEERZEICHEN (SP)	11
6.2 NICHT-TRENNBARES LEERZEICHEN (NBSP).....	11
6.3 SILBENTRENNUNGSSTRICH (SHY).....	11
7 Zusammensetzung des Zeichenvorrats	12
8 Festlegung des codierten Zeichensatzes	12
8.1 Zeichensätze	12
8.2 Erklärung der Code-Tabelle.....	13
8.3 Codierte Darstellung der Schriftzeichen des Zeichenvorrats.....	13
9 Untermengen des Zeichenvorrats.....	14
10 Identifizierung der Wahlmöglichkeiten.....	15
10.1 Zweck und Kontext der Identifizierung.....	15
10.2 Identifizierung der Codierungsmethode.....	15
10.3 Identifizierung von Grund- und Ergänzungszeichensätze	15
10.4 Identifikation von Untermengen.....	16
Anhang A (normativ) 7-Bit-Code	27
Anhang B (informativ) Methode zur Bildung von Kurzbezeichnungen für diese Norm.....	30
Anhang C (informativ) Benutzung von diakritischen Zeichen ohne Zeichenschritt	40
Anhang D (informativ) Benutzung der lateinischen Buchstaben in verschiedenen Sprachen.....	41
Anhang E (informativ) Alternative codierte Darstellung des Zeichenvorrats ohne diakritische Zeichen ohne Zeichenschritt	45
Anhang F (informativ) Zusammenfassung der Änderungen zwischen der zweiten (1994) und der hier vorliegenden dritten Ausgabe dieser Norm.....	46
Literaturhinweise	47

Nationales Vorwort

Dieses Dokument (ISO/IEC 6937:2001) wurde vom Technischen Komitee ISO/IEC JTC 1 (Information technology, Subcommittee SC 2, Coded character sets) erarbeitet.

Das zuständige deutsche Gremium ist der Arbeitskreis „Codierte Zeichensätze“ im Normenausschuss „Informationstechnik und Anwendungen (NIA)“.

Diese dritte Ausgabe der ISO/IEC 6937 ersetzt die zweite Ausgabe von 1994, die technisch überarbeitet wurde.

Anhang A ist ein normativer Teil dieser Norm, die Anhänge B, C, D, E und F sind nur zur Information.

Für die in diesem Dokument zitierten Internationale Normen wird im Folgenden auf die entsprechenden Deutschen Normen hingewiesen:

ISO/IEC 2022 siehe DIN ISO/IEC 2022

Die im Dokument zitierte Internationale Norm ISO/IEC 10646-1 wurde inzwischen durch die Internationale Norm ISO/IEC 10646 ersetzt.

Änderungen

Gegenüber DIN ISO/IEC 6937:1996-11 wurden folgende Änderungen vorgenommen:

- a) neue ISO/IEC-Ausgabe übernommen und ins Deutsche übersetzt;
- b) die Namen KLEINBUCHSTABE Æ LIGATUR und GROSSBUCHSTABE Æ LIGATUR durch KLEINBUCHSTABE Æ und GROSSBUCHSTABE Æ ersetzt und damit an ISO/IEC 10646-1 angepasst;
- c) die Namen NOTENZEICHEN in ACHTEL NOTENZEICHEN und TRADEMARKSIGN in TRADE MARK SIGN geändert und damit an ISO/IEC 10646-1 angepasst;
- d) eine Anzahl von Kurzbezeichnungen geändert, wie in Anhang B, Anmerkung 2 beschrieben;
- e) eine Anzahl von kleinen Korrekturen und Klarstellungen durchgeführt.

Frühere Ausgaben

DIN ISO/IEC 6937: 1996-11

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN 2137-2, *Büro und Datentechnik — Tastaturen — Teil 2: Deutsche Tastatur für Daten- und Textverarbeitung, Tastenanordnung und Belegung mit Schriftzeichen*

DIN 66303, *Informationstechnik — 8-Bit-Code*

DIN ISO/IEC 2022, *Informationstechnik — Zeichencodestruktur und Erweiterungstechniken*

Informationstechnik — Zeichensätze für die Textkommunikation — Lateinisches Alphabet

Einleitung

Diese Norm legt einen Zeichenvorrat und dessen codierte Darstellung für den Gebrauch in der Textübertragung fest.

Im Allgemeinen besteht Text aus Zeichen und Bildern (siehe 4.16), diese Norm bezieht sich aber nur auf Texte, die ausschließlich aus Zeichen bestehen.

Diese Norm basiert auf 8-Bit-Codierung, wobei der normative Anhang A den 7-Bit-Code für den Zeichensatz dieser Norm festlegt.

Die weiteren Anhänge beinhalten:

- a) die Beschreibung einer Methode um Kurzbezeichnungen für alle Zeichen dieser Norm festzulegen (Anhang B);
- b) eine Zusammenfassung von diakritischen Zeichen in Verbindung mit den Basiszeichen des lateinischen Alphabets (Anhang C);
- c) eine Zusammenfassung über die Benutzung der Zeichen des lateinischen Alphabets in verschiedenen Sprachen (Anhang D);
- d) eine alternative codierte Darstellung des Zeichenvorrats ohne Verwendung diakritischer Zeichen ohne Zeichenschritt (Anhang E);
- e) eine Zusammenfassung der Änderungen zwischen der zweiten (1994) und der hier vorliegenden dritten Ausgabe von ISO/IEC 6937 (Anhang F);
- f) ein Literaturverzeichnis.

1 Anwendungsbereich

Diese Norm legt Folgendes fest:

- a) die codierte Darstellung von Zeichen;
- b) einen Zeichenvorrat von lateinischen Schrift- und Sonderzeichen für die Textkommunikation in vielen europäischen Sprachen, die das lateinische Alphabet verwenden;
- c) Regeln für die Definition und den Gebrauch von Untermengen von Zeichenvorräten, z. B. Untermengen des in dieser Norm festgelegten Zeichenvorrats.

2 Normkonformität und Implementierung

2.1 Normkonformität

2.1.1 Normkonformität des Informationsaustauschs

Ein Codiertes-Zeichen-Datenelement (CC-Datenelement) innerhalb einer codierten Information für den Datenaustausch ist in Übereinstimmung mit dieser Norm, wenn alle codierten Darstellungen von Zeichen innerhalb dieses CC-Datenelementes den vorgeschriebenen Regeln dieser Norm entsprechen.

Eine Aussage über die Normkonformität muss enthalten:

- die Untermenge des Zeichenvorrats in Übereinstimmung mit Abschnitt 9, wenn eine verwendet wird;
- die 7-Bit-Codierung in Übereinstimmung mit Anhang A, wenn sie verwendet wird.

2.1.2 Gerätekonformität

Ein Gerät ist in Übereinstimmung mit dieser Norm, wenn es den Anforderungen nach 2.1.2.1 und entweder 2.1.2.2 und/oder 2.1.2.3 entspricht.

2.1.2.1 Gerätebeschreibung

Ein Gerät, das mit dieser Norm übereinstimmt, muss Gegenstand einer Beschreibung sein, welche die Mittel beschreibt, mit denen der Benutzer die Zeichen dem Gerät übergeben kann oder diese erkennen kann, wenn sie ihm verfügbar gemacht werden, wie es in 2.1.2.2 und 2.1.2.3 festgelegt ist.

2.1.2.2 Abgebende Geräte

Ein abgebendes Gerät muss es seinem Benutzer gestatten, jede beliebige Folge von Zeichen gemäß der gewählten Festlegung zu übergeben, und es muss in der Lage sein, die codierten Darstellungen der Zeichen innerhalb eines CC-Datenelementes zu senden.

2.1.2.3 Annehmende Geräte

Ein annehmendes Gerät muss in der Lage sein, alle codierten Darstellungen von Zeichen, die in einem CC-Datenelement enthalten sind und 2.1.1 entsprechen, zu empfangen und zu interpretieren. Das Gerät muss die entsprechenden Zeichen dem Benutzer in einer Weise verfügbar machen, dass dieser sie unter denen erkennen kann, die der gewählten Festlegung entsprechen, und dass er sie voneinander unterscheiden kann.

2.2 Implementierung

Für die Verwendung des in dieser Norm festgelegten Zeichensatzes müssen Implementierungen für verschiedene Datenträger festgelegt werden. Diese Datenträger können z. B. magnetische oder optische austauschbare Datenträger oder Übertragungskanäle sein. Dieser Datenaustausch geschieht entweder indirekt durch den Austausch von körperlichen Datenträgern oder durch örtliche Verbindungen von verschiedenen Einheiten (wie Ein-/Ausgabegeräte oder Rechenanlagen) oder durch Datenübermittlungseinheiten.

Die Implementierung der Zeichensätze auf körperlichen Datenträgern und für die Datenübermittlung unter Berücksichtigung der benötigten Fehlerbehandlungen wird in anderen internationalen und nationalen Normen festgelegt.

3 Normative Verweisungen

Diese Norm enthält durch datierte oder undatierte Verweisungen Festlegungen aus anderen Publikationen. Diese normativen Verweisungen sind an den jeweiligen Stellen im Text zitiert, und die Publikationen sind nachstehend aufgeführt. Bei datierten Verweisungen gehören spätere Änderungen oder Überarbeitungen nur zu dieser Norm, falls sie durch Änderung und Überarbeitung eingearbeitet sind. Bei undatierten Verweisungen gilt die letzte Ausgabe der in Bezug genommenen Publikation (einschließlich Änderungen).

ISO/IEC 2022:1994, *Information technology — Character code structure and extension*

ISO 2375:1985, *Data processing — Procedure for registration of escape sequence*

ISO/IEC 7350:1991, *Information technology — Registration of repertoires of graphic Characters from ISO/IEC 10367*

ISO/IEC 10367:1991, *Information technology — Standardized coded graphic sets for use in 8-bit-codes*

ISO/IEC 10538:1991, *Information technology — Control functions for text communication*

ISO/IEC 10646-1:2000, *Information technology — Universal Multiple-Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane*

4 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

4.1

aktuelle Zeichenstelle

(en: active position)

Zeichenstelle, die für das nächste Zeichen bereitsteht oder an der die nächste Steuerfunktion ausgeführt wird; auf Bildschirmgeräten meist durch eine Positionsmarke gekennzeichnet

4.2

Bitkombination

(en: bit combination)

geordnete Menge von Bits, die für die codierte Darstellung benutzt wird

4.3

Zeichen

(en: character)

Element aus einer festgelegten Menge von Elementen, das für die Organisation, die Steuerung oder die Darstellung von Daten verwendet wird

4.4

Zeichenstelle

(en: character position)

Stelle in einer Anzeige, an der ein Zeichen für das Auge sichtbar gemacht werden kann

4.5

CC-Datenelement

Codiertes-Zeichen-Datenelement

(en: coded character data element)

Element der ausgetauschten Information, das aus Folgen codierter Darstellungen von Zeichen besteht, entsprechend einer oder mehrerer Normen für codierte Zeichensätze

ANMERKUNG 1 In einer Übertragungsumgebung, die mit dem Referenzmodell für Offene-Systeme-Verbindung nach ISO 7498 übereinstimmt, stellt ein CC-Datenelement die gesamte oder einen Teil der Information dar, die mit der Darstellungs-Protokoll-Dateneinheit (PPDU), wie sie in der Internationalen Norm festgelegt ist, verkehrt.

ANMERKUNG 2 Wenn der Informationsaustausch mit austauschbaren Datenträgern erfolgt, stellt ein CC-Datenelement die gesamten oder einen Teil der Benutzerdaten dar und nicht diejenigen, die während der Formatierung und Initialisierung entstehen.

4.6

Codierter Zeichensatz

Code

(en: coded character set; code)

Satz von eindeutigen Regeln, die einen Zeichensatz und die Eins-zu-Eins-Beziehung zwischen den Zeichen des Vorrats und deren Bit-Kombinationen festsetzen

4.7

Code-Erweiterung

(en: code extension)

Verfahren für die Festlegung einer Codierung für Zeichen, die in einem Zeichensatz eines vorgegebenen Codes nicht enthalten sind

4.8

Code-Tabelle

(en: code table)

Tabelle, die die Anordnung von Bitkombinationen, denen Zeichen zugeordnet sind, zeigt

4.9

Steuerzeichen

(en: control character)

Steuerfunktion, deren codierte Darstellung aus einer Bitkombination besteht

4.10**Steuerfunktion**

(en: control function)

Funktion, die sich auf die Aufzeichnung, die Verarbeitung, die Übermittlung und Übertragung oder die Interpretation von Daten auswirkt, und deren codierte Darstellung aus einer oder mehreren Bitkombinationen besteht

4.11**Gerät**

(en: device)

Teil einer Informationsverarbeitungseinrichtung, das codierte Informationen innerhalb von CC-Datenelementen senden und/oder empfangen kann

ANMERKUNG Es kann ein Ein-/Ausgabegerät im herkömmlichen Sinne sein oder ein Prozess wie ein Anwendungsprogramm oder eine Übergangsfunktion.

4.12**Escape-Folge**

(en: escape sequence)

Folge von Bitkombinationen, die für Steuerzwecke in Code-Erweiterungsverfahren benutzt werden. Die erste dieser Bitkombinationen stellt die Steuerfunktion ESCAPE (ESC) dar

4.13**Schriftzeichen**

(en: graphic character)

Zeichen, das keine Steuerfunktion ist und normalerweise in handschriftlicher oder gedruckter Form oder als Anzeige auf einem Bildschirm dargestellt wird und dessen codierte Darstellung aus einer oder mehreren Bitkombinationen besteht

4.14**Bildzeichen**

(en: graphic symbol)

sichtbare Darstellung eines Schriftzeichens oder einer Steuerfunktion

4.15**Zeichenvorrat**

(en: repertoire)

festgelegte Menge von Zeichen, die mit einer oder mehreren Bitkombinationen eines festgelegten Zeichensatzes dargestellt werden können

4.16**Text**

(en: text)

Darstellung von Informationen für die menschliche Wahrnehmung, die für die Wiedergabe in zweidimensionaler Form vorgesehen ist, zum Beispiel auf Papier gedruckt oder auf einem Bildschirm angezeigt

Text besteht aus graphischen Symbolen, Worten oder Sätzen in natürlichen oder künstlichen Sprachen, aus Bildern, Diagrammen und Tabellen.

ANMERKUNG Diese Norm bezieht sich nur auf Text, der aus Schriftzeichen besteht.

4.17**Textkommunikation**

(en: text communication; communication of text)

Übermittlung von Text mit Mitteln der Telekommunikation

ANMERKUNG Im Sinne dieser Norm bedeutet Textkommunikation die binärcodierte Übertragung von Zeichen.

4.18**Benutzer**

(en: user)

Person oder eine andere Instanz, welche die von einem Gerät bereitgestellten Dienste abruft

ANMERKUNG 1 Diese Instanz kann zum Beispiel ein Prozess wie ein Anwendungsprogramm sein, wenn das „Gerät“ ein Codeumsetzer ist, oder eine Übertragungsfunktion.

ANMERKUNG 2 Die Zeichen, die von einem Benutzer übergeben oder ihm zugänglich gemacht werden, können in geräteeigener Form oder in nicht-herkömmlicher sichtbarer Darstellung vorliegen, vorausgesetzt, die Bedingungen nach 2.1.2 werden erfüllt.

5 Kennzeichnung, Code-Tabelle und Namen**5.1 Kennzeichnung**

Die Bits der Bitkombinationen des 8-Bit-Codes werden mit b8, b7, b6, b5, b4, b3, b2, und b1 gekennzeichnet, wobei b8 das Bit mit der höchsten Wertigkeit oder linksbündige Bit und b1 das Bit mit der niedrigsten Wertigkeit oder rechtsbündige Bit ist.

Die Bitkombinationen können als Binärzahl im Bereich von 0 bis 255 interpretiert werden mit der folgenden Wertigkeit für die einzelnen Bits:

Bit:	b8	b7	b6	b5	b4	b3	b2	b1
Wertigkeit:	128	64	32	16	8	4	2	1

In dieser Norm werden die Bitkombinationen mit xx/yy bezeichnet, wobei xx und yy einen Bereich von 00 bis 15 haben.

Die Zuordnung der Bezeichnung in der Form xx/yy und der Bitkombination mit der Bitbezeichnung b8 bis b1 ist wie folgt:

- xx ist die Zahl, dargestellt durch die Bits b8, b7, b6 und b5, denen die Wertigkeiten 8, 4, 2 und 1 zugeordnet sind.
- yy ist die Zahl, dargestellt durch die Bits b4, b3, b2 und b1, denen die Wertigkeiten 8, 4, 2 und 1 zugeordnet sind.

Die Kennzeichnung in der Form xx/yy ist die gleiche, die verwendet wird, um einen Platz innerhalb von Code-Tabellen zu bezeichnen, wobei xx die Spaltennummer und yy die Zeilennummer ist (siehe 5.2).

5.2 Code-Tabelle

Eine 8-Bit-Code-Tabelle enthält 256 Plätze, aufgeteilt in 16 Spalten und 16 Zeilen. Die Spalten und Zeilen werden von 00 bis 15 benummert.

Die Plätze in der Code-Tabelle (Code-Plätze) werden mit der Bezeichnung in der Form xx/yy gekennzeichnet, wobei xx die Spaltennummer und yy die Zeilennummer angibt. Die Spaltennummern werden oberhalb der Tabelle, die Zeilennummern werden links der Tabelle gezeigt.

Die Plätze in der Code-Tabelle haben eine Eins-zu-Eins-Zuordnung zu den entsprechenden Bitkombinationen des Codes. Die Kennzeichnung eines Code-Platzes in der Form xx/yy ist die gleiche wie die der zugeordneten Bitkombination.

5.3 Namen

Diese Norm ordnet jedem Zeichen einen eindeutigen Namen zu. Zusätzlich werden für die Zeichen LEERZEICHEN, NICHT-TRENNBARES LEERZEICHEN und SILBENTRENNUNGSSTRICH je eine Kurzbezeichnung festgelegt. Für Kurzbezeichnungen werden nur lateinische Buchstaben von A bis Z verwendet. Die Kurzbezeichnungen sollen in allen Übersetzungen des entsprechenden Internationalen Standards erhalten bleiben.

Diese Norm legt zudem für jedes Schriftzeichen ein Bildzeichen fest. Dieses Bildzeichen wird auf dem zugeordneten Platz der Code-Tabelle gezeigt. Diese Norm beschränkt aber weder die Verwendung eines Zeichens, noch legt sie für die Darstellung einen bestimmten Stil oder eine bestimmte Schriftart fest.

Die Namen der Zeichen wurden an ISO/IEC 10646-1 angepasst.

6 Festlegung von LEERZEICHEN, NICHT-TRENNBARES LEERZEICHEN und SILBENTRENNUNGSSTRICH

6.1 LEERZEICHEN (SP)

(SPACE)

Ein Schriftzeichen, die sichtbare Darstellung geschieht durch die Abwesenheit eines Bildzeichens. Die codierte Darstellung ist 02/00.

6.2 NICHT-TRENNBARES LEERZEICHEN (NBSP)

(NO-BREAK SPACE)

Ein Schriftzeichen, die sichtbare Darstellung geschieht durch die Abwesenheit eines Schriftzeichens. Es wird verwendet, wenn an der Stelle seines Auftretens ein Zeilenumbruch verhindert werden soll.

6.3 SILBENTRENNUNGSSTRICH (SHY)

(SOFT HYPHEN)

Ein Schriftzeichen, das gleich oder ähnlich dem Mittestrich dargestellt wird. Es wird verwendet, wenn ein Zeilenumbruch eine Silbentrennung in einem Wort erfordert.

7 Zusammensetzung des Zeichenvorrats

Der in dieser Norm festgelegte Zeichenvorrat enthält

a) LEERZEICHEN (SP)

und 332 folgende Zeichen

g) Lateinische alphabetische Zeichen umfassen

- a. die 52 großen und kleinen Schriftzeichen des lateinischen Basialphabets,
- b. zusammengesetzte Schriftzeichen, die graphische Darstellung der Zeichen, die aus einem Zeichen des lateinischen Basialphabets und einem diakritischen Zeichen zusammengesetzt sind,
- c. spezielle lateinische Zeichen, die weder Grundzeichen, noch zusammengesetzte Zeichen sind.

h) nicht alphabetische Zeichen, wie Zahlen, Bruchzahlen, Interpunktionszeichen und diakritische Zeichen, Währungssymbole usw.

Der Zeichenvorrat, ausgenommen LEERZEICHEN, ist in Tabelle 4 festgelegt. In jedem Tabelleneintrag legt die erste Spalte den Namen des Zeichens fest. Die zweite Spalte legt die codierte Darstellung fest (siehe 8.3).

ANMERKUNG 1 Einen Überblick über die Verwendung der lateinischen Zeichen in den verschiedenen Sprachen enthält Anhang D.

ANMERKUNG 2 Die Benutzung der folgenden Zeichen: GROSSBUCHSTABE L MIT MITTEPUNKT, KLEINBUCHSTABE L MIT MITTEPUNKT und KLEINBUCHSTABE N MIT VORANGEGANGENEM APOSTROPH ist verboten und sollten besser als „L“ / „l“ mit nachfolgendem MITTEPUNKT und APOSTROPH gefolgt von „n“ codiert werden.

8 Festlegung des codierten Zeichensatzes

8.1 Zeichensätze

Die codierte Darstellung des Zeichenvorrates, der in dieser Norm festgelegt ist, erlaubt die Benutzung von LEERZEICHEN und zwei Zeichensätzen, dem „Grundzeichensatz“ und dem „Ergänzungszeichensatz“.

Der Grundzeichensatz muss aus den Schriftzeichen des Basis G0-Zeichensatzes bestehen, der durch die internationale Registrierung Nr. 6 festgelegt ist und die Bitkombinationen 02/01 bis 07/15 enthält. Die Zeichen des Grundzeichensatzes dürfen nicht benutzt werden, um zusammengesetzte Zeichen zu generieren die in dieser Internationalen Norm festgelegt wird. Der Grundzeichensatz enthält die Buchstaben des lateinischen Basialphabets, einige diakritische Zeichen, die nur eigenständig verwendet werden dürfen (diakritische Zeichen mit Zeichenschritt), und einige nichtalphabetische Zeichen.

Der Ergänzungszeichensatz (G1-Zeichensatz) besteht aus den Schriftzeichen, die durch die internationale Registrierung Nr. 156 festgelegt sind und die Bitkombinationen 10/00 bis 11/15 und 13/00 bis 15/15 und diakritische Zeichen ohne Zeichenschritt zur Erzeugung von zusammengesetzten Zeichen in den Bitkombinationen 12/00 bis 12/15 enthält. Die Schriftzeichen enthalten eine Anzahl von Zeichen, die den Grundzeichensatz ergänzen.

Ein diakritisches Zeichen ohne Zeichenschritt darf nur in Verbindung mit den Buchstaben des Basiszeichensatzes und mit LEERZEICHEN verwendet werden.

Die erlaubten Kombinationen von diakritischen Zeichen ohne Zeichenschritt mit Buchstaben sind in Tabelle 4 festgelegt. Dieser Satz von Kombinationen ist in Anhang C zusammengefasst.

Die Code-Tabelle des Grund- und des Ergänzungszeichensatzes ist in Tabelle 1 dargestellt. Die schattierten Positionen der Bitkombination sind reserviert, wie in 8.2 festgelegt ist.

Die Namen der Zeichen des Grundzeichensatzes sind in Tabelle 2 festgelegt.

Die Namen der Schriftzeichen und der diakritischen Zeichen ohne Zeichenschritt des Ergänzungszeichensatzes sind in Tabelle 3 festgelegt. Da die diakritischen Zeichen ohne Zeichenschritt in dieser Norm nicht als eigenständige Zeichen betrachtet werden, wurden sie in Tabelle 3 klein und kursiv geschrieben.

ANMERKUNG Die schattierten Codepositionen 00/00 bis 01/15 und 07/15 bis 09/15 sind nicht Gegenstand dieser Norm.

8.2 Erklärung der Code-Tabelle

8.2.1 Die Bitkombinationen 10/04 und 10/06 sind reserviert für zukünftige Normung und sollen nicht benutzt werden.

8.2.2 Die diakritischen Zeichen ohne Zeichenschritt Spalte 12 sind nur in Verbindung mit den lateinischen Buchstaben des Grundzeichensatzes und LEERZEICHEN zu verwenden (siehe Anhang C). Die Bildzeichen, die in Spalte 12 dargestellt werden, zeigen die diakritischen Zeichen als eigenständige Zeichen.

8.2.3 Die Bitkombinationen 12/00, 12/09 und 12/12 sind für mögliche zusätzliche diakritische Zeichen reserviert und dürfen nicht benutzt werden.

8.2.4 Die Bitkombinationen 13/08 bis 13/11 und 14/05 sind für zukünftige Normung reserviert und sollen nicht benutzt werden.

8.3 Codierte Darstellung der Schriftzeichen des Zeichenvorrats

Die codierte Darstellung der Schriftzeichen des Zeichenvorrats, der in dieser Internationalen Norm definiert ist, ist in Tabelle 4 festgelegt. Die Formate der codierten Darstellung sind folgendermaßen:

a) Buchstaben mit Akzent

Jeder Buchstabe mit Akzent wird dargestellt durch eine Folge von Bitkombinationen, die codierte Darstellung eines diakritischen Zeichens ohne Zeichenschritt (ein Element des Ergänzungszeichensatzes) gefolgt von der codierten Darstellung eines lateinischen Buchstabens des Basiszeichensatzes (ein Element des Grundzeichensatzes).

b) Diakritische Zeichen als eigenständige Schriftzeichen

Die diakritischen Zeichen, die Elemente des Grundzeichensatzes sind (GRAVIS, ZIRKUMFLEX und TILDE), werden dargestellt als separate Schriftzeichen mit der betreffenden einzigen Bitkombination in dem Bereich von 02/01 bis 07/14.

Die übrigen 13 diakritischen Zeichen der Spalte 12 werden als eigenständige Schriftzeichen durch eine Folge von Bitkombinationen, bestehend aus der codierten Darstellung des diakritischen Zeichens ohne Zeichenschritt (ein Element des Ergänzungsschriftzeichensatzes), gefolgt von der codierten Darstellung des LEERZEICHENS, dargestellt, z. B. die Bitkombination 02/00.

c) Alle sonstigen Schriftzeichen des Zeichenvorrats

Alle sonstigen Schriftzeichen, die keine Buchstaben mit Akzent und keine eigenständigen diakritischen Zeichen sind, werden entweder durch den Grund- oder den Ergänzungsschriftzeichensatz dargestellt und bestehen aus der betreffenden einzelnen Bitkombination in den Bereichen 02/01 bis 07/14 oder 10/00 bis 15/15.

Abhängig von der Benutzung der Code-Erweiterungstechnik kann es notwendig sein, dass vor einem Zeichen des Grund- oder Ergänzungsschriftzeichensatzes eine Code-Erweiterungsfunktion vorweggeschickt werden muss, um den Zeichensatz aufzurufen.

ANMERKUNG Erklärungen über einige Buchstaben:

Der Buchstabe mit Akzent KLEINBUCHSTABE G MIT Cédille wurde in der Ausgabe 1983 dieser Norm „kleines g mit Akut“ genannt. Aus Kompatibilitätsgründen wurde die codierte Darstellung nicht geändert. Der Name wurde an ISO/IEC 10646-1 angepasst. Die Cédille ist aus Darstellungsgründen über dem „g“ platziert. Es ist vorgesehen, den Buchstaben in der lettischen Sprache zu verwenden und er korrespondiert mit dem Zeichen GROSSBUCHSTABE G MIT CÉDILLE.

In dieser Norm ist kein GROSSBUCHSTABE ETH vorhanden. Der Buchstabe GROSSBUCHSTABE D MIT STRICH kann als Großbuchstabe des isländischen ETH verwendet werden. Es soll angemerkt werden, dass in ISO/IEC 10646 und ISO/IEC 10367 beide Zeichen GROSSBUCHSTABE ETH und GROSSBUCHSTABE D MIT STRICH vorhanden sind.

9 Untermengen des Zeichenvorrats

Der Grund für die Festlegung von Untermengen ist, mit Geräten mit eingeschränktem Zeichenvorrat kommunizieren zu können. Ein Beispiel für ein Gerät der Textkommunikation mit eingeschränktem Zeichenvorrat ist ein Terminal, das ein Ausgabegerät mit austauschbarem Druckkopf besitzt (physikalisch oder anders). Jedoch, um die Regeln dieser Norm zu erfüllen, muss das Terminal alle Schriftzeichen dieser Norm empfangen und darstellen können, evtl. durch Benutzung einer oder mehrerer Druckköpfe.

Untermengen werden nach den folgenden Regeln gebildet:

- a) Eine Untermenge muss das LEERZEICHEN, die 26 kleinen und die 26 großen lateinischen Buchstaben ohne Akzent enthalten.
- b) Eine Untermenge muss 10 Ziffern enthalten.
- c) Eine Untermenge muss die folgenden Zeichen enthalten:

graphische Darstellung	Name
'	APOSTROPH (Auslassungszeichen)
(	RUNDE KLAMMER LINKS
)	RUNDE KLAMMER RECHTS
,	KOMMA (Beistrich)
-	MITTESTRICH (Bindestrich, Gedankenstrich, Minuszeichen)
.	PUNKT
/	SCHRÄGSTRICH
:	DOPPELPUNKT
?	FRAGEZEICHEN
+	PLUSZEICHEN
=	GLEICHHEITSZEICHEN

- d) Eine Untermenge darf weitere Schriftzeichen aus dem Zeichenvorrat dieser Norm enthalten.

- e) Eine Untermengen darf keine Zeichen enthalten, die in dieser Norm nicht festgelegt sind.
- f) Zwei oder mehr Schriftzeichen des Zeichenvorrats dürfen nicht als Einzelzeichen in der Untermenge enthalten sein.

Die Prozedur für die Registrierung der Untermengen des Zeichenvorrats ist in ISO/IEC 7350 festgelegt.

Es wird angestrebt, dass die Kennzeichnungsparameter der registrierten Untermengen als Parameterwert für die Steuerfunktion IDENTIFY GRAPHIC SUBREPERTOIRE (IGS), wie in ISO/IEC 10538 festgelegt, verwendet werden können.

10 Identifizierung der Wahlmöglichkeiten

10.1 Zweck und Kontext der Identifizierung

CC-Datenelemente, die mit einer Wahlmöglichkeit dieser Norm konform sind, müssen alle oder einen Teil der in dieser Norm festgelegten codierten Informationen enthalten, die zwischen Sender und Empfänger übermittelt werden. Die Identifizierung der ausgewählten Wahlmöglichkeit dieser Norm muss auch dem Empfänger zur Verfügung stehen. Der Weg, wie die geschieht, ist nicht durch diese Norm festgelegt.

Jedoch, einige Normen für die Übermittlung von codierten Informationen können erlauben oder regeln, dass die Identifizierung in den CC-Datenelementen enthalten ist und ein Teil der Übermittlungsinformation ist. Dieser Absatz legt eine codierte Darstellung für die Wahlmöglichkeiten dieser Norm fest. Solche codierten Darstellungen formen ein ganzes oder einen Teil eines Erkennungs-Datenelements, das in der Übermittlungsinformation in Übereinstimmung mit den relevanten Normen enthalten ist.

10.2 Identifizierung der Codierungsmethode

Die verwendete Codierungsmethode muss durch eine der folgenden Ankündigungsfolgen festgelegt werden:

ESC 02/00 04/10 muss den 7-Bit-Code festlegen (wie in Anhang A beschrieben);

ESC 02/00 04/11 muss den 8-Bit-Code festlegen.

10.3 Identifizierung von Grund- und Ergänzungszeichensätze

Die Escape-Folgen, die die Grund- und Ergänzungszeichensätze bereitstellen sind:

- ESC 02/08 04/02 : stellt den Grundzeichensatz der gegenwärtigen Ausgabe dieser Norm (ISO-IR 6) als den G0-Satz bereit;
- ESC 02/13 05/02 : stellt den Ergänzungszeichensatz der gegenwärtigen Ausgabe dieser Norm (ISO-IR156) als den G1-Satz bereit;
- ESC 02/14 05/02 : stellt den Ergänzungszeichensatz der gegenwärtigen Ausgabe dieser Norm als den G2-Satz bereit;
- ESC 02/15 05/02 : stellt den Ergänzungszeichensatz der gegenwärtigen Ausgabe dieser Norm als den G3-Satz bereit.