

Selbstbeschreibende Austauschdatei für den Informationsaustausch

DIN
66 249

Information processing; data descriptive file for information interchange

Einsprüche bis 31. Aug 1987

Anwendungswarnvermerk
auf der letzten Seite beachten!

Zusammenhang mit der von der International Organization for Standardization (ISO) herausgegebenen Internationalen Norm ISO 8211-1985, siehe Erläuterungen.

Inhalt

	Seite
1 Anwendungsbereich und Zweck	1
2 Normkonformität	2
3 Bezeichnung	2
4 Begriffe	3
5 Selbstbeschreibende Austauschdatei	5
6 Beschreibung der Datentypen und -strukturen	22
7 Erweiterung von Zeichensätzen	30
8 Austauschstufen	32
Anhang A Hinweise für die Implementierung	33
Anhang B Beispiele für Datenfelder	37
Anhang C Dateiaustauschmodell und logische Satzstruktur	41
Zitierte Normen	44
Erläuterungen	44

1 Anwendungsbereich und Zweck

Diese Norm legt ein Dateiaustauschformat fest, das dazu bestimmt ist, den Austausch von Dateien zwischen Datenverarbeitungsanlagen zu erleichtern. Es ist nicht dazu bestimmt, als Satzformat zur Speicherung systemspezifischer Dateien bestimmter Systeme zu dienen. Es beschreibt eine allgemeine Struktur, die verwendet werden kann, um zwischen Rechnersystemen Datensätze auszutauschen, die eine Vielfalt von Datentypen und Strukturen enthalten. Es dient als Mittel zur Beschreibung des Inhalts von Datensätzen, legt aber nicht deren spezifischen Inhalt fest.

Fortsetzung Seite 2 bis 44

In Abhängigkeit von der Komplexität der zugelassenen Strukturen werden drei Austauschstufen definiert. Diese Norm beschreibt:

- a) Eine medienunabhängige Datei und Datensatzbeschreibungen für den Informationsaustausch unter Berücksichtigung der Normen für Kennsätze und Dateianordnung, siehe DIN 66 029, DIN 66 229 und DIN 66 239.
- b) Die Beschreibung von Datenelementen, Vektoren, Matrizen und Hierarchien, die Zeichenfolgen, Bitfolgen und Zahlen sind. Die Zahlendarstellung ist in DIN 66 250 festgelegt.
- c) Eine selbstbeschreibende Austauschdatei, die einen Datenbeschreibungssatz und die durch ihn beschriebenen Datensätze umfaßt und die einen Informationsaustausch unter Zuhilfenahme möglichst weniger externer Beschreibungselemente ermöglicht.
- d) Den Datenbeschreibungssatz, der die Eigenschaften jedes Datenfeldes innerhalb eines Datensatzes festlegt.

2 Normkonformität

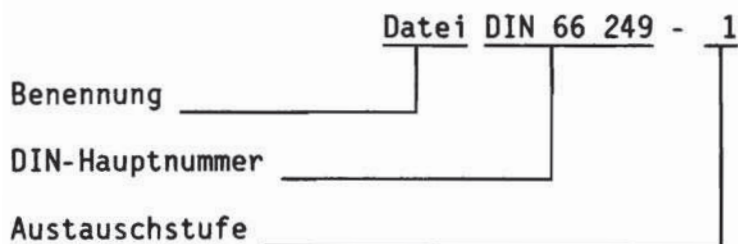
Eine Austauschdatei stimmt mit dieser Norm überein, wenn der Datenbeschreibungssatz und die Datensätze den Festlegungen in dieser Norm entsprechen.

Eine Aussage zur Normkonformität muß die Austauschstufe angeben, mit der der Inhalt einer Datei übereinstimmt.

Voraussetzung ist die Normkonformität mit den zitierten Normen für die Kennsätze und Dateianordnung sowie die Zahlendarstellung.

3 Bezeichnung

Bezeichnung einer Datei, die die Anforderungen dieser Norm, z. B. in der Austauschstufe 1, erfüllt:



4 Begriffe

Anmerkung: Die in diesem Abschnitt gegebenen Definitionen dienen zum Verständnis der vorliegenden Norm. Der Festlegung dieser Begriffe für die allgemeine Anwendung in der Informationsverarbeitung soll dadurch nicht vorgegriffen werden.

Zu den deutschen Benennungen sind die englischen Fachausdrücke in Klammern angegeben. Sie sind nicht Bestandteil der Norm.

4.1 Abbilden (to map)

Die Beziehung zwischen den Elementen zweier Strukturen herstellen.

4.2 Alphanumerisches Zeichen (alphanumeric character)

Ein Zeichen aus den Spalten 02 bis 07 (mit Ausnahme des Zeichens 07/15) der Tabelle 2.

4.3 Bit-Feld (bit field)

Ein Datenfeld bestehend aus binären Ziffern, das falls notwendig, nach rechts hin mit binären Nullen bis zu einem Byte aufgefüllt wird.

Anmerkung: Siehe auch Bitfolge in Zeichendarstellung.

4.4 Zeichendarstellung für eine Bitfolge (character mode bit string)

Eine Folge von Zeichen (0 oder 1), die eine Folge von Ziffern darstellt.

4.5 Code-Umschaltungs-Zeichen (escape character)

Zeichen zur Code-Erweiterung, das mit einem oder mehreren Folgezeichen benutzt wird, um die Bedeutung eines oder mehrerer aufeinanderfolgender Zeichen zu ändern.

4.6 Datei (file)

Eine Zusammenfassung von Sätzen, die in einem verarbeitungstechnischen Zusammenhang stehen.

4.7 Datenbeschreibungssatz (data descriptive record)

Ein Satz, der logisch den Datensätzen vorangeht und der Steuerparameter und Datenbeschreibungen enthält, die zur Interpretation der Datensätze und Datenelemente notwendig sind. Der Datenbeschreibungssatz ist logisch der erste Satz der Datei nach den Anfangskennsätzen (sofern vorhanden).

4.8 Datensatz (data record)

Ein Satz, der eine Zusammenfassung von Datenfeldern enthält.

4.9 elementar (elementary)

Eigenschaft eines Feldinhaltes, der ohne Bedeutungsverlust nicht teilbar ist.

4.10 Feld-Trennzeichen (field terminator)

Ein Zeichen, das variable Felder innerhalb eines Satzes beendet. Siehe Platz 01/14 in Tabelle 2.

4.11 Hierarchie (hierarchy)

Eine geordnete Baumstruktur, die sich aus einem Übergeordneten Wurzelknoten mit mehreren nachfolgenden geordneten Teilbäumen zusammensetzt, deren Wurzelknoten Zwischenknoten darstellen und die mit Blattknoten enden.

4.12 Inhaltsverzeichnis (directory)

Eine Tabelle von Einträgen, in denen die Datenfelder eines Satzes beschrieben sind.

4.13 Kennung (tag)

Kennzeichen eines Eintrags im Inhaltsverzeichnis, das den internen Namen eines zugeordneten Feldes spezifiziert.

4.14 Bezeichner (label)

Eine Folge von Zeichen, um ein Feld und dessen Inhalt zu identifizieren oder zu kennzeichnen.

4.15 Linksseitige Verarbeitungsfolge (preorder traversal sequence)

Eine Folge von Knoten einer Hierarchie, die ihre Struktur (teilweise) bestimmt.

4.16 Satz (record)

Eine Zusammenfassung von miteinander in Verbindung stehenden Daten, unabhängig von ihrer physikalischen Umgebung.

4.17 Selbstbeschreibende Austauschdatei (data descriptive file)

Eine Datei, die einen Datenbeschreibungssatz und die durch ihn beschriebenen Datensätze enthält.

4.18 Trennzeichen (delimiter)

Ein einzelnes Zeichen, das Daten unterteilt und organisiert. Siehe Abschnitte 4.10 und 4.19.

4.19 Unterfeld-Trennzeichen (unit terminator)

Ein Zeichen, mit dem variable Felder in Teilfelder geteilt werden. Siehe Platz 01/15 in Tabelle 2.

4.20 Variables Feld (variable-length field)

Ein Feld, dessen Länge durch die Anzahl der Zeichen, die in ihm vorkommen, bestimmt wird. Die Länge kann von einem zum anderen auftreten verschieden sein.

4.21 Vorspann (leader)

Ein Feld fester Länge, welches am Anfang von jedem Satz steht und Parameter für die Behandlung des Satzes beinhaltet.

5 Selbstbeschreibende Austauschdatei

5.1 Allgemeine Struktur

Dieser Abschnitt beschreibt die allgemeine Struktur der selbstbeschreibenden Austauschdatei. In weiteren Abschnitten wird die Struktur im einzelnen dargestellt. Tabelle 1 zeigt eine Datei mit Kennsätzen.

Tabelle 1. Datei und Kennsätze

Dateianfangskennsätze
Selbstbeschreibende Austauschdatei: Datenbeschreibungssatz Datensätze
Dateiendekennsätze

Diese Norm beschreibt eine selbstbeschreibende Austauschdatei, die zusätzliche Kennsätze für spezielle Medien benötigt. Jede Datei besteht aus

- dem Datenbeschreibungssatz und
- den Datensätzen.

Die gesamte Struktur ist im Bild 1 abgebildet. Bild 1 stellt die hierarchische Struktur von Datenbeschreibungssätzen und Datensätzen zusammen mit den Vorspannen und Inhaltsverzeichnissen jedes Satzes, typischer Sätze und ein typisches variables Datenfeld für jeden Satz dar.

Die Struktur von Vorspann, Inhaltsverzeichnis sowie von Feldern und Sätzen ist immer gleich und unabhängig von ihrem Inhalt.

Jedes variable Feld des Datenbeschreibungssatzes beschreibt die Daten der gleichartigen variablen Datenfelder in den Datensätzen. Die Gleichartigkeit und der Zusammenhang wird durch eine gleiche Kennung in dem variablen Feld des Datenbeschreibungssatzes und den variablen Datenfeldern der Datensätze gekennzeichnet.

Felder, in denen "a"-Zeichen zugelassen sind, werden linksbündig beschrieben und nach rechts mit Leerzeichen aufgefüllt. Felder, in denen "n"-Zeichen zugelassen sind, werden rechtsbündig beschrieben und nach links entweder mit Nullen oder Leerzeichen aufgefüllt.

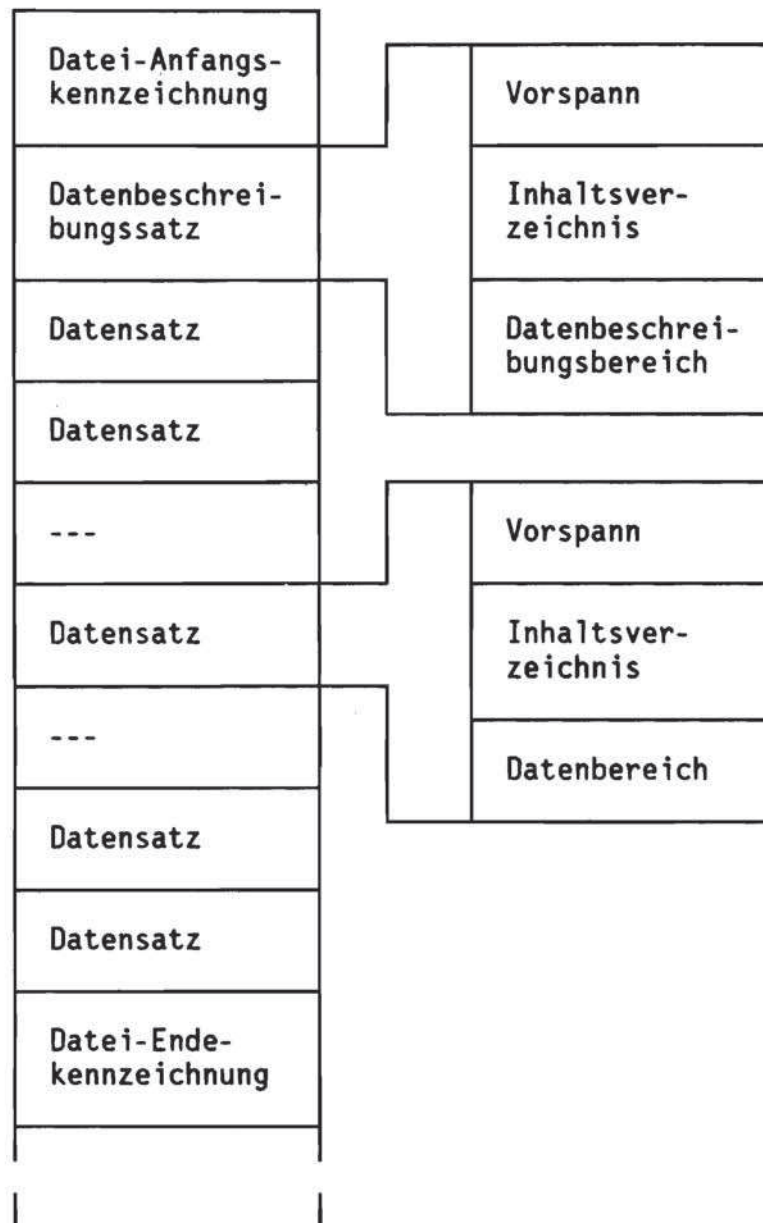
In dieser Norm bedeutet "n"-Zeichen eine Ziffer von 0 bis 9. Ein "a"-Zeichen bedeutet irgendein Zeichen des 8-Bit-Code aus den Spalten 02 bis 05 aus Tabelle 2.

Tabelle 2. 8-Bit-Code-Tabelle der "Allgemeinen Referenzversion" nach DIN 66 303, Ausgabe 11.86, Code-Tabelle 1; ergänzt mit Untergruppen-Trennzeichen auf dem Platz 01/14 und dem Teilgruppen-Trennzeichen auf dem Platz 01/15

					b_8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
					b_7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
					b_6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
					b_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	
0	0	0	0	00			SP	0	@	P		p			NBSP	°	À	Đ	à	ò	
0	0	0	1	01			!	1	A	Q	a	q			i	±	Á	Ñ	á	ñ	
0	0	1	0	02			"	2	B	R	b	r			¢	²	Â	Ò	â	ò	
0	0	1	1	03			#	3	C	S	c	s			£	³	Ã	Ó	ã	ó	
0	1	0	0	04			\$	4	D	T	d	t			¤	'	Ä	Ö	ä	ô	
0	1	0	1	05			%	5	E	U	e	u			¥	µ	Å	Õ	å	ö	
0	1	1	0	06			&	6	F	V	f	v			¦	¶	Æ	Ö	æ	ö	
0	1	1	1	07			'	7	G	W	g	w			§	·	Ç	×	ç	÷	
1	0	0	0	08			(	8	H	X	h	x			¨	,	È	Ø	è	ø	
1	0	0	1	09			)	9	I	Y	i	y			©	¹	É	Ù	é	ù	
1	0	1	0	10			*	:	J	Z	j	z			ª	º	Ê	Ú	ê	ú	
1	0	1	1	11			+	;	K	[	k	{			«	»	Ë	Û	ë	û	
1	1	0	0	12			,	<	L	\	l				¬	¼	Ì	Ü	ì	ü	
1	1	0	1	13			-	=	M	]	m	}			SHY	½	Í	Ý	í	ý	
1	1	1	0	14		IS_2	.	>	N	^	n	~			®	¾	Î	Ï	î	ï	
1	1	1	1	15		IS_1	/	?	O	_	o				™	¿	Ï	ß	ï	ÿ	

**Kennsatz und
Dateistruktur**

Grobdarstellung



zur nächsten
Datei

Bild 1. Dateien und Kennsätze

5.2 Datenbeschreibungssatz

Der Datenbeschreibungssatz besteht aus den Bereichen und Trennzeichen, siehe Tabelle 3.

Tabelle 3. Dateibeschreibungssatz

Name des Bereiches	Länge
Vorspann	24
Inhaltsverzeichnis	$k \cdot p$
Feld-Trennzeichen	1
Datenbeschreibungsbereich	variabel
Feld-Trennzeichen	1

Der Datenbeschreibungssatz ist der erste Satz der Datei. Er besteht aus

- einem Vorspann aus 24 Zeichen,
- einem Inhaltsverzeichnis der Länge $k \cdot p$, das durch das Feld-Trennzeichen (01/14) beendet wird, wobei k die Anzahl der Einträge im Inhaltsverzeichnis und p die Länge jedes Eintrages bedeutet, siehe Abschnitt 5.2.2 und
- dem Datenbeschreibungsbereich, einer Anzahl von k variablen Feldern, von denen jedes durch ein Feld-Trennzeichen beendet wird.

5.2.1 Vorspann des Datenbeschreibungssatzes

Der Vorspann des Datenbeschreibungssatzes besteht aus den in Tabelle 4 beschriebenen Feldern.

Tabelle 4. Vorspann des Datenbeschreibungssatzes

Stelle	Feldname	Länge	Inhalt
0 bis 4	Satzlänge	5	"n"-Zeichen. Gibt die Gesamtanzahl von Zeichen im Datenbeschreibungssatz an. Der Inhalt 0 bedeutet, daß der Datenbeschreibungssatz eine Länge von mehr als 99999 Zeichen hat.
5	Austauschstufe	1	"n"-Zeichen. Gibt die Austauschstufe, wie im Abschnitt 8 beschrieben, an. 1 bedeutet: Austauschstufe 1 2 bedeutet: Austauschstufe 2 3 bedeutet: Austauschstufe 3
6	Vorspann-Kennzeichen	1	"a"-Zeichen. Enthält das Zeichen L, das anzeigt, daß dieser Vorspann zum Datenbeschreibungssatz gehört.
7	Code-Erweiterungsanzeiger in Datenfeldern	1	"a"-Zeichen. Gibt an, ob in Datenfeldern ESC-Folgen auftreten, durch die ein erweiterter Zeichenvorrat, siehe DIN 66 203, bereitgestellt wurde. Leerzeichen bedeutet keine Erweiterung. E bedeutet eine Erweiterung.
8	Reserviert für spätere Normung	1	Leerzeichen
9	Anwendungssystem-Anzeiger	1	Leerzeichen. Dieses Feld ist reserviert für spätere Normung und muß in dieser Version ein Leerzeichen enthalten.
10 bis 11	Länge des Feldes Feldsteuerzeichen	2	"n"-Zeichen. Gibt die Anzahl der Zeichen des Feldes "Feldsteuerzeichen" an. Das Feld darf nur die Zeichen "00", "06" oder "09" enthalten, siehe Abschnitte 5.3.3.2 und 7 sowie Anhang B.2.1.
12 bis 16	Basisposition der Daten	5	"n"-Zeichen. Gibt die Position des ersten Bytes des Datenbeschreibungsbereiches an. Sie ist gleich der Länge in Bytes des Vorspanns und des Inhaltsverzeichnisses einschließlich des Feld-Trennzeichens am Ende des Inhaltsverzeichnisses.

Tabelle 4. (Fortsetzung)

Stelle	Feldname	Länge	Inhalt
17 bis 19	Anzeiger für den codierten Zeichen-vorrat	3	"a"-Zeichen. Gibt an, ob eine vor-vereinbarte Code-Erweiterung in der Datei benutzt wird. Die Werte in diesem Feld haben folgende Bedeutung: a) (02/00) (02/00) (02/00) Gibt an, daß als vorvereinbarter Zeichensatz die Internationale Referenz-Version nach DIN 66 003 bereitgestellt wird. b) (02/00) (02/01) (02/00) Gibt an, daß vorvereinbarte Zeichensätze für ein oder mehrere Datenfelder bereitgestellt sind, siehe Abschnitt 7. c) Andere "a"-Zeichen für Umschaltfolgen geben an, daß ein vorvereinbarter Zeichensatz für die ganze Datei bereitgestellt wird, siehe Abschnitt 7.
20 bis 23	Feldlängentabelle für Einträge	4	"n"-Zeichen. Gibt die Länge der Einträge in das Inhaltsverzeichnis an und besteht aus den in Tabelle 5 beschriebenen Feldern.

Tabelle 5. Feldlängentabelle für Einträge des Datenbeschreibungssatzes

Stelle	Feldname	Länge	Inhalt
20	Länge des Feldes Feldlänge	1	"n"-Zeichen. Gibt die Länge in Bytes des Feldes "Feldlänge" eines Eintrages im Inhaltsverzeichnis an. Der Wertebereich ist 1 bis 9.
21	Länge des Feldes Feldposition	1	"n"-Zeichen. Gibt die Länge in Bytes des Feldes "Feldposition" eines Eintrages im Inhaltsverzeichnis an. Der Wertebereich ist 1 bis 9.
22	Reserviert für spätere Normung	1	"n"-Zeichen. Reserviert für spätere Normung. Der Inhalt muß 0 sein.
23	Länge des Feldes Feldkennung	1	"n"-Zeichen. Gibt die Länge in Bytes des Feldes "Feldkennung" eines Eintrages im Inhaltsverzeichnis an. Der Wertebereich ist 1 bis 7.

5.2.2 Inhaltsverzeichnis des Datenbeschreibungssatzes

Das Inhaltsverzeichnis des Datenbeschreibungssatzes besteht aus wiederholten Einträgen, deren Feldlängen in der Feldlängentabelle für Einträge beschrieben sind. Das Inhaltsverzeichnis muß für jedes Feld des nachfolgenden Datenbeschreibungsbereiches einen Eintrag haben und endet mit dem Feld-Trennzeichen (01/14). Der Datenbeschreibungssatz muß die Feldkennungen aller Datensätze definieren.

Jeder Eintrag im Inhaltsverzeichnis gibt die Kennung, Länge und Position des zugehörigen Datenbeschreibungsfeldes an. Er besteht aus $m + n + t$ Bytes, wobei

- m die Länge des Feldes "Feldlänge", entsprechend dem Inhalt der Stelle 20 in Tabelle 4, ist;
- n die Länge des Feldes "Feldposition", entsprechend dem Inhalt der Stelle 21 in Tabelle 4, ist;
- t die Länge des Feldes "Feldkennung", entsprechend dem Inhalt der Stelle 23 in Tabelle 4, ist.

Aufbau eines Eintrages siehe Tabelle 6, dabei ist i der Index des Eintrags in das Inhaltsverzeichnis und $p = t + m + n$.

Die Einträge im Inhaltsverzeichnis des Datenbeschreibungssatzes müssen eindeutig den Datenbeschreibungsfeldern entsprechen.

Für hierarchische Datenstrukturen, siehe Tabelle 8, müssen die Einträge im Inhaltsverzeichnis des Datenbeschreibungssatzes in der gleichen Ordnung sein, wie die linksseitige Verarbeitungsfolge des generischen Datenbaumes.

Tabelle 6. Eintrag im Inhaltsverzeichnis des Datenbeschreibungssatzes

Stelle	Feldname	Länge	Inhalt
$p \cdot (i-1)$	Feldkennung	t	"a"-Zeichen. Dient zur Identifizierung eines Datenbeschreibungsfeldes und besteht aus 1 bis 7 "a"-Zeichen. Die gleiche Feldkennung darf im Datenbeschreibungssatz nur einmal vorkommen. Anmerkung: Die Folge 0 ... bedeutet t - 1 führende Nullen. Feldkennung 0 ... 0, sofern vorhanden, identifiziert das Feld Dateikontrollfeld und kommt nur im Datenbeschreibungssatz vor. Feldkennung 0 ... 1 identifiziert das Feld "Beschreibung des Satz-kennungsfeldes" und muß vorhanden sein. Feldkennung 0 ... 2, sofern vorhanden, identifiziert ein wahlfreies Benutzerfeld im Datenbeschreibungssatz ohne zugehöriges Datenfeld im Datensatz. Dieses Benutzerfeld wird dem Benutzer zur Verfügung gestellt. Anmerkung: Der Inhalt dieses Feldes ist durch den Anwender festgelegt und kann dazu benutzt werden, eine erweiterte Dateibeschreibung oder zusätzliche Verarbeitungskontrollen für die Datei zu übermitteln. Feldkennungen 0 ... 3 bis 0 ... 9 sind für spätere Normung reserviert. Die Feldkennungen 0 ... 0 bis 0 ... 9, sofern vorhanden, müssen als erste Felder im Inhaltsverzeichnis in numerisch aufsteigender Folge auftreten.
$p \cdot (i-1) + t$	Feldlänge	m	"n"-Zeichen. Gibt die Länge in Bytes des zugehörigen Datenbeschreibungsfeldes einschließlich Feld-Trennzeichen an und wird rechtsbündig mit führenden Nullen eingetragen.

Tabelle 6. (Fortsetzung)

Stelle	Feldname	Länge	Inhalt
$p \cdot (i-1) + t + m$	Feldposition	n	"n"-Zeichen. Gibt die relative Position des ersten Bytes im Datenbeschreibungsfeld an, auf den sich der Eintrag bezieht. "n"-Zeichen sind rechtsbündig zu schreiben und mit Nullen aufzufüllen. Das erste Byte des Datenbeschreibungsfeldes, das auf das Inhaltsverzeichnis folgt, soll mit 0 benummert werden.

5.2.3 Datenbeschreibungsbereich

Der Datenbeschreibungsbereich enthält in seinen Datenfeldern die Information, die die zugehörigen Datenfelder der Datensätze beschreibt. Sie haben die gleiche Kennung und stellen die Steuerparameter für eine automatische Verarbeitung zur Verfügung. Diese Felder haben ein vordefiniertes Format, welches durch den Vorspann des Datenbeschreibungssatzes bestimmt ist. Alle Felder enden mit einem Feld-Trennzeichen (01/14).

Tabelle 7. Aufbau des Datenbeschreibungsbereichs

Name	Kennung
Dateikontrollfeld	0...0
Beschreibung des Satz kennungsfeldes	0...1
Datenbeschreibungsfelder	"a"-Zeichen

5.2.3.1 Dateikontrollfeld

Das Dateikontrollfeld (Kennung 0...0) enthält folgende Unterfelder:

- a) Feldsteuerzeichen, sofern vorhanden