

---

---

**Information technology — Automatic  
identification and data capture  
techniques — PDF417 bar code  
symbology specification**

*Technologies de l'information — Techniques d'identification  
automatique et de capture des données — Spécifications pour la  
symbologie de code à barres PDF417*

**iTeh STANDARD PREVIEW  
(standards.iteh.ai)**

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC 15438:2006](#)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

© ISO/IEC 2006

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

Foreword.....	vi
Introduction .....	vii
<b>1 Scope .....</b>	<b>1</b>
<b>2 Normative references .....</b>	<b>1</b>
<b>3 Terms and definitions.....</b>	<b>2</b>
<b>4 Symbols, operations and abbreviated terms .....</b>	<b>3</b>
4.1 Symbols .....	3
4.2 Mathematical operations.....	4
4.3 Abbreviated terms .....	4
<b>5 Requirements .....</b>	<b>5</b>
5.1 Symbology characteristics .....	5
5.1.1 Basic characteristics .....	5
5.1.2 Summary of additional features .....	6
5.2 Symbol structure .....	7
5.2.1 PDF417 symbol parameters.....	7
5.2.2 Row parameters.....	7
5.2.3 Codeword sequence.....	7
5.3 Basic encodation .....	8
5.3.1 Symbol character structure .....	8
5.3.2 Start and stop characters .....	9
5.4 High level (data) encodation.....	10
5.4.1 Function codewords.....	10
5.4.2 Text Compaction mode .....	13
5.4.3 Byte Compaction mode.....	17
5.4.4 Numeric Compaction mode .....	19
5.4.5 Advice to select the appropriate compaction mode .....	21
5.4.6 Treatment of PDF417 reserved codewords.....	21
5.5 Extended Channel Interpretation .....	21
5.5.1 Encoding the ECI assignment number.....	22
5.5.2 Pre-assigned and default Extended Channel Interpretations .....	23
5.5.3 Encoding ECI sequences within compaction modes .....	23
5.5.4 Post-decode protocol .....	25
5.6 Determining the codeword sequence.....	25
5.7 Error detection and correction .....	26
5.7.1 Error correction level.....	26
5.7.2 Error correction capacity .....	26
5.7.3 Defining the error correction codewords .....	27
5.8 Dimensions.....	27
5.8.1 Minimum width of a module (X).....	27
5.8.2 Row height (Y).....	28
5.8.3 Quiet zones.....	28
5.9 Defining the symbol format .....	28
5.9.1 Defining the aspect ratio of the module .....	28
5.9.2 Defining the symbol matrix of rows and columns .....	28
5.10 Generating the error correction codewords .....	30
5.11 Low level encodation.....	31
5.11.1 Clusters.....	32
5.11.2 Determining the symbol matrix .....	32
5.11.3 Determining the values of the left and right row indicators.....	32

5.11.4	Row encoding.....	33
5.12	Compact PDF417.....	33
5.13	Macro PDF417.....	33
5.13.1	Compaction modes and Macro PDF417 .....	34
5.13.2	ECIs and Macro PDF417 .....	34
5.14	User guidelines .....	34
5.14.1	Human readable interpretation.....	34
5.14.2	Autodiscrimination capability.....	34
5.14.3	User-defined application parameters.....	34
5.14.4	PDF417 symbol quality.....	35
5.15	Reference decode algorithm.....	35
5.16	Error detection and error correction procedure .....	35
5.17	Transmitted data .....	35
5.17.1	Transmitted data in the basic (default) interpretation.....	35
5.17.2	Transmission protocol for Extended Channel Interpretation (ECI) .....	36
5.17.3	Transmitted data for Macro PDF417 .....	37
5.17.4	Transmission of reserved codewords using the ECI protocol.....	37
5.17.5	Symbology identifier.....	37
5.17.6	Transmission using older protocols.....	37
Annex A (normative) Encoding/decoding table of PDF417 symbol character bar-space sequences.....		39
Annex B (normative) The default character set for Byte Compaction mode .....		55
Annex C (normative) Byte Compaction mode encoding algorithm .....		56
Annex D (normative) Numeric Compaction mode encoding algorithm.....		58
Annex E (normative) User selection of error correction level .....		60
E.1	Recommended minimum error correction level.....	60
E.2	Other user consideration of the error correction level.....	60
Annex F (normative) Tables of coefficients for calculating PDF417 error correction codewords .....		61
Annex G (normative) Compact PDF417.....		66
G.1	Description.....	66
G.2	Print quality.....	66
Annex H (normative) Macro PDF417.....		67
H.1	Macro PDF417 overview .....	67
H.2	Macro PDF417 syntax .....	67
H.3	High level encoding considerations .....	70
H.4	Encodation example .....	70
H.5	Macro PDF417 and the Extended Channel Interpretation protocol .....	71
H.6	Macro PDF417 data transmission .....	72
Annex I (normative) Testing PDF417 symbol quality.....		75
Annex J (normative) Reference decode algorithm for PDF417 .....		76
J.1	Initialisation .....	76
J.2	Reference decode algorithm for line decoding.....	76
J.3	Filling the matrix .....	78
J.4	Interpretation .....	79
Annex K (normative) Error correction procedures .....		80
Annex L (normative) Symbology identifier .....		82
Annex M (normative) Transmission protocol for decoders conforming with original PDF417 standards .....		83
M.1	Basic Channel mode.....	83
M.2	GLI encoded symbols.....	83
M.3	Macro PDF417 symbols.....	85
M.4	Transmission of reserved codewords using the original PDF417 protocol .....	86
M.5	Achieving compatibility between old and new PDF417 equipment.....	86
Annex N (informative) Algorithm to minimise the number of codewords .....		89

<b>Annex O</b> (informative) <b>Guidelines to determine the symbol matrix</b> .....	<b>91</b>
<b>O.1</b> <b>Parameters affecting the determination of the matrix</b> .....	<b>91</b>
<b>O.2</b> <b>Guidelines should any parameters not be achieved</b> .....	<b>94</b>
<b>Annex P</b> (informative) <b>Calculating the coefficients for generating the error correction codewords</b> <b>– worked example</b> .....	<b>95</b>
<b>Annex Q</b> (informative) <b>Generating the error correction codewords - worked example</b> .....	<b>96</b>
<b>Annex R</b> (informative) <b>Division circuit procedure for generating error correction codewords</b> .....	<b>99</b>
<b>Annex S</b> (informative) <b>Additional guidelines for the use of PDF417</b> .....	<b>100</b>
<b>S.1</b> <b>Autodiscrimination compatibility</b> .....	<b>100</b>
<b>S.2</b> <b>Pixel-based printing</b> .....	<b>100</b>
<b>Bibliography</b> .....	<b>102</b>

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 15438 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 31, *Automatic identification and data capture techniques*.

This second edition cancels and replaces the first edition (ISO/IEC 15438:2001), which has been technically revised.

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

## Introduction

The technology of bar coding is based on the recognition of patterns of bars and spaces of defined dimensions. There are various methods of encoding information in bar code form, known as symbologies, and the rules defining the translation of characters into bars and space patterns and other essential features are known as the symbology specification.

Manufacturers of bar code equipment and users of bar code technology require publicly available standard symbology specifications to which they can refer when developing equipment and application standards. It is the intent and understanding of ISO/IEC that the symbology presented in this International Standard is entirely in the public domain and free of all user restrictions, licences and fees.

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO/IEC 15438:2006](#)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>



# Information technology — Automatic identification and data capture techniques — PDF417 bar code symbology specification

## 1 Scope

This International Standard specifies the requirements for the bar code symbology known as PDF417. It specifies PDF417 symbology characteristics, data character encodation, symbol formats, dimensions, error correction rules, reference decoding algorithm, and a number of application parameters.

## 2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 646:1991, *Information technology — ISO 7-bit coded character set for information interchange*

ISO/IEC 8859-1, *Information technology — 8-bit single-byte coded graphic character sets — Part 1: Latin alphabet No. 1*

ISO/IEC 15415, *Information technology — Automatic identification and data capture techniques — Bar code print quality test specification — Two-dimensional symbols*

ISO/IEC 15424, *Information technology — Automatic identification and data capture techniques — Data Carrier Identifiers (including Symbology Identifiers)*

ISO/IEC 19762-1, *Information technology — Automatic identification and data capture (AIDC) techniques — Harmonized vocabulary — Part 1: General terms relating to AIDC*

ISO/IEC 19762-2, *Information technology — Automatic identification and data capture (AIDC) techniques — Part 2: Optically readable media (ORM)*

ISO/IEC 24723, *Information technology — Automatic identification and data capture techniques — EAN.UCC Composite bar code symbology specification*

AIM Inc. International Technical Standard: ITS/04-001, *Extended Channel Interpretations — Part 1: Identification Schemes and Protocols*<sup>1</sup>

<sup>1</sup> Published by AIM Global, 125 Warrendale-Bayne Road, Suite 100, Warrendale, PA 15086, USA.

### 3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 19762-1, ISO/IEC 19762-2 and the following apply.

#### 3.1 basic channel model

standard system for encoding and transmitting bar code data where data message bytes are output from the decoder but no control information about the message is transmitted

NOTE A decoder complying with this model operates in Basic Channel Mode.

#### 3.2 bar-space sequence

sequence which represents the module widths of the elements of a symbol character

#### 3.3 cluster

any of three mutually exclusive subsets of PDF417 symbol characters; the symbol characters in a given cluster conform with particular structural rules which are used in decoding the symbology

#### 3.4 compaction mode

any of three data compaction algorithms in PDF417 (Text, Numeric and Byte Compaction modes) which are used to map 8-bit data bytes efficiently to PDF417 codewords

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

#### 3.5 e-distance

distance from leading edge of an element to leading edge of the next similar element, or from trailing edge to trailing edge

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

#### 3.6 error correction codeword

codeword which encodes a value derived from the error correction codeword algorithm to enable decode errors to be detected and, depending on the error correction level, to be corrected

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

#### 3.7 Extended Channel Interpretation ECI

procedure within some symbologies, including PDF417, to replace the default interpretation with another interpretation in a reliable manner

NOTE The interpretation intended prior to producing the symbol can be retrieved after decoding the scanned symbol to recreate the data message in its original format.

#### 3.8 Extended Channel Model

system for encoding and transmitting both data message bytes and control information about the message, the control information being communicated using Extended Channel Interpretation (ECI) escape sequences

NOTE A decoder complying with this model operates in Extended Channel Mode.

#### 3.9 function codeword

codeword which initiates a particular operation within a symbology, for example to switch between data encoding sets, to invoke a compaction scheme, to program the reader, or to invoke Extended Channel Interpretations

**3.10****Global Label Identifier****GLI**

procedure in the PDF417 symbology which behaves in a similar manner to Extended Channel Interpretation

NOTE The GLI system was the PDF417-specific precursor to the symbology-independent ECI system.

**3.11****Macro PDF417**

procedure in the PDF417 symbology logically to distribute data from a computer file across a number of related PDF417 symbols

NOTE 1 The procedure considerably extends the data capacity beyond that of a single symbol.

NOTE 2 This procedure is similar to the Structured Append feature in other symbologies.

**3.12****Mode Latch codeword**

codeword which is used to switch from one mode to another mode, which stays in effect until another latch or shift codeword is implicitly or explicitly brought into use, or until the end of the symbol is reached

**3.13****Mode Shift codeword**

codeword which is used to switch from one mode to another for one codeword, after which encoding returns to the original mode

**3.14****Row Indicator codeword**

PDF417 codeword adjacent to the start or stop character in a row, which encodes information about the structure of the PDF417 symbol in terms of the row identification, total number of rows and columns, and the error correction level

[ISO/IEC 15438:2006](https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006)

<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

**3.15****Symbol Length Descriptor**

first codeword in a PDF417 symbol, which encodes the total number of data codewords in the symbol

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

**4 Symbols, operations and abbreviated terms****4.1 Symbols**

For the purposes of this document, the following mathematical symbols apply. There are some cases where the symbols below have been used in a different manner in an equation. This has been done for consistency with a more general use of the notation and is always clearly defined in the text.

- A* symbol aspect ratio (height to width) of a PDF417 symbol
- b* element width in a symbol character
- c* number of columns in the symbol in the data region (excluding start, stop and row indicator codewords)
- d* data codeword including all function codewords
- E* error correction codeword
- e* edge to similar edge dimension in a symbol character
- F* row number

## ISO/IEC 15438:2006(E)

<i>f</i>	number of substitution errors
<i>H</i>	height of symbol including quiet zone
<i>K</i>	cluster number
<i>k</i>	number of error correction codewords
<i>L</i>	left row indicator
<i>l</i>	number of erasures
<i>m</i>	number of source data codewords prior to the addition of the Symbol Length Descriptor and any pad codewords
<i>n</i>	total number of data codewords including Symbol Length Descriptor and any pad codewords
<i>p</i>	pitch or width of a symbol character
<i>Q<sub>H</sub></i>	horizontal quiet zone
<i>Q<sub>V</sub></i>	vertical quiet zone
<i>R</i>	right row indicator
<i>r</i>	number of rows in the symbol
<i>s</i>	error correction level
<i>W</i>	width of symbol including quiet zone
<i>X</i>	X-dimension or module width
<i>Y</i>	module height (also called row height)

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

ISO/IEC 15438:2006  
<https://standards.iteh.ai/catalog/standards/sist/100bb174-498d-4fd1-a2f3-87e25655dcd1/iso-iec-15438-2006>

### 4.2 Mathematical operations

For the purposes of this document, the following mathematical operations apply.

<i>div</i>	is the integer division operator, rounding down
<i>INT</i>	is the integer value i.e. where a number is rounded down to its whole number component, ignoring its decimal fractions
<i>mod</i>	is the positive integer remainder after division. If the remainder is negative, add the value of the divisor to make it positive. For example, the remainder of -29 160 divided by 929 is -361 which when added to 929 yields 568.

### 4.3 Abbreviated terms

For the purposes of this document, the following abbreviated terms apply.

ECI	Extended Channel Interpretation
GLI	Global Label Identifier

## 5 Requirements

### 5.1 Symbology characteristics

#### 5.1.1 Basic characteristics

PDF417 is a bar code symbology with the following basic characteristics:

- a) Encodable character set:
- 1) Text Compaction mode (see 5.4.1.5) permits all printable ASCII characters to be encoded, i.e. values 32 - 126 inclusive in accordance with ISO/IEC 646 (IRV), as well as selected control characters.
  - 2) Byte Compaction mode (see 5.4.3) permits all 256 possible 8-bit byte values to be encoded. This includes all ASCII characters value 0 to 127 inclusive and provides for international character set support.
  - 3) Numeric Compaction mode (see 5.4.4) permits efficient encoding of numeric data strings.
  - 4) Up to 811 800 different character sets or data interpretations.
  - 5) Various function codewords for control purposes.
- b) Symbol character structure:  $(n, k, m)$  characters of 17 modules ( $n$ ), 4 bar and 4 space elements ( $k$ ), with the largest element 6 modules wide ( $m$ ).
- c) Maximum possible number of data characters per symbol (at error correction level 0): 925 data codewords which can encode:
- 1) Text Compaction mode: 1 850 characters (at 2 data characters per codeword).
  - 2) Byte Compaction mode: 1 108 characters (at 1,2 data characters per codeword).
  - 3) Numeric Compaction mode: 2 710 characters (at 2,93 data characters per codeword)
- At the minimum recommended error correction level, there is a maximum of 863 data codewords which can encode:
- 4) Text Compaction mode: 1 726 characters (at 2 data characters per codeword).
  - 5) Byte Compaction mode: 1 033 characters (at 1,2 data characters per codeword).
  - 6) Numeric Compaction mode: 2 528 characters (at 2,93 data characters per codeword).
- d) Symbol size:
- 1) Number of rows: 3 to 90.
  - 2) Number of columns: 1 to 30.
  - 3) Width in modules: 90X to 583X including quiet zones.
  - 4) Maximum codeword capacity: 928 codewords.
  - 5) Maximum data codeword capacity: 925 codewords.

Since the number of rows and the number of columns are selectable, the aspect ratio of a PDF417 symbol may be varied when printing to suit the spatial requirements of the application.

- e) Selectable error correction: 2 to 512 codewords per symbol (see 5.7).
- f) Non-data overhead:
  - 1) Per row: 73 modules, including quiet zones.
  - 2) Per symbol: a minimum of 3 codewords, represented as symbol characters.
- g) Code type: continuous, multi-row two-dimensional.
- h) Character self-checking: Yes.
- i) Bi-directionally decodable: Yes.

### 5.1.2 Summary of additional features

Additional features which are inherent or optional in PDF417 are summarised below:

- a) **Data compaction:** (inherent) Three schemes are defined to compact a number of data characters into codewords. Generally data is not directly represented on a one character for one codeword basis (see 5.4.1.5 to 5.4.4).
- b) **Extended Channel Interpretations:** (optional) These mechanisms allow up to 811 800 different data character sets or interpretations to be encoded (see 5.5).
- c) **Macro PDF417:** (optional) This mechanism allows files of data to be represented logically and consecutively in a number of PDF417 symbols. Up to 99 999 different PDF417 symbols can be so linked or concatenated and be scanned in any sequence to enable the original data file to be correctly reconstructed (see 5.13).
- d) **Edge to edge decodable:** (inherent) PDF417 can be decoded by measuring elements from edge to similar edge (see 5.3.1).
- e) **Cross row scanning:** (inherent) The combination of the following three characteristics in PDF417 facilitates cross row scanning:
  - being synchronised horizontally, or self clocking
  - row identification
  - being synchronised vertically, by using the cluster values to achieve local row discrimination.

This combination allows a single linear scan to cross a number of rows and achieve a partial decode of the data so long as at least one complete symbol character per row is decoded into its codeword. The decoding algorithm can then place the individual codewords into a meaningful matrix.
- f) **Error correction:** (inherent) A user may define one of 9 error correction levels. All but Level 0 not only detect errors but also can correct erroneously decoded or missing codewords (see 5.7).
- g) **Compact PDF417:** (optional) In relatively 'clean' environments, it is possible to reduce some of the row overhead to improve the symbol density (see 5.12).

NOTE In earlier specifications of PDF417, Compact PDF417 was called Truncated PDF417. Compact PDF417 is the preferred term to avoid confusion with the more general use of the term 'truncated'.

## 5.2 Symbol structure

### 5.2.1 PDF417 symbol parameters

Each PDF417 symbol consists of a stack of vertically aligned rows with a minimum of 3 rows (maximum 90 rows). Each row shall include a minimum of 1 symbol character (maximum 30 symbol characters), excluding start, stop and row indicator columns. The symbol shall include a quiet zone on all four sides. Figure 1 illustrates a PDF417 symbol encoding the text: PDF417 Symbology Standard.

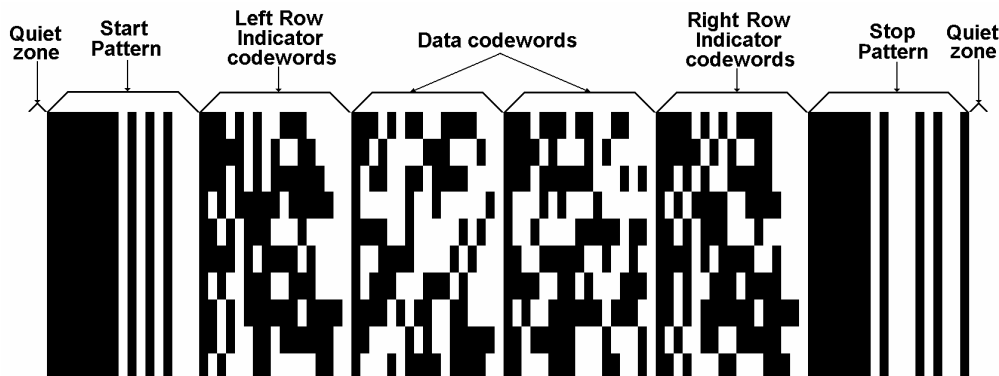


Figure 1 — PDF417 symbol structure

### 5.2.2 Row parameters

Each PDF417 row shall comprise:

- a) leading quiet zone
- b) start character
- c) left row indicator symbol character
- d) 1 to 30 symbol characters
- e) right row indicator symbol character
- f) stop character
- g) trailing quiet zone

NOTE The number of symbol characters (or codewords) defined in item 'd' above is equal to the number of data columns in the PDF417 symbol.

### 5.2.3 Codeword sequence

A PDF417 symbol may contain up to 928 symbol characters or codewords. Symbol character is the more appropriate term to refer to the printed bar/space pattern; codeword is more appropriate for the numeric value of the symbol character. The codewords shall follow this sequence:

- a) The first codeword, the Symbol Length Descriptor, shall always encode the total number of data codewords in the symbol, including the Symbol Length Descriptor itself, data codewords and pad codewords, but excluding the number of error correction codewords.
- b) The data codewords shall follow, from the most significant encodable character. Function codewords may be inserted to achieve data compaction.