Section 5.6:

Data Matrix Symbology

Table of Contents

5.6.1 Introduction....2

5.6.2 Data Matrix Features and Symbol Basics....3

5.6.3 Data Matrix Symbology....4

5.6.3.1 Square and Rectangular Formats....4

5.6.3.2 Data Matrix Symbol Sizes....5

5.6.3.3 Data Transmission and Symbology Identifier Prefixes....7

5.6.3.4 Width and Height of a Module (X)....7

5.6.3.5 Symbol Quality Grade....7

5.6.3.6 Advice for Selecting the Symbology....8

5.6.3.7 Human Readable Interpretation of Data Matrix Symbols....9

5.6.1Introduction

This section of the *GS1 General Specifications* addresses some of the technical aspects of the two-dimensional bar code symbology called Data Matrix. Data Matrix is a standalone, two-dimensional matrix symbology that is made up of square modules arranged within a perimeter finder pattern. Unlike a Composite Component® symbol (see Section 5.5), Data Matrix does not require a linear symbol. Data Matrix has been used in the public domain since 1994.

This section provides only a brief technical description and overview of the Data Matrix symbology. A more detailed technical specification can be found in the International Standard *ISO/IEC 16022 Information technology - International Symbology Specification - Data Matrix*. The GS1 System has adopted Data Matrix partly because it can encode GS1 System data structures and offers other technical advantages. Its compact design and the existence of various production methods that accommodate placing the symbology onto various substrates offer certain advantages over other symbologies currently in the GS1 System.

Data Matrix ISO version ECC 200 is the only version that supports GS1 System data structures, including Function Code 1. The ECC 200 version of Data Matrix uses Reed-Solomon error correction, and this feature helps correct for partially damaged symbols. In the remainder of this section, the ECC 200 version of Data Matrix is assumed when the symbology is described as Data Matrix or DM. This version of Data Matrix is similar in stability to ISO versions of current GS1 System symbologies.

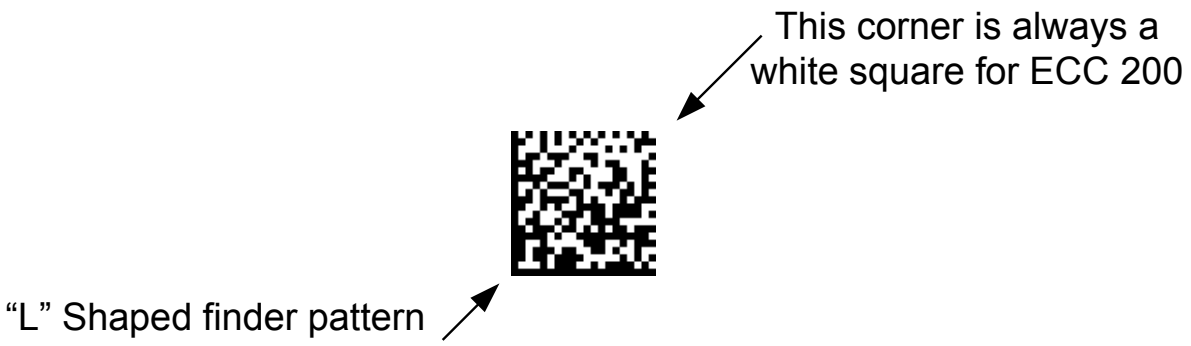
Implementation of Data Matrix shall be done per approved GS1 System application guidelines. This section will not describe the specific applications. The user needs to refer to specific application standards and guidelines in other sections of these *GS1 General Specifications* as they are approved for use. However, some of the production processes that can be used to produce Data Matrix Symbols are as follows:

- Direct part marking, such as is done by dot peening on items, such as automotive, aircraft metal parts, medical instruments, and surgical implants
- Laser or chemically etched parts with low contrast or light marked elements on a dark background (e.g., circuit boards and electronic components, medical instruments, and surgical implants)
- High-speed ink jet printed parts and components where the marked dots cannot form a scannable linear symbol
- Very small items that require a symbology with a square aspect ratio and/or cannot be marked within the allocated packaging space by existing Reduced Space Symbology® (RSS) and Composite Symbols

Data Matrix Symbols are read by two-dimensional imaging scanners or vision systems. Most other scanners that are not two-dimensional imagers cannot read Data Matrix. Data Matrix Symbols are restricted for use with new niche applications that will involve imaging scanners throughout the supply chain.

5.6.2Data Matrix Features and Symbol Basics

Figure 5.6.2 – 1: Data Matrix Symbol



- Figure 5.6.2 – 1 represents a Data Matrix Symbol with 20 rows and 20 columns (including the perimeter finder pattern but not including Quiet Zones).
- Data Matrix solid "L" shaped finder or alignment pattern is one module wide.
- Data Matrix Quiet Zone is one module wide on all four sides. As with other bar code Quiet Zones, do not print in this area.
- ECC 200 symbols can always be recognized from older versions of Data Matrix because the corner opposite the middle of the finder pattern is a zero module or white in normal print.
- For Square Data Matrix symbols, only an even number of rows and columns exist. Depending on data requirements, symbols can range from 10 row by 10 columns (10 x10) to 144x144 (including finder pattern but not the Quiet Zone)
- For normal printing, a module is one X by one X in dimension. Representation of data: A dark module is a binary one and a light module is a binary zero (or a light module is a binary one and a dark module is a binary zero for a symbol with reflectance reversal).
- ECC 200 (ECC = Error Checking and Correction) that uses Reed-Solomon error correction. Figure 5.6.3.2 – 1: ECC 200 Square Symbol Attributes shows the fixed amounts of error correction associated for each allowable Data Matrix Symbol size.
- FNC1 for GS1 System compatibility can be encoded at the beginning of the data string and as a group separator. When a FNC1 is used as a group separator, it shall be represented in the transmitted message by the ASCII character <GS> (ASCII value 29).
- Encodable character set:
  - Values 0 - 127 in accordance with ISO/IEC 646 International Reference Version (e.g., all 128 ASCII characters)
  - Values 128 - 255 in accordance with ISO/IEC 8859-1; Latin alphabet No. 1. These are referred to as extended ASCII.
  - The GS1 System requires that only the subset of ISO/IEC 646 International Reference Version defined in these *GS1 General Specifications* be used for Application Identifier (AI) Element Strings.

- Data characters per symbol (for the maximum symbol size):
  - Alphanumeric data: up to 2335 characters
  - Eight-bit byte data: 1556 characters
  - Numeric data: 3116 digits
- Large, square ECC symbols (at least 32 X32) will include alignment patterns to separate the data regions.
- Code type: matrix (Composite Component® is a stacked type)
- Orientation independence: Yes (requires a two-dimensional imaging scanner)
- Summary of additional features inherent or optional in Data Matrix:
  - Reflectance reversal: (Inherent) Symbols can be read when marked so that the image is either dark on light or light on dark.
  - Rectangular symbols: Six symbol formats are specified in a rectangular form.
  - Extended Channel Interpretation (ECI) capability allows Data Matrix to encode data from other alphabets.

5.6.3Data Matrix Symbology

The technical description of Data Matrix contained within this section provides additional information based on ISO technical specification 16022, and it is provided as a further aid in the development of specific applications. Data Matrix Symbols shown in the following subsections have been magnified to show detail; otherwise, they might not be seen clearly.

5.6.3.1Square and Rectangular Formats

Data Matrix may be printed in a square or rectangular format. The square format is usually used as it has a larger range of sizes and is the only format available for symbols encoding a large amount of data. The largest rectangular symbol can encode 98 digits, while the largest square symbol can encode 3,116 digits. An enlarged rectangular symbol and an equivalent square symbol are shown in Figure 5.6.3.1 – 1.

Figure 5.6.3.1 – 1: Rectangular and Square Data Matrix Symbols

(Specific applications are not used in the data encodation. Both symbols contain the same data)

5.6.3.2Data Matrix Symbol Sizes

Data Matrix Symbology has multiple sizes to match various data content (see Figure 5.6.3.2 – 1). Data Matrix Symbols have 24 sizes of the square format ranging from 10 by 10 modules up to 144 by 144 modules, not including the 1-X surrounding Quiet Zone. The rectangular format has 6 sizes from 8 by 18 modules up to 16 by 48 modules not including the 1-X surrounding Quiet Zone. Data Matrix sizes of 52 by 52 or larger have 2 to 10 interleaved blocks of Reed-Solomon error correction codewords.

The term “codeword” is used often to describe attributes concerning the encodation of data into Data Matrix Symbols. *ISO 16022* defines codeword as “A symbol character value. An intermediate level of coding between source data and the graphical encodation in the symbol.” Codewords are typically eight bits of data. FNC1, two numerics, and one alpha all take up one codeword each.

Figure 5.6.3.2 – 1: ECC 200 Square Symbol Attributes***

Symbol Size*		Data Region		Mapping	Total		Reed-Solomon		Inter-	Data Capacity			Error	Max. Correctable
				Matrix	Codewords		Block		leaved	Num.	Alphanum.	Byte	Correction	Codeword
Row	Col	Size	No.	Size	Data	Error	Data	Error	Blocks	Cap.	Cap.	Cap.	Overhead %	Error/Erasure
10	10	8x8	1	8x8	3	5	3	5	1	6	3	1	62.5	2/0
12	12	10x10	1	10x10	5	7	5	7	1	10	6	3	58.3	3/0
14	14	12x12	1	12x12	8	10	8	10	1	16	10	6	55.6	5/7
16	16	14x14	1	14x14	12	12	12	12	1	24	16	10	50	6/9
18	18	16x16	1	16x16	18	14	18	14	1	36	25	16	43.8	7/11
20	20	18x18	1	18x18	22	18	22	18	1	44	31	20	45	9/15
22	22	20x20	1	20x20	30	20	30	20	1	60	43	28	40	10/17
24	24	22x22	1	22x22	36	24	36	24	1	72	52	34	40	12/21
26	26	24x24	1	24x24	44	28	44	28	1	88	64	42	38.9	14/25
32	32	14x14	4	28x28	62	36	62	36	1	124	91	60	36.7	18/33
36	36	16x16	4	32x32	86	42	86	42	1	172	127	84	32.8	21/39
40	40	18x18	4	36x36	114	48	114	48	1	228	169	112	29.6	24/45
44	44	20x20	4	40x40	144	56	144	56	1	288	214	142	28	28/53
48	48	22x22	4	44x44	174	68	174	68	1	348	259	172	28.1	34/65
52	52	24x24	4	48x48	204	84	102	42	2	408	304	202	29.2	42/78
64	64	14x14	16	56x56	280	112	140	56	2	560	418	277	28.6	56/106
72	72	16x16	16	64x64	368	144	92	36	4	736	550	365	28.1	72/132
80	80	18x18	16	72x72	456	192	114	48	4	912	682	453	29.6	96/180
88	88	20x20	16	80x80	576	224	144	56	4	1152	862	573	28	112/212
96	96	22x22	16	88x88	696	272	174	68	4	1392	1042	693	28.1	136/260

104	104	24x24	16	96x96	816	336	136	56	6	1632	1222	813	29.2	168/318
120	120	18x18	36	108x108	1050	408	175	68	6	2100	1573	1047	28	204/390
132	132	20x20	36	120x120	1304	496	163	62	8	2608	1954	1301	27.6	248/472
144	144	22x22	36	132x132	1558	620	156	62	8**	3116	2335	1556	28.5	310/590
							155	62	2**					

Figure 5.6.3.2 – 2: ECC 200 Rectangular Symbol Attributes***

Symbol Size*		Data Region		Mapping	Total		Reed-Solomon		Inter-	Data Capacity			Error	Max. Correctable
				Matrix	Codewords		Block		leaved	Num.	Alphanum.	Byte	Correction	Codeword
Row	Col	Size	No.	Size	Blocks	Cap.	Cap.	Cap.	Blocks	Cap.	Cap.	Cap.	Overhead %	Error/Erasure
8	18	6x16	1	6x16	5	7	5	7	1	10	6	3	58.3	3/+
8	32	6x14	2	6x28	10	11	10	11	1	20	13	8	52.4	5/+
12	26	10x24	1	10x24	16	14	16	14	1	32	22	14	46.7	7/11
12	36	10x16	2	10x32	22	18	22	18	1	44	31	20	45.0	9/15
16	36	14x16	2	14x32	32	24	32	24	1	64	46	30	42.9	12/21
16	48	14x22	2	14x44	49	28	49	28	1	98	72	47	36.4	14/25

* Note: Symbol size does not include Quiet Zones.

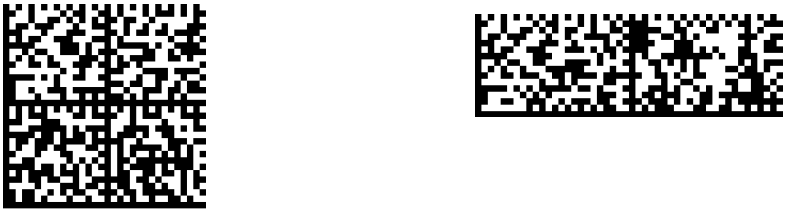
** Note: In the largest symbol (144x144), the first eight Reed-Solomon blocks shall be 218 codewords long encoding 156 data codewords. The last two blocks shall encode 217 codewords (155 data codewords). All the blocks have 62 error correction codewords.

*** Equivalent to Table 11 in the international standard *ISO-16022*, first edition, 2000-05-01.

The square format is divided into 4 to 36 data regions for symbols sized 32 by 32 modules and larger. The rectangular format symbols may also be divided into two data regions. Each data region is separated from the other regions by alignment patterns that consist of an alternating pattern of ones and zeroes and a solid line of ones (a dark line when there is no reflectance reversal). Figure 5.6.3.2-3 shows a four-segment square symbol on the left and a two-segment rectangular symbol on the right, each with hypothetical data shown to create the effect.

Figure 5.6.3.2 – 3: Segmented Data Matrix Symbols: Square and Rectangular Formats

(Sizes of these Data Matrix Symbols are larger than what would be used in a typical application so that typical alignment patterns can be easily seen.)



5.6.3.3Data Transmission and Symbology Identifier Prefixes

The GS1 System requires the use of Symbology Identifiers. Data Matrix uses the Symbology Identifier of "]d2" (see Figure 5.6.3.3 - 1) for GS1 System compliant symbols that have a leading FNC1 character. This indicates that Application Identifier (AI) data is encoded equivalent to the Symbology Identifier "]C1" for GS1-128 Symbols and “]e0” for Reduced Space Symbology® (RSS) and Composite Symbols. For more information on Symbology Identifiers, see the International standard *ISO/IEC 15424 Information technology — Automatic identification and data capture techniques — Data Carrier Identifiers.*

For example, a Data Matrix Symbol encoding AI (01) Element String 10012345678902 produces the transmitted data string “]d20110012345678902.” Data transmission follows the same principles that apply to the concatenation of AI Element Strings from GS1-128 Symbols or RSS Expanded®.

Figure 5.6.3.3 – 1: Symbology Identifier for Data Matrix ECC 200

	Message Content	Separator
]d2	Standard AI Element Strings	None

5.6.3.4Width and Height of a Module (X)

The range of the X-dimensions will be defined by the application specification, having due regard to the availability of equipment for the production and reading of symbols and complying with the general requirements of the application.

The X-dimension shall be constant throughout a given symbol. The X-dimension should apply to both the width and height of the modules.

5.6.3.5Symbol Quality Grade

The International Standard *ISO/IEC 15415 Information technology - Automatic identification and data capture techniques - Bar code symbol print quality test specification - Two-dimensional symbols* methodology shall be used for measuring and grading Data Matrix. The print quality grade is measured by verifiers that comply with the standard. The grade includes a grade level, measuring aperture, the wavelength of light used for the measurement, and the illumination angle relative to the symbol.

A symbol grade is only meaningful if it is reported in conjunction with the illumination and aperture used. It should be shown in the format *grade/aperture/light/angle*, where:

- "grade" is the overall symbol grade as defined in *ISO/IEC 15415 Information technology - Automatic identification and data capture techniques - Bar code symbol print quality test specification - Two-dimensional symbols* (e.g., the arithmetic mean to one decimal place of the Scan Reflectance Profile or scan grades). For Data Matrix, the grade number may be followed by an asterisk, *, which indicates that the surroundings of the symbol contain extremes of reflectance that may interfere with reading. For most applications, this should be specified as causing the symbol to fail.

- "aperture" is the diameter in thousandths of an inch (to the nearest thousandth) of the synthetic aperture defined in *ISO/IEC 15415 Information technology - Automatic identification and data capture techniques - Bar code symbol print quality test specification - Two-dimensional symbols*.

- "light" defines the illumination: A numeric value indicates the peak light wavelength in nanometres (for narrow band illumination); the alphabetic character W indicates that the symbol has been measured with broadband illumination (white light) the spectral response characteristics of which must imperatively be defined or have their source specification clearly referenced.

- "angle" is an additional parameter defining the angle of incidence (relative to the plane of the symbol) of the illumination. It shall be included in the reporting of the overall symbol grade when the angle of incidence is other than 45 degrees. Its absence indicates that the angle of incidence is 45 degrees.

Note: This international standard provides for 30 degrees and 90 degrees illumination in addition to the default 45 degrees.

The aperture is normally specified as being 80% of the minimum X-dimension allowed for the application. The printing method must produce the Data Matrix "L" pattern with gaps between the dots less than 25% of the specified aperture. If symbols with greater than the minimum X dimension are allowed by the application, the same absolute maximum gap dimension must be maintained.

Examples:

- 2,8/05/660 would indicate that the average of the grades of the Scan Reflectance Profiles, or of the scan grades, was 2,8 when these were obtained with the use of a 0,125 mm aperture (ref. no. 05) and a 660 nm light source, incident at 45 degrees.

- 2,8/10/W/30 would indicate the grade of a symbol intended to be read in broadband light, measured with light incident at 30 degrees and using a 0,250 mm aperture (ref. no. 10), but would need to be accompanied either by a reference to the application specification defining the reference spectral characteristics used for measurement or a definition of the spectral characteristics themselves.

- 2,8/10/670* would indicate the grade of a symbol measured using a 0,250 mm aperture (ref. no. 10), and a 670 nm light source, and indicates the presence of a potentially interfering extreme reflectance value in the surroundings of the symbol.

Recommended symbol grades for Data Matrix are identified in individual applications in Section 5.4.

5.6.3.6Advice for Selecting the Symbology

Any use of Data Matrix should comply with GS1 System global application guidelines and be restricted to those applications defined by the GS1 System for Data Matrix. Data Matrix will not replace other GS1 System Symbologies. Existing applications that are satisfactorily utilising EAN/UPC Symbols, ITF-14 Symbols, GS1-128 Symbols, Reduced Space Symbology® (RSS) Symbols, or Composite Symbols should continue to use them.

Note: Scanning systems that need to read Data Matrix Symbols must be 2D imaging scanners and be appropriately programmed to read the GS1 System version of Data Matrix or ECC 200.

5.6.3.7Human Readable Interpretation of Data Matrix Symbols

The Human Readable Interpretation of the primary Application Identifier (AI) Element String encoded in the Data Matrix Symbol should be shown with the symbol. How the human readable data will be shown shall be determined by the specific application guidelines. Typical conventions, as used for Reduced Space Symbology® (RSS) and Composite Component® Symbols, place the key information, such as the Global Trade Item Number™ (GTIN™), in the human readable data underneath the bar code symbol, while secondary information is placed above. The characters should be clearly legible (such as OCR-B) and must be obviously associated with the symbol.

AIs should be clearly recognisable to facilitate key entry. This is achieved by putting the AI between parentheses in the Human Readable Interpretation. Note: The parentheses are not part of the data and are not encoded in the bar code symbol, following the same principles that apply to GS1-128 Symbols and RSS Expanded® Symbols.

For Data Matrix Symbols encoding large amounts of data, it may not be practical to display all the data in Human Readable Interpretation form. Even if there is space to show it in this form, it may not be practical to key enter that much data. In these instances, some of the data may be omitted from the Human Readable Interpretation. However, primary identification data (GS1 System keys), such as the GTIN, must always be shown. Application specifications may provide additional guidance on Human Readable Interpretation.