



DPC: 24 / 30497453 DC

BSI Group Headquarters

389 Chiswick High Road London W4 4AL

Tel: +44 (0)20 8996 9000

Fax: +44 (0)20 8996 7400

www.bsigroup.com

Date: 25 July 2024
Origin: International**Latest date for receipt of comments: 14 August 2024**

Project No. 2023/01520

Responsible committee: IOT/1 Internet of Things and Digital Twin

Interested committees:

Title: BS ISO/IEC 30178 Internet of Things (IoT). Data format, value and coding

Please notify the secretary if you are aware of any keywords that might assist in classifying or identifying the standard or if the content of this standard

- i) has any issues related to 3rd party IPR, patent or copyright
- ii) affects other national standard(s)
- iii) requires additional national guidance or information

WARNING: THIS IS A DRAFT AND MUST NOT BE REGARDED OR USED AS A BRITISH STANDARD.
THIS DRAFT IS NOT CURRENT BEYOND 14 August 2024

This draft is issued to allow comments from interested parties; all comments will be given consideration prior to publication. No acknowledgement will normally be sent. **See overleaf for information on the submission of comments.**

No copying is allowed, in any form, without prior written permission from BSI except as permitted under the Copyright, Designs and Patent Act 1988 or for circulation within a nominating organization for briefing purposes. Electronic circulation is limited to dissemination by e-mail within such an organization by committee members.

Further copies of this draft may be purchased from BSI Shop <http://shop.bsigroup.com> or from BSI Customer Services, Tel: +44(0) 20 8996 9001 or email cservices@bsigroup.com. British, International and foreign standards are also available from BSI Customer Services.

Information on the co-operating organizations represented on the committees referenced above may be obtained from <http://standardsdevelopment.bsigroup.com>

Responsible Committee Secretary: **Ms Sophie Erskine (BSI)**
Direct tel: **0208 996 7254**
E-mail: sophie.erskine@bsigroup.com

Introduction

This draft standard is based on national and international discussions. Your comments on this draft are invited and will assist in the preparation of the consequent standard.

For international standards, comments will be reviewed by the relevant UK national committee before sending the consensus UK vote and comments to the international committee, which will then decide appropriate action. If the international standard is approved, it is usual for the text to be published as a British Standard.

For national standards, comments will be reviewed by the relevant UK national committee and the resulting standards published as a British Standard.

UK Vote

Please indicate whether you consider the UK should submit a negative (with supporting technical reasons) or positive vote on this draft. Please indicate if you are aware of any reason why this draft standard should not be published as a British Standard.

Submission of Comments

- **Annotated drafts are not acceptable and will be rejected.**
- All comments should be submitted online at <https://standardsdevelopment.bsigroup.com/>. You will need to register in order to comment.

Template for comments and secretariat observations

Date: xx/xx/20xx	Document: ISO/DIS xxxx
------------------	------------------------

1	2	(3)	4	5	(6)	7
MB	Clause No./ Subclause No./Annex (e.g. 3.1)	Paragraph/Figure/ Table/Note	Type of comment	Commend (justification for change) by the MB	Proposed change by the MB	Secretariat observations on each comment submitted
	3.1	Definition 1	ed	Definition is ambiguous and needs clarifying.	Amend to read '...so that the mains connector to which no connection...'	
	6.4	Paragraph 2	te	The use of the UV photometer as an alternative cannot be supported as serious problems have been encountered in its use in the UK.	Delete reference to UV photometer.	



JTC1-SC41/447/CD

COMMITTEE DRAFT (CD)

PROJECT NUMBER:

ISO/IEC 30178 ED1

DATE OF CIRCULATION:

2024-07-19

CLOSING DATE FOR COMMENTS:

2024-09-13

SUPERSEDES DOCUMENTS:

JTC1-SC41/392/CD, JTC1-SC41/414A/CC

IEC ISO/IEC JTC 1/SC 41 : INTERNET OF THINGS AND DIGITAL TWIN

SECRETARIAT:

Korea, Republic of

SECRETARY:

Ms Jooran Lee

OF INTEREST TO THE FOLLOWING COMMITTEES:

HORIZONTAL FUNCTION(S):

ASPECTS CONCERNED:

This document is still under study and subject to change. It should not be used for reference purposes.

Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Recipients of this document are invited to submit, with their comments, notification of any local regulations or technical reasons that may exist and should be considered should this proposal proceed, recognizing that failure to address such requirements could result in the need for new or existing "In Some Countries" clauses. (SEE AC/22/2007 OR NEW GUIDANCE DOC).

TITLE:

Internet of Things (IoT) - Data format, value and coding

NOTE FROM TC/SC OFFICERS:

As per SC 41 Helsinki Resolution #505, this document is circulated for 3rd CD consultation (8-week period).

CONTENTS

1		
2		
3	FOREWORD.....	5
4	Introduction.....	7
5	1 Scope.....	8
6	2 Normative references	8
7	3 Terms and definitions.....	8
8	4 Abbreviated terms	11
9	5 Background	12
10	5.1 Overview	12
11	5.1.1 General	12
12	5.1.2 Sensor-related data	13
13	5.1.3 Source data collection	14
14	5.1.4 Time-resolution and sample rate	14
15	5.1.5 Signal-processing implications.....	14
16	5.1.6 Low-power sensors.....	14
17	5.1.7 Measurement data	15
18	5.1.8 Numerical component	15
19	5.1.9 Semantic component	15
20	5.1.10 Unit	15
21	5.1.11 Prefix.....	16
22	5.2 Sensor value metadata	17
23	5.2.1 General	17
24	5.2.2 Resolution	18
25	5.2.3 Precision (measurement).....	18
26	5.2.4 Sampling time (time resolution)	18
27	5.2.5 Sample rate	19
28	5.2.6 Range.....	19
29	5.2.7 Accuracy	19
30	5.2.8 Physical quantity	19
31	5.2.9 Unit	19
32	5.2.10 Data type.....	19
33	5.2.11 Physical quantity	20
34	5.3 Physical quantities	20
35	5.3.1 Overview	20
36	5.3.2 Occurrences	20
37	5.4 Examples of data models	22
38	5.4.1 Overview	22
39	5.4.2 Exchange formats.....	22
40	5.4.3 File formats	22
41	5.4.4 Information modelling formats.....	22
42	5.5 Interoperability challenges	22
43	5.5.1 System integration.....	22
44	5.5.2 Maximizing the economic value of data through reusability	23
45	5.5.3 Sensor lifecycle cost.....	24
46	5.5.4 Sensor data capital value	25
47	5.5.5 Scalability.....	25

48	5.5.6	Semantics	28
49	5.5.7	Symbol mapping	29
50	5.5.8	Sensor Semantics.....	29
51	5.5.9	Representation vs. presentation	31
52	5.5.10	Human Readability	31
53	5.5.11	Resolving data points	32
54	5.5.12	Type Inference & Casting.....	34
55	5.5.13	Hard Coupling	34
56	5.5.14	Machine-to-machine Exchange	34
57	5.5.15	Representation	34
58	5.5.16	Presentation	34
59	5.5.17	Data translation	35
60	5.5.18	Errors and quality	37
61	6	Core profile	38
62	6.1	Overview	38
63	6.2	Sensor	38
64	6.2.1	Unit	38
65	6.2.2	The “unit-word”	38
66	6.2.3	Dimensionless Units	39
67	6.2.4	Range.....	39
68	6.3	Physical quantity	40
69	6.3.1	Overview	40
70	6.3.2	Operations.....	41
71	6.3.3	Significant Comparison	42
72	7	High-level system design	42
73	7.1	Overview	42
74	7.2	Type safety	44
75	7.3	Conversion mechanics	44
76	7.3.1	Overview	44
77	7.3.2	Single-step conversion	44
78	7.3.3	Unit resolver	44
79	7.3.4	Magnitude resolver	44
80	7.3.5	Precision estimation	45
81	7.3.6	Time-resolution and oversampling	45
82	7.4	Sanity-check mechanisms.....	45
83	7.5	Component manufactory	46
84	7.6	Digitized specification	46
85	7.6.1	Measurement range.....	46
86	7.6.2	Accuracy	46
87	7.6.3	Precision	46
88	Annex A	: Single-step Conversion	47
89	A.1	Overview	47
90	A.2	Example 1: Forward temperature conversion	47
91	A.3	Example 2: Forward energy conversion	48
92	A.4	List of unit-word value examples	50
93	A.5	Floating-point format.....	52
94	ANNEX B	Unifying Sensor Data and Mapping with Standards	53
95	RFC 8428:	Sensor Measurement Lists (SenML),	
96		https://datatracker.ietf.org/doc/rfc8428/	Error! Book mark not defined.

97	ISO/IEC 1539-1:202n Information technology — Programming languages — Fortran —	
98	Part 1: Base language	57
99	ISO/IEC 60559 Floating-point arithmetic	57
100	ISO/IEC CD 11404 - Information technology — General-Purpose Datatypes (GPD)	Error!
101	Bookmark not defined.	
102	SAREF-Compliant Knowledge Discovery for Semantic Energy and Grid Interoperability.	
103	Amelie Gyrard, Antonio Kung, Olivier Genest, Alain Moreau. IEEE World Forum on	
104	Internet of Things (WF-IoT 2021).....	57
105		
106		

INTERNATIONAL ELECTROTECHNICAL COMMISSION

Information Technology - Internet of Things - Data format, value and coding**FOREWORD**

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

152

Draft	Report on voting
XX/XX/FDIS	XX/XX/RVD

153

154 Full information on the voting for its approval can be found in the report on voting indicated in
155 the above table.

156 The language used for the development of this International Standard, is English

157 This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in
158 accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available
159 at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are
160 described in greater detail at www.iec.ch/standardsdev/publications.

161 The committee has decided that the contents of this document will remain unchanged until the
162 stability date indicated on the IEC website under "<http://webstore.iec.ch>" in the data related to
163 the specific document. At this date, the document will be

- 164 • reconfirmed,
- 165 • withdrawn,
- 166 • replaced by a revised edition, or
- 167 • amended.

168

169 The International Electrotechnical Commission (IEC) draws attention to the fact that it is claimed
170 that compliance with this document may involve the use of a patent. IEC takes no position
171 concerning the evidence, validity, and scope of this patent right.

172 The holder of this patent right has assured IEC that s/he is willing to negotiate licences under
173 reasonable and non-discriminatory terms and conditions with applicants throughout the world.
174 In this respect, the statement of the holder of this patent right is registered with IEC. Information
175 may be obtained from the patent database available at <http://patents.iec.ch>.

176 Attention is drawn to the possibility that some of the elements of this document may be the
177 subject of patent rights other than those in the patent database. IEC shall not be held
178 responsible for identifying any or all such patent rights.

179

Introduction

Digital transformation is an inherently data-driven process. Within the field of IoT, many actors in the forefront of technological development seek to model the physical world with the help of data.

Digital twins, city and building information models, process industry, automation, automotive and even healthcare, are all technologies or fields in which data models are developed to accommodate domain-specific needs.

Sensors form the backbone of live data generation, and help synchronize the digital world with the real world. Numerous data and information models occur in the IT/OT landscape, at various layers in digital systems. Some are verbose, text-based formats that come with more or less rich metadata, while some are more constrained.

The huge variability in how sensory data is represented poses a problem for all parties in the value chain – not least for the data producers themselves. Variability is often regarded as an interoperability problem but is, in essence, an economic problem since it raises the integration thresholds everywhere for the exchange of formatted data. This often puts the economic feasibility of data collection in jeopardy, since low data accessibility for data consumers also means poor return of economic value for the data producer.

This has often limited large-scale data collection to domains in which there is a strong need for very specific data collection (such as the automotive industry), and favoured applications where a single data producer has a direct business agreement with the data consumer, and full control over the data model. The design of the data gathering process has often been guided by a single customer's need, with a single benchmark for data quality.

Big Data still remains somewhat elusive for general-purpose data collection.

The variability in format and inconsistency in data quality are issues that undermine new stakeholders' trust in third-party value chains. Where data actually might have been accurate enough, an insufficient amount of consistent data-source documentation dissuaded any potential new data consumers – further limiting the capital value of data.

In the end, it has proven difficult for data producers to generate economically viable data – especially about physical environments.

This document outlines principles and technical solutions for overcoming this variability, and proposes a fundamental representation of data points for sensors and physical quantities that can be used everywhere in the digital landscape to improve the consistency of how both data and metadata are to be interpreted by application logic. When applied correctly, with an agnostic attitude towards application requirements, these principles also help form a baseline for data quality and accessibility.

With the perspective taken in this document, sensory data (after its collection) is regarded as a description of a physical entity. This also includes any other physical measurement, such as the physical parameters of real-world objects and systems. The intention is to allow for sensory data to be possible to integrate seamlessly with other data points in physical models under a shared, homogeneous interpretation.

INTERNET OF THINGS (IoT) - DATA FORMAT, VALUE AND CODING

1 Scope

This document defines common formats, values, and codings for data exchange among systems in the internet of things (IoT).

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 20924, Information technology – Internet of Things (IoT) - Definition and vocabulary

ISO/IEC 60559 Floating-point arithmetic¹

ISO/IEC Guide 99 International vocabulary of metrology - Basic and general concepts and associated terms (VIM)

BIPM, *The international System of Units (SI)*

Available at: <https://www.bipm.org/en/publications/si-brochure/>

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 20924, ISO/IEC Guide 99 and the following apply.

3.1

SI special names

Special names for the derived SI units

Example 1: Newton [N] equivalent to kilogram meter per second squared [kg·m/s²]

Example 2: Pascal [Pa] equivalent to newton per meter squared [N/m²] or kilogram per meter times second squared [kg/m·s²]

Example 3: Joule [J] equivalent to kilogram meter squared per second squared [kg·m²/s²]

Note 1 to entry: See BIPM The international System of Units (SI) 9th edition (2019), 2.3.4 for further information.

3.2

data type

named category of data characterized by a set of values, a syntax for denoting these values, and a set of operations that interpret and manipulate the values

Example 1: uint16 – A 16-bit wide, unsigned integer capable of representing the whole-number range 0 to 4294967295.

Example 2: bool – A boolean value with only two states (true and false)

Example 3: float – A 32-bit approximation of a real number

Note 1 to entry: Standardized representation of data types is essential for cross-platform interoperability.

¹ Equivalent with the IEEE-754 standard

Note 2 to entry: The total number of bits determine the maximum number of unique values the data type can encode.

Note 3 to entry: For further information on data types see ISO/IEC 11404.

[SOURCE: 1539-1:2018, 3.147, modified – “data type” made the preferred term and notes to entry and examples added.]

3.3 type safety

software design principle that protects against representation-related errors such as unintended rounding or corruption of data by enforcing the consistent use of a *data type* (**Error! Reference source not found.**) for some variable

Note 1 to entry: Type safety allows scalability by decoupling design of separate software components while guaranteeing correctness within functional interfaces (such as APIs).

Note 2 to entry: Type safety is enforced by some programming languages.

3.4 floating-point format

digital *data type* (**Error! Reference source not found.**) representing a fixed-precision approximation of a real number with a well-defined range and precision, composed of an exponential and fractional component, and a sign

Note 1 to entry: Floating-point format is supported by all major programming languages.

Note 2 to entry: Depending on the variant, floating-point format typically features at least 6-7 significant decimal digits for single precision, or 15 significant decimal digits for double precision.

Note 3 to entry: Definition based on ISO/IEC 60559:2020, 3.6.

3.5 floating-point representation

Unencoded member of a floating-point format, representing a finite number, a signed infinity, a quiet NaN, or a signaling NaN

Note 1 to entry: A representation of a finite number has three components: a sign, an exponent, and a significand; its numerical value is the signed product of its significand; where its numerical value is the signed product of its significand and its radix raised to the power of its exponent.

3.6 physical quantity

property of some physical entity that can be measured in some physical dimension, characterized by a number and a unit

3.7 serialization format

format of representing structured data in a serialized sequence of bits that can be exchanged between two computer systems

Example 1: XML – Extensible Markup Language, a text-based format using tags to encode nested elements and structures.

Example 2: JSON – JavaScript Object Notation, a format of representing basic values with named identities and nested structures, in cleartext. JSON is considered a less verbose alternative to XML.

Note 1 to entry: Serialization format is sometimes based on cleartext for representation, with a syntax to encode structure.

Note 2 to entry: Serialization format is sometimes based on concrete data types for representation.

309

310 **3.8**311 **sensor**

312 <internet of things> digital device that converts a physical measurement into a digital sample
313 carrying a quantity and a unit

314 Note 1 to entry: Usually transmits the sample coded in some *serialization format* (0).

315

316 **3.9**317 **data point**

318 datum that carries an individual meaning regardless of representation

319

320 **3.10**321 **runtime**

322 period of time during which an instance of some software or hardware, such as an IoT sensor,
323 operates after deployment

324 Note 1 to entry: Runtime is typically separate from, and non-overlapping with, design-time.

325 Note 2 to entry: Runtime constitutes part of that instance's operational lifecycle.

326

327 **3.11**328 **design-time**

329 period of time during which a piece of software or hardware is designed

330 Note 1 to entry: Design-time precedes, is typically separate from, and non-overlapping with, runtime.

331 Note 2 to entry: After the end of design-time, the design and its dependencies are not possible to modify.

332

333 **3.12**334 **oversampling**

335 collecting multiple measurement samples and averaging their values into a single sample

336 Note 1 to entry: Oversampling typically decreases temporal resolution,

337 Note 2 to entry: The number of samples converted into a single sample is called OSR, or "oversampling ratio"

338

339 **3.13**340 **discrimination (measurement)**

341 interval of a measured signal smaller than which a sensor is unable to uniquely register

342 Note 1 to entry: Is often expressed interchangeably with measurement resolution, which means segmenting a pre-
343 defined measurement range into discrete steps.

344

4 Abbreviated terms

ADC	Analog-to-Digital Converter
AI	Artificial Intelligence
API	Application Programming Interface
CPU	Central Processing Unit
DTw	Digital Twin
I2C	Inter-Integrated Circuit
I3C	Improved Inter-Integrated Circuit
IT	Information Technology
IoT	Internet of Things
JSON	JavaScript Object Notation
MEMS	Micro-electromechanical systems
OPC	Open Platform Communications
OPC-UA	OPC Unified Architecture
OSR	Oversampling Ratio
OT	Operational Technology
PCB	Printed Circuit Board
SI	SI-UNIT, <i>Le Système International d'Unités</i>
SPI	Serial Peripheral Interface
SenML	Sensor Measurement Lists
SiP	System-in-package
XML	Extensible Markup Language

5 Background

5.1 Overview

5.1.1 General

Today's increasingly digitized technology is founded on sensory measurement. Among the IT/OT environments underpinning an information-based society, there is a wide flora of data models that characterize the data shared between IoT-based systems.

While the sensors that operate out in the field measure physical quantities according to well-defined specifications, most often neither their measurements nor their specifications can be guaranteed to be available and complete within their digital context.

Sensors operating in the field also have a limited interoperability with most other sensors – either in other domains or other digital ecosystems. Sometimes this limited interoperability is due to an inherent lack of data overlap, or even a lack of data completeness.

The range of IoT sensors vary in many respects, and a core assumption that can be made is that, at runtime, no one sensor has complete compatibility with every other sensor. This is both due to the actual information content (i.e. meaning) of data, and a lack of consistent interpretation. When IoT sensors operate in separate ecosystems as individuals, they very often adhere to different conceptual domains or product families. It is therefore necessary to ensure that a baseline for interoperability can be created, and relied on, when embarking on system or data integration projects.

Reducing the engineering cost revolves to a significant degree around predictability, since being able to plan ahead allows system architects to reduce development and integration costs over long timespans.

Many IoT-based exchange formats define semantics and data types to cover the most essential primary data, and accompanying metadata to some degree. For sensory data, this often includes some representation of the sensory value (magnitude) itself and its unit of measurement - which can belong to any standard or non-standard choice of scale.

Variations in how data is encoded, structured, or serialized, brings with it a need for data to be translated or adapted (through interpretation) before it can be used by a third party. This is sometimes hampered by how data is recorded, which can limit how much useful information it can carry.

Common drawbacks of many clear textbased serialization formats include arbitrary rounding or truncation of measurement values (which can permanently strip away numerical precision), and a lack of unambiguous representation of measurement units.

Translating data between different formats therefore often requires ad-hoc integrations that perform conversion with or without safeguards for data integrity.

Variations to data representation in general can also constitute a challenge, since aggregating a heterogeneous set of data sources requires many separate conversions.

Taken together, it is therefore hard for IoT data models, which often are very well defined within their own domain (or scope of application), to make assumptions about what data points:

1. need to be provided to satisfy any future integration with a third-party data model,
2. will be available for import from third-party data sets,
3. meet the quality requirements for every specific implementation.

Essentially, the lack of universally available data points in an arbitrary sensor's data output is limiting its potential value to future applications. Consequently, any applications to which its data might be reusable will need to conduct an assessment of the data sources' suitability – often through human intervention.

This is a costly process, and scales in proportion to the number of integrations with that data source. For data with indefinite lifecycles (such as statistical, historic data about environments) that need to be kept for posterity, this is a cost that accumulates over time.

Reducing the cost of each integration beforehand is therefore in the interest of every involved party.

5.1.2 Sensor-related data

While IoT data models vary greatly in characteristics and scope depending on domain and application, every sensor with a digital interface that measures a physical quantity constitutes a data source with some universally applicable characteristics:

Sensor characteristic	Data analysis impact
Sensors measure a physical quantity, which by definition will include a numerical magnitude and a unit	Determines the dimension of measurement in which an observed variable can vary
Sensors have an inherent accuracy of measurement	Determines the minimum error of a sample's value from the actual value
Sensor measurement samples are recorded with some numerical precision	Determines the smallest possible variation of a sample's value
Sensor measurements are collected over some period of time	Determines the maximum time-resolution ² of the sample's value
Sensors output measurement samples at some maximum sample rate	Determines the time spacing between measurements
Sensors perform measurements within some maximum measurement range	Determines the range within which a sample's value can be considered valid

Table 1: Data source characteristics of a sensor

Each of these characteristics help implementers determine how well a sensor suits a given application and, by extension, also how well its data do. The requirements imposed by the application are carried over to the physical design of the system at design-time. For one-off implementations, the characteristics of the data are merely determined to be sufficient for its application, and often just implicitly guaranteed by the design of the system itself.

Data without an explicit quality and characteristics benchmark makes any quality guarantees implicit to the design of the system. This in turn makes a retrospective analysis of the system's quality metrics necessary to extract any quality guarantees downstream in the value-chain. Only by explicitly specifying quality and characteristics in digitally accessible metadata can it allow for future suitability analysis decoupled from the system's design.

To retain a high interoperability for a sensor and its data long-term, metadata need to be easily accessible for authorized evaluation when needed.

² When multiple samples are combined into an average value, the time resolution of the average is constructed from the sum of the time resolution of all the samples. In contrast, when the median value is extracted, it retains the time resolution of its original sample.

5.1.3 Source data collection

It is not unusual for IoT data to be collected by microcontroller devices – sometimes also with real-time computing requirements.

Measurements performed by peripheral circuits of such devices typically implement analog-to-digital conversion with raw measurement values expressed as an integer value, where the least significant bit (LSB) corresponds to a step-wise discrimination between segments of the measurement range, and the measurement resolution. This equates to an upper bound on the level of theoretical precision that is possible.

It is therefore important for device manufacturers to take such design constraints into account when determining the essential sensor characteristics.

5.1.4 Time-resolution and sample rate

Sample rate is often limited by the maximum time resolution of measurement – new samples cannot be produced faster than the time it takes to complete a sample.

A specification for time-resolution can also be unavailable for a given sensor; in which case it impacts the ability to correctly determine how its value varies over time.

Sensors that leverage oversampling and averaging to improve signal-to-noise ratio (and increase effective precision) will acquire a slower sample rate and lower time resolution.

5.1.5 Signal-processing implications

For sensors with a high amount of noise or interference, signal processing is sometimes required to clean unwanted errors from measurement data. It is therefore important for a consumer of sensor data to know if this is feasible.

For data from real-time embedded systems, where sensor samples can be collected at an interval on the order of micro-seconds, software-based noise filtering is sometimes necessary to improve measurement accuracy.

A hard limitation on the extent to which a signal can be accurately captured is formulated by the Nyquist sampling theorem³, which puts an emphasis on a sensor's sample rate and time resolution.

5.1.6 Low-power sensors

Low-power sensors, such as those utilizing LPWAN (Low Power Wide Area Network) technologies like Sigfox or LoRaWAN, are designed to conserve energy and extend battery life.

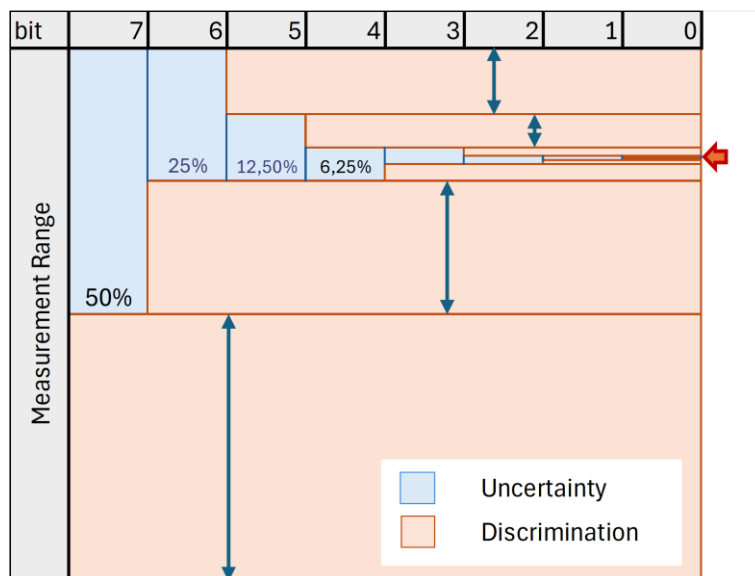


Figure 1: Discrimination and reduction of uncertainty of each bit within the range of an 8-bit ADC

³ Nyquist sampling theorem: Named after the Swedish-American engineer Harry Nyquist. According to the Nyquist sampling theorem, in order to accurately reconstruct a continuous-time signal from its discrete samples without any loss of information, the signal must be sampled at a rate that is equal to or greater than the Nyquist rate. In other words, the sampling rate must be at least twice the rate of the highest frequency component of the signal. If the signal is sampled below this rate, a phenomenon known as aliasing may occur, which leads to distortion and loss of information in the reconstructed signal. The Nyquist rate is a fundamental concept in signal processing and sampling theory. It is defined as twice the highest frequency component present in a given continuous-time signal.

<https://www.electropedia.org/iev/iev.nsf/display?openform&ievref=723-10-28>

This is achieved by incorporating a sleep mode in which the sensors remain inactive for extended periods of time.

During sleep mode, the sensors consume minimal power, and they wake up at predetermined intervals to take measurements and transmit data. This approach enables these devices to have a long operational life on a single battery charge, making them ideal for remote or hard-to-reach monitoring applications where frequent battery replacement is impractical.

The trade-off for this extended battery life is a relatively fine time resolution but a low sample rate. By waking up at longer intervals to create a sample, the sensors collect data less frequently than their high-power counterparts.

This is not necessarily a concern for applications where the monitored parameter changes slowly or does not require high-frequency measurements.

However, it is essential to consider the sensor's time-resolution and sample rate requirements when selecting low-power sensors for specific applications, as these characteristics can impact the ability to accurately capture and analyse rapidly changing signals or conditions.

5.1.7 Measurement data

As established in 5.1.1, important factors for the interoperability, reliability and reusability of measurement data can be summarized by avoiding loss of information, and reducing variability in how data is represented.

Apart from the obvious need for data to be sufficiently accurate, the top priority when making data as useful as possible is to make them easy to understand throughout their lifecycle. Well-documented data will remain useful indefinitely.

5.1.8 Numerical component

Sensors, as opposed to detectors (which count occurrences and discrete numbers, and can themselves incorporate sensors as part of their design), measure numerical quantities over a continuous range.

Across all domains in which sensors are used, digital measurement samples are represented in a wide variety of formats. As explained in 5.1.1, these can come in the form of human-readable text (which is typically presented in decimal notation), or in binary encodings (recognized in digital systems through data types).

By definition, binary computing machines require binary data types to be able to perform mathematical operations – this means that any textual representation (decimal or otherwise) first needs to be converted back into a binary data type before it can be processed as a numerical value. As explained in 5.1.1, this conversion can lead to loss of information if not made carefully.

5.1.9 Semantic component

Since sensors also vary greatly in their scale of measurement, and in what physical quantity is measured, this necessitates semantic qualifiers to properly characterize the numerical component. These most often map to a unit with an optional prefix.

5.1.10 Unit

The unit of measurement is the most essential characterization of which physical quantity (or dimension) is being measured. Without a unit of measurement, any numerical sample is meaningless.

Historically, units of measurement have been defined according to amounts with practical benefits to particular applications, depending on domain and on cultural heritage. Many with colloquial definitions attached to domain-specific phenomena or objects. A list of examples of historical units is shown in

Table 2.

Table 2: Examples of historical units**5.1.11 Prefix**

While range specifies the absolute bounds inside which the numerical component can vary, scale specifies a relative proportion to the unit of measurement, and is represented by a prefix (i.e. a symbol prepended to the unit's symbol to represent a multiple in the decimal system).

Unit	Dimension	Description/Origin
Day	Time	Time interval of a full rotation of the earth around its axis
Foot	Length	The length of a human foot, with definitions varying between cultures
Knot	Length/Time	Number of knots spaced ~14.4m apart passing through the hands of a sailor over a time period of 28 seconds
Stone	Mass	Stone weights crafted to multiples of the Roman pound [4]
Horsepower	Power	Power-output equivalence of steam engines to that of typical work horses

The 20 prefixes used in the metric system range from 10^{24} to 10^{-24} :

Name	Symbol	Base 10	Decimal
yotta	Y	10^{24}	1000000000000000000000000
zetta	Z	10^{21}	100000000000000000000000
exa	E	10^{18}	100000000000000000000000
peta	P	10^{15}	100000000000000000000000
tera	T	10^{12}	100000000000000000000000
giga	G	10^9	1000000000
mega	M	10^6	1000000
kilo	k	10^3	1000
hecto	h	10^2	100
deca	da	10^1	10
		10^0	1
deci	d	10^{-1}	0.1
centi	c	10^{-2}	0.01
milli	m	10^{-3}	0.001
micro	μ	10^{-6}	0.000001
nano	n	10^{-9}	0.000000001
pico	p	10^{-12}	0.000000000001
femto	f	10^{-15}	0.000000000000001
atto	a	10^{-18}	0.000000000000000001
zepto	z	10^{-21}	0.000000000000000000001
yocto	y	10^{-24}	0.000000000000000000000001

Table 3: Prefixes of the Metric System

When combined with a unit, a prefix allows for representing very large and very small values without an excess of non-significant digits.

⁴ [https://en.wikipedia.org/wiki/Stone_\(unit\)](https://en.wikipedia.org/wiki/Stone_(unit))