SmartCom

Digital System of Units

Digital Calibration Certificate

DCC – XML Schema

D-SI

EN

DOI: xxxx.yyyy.zzz

DOI: xxxx.yyyyy.zzz

Digital System of Units

D-SI

The meta data model for the exchange of

measurement results

Version 1.3

DOI: xxxx.yyyyy.zzz

Editors

Physikalisch Technische Bundesanstalt, Germany:
D. Hutzschenreuter, F. Härtig, W. Heeren, Th. Wiedenhöfer

National Physical Laboratory, United Kingdom:
C. Brown, A. Forbes, I. Smith, S. Rhodes

Czech Metrology Institute, Czech Republic:
I. Linkeova, J. Sykora, V. Zeleny

University Maribor, Slovenia:
B. Acko, R. Klobucar

Aalto University, Finland:
P. Nikkander, T. Elo, Z. Mustapää, K. Petri

Tallinn University of Technology, Estonia:
O. Maennel, K. Hovhannisyan

Ostfalia University, Germany:
B. Müller, L. Heindorf

University of Cassino and Southern Lazio, Italy:
V. Paciello

Comprising the results from our research and the fruitful and intensive discussions with all our other project partners and stakeholders worldwide.

Contact: smartcom@ptb.de

Braunschweig November 19

DOI: xxxx.yyyyy.zzz

Table of Contents

DOI: xxxx.yyyyy.zzz

1 Introduction and motivation

This document specifies the principles for the exchange of machine-readable data for all applications that transfer or require measurement data according to the specifications of the **Système International d'Unités (SI)** [1]. The document thus provides the basis for the harmonised, clear, secure and economical exchange of digital measurement values for a universe in which digital data is being transferred in accordance with the specifications introduced below.

For the digital exchange of metrological data, it is essential to associate at least each numerical value to a corresponding unit. These two pieces of information enable a statement to be made about the value of a quantity that can be interpreted according to the SI unit system. Because of its indivisibility and fundamental importance, this form of representation is called **atomic representation**. One example is:

1 kg

Here, "1" corresponds to the numerical value and "kg" to the assigned SI unit of kilogram. Together, both pieces of information characterise a mass quantity.

The complete indication of a measured quantity may contain additional information, for example, the specification of **measurement uncertainty** and a **time stamp**. Measurement uncertainty is information assigned to a measured quantity that provides an indication of its reliability. Usually this information is expressed by a **coverage interval** corresponding to a specified **coverage factor**. Conventions for the evaluation and expression of measurement uncertainty are laid down in the internationally recognised **Guide to the Expression of Uncertainty in Measurement (GUM)** [2]. A time stamp is required if a measured quantity or a constant is interpreted over a longer period, as some

DOI: xxxx.yyyyy.zzz

constants, for example, the Planck constant, have changed several times since the introduction of the SI.

In a digital network in which existing and new applications communicate with each other, even greater importance than before must be assigned to the SI. The ability of the SI to describe all physical processes using only seven base units leads to unprecedented clarity. This clarity is fundamental for the secure, harmonised and economical exchange of measurement values in the emerging digitalisation of metrology.

It is important to distinguish between **human-to-human** and **machine-to-machine** interfaces. The specifications presented here primarily relate to an automated communication and are essential for communication between machines and algorithms operating in an innovative digitised value chain. Stating measurement data only by using the 7 SI-base units will be the fundament for this development.

The change from units that are familiarly used by humans towards a communication only using SI-base units is supported by a **transition period**. In this period, it is permitted to use SI-derived units with an own symbol and other units that are stated in the SI-brochure from BIPM [1] like **Newton** or **Pascal**. It is expected that these units will have a less important role in future machine communication of metrological data, as each of these units can easily be expressed by a combination of SI-base units.

National legitimate units such as the Oechsle in Germany **or the imperial units** such as nautical miles, air miles, gallons, inches, etc., are also expected to continue to be used for a long time. Such insistence on the use of familiar systems often continues for generations until the advantages of alternative systems prevail. This behaviour also occurred when the base SI units of metre, kilogram and second were introduced in 1889. At that time, almost every town had its own units of length and weight and was reluctant to give up those particular units. However, with the spread of trade, the advantages of a uniform system of measurement became established.

DOI: xxxx.yyyyy.zzz

Against this background, the metadata model presented here also allows the use of familiar units. As a limitation, however, non-SI units can only be accepted if the measured quantities are specified in parallel using SI base units or if the non-SI units are accepted for use with the SI by the BIPM [1]. The following general conditions apply to the meta data model presented in this document:

- **Atom**: Atomic data are understood to contain at least one **number** (**numerical value)** and a corresponding **unit**. Optionally, **(measurement) uncertain**ty **information**, a **label** and a **time stamp** can also be specified.
- **Uniqueness**: Units are expressed in **SI base units**. This representation enables a clear, error-resistant and economical exchange of digital metrological data.
- **Numbers**: Numbers (numerical values) are displayed in the **decimal system**. Exponential representation is permitted. The separator for decimal places shall exclusively be the dot. Data such as NaN and INF are not permitted as they have no metrological significance.
- **Character representation:** Characters are represented in **UTF-8 format** [3], allowing all official languages to be mapped.
- **Accuracy**: When exchanging data, the accuracy of numbers is of paramount importance. Hence, some non-SI units for angle and time are permitted. Using angle as an example, circular closure can be expressed as both 360 degrees (°) and 2 pi radians. Radian is the appropriate form from both the metrological and the mathematical point of view. However, since pi is an irrational number with its representation limited to finite precision, greater accuracy can be obtained using degrees to represent circular closure.
- **Representation:** The representation of metrological data is based on existing internationally recognised documents. The most important of these are the **BIPM SI brochure BIPM** [1], the **Guide to the Expression of Uncertainty in Measurement**

(GUM) [2], the list of **fundamental physical constants** (**CODATA)** [4], the **International Vocabulary of Metrology** (**VIM)** [5] and the **ISO 80000 Part 1** [6].
- **Deviating units**: SI-derived units or units not listed in the SI brochure [1] can be specified in parallel to the SI units. In the event of contradictions, information provided using SI units shall take precedence.
- **Responsibility:** The responsibility, e.g. for the metrological content and its application, lies with the user of the meta data model.
- **Metadata model:** This document provides a **metadata model** for the exchange of measurement data which helps to establish the use of units in digitalised communication. The model comprises **identifiers** and **simple data types** for mandatory and optional components designed for a reliable exchange of metrological data. Reference implementations like in XML [7] or other formats, are not specified.

This document provides data models for the transfer of real quantities (including constants), complex quantities and vector quantities (lists of quantities). The fundamental and minimum required information that must be provided with each type of quantity are denoted as atomic types.

DOI: xxxx.yyyyy.zzz

2 Basic guidelines

The format described is essential for the exchange of metrological information in all fields of **Information and Communication Technologies (ICT)**. In terms of content, the format considers essential international metrology guidelines and standards.

The **metrology standards** considered for the definition of the exchange format are listed below in order of importance.

- **SI Brochure:** The SI Brochure [1], written by the **Bureau International des Poids et Mesures (BIPM)**, is an important guide on the use of **SI units**. It provides a compilation of quantities and units from the ISO 80000 Part 1 [6] that are recommended when using the SI. The SI Brochure also includes lists of atomic and natural constants.

- **GUM:** The **Guide to the Expression of Uncertainty in Measurement** (GUM) defines requirements to represent deviations of measured values from theoretically accurate measures [2, 8, 9]. Hereby the value of the **expanded measurement uncertainty** is elementary for real valued measurands [2, par. 6.2.1 and 5, def. 2.35]. This uncertainty indicates half the length of the **coverage interval** symmetrical to the measured value, which contains the scatter of the measured values with a defined coverage probability [8, par. 3.23 and 5, Def 2.36, Def 2.37]. The value of the uncertainty is calculated from a (combined) standard uncertainty of measurement by multiplication with a coverage factor [5, Def. 2.38]. In addition to the expanded measurement uncertainty, non-symmetrical coverage intervals are also possible.

- **CODATA:** The CODATA organization regularly publishes lists containing current values for **fundamental physical constants** [4]. These constants include the natural constants that will form the basis for the definition of all SI units in the future. For each real measurement result, the underlying CODATA values are those valid at the **time of the measurement**.

- **VIM:** The **International Vocabulary of Metrology (VIM)** defines the concept of quantities as a basic measurable property of objects, materials and substances [5, Def. 1.1]. Each quantity is expressed by a quantity value**.** A real measurement result is represented by a value**,** also called a **measured value** [5, Def. 2.10]. The measured value contains the indication of a **number** and a **unit of measurement** as reference [5, Def. 1.19]. The unit of measurement is also a real scalar quantity with the general abbreviation Unit [5, Def. 1.9].

- **ISO 80000 series:** The series of standards - in this document represented by ISO 80000 Part 1 [6] - defines requirements for the indication of quantities and units. The most important part is the specification of the **International System of Quantities (ISQ)** and the **International System of Units (SI)** assigned to it [6, Def. 3.6 and Def. 3.16].

The following example for the measurand temperature shows the application of the metrological specifications given above.

Example 1:

20.1(5) °C (k=2 and p=0.95)

*As required by **VIM**, the correct indication of the measured value is guaranteed by the number "20.1" and the unit of measurement "°C" (degrees Celsius). The symbolic unit "°C" is assigned to the temperature by **ISO 80000** and likewise the **SI brochure**. In addition, the accuracy of the measured value is indicated by the uncertainty " (5) " which gives the absolute uncertainty value "0.5 °C" according to the ISO 80000. As in the **GUM,** the uncertainty is specified as an expanded uncertainty corresponding to a coverage probability of "0.95" and calculated using a coverage factor of "2".*

The unit of measurement degrees Celsius is attributed to the SI base unit Kelvin, which is defined by the value of the Boltzmann constant from ***CODATA*** *today. Temperature measures before 20 May 2019 were made with the Kelvin defined by the triple point of isotopically pure water. Temperature measurements before and after the redefinition can only be compared correctly by conversion. Against this background, it will be important for all quantities to indicate measured values alongside the time of the measurement.*

In the next step, the measured value for the temperature is transferred from the human readable format to a machine-readable format. The next example shows how this could be done in the XML format.

Example 2:

```
<si:real>
  <si:label>temperature</si:label>
  <si:value>20.10</si:value>
  <si:unit>\degreecelsius</si:unit>
  <si:expandedUnc>
    <si:uncertainty>0.50</si:uncertainty>
    <si:coverageFactor>2</si:coverageFactor>
    <si:coverageProbability>0.95</si:coverageProbability>
    <si:distribution>normal</si:distribution>
  </si:expandedUnc>
</si:real>
```

Each value of the original temperature is encapsulated in a structural element. The designation of each structural element is unique and named according to the definitions of the metrological guidelines. The encapsulation of all structural elements in the superordinate element “si:real” indicates that it is a real measured value and with “si” indicating that all information comes from a controlled and agreed namespace based on the SI.

To make the structured data unmistakably interchangeable between machines, further ICT standards are used for the presentation of content, as described below.

- **IEEE 754:** The data types for the representation of numerical numbers such as the value of a quantity or components of its associated uncertainty shall at least be compatible with decimal floating-point numbers in the ANSI/IEEE 754 double precision format [10]. It includes a scientific statement of decimal exponents. By way of derogation from this standard, the indications NaN and INF are not allowed.

- **SI unit format:** As with numerical values, a reliable format is defined for the representation of units. This is founded on the units in the SI brochure of BIPM [1]. The underlying syntax for building units was partly took from siunitx [11]. Additional syntax rules were adapted from IEC TS 62720 [16].

- **ISO 8601:** Statements of the date and time when measurement data was recorded shall be made with reference to **Universal Coordinated Time (UTC)**. The format should comply with the ISO 8601 [12] format for legal local date and time with a difference to UTC (e.g. "2018-01-09T02:42:14.5+01:00").

- **Parsing (UTF-8):** Irrespective of whether saving them to a file or transferring them as a binary data packet between sensors, the character strings for all measured values and the surrounding structure require defined coding rules. The standard UTF-8 (Unicode Transfer Format 8-bit) is to be used here [3]. This format is provided by XML and numerous other software tools.

DOI: xxxx.yyyyy.zzz

3 SI unit format

Any format designed for the globally uniform mechanical exchange of metrological data is of added value if it is based on the International System of Units (SI system). The SI brochure of BIPM [1] provides the framework for the use of this system. The SI includes the use of SI base units and derived units as well as recognised units outside the SI that may be used with the SI system. Since the SI brochure is recognised worldwide, the definitions set therein should be the starting point of the unit format.

The technical implementation of the SI system in digital formats requires a unique definition of the semantics of units (identifier) and of the permitted syntax (combination of identifiers). This necessary machine-readable definition of a SI unit representation is adopted from a formal language for units that was used by Joseph Wright in the siunitx package [11]. The original work of Wright was reduced to the minimum set of syntax elements and unit identifiers that are necessary to represent the units in the 2019 SI brochure [1]. In addition, a comprehensive set of additional syntax rules from IEC TS 62720 [16] is recommended to be considered when building units in the machine-readable SI unit format.

Syntax for units

The syntax for specifying unit terms is shown in Figure 3.1.

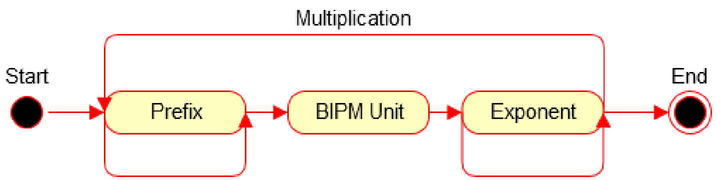

Figure 3.1 *Syntax for specifying SI unit terms*

Each unit term specifies at least one BIPM unit which can be any unit from the BIPM SI brochure with an own symbol. This unit can be extended by a prefix specifying multiples or submultiples of the unit. It is also permitted to form powers of the unit by appending an exponent. Furthermore, several units can algebraically be linked by multiplication.

The main syntax of BIPM unit, prefix, exponent and operator can be used to define different types of derived (coherent) unit terms. Exceptions to the general scheme can be noticed for the SI base unit for mass - the kilogram. This unit may not be combined with an additional prefix. Furthermore, the unit one for quantities with dimension number may not be combined with either a prefix nor an exponent.

In a practical application, the elements unit, prefix, exponent, and operator are specified by fixed strings and linked according to the syntax in Figure 3.1. The displayed format of the characters must be compatible with the UTF-8 encoding. Since the following definition of the semantics for the character strings is based exclusively on ASCII symbols, this is fulfilled by default for most available editors and development environments.

Semantics for units

The identifiers for BIPM units are separated into three groups. These are listed in Table 3.1. All identifiers are built by a combination of the textual unit name that is provided in the SI brochure [1] led by the backslash as starting character. Furthermore, the identifiers are written in lower case and blank spaces have been removed from the unit names.

The first group of identifiers comprises the seven physical SI base units for the measurands length (meter), mass (kilogram), time (second), thermodynamic temperature (Kelvin), current (ampere), amount of substance (mole) and luminous intensity (candela). By combining these base units, it is possible to express all physical units used today.

The second group of identifiers provides SI-derived units that have their own unit symbol (or likewise unit name) and that are of great historical importance. This comprises units like the Newton for force, the unit degree Celsius for temperature and the unit gram for mass.

The third group of BIPM units includes non-SI units that are accepted to be used with the SI and that have great historical importance. This includes units for practical time and angle measurements such as hours and degrees.

The unit one for quantities of dimension number is also provided as it is the integral element of the SI unit system. It is listed under the group of SI-derived units in this work.

Table 3.1 Groups of BIPM units and examples of identifiers in the SI unit format

Group of units	**Name**	**Symbol**	**Identifier**
SI base units[a]	metre kilogram	m kg	\metre \kilogram
SI derived units with their own symbol[b,c]	Newton degree Celsius one	N °C 1	\newton \degreecelsius \one
Units allowed for use with SI units[b,d]	hour ton electron volt	h t eV	\hour \tonne \electronvolt

a: Annex A, Table A.1; b: Annex A, Table A.2; c: Annex A, Table A.3; d: Annex A, Table A.4

Semantics of prefixes

The identifiers for prefixes provide all multiples and submultiples to the base 10 that are defined in the SI [1, Table 7]. Table A.5 in Appendix A lists the corresponding identifiers for prefixes in the SI unit format.

DOI: xxxx.yyyyy.zzz

Example 3:
The unit micrometre (symbolic µm) has the SI unit format representation

"\micro\metre"

with "\micro" *being clearly identifiable as a prefix and* "\metre" being the BIPM unit *as BIPM unit.*

Semantics of exponents
The power of a unit is indicated by the designation "\tothe{EXPONENT}". A decimal numerical value for the exponent is inserted in the field marked EXPONENT. The exponent value must be written according to the following properties:

- optionally one of the signs "+" or "-" followed by
- either an integer exponent value with the digits "0" to "9" (avoid leading zeros)
- or "0.5" for square roots
- Blank spaces are not allowed in the exponent.

Example 4:
The unit square meter for area (symbol m²) is written by using the exponent operator with power 2. The unit term in the SI unit format is,

"\metre\tothe{2}".

Semantics of unit multiplication
The multiplication of units is realized by appending BIPM unit terms to sequences of units. This shall be outlined by the following example.

Example 5:
The unit kilometre per hour (symbolic km/h) for speed is implemented by the unit term

DOI: xxxx.yyyyy.zzz

“\kilo\metre\hour\tothe{-1}”

where the unit “\kilo\metre” *is multiplied with the unit* “\hour\tothe{-1}”.

Table 3.2 lists additional examples of coherent derived units where multiple multiplications are made.

Table 3.2 Examples of derived units in the SI unit

Measured unit	Symbolic represen-tation	Representation in SI unit format
Radian	$\frac{\mathrm{m}}{\mathrm{m}}$	\metre\metre\tothe{-1}
Pascal	$\frac{\mathrm{kg}}{\mathrm{m\ s^2}}$	\kilogram\metre\tothe{-1}\second \tothe{-2}
Newton	$\frac{\mathrm{m\ kg}}{\mathrm{s^2}}$	\metre\kilogram\second\tothe{-2}
Ohm	$\frac{\mathrm{m^2kg}}{\mathrm{s^3A^2}}$	\metre\tothe{2}\kilogram\second \tothe{-3}\ampere\tothe{-2}

Further recommendations for syntax

The following list of constraints from the SI brochure [1] and IEC TS 62720 [16] for the combination of prefix, BIPM unit and exponent identifiers shall be considered when building unit terms.

- Each PIBM unit shall have only one prefix.
- Multiples and submultiples of the kilogram are built with the BIPM unit gram. The unit gram shall not be combined with the prefix kilo.
- The units kilogram, degree, degree Celsius, minute, hour, day, the reciprocal of the second, the reciprocal of the minute and the unit one shall not be combined with a prefix.

- If a BIPM unit has a prefix and an exponent, then the exponent is also applied to the prefix.
- The exponent is not allowed for the unit one.
- If a unit is combined by a multiplication of two or more units that can be interpreted as a division of units, then it can have either one prefix in the nominator, one prefix in the denominator or one prefix in each nominator and denominator.

DOI: xxxx.yyyyy.zzz

4 Meta data model real

Values of real quantities are the most fundamental data that are recorded and transmitted in practical applications. They play an eminent role in important areas like engineering, trade, the health system and calibration. Furthermore, real quantities are the basis on which complex quantities and vector quantities build upon.

This section gives the general definition of a real quantity in the D-SI meta data model. An **atomic quantity type** is presented that comprises the **minimum required information** for a real quantity value. Furthermore, an **extended quantity type** is introduced that adds information about the **uncertainty of measurement** to the atomic real quantity. The choice of the implementation of the data model, such as in XML, JSON or binary, is left to the user. A reference implementation of the D-SI data model of real quantities is provided for XML users by the EMPIR project 17IND02 SmartCom [18].

The following descriptions specify the atomic and extended real quantity types and their components in detail.

Atomic real quantity values

The atomic real measured value is the smallest entity to represent a measurement result. Its meta data model structure is shown in Figure 4.1.

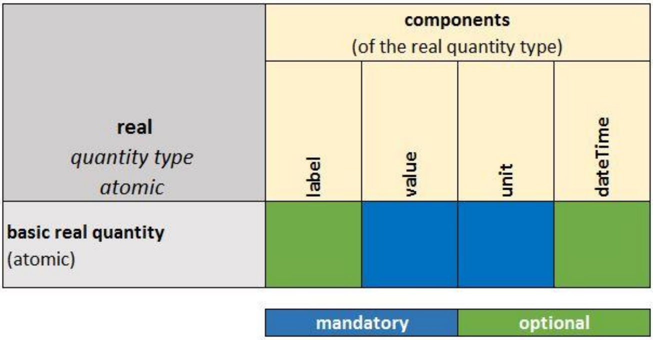

Figure 4.1 *Meta data model for the atomic specification of real quantity values*

The real quantity value must consist at least of a numerical value and the assigned unit. The following information is permitted:

- **Component "value":** The numerical value of the quantity must be provided. It must be written as a decimal floating-point number, which consists of a sign, the mantissa, an exponent separator, and a value for the exponent. Both the sign and the parts of the exponent are optional. The signs "+" and "-" are allowed. The mantissa comprises the digits "0" to "9". It may also contain the dot "." as a decimal separator. The dot can be anywhere before, after or within the mantissa. The exponent separator is "e" or "E". The exponent value is an integer with the digits "0" to "9". It can optionally have the sign "+" or "-". The value must not contain blank spaces and it is recommended to write the mantissa and the exponent without leading zeros.

  Example 6:
  The data type for numerical values permits for example
  - *decimal floating-point numbers like "1.0" or "-0.0045",*
  - *decimal numbers with scientific notation like "1.3e-2" or "2.E-1" and*
  - *integer numbers like "15" or "-0".*

- **Component "unit":** The unit must be a UTF-8 compatible string. A high level of machine readability of the unit is gained when using only SI-base units and combinations of SI-base units. A SI unit format that is fit for this purpose is presented in Section 3.

- **Component "label":** A label can optionally be assigned to the metrological data of the atomic quantity. Labels can provide information on the kind of measurement and/or the name of the quantity. As with the unit, it must be a UTF-8 compatible string.

- **Component "dateTime":** This element is used to assign a time stamp for the measurement to the quantity value. All time stamps must be recorded in legal local time with a difference (offset) to the Universal Coordinated World Time (UTC). The extended UTC format from ISO 8601 2004 [12, section 4.3.2.a] is used for implementation.

  Example 7:
  The extended ISO 8601 format consists of a date followed by a time followed by an offset of the time-zone to UTC time as shown in the example below.

  "2018-09-05T16:15:03.09-00:01"

  The date "2018-10-17" includes the year, the month and the day. Date and time are separated by the character "T". The time "16:15:23.09" consists of hours, minutes and seconds. The seconds can be stated as a two-digit integer value or with additional decimal places like "23.09". The number of decimal places is not limited. The offset to UTC time "-00:01" consists of a positive or negative sign which is followed by hours and minutes. The offset of the example is a time of minus one minute. Statements of date and time in UTC are indicated by the symbol "Z" instead of an offset "+00:00" (i.e. "2018-09-05T16:15:03Z").

DOI: xxxx.yyyyy.zzz

Real quantity value with measurement uncertainties

The inclusion of measurement uncertainties to the atomic real metadata type leads to extended real measurement value formats in Figure 4.2.

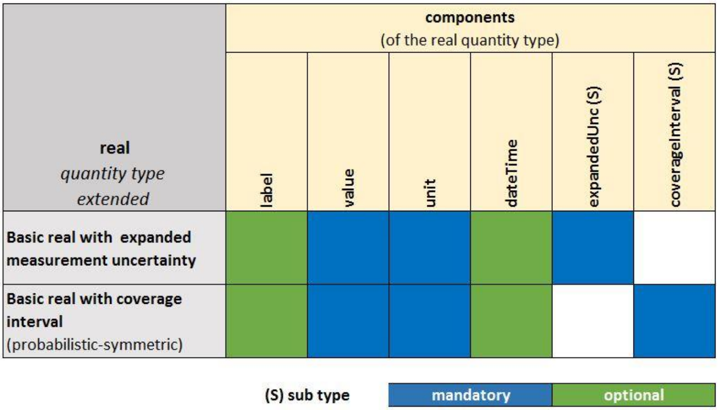

Figure 4.2 *Meta data model for the extension of the atomic quantity value type with measurement uncertainties*

The definitions of value, unit, label, and dateTime remain identical to the atomic type. The uncertainty of the numerical value of the real quantity can be chosen from two sub types: The first sub type is “expandedUnc”. It provides a data model for the **expanded measurement uncertainty**. The second sub type is “coverageInterval” which provides the data of a **probabilistic symmetric coverage interval**.

The data models of the uncertainty sub types are defined in more detail below.

DOI: xxxx.yyyyy.zzz

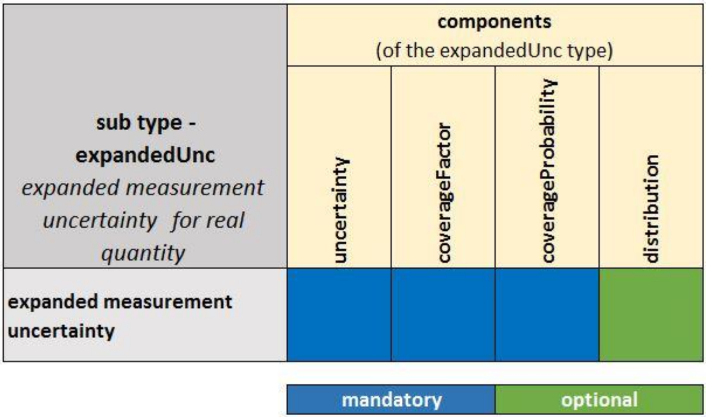

Figure 4.3 *Meta data model of the expanded measurement uncertainty for real quantities*

The **expanded measurement uncertainty** specifies an interval that is symmetric to the underlying real measurement value and that covers a certain percentage of the range in which the measured value is expected to vary. The D-SI data model for this type of uncertainty is shown in Figure 4.3. It requires to provide the value of the uncertainty, a coverage factor and the coverage probability. This data is characteristic for the expression of expanded measurement uncertainty in practical applications. In addition, the distribution of the quantity can be specified. The components of the expanded measurement uncertainty are defined as follows in the D-SI data model:

- **Component "uncertainty":** The value of "uncertainty" provides the half width of the coverage interval that is defined by the expanded measurement uncertainty. The coverage interval shall cover the part of the distribution of the underlying real measurement value according to the coverage probability value from the component "coverageProbability". The unit of the "uncertainty" value is the same unit as defined for the underlying real quantity in component "unit" (see Figure 4.2). The value type of "uncertainty" is a positive

DOI: xxxx.yyyyy.zzz

decimal number. Its definition is a restricted variant of the decimal number format of the component “value” that prohibits the sign minus (“-”) for the mantissa.

- **Component “coverageFactor”:** The expanded measurement uncertainty (component “uncertainty”) is calculated from the standard uncertainty of the underlying real measurement by multiplication with an expansion factor. The value of the expansion factor is provided by the component “coverageFactor”. The data format for the coverage factor is the decimal format of the component “value”. Values smaller than 1 are not permitted.
  Note: The value of the standard uncertainty is accessed by dividing “uncertainty” by “coverageFactor”.

- **Component “coverageProbability”:** This component specifies the coverage probability with which the uncertainty covers the distribution of the measurand. It is a decimal value based on the component “value”. Values less than 0 or greater than 1 are not permitted.

- **Component “distribution”:** If further information on the type of the distribution of the measurand is available, this information can be specified with the optional component “distribution”. This component must be written as a UTF-8 encoded string. If the “distribution” component is not provided, then, as a default, real quantities shall be interpreted as random variable with a normal distribution (commonly called Gaussian distribution).

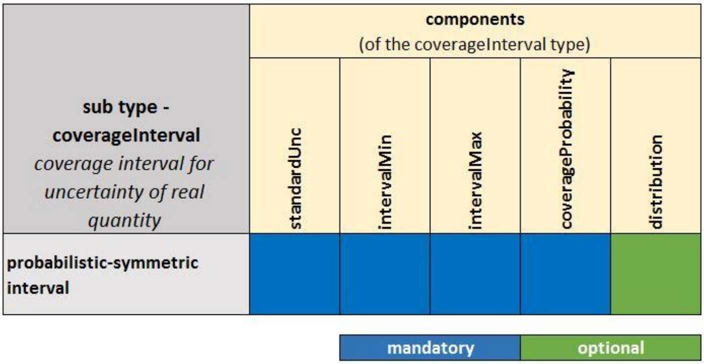

Figure 4.4 *Meta data model for the expression of uncertainty of a real quantity by a probabilistic-symmetric coverage interval*

The **probabilistic-symmetric coverage interval** defines the uncertainty of a real measurement value symmetric to the probability density function of the random distribution of the measurand. The D-SI data model for this type of uncertainty is shown in Figure 4.4. It requires to provide values for the bounds of the associated coverage interval which is not necessarily symmetric to the real measurement value. Furthermore, a coverage probability value and the standard uncertainty value of the measurement must be provided. In addition, the distribution of the quantity can be specified. The components of the probabilistic-symmetric coverage interval are defined as follows in the D-SI data model:

- **Component "standardUnc":** This component provides the standard uncertainty value of the underlying real measurand. Its value is assigned with the unit (component "unit") of the measured value (see Figure 4.2) The value type of "standardUnc" is a positive decimal number. Its definition is a restricted variant of the decimal number format of the component "value" that prohibits the sign minus ("-") for the mantissa.

- **Component "intervalMin":** This component specifies the value of the lower limit of a coverage interval for the measurement uncertainty. The value of the lower limit always refers to the unit defined under component "unit" of the underlying real quantity (see Figure 4.2). Its specification together with the component "intervallMax" is mandatory. It must be smaller than the value of the interval upper limit "intervalMax". The interval defined by the lower and upper limits of the interval contains values of the measurand probability "coverageProbability".
  The format for the component is a decimal value of the same type as the component "value".

- **Component "intervalMax":** This component specifies the value of the upper limit of a coverage interval for the measurement uncertainty. The value of the upper limit always refers to the unit defined under "unit" of the underlying real quantity (see Figure 4.2). Its specification is mandatory together with the component "intervallMin". It must be greater than the value of the lower interval limit "intervalMin". The interval defined by the lower and upper limits of the interval contains values of the measurand probability "coverageProbability".
  The format for the component is a decimal value of the same type as the component "value".

- **Component "coverageProbability":** This component specifies the size of the area of the underlying probability density function that is covered by the interval boundaries "intervalMin" and "intervalMax". It is a decimal value based on the component "value". Values less than 0 or greater than 1 are not permitted.

- **Component "distribution":** The "distribution" component has the same definition as in the case of the expanded measurement uncertainty (see Figure 4.3).

DOI: xxxx.yyyyy.zzz

5 Meta data model constants

Fundamental physical constants from the CODATA listing [4] and mathematical constants such as pi have always played a significant role for many measurements. The usage of constants became even more substantial, as the whole SI system of units was redefined on the pure basis of natural constants which went into force on 20 May 2019 [1]. Therefore, the D-SI data model provides a data element for the exchange of quantities representing constants. This data element is shown in Figure 5.1.

The **"constant" quantity** data model has the same basic structure as the atomic **"real"** quantity. This comprises the mandatory statement of a **"value" component together with the associated "unit" component**.

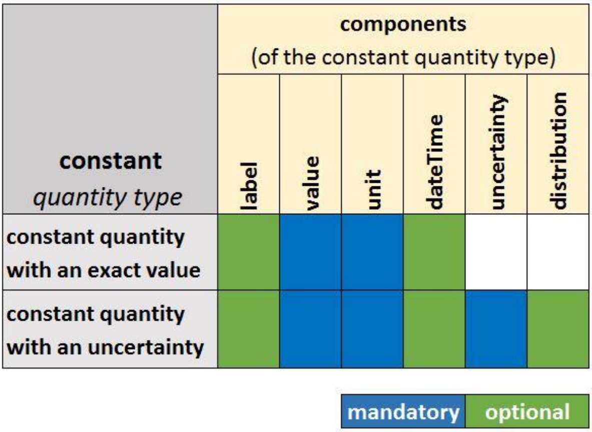

Figure 5.1 *Metadata model for quantities representing fundamental physical constants and mathematical constants.*

The **"label"** component is used to identify the constant. The **"dataTime"** component is of great importance for the definition of fundamental physical constants. The latest values of these fundamental constants are regularly published by CODATA. The

DOI: xxxx.yyyyy.zzz

“dateTime” element allows to identify to which release of CODATA the constant is associated.

The constants are distinguished into those constants with an exact numerical value and constants with an uncertainty attributed to the numerical value. The uncertainty value is implemented as component **“uncertainty”** in the constant data model. Its value refers to the unit element that is also valid for the value element.

The value of the **“uncertainty”** component shall provide the **standard deviation for values from CODATA** [4]. In addition, the distribution of the uncertainty, if known, can be provided in the component **“distribution”**

Mathematical constants play a special role. Numbers like pi or the Euler number are irrational numbers. They require rounding to a finite number of digits for a machine representation. The rounding leads to a rounding error that shall be expressed by a rectangular distribution that contains the true value of the constant with 100% probability. The rectangular distribution shall be centred at the numerical value that is provided by the element value. The component **“uncertainty”** shall give the **standard uncertainty value of the rectangular distribution**.

The statement of the rounding error by means of a statement of uncertainty has far-reaching consequences. It enables machines to understand the rounding error within the universal framework of uncertainty that is established and proven technology in measurement science since many decades. A bridge is fostered between computer science and metrology for a lasting support of the reliable exchange of data between different hardware and software applications. The concept is not only limited to the exchange of fundamental constants in the D-SI data model, but it can be enhanced to real and complex quantities as well.

Example 8:
The value of the constant pi shall be provided with eight digits of precision in the constant format. This example shows one way how to provide the necessary data in five steps.

<u>Step 1:</u> The label components is assigned with the constant name "pi".

<u>Step 2:</u> The value component is assigned with the value of pi rounded to eight decimal places by application of the rounding rule "half-even" (also known as "Bankers rounding"). The rounded value is "3.14159265".

<u>Step 3:</u> The unit component is assigned with the unit "\one" expressing that the value of pi can be interpreted as a quantity of dimension number.

<u>Step 4:</u> The uncertainty of the rounded value of pi is calculated by means of a rectangular distribution. Starting point is the rounding error which is $e = 5 \cdot 10^{-9}$*. The true value of pi lies within the interval* $3.14159265 \pm e$*. The rectangular distribution is defined on top of this interval by calculation of its standard uncertainty* $u = e/\sqrt{3}$*. The value of* u *is assigned to the uncertainty component of the data model for constants.*

<u>Step 5:</u> The distribution component is defined with the text "rectangular" indicating an uncertainty for a rectangular distributed quantity.

DOI: xxxx.yyyyy.zzz

6 Metadata model complex and list

This version of the D-SI documentation provides the basic concepts behind the meta data model for vector quantities. Thereby, the term vector quantity refers to complex quantities (vector of dimension two), vectors of real quantities and vectors of complex quantities in our context.

The meta data model includes an atomic type definition of a complex quantity value which can choose between a Cartesian coordinate form and a polar coordinate form as specified in GUM supplement 2 [9].

list *of quantities type* atomic	**components** (of the list type)				
	label	real (S)	complex (S)	list (S)	dateTime
basic list of real (atomic)		n times			
basic list of complex (atomic)			n times		
recursive list of lists				n times	

(S) sub type | mandatory | optional

Figure 6.1 *D-SI meta data model of the atomic list for the representation of data from vector quantities*

The vectors of real quantities and complex quantities are also formally defined in the GUM supplement 2 [9]. The D-SI meta data model provides these quantities by means of a generic list structure. The atomic data model of a list can either comprise a set of real quantities, a set of complex quantities or a set of lists (see Figure 6.1). The latter allows to model individual structures of data, such as a matrix, through nesting of lists.

The atomic types for complex quantities and lists can be extended by a hyper-ellipsoidal coverage region and a hyper-rectangular coverage region for providing information on multivariate uncertainty. The required structures are also defined in GUM supplement 2 [18] which was transferred to the D-SI. The central element of a multivariate uncertainty statement is a covariance matrix. It quantifies the stochastic dependency between the components of the underlying vector quantity.

An additional feature of the extended data model type for lists is the availability of structures to avoid an overhead of redundant data. In the case of a list of real quantities, it is possible to specify a list unit that is the designated unit for all quantities in the list. A single univariate uncertainty can be designated to all real quantities in the list as well. In a similar way, superior units and bivariate uncertainty statements can be provided for lists of complex quantities.

Please, refer to the XML reference implementation of the D-SI meta data model [18] for more detailed information about the structure and data types of the data models and components that were mentioned above.

7 Non-SI unit quantities (hybrid)

The use of non-SI units increases the possibility of confusion and errors that may have serious consequences. For example, in 1999 the NASA Mars Climate Orbiter was lost when software generated force values in the non-SI unit of pound second rather than the SI unit of Newton.

In [1] it is recommended that when a non-SI unit is used, the corresponding definition of the unit in terms of SI units is provided. For many applications, it is seen as good practice for values of quantities reported using non-SI units to be accompanied by corresponding values in SI units.

Recommended units in D-SI data Model

The D-SI data model for the exchange of metrological data recommends that the units in the SI brochure [1] be used. As defined by the International Committee for Weights and Measures (CIPM), the base quantities in the SI are length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity. The corresponding **base units** of the SI are the metre, the kilogram, the second, the ampere, the kelvin, the mole, and the candela [1, table 2].
Derived SI units are units that are expressed algebraically in terms of powers of the base units or other derived units [1, tables 4-6]. Furthermore, **SI prefixes** can be combined with SI base units and derived SI units to build multiples and submultiples of the units [1, table 7].

Units that are neither base units nor derived SI units (including the SI prefixes) are referred to as **non-SI units**.

The CIPM recognises that some non-SI units are widely used and will continue to be used for many years. The list of commonly-used non-SI units comprises the minute, the hour and the day

DOI: xxxx.yyyyy.zzz

(measures of time), the astronomical unit (length), the degree, the minute and the second (angle), hectare (area), the litre (volume), the tonne and the dalton (mass), the electronvolt (energy) and the neper, the bel and the decibel (logarithmic ratio quantity) [1, table 8].

The use of units that are expressed only as an algebraical combination of SI base units should be given priority to any other unit allowed in the SI and non-SI units. A unit that only requires to interpret SI base units has the greatest importance for a highly machine-readable data representation. Quantities with such units can easily be compared and interoperation is possible without conversion of data. Such comparison and interoperation of data requires much more effort for SI derived units with their own symbols, units with a prefix or the non-SI units. These units must be interpreted by means of their underlying definition in SI base units at first. Then conversion may be needed for interoperation of quantity values.

Unrecommended units in D-SI data model

Units that are not listed in the SI brochure [1] are not recommended for a machine-readable data exchange. Some of these unrecommended non-SI units are used mainly by particular groups, often to meet specialised requirements. For example, in civil aviation, the nautical mile (a measure of length) and the knot (speed) are widely used. Although it is intended that the use of these non-SI units in this area will eventually be discontinued, no "termination date" has yet been established [14, table 4-1]. In addition, the list of other unrecommended units includes the bar and millimetre of mercury (measures of pressure), the ångström (length), the barn (area) and the CGS (centimetre-gram-second) system of units

Further to the non-SI units listed in [1], there are numerous non-SI units in use, in some cases to meet specialised requirements, or in specific countries. A list of factors to allow

DOI: xxxx.yyyyy.zzz

conversion between such units and SI units can be found in the Guide for the Use of the International System of Units (SI) [13].

A comprehensive listing of other non-SI units is also provided by the so-called ISO/IEC 80000 group of standards [6]. It consists of 14 parts that define the international System of Quantities (ISQ) and the associated SI and non-SI units. Quantities and units for computer science applications are a special feature [15]. The definition of these units comprises counted quantities and binary prefixes outside the SI.

An exception to this recommendation is recognized for internationally accepted systems of units and scales in the area of reference materials and in the area of reference procedures. The discussion on how to treat such units and their inclusion in the D-SI data model like SI units is subject to a future revision. For now, the units for reference materials and reference procedures will be handled in the same way as unrecommended non-SI units.

The hybrid data model for non-SI units

While the recommended units can be directly used as reference for real and complex quantities in the D-SI data model, it is not allowed for the unrecommended units.

An adapter is offered for those quantities with an unrecommended unit in order to integrate them into the machine-readable D-SI data model. This adapter is denoted as the hybrid data model or in short “hybrid”.

The application of the hybrid data model requires a conversion of the quantity with the non-SI unit into a quantity with an adequate SI unit. Both quantities are then put together into one data element – the hybrid element. The hybrid data element must contain at least one quantity with an SI unit. The number of additional quantities with other units can be one or more.

DOI: xxxx.yyyyy.zzz

While machines will need and use the SI components of hybrid for a safe operation with the data, humans can still refer to customary non-SI.

A real quantity in hybrid comprises one real component that must state the quantity value in an SI-base unit. Furthermore, it can provide additional real quantities with SI derived units or non-SI units that convert to the real quantity with the SI-base unit. Figure 7.1 gives an example for such a hybrid element by means of XML structured data.

```
<si:hybrid>
   <!-- A: length from B converted to SI -->
   <si:real>
      <si:value>0.3048006</si:value>
      <si:unit>\metre</si:unit>
   </si:real>
   <!-- B: length with imperial unit foot -->
   <si:real>
      <si:value>1</si:value>
      <si:unit>ft(U.S. survey)</si:unit>
   </si:real>
</si:hybrid>
```

Figure 7.1 *XML example for the hybrid adapter for real quantities with the unrecommended non-SI unit foot (ft)*

For all quantities in a hybrid element, the model of measurement uncertainty must be identical. In the case of a real quantity it is not allowed to mix the expanded measurement uncertainty statement with the statement of a probabilistic symmetric coverage interval. In the case of complex quantities, the Cartesian coordinate form must not be mixed with the polar coordinate form. The mixture of coverage region types is also forbidden for complex quantities and lists of quantities.

Figure 7.2 shows a hybrid element that contains two lists. The list marked with the letter A provides quantities with SI units. List B contains quantities with non-SI units. List A was derived from list B by a conversion of the non-SI quantities to SI quantities. Hence,

DOI: xxxx.yyyyy.zzz

both lists must have the same amount of quantities. The upper real quantities in list A are the SI representation of the upper real quantity with the non-SI unit gallon in list B. The real quantity values at the lower position in lists A and B are respectively the second pair of equivalent quantities with a SI unit and a non-SI unit. Both quantities state an expanded measurement uncertainty which must be converted from list B to list A too. Thereby, the type of the uncertainty statement must not change but the numerical value of the uncertainty has to be converted.

```xml
<si:hybrid>
    <!-- A: list from B converted to SI -->
    <si:list>
        <si:real>
            <si:value>0.00454609</si:value>
            <si:unit>\metre\tothe{3}</si:unit>
        </si:real>
        <si:real>
            <si:label>hardness Rockwell C scale</si:label>
            <si:value>63.00</si:value>
            <si:unit>\metre\metre\tothe{-1}</si:unit>
            <si:expandedUnc>
                <si:uncertainty>1.56</si:uncertainty>
                <si:coverageFactor>2</si:coverageFactor>
                <si:coverageProbability>0.95</si:coverageProbability>
            </si:expandedUnc>
        </si:real>
    </si:list>
    <!-- B: list with non-SI units -->
    <si:list>
        <si:real>
            <si:value>1</si:value>
            <si:unit>gallon(U.K.)</si:unit>
        </si:real>
        <si:real>
            <si:label>hardness Rockwell C scale</si:label>
            <si:value>63.00</si:value>
            <si:unit>HRC</si:unit>
            <si:expandedUnc>
                <si:uncertainty>1.56</si:uncertainty>
                <si:coverageFactor>2</si:coverageFactor>
                <si:coverageFactor>0.95</si:coverageFactor>
            </si:expandedUnc>
        </si:real>
    </si:list>
</si:hybrid>
```

Figure 7.2 *XML example for the hybrid adapter for lists of real quantities with the unrecommended non-SI unit gallon and a quantity value expressed by the Rockwell hardness C scale (HRC) reference procedure*

Conversion between SI units and non-SI units for univariate quantities

In most cases, a non-SI unit $\mathrm{U_{NSI}}$ and the corresponding SI unit (or SI derived unit) $\mathrm{U_{SI}}$ are related through a (dimensionless) multiplicative conversion factor ρ. A quantity value q expressed as the product of the number m and the non-SI unit $\mathrm{U_{NSI}}$ may be written as the quantity value Q expressed as the product of the number M and the SI unit (or SI derived unit) $\mathrm{U_{SI}}$, where

$$Q/\mathrm{U_{SI}} = \rho q/\mathrm{U_{NSI}},$$

i.e.,

$$M = \rho m.$$

Uncertainty is propagated using the law of propagation of uncertainty as described in the Guide to the Expression of Uncertainty in Measurement (GUM) [2]. If the standard uncertainty associated with the quantity value is expressed as the product of the number $u(m)$ and the non-SI unit $\mathrm{U_{NSI}}$, the standard uncertainty may be expressed as the product of the number $u(M)$ and the SI unit (or SI derived unit) $\mathrm{U_{SI}}$, where

$$u(M) = \rho u(m).$$

Conversion between SI units and non-SI units for units of temperature

The quantity value t for temperature may be expressed as the product of the number t_F and the non-SI unit degree Fahrenheit (°F). The unit degree Fahrenheit (°F) can be related to the corresponding SI unit of Kelvin (K) through a linear conversion function, i.e.,

$$T/K = \rho_1 t/{}^\circ\mathrm{F} + \rho_2,$$

where $\rho_1 = 5/9$ and $\rho_2 = 45967/180$ are, respectively, the (dimensionless) multiplicative and additive terms in the linear conversion function.

The quantity value T may be expressed as the product of the number T_K and the SI unit Kelvin (K), where

$$T_\mathrm{K} = \rho_1 t_\mathrm{F} + \rho_2.$$

Uncertainty is propagated using the law of propagation of uncertainty as described in the GUM [2]. If the standard uncertainty associated with the quantity value is expressed as the product of the number $u(t_\mathrm{F})$ and the non-SI unit °F, the standard uncertainty may be expressed as the product of the number $u(T_\mathrm{K})$ and the SI unit K, where

$$u(T_\mathrm{K}) = \rho_1 u(t_\mathrm{F}).$$

Similarly, the SI coherent derived unit of degree Celsius (°C) and the corresponding SI unit of Kelvin (K) are related through a linear conversion function, i.e.,

$$T/\mathrm{K} = \rho_1 t/{}^\circ\mathrm{C} + \rho_2,$$

where $\rho_1 = 1$ and $\rho_2 = 273.15$ are, respectively, the (dimensionless) multiplicative and additive terms in the linear conversion function.

For a quantity value t expressed as the product of the number t_C and the SI coherent derived unit °C, the quantity value T may be expressed as the product of the number T_C and the SI unit Kelvin (K), where

$$T_\mathrm{C} = \rho_1 t_\mathrm{C} + \rho_2.$$

Uncertainty is propagated using the law of propagation of uncertainty as described in the GUM [2]. If the standard uncertainty associated with the quantity value is expressed as the product of the number $u(t_\mathrm{C})$ and the SI coherent derived unit °C, the standard uncertainty may be expressed as the product of the number $u(T_\mathrm{K})$ and the SI unit K, where

$$u(T_\mathrm{K}) = \rho_1 u(t_\mathrm{C}) = u(t_\mathrm{C}).$$

Conversion factors

Appendix B provides a table of non-SI units that are important for industrial applications and economic affairs. Contextual information on the application area, definition of the unit and conversion factors are provided. In addition, lists of conversion factors are provided for many more non-SI units in sections B.8 and B.9 of the Guide for the Use of the International System of Units (SI) [13].

DOI: xxxx.yyyyy.zzz

8 Medal system

Metrological data that are provided by the D-SI data model are categorized into different quality classes of machine-readability. Table 5.1 gives an overview.

Table 5.1 Quality classes of machine-readable metrology data

Requirement	Quality class				
	platinum	**gold**	**silver**	**bronze**	**improvable**
SI++ units (7 SI base units and 7 important units allowed with the SI)	x	x	x	x	x
SI++ units in hybrid element		x	x	x	x
SI++ units with SI prefix or SI-derived units1)		x	x	x	x
Non-SI units from BIPM SI brochure that are not part of the SI++ units1)			x	x	x
Units from the previous edition of the SI brochure that are deprecated in the latest edition of the SI1,2)				x	x
Unit not part of SI, missing components, wrong data types, invalid lists, …					x

1: Applied for all unit components including units in hybrid elements.

2: The 9th edition [1] of the SI brochure is the current edition and the 8th edition [17] is the previous one.

The platinum quality class corresponds to the strongest eligibility for an unambiguous and safe exchange of metrological data in machine communication. This eligibility decreases more and more with the medal designations gold, silver, and bronze. The transfer of metrological data from the class “improvable” is erroneous and chaotic as the exchanged information is incomplete and ambiguous. A machine may not be able to interpret this metrological data without human interference. Attributes such as “Next Generation” or dates are also associated with the quality classes. These attributes symbolise to users that the conversion of established data structures can take place in appropriate periods. New users are recommended to strive for the platinum quality class. A quality class is only reached if all metrological data that are digitally transmitted meet the requirements of that class.

Platinum - also known as “Next Generation” - is the purest form of mechanical exchange of metrological data. In addition to the specifications of the seven SI base units, the specification of the unit “one”, the units “degree”, “minute” and “second” for angles as well as the units “day” and “minute” for quantities of time are permitted in this quality class. This set of units is denoted as “SI++ units” in the context of the D-SI data model and it is listed in detail in the appendix tables A.1 and A.2. Prefixes, on the other hand, are not allowed in the Platinum class because they represent redundant information about the exponential notation of numerical values of a quantity.

Note: Due to the expected long time for a transfer to a communication of metrological data only relying on SI++ units, this quality class is also referred to as “Next Generation”.

Gold - also called “2030” – basically enables the exchange of values with the specification of SI++units extended by using SI prefixes and SI derived units, according to the specifications of the SI brochure [1].

Silver - also called 2024 - allows the exchange of values with the specification of SI++ units, SI prefixes, SI derived units and other units outside of the SI that are permitted to be used together with the SI [1].

Bronze - also known as 2020 - enables the exchange of digital measurement data by specifying a measured value and an outdated unit from the previous edition of the SI brochure [17] in the SI unit format. The units that can be used for the bronze medal are non-SI units that have once been allowed to be used with the SI but were deprecated in the latest edition of the SI brochure [1].

Improvable allows the exchange of measurement data without adhering to the formats described above. This includes the indication of measured values without the indication of a unit, the indication of measured values in a number system outside the decimal system or the separation of decimal and decimal places by a comma or other non-permitted characters.

Note (quality class in si:hybrid environment): The best medal ranking for si:hybrid elements is gold. It is achieved, if the unit of the first quantity element listed in si:hybrid is composed of SI++ units, SI++ units with SI prefixes or SI derived units. The silver medal ranking is achieved in si:hybrid if the first quantity contains a non-SI unit that is permitted by the SI brochure [1] but that is not part of the SI++ units. If the first quantity in si:hybrid has a unit with bronze ranking, than the bronze medal is achieved. The medal improvable is associated if any of the quantities in si:hybrid shows properties of the improvable quality classes.

The preceding definitions make clear, that, in this system, platinum meets the requirements with the most universal interchangeability. In contrast, national units such as the “Oechsle” in Germany, the “Chi” in China or the “imperial mile” in the USA are not supported by means of machine-readability. Hence, such national units are classified as “improvable”.

9 Benefits and outlook

A central challenge in the field of digitalisation is the reliability, uniqueness, security and objective evaluability of transmitted metrological information. The foundations for addressing this challenge are provided in the approach outlined in this brochure.
The presented meta data model will be the starting point for reliable and trustworthy new technologies and services for the largely digitalised world of tomorrow, in which metrological information will be available to all at any place and at any time.

Basically, there are two approaches to use the presented meta data format. It can either be transferred into existing data exchange formats or form the basis for establishing the use of units in new exchange formats. Digital calibration certificates are one example of new exchange formats. In the second volume of this series, a corresponding DIN EN ISO/IEC 17025 compliant design is presented.

In the future, networked sensors will record all aspects of the production process and make them available to a comprehensive quality management system. With these complete data sets, the performance of systems and processes can then be captured effectively and efficiently, allowing data analytic methods to provide information on optimised system performance. This activity leads to reduced downtime, less waste, significant improvement in quality, and ultimately greater economic success.

Current cloud storage and services offer ways to store data, but do not provide information about the origin of the data or how to interpret the data. Corresponding uniform standards and methodical procedures are lacking. Here the presented consistent SI-based data communication, the creation of a validation possibility of XML-based measurement result representation, and the proposed medal system for the objective quality evaluation of

DOI: xxxx.yyyyy.zzz

measurement results will contribute to a clear increase of comprehensibility, reliability and confidence in existing measurement results.

E-government concepts, which will be established in many industrialised countries in the near future, will benefit from the proposed uniform presentation of results, as will universities, research institutes and industry. Smart Products, Smart Logistic, Smart Grids, Smart Mobility and Smart Health will find a common, globally coordinated communication standard and so will be able to be reliably integrated into everyday life.

In the future, it is planned to extend the D-SI data model with structures for large amounts of data (e.g. Cartesian coordinates of CT measurements). The electronic transmission of measurement results for quantities that are not specified by a combination of a number and a unit of measurement, but are based, for example, on a reference material or a measurement method, is also part of further development.

DOI: xxxx.yyyyy.zzz

10 References

[1] BIPM Brochure, *The international System of Units (SI) 9th edition 2019*, https://www.bipm.org/en/publications/si-brochure/ (accessed August 2019)

[2] JCGM 100:2008 *Evaluation of Measurement Data – Guide to the Expression of uncertainty in Measurement*, GUM 1995 with minor corrections, first edition, September 2008

[3] RFC 3629 *UTF-8, A transformation format of ISO 10464*, https://tools.ietf.org/html/rfc3629 (accessed November 2018)

[4] P. J. Mohr, D. B. Newell, B. N. Taylor CODATA Recommended values of the Fundamental Physical Constants: 2014, DOI 10.5281/zenodo.22826, June 2015

[5] JCGM 200:2012 *International Vocabulary of Metrology – Basic and general concepts and associated terms (VIM)*, 3rd edition, 2008 version with minor corrections, 2012

[6] ISO 80000-1 *Quantities and units- Part 1: General*, first edition, November 2011

[7] World Wide Web Consortium *Extensible Markup Language (XML)*; version 1.0; fifth edition https://www.w3.org/TR/xml/ (accessed February 2019)

[8] JCGM 101:2008 *Evaluation of Measurement Data Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method*, first edition, 2008; https://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf (accessed October 2019)

[9] JCGM 102:2011 *Evaluation of Measurement Data – Supplement 2 to the “Guide to the expression of uncertainty in measurement” — Extension to any number of output quantities*, October 2011; https://www.bipm.org/utils/common/documents/jcgm/JCGM_102_2011_E.pdf (accessed October 2019)

[10] ANSI/IEEE Std 754-2008 *IEEE Standard for Floating-Point Arithmetic*, electronic ISBN 978-0-7381-5752-8, DOI 10.1109/IEEESTD.2008.4610935, August 2008

[11] J. Wright *siunitx- A comprehensive (SI) units package*, Version v2.7s, http://mirrors.ctan.org/macros/latex/contrib/siunitx/siunitx.pdf, Revision May 2018, (accessed November 2018)

[12] DIN ISO 8601:2004 Data elements and interchange formats – Information interchange – Representation of dates and times (ISO 8601:2004), September 2004

[13] A. Thompson and B. N. Taylor. Guide for the Use of the International System of Units (SI). NIST Special Publication 811, 2008 edition.

[14] Annex 5 to the Convention on International Civil Aviation: Units of Measurement to be Used in Air and Ground Operations. International Civil Aviation Organization, 2010.

[15] IEC 80000-13:2008 Quantities and units -- Part 13: Information science and technology

[16] IEC TS 62720:2017 Identification of units of measurements for computer-based processing

[17] Bureau International des Poids et Mesures, *The international System of Units (SI) 8th edition 2006*

[18] EMPIR 17IND02 SmartCom, *D-SI XML reference implementation v1.3.0*, https://zenodo.org/record/3366902#.XabORWbgpaQ (accessed October 2019)

DOI: xxxx.yyyyy.zzz

Appendix A Identifiers for SI unit format

Table A.1 SI base units and identifiers for the SI unit format

Quantity	**Unit name**	**Symbol**	**Identifier**	**Class**	**Reference**
length	meter	m	\metre	platinum	[1, Table 2]
mass	kilogram[a]	kg	\kilogram	platinum	[1, Table 2]
time	second	s	\second	platinum	[1, Table 2]
current	ampere	A	\ampere	platinum	[1, Table 2]
thermodynamic temperature	kelvin	K	\kelvin	platinum	[1, Table 2]
amount of substance	mole	mol	\mole	platinum	[1, Table 2]
luminous intensity	candela	cd	\candela	platinum	[1, Table 2]

a: The kilogram is not allowed to be used with a prefix.

Table A.2 Extended SI base units and identifiers for the SI unit format

Quantity	Unit name	Symbol	Identifier	Class	Reference
Dimension number	one[a]	1	\one	platinum	[1, Sec. 2.2.1]
time	day	d	\day	platinum	[1, Table 8]
time	hour	h	\hour	platinum	[1, Table 8]
time	minute	min	\minute	platinum	[1, Table 8]
Plane angle	degree	°	\degree	platinum	[1, Table 8]
Plane angle	arcminute	‘	\arcminute	platinum	[1, Table 8]
Plane angle	arcsecond	“	\arcsecond	platinum	[1, Table 8]

a: Unit one must not be used with a prefix or an exponent.

DOI: xxxx.yyyyy.zzz

Table A.3 Derived SI units and identifiers for the SI unit format

Unit name	Symbol	Identifier	Class	Reference	Base SI
gram[a]	g	\gram	gold	[1, Sec. 3]	10^{-3} kg
radian	rad	\radian	gold	[1, Table 4]	m/m
steradian	sr	\steradian	gold	[1, Table 4]	m^2/m2
hertz	Hz	\hertz	gold	[1, Table 4]	s^{-1}
newton	N	\newton	gold	[1, Table 4]	m kg s^{-2}
pascal	Pa	\pascal	gold	[1, Table 4]	m^{-1} kg s^{-2}
joule	J	\joule	gold	[1, Table 4]	m^2 kg s^{-2}

a: The unit gram must not be combined with the prefix kilo.

Continued on next page

DOI: xxxx.yyyyy.zzz

Table A.3 Derived SI units - continued

Unit name	**Symbol**	**Identifier**	**Class**	**Reference**	**Base SI**
watt	W	\watt	gold	[1, Table 4]	$m^2\ kg\ s^{-3}$
coulomb	C	\coulomb	gold	[1, Table 4]	$s\ A$
volt	V	\volt	gold	[1, Table 4]	$m^2\ kg\ s^{-3}\ A^{-1}$
farad	F	\farad	gold	[1, Table 4]	$m^{-2}\ kg^{-1}\ s^4\ A^2$
ohm	Ω	\ohm	gold	[1, Table 4]	$m^2\ kg\ s^{-3}\ A^{-2}$
siemens	S	\siemens	gold	[1, Table 4]	$m^{-2}\ kg^{-1}\ s^3\ A^2$
weber	Wb	\weber	gold	[1, Table 4]	$m^2\ kg\ s^{-2}\ A^{-1}$
tesla	T	\tesla	gold	[1, Table 4]	$kg\ s^{-2}\ A^{-1}$

Continued on next page

DOI: xxxx.yyyyy.zzz

Table A.3 Derived SI units - continued

Unit name	Symbol	Identifier	Class	Reference	Base SI
henry	H	\henry	gold	[1, Table 4]	m^2 kg s^{-2} A^{-2}
degree Celsius	°C	\degreecelsius	gold	[1, Table 4]	K
lumen	lm	\lumen	gold	[1, Table 4]	m^2 m^{-2} cd
lux	lx	\lux	gold	[1, Table 4]	m^{-2} cd
becquerel	Bq	\becquerel	gold	[1, Table 4]	s^{-1}
sievert	Sv	\sievert	gold	[1, Table 4]	m^2 s^{-2}
gray	Gy	\gray	gold	[1, Table 4]	m^2 s^{-2}
katal	kat	\katal	gold	[1, Table 4]	s^{-1} mol

DOI: xxxx.yyyyy.zzz

Table A.4 Further non-SI units with their own identifier in the SI unit format

Unit name	Symbol	Identifier	Class	Reference	Basis SI recommended
hectare	ha	\hectare	silver	[1, Table 8]	m^2
litre	l	\litre	silver	[1, Table 8]	m^3
tonne	t	\tonne	silver	[1, Table 8]	kg
electronvolt	eV	\electronvolt	silver	[1, Table 8]	$m^2\ kg\ s^{-2}$
dalton	Da	\dalton	silver	[1, Table 8]	kg
astronomical Unit	au	\astronomicalunit	silver	[1, Table 8]	m
neper	Np	\neper	silver	[1, Table 8]	1
bel	B	\bel	silver	[1, Table 8]	1
decibel(b	dB	\decibel	silver	[1, Table 8]	1

DOI: xxxx.yyyyy.zzz

Table A.5 SI prefixes and identifiers for the SI unit format

Prefix name [a]	**Symbol**	**Multiplier**	**Identifier**	**Class**	**Reference**
deca[a]	da	10^{1}	\deca	gold	[1, Table 7]
hecto[a]	h	10^{2}	\hecto	gold	[1, Table 7]
kilo[a,b]	k	10^{3}	\kilo	gold	[1, Table 7]
mega[a]	M	10^{6}	\mega	gold	[1, Table 7]
giga[a]	G	10^{9}	\giga	gold	[1, Table 7]
tera[a]	T	10^{12}	\tera	gold	[1, Table 7]
peta[a]	P	10^{15}	\peta	gold	[1, Table 7]
exa[a]	E	10^{18}	\exa	gold	[1, Table 7]
zetta[a]	Z	10^{21}	\zetta	gold	[1, Table 7]
yotta[a]	Y	10^{24}	\yotta	gold	[1, Table 7]

a: All prefixes must not be combined with the unit kilogram, decibel and one.
b: The prefix must no be combined with the unit gram.

Continued on next page

DOI: xxxx.yyyyy.zzz

Table A.5 SI prefixes and identifiers - continued

Prefix name[a]	**Symbol**	**Multiplier**	**Identifier**	**Class**	**Reference**
deci[a,b]	d	10^{-1}	\deci	gold	[1, Table 7]
centi[a]	c	10^{-2}	\centi	gold	[1, Table 7]
milli[a]	m	10^{-3}	\milli	gold	[1, Table 7]
micro[a]	µ	10^{-6}	\micro	gold	[1, Table 7]
nano[a]	n	10^{-9}	\nano	gold	[1, Table 7]
pico[a]	p	10^{-12}	\pico	gold	[1, Table 7]
femto[a]	f	10^{-15}	\femto	gold	[1, Table 7]
atto[a]	a	10^{-18}	\atto	gold	[1, Table 7]
zepto[a]	z	10^{-21}	\zepto	gold	[1, Table 7]
yocto[a]	y	10^{-24}	\yocto	gold	[1, Table 7]

a: All prefixes must not be combined with the unit kilogram, decibel and one.
b: The prefix must no be combined with the unit bel.

DOI: xxxx.yyyyy.zzz

Table A.6 Other identifiers in SI unit format

Type	**Identifier**	**Class**	**Remarks**
Operator for Division	\per	platinum	Divide the left unit by the right unit
Operator for potentiation	\tothe{<EXPONENT>}	platinum	Increments the unit to the left of the operator by the value given by the exponent. The expression <EXPONENT> must be replaced with a decimal value on application.

DOI: xxxx.yyyyy.zzz

Appendix B Common non-SI units

Non-SI units existed long before the International system of units (SI units) was introduced. For economic, business and cultural reasons, many non-SI units still appear not only in scientific, technical and commercial literature but also in industrial metrology.

This appendix brings a list of non-SI units used in several industrial areas such as marine, aviation, ship industry, etc. The following parameters of the selected non-SI units are given in the tables A.1 to A.12: the name, symbol, area of usage, definition, short description, quantity, parent SI unit, conversion rule to SI unit and unit identifiers in the D-SI unit format.

DOI: xxxx.yyyyy.zzz

Table B.1 – Contextual information for important non-SI units

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
nautical mile (international) mni, nq	length shipping industry, marine	length 1852 m	Common unit for distance in marine area.	m multiplier 1852	not defined \metre
nautical mile (British) M, NM, nm, mni, nq	length shipping industry, marine	length 1853.2 m	One nautical mile on the surface of the earth corresponds to an angle of one minute of arc at its centre.	m multiplier 1853.2	not defined \metre
foot ft	length distance (vertical), altitude, elevation, height, vertical speed	length 0.3048 m	Imperial unit for length of 0.3048 m.	m multiplier 1.3048	not defined \metre

DOI: xxxx.yyyyy.zzz

Table B.2 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
inch ", in	length	1/36 yard or 1/12 of a foot	0.0254 m	m multiplier 0.0254	not defined \metre
thousandth of an inch mil, thou	length pcb design (microelectronics industry)	1/1000 of an inch	0.0000254 m	m multiplier 0.0000254	not defined \metre
litre L, l	volume shipping industry	1/1000 of a cubic metre	Unit of volume the cube of length of one deci-metre, known most commonly as a litre (or litre) but may also be known as a cubic decimetre.	m^3 multiplier 0.001	\litre \metre\tothe{3}

DOI: xxxx.yyyyy.zzz

Table B.3 – Contextual information for important non-SI units - continued

Name **Symbol**	**Quantity** **Area of usage**	**Definition**	**Description**	**Parent SI** **Conversion to SI**	**D-SI unit** **Parent SI D-SI unit**
hectare ha	area agriculture	10000 m^2	Area of 100 m x 100 m	m^2 multiplier 10000	\hectare \metre\tothe{2}
hour h	time daytime	3600 s	A time period consisting of 3600 seconds (60 minutes)	s multiplier 3600	\hour \second
minute min	time daytime	60 s	A time period consisting of 60 seconds	s multiplier 60	\minute \second
imperial gallon gal	volume purchase fuel	volume 4.54609 litres	0.0045 m^3 , not to be confused with US gallon, used in UK and many former UK colonies	m^3 multiplier 0.0045469	not defined \metre\tothe{3}

Table B.4 – Contextual information for important non-SI units - continued

Name **Symbol**	**Quantity** **Area of usage**	**Definition**	**Description**	**Parent SI** **Conversion to SI**	**D-SI unit** **Parent SI D-SI unit**
US gallon gal	volume petrochemical industry, purchase fuel	volume of 3.785412 l	0.003785 m^3 , not to be confused with imperial gallon, used in USA	m^3 multiplier 0.003785412	not defined \metre\tothe{3}
barrel of oil bbl	volume petrochemical industry	42 US gallons	The volume of 1 standard oil barrel	m^3 multiplier 0.1589873	not defined \metre\tothe{3}
cubic foot ft^3	volume petrochemical industry	volume of cube with sides of 1 ft	0.0283 m^3, used for measurement of natural gas volume	m^3 multiplier 0.02831685	not defined \metre\tothe{3}
pound (avoirdupois) lb	mass airplane fuel quantity	Mass 0.45359237 kg	Unit for mass with historical importance.	kg multiplier 0.45359237	not defined \kilogram

DOI: xxxx.yyyyy.zzz

Table B.5 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
ounce (avoirdupois) oz	mass pcb design (micro-electronic industry)	1/16 of an pound (avoirdupois)	In pcb design ounces are used to measure thickness of copper layer on the pcb. A thickness of 1 oz is defined as thickness of copper if 1 oz of copper is pressed flat and spread evenly across surface of 1 ft^2. This is equivalent to a thickness of 1.37 mil or 34.79 µm.	kg multiplier 0.02834959	not defined \kilogram
pound per hour PPH, lb/h	flow fuel flow	Mass flow 0.4535927 kg/h = 126.00 mg/s	0.4535927 kg/h = 126.00 mg/s	kg multiplier 1.259979E-4	not defined \kilogram \second \tothe{-1}

DOI: xxxx.yyyyy.zzz

Table B.6 – Contextual information for important non-SI units - continued

Name **Symbol**	**Quantity** **Area of usage**	**Definition**	**Description**	**Parent SI** **Conversion to SI**	**D-SI unit** **Parent SI D-SI unit**
degree deg, °	angle aerospace, shipping industry	1/3600th of a full rotation in a plane	A degree is 1/360th of a full circle. 1 degree is equal to Pi/180 radians.	radian multiplier Pi/180 ≈ 0.01745329	\degree \metre\metre \tothe{-1} or \radian
minute (angle) '	angle aerospace, shipping industry	1/21600th of a full rotation in a plane	A minute is 1/21600th of a full circle 1 minute is equal to Pi/10800 radians.	radian multiplier Pi/10800 ≈0.0002908882	\arcminute \metre\metre \tothe{-1} or \radian
second (angle) "	angle aerospace, shipping industry	1/1296000th of a full rotation in a plane	A second is a 1/1296000th of a full circle. 1 second is equal to Pi/648000 radians.	radian multiplier Pi/648000 ≈4.848136E-06	\arcsecond \metre\metre \tothe{-1} or \radian

DOI: xxxx.yyyyy.zzz

Table B.7 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
degree Celsius °C	temperature universal	*T*/ K = *t*/°C + 273.15 *t* is the value of the temperature quantity with the unit °C	Celsius is a scale and unit of measurement for temperature. The unit was known until 1948 as centigrade.	K advanced calculation T / K = t / °C + 273.15	\degreecelsius \kelvin
degree Fahrenheit °F	temperature universal	*T* / K = (*t* / ° F + 459.67)/1.8 *t* is the value of the temperature quantity with the unit °F	On the Fahrenheit scale, the freezing point of water is 32 °F and the boiling point is 212 °F (at standard atmospheric pressure).	K advanced calculation T / K = (t / °F + 459.67) /1.8	not defined \kelvin

DOI: xxxx.yyyyy.zzz

Table B.8 – Contextual information for important non-SI units - continued

Name **Symbol**	**Quantity** **Area of usage**	**Definition**	**Description**	**Parent SI** **Conversion to SI**	**D-SI unit** **Parent SI D-SI unit**
tonne, in US metric ton t	mass shipping & air industry	1000 kg	Unit for large masses	kg multiplier 1000	\tonne \kilogram
Knot kt	speed shipping industry, marine	1 nautical mile per hour	Unit for speed of ships	m/s multiplier 0.5144444	Not defined \metre\second \tothe{-1}
millimetre of mercury, conventional mmHg	pressure medicine, aviation	1 mmHg = 133.322387415 Pa	Defined as exactly 133.322387415 pascals	Pa Multiplier 133.322337415	not defined \kilogram\metre \tothe{-1} \second \tothe {-2} or \pascal

DOI: xxxx.yyyyy.zzz

Table B.9 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
torr Torr	pressure metrology, high-vacuum physics and energy	1/760 * 101325 Pa	Defines as exactly 1/760 of a standard atmosphere (101325 Pa).	Pa multiplier 133.3224	not defined \kilogram\metre \tothe{-1}\second \tothe{-2} or \pascal
atmosphere standard atm	pressure aerospace industry	101325 Pa	The standard atmosphere is an international reference pressure defined as 101325 Pa.	Pa multiplier 101325	not defined \kilogram\metre \tothe{-1}\second \tothe{-2} or \pascal
bar bar	pressure automotive industry, hydraulic systems	100000 Pa	hundred thousand Pascal	Pa multiplier 100000	not defined \kilogram\metre \tothe{-1}\second \tothe{-2} or \pascal

DOI: xxxx.yyyyy.zzz

Table B.10 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
ampere hour Ah	electric charge batteries	A charge transported by a constant current of 1 ampere in 1 hour.	A unit of electric charge often used to measure charge in large batteries.	C multiplier 3600	\ampere\hour \ampere\second
milliampere hour mAh	electric charge batteries	A charge transported by a constant current of 1 milliampere in 1 hour.	A unit of electric charge often used to measure charges in small batteries.	C multiplier 3.6	\milli\ampere\hour ampere\second
kilowatt hour kWh	energy electrical energy consumption	Energy consumed when power of 1 kW is consumed constantly for 1 hour.	3.6 MJ	J multiplier 3600000	\kilo\watt\hour \metre\tothe{2} \kilogram\second \tothe{-2}

DOI: xxxx.yyyyy.zzz

Table B.11 – Contextual information for important non-SI units - continued

Name Symbol	Quantity Area of usage	Definition	Description	Parent SI Conversion to SI	D-SI unit Parent SI D-SI unit
neper Np	logarithmic ratio acoustics, shipping industry, air industry	ln(a/b) a, b are values of same unit	The ratio of values of power or field quantities on a logarithmic scale.	individual advanced calculation	\neper individual
bel B	logarithmic ratio acoustics, shipping industry, air industry	log(a/b) a, b are values of same unit	The ratio of values of power or field quantities on a logarithmic scale.	individual advanced calculation	\bel individual
decibel dB	logarithmic ratio acoustics, shipping industry, air industry	10 log (a/b) a, b are values of same unit	The ratio of values of power or field quantities on a logarithmic scale.	individual advanced calculation	\decibel individual

DOI: xxxx.yyyyy.zzz

Table B.12 – Contextual information for important non-SI units - continued

Name **Symbol**	**Quantity** **Area of usage**	**Definition**	**Description**	**Parent SI** **Conversion to SI**	**D-SI unit** **Parent SI D-SI unit**
British thermal unit Btu, Btu_{IT}	energy petrochemical industry	Energy needed to heat one lb of water for one °F.	Traditional unit of heat in imperial unit system, used to measure energy content of various fuels.	J multiplier 1055.056	not defined \metre\tothe{2} \kilogram\second \tothe{-2}
mach Ma	flow velocity high speed flight	ratio of flow velocity past a boundary to the local speed of sound	The local flow velocity 1 Ma is equal to the speed of sound.	m/s per m/s multiplier 1	not defined \metre\second \tothe{-1}\metre \tothe{-1}\second or \one

DOI: xxxx.yyyyy.zzz

DOI: xxxx.yyyyy.zzz

The content presented was developed within the framework of the EU-funded project SmartCom *“Communication and validation of smart data in IoT-networks*” *with the* support of international partners from science and industry.

https://www.ptb.de/empir2018/smartcom
(accessed November 2018)

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

DOI: xxxx.yyyyy.zzz