# Advanced metrological knowledge representation in the D-SI metadata model

Daniel Hutzschenreuter[1], Matthias Bernien[2], Frauke Gellersen[1], Moritz Jordan[1], Benedikt Seeger[1]

[1] *Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38112 Braunschweig, Germany,*
*daniel.hutzschenreuter@ptb,de*
[2] *Physikalisch-Technische Bundesanstalt, Abbestraße 2-12, 10587 Berlin, Germany*

*Abstract* – **Universal, unambiguous, safe, and easy to use metadata models are a critical prerequisite for the efficient digitalisation of services and tools in metrology. The D-SI is providing users in metrology and their stakeholders with a simple, yet powerful framework of metadata models for an efficient implementation of data based on the International System of Units. Recent extensions of the D-SI address needs of advanced application such as digital certificates of calibration and findings in respect to the ongoing digitalization under the International Committee of Weights and Measures. Updates are presented comprising improvements of semantics, kinds of quantities, statistical distributions of measurement uncertainties and representations of dependencies of uncertainties.**

## I. INTRODUCTION

Universal, unambiguous, safe, and easy to use metadata models are a critical prerequisite for the efficient digitalisation of services and tools in metrology. The General Conference on Weights and Measures has declared in its 26th meeting in 2022 [1] that it is a core mission for the metrology organisations and their partner organisations in standardisation, accreditation, and science to work towards commonly accepted digital representations of metrological data. These representations shall be based on the International System of Units (SI) [2] and implemented by the use of the FAIR (findable, accessible, interoperable, reusable) data principles [3]. A SI Digital Framework [4] and digital representations for the units of measurement of the SI-brochure [5] are under development in respect to this mission by the International Bureau of Weights and Measures and the Forum on Metrology and Digitalization of the International Committee of Weights and Measures [6]. These developments build an authoritative and trusted international digital reference for all knowledge related to the realization of and metrological traceability of measurements to the SI.

Metadata models such as the D-SI [7] are important tools to implement data exchange formats for humans and machines, that are linking the international references with a wide range of practical applications. These are, for example, measurement data in Digital Calibration Certificates [8], assets of research data, data ready for the training and verification of Machine-Learning, or data being exchanged through programming interfaces of software. The D-SI is in particular providing a consistent framework of metadata models combining the representation of measured quantity values (numbers, units) with semantics of the international Vocabulary in Metrology [9] and the Guide on the Expression of Uncertainties in Measurement [10] based on needs from both, the international perspective of metrology and users in industry.

Code Example 1 shows a quantity value for real and scalar data from the D-SI in Extensive Markup Language (XML) format [11]. The root element is si:real which is encapsulating data and metadata that are common to a majority of applications. Mandatory information are the numerical value of the quantity (si:value) and the associated unit of measurement in the SI (si:unit). Additional data are a text label (si:label), a time-stamp for the date and time attributed to the quantity (si:dateTime), and metadata on the uncertainty of the measurement (si:univariateMeasurementUncertainty). The uncertainty is provided by a standard measurement uncertainty (si:standardMU).

*Example 1. Real quantity value*

```
<si:real>
  <si:label>effective area of displacer
      </si:label>
  <si:value>0.01</si:value>
  <si:unit>\metre\tothe{2}</si:unit>
  <si:dateTime>2025-09-03 T13:42:00Z
      </si:dateTime>
  <si:univariateMeasurementUncertainty>
    <si:standardMU>
      <si:valueStandardMU>0.002
          </si:valueStandardMU>
      <si:distribution>normal
          </si:distribution>
    </si:standardMU>
  </si:univariateMeasurementUncertainty>
</si:real>
```

It defines the value of the standard uncertainty (si:valueStandardMU) and has the optional field (si:distribution) to define a Probability Density Function (PDF) of the uncertainty. Simple strings such as 'normal' and 'uniform' were used to annotate basic kinds of PDFs. Other elements in the D-SI are constant, complex and multivariate quantity values.

This paper presents recent advancements of the D-SI considering needs of priority that we have identified in discussions with partners from the international committees of metrology (CIPM FORUM-MD) and stakeholders of the application in DCCs. Section II will give an outline of general concepts that have been added to the D-SI metadata model in 2023. Section III gives an introduction to the future representation of Probability Density Functions (PDFs) for a better description of statistical properties of uncertainties of measurement. Furthermore, Section IV will outline the inclusion of uncertainty models of dependencies and resulting benefits for the propagation of measurement uncertainties across chain-links of calibration. This work will be concluded by a discussion of the updates in Section V and an outlook to future developments is given in Section VI.

## II. NEW GENERAL CONCEPTS IN THE D-SI

New general concepts comprise extensions to the existing data models that have been requested by user communities of the D-SI. These extensions have been identified to be of high priority to improve the interoperability and useability of metrological data for a wider number of users and are available in D-SI version 2.2.0 since 2023.

### A. Kinds of quantities

Quantity values using SI-base units as reference for the magnitude can simplify calculations such as the conversion and interoperation between data. To enable higher levels of an unambiguous interoperation it is however necessary to extend the information on the unit by metadata on the underlying kinds of quantities and scales of measurement [12]. One example are data for radioactive decay with the unit Becquerel and frequency with the unit Hertz. Both units have the SI-base representation 1/s and adding a mix of both quantity values only based on seeing, it is the same units of measurement would be an incorrect choice. Another example is the interoperation of quantities representing absolute temperature and differences of temperature. Depending on the kind of quantity (aspects) and units such as Degree Celsius and Kelvin, conversion of data needs to be able to distinguish between different measurement scale types and apply the correct calculations for these scales. The need to be able to assess information on the quantity for interoperation is in general relevant, if quantities of dimension one are present such as ratios of quantities of the same kind and counting quantities.

*Example 2. Quantity and significant digit*

```
<si:real>
  <si:label>effective area of displacer
      </si:label>
  <si:quantityTypeQUDT>Area
      </si:quantityTypeQUDT>
  <si:value>0.01143</si:value>
  <si:unit>\metre\tothe{2}</si:unit>
  <si:signifcantDigit>-3
      </si:signifcantDigit>
  <si:dateTime>2025-09-03 T13:42:00Z
      </si:dateTime>
</si:real>
```

The D-SI is addressing this need by introducing an abstract element for the quantity kind [9, def. 1.2] and the list of identifiers for kinds of quantities provided by the QUDT ontology [13] was adopted to the D-SI. Example 2 shows how the QUDT identifier can be written as element si:quantityTypeQUDT. Users can access metadata on the definition of the kind of quantity from QUDT by calling the resolvable persistent identifiers <prefix>/<ID>.ttl with the prefix qudt.org/vocab/quantitykind and ID 'Area' from the example.

### B. Significant digits

Knowledge on the significant digits of the magnitude of a quantity value is a requirement in practical applications where inference on the accuracy of data is deduced from printed (significant) digits in addition to values of uncertainty. The D-SI recommends to directly include contributions of rounding of values related to significant digits within the measurement uncertainty value as a clear way to enable a consistent and easy propagation of measurement uncertainty by machines. For cases, where users are required to provide additional information, an option was added to identify the significant digits. Using this option shall not replace the requirement to properly handle all uncertainty contributions within a budget leading to the reported uncertainty value.

A review of existing definitions for significant digits for an adoption in the D-SI turned out to lead to no useful reference. IEC [14] on basic terms in IT defines "least significant digit". However, it is not disclosed how to identify this digit in data. DKD Guideline L-13 [15] defines significant digits from the metrological perspective, where the leading two left digits of the uncertainty value are used to identify the significant digits of the value. However, it considers to provide the number of significant digits which is ambiguous, as these are counted starting from non-zero digits which will not work for digital numbers close to zero. ISO 80000 [16] addresses the need to identify significant digits by introducing the property of a rounding range. This rounding range is a power of ten defining the digit where rounding is applied which is more robust regarding the

digital representation of numbers than counting digits.

It was decided to introduce the element si:significantDigit in the D-SI which is defined to provide the integer number of the exponent of the power of ten of the rounding range as it is invariant considering different representations of numerical numbers and as users can quickly access the properties of significant digits considering the before mentioned specifications of IEC, ISO and DKD.

Example 2 shows the significant digit number -3. It implies that the value 0.01142 has the significant digits 0.011 and the rounding range is $10^{-3}$. Even if the value -3 is an exponent and not an actual identification of the significant digit in the sense of a specific place in the mantissa of a number, it was decided to still call it the significant digit to increase the recognition by users.

### C. Semantics of uncertainty of measurement

As the development of the SI Digital Framework is progressing and as new implementations in the form of ontologies become increasingly important, the role of semantics was carefully reviewed with the aim to close potential gaps that would limit a future interoperability. Semantics do on the one hand consider the terms used in data, such as identifiers for the units and kinds of quantities. Using ontologies such as the SI Reference Point and QUDT are important resources to enhance the semantic interoperability. On the other hand, element names and their underlying data types in the data model carry additional information on concepts that need to be interoperable across applications and domains. To be able to adhere to this requirement, it is important to take care when implementing the D-SI metadata model in specific formats such as XML. The XML implementation is based on an XML Schema Definition (XSD) file which sets up the element names and all underlying data types, whereby XML is coming with a set of universal primitive types such as integers, decimal (floating point numbers), strings, and Booleans. Using these primitive types directly to represent data creates only a weak semantic representation. Take for example the property of the label and the PDF of the measurement uncertainty. If both values would only be of the primitive type string, a machine would not be able to identify the difference in the meaning. Also, it is more difficult in this case to inform humans and machines in a machine-actionable way about links between these types and elements in XML and semantics form other sources such as controlled vocabularies for semantics of the VIM and GUM. To circumvent such kinds of limitations, the D-SI follows a strict strategy of typing where each element and property is provided by a unique and dedicated data type [11]. The real quantity value element has for example the type si:realQuantityType in the XSD which represents a subclass of a quantity value. All primitive data types are encapsulated by metrology specific types such as a type for the unit, the value of the magnitude and the PDF.

*Example 3. Alignment of semantics of D-SI implementations for real quantity value.*

```
// XML
<element name="real" type="realQuatityType">

// TURTL (ontology)
https://ptb.de/si#Real rdf:type owl:class,
        rdfs:subClassf dsio:QuantityValue;

// JSON-LD context
"real": {
  "@id": "https://ptb.de/si#Real",
  "@type": "https://ptb.de/si/realQuantityType"
}
```

Furthermore, uncertainty statements were semantically aligned regarding their dimensions as univariate measurements uncertainties (standard, expanded, coverage interval uncertainties) and bi- and multivariate uncertainties (coverage regions). To establish type definitions of all uncertainty models in the D-SI that are ready for future semantical interoperation, the new elements si:standardMU, si:expandedMU, and si:coverageIntervalMU were introduced. At the same time, we deprecated the old definitions si:expandedUncertainty and si:coverageInterval as their use of primitive types was not in line with the needs for future semantics.

As one side-effect, these updates also facilitated the alignment of various implementations of the D-SI. The XML implementation defines basic element names and semantic data types supporting the machine-readability of data. The implementation of the D-SI ontology instantiates a comprehensive knowledge graph adding meaning to the elements (classes of entities) and their properties in relation to other knowledge graphs such as the SI Digital Framework for reasoning by machines [17]. A JSON-LD (Java Script Object Notation – Linked Data) context definition of the D-SI [18] requires to associated element names with unique IDs of entities and data types which aligns with both, the ontological and XML representations. Example 3 outlines the links between the elements.

### III. PROBABILITY DENSITY FUNCTIONS

Using PDFs to model measurement uncertainties is a fundamental practice in metrology. To increase the digital maturity of interoperability, it is necessary to introduce parameterized entities for PDFs within data models. Semantics of PDFs require to identify the kind of a distribution such as the normal or student-t distribution and to align unambiguous types of parameters such as expectation values, variance, or degrees of freedom.

### A. Requirements of D-SI applications

The D-SI metadata model aims for an easy-to-use and universal implementation of PDFs that is not dependent on

any specific formatting of data such as XML or JSON. A representation should be compatible to the current practice in the D-SI where information can be included in a single line of code. PDFs defined by the GUM and its supplements are of relevance and the implementation should also support distributions from a wider range of applications. Furthermore, the representations need to adhere to FAIR data principles and allow semantic interoperability. This requirement includes the availability of persistent identifiers (PIDs) for all PDFs. The PIDs should allow to link to metadata on the distributions such as knowledge graphs and machine-actionable formulas that support a use and evaluation within software. Finally, existing and established standards should be reused where possible.

### B. Approach for representing PDFs

A review of existing standards for representing PDFs in a machine-usable format comprised openly available knowledge bases, ontologies, and software libraries. The number of existing approaches is surprisingly huge, as many specific solutions have been established. The reason for this development is probably the lack of commonly adopted international standards which poses a domain-overarching interoperability challenge. Thus, our investigation focused on comprehensive systems such as PDFs in METAS's UncLib [19], Wolfram software [20], UncertML [21], or ProbOnto [22]. The later implementation, ProbOnto, is an ontology for more than 150 different types of univariate and multivariate distributions. It gives a structured listing of codes for distributions and their parameters which are considered to be used within the D-SI. Examples of the codes are shown in Table 1.

*Table 1. ProbOnto PDF code examples.*

| PDF type code | Parameter codes |
|---|---|
| Uniform1 | minimum, maximum |
| Normal1 | mean, stdev |
| StudentT2 | mean, scale, degreesOfFreedom |

The proposal for the representation of the PDFs in the D-SI will use the codes of ProbOnto to define PIDs for distributions. The formalism to write the PID is the following:

$$<T>?<P>=<V>;\ldots;<P>=<V>$$

where $<T>$ is a placeholder to enter the PDF type code of the distribution, "?" is a separator indicating that the list of parameters begins, the entries $<P>$ are codes of distribution parameters, "=" is a value separator, $<V>$ are the numerical values of the parameters (double precision

floating point numbers allowing scientific exponent), and ";" is used for separating two or more parameters. This notation is a standard for parameterized URLs.

*Example 4. PID of normal distribution in Example 1.*

```
// PID
Normal1?mean=0.01;stdev=0.002
//D-SI element notation
<si:distribution>
    Normal1?mean=0.01;stdev=0.002
        </si:distribution>
```

Example 4 outlines a PID for a normal distribution. The tag 'Normal1' indicates, that the PDF is a normal distribution. Parameter 'mean=0.01' is the mean value of this distribution and 'stdev=0.002' is the value of its standard deviation (square root of variance value).

### C. FAIR services to resolve PDFs

The PID is only a simple semantic notation of the distribution of a measurement. To increase interoperability, it is required to establish additional FAIR services that link substantial metadata to the PIDs. Such services could be web services for example, that are looking up properties and metadata from ProbOnto and provide it to users in a semantically meaningful way by using ontologies or similar encodings. As the PID is implemented on the formalism of parametrized URLs, resolvers for the metadata can be implemented as simple RESTful Get requests [23] in the form of

<web service root URL><PID>.

## IV. DEPENDENCIES OF UNCERTAINTIES

So far, the D-SI models for measurement uncertainty are based on the reporting of simple uncertainty numbers, which are outcomes of the evaluation of uncertainty budgets. The budget-specific mathematical combination of different components of uncertainty is known and defined in the GUM under the principle of the linear propagation of uncertainties [10, section 7.1.4]. A better identification and ability to reuse information on individual components of uncertainty can help to improve the accuracy of data processing along chains of subsequent measurements and measurement models. It also helps to uncover hidden correlation between data. By introducing dependencies of uncertainty to the D-SI, the use of such information on components of uncertainties will be enabled.

### A. Comparison between conventional approach and measurement data considering dependencies

The hands-on example of calibrating leaks in gas flow measurement will be used to explain the concept of dependencies, how it is implemented in the D-SI, and how

it could improve the propagation of measurement uncertainties across DCC in calibration chains. In the calibration of test standards for low gas flow leaks, a model based measurand is established by the interpolation of flows coming from two reference leaks. Let $Q_1$ and $Q_2$ be the gas flows of the reference leaks $R_1$ and $R_2$. Then the interpolated gas flow of the test leak is

$$Q = \alpha(Q_2 - Q_1) + Q_1 \tag{1}$$

The parameter $\alpha$ depends on the readings of the reference leaks using a mass spectrometer, but for simplicity of the example, we consider it to be a constant value without uncertainty. The values of the gas flows of the references are calibrated individually. The calibrations are made using the same primary standard for flow. The flows are calculated according to Equation 2 where $p_1$ and $p_2$ are pressure measurements that are multiplied with the volume difference of the primary standard over time. The differences of time are $\Delta t_1$ and $\Delta t_2$. The volume difference is calculated by multiplying the effective area $A = A_1 = A_2$ of a calibrated displacer by its change in length $\Delta l_1$ and $\Delta l_2$.

$$Q_1 = p_1 A_1 \frac{\Delta l_1}{\Delta t_1}, \qquad Q_2 = p_2 A_2 \frac{\Delta l_2}{\Delta t_2} \tag{2}$$

The combined standard measurement uncertainties of $Q_1$ and $Q_2$ are calculated according to Equation 3 following the law of the propagation of uncertainties. The coefficients $c_*$ are the partial derivatives of $Q_i$ with respect to the variables $p_i$, $A$, $\Delta t_i$ and $\Delta l_i$. The standard measurement uncertainties of the inputs are the values $u(*)$.

$$u^2(Q_i) = c_{p_i}^2 u^2(p_i) + c_{A_i}^2 u^2(A) \\ + c_{\Delta l_i}^2 u^2(\Delta l_i) + c_{\Delta t_i}^2 u^2(\Delta t_i), \tag{3} \\ i = 1,2$$

It is obvious, that $Q_1$ and $Q_2$ are correlated through the component A, but also through the reference instruments that are used to measure the pressure, change of length and time. So far, this correlation has not been taken into account explicitly in calibrations of $Q$. We are assessing the influence of $A$ in the example in this paper. The other contributions of correlation are not considered for simplicity.

If the calibration model for $Q$ does not consider correlation, then its combined standard measurement uncertainty is

$$u_{uncor}^2(Q) = (1 - \alpha)^2 u^2(Q_1) + \alpha^2 u^2(Q_2). \tag{4}$$

The calculation of the correlated uncertainty is done by

$$u_{cor}^2(Q) = (1 - \alpha)^2 u^2(Q_1) + \alpha^2 u^2(Q_2) \\ + 2\alpha(1 - \alpha)\, u(Q_1, Q_2) \tag{5}$$

using the covariance value

$$u(Q_1, Q_2) = c_{A_1} c_{A_2} u^2(A). \tag{6}$$

Table 2 shows data of a calculation close to a real-world application. There is a clear difference in the uncertainty value considering correlation.

*Table 2. Example of values of uncertainty with and without correlation.*

| Input quantity values | |
|---|---|
| $p_1 = 3.250000$ Pa, <br> $\Delta l_1 = 4.76000 \cdot 10^{-3}$ m, <br> $\Delta t_1 = 234.000$ s, <br> $p_2 = 35.300$ Pa, <br> $\Delta l_2 = 4.08000 \cdot 10^{-3}$, <br> $\Delta t_2 = 200.600$ s, <br> $A = 9.0256 \cdot 10^{-5}$ m² | $u(p_1) = 3.3 \cdot 10^{-5}$ Pa, <br> $u(\Delta l_1) = 2.9 \cdot 10^{-7}$ m, <br> $u(\Delta t_1) = 5.0 \cdot 10^{-2}$ s, <br> $u(p_2) = 3.5 \cdot 10^{-2}$ Pa, <br> $u(\Delta l_2) = 2.9 \cdot 10^{-7}$ m, <br> $u(\Delta t_2) = 5.0 \cdot 10^{-2}$ s, <br> $u(A) = 6.9 \cdot 10^{-8}$ m² |
| Model-based output quantity values | |
| $Q_1 =$ <br> $5.9669 \cdot 10^{-9}$ N·m/s <br> $Q_2 =$ <br> $6.48 \cdot 10^{-8}$ N·m/s | $u(Q_1) =$ <br> $7.7 \cdot 10^{-12}$ N·m/s <br> $u(Q_2) =$ <br> $8.4 \cdot 10^{-11}$ N·m/s |
| $Q = 2.0799 \cdot 10^{-8}$ N·m/s, <br> ($\alpha = 0.2521008$) | $u_{uncor}(Q) =$ <br> $2.2 \cdot 10^{-11}$ N·m/s <br> $u_{cor}(Q) =$ <br> $2.4 \cdot 10^{-11}$ N·m/s |

*B. Existing approaches*

Two important tools explicitly applying dependencies are METAS's UncLib which is developed by Wollensack [19] and the GUM Tree Calculator (GTC) which is developed by Hall [24]. Common to both tools is the ability to propagate measurement uncertainty and correlation along measurement models that relate the measurement uncertainty of calculated quantities to input quantities. Uncertainties of input quantities in the models are standard measurement uncertainties; each uncertainty has a unique identifier and it is possible to quantify correlation between inputs. These inputs are in our example the properties $p_i$, $A$, $\Delta t_i$ and $\Delta l_i$. The uncertainties of output quantities in the measurement models are combined standard measurement uncertainties which are provided as lists of the components of the uncertainty budget. In our example, these components are $c_{p_i} u(p_i)$, $c_{A_i} u(A)$, etc. Each of the listed items has as additional information the identifier of the uncertainty of the input quantity it is depending on.

Despite the conceptual similarities of both tools, there are also differences. The latest version of UncLib considers in addition the usage of non-linear models to

propagate uncertainty and the application of Monte-Carlo-type methods for a simulation of properties of uncertainty. Thus, the uncertainty contributions are implemented by the sensitivity coefficients (Jacobian matrix values $c_*$) and providing the PDFs related to the input quantities. Correlation between input quantities is implemented by decomposing the correlation matrix into a set of virtual input quantities whose dependencies to the actual input quantities are used to reconstruct correlation data. A specific Eigenvalue decomposition method is required to calculate the virtual inputs. GTC is focusing on the linear propagation of uncertainties. It provides the dependencies by the values $c_* u(*)$. In addition, correlation between inputs can be expressed through a separate list of correlation values for each input quantity. GTC also allows explicit identification of interim quantities and their dependencies.

### C. Modelling of Dependencies in the D-SI

The requirements to ensure an easy-to-use D-SI and the integration with existing models for uncertainty are guiding the modelling of dependencies. User will need calculation software (middleware) to create and process data. The software must be able to handle correlation between input quantities, underlying units of measurements, and types of quantities in a straightforward way and, where appropriate, preventing redundancy of information.

In this respect, the existing D-SI models for univariate uncertainties are extended by an optional identifier field (si:uniqueID) to provide unique identifiers of the uncertainty components of input quantity values. Furthermore, a list of correlation values (si:correlations) can be added to univariate uncertainty statements. Example 5 outlines the input quantity $A_1$ in respect to the before mentioned properties. The quantity value is providing a standard measurement uncertainty element with the unique identifier 'ID:A1'. As $A_1$ and input quantity $A_2$ ('ID:A2') are identic, the list of correlations is providing correlation value one. This information is written in the D-SI as a list item (si:correlation) that has the numerical correlation value (si:valueCorrelation) and the unique identifier of the input quantity it is related to (si:correlatedQuantityID). Correlation elements are only written in the D-SI if their values are non-zero and for identifier pairs where si:uniqueID is not equal to si:correlatedQuantityID. Correlation between inputs would typically be specified within one DCC.

For writing the uncertainty of an output quantity that is derived from dependencies, a new univariate measurement uncertainty model is introduced (si:dependenciesMU). This model provides a list of the components (si:dependency) of the uncertainty. Each dependency consists of a numerical value (si:valueComponentMU) which is the product $c \cdot u$ (uncertainty of an input quantity times the sensitivity coefficient from the uncertainty

budget) and a reference to the unique identifier of the input quantity (si:inputQuantityID). The unit of measurement and type of quantity of the parent D-SI element (si:real) apply to the numerical value which is in line with other uncertainty models. Example 6 outlines this new model to represent the value and uncertainty of gas flow $Q$ including correlation. Within this approach unique identifiers of the input quantities and their uncertainty component provide a way of identifying and quantifying correlations between uncertainties stemming from different DCCs.

*Example 5. Dependencies in D-SI – input quantity*

```
<si:real>
  <si:label>effective area of displacer A1
      </si:label>
  <si:value>0.01</si:value>
  <si:unit>\metre\tothe{2}</si:unit>
  <si:univariateMeasurementUncertainty>
    <si:standardMU>
      <si:valueStandardMU>0.002
          </si:valueStandardMU>
      <si:distribution>normal
          </si:distribution>
      <si:uniqueID>ID:A1</si:uniqueID>
    </si:standardMU>
    <si:correlations>
      <si:correlation>
        <si:valueCorrelation>1
            </si:valueCorrelation>
        <si:correlatedQuantityID>ID:A2
            </si:correlatedQuantityID>
      </si:correlation>
    </si:correlations>
  </si:univariateMeasurementUncertainty>
</si:real>
```

*Example 6. Dependencies in D-SI – output quantity*

```
<si:real>
  <si:label>gas flow Q</si:label>
  <si:value>2.50</si:value>
  <si:unit>\metre\tothe{2}</si:unit>
  <si:univariateMeasurementUncertainty>
    <si:dependenciesMU>
      <si:dependency>
        <si:inputQuantityID>ID:1
            </si:inputQuantityID>
        <si:valueComponentMU>0.025
            </si:valueComponentMU>
      </si:dependency>
      <si:dependency>
        <si:inputQuantityID>ID:2
            </si:inputQuantityID>
        <si:valueComponentMU>0.025
            </si:valueComponentMU>
      </si:dependency>
      […]
    </si:dependenciesMU>
  </si:univariateMeasurementUncertainty>
</si:real>
```

A complete example of input and output quantity data in D-SI XML format and example calculations based on the R application are available [26].

## V.    DISCUSSION

The recent updates of the D-SI lay a foundation for applications that need higher levels of machine-interoperability in the future. These comprise the ability to interoperate units of measurement and kinds of quantities in a common framework but also to integrate with digital approaches for representing semantics in metrology and chains of traceability.

Metrological communities that are fostering the harmonization of the use of kinds of quantities are still at the beginning of their work. There is a clear need for a machine-actionable representation of kinds of quantities defined by the International System of Quantities [16], however, no digital representation exists today that is fully covering the whole system. For this reason, the D-SI may adopt further representation systems in addition to QUDT in the future. Important and emerging developments are the kinds of quantities from BIPM's SI Digital Framework [4], IEC's Common Data Dictionary codes [25], and developments such as the M-Layer registry of units, quantities and measurement scales [12].

It was already mentioned that the provision of PDFs is posing a new interoperability challenge. In addition to the suggested inline codes in the D-SI, proper services are required to provide machine-actionable metadata. The final technical implementation of the representation of the PDF's parameters is still to be determined and consultations with stakeholders are ongoing. While the URL-based string representation has advantageous regarding querying of metadata, an alternative hierarchical representation of key-value-pairs of parameters is considered to ease parsing of data in software.

The D-SI is aiming to establish a basic web service covering distributions from the GUM [10]. Initial ideas on this include the creation of PIDs for each distribution that can be resolved to machine-readable metadata derived from a common ontology. Additional APIs could be added to allow to draw random samples of the distribution, to create plots, and to interoperate information across different formats of representation. Moreover, some statistical analysis also requires the provision of products of PDFs (multiplication). The digital representations and web services will need to allow this in future.

Modelling of uncertainties based on dependencies is posing a great opportunity but also a great challenge to the whole user community. The technical implementation is accompanied by many additional questions that need to be solved. A thoughtful definition of the unique identifiers is required for example. One requirement is to limit the length of the identifiers to a fixed number of characters to enable efficient traversing of data by software. At the same time, the identifiers can act as persistent and resolvable links in the sense of FAIR data. Furthermore, in the case of UncLib, the identifiers can also serve as random seeds for Monte Carlo simulation. Another challenge is the need of additional training capacities to enable users from a wider number of communities to properly apply the dependencies. Semantic representations of knowledge in the GUM can help to make the concepts of the linear propagation of uncertainties more transparent to humans and machines. However, a lot more easy-to-understand and hands-on training courses will be needed.

Software enabling an easy use of the D-SI and the dependency-based treatment of measurement uncertainties based on GTC and UncLib is another important component supporting users. Seeger et al. are developing for example the 'dccQuantities' open-source software library for Python, which will allow the automated application of relevant mathematical operations [27].

Finally, it is left open, if an additional model for the representation of uncertainties based on Monte-Carlo simulation data will be introduced. Storing lists of full samples is already possible in applications such as DCCs which are based on the D-SI. However, users are hesitating to store the potentially large amounts of data. Investigation and exchange with users will continue to discuss, if alternative approaches such as the approximation of PDFs and summary models (e.g., histograms) of Monte-Carlo data may be adopted in the future.

## VI.    SUMMARY AND CONCLUSION

The presented updates of general concepts such as types of quantities and semantics, the introduction of PDFs and modelling of dependencies lead to a substantial advancement of the capabilities of the D-SI metadata model for exchanging values of quantities in digital applications. By these updates, the development of core models in the D-SI are completed. Future developments will provide maintenance of the core-models and extensions such as introducing additional representation systems for kinds of quantities or advancing the services helping to interoperate units, quantities, PDFs, and semantics, as well as implementations in various formats such as XML, JSON-LD, and ontologies.

## VII.    ACKNOWLEDGEMENTS

## REFERENCES

[1] General Conference on Weights and Measures, 'Resolution 2 of the 27th CGPM (2022), On the

global digital transformation and the International System of Units'. BIPM. [Online]. Available: https://doi.org/10.59161/CGPM2022RES2E

[2] BIPM, 'The International System of Units (SI), 9th edition'. 2019. [Online]. Available: https://www.bipm.org/documents/20126/41483022/SI-Brochure-9.pdf

[3] M. D. Wilkinson et al., 'The FAIR Guiding Principles for scientific data management and stewardship', Sci Data, vol. 3, p. 160018, 2016, doi: 10.1038/sdata.2016.18.

[4] BIPM, 'SI Digital Framework, Digital references for FAIR measurement data'. [Online]. Available: https://si-digital-framework.org/.

[5] BIPM, 'SI Reference Point'. [Online]. Available: https://si-digital-framework.org/SI.

[6] CIPM, 'Forum on Metrology and Digitalization (FORUM-MD)'. [Online]. Available: https://www.bipm.org/en/committees/fo/forum-md

[7] [1] Daniel Hutzschenreuter et al., 'SmartCom Digital System of Units (D-SI) Guide for the use of the metadata-format used in metrology for the easy-to-use, safe, harmonised and unambiguous digital transfer of metrological data - Second Edition', Zenodo, EMPIR, 2020, doi: 10.5281/zenodo.3816686.

[8] [1] S. G. Hackel et al., 'The Digital Calibration Certificate (DCC) for an End-to-End Digital Quality Infrastructure for Industry 4.0', Sci., vol. 5(1), p. 11, 2023, doi: 10.3390/sci5010011.

[9] JCGM, 'JCGM 200:2012 - International vocabulary of metrology – Basic and general concepts and associated terms (VIM) - 3rd edition'. BIPM, 2008. [Online]. Available: https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf.

[10] JCGM, 'JCGM 100:2008(E), Evaluation of measurement data - Guide to the expression of uncertainty in measurement (2008)'. BIPM, 2008. [Online]. Available: https://doi.org/10.59161/JCGM100-2008E.

[11] PTB, 'D-SI XML GitLab repository'. [Online]. Available: https://gitlab1.ptb.de/d-ptb/d-si/xsd-d-si.

[12] M. J. Kuster, 'Toward a metrology-information layer for digital systems', ACTA IMEKO, vol. 12, no. 1(18), pp. 1–4, 2023.

[13] QUDT.org, 'Website of Quantities Units Dimensions and Types (QUDT) Ontology'. [Online]. Available: https://qudt.org/

[14] ISO/IEC JTC 1, ISO/IEC 2382:2015 Information technology -Vocabulary, 2015.

[15] DKD, DKD-L 13-1 Rounding of Results and Measurement uncertainties in Calibration Certificates, Nov. 2021. doi: 10.7795/550.20220114.

[16] ISO/TC 12, ISO 80000-1:2022 Quantities and units - Part1: General, 2022.

[17] M. Jordan, 'D-SI Ontology GitLab repository'. [Online]. Available: https://gitlab1.ptb.de/d-ptb/d-si/ontology-d-si.

[18] D. Hutzschenreuter, 'D-SI JSON-LD context'. [Online]. Available: https://www.ptb.de/si/SI_Format.jsonld.

[19] M. Wollensack, METAS UncLib. (2025). [Online]. Available: https://www.metas.ch/metas/en/home/fabe/hochfrequenz/unclib.html.

[20] Wolfram, 'The Ultimate Univariate Probability Distribution Explorer'. [Online]. Available: https://blog.wolfram.com/2013/02/01/the-ultimate-univariate-probability-distribution-explorer

[21] 'Uncertainty Markup Language: UncertML'. [Online]. Available: https://web.archive.org/web/20161029215725/http://www.uncertml.org/.

[22] 'ProbOnto - the Ontology and Knowledge Base of Probability Distributions'. [Online]. Available: https://sites.google.com/site/probonto.

[23] [1] Wikipedia, 'REST (Representational State Transfer)'. [Online]. Available: https://en.wikipedia.org/wiki/REST.

[24] B. Hall and J. Borbely, GUM Tree Calculator. (2024). MSL. [Online]. Available: https://gtc.readthedocs.io/.

[25] IEC/SC 3D, 'IEC 61360-4 - IEC/SC 3D - Common Data Dictionary'. [Online]. Available: https://cdd.iec.ch/.

[26] D. Hutzschenreuter, 'D-SI Metadata Model – example data and software scripts for handling uncertainty dependencies'. Zenodo, 2025. doi: 10.5281/zenodo.15276391.

[27] B. Seeger, E. V. Stehr, and T. Bruns, 'Python-based development libraries and web applications to process and store data using a JSON representation XML-DCC', Measurement: Sensors, p. 101480, 2025, doi: 10.1016/j.measen.2024.101480.