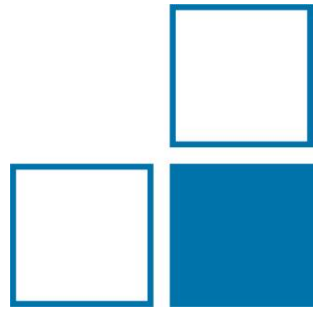
Physikalisch-Technische Bundesanstalt
Braunschweig and Berlin
National Metrology Institute

D-SI 2.2.0

Overview extension

2023-11-30



Thanks to all colleagues and partners who have proposed many valuable ideas for improvements and additions to the D-SI in the past months.

You helped the D-SI to make a major step into the future increasing its quality and useability for a wider range of (metrological) applications.

D-SI

D-SI Metadata Modell

Machine-2-machine exchange
- universal
- unambiguous
- save
- Easy-to-understand

Broschure
DOI: 10.5281/zenodo.3816686

Implementation Metadata Modell

- XML Schema
- (JSON Schema)
- (Ontology OWL2)

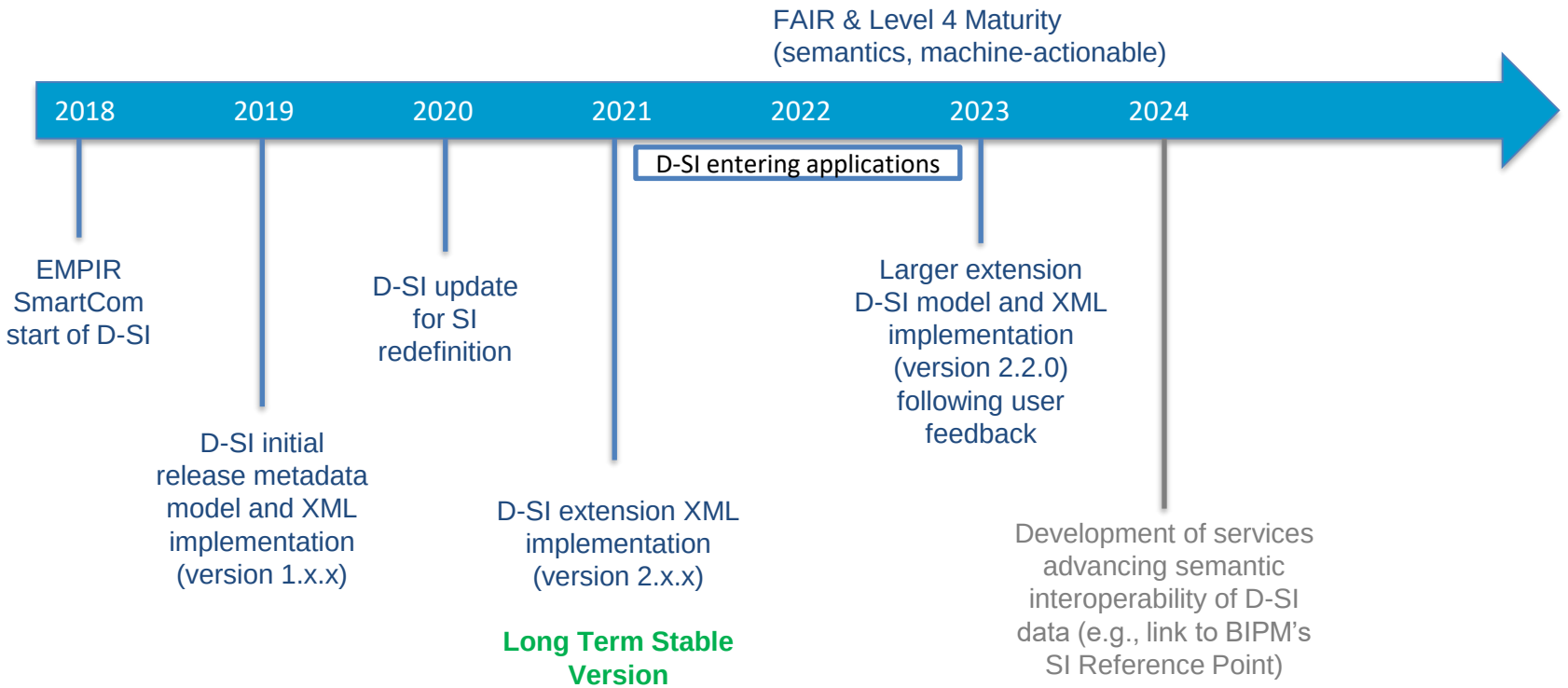
Repository
https://gitlab1.ptb.de/d-ptb/d-si/xsd-d-si

Instances of data with D-SI elements

- XML element, e.g., in Digital Calibration Certificate
- JSON, HDF5, CSV,…

```
<si:real>
  <si:label>length</si:label>
  <si:value>1.34</si:value>
  <si:unit>\metre</si:unit>
  <si:dateTime>
     2021-10-28T17:00:00.00+02:00
  </si:dateTime>
</si:real>
```

D-SI timeline



Extensions

1. Type of quantity
2. Standard measurement uncertainty
3. Element names for uncertainty
4. Disambiguating uncertainty statements
5. Voids (NaN, undefined values)
6. Significant Digit
7. Complex as XML list
8. Additional components in units
9. Preparing semantics for future
10. Sanitizing XSD

Overview

1. Type of quantity

si:quantityType

- Information on underlying quantity is essential to achieve ISO Smart level 4 of machine interoperable data
- Important: Identifiers for basic physical quantities (length, mass, ratios,…) not measurands (radius of circle, voltage at input, …)
- Thus, introduction of quantity kind (type) using PIDs from QUDT.org ontology
- Implementation considering extensions in the future such as using additional vocabularies from ISO, IEC etc.

si:quantityType - implementation

Extensible XSD

```
<xs:element name="quantityType" type="xs:string" abstract="true"/>
<xs:element name="quantityTypeQUDT" substitutionGroup="si:quantityType" type="qudt:quantityKind"/>
<xs:element name="quantityTypeIEC" substitutionGroup="si:quantityType"/>
```

Allowing

```
<si:real>
  …
  <si:quantityTypeQUDT>Length</si:quantityTypeQUDT> …
</si:real>
```

```
<xs:element name="real">
  …
  <xs:element ref="si:quantityType" minOcc
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema version="1.0.0" xmlns:xs="http://www.w3.org/2001/XMLSchema"
    targetNamespace="http://qudt.org/vocab/"
    xmlns:qudt="http://qudt.org/vocab/"
    elementFormDefault="qualified">

  <xs:simpleType name="quantityKind">
    <xs:restriction base="xs:string"/>
```

PID: https://qudt.org/vocab/quantitykind/Length(.ttl)

Overview

2. Standard measurement uncertainty

Standard measurement uncertainty

- Introducing element to provide minimum (initial) information on standard measurement uncertainty

```
<si:standardMU>
  <si:valueStandardMU>0.05</si:valueStandardMU>          (mandatory)
  <si:distribution>normal</si:distribution>              (optional)
</si:standardMU>
```

Overview

3. Element names for uncertainty

Element names for uncertainties

Need

- Improvement of naming for users and applications
- Improvement of underlying types for future semantics

Realization

- Extension of D-SI with uncertainty elements using improved (new) naming
- Preservation of old uncertainty element names to guarantee backwards compatibility (deprecated elements)

Example: expanded uncertainty

New:

- Top level element si:measurementUncertaintyUnivariate
- Individual MU value element

```
<si:real>
  <si:value>1.00</si:value>
  <si:unit>\metre</si:unit>
  <si:expandedUnc>
    <si:uncertainty>0.10</si:uncertainty>
    <si:coverageFactor>1.96</si:coverageFactor>
    <si:coverageProbability>0.95</si:coverageProbability>
    <si:distribution>normal</si:distribution>
  </si:expandedUnc>
</si:real>
```

becomes

```
<si:real>
  <si:value>1.00</si:value>
  <si:unit>\metre</si:unit>
  <si:measurementUncertaintyUnivariate>
    <si:expandedMU>
      <si:valueExpandedMU>0.10</si:valueExpandedMU>
      <si:coverageFactor>1.96</si:coverageFactor>
      <si:coverageProbability>0.95</si:coverageProbability>
      <si:distribution>normal</si:distribution>
    </si:expandedMU>
  </si:measurementUncertaintyUnivariate>
</si:real>
```

Change of element names and types

New tag	Old tag	New type	Old type
si:expandedMU	si:expandedUnc	si:expandedMUType	si:expandedUncType
si:coverageIntervalMU	si:coverageInterval	si:coverageIntervalMUType	si:coverageIntervalType
si:standardMU	-	si:standardMUType	-
si:valueExpandedMU	si:uncertainty	si:valueExpandedMUType	si:uncertaintyValueType
si:valueStandardMU	si:standardUnc	si:valueStandardMUType	si:uncertaintyValueType
si:intervalMin	si:intervalMin	si:intervalMinType	si:decimalType
si:intervalMax	si:intervalMax	si:intervalMaxType	si:decimalType

Overview

4. Disambiguating uncertainty statements

Refinement of documentation

k-value statement

consistent: k=1.96, p=0.95, normal distribution

inconsistent: k=2.00, p=0.95, normal distribution

a) Consistency of data responsibility of data creator

b) Use k-value and distribution information if p-value is inconsistent

k=2.00, normal distribution → p > 0.95

Statement of distribution

- Recommendation to always provide information on distribution
- If si:distribution not provided, then distribution unknown (not default „normal distribution“

Overview

5. Voids (NaN, undefined values)

Voids - basics

Background:
- Need in pharma (ALCOA+ principles) and several applications vs.
- D-SI: high risk of ambiguous statements if improperly introduced

solution

- “NaN” in xs:double
- Useable in XML, JSON, …
- documentation in D-SI as missing value. Definition of reason for missing value not part of D-SI.
- Medal “BRONZE”

```
<si:real>
  <si:value>NaN</si:value>
  <si:unit>\metre</si:unit>
<si:real>
```

… in real, complex, all lists, and uncertainty statements

Voids – in uncertainty statements

Voids supported in uncertainty statements

allowed

```
<si:uncertainty>NaN</si:uncertainty>
<si:coverageFactor>NaN</si:coverageFactor>
<si:coverageProbability>NaN</si:coverageProbability>
```

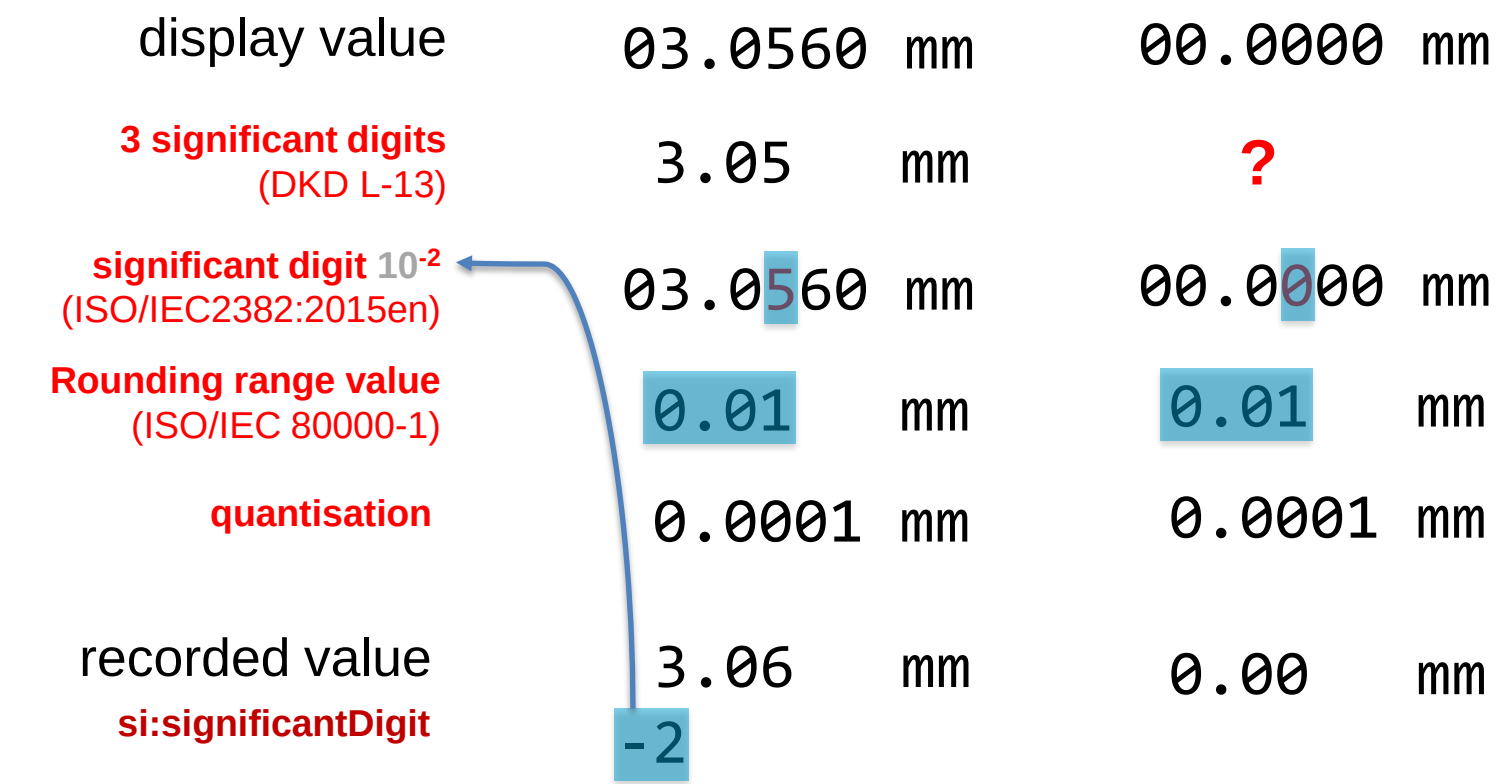
forbidden

```
<si:uncertainty>NaN</si:uncertainty>
<si:coverageFactor>1.96</si:coverageFactor>
<si:coverageProbability>0.95</si:coverageProbability>
```

Overview

6. Significant Digit

Significant Digit - Standardization



Significant Digit - Remarks

- Recommendation by D-SI to encode influences from significant digits and rounding as component contributing to the value of the measurement uncertainty for unambiguous use of data in digital systems.

- Provision of the significant digit is optional and can be applied by users, who are required to provide it. The **D-SI solution is compliant to ISO/IEC 80000-1, ISO/IEC2382:2015en, and calibration guideline requirements**.

- Provision of more metadata, e.g., rounding methods etc. is subject to higher level data structures such as calibration metadata, etc.

Overview

7. Complex as XML list

Complex as XML List

Concept the same as XML list for real & improved names of uncertainty elements

```
<si:complexListXMLList>
  <si:valueRealXMLList>1 2 3</si:valueRealXMLList>
  <si:valueImagXMLList>4 5 6</si:valueImagXMLList>
  <si:unitXMLList>\metre \second \kilogram</si:unitXMLList>
  <si:measurementUncertaintyBivariateXMLList>
    <si:ellipsoidalRegionMUXMLList>
      <si:covarianceMatrixXMLList>
        <si:columnXMLList>
          <si:covarianceXMLList>
            <si:valueXMLList>0.3 0.4 0.5</si:valueRealXMLList>
            <si:unitXMLList>\metre\metre</si:unitXMLList>
          </si:covarianceXMLList>
          ...
  </si:measurementUncertaintyBivariateXMLList>
</si:complexListXMLList>
```

Overview

8. Additional components in units

Units – additional units and components

\bit	gold	bit	bit	1	\one
\byte	gold	byte	B	1	\one
\ppm	gold	parts per million	ppm	1	\one
\percent	gold	percent	%	1	\one

Dimension one units

Needed, as allowed in 9th addition of SI brochure

Units – additional units and components

\yotta	gold	yotta(a	Y	1000000000000000000000000	[1, Table 7]
\ronna	gold	ronna	R	1000000000000000000000000000	[2, Resolution 3]
\ronto	gold	ronto	r	0.000000000000000000000000001	[2, Resolution 3]
\quecto	gold	quecto	q	0.000000000000000000000000000001	[2, Resolution 3]

\kibi	gold	kibi	Ki	1024	[1, p. 143]
\mebi	gold	mebi	Mi	1048576	[1, p. 143]
\gibi	gold	gibi	Gi	1073741824	[1, p. 143]

New SI prefixes and binary prefixes

Units – additional operator “\per”

D-SI Identifyer	Description	Examples	D-SI Medal
	Multiplication: simply append one unit to another	\metre\metre	PLATINUM
\tothe{n}	Exponent: following a unit where n is an integer or (+-)0.5	\metre\tothe{2}, \volt\tothe{-0.5}	PLATINUM
\per	Division: All units on the left side of the operator devided by all units on its right side	\metre\per\metre	SILVER

$$\frac{m}{s}$$

```
\metre\per\second
```

$$\frac{1}{s}$$

```
\per\second
\one\per\second
```

$$\frac{ms^{-1}}{kg}$$

```
\metre\second\tothe{-1}\per\kilogram
```

~~\metre\per\second\per\kilogram~~

Overview

9. Preparing semantics for future

Semantics

Interoperation with Knowledge Graphs, linked data, ontologies, …

BIPM SI Ref. Point

VIM, GUM vocab., ...

D-SI Element Types & URLs (PIDs)

Semantic Data Model (semantic interoperability layer)

Level 3-4

FAIR

D-SI extension enabling Semantic Data Model on top of Common Data Model

D-SI Metadata Model & Implementations

Common Data Model

Level 2-3

FAIR

Semantics – What are the issues?

Each component (type) in the D-SI metadata model needs to allow a machine-actionable link to metadata providing its meaning (semantics).

However, the initial design did not have this need in mind and the definition and use of some types is too ambiguous to establish unambiguous semantic links, e.g.

si:label and si:distribution have different meaning but are of the same type xs:string.

Semantics Example

```
<xs:complexType name="expandedMUType">
            <xs:annotation>
        <xs:documentation xml:lang="en">Definition of expanded measurement uncertainty data.</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="valueExpandedMU" type="si:valueExpandedMUType"/>
        <xs:element name="coverageFactor" type="si:kValueType"/>
        <xs:element name="coverageProbability" type="si:probabilityValueType"/>
        <xs:element name="distribution" type="si:distributionType" minOccurs="0"/>
    </xs:sequence>
</xs:complexType>
```

becomes

```
<xs:complexType name="expandedUncType">
            <xs:annotation>
        <xs:documentation xml:lang="en">[deprecated element] use si:expandedMU</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="uncertainty" type="si:uncertaintyValueType"/>
        <xs:element name="coverageFactor" type="si:kValueType"/>
        <xs:element name="coverageProbability" type="si:probabilityValueType"/>
        <xs:element name="distribution" type="xs:string" minOccurs="0"/>
    </xs:sequence>
</xs:complexType>
```

Semantics - improvements

All updates regarding quantity types and improved uncertainty elements establishing ability for unique semantic linking.

- Unique element names and types whenever data has different meaning.
- Specific simple types based on String and Double adding meaning like “min value type”, “label type”, “quantityTypeQUDT”, …

- However, ensuring backwards compatibility of the D-SI leads to some limitation to the ability to disambiguate all types for better semantics.

Overview

10. Sanitizing XSD

Sanitizing XSD

Old structure

<element1>
<element2>
…
<element1Type>
<element2Type>
…
<simpleType1>
…

New structure (easy to read)

<element1>
<element1Type>
<element2>
<element2Type>
…
<simpleType1>
…

Summary

1. Type of quantity
2. Standard measurement uncertainty
3. Element names for uncertainty
4. Disambiguating uncertainty statements
5. Voids (NaN, undefined values)
6. Significant Digit
7. Complex as XML list
8. Additional components in units
9. Preparing semantics for future
10. Sanitizing XSD