The Semantic SI Ontology: Engineering, Alignment, and Validation of a Semantic SI Model

Moritz Jordan[1][a], Giacomo Lanza[1][b], Shanna Schönhals[1][c] and Sören Auer[2,3][d]
[1]Physikalisch-Technische Bundesanstalt (PTB), Braunschweig und Berlin, Germany
[2]TIB - Leibniz Information Centre for Science and Technology, Hannover, Germany
[3]L3S Research Center, Leibniz Universität Hannover (LUH), Germany

Keywords: Ontology Engineering, Digital System of Units (D-SI), FAIR Data, Semantic Metrology.

Abstract: The Semantic SI Ontology (SIS) provides a semantic model for representing the value of a physical quantity, including kind of quantity, International System of Units (SI) measurement unit, and measurement uncertainty, in accordance with the official Bureau International des Poid et Mesures (BIPM) recommendations. Developed as a formal counterpart to the Digital System of Units XML schema definition (D-SI XSD), the ontology enables harmonized, machine-readable, and interoperable representation of metrological data. This paper outlines the design rationale, the transformation methodology from eXtensible Markup Language (XML) Schema to Web Ontology Language (OWL), and its alignment with existing semantic standards such as Quantities, Units, Dimensions, and Data Types Ontologies (QUDT) and the SI Reference Point. Core ontology structures - such as QuantityValue and MeasurementUncertainty - are discussed in detail. The paper also presents the validation framework, including a Python toolchain for generating OWL individuals from XML, Shapes Constraint Language (SHACL)-based shape validation, and reasoning with established OWL reasoners. Applications in ongoing projects, such as the Metadata4Ing and the Digital Calibration Certificate (DCC) ontology, demonstrate the practical relevance of the SIS as a foundational component in the digital transformation of metrology.

1 INTRODUCTION

The International System of Units (SI) is a globally recognized foundation for scientific measurements. It is also the common foundation for numerous norms and standards that are used in industrial applications and ensure interoperability between various actors of the global economic system. The Bureau International des Poids et Mesures (BIPM) provides and maintains the official description of the SI in its official brochure (BIPM, 2024) which is one of the fundamental standards in metrology.

The SI brochure, in combination with the International Vocabulary of Metrology (VIM) (BIPM, 2012) and the Guide to the Expression of Uncertainty in Measurement (GUM) (JCGM et al., 2023), also prescribes the correct way of reporting numerical information, which includes: the numerical value, the kind of quantity, the measurement unit and the measurement uncertainty. In the context of digital transformation and Findable, Accessible, Interoperable, and Reusable (FAIR) data principles, representing this compound information in a machine-interpretable and semantically rich format is an essential task, and currently one of the major challenges in metrology, as recognized by the highest authority in metrology, the Comité International des Poids et Mesures (CIPM), in its Vision on Transforming the International System of Units for a Digital World (CIPM, 2023).

In response to this need, the Semantic System of Units Ontology (SIS) has been developed as a semantic model that formalizes the representation of quantity values, units, and measurement uncertainties. This ontology provides a robust, logic-enabled foundation for expressing metrological data in alignment with the BIPM recommendations. It is conceived as the semantic complement to the Digital System of Units (D-SI) XML Schema Definition (XSD), which has been widely adopted for structured digital

[a] https://orcid.org/0000-0003-0941-5950
[b] https://orcid.org/0000-0002-2239-3955
[c] https://orcid.org/0000-0002-0475-0568
[d] https://orcid.org/0000-0002-0698-2864

Jordan, M., Lanza, G., Schönhals, S. and Auer, S.
The Semantic SI Ontology: Engineering, Alignment, and Validation of a Semantic SI Model.
DOI: 10.5220/0013707600004000

In *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025) - Volume 2: KEOD and KMIS*, pages 28-39
ISBN: 978-989-758-769-6; ISSN: 2184-3228

data exchange in the metrology community.

The SIS bridges the gap between eXtensible Markup Language (XML)-based data serialization and semantic modeling by leveraging the Web Ontology Language (OWL). This enables both syntactic validation and semantic reasoning over data involving physical measurements. It allows digital systems to interpret, validate, and interlink measurement data autonomously—thus unlocking new potentials for automated compliance, intelligent calibration workflows, and long-term traceability in smart manufacturing and laboratory environments.

Moreover, the SIS has been designed to be modular, extensible, and interoperable with other prominent semantic standards such as the Quantities, Units, Dimensions, and Data Types Ontologies (QUDT) ontologies and the SI Reference Point (Miles et al., 2022) from the BIPM. These synergies ensure that SIS is not an isolated artifact, but a foundational semantic infrastructure component compatible with international efforts toward digital metrology.

This paper presents the rationale, design process, and validation of the SIS. It elaborates on the methodology of converting XML schema constructs into OWL classes and properties, discusses alignment strategies with existing ontologies, and outlines the mechanisms for validating semantic integrity via Shapes Constraint Language (SHACL) and automated reasoners. Finally, it highlights the ontology's application in real-world initiatives such as Metadata4Ing and the Digital Calibration Certificate (DCC), demonstrating its relevance and adaptability in the digital transformation of metrology.

1.1 The "SI Brochure" and Its Digital Representations

The basic metrological knowledge, condensed from 150 years of research and international political agreements, is curated and disseminated by the BIPM in a collection of binding publications concerning denomination, relationships and usability of metrological terms and structures. Among these, the **BIPM brochure** is the one which deals with measurement units, namely regulating:

1. The International System of Units (SI), i. e. a list of standardized units, rules for multiples and algebraic combination of these units, and definitions for basic physical quantities.
2. Official recommendations for the use of the SI, including prescriptions how to use quantities and units (or their symbols) to report numerical values.

Other complementary recommendations are provided by other BIPM publications, e.g. the denomination of other fundamental metrological terms (VIM) and the calculation and representation of measurement uncertainties (GUM).

A plethora of digital representations for units exists, spacing from linear controlled vocabularies to mature semantic models including logical and mathematical relations: a selection of such approaches is reported in Section 2.2. Instead, a digital representation of the recommendations is still missing but it is increasingly needed for automated reasoning, traceability, and interoperability in scientific and industrial contexts. This is the focus of the work described in the present manuscript.

1.2 Previous Work: The D-SI XSD

The D-SI XSD, a comprehensive digital version of the BIPM recommendations, was implemented within an international cooperation in the context of DCC, involving also partners from the industry. The result is a metadata scheme for the structured representation of quantity values, including constants of nature, serialized in XSD. While machine-readable, it lacks semantic richness and reasoning capability. This schema served as the starting point for the SIS.

2 ONTOLOGIES IN METROLOGY

2.1 Semantic Technologies in Metrology

The increasing digitalization of metrology, driven by the need for automation, data interoperability, and traceable digital records, has encouraged the adoption of semantic technologies, which enable automated data interpretation and information processing by machines. Semantic technologies - including ontologies, Resource Description Framework (RDF), and reasoning tools - allow for the formal representation of domain knowledge in a machine-readable and logic-enabled way. In metrology, this enables a standardized and interoperable vocabulary for describing measurement results, uncertainty declarations, measurement units, measurement or calibration procedures, and reference values.

Ontologies are particularly powerful for modeling complex relationships among metrological concepts. By making explicit the semantics of terms and data structures, ontologies help bridge the gap between disparate systems and organizations, improving machine interpretability and reducing ambiguity in data exchange.

2.2 Existing Ontologies Covering Metrological Terms

Several ontology-driven efforts have emerged in recent years to address semantic interoperability and digital traceability in metrology and related domains:

- **SI Digital Framework (BIPM)**(Miles et al., 2022):
  This international initiative, coordinated by the International Bureau of Weights and Measures (BIPM), aims to create a semantic infrastructure for the SI in the digital age. One of its key components is the **SI Reference Point**, a machine-readable, RDF-based representation of the SI Brochure. The SI Reference Point provides canonical Uniform Resource Identifier (URI)s for SI units, prefixes, and constants and serves as a trusted semantic reference for digital metrology systems.
- **The M-Layer** (Hall, 2023):
  This international initiative, coordinated by the NCSLI consortium, aims to capture the semantic complexity of quantities, introducing the concepts of *aspect* (kind of quantity) and *scale* (ratio, interval, ordinal, and nominal).
- **QUDT** (QUDT Working Group, 2023):
  A general-purpose ontology widely used across science and engineering. It includes a comprehensive taxonomy of units and quantities, unit conversions, and relationships between physical dimensions. Although not "metrology native",it is well-aligned with SI and has been adopted in several industrial and research contexts.
- **Ontology of Units of Measure (OM)** (Rijgersberg et al., 2013):
  A domain-neutral ontology for representing quantities and measurement units. It provides reasoning support for unit compatibility and conversion, and is often used in scientific computing and Internet of Things (IoT) applications.
- **Extensible Observation Ontology (OBOE)** (Madin et al., 2007):
  Developed initially for ecological and environmental research, OBOE models entities, observations, and measurements in a flexible way. While not designed specifically for SI-based metrology, its conceptual structure overlaps with measurement modeling and serves as an inspiration for context-aware data models.

A somewhat longer collection of semantic descriptions for metrological information, with a call for their evaluation, is given in (Lanza et al., 2024), as well in (Keil and Schindler, 2018).

2.3 Positioning the Semantic SI Ontology (SIS)

The SIS distinguishes itself within this ecosystem as a highly targeted semantic model for the precise, harmonised, and safe digital representation of numerical variables conforming to the SI and its related concepts, such as constants, measurement uncertainties, and quantity values. The SIS does not contain a vocabulary or collection of quantities and units. Instead, it is conceived as a data model defining classes, properties, and validation rules to be used in operational metrology in combination with such a collection, such as QUDT, OM, or the more recent and authoritative SI Reference Point. The SIS brings an additional semantic layer, which guarantees, in addition, the machine *interpretability* of data.

By providing a modular, extensible, and machine-actionable model grounded in both formal ontological principles and practical XSD implementations, the SIS complements existing efforts and fills a crucial niche in the semantic infrastructure for modern metrology.

The SIS is available at the address https://www.ptb.de/sis/ .

3 CREATING THE SIS

3.1 Methodology: Transforming Schema Constructs into OWL Constructs

The SIS was designed trying to build a simple but rigorous semantical representation of the BIPM recommendations reported in the SI brochure, the VIM and the GUM. By logically grouping all required fields and considering the most relevant use cases occurring in practice, we defined the following groups of classes:

- Classes expressing occurrences of a quantity: single real value, single complex value, constant value, real list, complex list, nested lists (Section 4.1)
- Classes expressing uncertainty declarations: univariate/multivariate, standard/expanded, coverage interval (Section 4.2).

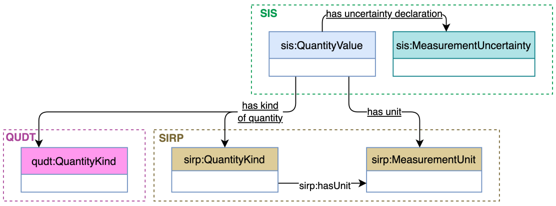

Figure 1: Basic structure of the SIS ontology showing the main parent classes `sis:QuantityValue` and `sis:MeasurementUncertainty`, as well as the interactions with the SI Reference Point (SIRP) and with QUDT.

The fields themselves were implemented either as datatype properties, having as value elementary data types (numeric, datetime, boolean, string) or as object properties, having as value composite data structures (i. e. instances of classes). We adopted the D-SI XSD as a blueprint for the choice of classes and properties, using the mapping approach described by (Wheeler et al., 2012). Some elements which were too much serialization-specific or which could be effectively described with the help of restrictions were left out in the mapping process.

In a subsequent step, a hierarchical structure was established by introducing subclass relationships where appropriate. In a further step, the concepts defined in the ontology were examined to define semantic connections to existing domain frameworks, such as the BIPM SI Digital Framework and QUDT, to enable reuse and ensure semantic interoperability.

3.2 Mapping of XSD Elements to Ontology Constructs

Following the general rule observed by (Wheeler et al., 2012), complex types were mapped to `owl:Class` constructs. Simple types were transformed into `rdfs:Datatype` to preserve their semantic meaning, as they often refine more general XSD types through additional restrictions.

Globally defined elements in the D-SI XSD - such as `six:Real`, which is of type RealQuantityType - were used to define the top-level concepts in the ontology. The name of the global element (e.g., `six:Real`) was adopted as the ontology class name (e.g., `sis:Real`), while the associated type (e.g. RealQuantityType - see Figure 2) determined the structure of the class. Consequently, the elements defined within RealQuantityType were mapped to object properties and datatype properties of the `sis:Real` class (see Table 1 for an overview of the mapping of types and elements).

```
<xs:complexType name="realQuantityType">
    <xs:sequence>
        <!-- optional label and quantity type-->
        <xs:element name="label" type="xs:string" minOccurs="0"/>
        <xs:element ref="si:quantityType" minOccurs="0"/>

        <!-- mandatory information -->
        <xs:element name="value" type="si:decimalType"/>
        <xs:element name="unit"  type="si:unitType"/>

        <!-- optional time stamp and significant digit -->
        <xs:element name="significantDigit" type="si:significantDigitType"
minOccurs="0"/>
        <xs:element name="dateTime" type="xs:dateTime" minOccurs="0"/>

        <!-- optional univariate measurement uncertainty -->
        <xs:element name="measurementUncertaintyUnivariate"
type="si:measurementUncertaintyUnivariateType" minOccurs="0"/>
    </xs:sequence>
</xs:complexType>
```

Figure 2: RealQuantityType in the D-SI XML schema definition.

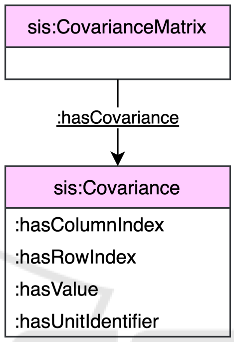

Figure 3: Class of `sis:CovarianceMatrix` with object property `hasCovariance` with range `sis:Covariance` with positional information preserved by indices.

3.3 Design Decisions and Challenges

The development of the SIS, while grounded in a well-defined XSD, presented conceptual and technical challenges discussed in the following section.

3.3.1 Loss of Order from `xs:sequence` and `xs:list`

One major limitation encountered during the schema-to-ontology transformation process was the inability of OWL (Web Ontology Language) to represent ordered sequences natively. The `xs:sequence` construct in XSD implies a strict ordering of child elements - information that is semantically meaningful in certain cases, such as:

- **Ordered lists of values**, such as numerical datasets using `xs:list`.
- **Covariance matrices**, where both the order of values and their multi-dimensional position are critical.

OWL does not preserve element order in standard property assertions. To address this, a dual strategy was adopted:

Table 1: Overview of the mapping of XML types to ontology constructs.

XSD	XSD example	Ontology	Ontology example
complexType	`six:realQuantityType`	class	`sis:Real`
simpleType	`six:unitPhaseType`	datatype	`---`
element with complexType	`six:covarianceMatrix`	object property	`sis:hasCovarianceMatrix`
element with simpleType	`six:unit`	datatype property	`sis:hasUnitIdentifier`

- In cases where ordering is not semantically important (e.g., sets of metadata properties or optional nested values), the ordering information was safely discarded during the transformation. If the knowledge graph data is transformed to XML data, the order information lies in the XSD.
- In cases where ordering is essential, a specialized class pattern was introduced. These classes (e.g. `sis:RealInList` or `sis:Covariance`) encapsulate:
  - The actual numerical value (e.g., a matrix entry or a real value in a list)
  - One or more index properties (e.g., `sis:hasRowIndex`, `sis:hasColumnIndex` or `sis:hasListIndex`) used to encode the position of the value within the list or matrix (see Figure 3 or `sis:RealInList` and `sis:ComplexInList` in Figure 4).

This pattern preserves the full structure of ordered collections in a way that is still compatible with OWL reasoning and SPARQL Protocol and RDF Query Language (SPARQL) querying, albeit at the cost of some verbosity.

3.3.2 From Schema Typing to Semantic Hierarchies

Another challenge was transitioning from a structural, schema-driven typing system to a semantically meaningful ontology class hierarchy. XML schema definitions tend to emphasize data validation, using complex types to group and constrain elements. However, in ontological modeling, the focus shifts to conceptual clarity and inheritance of meaning.

To this end, the mapping process required:

- Promoting XSD complex types to OWL classes.
- Introducing subclass relationships to reflect domain hierarchies not explicitly captured in the schema.
- Reinterpreting simple types with value restrictions as RDF Schema (RDFS) datatypes, while preserving their semantic constraints where possible.

This transition also required iterative domain validation to ensure that the resulting ontology remained faithful to both the formal structure of the XSD and the conceptual integrity of metrological knowledge.

3.3.3 Ontology Size and Granularity

A further consideration was the level of granularity for ontology classes and properties. The D-SI XSD uses a rich variety of named simple types, many of which are specific to particular use cases (e.g., decimalType, intervalMinType). Rather than collapsing these into overly generic types, the decision was made to preserve semantic granularity in the ontology.

3.4 Ontology Governance

The governance of the SIS is closely aligned with the underlying D-SI XSD, which serves as the primary source of truth for the ontology's conceptual scope. As such, the ontology will be updated to reflect changes in the XSD, ensuring consistency across both representations.

The next version of the D-SI XSD is being developed with an emphasis on long-term stability. Once the ontology has reached full coverage of the schema's scope, it too is intended to become a stable reference artefact. This stability is essential for enabling persistent identifiers, reproducible queries, and interoperability across systems that consume or reason over D-SI-based data.

Ongoing development, updates, and governance of the ontology will be managed by a dedicated team at the Physikalisch-Technische Bundesanstalt (PTB) in Germany. This team will be responsible for curating changes, maintaining alignment with the D-SI XSD, and ensuring the integrity and long-term sustainability of the ontology. Requests for change, extensions, or clarifications will be reviewed by this team in coordination with stakeholders and domain experts.

4 CORE STRUCTURE OF THE SIS

The SIS (Jordan and Lanza, 2025) is designed to provide a precise, semantic representation of the con-

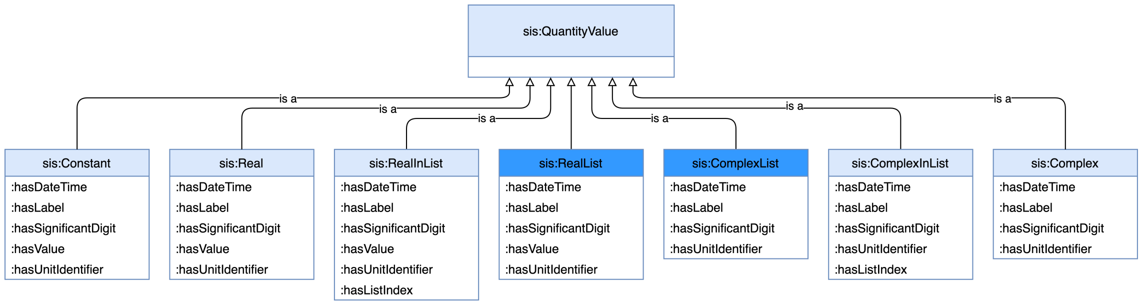

Figure 4: Hierarchy of `sis:QuantityValue` with the exemplary subclasses `sis:Real` `sis:Complex` and `sis:Constant` and some of their properties (e.g. `:hasValue`, `:hasUnitIdentifier`).

cepts used in the D-SI XSD. Its core is built around classes that represent quantities, constants, measurement values, and their associated uncertainties. These classes reflect the structure of the D-SI XSD, but are extended and aligned with broader metrological concepts through ontological modeling and alignment with existing vocabularies such as the VIM (BIPM, 2012).

The SIS in its present state (Version v0.2.1) comprehends 29 classes, 17 object properties, 19 datatype properties and 452 axioms.

4.1 QuantityValue: Central Abstraction for Measured and Defined Values

At the heart of the ontology is the abstract class `sis:QuantityValue`, which serves as the superclass for all types of numerical values and their associated units in the ontology. This includes both measured values (e.g., physical quantities) and defined constants.

`sis:QuantityValue` is defined as semantically close to the VIM concept of "quantity value" and is annotated with a `skos:closeMatch` to the corresponding concept in the VIM thesaurus (BIPM, 2017). It serves as a structural and semantic anchor for all types of value-bearing classes in the ontology.

The main subclasses of QuantityValue include:

- `sis:Constant` — A quantity with a fixed value (e.g., Planck constant), often used for defined values in the SI.
- `sis:Real` — A scalar real-valued quantity.
- `sis:RealList` — A collection of real values (e.g., for a data series).
- `sis:Complex` — A complex number with real and imaginary parts.
- `sis:ComplexList` — A collection of complex numbers.

Each subclass includes object and data properties such as hasUnit, hasValue (see Figure 4).

4.2 MeasurementUncertainty: Modeling Uncertainty in Measurement

A second major branch of the ontology is `sis:MeasurementUncertainty`, which represents uncertainty associated with quantity values. This class generalizes different types of uncertainty representations, from simple scalar standard uncertainties to multivariate confidence regions.

The class hierarchy under `sis:MeasurementUncertainty` is structured as follows (see Figure 5):

4.2.1 MeasurementUncertaintyUnivariate

This subclass represents one-dimensional (scalar) uncertainties. It is further specialized into:

- `sis:StandardMU` — A standard uncertainty expressed as a standard deviation, u.
- `sis:ExpandedMU` — An expanded uncertainty calculated by multiplying the standard deviation by a coverage factor $k > 1$ to achieve a high coverage probability $P > 90\%$.
- `sis:CoverageIntervalMU` — An uncertainty expressed in terms of a confidence interval, also associated with a coverage probability $P > 90\%$.

These classes reflect common uncertainty expressions found in calibration and conformity assessment documents.

4.2.2 MeasurementUncertaintyMultivariate

This subclass represents uncertainties over multidimensional quantities. It is further specialized into:

- `sis:EllipsoidalRegion` — A region defined by a covariance matrix and coverage probability, typically modeled as an uncertainty ellipsoid.
- `sis:RectangularRegion` — A region defined by bounds on individual components, such as an axis-aligned hyperrectangle.

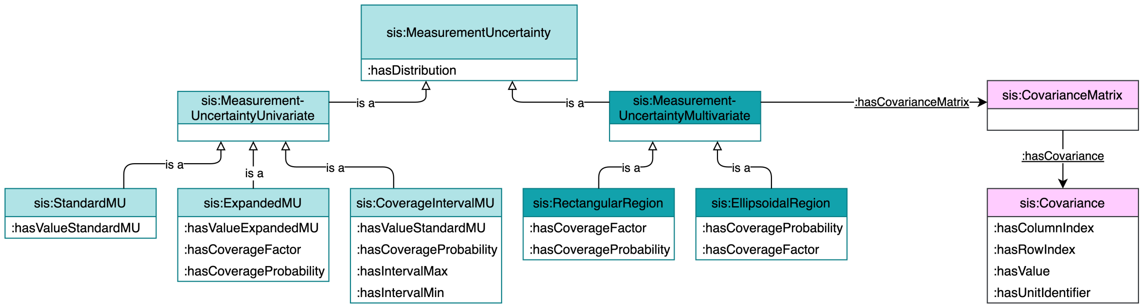

Figure 5: Hierarchy of `sis:MeasurementUncertainty` with its subclasses for univariate and multivariate uncertainties, some "grandchildren" classes and the relative datatype properties. The related classes for covariance matrices and covariance elements are also shown. The classes are connected by the object properties `:hasCovarIanceMatrix`, `:hasCovariance`.

These multivariate structures are essential for modeling correlations between components of vector-valued quantities, such as force vectors or tensor quantities.

4.2.3 ListMeasurementUncertaintyUnivariate

This class allows for the representation of lists of scalar uncertainties, where each element may correspond to one vector component. It is particularly useful for representing uncertainty in lists of values. Notably, both StandardMU and ExpandedMU are modeled as subclasses of this class as well, to support cases where scalar uncertainties are expressed in a list-like format.

4.3 Alignment and Reuse of Existing Semantic Resources

One key principle in the design of the SIS is semantic interoperability—the ability to integrate with and leverage existing, widely accepted ontologies in the metrology and quantity domain. Instead of developing an isolated model, the SIS reuses and aligns with established vocabularies to enhance compatibility, reduce redundancy, and promote harmonization across digital metrological resources.

4.3.1 Reuse of QuantityKind from QUDT and SI Reference Point

The SIS introduces the class `sis:QuantityType` to capture the conceptual dimension or kind of a physical quantity (e.g., length, energy, action). Rather than creating an entirely new controlled vocabulary, this class is mapped to and aligned with existing resources, specifically:

- `qudt:QuantityKind` from the QUDT Ontologies;
- `sirp:QuantityKind` from the SI Reference Point.

Instances of `sis:QuantityType` use OWL object properties to reference external instances in these ontologies. For example, an instance representing the Planck constant may include an object property `sis:hasQuantityType` linking it to `qudt:Action`, an instance of `qudt:QuantityKind` (see Figure 6). This alignment enables semantic clarity and facilitates interoperability in applications that already use QUDT or SI RDF models.

4.3.2 Unit Representation and Semantic Modeling of Compound Units

The SIS supports multiple approaches for encoding units:

1. `sis:hasUnit` [Datatype Property] — This is a string-based representation of a D-SI-compliant unit, as it is being used in the D-SI XSD (e.g. `\joule\hertz\tothe{-1}` for the unit of the Planck Constant).
2. `sis:hasSiMeasurementUnit` [Object Property] — This links a quantity to a structured semantic representation of a unit, using the SI Reference Point's semantic model for units, including compound units.

Compound units (e.g., joule per hertz) are modeled using semantic decomposition into unit terms realized in the SI Reference Point. An instance of a compound unit, such as `sirp:joule.hertz-1`, is composed of unit terms as seen in Figure 6. The SIS reuses this model directly, enabling detailed reasoning over unit structures and consistency checks.

5 VALIDATION AND REASONING

Ensuring the correctness, consistency, and completeness of the SIS is a key step toward its reliable use in practical applications. This chapter describes the validation and reasoning strategies used throughout the

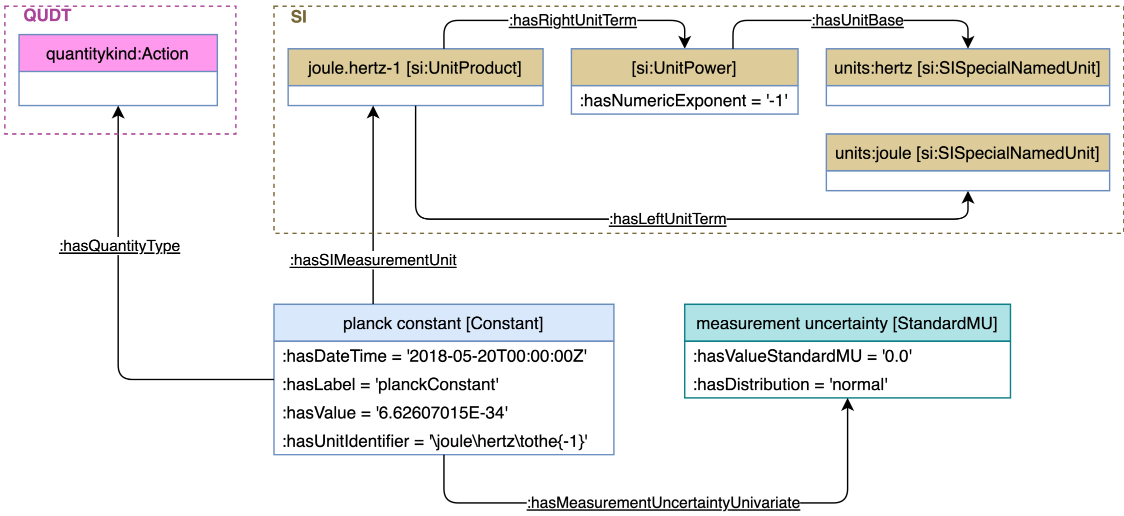

Figure 6: Instance of Planck Constant.

ontology's development process. The combination of data-driven validation using existing XML data, conformance testing via SHACL shapes, and semantic reasoning ensures that the SIS faithfully models the underlying D-SI data model while maintaining logical rigor.

5.1 XML-to-OWL Conversion Tool

To support large-scale validation with existing datasets, a custom Python tool was developed that transforms D-SI XML data into OWL individuals conforming to the SIS. This mapping algorithm systematically traverses XML instances that are valid according to the D-SI XSD and generates a corresponding RDF graph using OWL constructs.

By reusing real-world XML data as a source of truth, this tool enables

- **Rapid ontology validation** against large data volumes.
- **Robust regression testing**, ensuring that changes to the ontology do not break compatibility with existing XML data.
- **Data-driven ontology population**, facilitating the bootstrapping of knowledge graphs from already validated sources.

The generated RDF individuals retain all semantic information from the original XML representation, including nested structures and datatype values, enabling end-to-end validation of the ontology's modeling choices.

5.2 SHACL Shape Constraints

To enforce structural and semantic constraints at the RDF level, SHACL shapes were written for each class in the ontology. These shapes mirror the restrictions imposed by the XSD, including:

- **Cardinality restrictions** on properties,
- **Datatype and value ranges**,
- **Object property presence and relationships**,
- **Class hierarchies and expected typing**.

For example, a SHACL shape for the class `sis:Real` might ensure that it has exactly one `sis:hasValue` property of type `sis:decimal`, a `sis:hasUnit` property pointing to a recognized unit, and an optional `sis:hasStandardMU` (standard measurement uncertainty) property.

This dual modeling, with both OWL axioms and SHACL constraints, allows the ontology to be semantically expressive using OWL for reasoning and operationally verifiable using SHACL for structural conformance.

5.3 Validation Workflow and Graph Comparison

A comprehensive validation workflow was established to ensure that ontology instances generated from XML data and those authored directly in Turtle syntax are equivalent and conformant. The workflow consists of the following steps (see also Figure 7):

1. Generation: D-SI XML files are transformed to RDF/OWL individuals (in Turtle format) using the Python mapping tool

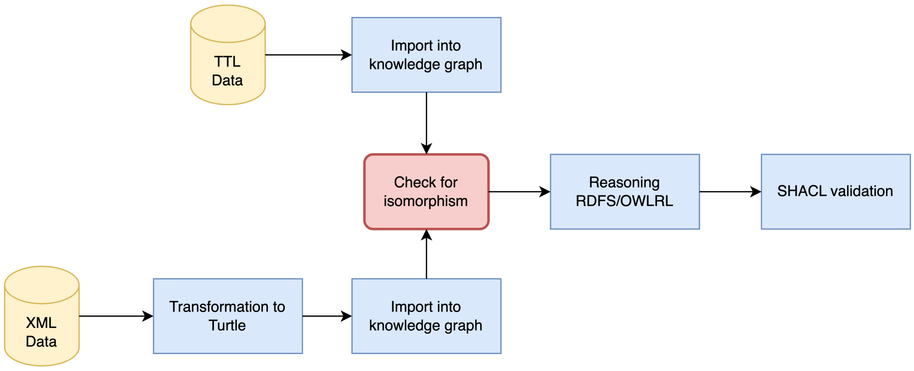

Figure 7: Validation workflow.

2. Import: Equivalent data authored directly in Turtle format and XML-sourced data are loaded into separate RDF graphs
3. Normalization: Both RDF graphs are serialized into a canonical form using RDFLib (Team, 2025), a Python library for handling RDF graphs. Blank node identifiers are standardized
4. Isomorphism Check: The two graphs are compared for isomorphism, ensuring structural and semantic equivalence across both representations
5. Logical Consistency: pySHACL also performs OWL reasoning and checks for logical consistency using built-in RDF/OWL reasoners.
6. SHACL Validation: The generated RDF graph is validated against SHACL shapes using the pySHACL Python library (Sommer, 2025).

This automated test suite is integrated into a pytest-based testing environment, providing rapid feedback during ontology development.

5.4 Advanced Reasoning with Protégé

Beyond automated validation, further semantic reasoning is performed using the Protégé ontology editor with two OWL 2 reasoners:

- **HermiT:** For checking logical consistency, class satisfiability, and inferred subclass relationships.
- **Pellet:** For additional support of datatype reasoning and explanation of inferred axioms.

These tools help to uncover hidden inconsistencies, unexpected inferences, and redundant axioms or design flaws.

Manual reasoning with Protégé ensures that the ontology behaves as intended under a standard OWL 2 DL reasoning regime, further enhancing the trustworthiness and robustness of the SIS.

5.5 Manual Validation with Domain Experts

In addition to automated validation and reasoning, manual review and feedback from domain experts played a crucial role in ensuring the semantic accuracy and practical relevance of the SIS. Metrology specialists were involved in iterative review cycles throughout the ontology development process to assess whether:

- The terminology and class structure reflected established metrological concepts.
- Relationships between quantities, units, and uncertainties were intuitive and aligned with real-world usage.
- Key distinctions (e.g., between different types of uncertainty or compound units) were properly modeled and unambiguous.
- The ontology remained understandable and usable for both human experts and machines.

These expert discussions led to several refinements, such as:

- Adjustments to class names to better reflect VIM terminology.
- Inclusion of `skos:closeMatch` annotations to clarify alignment with well-known definitions.
- Improvements to the handling of multivariate uncertainties and unit representations.

This collaborative validation step ensured that the SIS not only adheres to formal modeling standards but is also aligned with the expectations and needs of the metrology community.

6 APPLICATIONS OF THE SIS

The SIS was developed as a flexible, interoperable semantic data model for representing quantities, units, and uncertainties in accordance with the SI. Since its release, it has been the subject of discussions for its integration in multiple metrology-related initiatives and digital certification frameworks. This chapter presents current and emerging applications of the SIS.

6.1 Integration with Metadata4Ing

An ongoing discussion is underway with the Metadata for Engineering (M4I) workgroup within the National Research Infrastructure for Engineering Sciences (NFDI4Ing), which has created and maintained the homonymous (M4I) metadata schema for the description of data, investigations and data generation processes in the context of quantitative sciences, specially engineering and materials science. In future versions of the M4I ontology (Arndt et al., 2022), it is planned that numerical variables, which describe measurable quantities within engineering experiments and simulations, will be modeled as subclasses of `sis:QuantityValue`.

By inheriting from `sis:QuantityValue`, M4I will not need to define a separate uncertainty model. The SIS's inherent support for uncertainty (via the `MeasurementUncertainty` class and its subclasses) provides a robust and flexible foundation for representing error bounds, coverage intervals, and multivariate uncertainty. This reuse avoids redundancy and strengthens semantic alignment across domains, especially where scientific rigor and reproducibility are essential.

This integration exemplifies the ontology's ability to support broader scientific use cases beyond classical metrology, enabling fine-grained, semantically valid data annotations in digital research environments.

6.2 Use in the Digital Calibration Certificate (DCC) Ontology

Another major application of the SIS is its incorporation into the DCC Ontology, as described in (Jordan et al., 2024). The DCC Ontology serves as a semantic counterpart to the DCC XSD, which structures data for digital calibration certificates compliant with ISO/IEC 17025 — the international standard for the competence of testing and calibration laboratories.

The transition from analog to digital calibration certificates represents a paradigm shift toward machine-readable and machine-actionable metrological data. This transformation is essential for reducing manual processing effort, increasing accuracy, and enabling full integration into digital quality management systems.

The SIS is imported as the core model for quantity values within the DCC Ontology. It provides standardized classes for physical constants, measured values, and associated uncertainty, all of which are required in calibration data. Notably, this semantic integration guarantees that every numerical value in a certificate is precisely annotated with its unit, uncertainty model, and quantity type. Such clarity enables semantic reasoning, automatic data validation, and digital audit trails, which will be necessary for quality assurance and traceability in future digital metrology workflows.

6.3 Use in the Digital Document eXchange (DX) for ISO 170XX Documents

A planned future application of the SIS is its integration into the DX Ontology (Jagieniak et al., tted), which is being developed as a general semantic data model for ISO 170XX standard-based documents. These documents include but are not limited to:

- **DCC**
- **Digital Calibration Request (DCR)**
- **Digital Test Report (DTC)**
- **Digital Reference Material Certificate (DRMC)**

The Digital document eXchange (DX) Ontology provides a high-level, modular framework designed to cover the semantic representation of various standardized digital documents used across the metrology and quality infrastructure domains. For each document type, the DX Ontology defines a dedicated submodel tailored to the specific structural and semantic requirements of that document (e.g., a DCC submodel for calibration certificates, a DTC sub-model for test reports, etc.).

While the SIS is not the core foundation of the DX Ontology, it is a reusable component that will be incorporated whenever quantity values need to be represented. This design promotes modularity and reuse, enabling other parts of the DX Ontology to focus on document-specific metadata, relationships, and processes.

Especially in order to be used in industry applications and standardization documents, technology-agnostic yet semantically precise description of the metrological concepts described above are desperately needed. One application example is the submodel description currently being created by the Industrial Digital Twin Association (IDTA), a large industrial consortium defining the standards for data exchange in Industry 4.0, for Digital Quality Document (Digital Calibration Certificate). Here, the description of quantity values provided by SIS within the DCC ontology can be directly used as described in section 6.2 and will facilitate the creation of these sub-models (dig, 2025).

7 DISCUSSION AND CONCLUSION

The development of the SIS represents a significant step toward harmonized, semantically rich representations of quantitative metrological data. By deriving the ontology from the established D-SI XSD, we ensured compatibility with existing infrastructures while leveraging OWL's advantages for semantic expressiveness, reasoning, and data integration.

One of the key challenges addressed during development was the translation of XML-specific constructs—such as element ordering and schema restrictions—into OWL, which does not natively support sequence semantics. Where necessary, alternative modeling strategies were introduced, such as index-based representations for ordered values like covariance matrices. The reuse and alignment with existing ontologies, including QUDT, the SI Reference Point, and VIM, further enhance the interoperability and FAIRness of the ontology.

Validation was approached comprehensively, combining automated tools (e.g., SHACL, RDF reasoning, isomorphism tests) with manual review by metrology domain experts. This hybrid strategy ensures both technical correctness and domain adequacy. Moreover, the Python-based tooling for converting XML data into OWL individuals facilitates large-scale testing and adoption.

A more in-depth analysis of how the SIS will be performed when used for large-scale datasets as well as within distributed systems which will be part of the future work, e.g. when applied in the M4I community. The performance and scalability implications that may come up with the broader use of SIS will help to streamline and improve the ontology and contribute to its further development.

The SIS has already demonstrated its applicability in several contexts, including its upcoming integration into the DCC and M4I Ontology. Its design as a reusable, domain-specific module for representing quantity values with units and uncertainties makes it an ideal building block for broader semantic infrastructures such as the DX Ontology, which aims to support a wide range of ISO 170XX-compliant digital documents.

In conclusion, the SIS fills a critical gap in the semantic metrology landscape by providing a machine-interpretable, extensible, and interoperable model for core SI-based data. Future work will focus on broader adoption across metrological domains, further refinement through community feedback, and integration into national and international digital metrology strategies.

REFERENCES

(2025). Industrial digital twin.

Arndt, S., Farnbacher, B., Fuhrmans, M., Hachinger, S., Hickmann, J., Hoppe, N., Horsch, M. T., Iglezakis, D., Karmacharya, A., Lanza, G., Leimer, S., Munke, J., Terzijska, D., Theissen-Lipp, J., Wiljes, C., and Windeck, J. (2022). Metadata4ing: An ontology for describing the generation of research data within a scientific activity.

BIPM (2012). International vocabulary of metrology.

BIPM (2017). Vim definitions with informative annotations.

BIPM (2024). The international system of units (si).

CIPM (2023). Transforming the international system of units for a digital world, decision cipm/112-12 (2023).

Hall, B. (2023). Modelling Expressions of Physical Quantities. In *Proceedings of the 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 216–223, Rome, Italy. SCITEPRESS - Science and Technology Publications.

Jagieniak, J., Lin, S., Jordan, M., Demir, M.-A., Doering, L., Engel, T., Heeren, W., Loewe, J., Shanna, S., and Hackel, S. (Submitted). The dx schema as a modular concept for metrological certificates and reports.

JCGM, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML (2023). JCGM GUM-1:2023 Guide to the expression of uncertainty in measurement — Part 1: Introduction.

Jordan, M. and Lanza, G. (2025). D-si ontology.

Jordan, M., Schönhals, S., and Auer, S. (2024). Enhancing interoperability in digital calibration data exchange: A case for ontology development. *Measurement: Sensors*, page 101459.

Keil, J. M. and Schindler, S. (2018). Comparison and evaluation of ontologies for units of measurement. *Semantic Web*, 10(1):33–51. Publisher: SAGE Publications.

Lanza, G., Koval, M., Grasso Toro, F., Kuster, M., Gasca Aragón, H., Nahuel Coppa, D., Iturrate-García, M., Küpferling, M., Gruber, M., Hippolyte, J.-L., and Mari, L. (2024). Establishment of good practices for the usage of machine-actionable core metrological terminology. *Measurement: Sensors*, page 101456.

Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., and Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, 2(3):279–296. Meta-information systems and ontologies. A Special Feature from the 5th International Conference on Ecological Informatics ISEI5, Santa Barbara, CA, Dec. 4–7, 2006.

Miles, J., Abdallah, A. B., Chalk, S., Dudle, G., Gruber, M., Hippolyte, J.-L., and Meynadier, F. (2022). The SI Reference Point and Digital SI. Accessed: 2025-05-13.

QUDT Working Group (2023). QUDT Ontologies: Quantities, Units, Dimensions and Data Types. Version 2.1, Accessed: 2025-05-13.

Rijgersberg, H., van Assem, M., and Top, J. (2013). Ontology of units of measure and related concepts.

Sommer, A. (2025). pyshacl: Python validator for shacl. https://github.com/RDFLib/pySHACL. Version 0.30.1.

Team, R. (2025). Rdflib: A python package for working with rdf. https://github.com/RDFLib/rdflib. Version 7.1.4.

Wheeler, A., Dike, J., and Winburn, M. (2012). Xsd to owl: A case study.

List of Abbreviations

BIPM Bureau International des Poids et Mesures. 1–4

CIPM Comité International des Poids et Mesures. 1

D-SI Digital System of Units. 1, 2, 4–8, 11

DCC Digital Calibration Certificate. 2, 10, 11

DCR Digital Calibration Request. 10

DRMC Digital Reference Material Certificate. 10

DTC Digital Test Report. 10

DX Digital document eXchange. 10, 11

FAIR Findable, Accessible, Interoperable, and Reusable. 1, 11

GUM Guide to the Expression of Uncertainty in Measurement. 1, 2

IoT Internet of Things. 3

M4I Metadata for Engineering. 10, 11

NFDI4Ing National Research Infrastructure for Engineering Sciences. 10

OBOE Extensible Observation Ontology. 3

OM Ontology of Units of Measure. 3

OWL Web Ontology Language. 2–5, 7–9, 11

QUDT Quantities, Units, Dimensions, and Data Types Ontologies. 2–4, 7, 11

RDF Resource Description Framework. 2, 7–9, 11

RDFS RDF Schema. 5

SHACL Shapes Constraint Language. 2, 8, 9, 11

SI International System of Units. 1–3, 7, 10, 11

SIS Semantic System of Units Ontology. 1–11

SPARQL SPARQL Protocol and RDF Query Language. 5

URI Uniform Resource Identifier. 3

VIM International Vocabulary of Metrology. 1, 2, 6, 9, 11

XML eXtensible Markup Language. 2, 5, 8, 9, 11

XSD XML Schema Definition. 1–8, 10, 11