Semantic Data Charter (SDC) Specification

W3C Editor's Draft 27 September 2025

This version:
- https://www.google.com/search?q=https://semanticdatacharter.com/spec/sdc4-20250927/

Latest published version:
- https://www.google.com/search?q=https://semanticdatacharter.com/spec/sdc4/

Editors:
- Timothy W. Cook, Axius-SDC, Inc.

Abstract

The Semantic Data Charter (SDC) is a framework for creating semantically rich, machine-readable, and interoperable data models. It provides a set of principles and a reference model for defining data with clear, unambiguous meaning, ensuring that data is not only structured but also self-describing. This specification details the SDC Reference Model (RM), the core data types, and provides guidance on creating domain-specific data models.

Status of This Document

This document is an Editor's Draft and does not represent a consensus of any W3C Working Group. It is intended to be a starting point for discussion and is subject to change without notice.

1. Introduction

In an increasingly data-driven world, the ability to share, understand, and reuse data across different systems and domains is paramount. Traditional data modeling approaches often focus on the structure of data, leaving the meaning implicit and subject to misinterpretation. The Semantic Data Charter addresses this challenge by providing a framework for creating data models that are both structurally sound and semantically explicit.

The SDC is founded on three core pillars:

- **Enforce Governance:** Establish a formal, machine-readable contract for your data.
- **Embed Meaning:** Link your data to a universal business vocabulary.
- **Mandate Quality:** Formally define rules for handling imperfect data.

This specification provides the technical details of the SDC Reference Model, which is implemented in XML Schema (XSD) and OWL (Web Ontology Language).

2. Conformance

Conformance to the Semantic Data Charter is defined at two levels:

- **SDC Reference Model Conformance:** A data model conforms to the SDC Reference Model if its defining XML Schema is a valid xsd:restriction of the SDC Reference Model (sdc4.xsd). Data models **MUST NOT** use xsd:extension to add new elements or attributes to the SDC types. Conformance requires that both the data model schema itself is valid and that any data instance successfully validates against that restricted schema.
- **SDC Principles Conformance:** An organization's data practices conform to the SDC principles if they adhere to the governance, meaning, and quality pillars outlined in this document.

3. Architecture

The SDC architecture is composed of the following key components:

- **Reference Model (RM):** A set of core data types and structures that serve as the building blocks for all SDC-compliant data models. The RM is defined in sdc4.xsd and sdc4.ttl.
- **Data Models (DMs):** Domain-specific models created by constraining the components of the Reference Model.
- **Model Components (MCs):** The individual building blocks within a Data Model, derived from the RM components.
- **Semantic Annotations:** Metadata embedded within the data models that provide context and meaning, linking the data to ontologies and controlled vocabularies.

3.1. Core Data Types

The SDC Reference Model provides a rich set of extended data types (Xd* types) that go beyond the standard XML Schema data types. These include:

- XdAnyType: The base type for all SDC extended data types.
- XdStringType, XdTokenType: For textual data.
- XdBooleanType: For boolean values.
- XdCountType, XdQuantityType, XdFloatType, XdDoubleType: For numeric data.
- XdTemporalType: For date and time information.
- XdLinkType: For creating relationships between data models.
- XdFileType: For embedding or referencing binary data.
- ClusterType: For grouping related data elements.
- XdAdapterType: For adapting any Xd* type for use within a ClusterType.

3.2. Governance and Provenance Components

The SDC Reference Model includes a comprehensive set of components for capturing data governance, provenance, and audit information directly within the data model.

3.2.1. Governance

Governance components define the actors, roles, and responsibilities associated with the data.

- **PartyType**: Represents an actor involved with the data, which can be a person, organization, device, or software application.
- **ParticipationType**: Describes the role of a PartyType in a specific activity related to the data.
- **AttestationType**: Provides a formal mechanism for a party to attest to the content of the data.
- **AuditType**: Captures a detailed audit trail of every system and user that has interacted with the data.
- **Access Control**: The model supports linking to external access control systems via the acs element and allows for fine-grained control via the act (Access Control Tag) element.

3.2.2. Provenance

Provenance components provide a complete history of the data's origin, creation, and modification.

- **Instance Metadata**: The root DMType contains key provenance fields: instance_id, instance_version, creation_timestamp, subject, and provider.
- **Temporal Validity**: Every data element (XdAnyType) contains optional timestamp fields to track its lifecycle: vtb (Valid Time Begin), vte (Valid Time End), tr (Time Recorded), and modified.

3.3. The Data Model (DM) Wrapper and Semantic Grounding

A critical distinction exists between the metadata for a *data instance* and the metadata for the *data model definition* itself. The SDC architecture addresses both.

3.3.1. The DMType as an Instance Wrapper

The DMType serves as the root element for every SDC data instance. Its primary role is to be a container for the data payload and the provenance and governance metadata of that *specific instance*. The elements within the DMType, such as instance_id, creation_timestamp, provider, and subject, describe the "who, what, when, and where" of the data payload itself, not the abstract model it conforms to.

3.3.2. Semantic Grounding of the Data Model Definition

The semantic meaning and descriptive metadata of the *data model definition* (i.e., the schema) are established separately. This is where semantic vocabularies like Dublin Core are properly applied. This definition-level metadata is typically embedded directly within the schema files (e.g., using xsd:appinfo and RDF annotations in the XSD) or maintained in an external metadata registry.

3.4. Model Component Reusability and Immutability

A cornerstone of the SDC architecture is the ability to create reusable and immutable Model Components (MCs). This is achieved through a specific pattern of schema definition.
When a domain-specific structure is needed, it is defined by creating a new xsd:complexType that restricts a base type from the SDC Reference Model (e.g., sdc4:ClusterType). This new complexType is given a unique and permanent name using a Collision-Resistant Unique Identifier (CUID2), prefixed with mc-, such as mc-clj5x1g8f000008l09j7f6c3d.
To incorporate this component into a Data Model, an xsd:element is declared with a corresponding ms- prefixed name, assigned the mc-<CUID2> type, and placed into the xsd:substitutionGroup of a generic element from the Reference Model (e.g., sdc4:Item).
This mechanism provides three key advantages:

1. **Immutability and Reuse**: The mc-<CUID2> component can be imported and reused across any number of Data Models.
2. **Consistent Querying**: An application can reliably query for all instances of the ms-clj5x1g8f000008l09j7f6c3d element, knowing it will always find the same structure.
3. **Separation of Structure and Semantics**: This pattern decouples the immutable physical structure from its semantic meaning. The mc-<CUID2> and ms-<CUID2> define the structure, while the label element within that component's schema definition is *fixed* to provide the specific, immutable business context (e.g., fixed="Systolic").

4. Modeling Examples

Note: *The following XML snippets are instance documents. They are based on data model schemas where CUID-named elements (ms-... and dm-...) define the data structure. The semantic meaning, such as "Blood Pressure" or "Systolic", is not present in the instance label element; instead, it is a fixed attribute within the schema definition of the corresponding component, as described in Section 3.4.*

4.1. Healthcare

Use Case: Modeling a patient's vital signs. The dm- and ms- elements are non-semantic structural identifiers.

```
<dm-h7k2x...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
  <sdc4:dm-label>Patient Vitals</sdc4:dm-label>
  <sdc4:dm-language>en-US</sdc4:dm-language>
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4d>
      <sdc4:xdquantity-value>120</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
```

```
            <sdc4:xdstring-value>mmHg</sdc4:xdstring-value>
          </sdc4:xdquantity-units>
        </sdc4:ms-clj5x2p4k000108l01a2b3c4d>
        <sdc4:ms-clj5x2p4k000108l01a2b3c4e>
          <sdc4:xdquantity-value>80</sdc4:xdquantity-value>
          <sdc4:xdquantity-units>
            <sdc4:xdstring-value>mmHg</sdc4:xdstring-value>
          </sdc4:xdquantity-units>
        </sdc4:ms-clj5x2p4k000108l01a2b3c4e>
    </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-h7k2x...>
```

4.2. Agriculture

Use Case: Tracking soil moisture levels.

```
<dm-a4g9y...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
    <sdc4:dm-label>Soil Moisture Reading</sdc4:dm-label>
    <sdc4:dm-language>en-US</sdc4:dm-language>
    <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

    <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
        <sdc4:ms-clj5x2p4k000108l01a2b3c4f>
            <sdc4:xdquantity-value>35.5</sdc4:xdquantity-value>
            <sdc4:xdquantity-units>
                <sdc4:xdstring-value>%</sdc4:xdstring-value>
            </sdc4:xdquantity-units>
        </sdc4:ms-clj5x2p4k000108l01a2b3c4f>
        <sdc4:ms-clj5x3a9b000208l0e5f6g7h8>
            <sdc4:xdtemporal-datetime>2025-09-27T14:30:00Z</sdc4:xdtemporal-datetime>
        </sdc4:ms-clj5x3a9b000208l0e5f6g7h8>
    </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-a4g9y...>
```

4.3. Business

Use Case: Representing a customer order.

```
<dm-b8f3z...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
    <sdc4:dm-label>Customer Order</sdc4:dm-label>
    <sdc4:dm-language>en-US</sdc4:dm-language>
```

```
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
      <sdc4:xdstring-value>ORD-2025-12345</sdc4:xdstring-value>
    </sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4d>
      <sdc4:xdquantity-value>199.99</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
        <sdc4:xdstring-value>USD</sdc4:xdstring-value>
      </sdc4:xdquantity-units>
    </sdc4:ms-clj5x2p4k000108l01a2b3c4d>
  </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-b8f3z...>
```

4.4. Finance

Use Case: Modeling a stock trade.

```
<dm-f1d6w...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/
sdc4/)">
  <sdc4:dm-label>Stock Trade</sdc4:dm-label>
  <sdc4:dm-language>en-US</sdc4:dm-language>
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
      <sdc4:xdstring-value>AXS</sdc4:xdstring-value>
    </sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
    <sdc4:ms-clj5y5e2d000408l0m1n2o3p4>
      <sdc4:xdcount-value>100</sdc4:xdcount-value>
      <sdc4:xdcount-units>
        <sdc4:xdstring-value>shares</sdc4:xdstring-value>
      </sdc4:xdcount-units>
    </sdc4:ms-clj5y5e2d000408l0m1n2o3p4>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4d>
      <sdc4:xdquantity-value>50.25</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
        <sdc4:xdstring-value>USD</sdc4:xdstring-value>
      </sdc4:xdquantity-units>
    </sdc4:ms-clj5x2p4k000108l01a2b3c4d>
  </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-f1d6w...>
```

4.5. Threat & Fraud Detection

Use Case: Logging a suspicious login attempt.

```
<dm-t5c2v...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
  <sdc4:dm-label>Suspicious Login Attempt</sdc4:dm-label>
  <sdc4:dm-language>en-US</sdc4:dm-language>
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
      <sdc4:xdstring-value>198.51.100.1</sdc4:xdstring-value>
    </sdc4:ms-clj5x4b1c000308l0i9j8k7l6>
    <sdc4:ms-clj5x3a9b000208l0e5f6g7h8>
      <sdc4:xdtemporal-datetime>2025-09-27T18:00:00Z</sdc4:xdtemporal-datetime>
    </sdc4:ms-clj5x3a9b000208l0e5f6g7h8>
     <sdc4:ms-clj5x4b1c000308l0i9j8k7l7>
      <sdc4:xdstring-value>Multiple failed attempts</sdc4:xdstring-value>
    </sdc4:ms-clj5x4b1c000308l0i9j8k7l7>
  </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-t5c2v...>
```

4.6. Aerospace

Use Case: Recording telemetry from a satellite.

```
<dm-s9p8q...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
  <sdc4:dm-label>Satellite Telemetry</sdc4:dm-label>
  <sdc4:dm-language>en-US</sdc4:dm-language>
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4f>
      <sdc4:xdquantity-value>400</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
        <sdc4:xdstring-value>km</sdc4:xdstring-value>
      </sdc4:xdquantity-units>
    </sdc4:ms-clj5x2p4k000108l01a2b3c4f>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4g>
      <sdc4:xdquantity-value>-10.5</sdc4:xdquantity-value>
```

```
        <sdc4:xdquantity-units>
          <sdc4:xdstring-value>C</sdc4:xdstring-value>
        </sdc4:xdquantity-units>
      </sdc4:ms-clj5x2p4k000108l01a2b3c4g>
  </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-s9p8q...>
```

4.7. Engineering

Use Case: Defining the specifications for a mechanical part.

```
<dm-e3n7m...
xmlns:sdc4="[https://semanticdatacharter.com/ns/sdc4/](https://semanticdatacharter.com/ns/sdc4/)">
  <sdc4:dm-label>Mechanical Part Specification</sdc4:dm-label>
  <sdc4:dm-language>en-US</sdc4:dm-language>
  <sdc4:dm-encoding>UTF-8</sdc4:dm-encoding>

  <sdc4:ms-clj5x1g8f000008l09j7f6c3d>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4h>
      <sdc4:xdquantity-value>50</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
        <sdc4:xdstring-value>mm</sdc4:xdstring-value>
      </sdc4:xdquantity-units>
    </sdc4:ms-clj5x2p4k000108l01a2b3c4h>
    <sdc4:ms-clj5x2p4k000108l01a2b3c4i>
      <sdc4:xdquantity-value>10</sdc4:xdquantity-value>
      <sdc4:xdquantity-units>
        <sdc4:xdstring-value>mm</sdc4:xdstring-value>
      </sdc4:xdquantity-units>
    </sdc4:ms-clj5x2p4k000108l01a2b3c4i>
  </sdc4:ms-clj5x1g8f000008l09j7f6c3d>
</dm-e3n7m...>
```

5. Security Considerations

Implementers of the Semantic Data Charter should be aware of the following security considerations:

- **Data Privacy:** When modeling data that contains personally identifiable information (PII), care must be taken to ensure compliance with relevant privacy regulations (e.g., GDPR, HIPAA).
- **Access Control:** The act (Access Control Tag) element in XdAnyType should be used to enforce access control policies.

- **Data Integrity:** The hash-result and hash-function elements in XdFileType can be used to verify the integrity of binary data.

6. Reference Model Component Details

This section provides a detailed explanation of the primary complexType components available in the SDC Reference Model (sdc4.xsd).

6.1. Root and Structural Components

These components form the foundational structure of any SDC Data Model.

- **DMType**: The mandatory root element type for any SDC instance document. It acts as a wrapper for the data payload and contains instance-specific metadata.
  Data Model Schema Example: This example defines the root element for a "Patient Vitals" Data Model. It restricts the base DMType, fixes the dm-label to provide the model's semantic name, and specifies that the main data payload will be a specific "Vitals Cluster" component.

```
<xsd:complexType name="mc-h7k2x...">
 <xsd:complexContent>
  <xsd:restriction base="sdc4:DMType">
   <xsd:sequence>
    <xsd:element name="dm-label" type="xsd:string" fixed="Patient Vitals"/>
    <xsd:element name="dm-language" type="xsd:language" fixed="en-US"/>
    <xsd:element name="dm-encoding" type="xsd:string" fixed="UTF-8"/>
    ...
    <!-- The root item for this DM is a specific cluster -->
    <xsd:element ref="sdc4:ms-clj5x1g8f000008l09j7f6c3d"/>
    ...
   </xsd:sequence>
  </xsd:restriction>
 </xsd:complexContent>
</xsd:complexType>
```

- **ItemType**: An abstract type that serves as the base for ClusterType and XdAdapterType. It is not used directly but enables polymorphism in the model.
- **ClusterType**: A container component used to group other ItemType elements. This allows for the creation of arbitrarily complex hierarchical data structures.
  Data Model Schema Example: This defines a reusable "Blood Pressure" cluster. It contains references to two adapted quantity components, one for systolic and one for diastolic pressure. The label is fixed, making the semantics of this structure immutable.

```
<xsd:complexType name="mc-clj5x1g8f000008l09j7f6c3d">
 <xsd:complexContent>
  <xsd:restriction base="sdc4:ClusterType">
   <xsd:sequence>
```

```
      <xsd:element name="label" type="xsd:string" fixed="Blood Pressure" />
      <!-- Reference to the adapter for the Systolic component -->
      <xsd:element ref="sdc4:ms-..." minOccurs="1" maxOccurs="1" />
      <!-- Reference to the adapter for the Diastolic component -->
      <xsd:element ref="sdc4:ms-..." minOccurs="1" maxOccurs="1" />
     </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

- **XdAdapterType**: A "leaf" node in a ClusterType structure. It acts as a wrapper that holds a single data-typed component (XdAnyType), adapting it for inclusion in a cluster's item list.
  Data Model Schema Example: This defines an adapter for the "Systolic" component. It restricts XdAdapterType and specifies that it must contain exactly one ms-... element that represents the systolic blood pressure value.
```
<xsd:complexType name="mc-lyp1zhdjr31il1qdhvhs828k">
  <xsd:complexContent>
    <xsd:restriction base="sdc4:XdAdapterType">
      <xsd:sequence>
        <!-- The adapter holds a single reference to the data component -->
        <xsd:element ref="sdc4:ms-clj5x2p4k000108l01a2b3c4d"/>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

6.2. Core Data Types (Xd* Types)

These types provide semantically rich representations for common data values.

- **XdStringType**: A general-purpose type for character strings.
  Data Model Schema Example: This creates a reusable component for a "Customer ID". It restricts XdStringType, fixes the label, and applies a regex pattern to the value, enforcing a specific format.
```
<xsd:complexType name="mc-clj5x4b1c000308l0i9j8k7l6">
  <xsd:complexContent>
    <xsd:restriction base="sdc4:XdStringType">
      <xsd:sequence>
        <xsd:element name="label" type="xsd:string" fixed="Customer ID"/>
        ...
        <xsd:element name="xdstring-value">
          <xsd:simpleType>
            <xsd:restriction base="xsd:string">
```

```
            <xsd:pattern value="CUST-[0-9]{5}"/>
          </xsd:restriction>
        </xsd:simpleType>
      </xsd:element>
      ...
    </xsd:sequence>
  </xsd:restriction>
 </xsd:complexContent>
</xsd:complexType>
```

- **XdQuantityType**: Represents a physical quantity with a decimal value and units.
  Data Model Schema Example: This defines a "Systolic Blood Pressure" component. It restricts XdQuantityType, fixes the label, and requires the units to be "mmHg".

```
<xsd:complexType name="mc-clj5x2p4k000108l01a2b3c4d">
 <xsd:complexContent>
  <xsd:restriction base="sdc4:XdQuantityType">
   <xsd:sequence>
    <xsd:element name="label" type="xsd:string" fixed="Systolic"/>
    ...
    <xsd:element name="xdquantity-value" type="xsd:decimal"/>
    <xsd:element name="xdquantity-units">
     <xsd:complexType>
      <xsd:complexContent>
       <xsd:restriction base="sdc4:XdStringType">
        <xsd:sequence>
         <xsd:element name="xdstring-value" type="xsd:string" fixed="mmHg"/>
        </xsd:sequence>
       </xsd:restriction>
      </xsd:complexContent>
     </xsd:complexType>
    </xsd:element>
   </xsd:sequence>
  </xsd:restriction>
 </xsd:complexContent>
</xsd:complexType>
```

- **XdTemporalType**: A flexible type for representing date and time.
  Data Model Schema Example: This defines a "Measurement Time" component. It restricts XdTemporalType to allow *only* an xsd:dateTime value, ensuring full temporal precision.

```
<xsd:complexType name="mc-clj5x3a9b000208l0e5f6g7h8">
 <xsd:complexContent>
  <xsd:restriction base="sdc4:XdTemporalType">
```

```
      <xsd:sequence>
        <xsd:element name="label" type="xsd:string" fixed="Measurement Time"/>
        ...
        <xsd:choice>
          <xsd:element name="xdtemporal-datetime" type="xsd:dateTime"/>
        </xsd:choice>
      </xsd:sequence>
    </xsd:restriction>
  </xsd:complexContent>
</xsd:complexType>
```

- **Other Xd* Types**: The same restriction pattern applies to all other Xd* types, allowing modelers to create specific, reusable components like XdLinkType with a fixed relation, XdCountType with fixed units, or XdOrdinalType with a fixed set of symbols.

6.3. Governance and Provenance Types

- **PartyType**: Describes a person, organization, or device involved with the data.
- **ParticipationType**: Models the involvement of a Party in an activity.
- **AuditType**: Records a single audit event.
- **AttestationType**: Provides a formal record of attestation.

6.4. Data Quality and Exceptional Value Types

- **ExceptionalValueType**: The abstract base for a set of "null flavor" types.
- **Specific Exceptional Values**: The reference model provides pre-defined restrictions of ExceptionalValueType for common cases (NIType, UNKType, MSKType, NAType, INVType, OTHType).

7. References

- **Semantic Data Charter Reference Model Schema:** https://semanticdatacharter.com/ns/sdc4/sdc4.xsd
- **XML Schema Part 2: Datatypes Second Edition:** https://www.w3.org/TR/xmlschema-2/
- **Web Ontology Language (OWL):** https://www.w3.org/OWL/
- **RFC 2119:** https://www.rfc-editor.org/rfc/rfc2119
- **Dublin Core Metadata Element Set:** https://www.dublincore.org/specifications/dublin-core/dces/