ISO/TC 295

Date: 2022-~~10-12~~11-17

Exchange formats for the audit data collection standard: XML and JSON

~~For Publication~~

~~**Warning for WDs and CDs**~~

~~*A model manuscript of a draft International Standard (known as "The Rice Model") is available at*~~

ISO copyright office

CP 401 • Ch. de Blandonnet 8

CH-1214 Vernier, Geneva

Phone: +41 22 749 01 11

Email: copyright@iso.org

Website: www.iso.orgwww.iso.org

Published in Switzerland

Contents

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 295, *Audit* ~~Data Services~~*data services*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Accounting and enterprise resource planning (ERP) software packages are widely used in businesses and by various government organizations to manage and track business processes, post transactions and produce financial reports. Because of the nature of the information contained within the ERP systems, the data ~~is~~are also leveraged by internal and external auditors to assess the business controls, processes and financial reporting. There are numerous ERP packages that are used by businesses and government organizations, which can vary greatly in design (e.g~~.~~, interfaces, data content, data formats, operational reports, management reports, financial reports). These and other design differences present challenges in the collection of data for auditing supervision management purposes.

This document concerns the specification of technical exchange formats as output file formats for the functional content defined in ISO 21378.

Exchange formats for audit data collection standard: XML and JSON

1 Scope

ISO 21378, the audit data collection standard (ADCS)), defines the functional requirements for exchanging audit data in flat file format. This document concerns the specification of technical exchange formats in extensible markup language (XML), javascriptJavaScript object notation (JSON) and flat filecomma-separated values (CSV) as output file formats for the functional content defined in ISO 21378.

This document also contains the following schemas and sample files for tables in ISO 21378.

— XML schema,;

— XML sample file,;

— JSON schema,;

— JSON sample file.

To keep the three exchange formats (XML, JSON, and CSV) consistent, this document also specifies how to use the technical solution in the CSV format.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 21378:2019, *Audit data collection*

ISO 8601-1, *Date and time — Representations for information interchange — Part 1: Basic Rules rules*

ISO/IEC 21778:2017, *Information technology — The JSON data interchange syntax*

W3C *Extensible Markup Language (XML) 1.0 (Fifth Edition).* Available at https://www.w3.org/TR/2008/REC-xml-20081126/https://www.w3.org/TR/2008/REC-xml-20081126/

W3C *XML Schema Part 1: Structures Second Edition*. Available at https://www.w3.org/TR/2004/REC-xmlschema-1-20041028/https://www.w3.org/TR/2004/REC-xmlschema-1-20041028/

W3C *XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes*. Available at https://www.w3.org/TR/2012/REC-xmlschema11-2-20120405/https://www.w3.org/TR/2012/REC-xmlschema11-2-20120405/

JSON Schema: A Media Type for Describing JSON Documents. Available at https://json-schema.org/draft/2020-12/json-schema-core.htmlhttps://json-schema.org/draft/2020-12/json-schema-core.html

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
data
reinterpretable representation of information in a formalized manner suitable for communication, interpretation or processing

[SOURCE: ISO/IEC 11179-1:2015(en), 3.2.6], 3.2.6, modified — Notes to entry have been deleted.]

3.2
data element
basic unit of identifiable and definable *data* (3.1)

[SOURCE: ISO 2146:2010, 3.4, modified — The admitted term "element" has been deleted.]

3.3
data file
collection of *data* (3.1) records having a homogeneous structure

[SOURCE: ISO 14825:201121378:2019, 3.1.2, modified — The word "related" and the note to entry have been deleted.]4]

3.4
data structure
framework comprising a number of *data elements* (3.2) in a prescribed form

[SOURCE: ISO 21007-1:2005, 2.16, modified —"— The word "element" has been deleted from the term.]

3.5
syntax
set of rules, principles and processes that govern the *data structure* (3.4)

3.6
data model
graphical and/or lexical representation of *data* (3.1), specifying their properties, structure, and inter-relationships

[SOURCE: ISO/IEC 11179-1:2015~~(en),~~ 3.2.7]

3.7
entity
group of *data elements* (3.2) describing an object

Note 1 to entry: ~~Equivalent~~It is equivalent to "~~Object Class~~object class" in ISO/IEC 11179-1~~:2015~~.

3.8
relation
relationship between two *entities* (3.7)

3.9
attribute
data element (3.2) describing an object

Note 1 to entry: ~~Equivalent~~It is equivalent to "~~Property~~property" in ISO/IEC 11179-1~~:2015~~.

3.10
domain
set of properties to define the value space of *attributes* (3.9)

Note 1 to entry: ~~contain~~A domain contains *code lists* (3.11) and *code values*~~.~~ ~~Equivalent~~ (3.12). It is equivalent to "~~Representation~~representation" in ISO/IEC 11179-1~~:2015~~.

3.11
code list
standardized list of *code values* (3.12) with a common scope

3.12
code value
one value from a *code list* (3.11)

4 Exchange formats

4.1 General

The data format is the carrier of data exchange between auditor and auditee. So, it is necessary to make an agreement on data format between the two sides in data exchange. There are multiple options for output data formats.

In the case of ADCS, three exchange formats are specified:

— XML data files defined by XML schema specification (W3C);

— JSON data files defined by JSON schema specification (json-schema.org);

— CSV data files.

These three formats shall contain the same functional ADCS content specification in accordance with ISO 21378 but differ on the technical level only.

This means that these exchange formats are convertible from one to the other. For instance, ~~convertingit~~ is possible to convert XML data files to CSV data files or CSV data files to JSON data files.

To generate the XML and JSON schemas in a consistent way, a data model is set up from ISO 21378 according to ISO/IEC 11179-1~~:2015~~. This data model is used to create consistent exchange format specifications.

Each ADCS table is modelled into a "Parent Entity" and reusable data groups within a table are modelled into "Child Entities", ~~e.g.,~~for example, "Physical Address", "Billing Address", "Tax", "Created", "Modified", "Posted~~" etc.,"~~, completely in line with ISO 21378. This way guarantees that all reusable groups that occur in multiple tables are defined in the same way.

As the data model is not in scope of this document, it is only documented for reference purposes in Annex D.

The various technical specifications are explained in this document~~,~~; and the additional packaging and communication agreements are defined.

4.2 XML

4.2.1 General

This subclause describes how to exchange the ADCS tables using XML data files.

XML data files shall be created in accordance with the syntax specifications written in W3C Extensible Markup Language (XML) 1.0 (Fifth Edition) and also in accordance with the functional requirements written in ISO 21378.

4.2.2 Technical requirements

4.2.2.1 General

Only one ADCS table, with a number of lines, shall be sent in one XML data file.

The filename shall comply with the rules described in ~~Subclause~~ 4.6.5.

Each ADCS table has its own XML schema to define its XML file. When there are 71 tables, there are 71 XML schemas.

Such an XML data file contains (schematic):

```
<?xml version='1.0' encoding='UTF-8'?>
<root-tag
    xsi:schemaLocation='name-space xml-schema-name.xsd'
    xmlns='name-space'
    xmlns:xsi='http://www.w3.org/2001/XMLSchema-instance'>
    <table-line-tag>
        <data-element-tag>value</data-element-tag>
        <group-tag>
            <data-element-tag>value</data-element-tag>
        </group-tag>
    </table-line-tag>
</root-tag>
```

~~Italics printed~~

Italicized text shall be replaced by the definitions in this document.

~~4.2.2.1~~4.2.2.2~~4.2.2.1~~ Root tag

The root tag equals "Adc" followed by the ADCS table name like: “AdcApAdjustments”.

~~4.2.2.2~~4.2.2.3~~4.2.2.2~~ Target name space definition

The target name space equals "http://schemas.iso.org/AdcsML/Messages/" followed by the message name and message version, like "http://schemas.iso.org/AdcsML/Messages/AdcApAdjustments-v1".

~~4.2.2.3~~4.2.2.4~~4.2.2.3~~ XML tags

Table-line-tag, group-tag and data-element-tag are XML tags which are defined in the related XML schema.

XML tags are in the first step derived from the full data element names, and in the second step shortened in a consistent manner, according to an abbreviation list (see Annex A).

Because auditors and auditees usually exchange bulk files according to ADCS, abbreviated tags can reduce the file sizes by 30~~% -~~ % to 40~~%.~~ %.

Full data element names and abbreviated XML tags are together specified in ADCS to XML mapping tables (see explanation and example of mapping tables in ~~Subclause~~ 4.5).

In XML schemas the full data element names are documented as annotations.

This means that programmers can have easy access to the full data element names, either from the mapping tables or from the XML schemas.

~~4.2.2.4~~4.2.2.5~~4.2.2.4~~ Optional fields

Optional fields and optional groups of fields can be omitted from the XML data file, only if these data fields are not available in the source system that delivers the audit data.

~~4.2.2.5~~4.2.2.6~~4.2.2.5~~ Repeating groups

Repeating groups like “Tax” can occur up to the maximum number that is defined in the XML schema.

~~4.2.2.6~~4.2.2.7~~4.2.2.6~~ Special characters to be escaped

The following special characters shall be escaped in XML data files as follows:

< (less than) to be replaced with <

> (greater than) to be replaced with >

& (ampersand) to be replaced with &

‘ (apostrophe) to be replaced with '

“ (quotation mark) to be replaced with "

4.2.3 XML schema

4.2.3.1 ~~4.2.3.1~~ General

Each ADCS file is technically specified by an XML schema and can also be validated by using that schema.

XML schema files shall be created in accordance with the syntax specifications written in W3C Extensible Markup Language (XML) 1.0 (Fifth Edition), W3C XML Schema Part 1: Structures Second Edition, W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes and also in accordance with the functional requirements written in ISO 21378.

Table 1 shows ISO 21378~~:2019~~ representation formats and the equivalents used in XML schema.

Table 1 — Representation specification in XML schema

ISO 21378~~:2019~~ representation	XML schema representation
%ns	<xsd:restriction base="xsd:string"> <xsd:maxLength value="n"/> </xsd:restriction>
%nc	<xsd:restriction base="xsd:string"> <xsd:length value="n"/> </xsd:restriction>
%m.nf	<xsd:restriction base="xsd:decimal"> <xsd:totalDigits value="m"/> <xsd:fractionDigits value="n"/> </xsd:restriction>
%nd	<xsd:restriction base="xsd:integer"> <xsd:totalDigits value="n"/> </xsd:restriction>

Table 2 shows ISO 21378~~:2019~~ datatypes and the equivalents used in XML schema.

Table 2 — Datatypes in XML schema

ISO 21378~~:2019~~ datatype	XML schema datatype
Date	<xsd:restriction base="xsd:date">
Time	<xsd:restriction base="xsd:time">
String	<xsd:restriction base="xsd:string">
Decimal	<xsd:restriction base="xsd:decimal">
Integer	<xsd:restriction base="xsd:integer">
Boolean	<xsd:restriction base="xsd:boolean"/>

All schema files are bundled in one official ISO delivery package. Another development package is also available.

a) Official ISO delivery package: This package contains the schemas established ~~as~~in this document and ISO 21378 for exchanging the audit data.

b) Development package: This package contains the schemas that are used to develop new versions of the existing schemas and for the development of extensions.

The electronic version of these XML schemas can be downloaded from the URL which can be found in Annex E.

4.2.3.2 ~~4.2.3.2~~ Official ISO delivery package

All schema files in this package have one and the same version number, which can be found in the schema file itself.

The first delivery of the package has version number 1. After a certain period, if there are sufficient reasons for this, a new package of schemas will be delivered with a successor version number, for example, version number 2.

This package contains:

a) All ADCS table schemas. The name of the xml schema file equals “Adc” followed by the ADCS table name and then followed by “.xsd”, such as "AdcApAdjustments.xsd".

b) One data types schema named “Adc_DataTypes.xsd”. This schema is imported into all ADCS table schemas.

A list of these XML schemas can be found in Annex B.

~~Example:~~ EXAMPLE 1 Part of “AdcApAdjustments.xsd”:


```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML Definition of message hierarchy
-->
<!-- Message:         ADC AP Adjustments v1.0
-->
<!-- Generated at:    25-05-2022 16:53:41
-->
<!-- Organisation:    This schema is part of ISO 21377 Technical
Specification -->

<xsd:schema
xmlns:str="http://schemas.iso.org/AdcsML/Messages/AdcApAdjustments-v1"
                 xmlns:cdt="http://schemas.iso.org/AdcsML/Adc_DataTypes-
v1"
                 xmlns:xsd="http://www.w3.org/2001/XMLSchema"

targetNamespace="http://schemas.iso.org/AdcsML/Messages/AdcApAdjustments-
v1"
                 elementFormDefault="qualified"
attributeFormDefault="unqualified" version="1.0">

   <xsd:import namespace="http://schemas.iso.org/AdcsML/Adc_DataTypes-v1"
schemaLocation="Adc_DataTypes.xsd"/>
   <xsd:element name="AdcApAdjustments">
     <xsd:annotation>
        <xsd:documentation>MESSAGE</xsd:documentation>
     </xsd:annotation>
```

```
      <xsd:complexType>
         <xsd:sequence>
            <xsd:element name="ApAdj" minOccurs="0" maxOccurs="unbounded">
               <xsd:annotation>
                 <xsd:documentation>AP ADJUSTMENT</xsd:documentation>
               </xsd:annotation>
               <xsd:complexType>
                 <xsd:sequence>

                     <xsd:element name="AdjId" type="cdt:TypId100"
minOccurs="1" maxOccurs="1">
                        <xsd:annotation>
                           <xsd:documentation>Adjustment
ID</xsd:documentation>
                        </xsd:annotation>
                     </xsd:element>

                     <xsd:element name="AdjNr" type="cdt:TypText100"
minOccurs="0" maxOccurs="1">
                        <xsd:annotation>
                           <xsd:documentation>Adjustment
Number</xsd:documentation>
                        </xsd:annotation>
                     </xsd:element>

                 </xsd:sequence>
               </xsd:complexType>
            </xsd:element>
         </xsd:sequence>
      </xsd:complexType>
   </xsd:element>
</xsd:schema>
```

~~Example:~~

EXAMPLE 2 Part of Adc_DataTypes.xsd:

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML Definition of elementary types
-->
<!-- Datamodel:     Audit Data Collection Datamodel 1, version HEAD
-->
<!-- Generated at:  25-05-2022 16:53:41
-->
<!-- Organisation:  This schema is part of ISO 21377 Technical
Specification   -->

<xsd:schema xmlns:cdt="http://schemas.iso.org/AdcsML/Adc_DataTypes-v1"
            xmlns:xsd="http://www.w3.org/2001/XMLSchema"
            targetNamespace="http://schemas.iso.org/AdcsML/Adc_DataTypes-
v1"
```

```
elementFormDefault="qualified"attributeFormDefault="unqualified"
version="1.0">

<!-- IDENTIFIER 100: FORMAT: AN..100 -->
<xsd:simpleType name="TypId100">
   <xsd:restriction base="xsd:string">
      <xsd:maxLength value="100"/>
      </xsd:restriction>
</xsd:simpleType>

<!-- TEXT 100: FORMAT: AN..100 -->
<xsd:simpleType name="TypText100">
   <xsd:restriction base="xsd:string">
         <xsd:maxLength value="100"/>
       </xsd:restriction>
   </xsd:simpleType>

</xsd:schema>
```

~~4.2.3.4~~4.2.3.3~~4.2.3.3~~ Development package

This package contains the schemas that are used to develop new versions of the existing schemas and for the development of extensions.

In this package each schema has its own version number, which is included in the version attribute and which is added to the name of the XML schema file. In the first delivery of this package, all schemas have version 1.0 like "AdcApAdjustments-v1.0.xsd". When a schema changes, the minor version number is incremented each time (e.g~~.~~, version number changes from 1.0 to 1.1).

In this package, all schemas contain their own data type definitions, so that they can be further developed independently and there is no common data type schema.

After a certain period of time, if there are sufficient reasons to do this, a new official ISO delivery package will be delivered with the next successor major version number, for example, version number 2 (see ~~Subclause~~ 4.2.3.2~~),~~); and the new common data type schema (Adc_DataTypes.xsd) will be delivered.

Also a new development package will be delivered, in which all schemas will have version 2.0.

~~Example:~~ EXAMPLE Part of "AdcApAdjustments-v1.1.xsd":

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- XML Definition of message hierarchy
-->
<!-- Message:  ADC AP Adjustments v1.1
-->
<!-- Generated at:    25-05-2022 16:53:41
-->
<!-- Organisation:  This schema is part of ISO 21377 Technical
Specification     -->

<xsd:schema
xmlns:str="http://schemas.iso.org/AdcsML/Messages/AdcApAdjustments-v1"
```