ISO/TC 295/WG 1 "Audit Data Collection for Non-financial Enterprises"
Convenorship: **SAC**
Convenor: **WANG Wenyu Ms**

Collated Comments for the proejct Exchange formats for the Audit Data Collection Standard XML and JSON

Document type	Related content	Document date	Expected action
Meeting / Working documents for discussion	Project: ISO/WD TS 5409 Ballot: N525 (restricted access)	2021-05-10	**INFO**

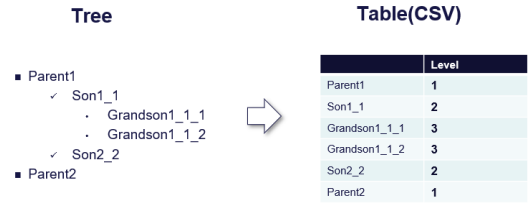
Template for comments and secretariat observations

Date:2021-05-01	Document: N525	Project: ISO/WD TS 5409 Exchange formats for the Audit Data Collection Standard: XML and JSON

SON

MB/ NC[1]	Line number	Clause/ Subclause	Paragraph/ Figure/Table	Type of comment[2]	Comments	Proposed change	Observations of the secretariat
CN-Lyu-001		3		ed	The number 3.2 is skipped	Modify the term number to consecutive number.	
CN-Lyu-002		3.6, 3.10		ed	The referenced number of data structures is wrong	The referenced number of data structures shall be 3.7.	
CN-Lyu-003		4.1	Paragraph 5	Te	What is the relationship between data model and current standards?	Data model is not included in the scope of this project.	
CN-Lyu-004		4.1	Paragraph 5	ed	To generate the schema's in a concistent way.	schema's shall be schemas, concistent shall be consistent. The mistakes appear several times.	
CN-Lyu-005		4.1	Paragraph 5	Te	For the functional content, see the specifications of the "Audit Data Collection Data Model". What is the Audit Data Collection Data Model?		
CN-Lyu-006		4.2.2.1		ed	Used table names are taken from the ISO 21378:2019 specifications.	The specifications shall be standard.	
CN-Shinewing-007		4.2.3		ge	In addition to viewing the final report, auditors often database or interface data directly, the nodes of the current XML file are defined by a short name, it is difficult for auditors to read xml files directly.	Auditors want the software to have a set of tools that automatically combine schema with xml files for full name presentation	
CN-Shinewing-		4.2.4		ge	Audit data collection involves different reports, generally exported from different systems, auditors want to be able to easily see which software the data	Auditors want to be able to define data sources in one place, such as from a software vendor	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

MB/ NC[1]	Line number	Clause/ Subclause	Paragraph/ Figure/Table	Type of comment[2]	Comments	Proposed change	Observations of the secretariat
008					is exported from		
CN-Lyu-009		4.4		Te	The specification for csv in this clause is not the same with ISO 21378:2019.	Adding the content to show this specification for csv is for conversion between csv and XML, JSON, and it is optional. If users don't need conversion to XML, JSON, it's OK to follow ISO 21378:2019.	
CN-Shinewing-010		4.4.1		ge	CSV files are in the form of tables and do not yet know how to define the data for tree structures	Add a field: level, and can define the tree structure through this field, e.g. 	
JISC-011		4.4.2		te	"Audit Data Collection Data Model" in 4.4.2 Technical guideline is not specified clearly. There is no requirement about what the Audit Data Collection Data Model is. Mapping to each physical file syntax SHALL be defined based on the Data Model.	We need both a "data dictionary" and a "common data structure" which documented and based on the CCTS. Since CCTS is not supporting all the features of ISO 21378:2019 (ADCS), we want to expand it with things like primary keys and references to the primary keys, as it is included in tables in the ADCS. Extend the identifier business information entity (IDBIE) and the related business information entity (RLBIE) to describe technical specifications based on CCTS. Next, write the Data Model based on the extended CCTS.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

SON

MB/ NC[1]	Line number	Clause/ Subclause	Paragraph/ Figure/Table	Type of comment[2]	Comments	Proposed change	Observations of the secretariat
CN-Lyu-012		4.4.2.3		ed	Numerical data can contain de dot character (.) as decimal seperator.	De dot shall be the dot? Seperator shall be separator.	
CN-Lyu-013		4.4.4	Figure 1	Te	This is an example of a CSV file, but the use of Excel to open the file. In this way, some technical details of CSV are hidden, such as separator (,), escape character (\), which is not conducive to readers' understanding.	Please open this file with a text reader.	
JISC-014		4.5		te	The mapping table is defined without a data model specification and definition. We cannot map CSV directly to XML, JSON, XBRL, and other formats. In addition to supporting ISO 21378, technical specifications are needed to support data model definitions and data model-to-physical file mappings to support new application areas for the following new project; ISO/WD 5401 Audit Data Collection — Customs and Indirect Taxes Extension, ISO/WD 5405 Audit Data Collection Extension: Government Regulated Financial Reports and Payroll.	ditto	
CN-Lyu-015		4.5		ed	The ISO 21378-2019 shall be ISO 21378:2019		

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

SON

MB/ NC[1]	Line number	Clause/ Subclause	Paragraph/ Figure/Table	Type of comment[2]	Comments	Proposed change	Observations of the secretariat
CN-Lyu-016		4.6.9	Paragraph 2	ed	It is not allowed to add own data elements into an ADC file	ADC shall be changed to ADCS. Similar change to other ADC.	
CN-Lyu-017		2, Bibliography		ed	In ISO DP2, Clause 15 requires that referenced documents listed in Normative references are not numbered. Some referenced documents in Clause 2 are duplicated with Bibliography.	Please delete the a, b, c, etc. used for numbering. Please avoid the repeat referenced documents in clause 2 with Bibliography.	
CN-Lyu-018		Foreword	Paragraph 6	ed	This document was prepared by Project Committee ISO/PC 295, Audit data collection.	This document was prepared by Technical Committee ISO/TC 295, Audit data services.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

MB/ NC[1]	Line number	Clause/ Subclause	Paragraph/ Figure/Table	Type of comment[2]	Comments	Propo

File: N525_XIAO, Ziwei Mr.docx

About the abbreviation problem: I am in view of the previously released standard 21378 has been through the consensus of various countries, to determine the naming rules of relevant data elements and other content, it is not recommended to abbreviate.

Main reasons:
One is that the original rules have been very mature, and applied in other format files, if the adjustment in XML, will lead to the invalidation of the previous results, but also bring the chain modification problem, if not to modify the naming rules of other formats, there will be a confusion of multiple naming rules under the same big standard.

Second, the scientific rationality after the abbreviation needs to be re-identified, which may greatly increase the workload.

Thirdly, regarding to the extent to which the abbreviation can bring performance optimization, the need for objective assessment, such as no fundamental solution to the file size problem, then the need to modify is not strong enough. Moreover, the size of the field names that play a role in the header of an XML file is very small compared to the size of the entire file, abbreviation is not fundamental solution to the file size problem.

Finally, I want to provide a scenario that we encountered when merging tables from various realms or locales, having two different formats, one abbreviated and the other not abbreviated, can cause a lot of trouble in merging tables, so it is recommended to choose a format that is widely agreed upon based on the previous good results.

In addition, it is necessary to consider the relationship between the work related to the CSV format in this project and the relevant conventions in ISO 21378:2019, but not related to CSV. In addition, we should carefully consider what tools and technical route to choose, and recommend testing relevant tools and carrying out data modeling work.

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

N525_LV, Tianyang Mr.doc: Collation successful

N525_Sambuichi, Nobuyuki Mr.doc: Collation successful

Collation of files was successful. Number of collated files: 3

SELECTED (number of files): 3

PASSED TEST (number of files conformed to CCT table model): 2

FAILED TEST (number of files conformed to CCT table model): 0

N525_XIAO, Ziwei Mr.docx: Collation successful

PASSED OTHER FILES (number of files to be collated at the end of the result file not conformed to CCT table model): 1

CCT - Version 2020.1