eForms – Future of data quality

Bertrand Lorentz, Publications Office
2020-12-01

Agenda

- Improving data quality
- Challenges of eForms
- Business rules
- Central Validation Service
- Rules implementation

Improving data quality

- eForms regulation :
  "In order to ... increase data reliability ... standard forms should be established"

- Better quality → more useful

- Data quality must be verified

Challenges of eForms (1)

- eForms regulation has requirements/restrictions
  - Data type (indicator, code, date, etc.)
  - Mandatory/optional presence

Level	ID	Name	Description	Data type	Planning									Competition															DAP			Result									Con. Mod.		
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
++	BT-757	Notice Version	The version of the notice. This helps, for example, to avoid errors caused by multiple change notices sent at a similar time.	Identifier	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
++	BT-01	Procedure Legal Basis	The legal basis (e.g. a European Directive or Regulation, national law) under which the procurement procedure takes place.	Code	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
++	BT-03	Form Type	The type of form according to procurement legislation.	Code	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
++	BT-02	Notice Type	The type of notice according to procurement legislation.	Code	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
++	BT-05	Notice Dispatch Date	The date and time when the notice has been sent to TED.	Date	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
++	BT-738	Notice Publication Date Preferred	The preferred date of publication of the notice on TED (e.g. to avoid publication during a national holiday).	Date	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
++	BT-702	Notice Official Language	The language(s) in which this notice is officially available. These linguistic versions are equally legally valid.	Code	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M

Challenges of eForms (2)

- UBL schemas are very permissive
  - For the presence of elements
  - For some values (code lists not in the schema)

```
<ContractNotice>
    <cbc:ContractFolderID></cbc:ContractFolderID>
    <cac:ContractingParty>
        <cac:Party></cac:Party>
    </cac:ContractingParty>
    <cac:ProcurementProject>
    </cac:ProcurementProject>
    <cac:ProcurementProjectLot>
        <cbc:ID></cbc:ID>
    </cac:ProcurementProjectLot>
</ContractNotice>
```

Challenges of eForms (3)

- Structure of Business Groups is different than structure of UBL

- Some Business Terms are generic/high level
  - BT-500 "Organisation Name"
    - Buyer name
    - Winner name
    - . . .
  - BG-2 Purpose / BT-21 Title
    - Title of the procurement
    - Title of lot
    - . . .

- From 282 Business Terms to > 1200 fields

Business rules

- Legal framework (public procurement directives) and "business" stay the same

- Examples of business rules

  "The total value of the lot cannot be higher than the total value of the procurement"

  "Conditions for opening of tenders are mandatory if open procedure"

New rules

- Consistency with previous notices in the same procedure

- Change notice instead of corrigendum

Central Validation Service (CVS)

- Single point of truth :
  Is this notice XML valid for publication on TED ?
- Only for eForms notices
- Used by all our systems, including eNotices2
- Available for eSenders as REST API

- Validates
  - XML is well formed
  - XML respects the schema
  - Schematron rules are followed

Central Validation Service API

- 1 operation: get validation report for a notice

- Input
  - XML notice
  - Desired language for the messages in the validation report
  - Name of the ruleset to use
    - Optional, defaults to the current ruleset
    - For example to validate against upcoming version of rules
  - Validation mode: static or dynamic
    - Optional, defaults to "dynamic"
    - "static" does not use information outside of the notice

Central Validation Service API

■ Output

- Validation report in SVRL (Schematron Validation Report Language)
    - Standardized XML language
    - Defined together with Schematron : ISO 19757 Annex D
    - For each failure
        - Exact location that triggered the failure
        - Test that failed
        - Text of the message

Schematron rules

Simplified example

```
<rule context="/*/cac:ProcurementProjectLot/cac:ProcurementProject
[/*/cbc:NoticeTypeCode = 'cn-standard']
[/*/cbc:RegulatoryDomain = '32014L0024']">

    <assert role="error" test="count(cbc:Name) > 0">
        The element "BT-21 - Title" must be present for this type of
notice.
    </assert>

    <assert role="error" test="count(cbc:Description) > 0">
        The element "BT-24 - Description" must be present for this type
of notice.
    </assert>

</rule>
```

SVRL Validation Report

Simplified example

```
<svrl:failed-assert
    location="/ContractNotice/cac:ProcurementProjectLot[2]/cac:Procur
ementProject"
    role="error"
    test="count(cbc:Description) > 0">

        <svrl:text>The element "BT-24 - Description" must be present
for this type of notice.</svrl:text>

</svrl:failed-assert>
```

Current status

- First iteration for all types of rules
  - Over 16 000 "rule"
  - Over 34 000 "assert"
  - On average, over 400 "assert" are checked on a notice

- Working on organisation and management of those rules

Questions ? Ideas ?

For any comments/suggestions, please contact TED-EVENTS@publications.europa.eu, specifying the subject "Questions for eSenders seminar – eForms - future of data quality".