Publications Office of the European Union

Law | Data | Publications

TED DATA QUALITY ASSISTANT

Karl Ferrand & Ana-Maria Babaligea

TED-together, 25 November 2026

TED Data Quality Assistant:
AI-assisted content check for EU procurement notices

Context

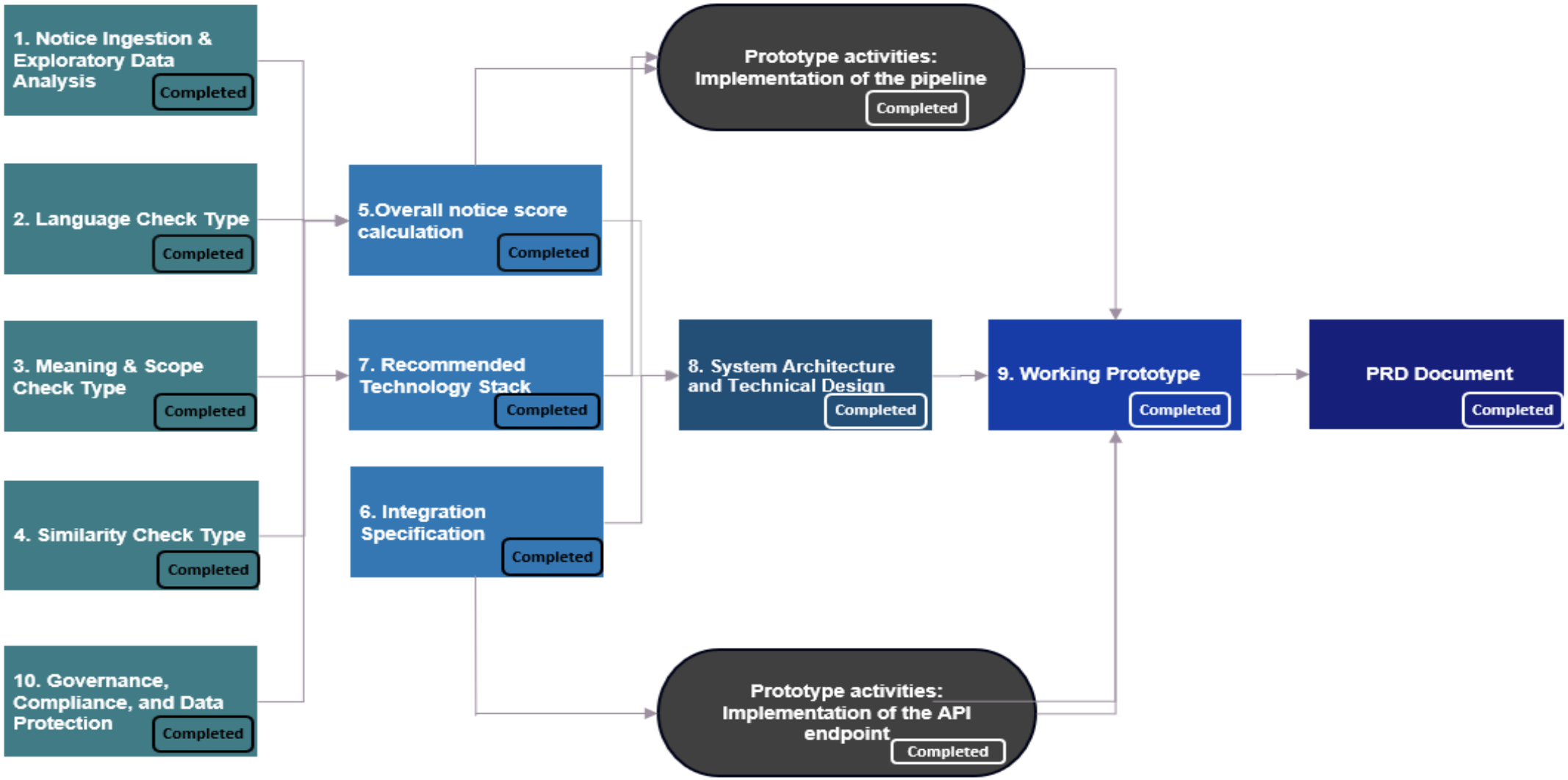
- Manual validation of every TED notice ended with eForms implementation
- Automatic validation by Central Validation Service (CVS) does not check text fields or assess content and meaning

Scope

TED Data Quality Assistant

- After CVS validation, it will check every notice submitted for publication on the TED Website
- it will generate a confidence score for each notice with location and explanation
- type of checks performed: language, meaning and scope on the core fields (titles, descriptions, additional information, organisation name)
- notices with low confidence score will be manually reviewed by OP and approved/rejected for publication (“human in the loop”)
- the review results will be captured and used to update the reference dataset

Proof of concept (POC) - work overview



TED Data Quality Assistant – POC conclusions

- **Feasible:** AI-assisted quality checking technically achievable
- **Proven scope:** 2 checks (Language, Meaning & Scope) on 4 core text fields
- **Advisory only – human in the loop:** TEDDQA flags and scores notices but never blocks. Final decision belongs to OP reviewer
- **Recommended model: Mistral Medium (mistral-medium-2508):** best benchmarking results, right size/cost balance
- **Hosting: AI@EC – European Commission internal sovereign platform**
- **Limitations:**

  - false positives/negatives

  - language variation

Next steps

- Product Requirements Document – finalisation and ordering
- Development phase and testing
- Implementation of DQA module in production with AI@EC platform
- Integration of DQA in the daily workflow of the internal OP application (TED Monitor)
- Include the current two CVS “lawfulness” checks
- Additional "processingStatus" (for notices in Submitted status)
- Inform buyer of rejection reasons
- Target in production by end 2027
- No public API is currently planned for DQA
- Checks will be documented and shared

TEDDQA
DEMO

Demo - Prototype pipeline

Notice extraction

- Parse XML content
- Load field mappings
- Extract notice-level metadata
- Extract configured notice fields

Check types

- Language
- Meaning and scope
  - Build language and meaning and scope check requests
  - Detect notice field languages
  - Parse check responses
  - Build notice field results
  - Build overall result

Overall score calculation

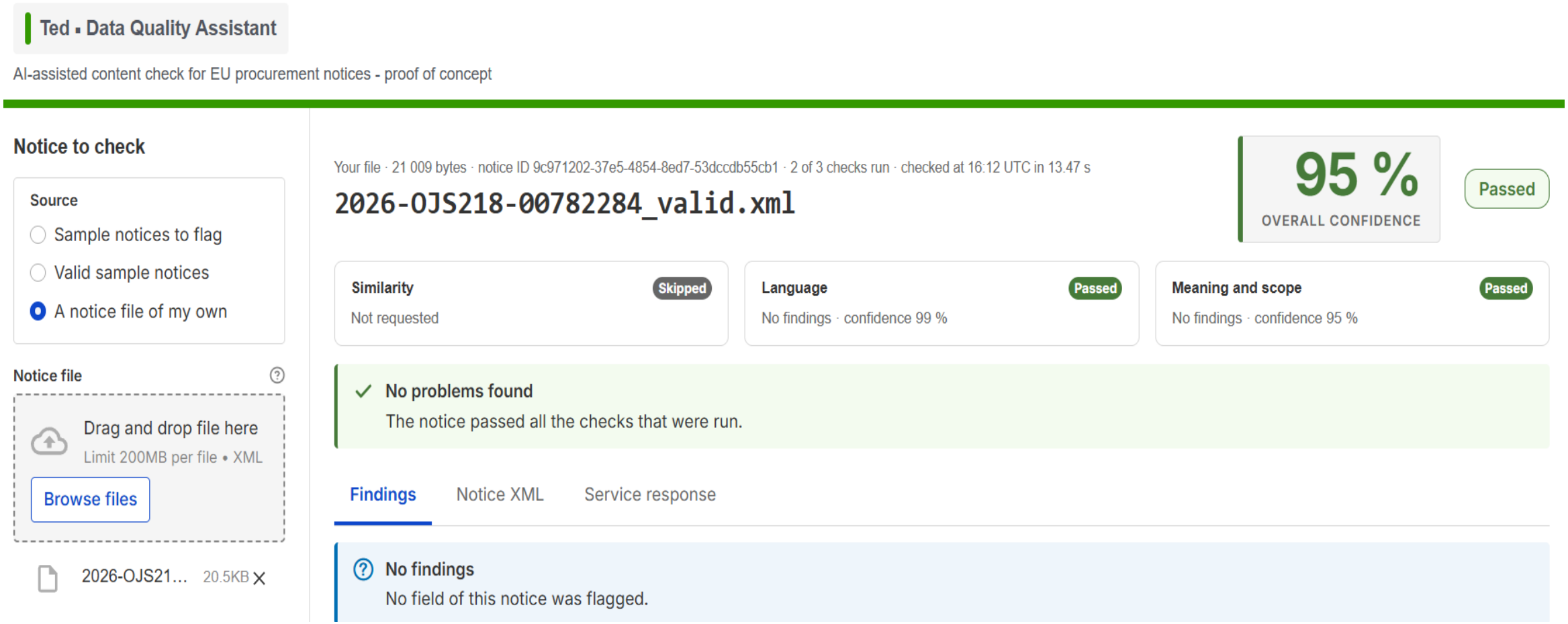
- Aggregate check scores
- Produce structured validation result

Output

Structured JSON result with per-check scores and field level findings

Demo

1. Valid notice



Demo

2. Non - Valid notice per language check type – wrong language

Wrong language · 33 565 bytes · notice ID 1eb71519-fbf0-4f36-a1d7-c05b694de36c · 2 of 3 checks run · checked at 16:42 UTC in 16.66 s

1eb71519-fbf0-4f36-a1d7-c05b694de36c-01.xml

1 %
OVERALL CONFIDENCE

Flagged

Similarity	Language	Meaning and scope
Skipped	Flagged	Passed
Not requested	5 findings · confidence 1 %	No findings · confidence 95 %

5 findings to review
They are listed under Findings.

Demo

2. Non - Valid notice per language check type – wrong languages

5 findings to review
They are listed under Findings.

Findings 5 | Notice XML | Service response

Showing the findings of all checks. Select a check above to show only its findings. Select a column heading to sort, and a row to see the finding below. Previous and next follow the order in which the service returned the findings.

Check	Finding	Field	Message	Confidence
Language	WRONG_LANGUAGE	BT-21-Part	The text is written in English, with characteristic English vocabulary, g	1 %
Language	WRONG_LANGUAGE	BT-21-Procedure	The text is written in English, with characteristic English vocabulary, g	1 %
Language	WRONG_LANGUAGE	BT-24-Part	The entire text is written in English with consistent English vocabulary	1 %
Language	WRONG_LANGUAGE	BT-24-Procedure	The entire text is written in English, with clear English vocabulary, grammar, and sentence structure.	1 %
Language	WRONG_LANGUAGE	BT-300-Procedure	The entire text is written in English with clear English vocabulary, grar	1 %

Demo

2. Non - Valid notice per language check type – wrong languages

Finding 4 of 5 · WRONG_LANGUAGE in BT-24-Procedure

The entire text is written in English, with clear English vocabulary, grammar, and sentence structure.

Check	Language
Confidence	1 %
Business entity	procedure (procedure)
XML location	/*/cac:ProcurementProject/cbc:Description

Show this field in the notice XML

```
274     <cbc:Description languageID="SPA">This PIN provides early information about the expected starting date
275     <cbc:ProcurementTypeCode listName="contract-nature">services</cbc:ProcurementTypeCode>
```

Demo

2. Non - Valid notice per language check type – wrong language


```
3 : {
  "code" : "WRONG_LANGUAGE"
  "message" :
  "The entire text is written in English, with clear English vocabulary, grammar, and sentence
  structure."
  "confidence_score" : 0.01
  "field_id" : "BT-24-Procedure"
  "field_type" : "Description"
  "xml_path" : "/*/cac:ProcurementProject/cbc:Description"
  "business_entity_id" : "procedure"
  "business_entity_label" : "procedure"
  "business_entity_instance_id" : NULL
  "evidence" : {
    "declared_language" : "SPA"
    "detected_languages" : [
      0 : {
        "language_code" : "ENG"
        "confidence" : 0.99
      }
    ]
    "language_check_label" : "wrong_language"
  }
```

Demo

3. Non - Valid notice per meaning and scope check type – test notices

With the word TEST · 19 942 bytes · notice ID 0f20ff45-38df-4a63-84c6-f96185cfa471 · 2 of 3 checks run · checked at 16:55 UTC in 9.29 s

2024-OJS016-00044677.xml

5 %
OVERALL CONFIDENCE

Flagged

Similarity	Meaning and scope	Language
Skipped	Flagged	Passed
Not requested	8 findings · confidence 5 %	No findings · confidence 60 %

8 findings to review
They are listed under Findings.

Demo

3. Non - Valid notice per meaning and scope check type – test notices

8 findings to review
They are listed under Findings.

Findings 8 | Notice XML | Service response

Showing the findings of all checks. Select a check above to show only its findings. Select a column heading to sort, and a row to see the finding below. Previous and next follow the order in which the service returned the findings.

Check	Finding	Field	Message
Meaning and scope	TESTING_CONTENT	BT-21-Lot	Text contains explicit test marker 'test' in a Lot Title field, indicating
Meaning and scope	TESTING_CONTENT	BT-21-Lot	Text contains explicit test marker 'test' in a Lot Title field, indicating
Meaning and scope	TESTING_CONTENT	BT-21-Procedure	Text contains explicit test marker 'test' indicating placeholder/dumr
Meaning and scope	TESTING_CONTENT	BT-24-Lot	The text 'test' is a clear placeholder/dummy entry, matching the def
Meaning and scope	TESTING_CONTENT	BT-24-Lot	The text 'test' is a clear placeholder/dummy entry, matching the definition of testing.
Meaning and scope	TESTING_CONTENT	BT-24-Procedure	The text 'test' is a clear placeholder/dummy entry, matching the def
Meaning and scope	TESTING_CONTENT	BT-300-Procedure	The text 'test' is a clear placeholder or dummy entry, matching the c
Meaning and scope	TESTING_CONTENT	BT-500-Organization-Company	Explicit test placeholder in an Organisation Name field.

Demo

3. Non - Valid notice per meaning and scope check type – test notices

Finding 6 of 8 · TESTING_CONTENT in BT-24-Procedure

The text 'test' is a clear placeholder/dummy entry, matching the definition of the testing category.

Check	Meaning and scope
Confidence	5 %
Business entity	procedure (procedure)
XML location	`/*/cac:ProcurementProject/cbc:Description`

Show this field in the notice XML

```
114      <cbc:Name languageID="ENG">CB - OJEU test 20/01</cbc:Name>
115      <cbc:Description languageID="ENG">test</cbc:Description>
```

Demo

3. Non - Valid notice per meaning and scope check type – test notices


```
6 : {
  "code" : "TESTING_CONTENT"
  "message" :
  "The text 'test' is a clear placeholder or dummy entry, matching the definition of the
  testing category."
  "confidence_score" : 0.05
  "field_id" : "BT-300-Procedure"
  "field_type" : "Additional Information"
  "xml_path" : "/*/cac:ProcurementProject/cbc:Note"
  "business_entity_id" : "procedure"
  "business_entity_label" : "procedure"
  "business_entity_instance_id" : NULL
  "evidence" : {
    "predictions" : [
      0 : {
        "label" : "testing"
        "confidence" : 0.95
      }
    ]
    "explanation" :
    "The text 'test' is a clear placeholder or dummy entry, matching the definition of
    the testing category."
  }
}
```

Demo

4. Non - Valid notice per meaning and scope check type – spam

Your file · 18 756 bytes · notice ID b5a71b8c-3363-4825-8f0e-b11ae244dfb2 · 2 of 3 checks run · checked at 16:34 UTC in 7.47 s

2026-OJS029-00098072_spam.xml

5 %
OVERALL CONFIDENCE

Flagged

Similarity	Language	Meaning and scope
Skipped	Passed	Flagged
Not requested	No findings · confidence 98 %	1 finding · confidence 5 %

1 finding to review
They are listed under Findings.

Demo

4. Non - Valid notice per meaning and scope check type – spam

1 finding to review

They are listed under Findings.

Findings 1 | Notice XML | Service response

Showing the findings of all checks. Select a check above to show only its findings. Select a column heading to sort, and a row to see the finding below. Previous and next follow the order in which the service returned the findings.

Check	Finding	Field	Message	Confidence
Meaning and scope	SPAM_CONTENT	BT-24-Procedure	The text contains a scam/prize hook ('Click quickly to confirm your p	5 %

Demo

4. Non - Valid notice per meaning and scope check type – spam

Finding 1 of 1 · SPAM_CONTENT in BT-24-Procedure

The text contains a scam/prize hook ('Click quickly to confirm your prize: today') attached to an otherwise legitimate-looking procurement notice, qualifying as spam.

Check	Meaning and scope
Confidence	5 %
Business entity	procedure (procedure)
XML location	`/*/cac:ProcurementProject/cbc:Description`

Show this field in the notice XML

```
121    <cbc:Description languageID="ENG">Click quickly to confirm your prize: today. Tender for the appointme
122    <cbc:ProcurementTypeCode listName="contract-nature">services</cbc:ProcurementTypeCode>
```

Demo

4. Non - Valid notice per meaning and scope check type – spam


```
2 : {
  "notice_id" : "b5a71b8c-3363-4825-8f0e-b11ae244dfb2"
  "validation_check_type" : "meaning_and_scope_check"
  "status" : "warning"
  "confidence_score" : 0.05
  "findings" : [
    0 : {
      "code" : "SPAM_CONTENT"
      "message" :
      "The text contains a scam/prize hook ('Click quickly to confirm your prize: today') attached to an otherwise legitimate-
      looking procurement notice, qualifying as spam."
      "confidence_score" : 0.05
      "field_id" : "BT-24-Procedure"
      "field_type" : "Description"
      "xml_path" : "/*/cac:ProcurementProject/cbc:Description"
      "business_entity_id" : "procedure"
      "business_entity_label" : "procedure"
      "business_entity_instance_id" : NULL
      "evidence" : {
        "predictions" : [
          0 : {
            "label" : "spam"
            "confidence" : 0.95
          }
        ]
        "explanation" :
        "The text contains a scam/prize hook ('Click quickly to confirm your prize: today') attached to an otherwise legitimate-
        looking procurement notice, qualifying as spam."
      }
```

THANK YOU

© European Union



