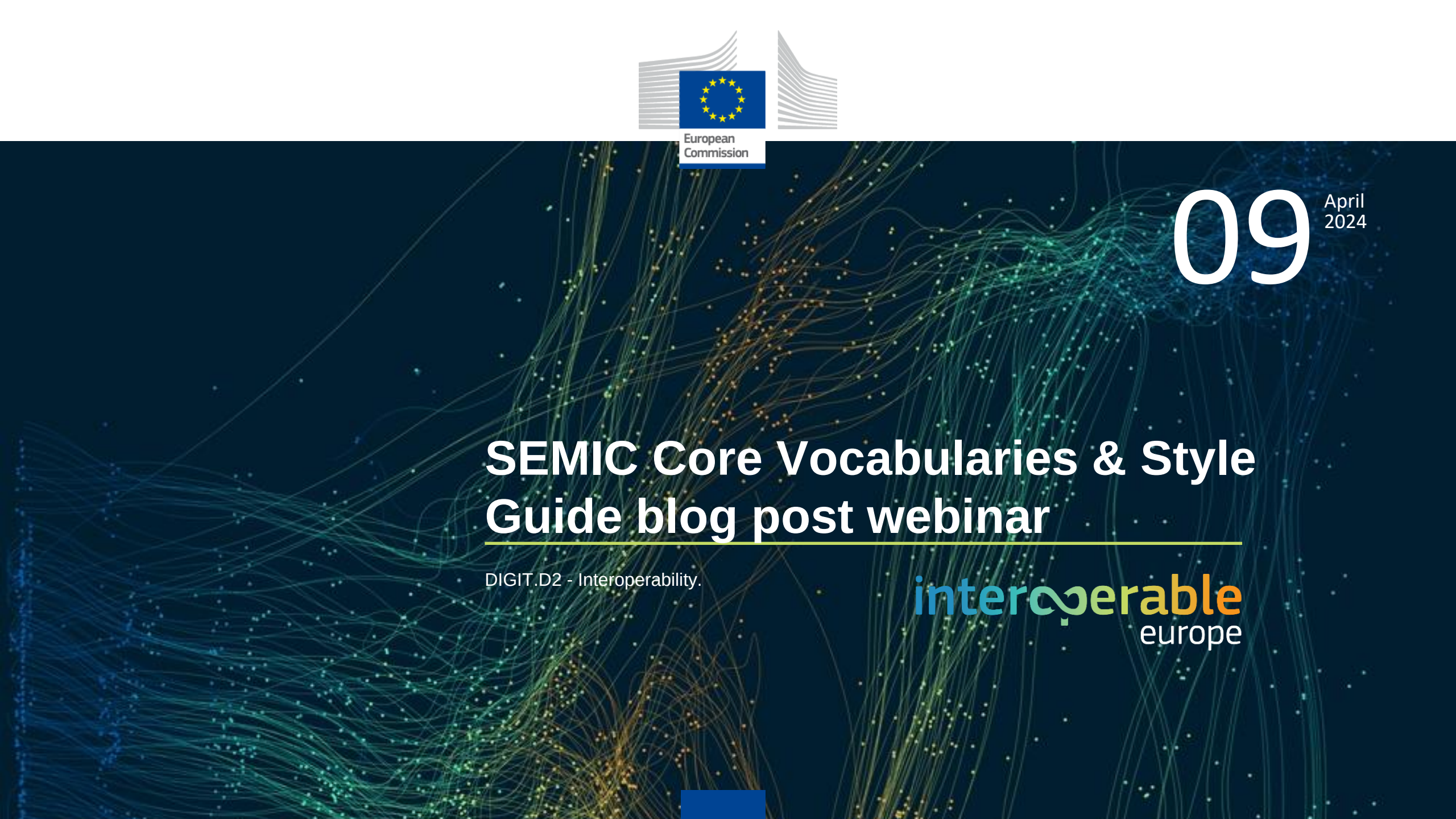

European Commission
09
April
2024
SEMIC Core Vocabularies & Style Guide blog post webinar
DIGIT.D2 - Interoperability.
interoperable europe

Objectives of this webinar

Close public review

Discuss various open issues related to the Core Vocabularies and CPSV-AP

Gather input for improvements

Gather input for those issues that could not be resolved in the latest release for public review.

Discuss the SG blog post

Discuss the Style Guide blog post on mapping the Core Vocabularies to XML

Workshop practicalities

Audio

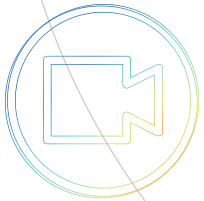
Click on ‘connect audio’ but please mute your microphones

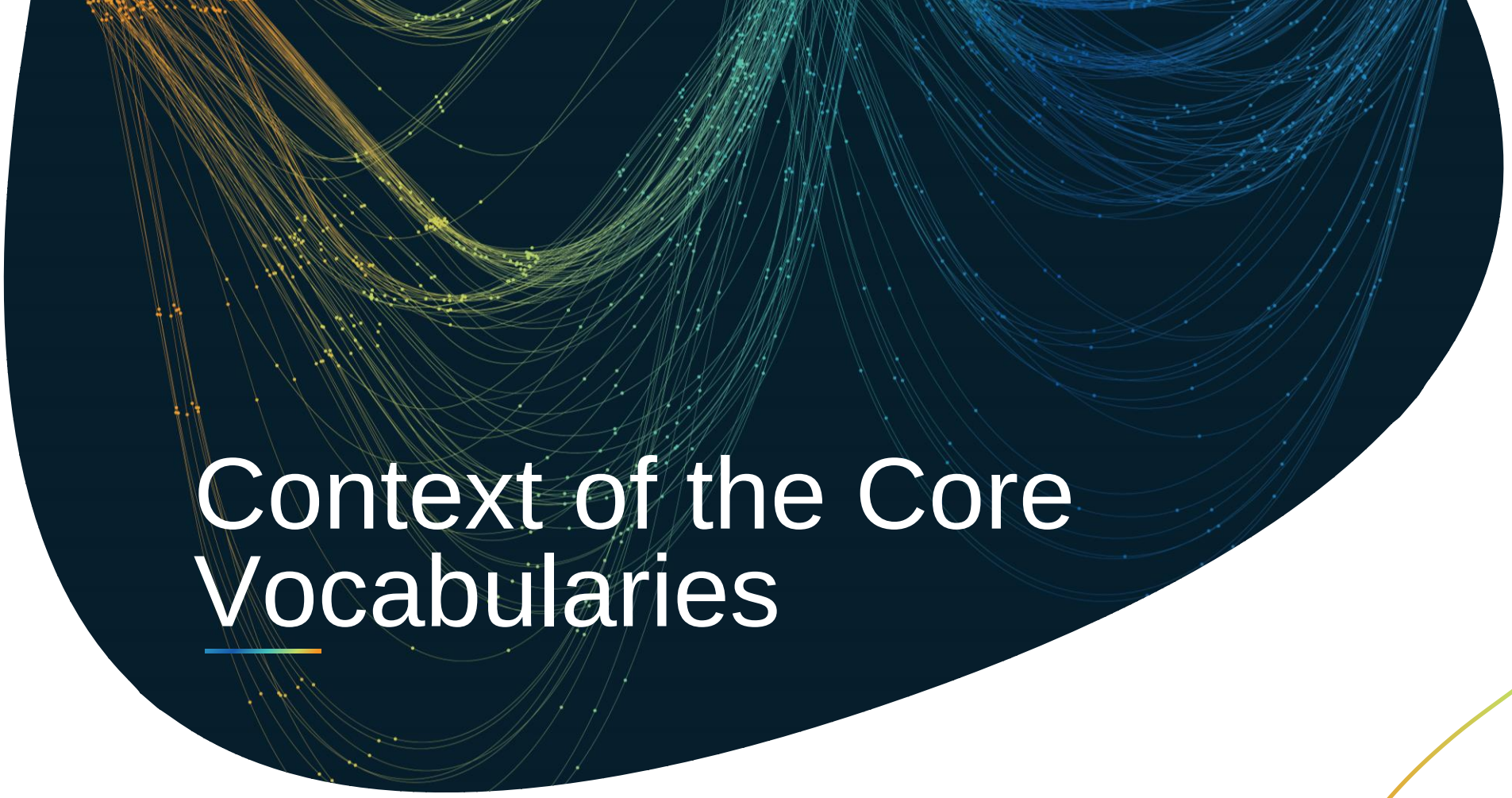
Chat

You can also share your questions for the Q&A session via the chat

Recording

The workshop will be recorded





Context of the Core Vocabularies

Introduction to SEMIC

The objectives of the SEMIC action are to promote Semantic Interoperability amongst the EU Member States by:

- Promoting the share and reuse of semantic assets, experience and tools and facilitating agreements in key areas.
- Identifying opportunities for alignment on semantic definitions, metadata and reference data sources with special focus on identification and definitions of Core Concepts / Vocabularies.
- Raising awareness on the importance of data and metadata management.

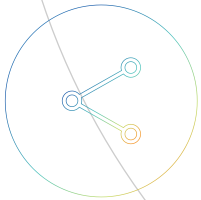
Objectives of the Core Vocabularies

The e-Government Core Vocabularies are data models that capture the **fundamental characteristics** of a data entity in a context-neutral and syntax-neutral fashion.

Simplified

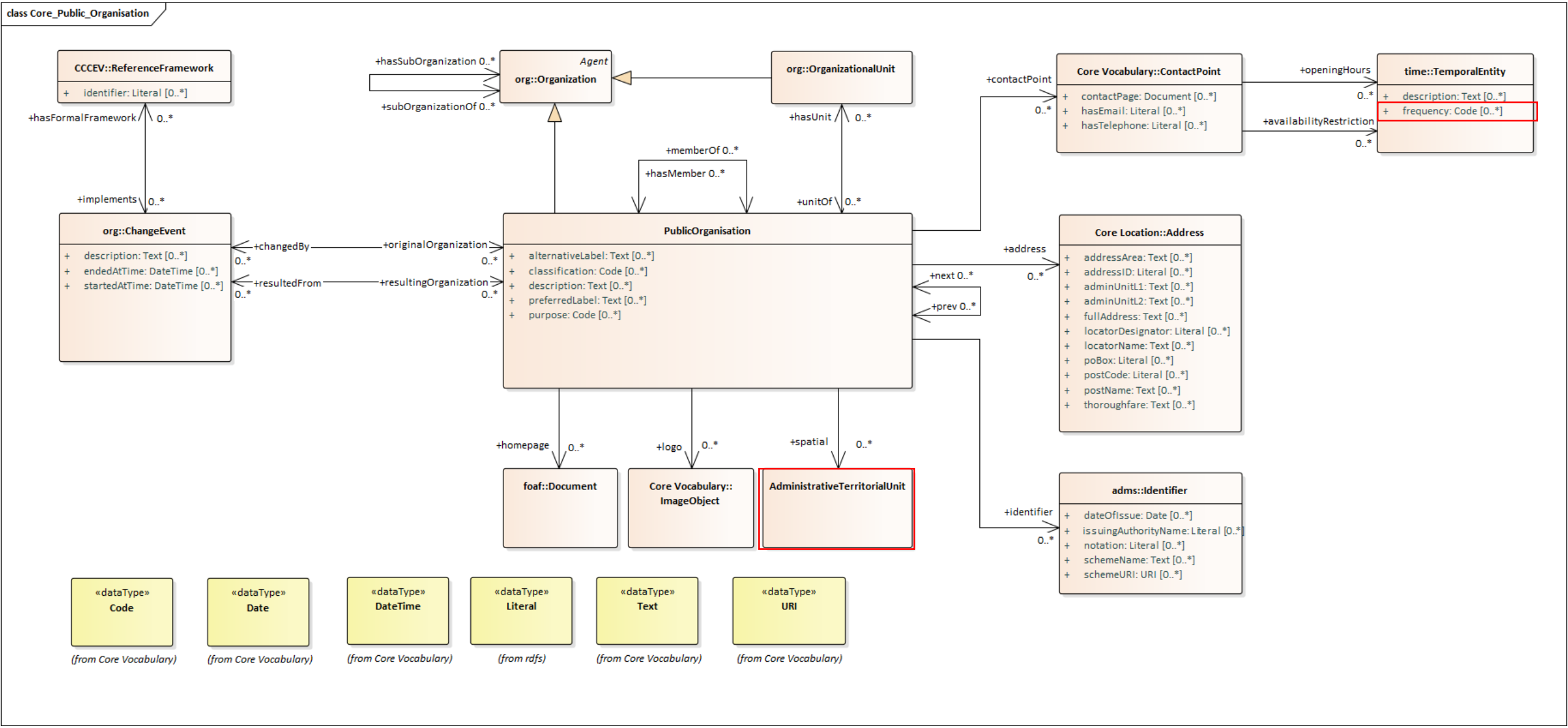
Re-usable

Extensible



Issues on CPOV

CPOV 2.1.1


class Core_Public_Organisation
CCCEV::ReferenceFramework
+ identifier: Literal [0..*]
+hasFormalFramework
0..*
+implements
0..*
org::ChangeEvent
+ description: Text [0..*]
+ endedAtTime: DateTime [0..*]
+ startedAtTime: DateTime [0..*]
+changedBy
0..*
+originalOrganization
0..*
+resultedFrom
0..*
+resultingOrganization
0..*
+hasSubOrganization 0..*
+subOrganizationOf 0..*
Agent
org::Organization
org::OrganizationalUnit
+hasUnit
0..*
+memberOf 0..*
+hasMember 0..*
+unitOf
0..*
PublicOrganisation
+ alternativeLabel: Text [0..*]
+ classification: Code [0..*]
+ description: Text [0..*]
+ preferredLabel: Text [0..*]
+ purpose: Code [0..*]
+contactPoint
0..*
Core Vocabulary::ContactPoint
+ contactPage: Document [0..*]
+ hasEmail: Literal [0..*]
+ hasTelephone: Literal [0..*]
+openingHours
0..*
+availabilityRestriction
0..*
time::TemporalEntity
+ description: Text [0..*]
+ frequency: Code [0..*]
+address
0..*
Core Location::Address
+ addressArea: Text [0..*]
+ addressID: Literal [0..*]
+ adminUnitL1: Text [0..*]
+ adminUnitL2: Text [0..*]
+ fullAddress: Text [0..*]
+ locatorDesignator: Literal [0..*]
+ locatorName: Text [0..*]
+ poBox: Literal [0..*]
+ postCode: Literal [0..*]
+ postName: Text [0..*]
+ thoroughfare: Text [0..*]
+next 0..*
+prev 0..*
+homepage
0..*
+logo
0..*
+spatial
0..*
foaf::Document
Core Vocabulary::
ImageObject
AdministrativeTerritorialUnit
+identifier
0..*
adms::Identifier
+ dateOfIssue: Date [0..*]
+ issuingAuthorityName: Literal [0..*]
+ notation: Literal [0..*]
+ schemeName: Text [0..*]
+ schemeURI: URI [0..*]
«dataType» Code
(from Core Vocabulary)
«dataType» Date
(from Core Vocabulary)
«dataType» DateTime
(from Core Vocabulary)
«dataType» Literal
(from rdfs)
«dataType» Text
(from Core Vocabulary)
«dataType» URI
(from Core Vocabulary)

Examples missing property (issue #30)

Request

In the example:

Should ex:shos have a cv:frequency?

To Be

```
ex:sohs a time:TemporalEntity ;
  dct:description "closed on weekend"@en;
  cv:frequency <http://publications.europa.eu/resource/authority/frequency/WEEKLY>;
  time:hasBeginning [ a time:Instant;
    time:inDateTime [ a time:DateTimeDescription;
      time:dayOfWeek time:Saturday;
    ];
  ];
```

Resolution

- This is an optional property
- cv:frequency has been added

Administrative Territorial Unit (issue #31, CPSV-AP #123)

Request

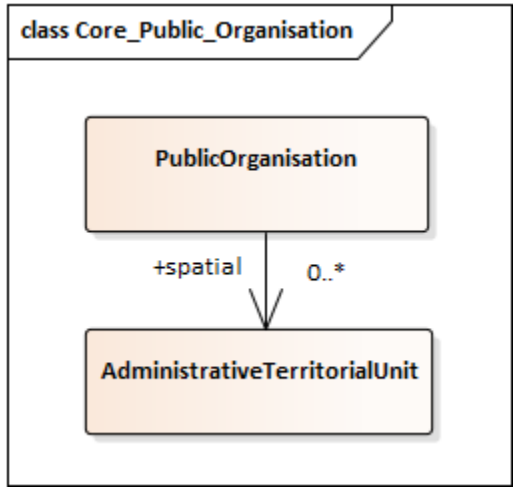
Administrative Territorial Unit is a specific code list.

- It does not align with the definition:
  - "A code from **a** list …"
- It does not align with the Usage note
  - "In Europe, this is **likely** to be the Administrative Territorial Units Named Authority List"
- It does not align with Style Guide rule CMC-R14
  - Enumeration should be used
  - Only when one specific code list is possible
- Within Public Service dct:spatial has range dct:Location

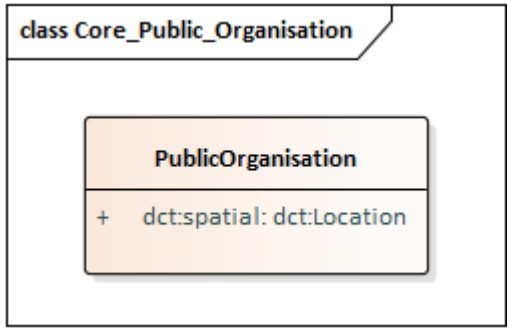
Proposition

- In CPOV and CPSV-AP change the range to dct:Location
- Maintain Usage Note with recommendation for ATU

As is

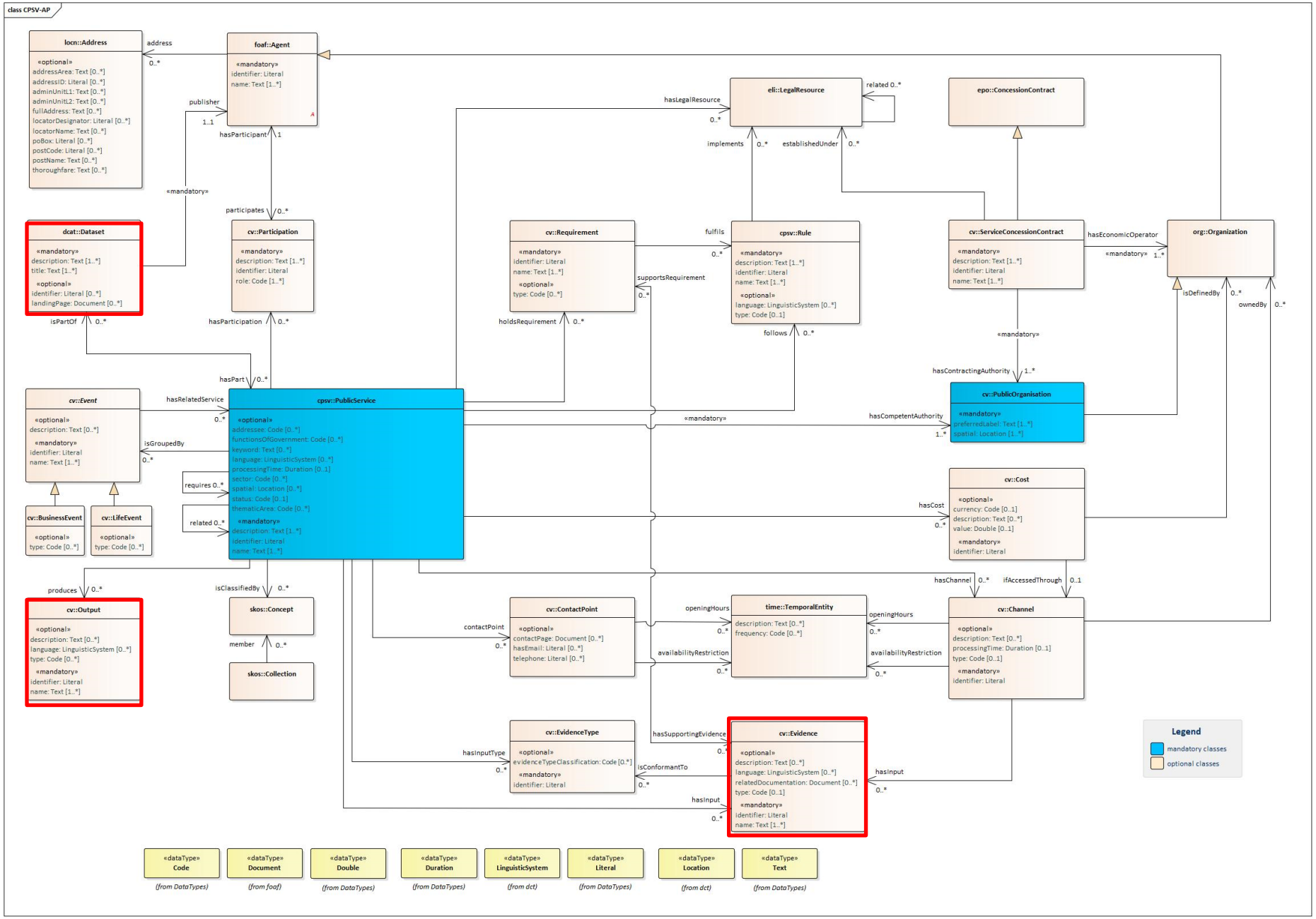


To Be



Issues on CPSV-AP

CPSV-AP 3.2.0


class CPSV-AP
locn::Address
«optional»
addressArea: Text [0..*]
addressID: Literal [0..*]
adminUnitL1: Text [0..*]
adminUnitL2: Text [0..*]
fullAddress: Text [0..*]
locatorDesignator: Literal [0..*]
locatorName: Text [0..*]
poBox: Literal [0..*]
postCode: Literal [0..*]
postName: Text [0..*]
thoroughfare: Text [0..*]
address
0..*
foaf::Agent
«mandatory»
identifier: Literal
name: Text [1..*]
publisher
1..1
hasParticipant
1
«mandatory»
participates
0..*
dcat::Dataset
«mandatory»
description: Text [1..*]
title: Text [1..*]
«optional»
identifier: Literal [0..*]
landingPage: Document [0..*]
isPartOf
0..*
cv::Participation
«mandatory»
description: Text [1..*]
identifier: Literal
role: Code [1..*]
hasParticipation
0..*
hasPart
0..*
cv::Event
«optional»
description: Text [0..*]
«mandatory»
identifier: Literal
name: Text [1..*]
cv::BusinessEvent
«optional»
type: Code [0..*]
cv::LifeEvent
«optional»
type: Code [0..*]
isGroupedBy
0..*
hasRelatedService
0..*
cpsv::PublicService
«optional»
addressee: Code [0..*]
functionsOfGovernment: Code [0..*]
keyword: Text [0..*]
language: LinguisticSystem [0..*]
processingTime: Duration [0..1]
sector: Code [0..*]
spatial: Location [0..*]
status: Code [0..1]
thematicArea: Code [0..*]
«mandatory»
description: Text [1..*]
identifier: Literal
name: Text [1..*]
requires 0..*
related 0..*
produces
0..*
cv::Output
«optional»
description: Text [0..*]
language: LinguisticSystem [0..*]
type: Code [0..*]
«mandatory»
identifier: Literal
name: Text [1..*]
isClassifiedBy
0..*
skos::Concept
member
0..*
skos::Collection
eli::LegalResource
related 0..*
hasLegalResource
0..*
implements
0..*
establishedUnder
0..*
epo::ConcessionContract
cv::Requirement
«mandatory»
identifier: Literal
name: Text [1..*]
«optional»
type: Code [0..*]
fulfils
0..*
supportsRequirement
0..*
holdsRequirement
0..*
cpsv::Rule
«mandatory»
description: Text [1..*]
identifier: Literal
name: Text [1..*]
«optional»
language: LinguisticSystem [0..*]
type: Code [0..1]
follows
0..*
cv::ServiceConcessionContract
«mandatory»
description: Text [1..*]
identifier: Literal
name: Text [1..*]
hasEconomicOperator
«mandatory»
1..*
org::Organization
isDefinedBy
0..*
ownedBy
0..*
«mandatory»
hasContractingAuthority
1..*
cv::PublicOrganisation
«mandatory»
preferredLabel: Text [1..*]
spatial: Location [1..*]
hasCompetentAuthority
1..*
«mandatory»
cv::Cost
«optional»
currency: Code [0..1]
description: Text [0..*]
value: Double [0..1]
«mandatory»
identifier: Literal
hasCost
0..*
hasChannel
0..*
ifAccessedThrough
0..1
cv::Channel
«optional»
description: Text [0..*]
processingTime: Duration [0..1]
type: Code [0..1]
«mandatory»
identifier: Literal
openingHours
0..*
availabilityRestriction
0..*
time::TemporalEntity
description: Text [0..*]
frequency: Code [0..*]
cv::ContactPoint
«optional»
contactPage: Document [0..*]
hasEmail: Literal [0..*]
telephone: Literal [0..*]
contactPoint
0..*
cv::EvidenceType
«optional»
evidenceTypeClassification: Code [0..*]
«mandatory»
identifier: Literal
hasInputType
0..*
isConformantTo
0..*
hasSupportingEvidence
0..*
cv::Evidence
«optional»
description: Text [0..*]
language: LinguisticSystem [0..*]
relatedDocumentation: Document [0..*]
type: Code [0..1]
«mandatory»
identifier: Literal
name: Text [1..*]
hasInput
0..*
hasInput
0..*
Legend
mandatory classes
optional classes
«dataType» Code (from DataTypes)
«dataType» Document (from foaf)
«dataType» Double (from DataTypes)
«dataType» Duration (from DataTypes)
«dataType» LinguisticSystem (from dct)
«dataType» Literal (from DataTypes)
«dataType» Location (from dct)
«dataType» Text (from DataTypes)

Need to relate Evidence and Output to Dataset (issue #125)

Request

- Relate cv:Output and cv:Evidence to dcat:Dataset using an optional relation
- A relation is required for Once Only implementations

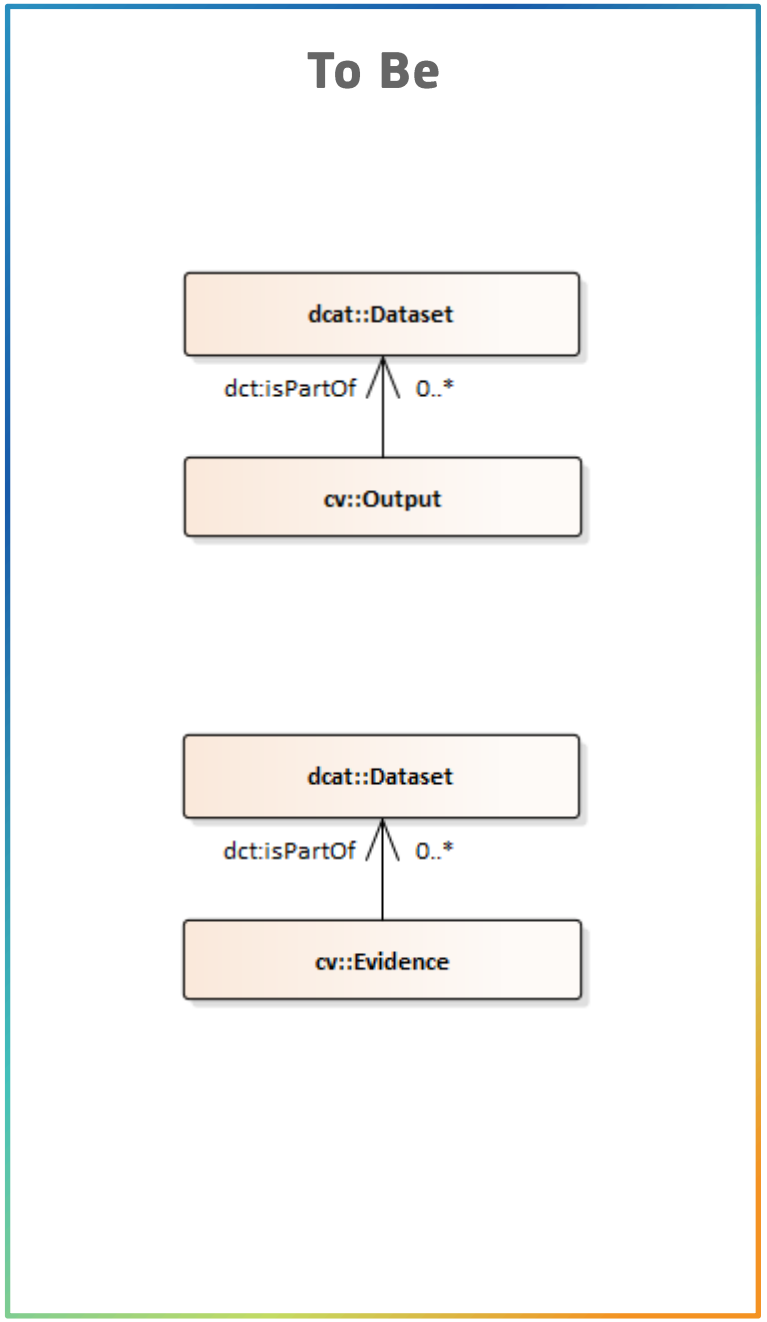
Context from previous webinar

- Removed mandatory subclassing: cv:Evidence is no longer a dcat:Dataset
- Changed relation between cpsv:PublicService and dcat:Dataset: from cv:isDescribedAt to dct:isPartOf

Proposition

Have optional dct:isPartOf towards dcat:Dataset from:

- cv:Evidence
- Cv:Output



CPSV-AP for “describing” and for “executing” (issue #128)

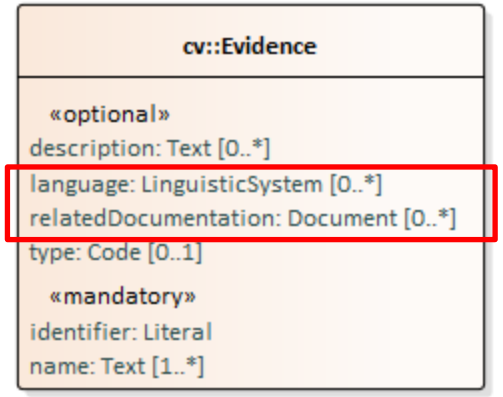
Request

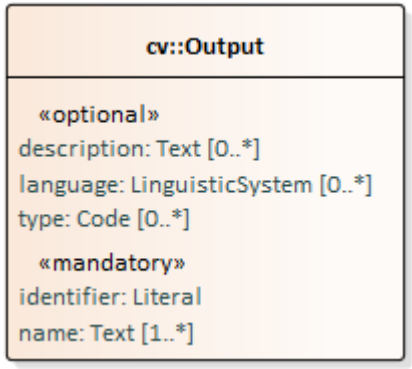
1. How to use cv:Evidence and cv:output when describing a cpsv:PublicService?
2. Update Language definition to allow both usages of CPSV-AP for the description and execution of Public Services
3. Update the relatedDocumentation usage note to allow the “execution” use case

Example

- Using CPSV-AP to describe a Public Service as needed for a Catalogue of service
  - The generic concept of a drivers license is a cv:EvidenceType
- Using CPSV-AP during the execution of a public service
  - John Doe’s drivers license is a cv:Evidence

As is





CPSV-AP for “describing” and for “executing” (issue #128)

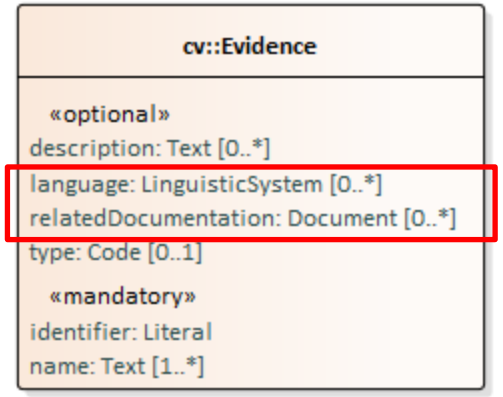
Context

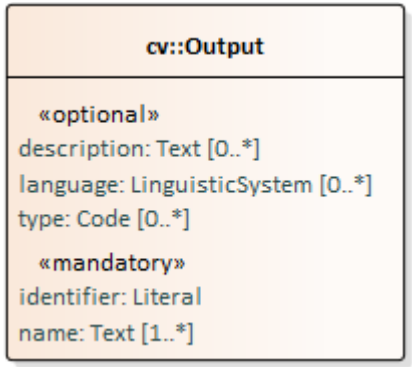
1. As described in issue 95: CPSV-AP can be used to describe a Public Service or describe the execution of a public service
   - For describing Public Services it is recommended to use cv:EvidenceType rather than cv:Evidence which is more suitable for the execution of Public Services
2. Language: The language(s) in which the Evidence must be provided.
   - This definition is catered to “describing” while the definition should allow “execution”
3. Related documentation: a particular template for an administrative document, an application or a guide on formatting the Evidence.
   - This usageNote is catered to “describing” while it should focus on “execution”

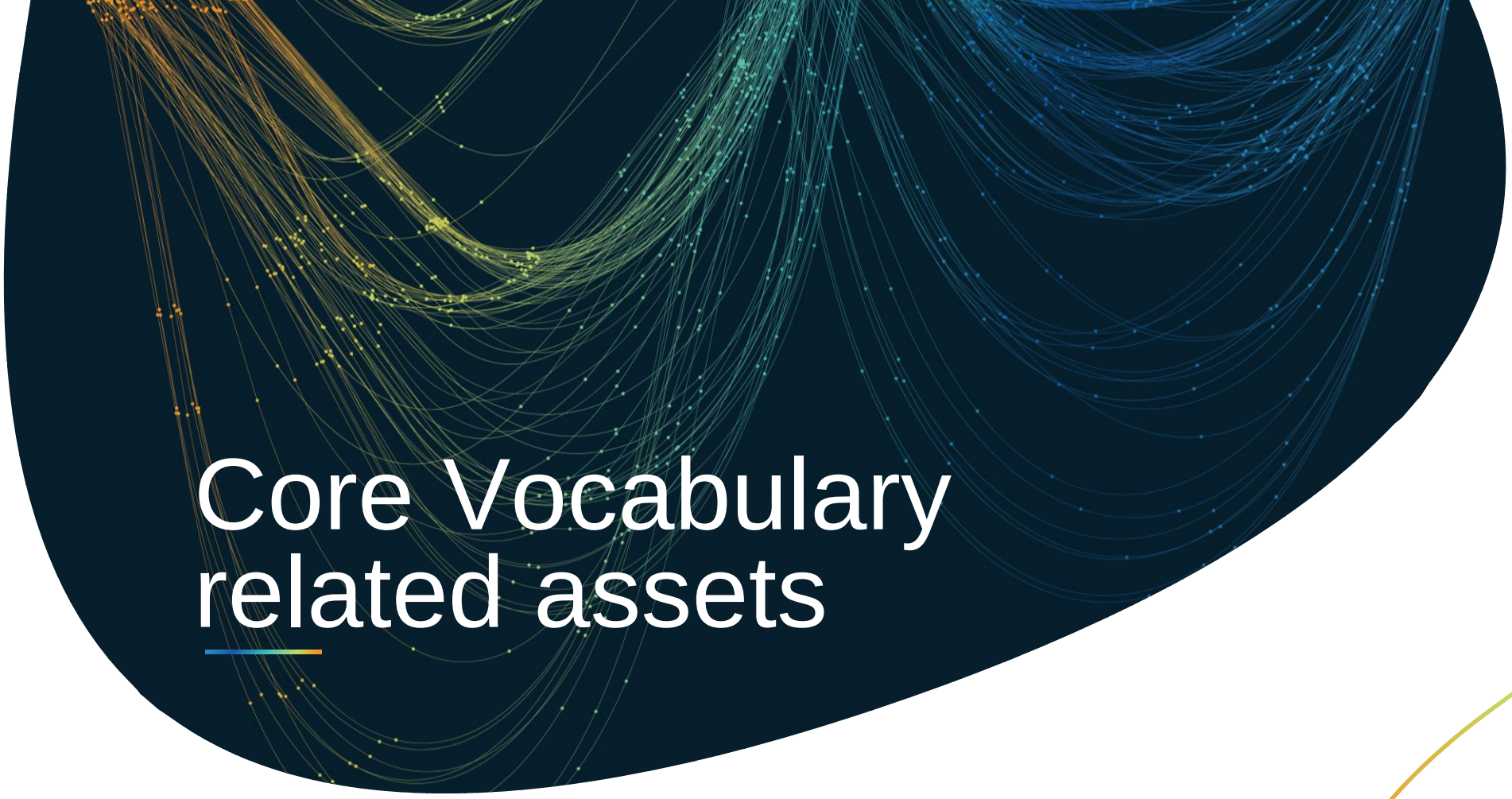
Proposition

1. Explicitly add the “execution” use case to the introduction of the specification
   - Reference the UC in cv:Evidence & cv:Output’s usage notes
2. Update Language definition to allow both usages
   - The language(s) the piece of Evidence is in.
3. Update the related documentation usage note
   - For instance, an energy audit report providing more context to the evidence of a home energy efficiency score.

As is







Core Vocabulary related assets

New template and examples

ReSpec

- From the public review of the Core Vocabularies onwards, all SEMIC specifications will be published using ReSpec

Examples

- Examples are provided to facilitate the use of the SEMIC specifications in Turtle and JSON-LD
- The examples are linked to:
  - a convertor that allows to perform inline transformation of the RDF into JSON-LD, Turtle, RDF/XML and Trig formats
  - a validator that reuses the SHACL shapes associated to the specification by performing a strict validation (expected values, cardinalities, etc).
  - The ITB validation service is currently being tested. The current platform will be Phased out if possible.

Consolidated Core Vocabularies

Consolidated Core Vocabularies

A HTML page will be published containing all m8g terms.

- It will contain all classes, properties and relations defined under the data.europa.eu/m8g/ name space
- It will contain a consolidated diagram combining all Core Vocabularies.

Our aim: to facilitate the reuse of m8g classes, properties and definitions.
By increasing findability.
No need to explore 6 Vocabularies and 1 Application profile.

Core Vocabularies
29 February 2024

▼ More details about this document

Latest published version:
https://semiceu.github.io/Consolidated-Core-Vocabularies/releases/2.2.0/

Latest editor's draft:
https://semiceu.github.io/Consolidated-Core-Vocabularies/releases/2.2.0/

History:
Commit history

Editors:

§ 4.1.18 Opening Hours Specification

Class	Opening Hours Specification
URI	http://data.europa.eu/m8g/OpeningHoursSpecification
label	Opening Hours Specification
definition	An exceptional circumstance that overrides the normal Opening Hours, such as being closed on public holidays.
usage	This class is aligned to Schema.org OpeningHoursSpecification class.

§ 4.1.19 Output

Class	Output
URI	http://data.europa.eu/m8g/Output
label	Output
definition	An output can be any resource - document, artefact - anything produced by the Public Service. In the context of a Public Service, the output provides an official document or other artefact of the Competent Authority (Public Organization) that permits/authorises/entitles an Agent to (do) something.

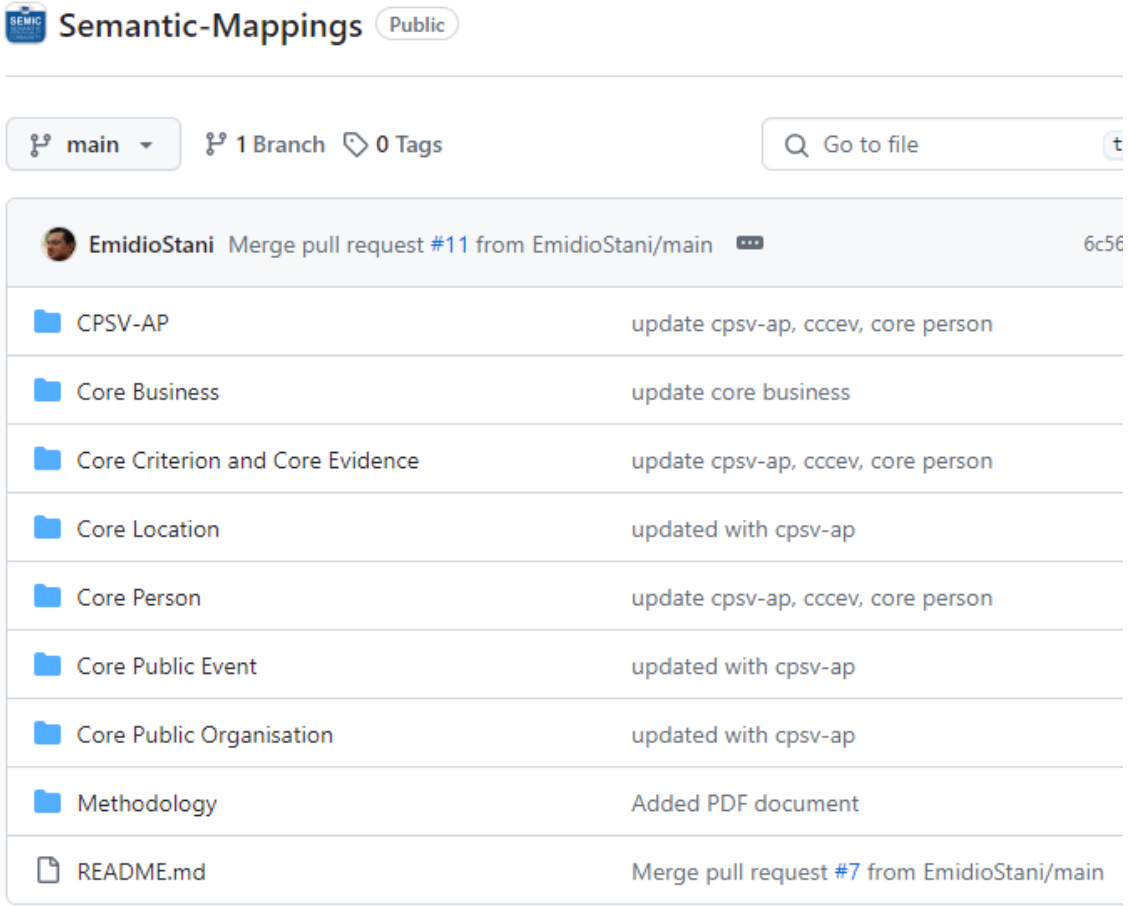
§ 4.1.20 Participation

Class	Participation
URI	http://data.europa.eu/m8g/Participation
label	Participation
definition	The Participation class allows to define roles within a certain context.

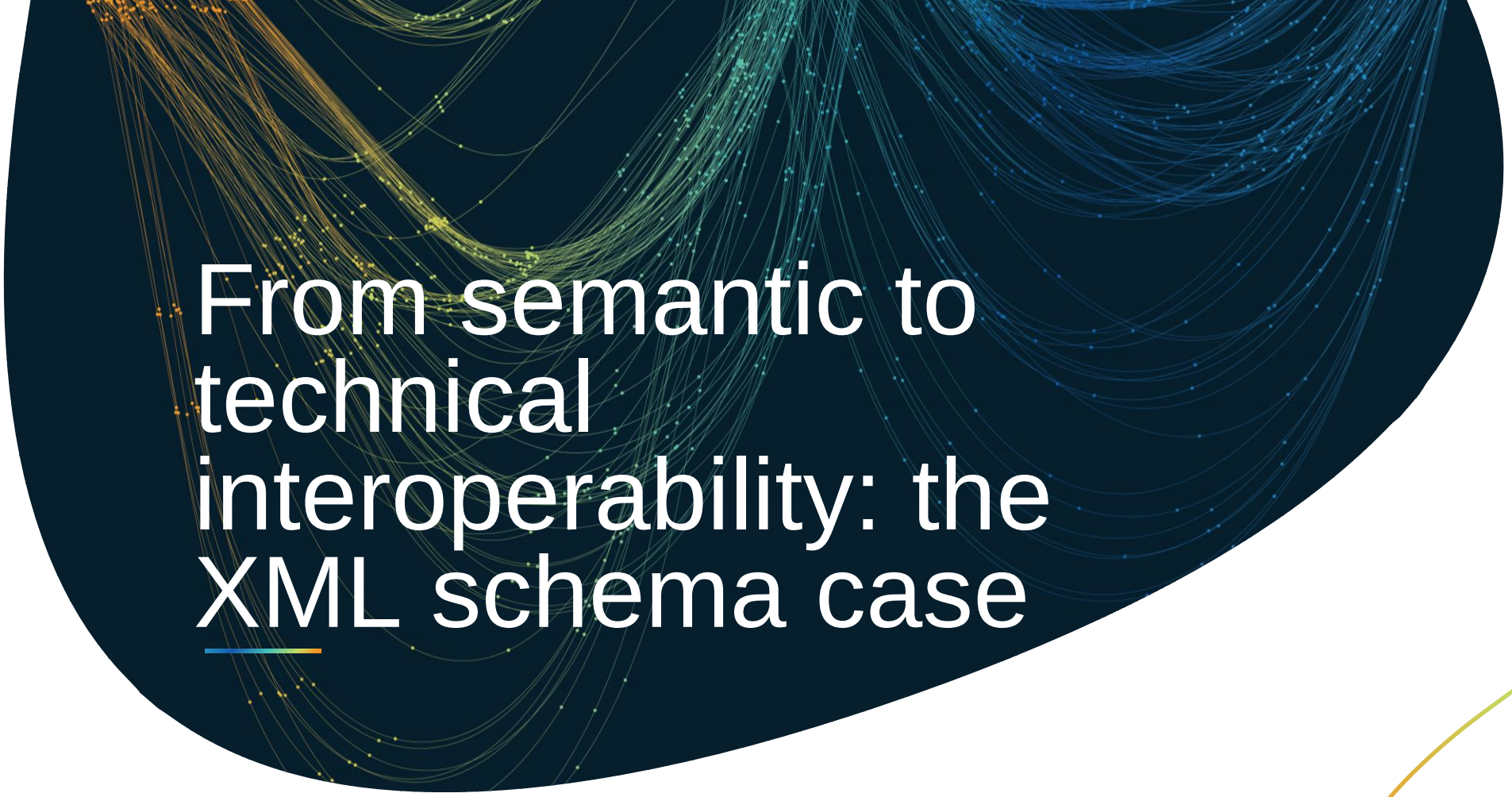
An updated mapping

Mapping with Schema.org

- The mapping has been updated to this release of the Core Vocabularies
- The mapping has been updated to v26 of schema.org
- The maping now also includes CPSV-AP



Short break (5 minutes)



From semantic to technical interoperability: the XML schema case

XML schema: a case study using the CoreVocabularies

The blog post

- In order to have both Semantic and Technical Interoperability a conceptual model will require conversion to a physical model.
- We explore the mapping between RDF and XML Schema using Core Vocabularies as starting point.
- The creation of a standard way to describe the Core Vocabularies as XML Schemas

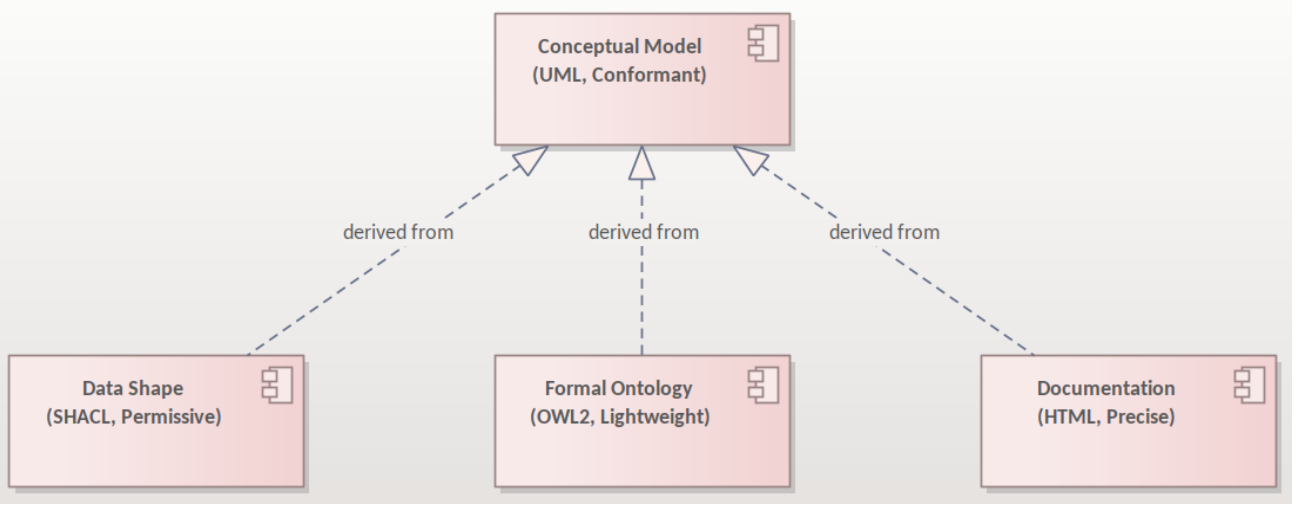
Today

- We will focus on the open question in the various sections.

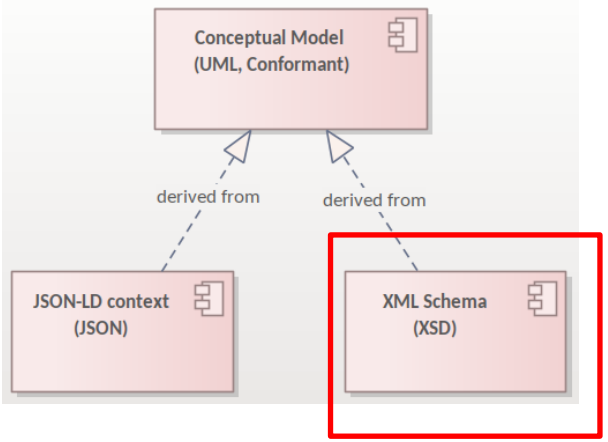
Table of Content

Style Guide assets generation

Following the process of the Style Guide, artefacts can be derived from the Conceptual models.



Following the same process, semantic engineers will move from a conceptual model to a physical data model such as an XML schema.



Need for metadata (including versioning) + Mural

```
<xs:schema xmlns="https://data.europa.eu/m8g/xml/"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="https://data.europa.eu/m8g/xml/"
version="2.0.0"
elementFormDefault="qualified"
attributeFormDefault="unqualified"
xmlns:cv="https://data.europa.eu/m8g/xml/">
```

Question for the community: Currently SEMIC does not provide an XML schema. What is the minimum metadata needed for an XML schema aside from versioning?

How to define data types

```
<xs:simpleType name="Literal">
  <xs:restriction base="xs:string"/>
</xs:simpleType>
<xs:complexType name="langstring">
  <xs:simpleContent>
    <xs:extension base="xs:string">
      <xs:attribute name="lang" type="xs:language"/>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>
```

Following the W3C SAWSDL convention:

```
<xs:simpleType name="Literal"
sawsdl:modelReference="http://www.w3.org/2000/01/rdf-schema#Literal">
<xs:restriction base="xs:string">
</xs:simpleType>
```

XML attribute that identifies or defines a concept in a semantic model

- SAWSDL does not establish rules for URI dereferencing
- Dereferencing depends on the model owner

Question for the community: Is it important to bring a mapping like the Literal expression above within the XML schema ?
Are there any Member States using SAWSDL, as for example Finland?

Reusing existing XML schemas vs re-implementing

Not all ontologies are also expressed in XML schema. For example, Core Vocabularies are based on DCTERMS and FOAF. While for DCTERMS there is one that could be reused, FOAF doesn't have it.

Reusing DCTERMS Agent

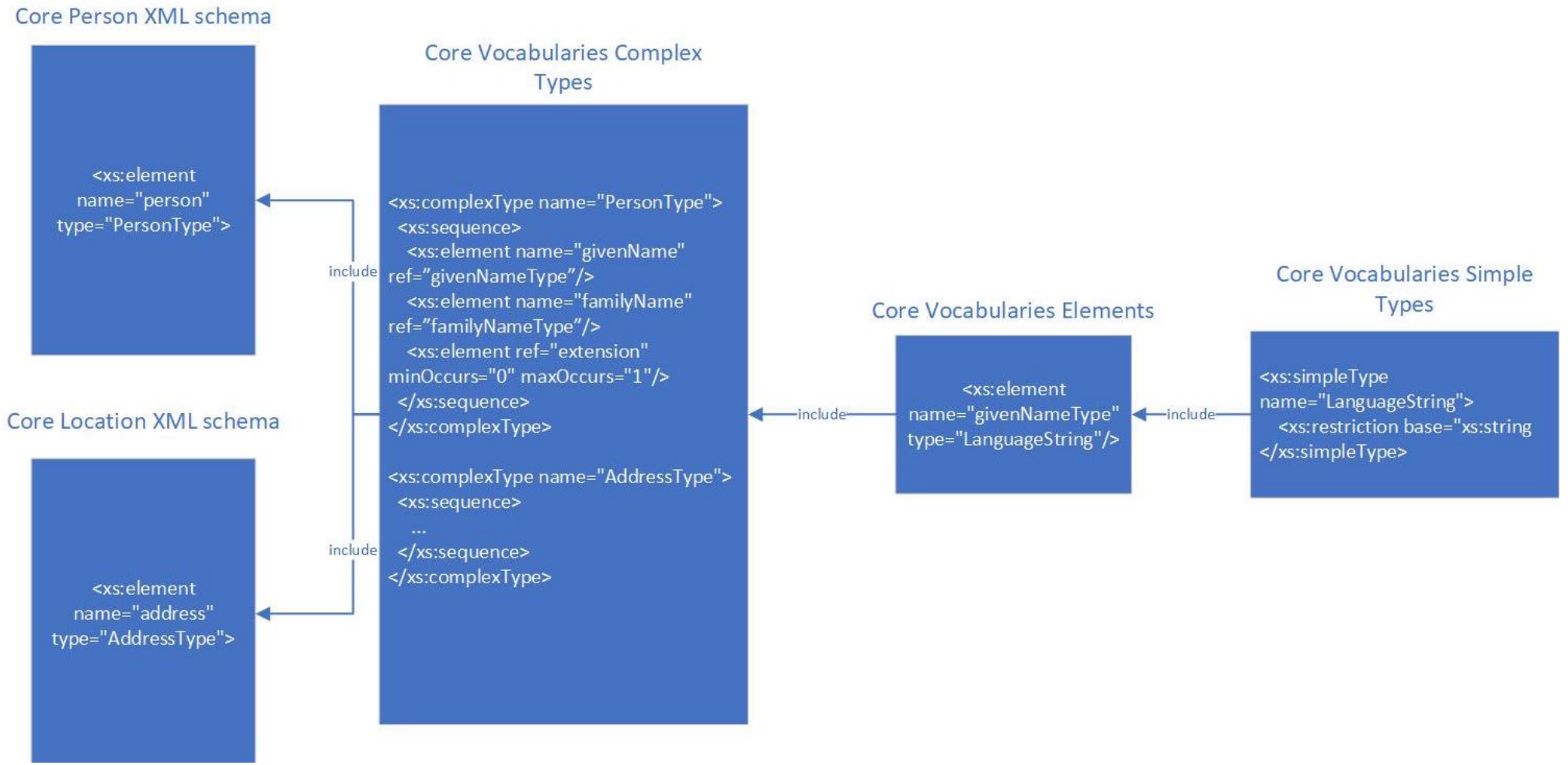
```
<xs:schema elementFormDefault="qualified"
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    xmlns:dct="http://purl.org/dc/terms/">

  <xs:import namespace="http://purl.org/dc/terms/" schemaLocation="dcterms.xsd"/>

 <xs:complexType name="AgentType">
  <xs:sequence>
   <xs:element ref="dct:title" maxOccurs="unbounded" minOccurs="0">
  </xs:sequence>
<xs:complexType>
```

Implementing FOAF Agent

```
<xs:complexType name="AgentType">
 <xs:sequence>
  <xs:element name="name" type="xs:string" maxOccurs="unbounded" minOccurs="0">
   <xs:annotation>
    <xs:documentation xml:lang="en">The noun given to the Agent.</xs:documentation>
   </xs:annotation>
  </xs:element>
  <xs:element name="type" type="ConceptType" maxOccurs="unbounded" minOccurs="0">
   <xs:annotation>
    <xs:documentation xml:lang="en">A classification assigned to an Agent.</xs:documentation>
   </xs:annotation>
  </xs:element>
```

Question for the community: What kind of approach should be considered when integrating external namespaces that do not have a respective XML schema ?

How the XML schema could be structured



Question for the community: Is the above proposed approach of structuring Core Vocabularies in different files useful?

Extension mechanism

```
<xs:element name="person" type="PersonType">

<xs:complexType name="PersonType">
 <xs:sequence>
  <xs:element name="givenName" type="xs:string"/>
  <xs:element name="familyName" type="xs:string"/>
  <xs:element ref="extension" minOccurs="0" maxOccurs="1"/>
 </xs:sequence>
</xs:complexType>

<xs:element name="extension" type="ExtensionType">

<xsd:complexType name="ExtensionType">
 <xsd:sequence>
  ...
  <xsd:element ref="ExtensionContent" minOccurs="1" maxOccurs="1"/>
 </xsd:sequence>
</xsd:complexType>

<xs:element name="ExtensionContent"" type="ExtensionContent"Type">

<xsd:complexType name="ExtensionContentType">
 <xsd:sequence>
  <xsd:any namespace="##other" processContents="lax"
        minOccurs="1" maxOccurs="1"/>
 </xsd:sequence>
</xsd:complexType>
```

VS

```
<xs:element name="person" type="PersonType">

<xs:complexType name="PersonType">
 <xs:sequence>
  <xs:element name="givenName" type="xs:string"/>
  <xs:element name="familyName" type="xs:string"/>
  <xs:element ref="PersonAugmentationPoint" minOccurs="0" maxOccurs="1"/>
 </xs:sequence>
</xs:complexType>

<xs:element name="PersonAugmentationPoint" abstract="true">
</xs:element>

--- Extending schema

<xs:complexType name="PersonExtensionType">

<xs:annotation>
  <xs:documentation>A data type for additional information about a person.</xs:documentation>
 </xs:annotation>
 <xs:sequence>
  <xs:element ref="j:PersonDateOfDeath" minOccurs="0" maxOccurs="unbounded"/>
 </xs:sequence>
</xs:complexType>

<xs:element name="PersonAugmentation" type="j:PersonExtensionType"
 substitutionGroup="xs:PersonAugmentationPoint">
</xs:element>
```

Questions for the community:

- Is the approach of providing an extension mechanism for each defined type useful for the SEMIC community?
- If yes, what kind of approach should be taken?

Multiple Inheritance and Multiple Instances

```
<xs:group name="Person">
  <xs:sequence>
    <xs:element name="firstName" type="xs:string"/>
    <xs:element name="lastName" type="xs:string"/>
  </xs:sequence>
</xs:group>

<xs:group name="LegalEntity">
  <xs:sequence>
    <xs:element name="legalName" type="xs:string"/>
    <xs:element name="identifier" type="xs:string"/>
  </xs:sequence>
</xs:group>

<xs:complexType name="Person">
  <xs:sequence>
    <xs:group ref="Person"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="LegalEntity">
  <xs:sequence>
    <xs:group ref="LegalEntity"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="Host">
  <xs:sequence>
    <xs:group ref="Person"/>
    <xs:group ref="LegalEntity"/>
    <xs:element name="address" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
```

Questions for the community:

- Currently SEMIC Core Vocabularies do not have classes inheriting from multiple classes but the community might have this multiple inheritance.
- Is the approach above needed for reusers, when expressing Core Vocabularies as an XML schema ?

Naming and Design Rules to be followed

Rules in common between NIEM and UBL:

- The XML schema must have a version, see NIEM, UBL;
- All elements declarations must be global, see NIEM and UBL;
- All types must be global, see NIEM and UBL;
- No use of xs:all, see NIEM and UBL;
- No use of xs:any, see NIEM and UBL (NIEM speaks about augmentation point while UBL provides its extension element);
- No use of xs:choice, see NIEM and UBL.

Other rules are different such as:

- On the use of substitution groups, see NIEM and UBL
- On the use of union, see NIEM and UBL

Questions for the community: Aside of the common rules that could already be adopted, what kind of approach could be followed up?

Need for XML validation service

Person.xml

```
<person xmlns="https://data.europa.eu/m8g/xml/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="https://data.europa.eu/m8g/xml/person.xsd">
   <givenName>René François Ghislain</givenName>
   <familyName>Magritte</familyName>
</person>
```

Person.xsd

```
<xs:complexType name="PersonType">
 <xs:sequence>
  <xs:element name="givenName"
ref="givenNameType"/>
  <xs:element name="familyName"
ref="familyNameType"/>
  <xs:element ref="extension"
minOccurs="0" maxOccurs="1"/>
 </xs:sequence>
</xs:complexType>

<xs:complexType name="AddressType">
 <xs:sequence>
  ...
 </xs:sequence>
</xs:complexType>
```

Question for the community: Currently SEMIC does not provide an XML schema validation service for the Core Vocabularies such as it is done for ITB CPSV-AP and DCAT-AP SHACL validators. Is this type of service needed ?

Wrap-up

Next steps

- The mural will remain open for 24h, after which the content is moved to the Style Guide GitHub Issue page.
- All feedback on the Core Vocabulary XML Schemas is to be provided by **15 May 2024**
- Close all issues posted during the Core Vocabulary public review
- Release Core Vocabulary assets as SEMIC Recommendations

Survey on usage of SEMIC assets

We invite you to fill in the following survey which gauges the usage of SEMIC assets:

https://ec.europa.eu/eusurvey/runner/SEMICSpecificationAdoption


Thank you

interoperable europe

innovation ∞ govtech ∞ community

Stay in touch

(@InteroperableEU) / Twitter

Interoperable Europe - YouTube

Interoperable Europe | LinkedIn

DIGIT-INTEROPERABILITY@ec.europa.eu

https://joinup.ec.europa.eu/collection/interoperable-europe/interoperable-europe