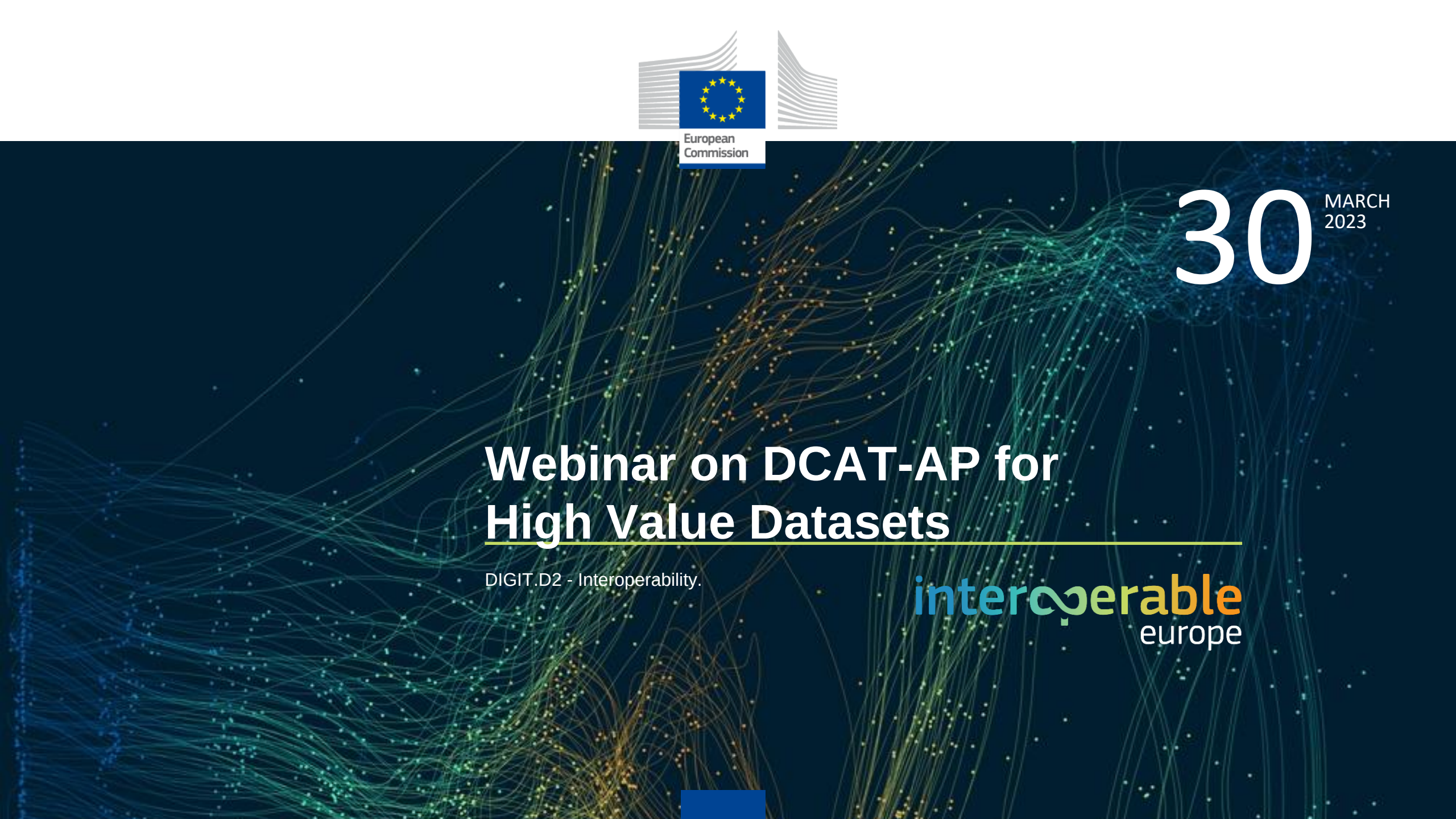

European Commission
30
MARCH
2023
Webinar on DCAT-AP for High Value Datasets
DIGIT.D2 - Interoperability.
interoperable europe

# Objective of this webinar

- Creating a common metadata descriptions for High Value Datasets in DCAT-AP in a collaborative way

# Agenda

1. **Welcome**
2. **Context of this webinar**
3. **General Implementation Approach**
4. **Proposal building DCAT-AP for HVD – Part 1**
5. **Short break**
6. **Proposal building DCAT-AP for HVD – Part 2**
7. **Wrap up and next steps**

# Workshop practicalities

Please mute your microphones

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

The workshop will be recorded, and will be shared publicly.

Introduce yourself in the chat:

- Name
- Affiliation
- Contact email address

# Context of this webinar

# Wrap up from previous webinar

- Governance of this activity
- Assessment of the High Value Dataset implementing regulation
- Introduction to DCAT-AP
- Discussion on topics

**Learnings**:

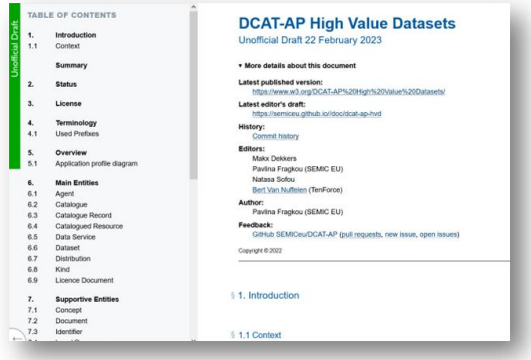
- Differences in implementation expectations for the HVD implementing regulation
- Community needs time to reflect on the different topics

GitHub: https://github.com/SEMICeu/DCAT-AP/issues

# Objective

A document that expresses how to apply DCAT-AP in order to satisfy the metadata requirements expressed in the Implementing Regulation for High Value Datasets.

Preliminary draft is shared as https://semiceu.github.io/DCAT-AP/releases/2.2.0-hvd/

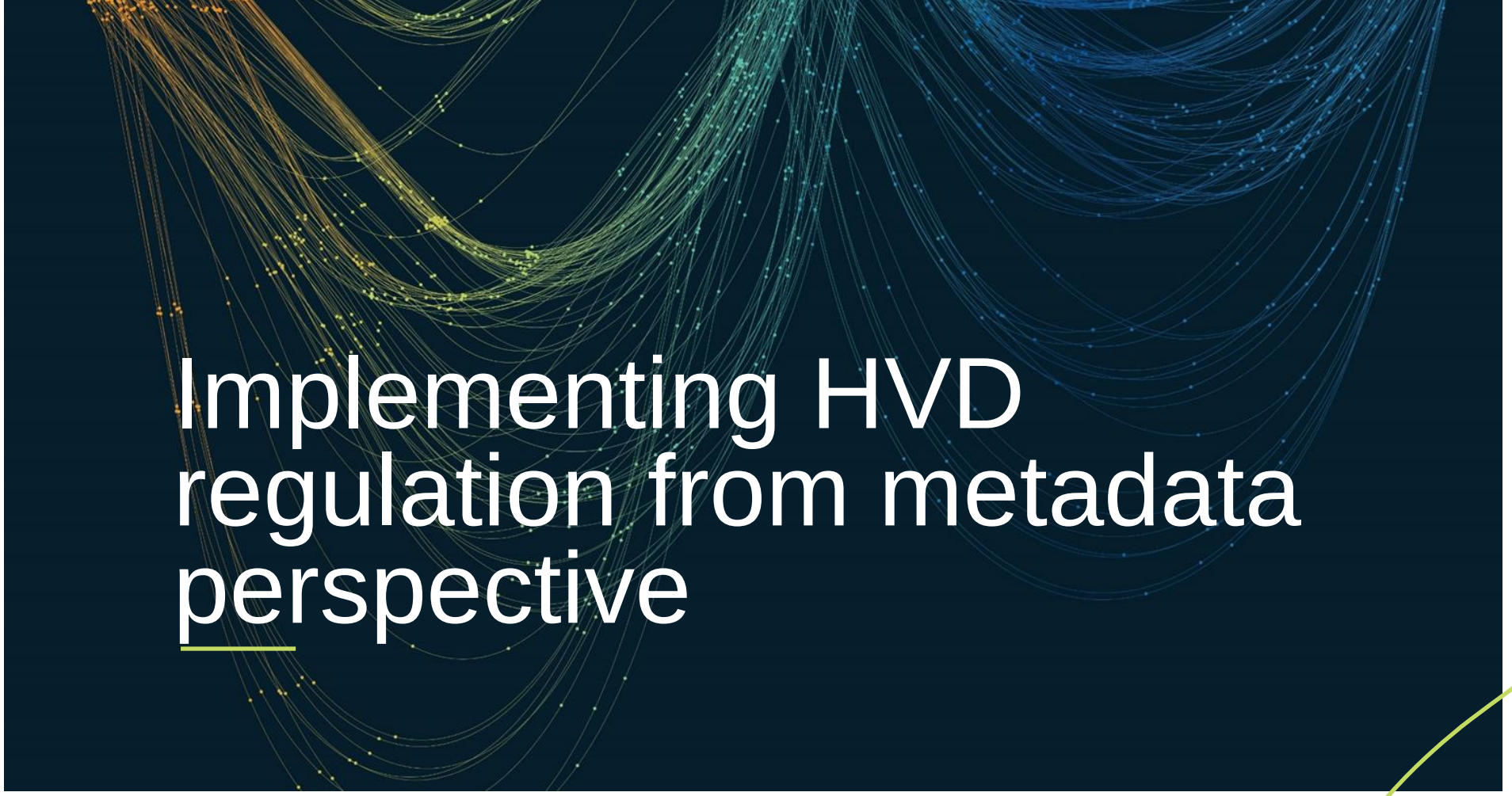



European Commission

# Governance

This activity is supported by

- DG CNECT for endorsement and alignment with policy implementation
- DIGIT/SEMIC for editorial and community facilitating services
- DG ENV & JRC for alignment with the geospatial community



# Implementing HVD regulation from metadata perspective

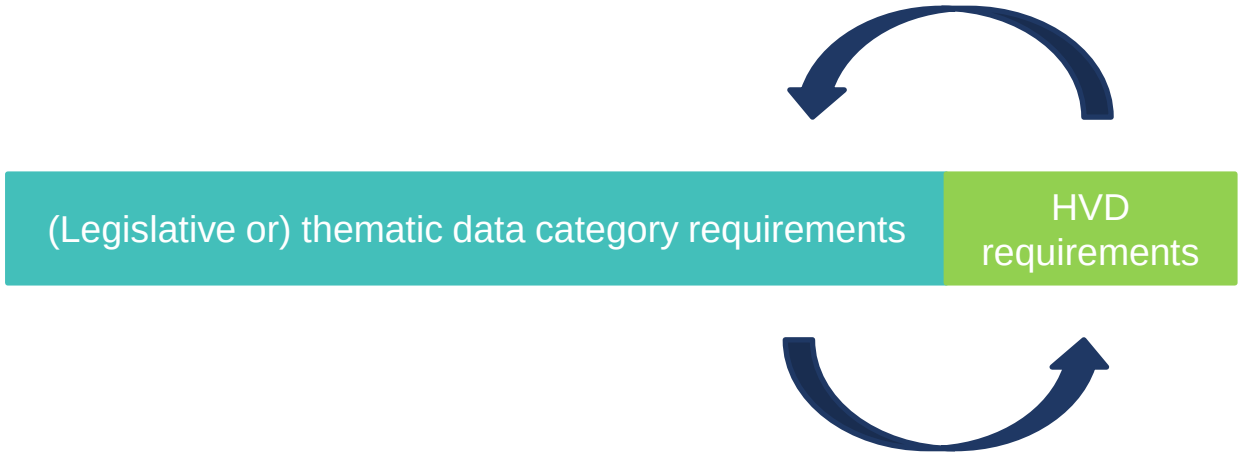
# HVD engagement

- HVD regulation calls for an (extra) engagement / commitment for the MS w.r.t. the identified datasets
- Identified datasets = a list of HVD provided by a MS

# A MS HVD list

- Each MS must organise itself to provide the list
- What is on the list cannot be retracted (persistent links)
- A collaboration across domains & portals:
  - Open data portals
  - INSPIRE (geo) portals
  - Statistical portals
  - …

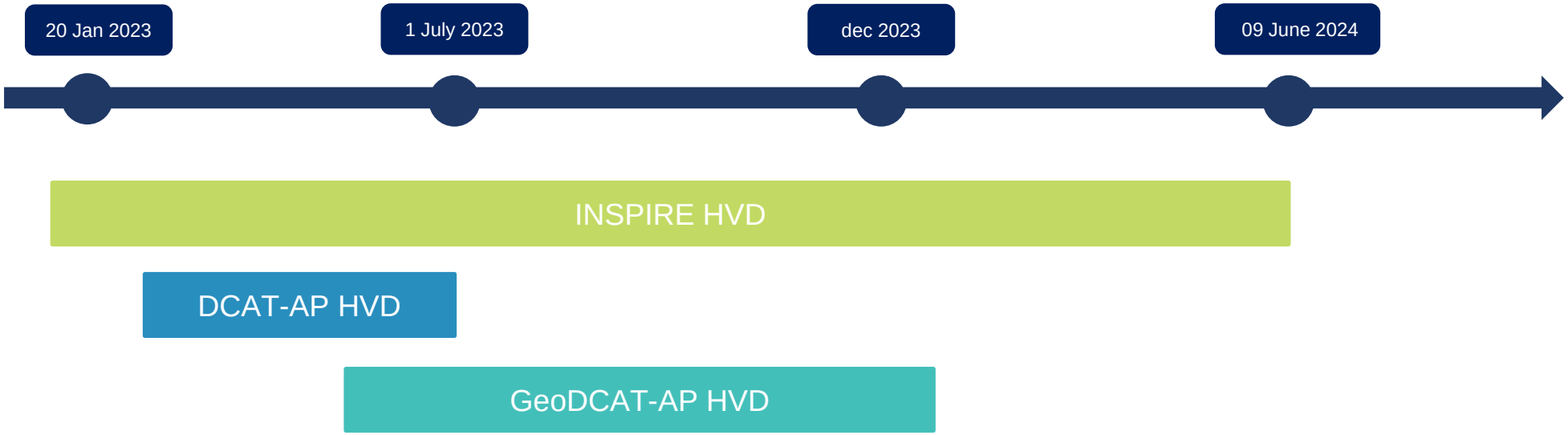
# Based on existing legal frameworks




HVD defines datasets organised by six thematic categories, each of which is defined by an existing legal framework.

- Interplay between domain and HVD requirements
- Specific requirements data category out of scope for today's discussion

# Geospatial data category as example




EC coordination (DG ENV,JRC,CNECT,DIGIT(SEMIC))

- Align implementation interpretations
- Improve collaboration between communities (i.e. geospatial data + (open) data policy + data portals)

# HVD reporting challenge

- The regulation states that the datasets must be documented via metadata according to the domain requirements
- MS should provide a report (implicitly in a common format)

→ What is the base information that can be used to create this report?

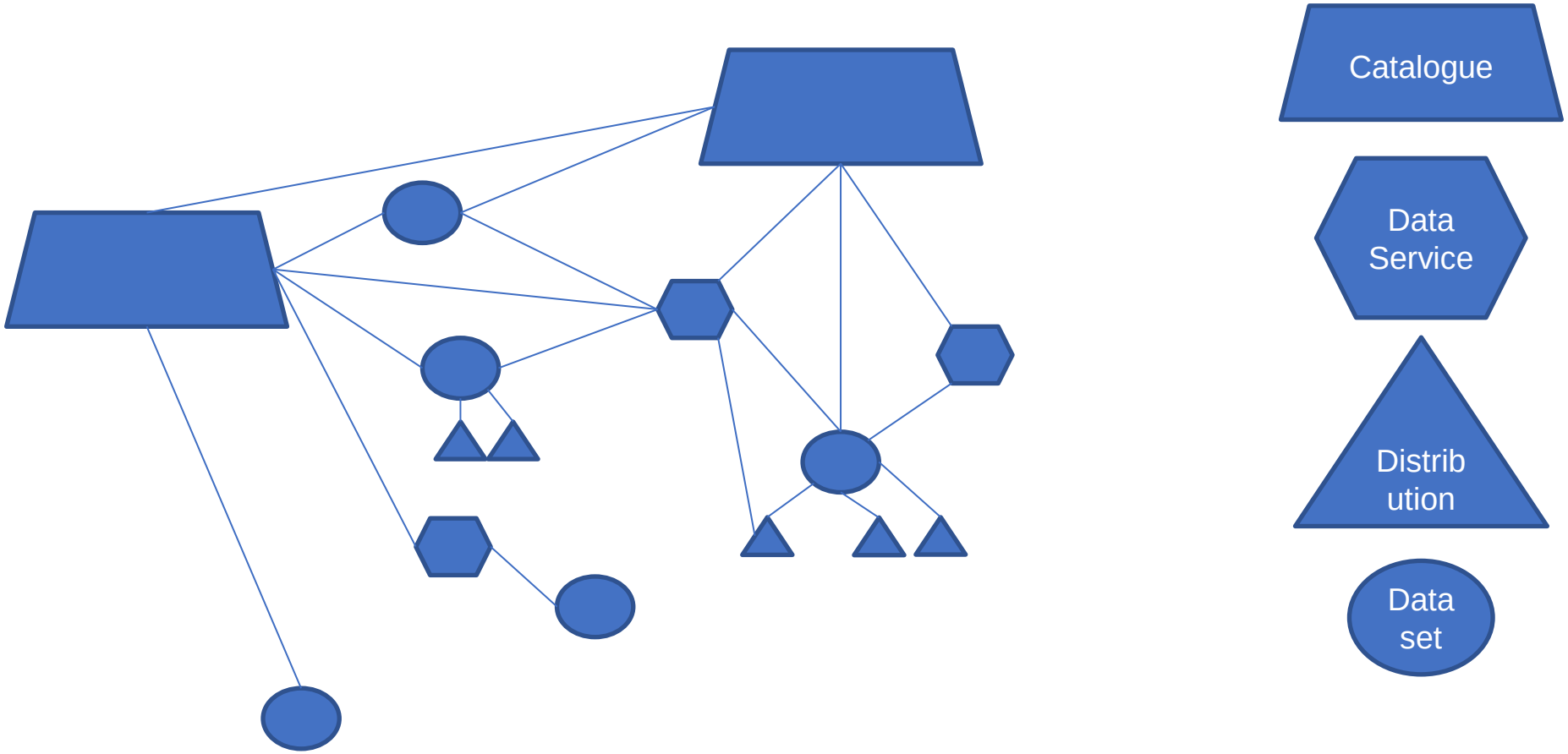
# DCAT-AP for HVD

- DCAT-AP is a metadata standard for describing catalogues of datasets and data services
- DCAT-AP is domain agnostic
- Objective of this activity: create DCAT-AP HVD which can capture the key notions of the HVD in a uniform way
  - To assist the implementation of the HVD regulation (helicopter view)
  - Link with real metadata (not composed at an office desk)
  - An EU-wide platform to align on the interpretation of HVD requirements

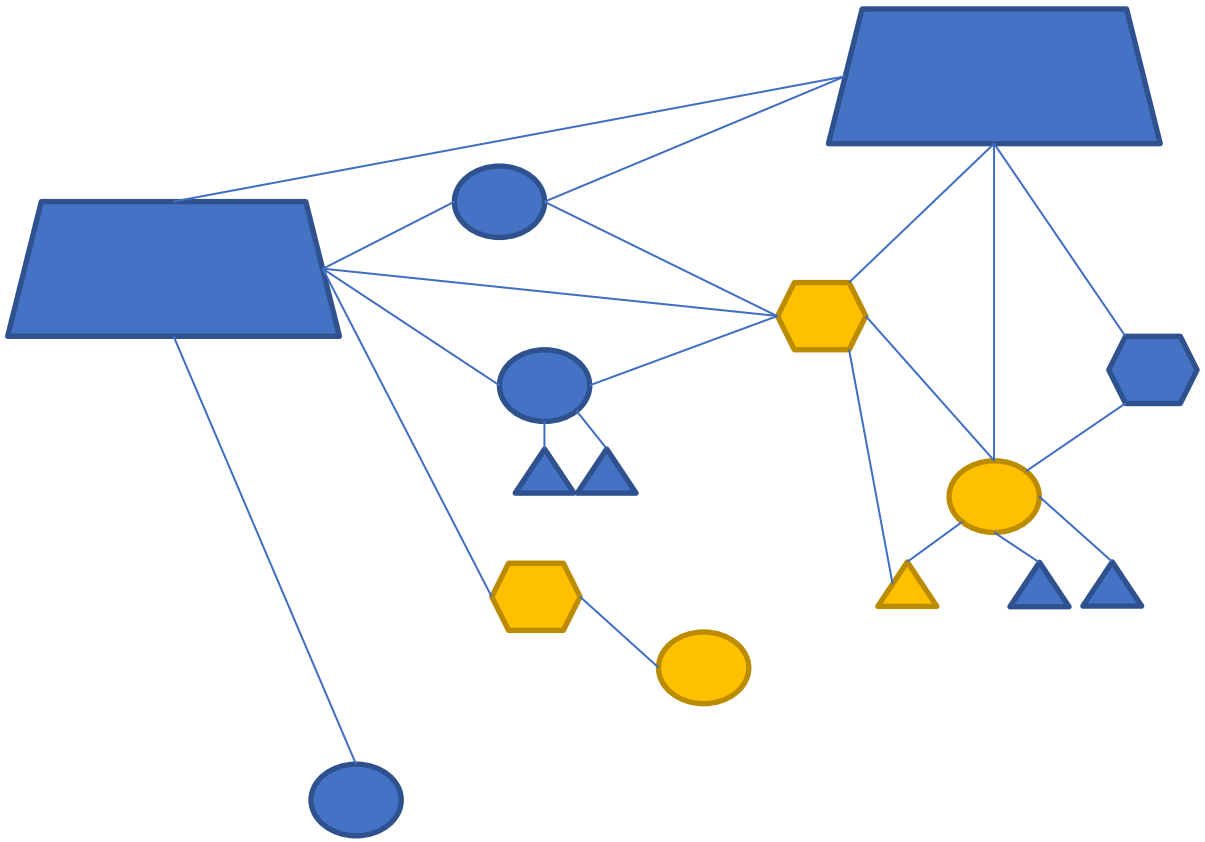
# Current state of affairs

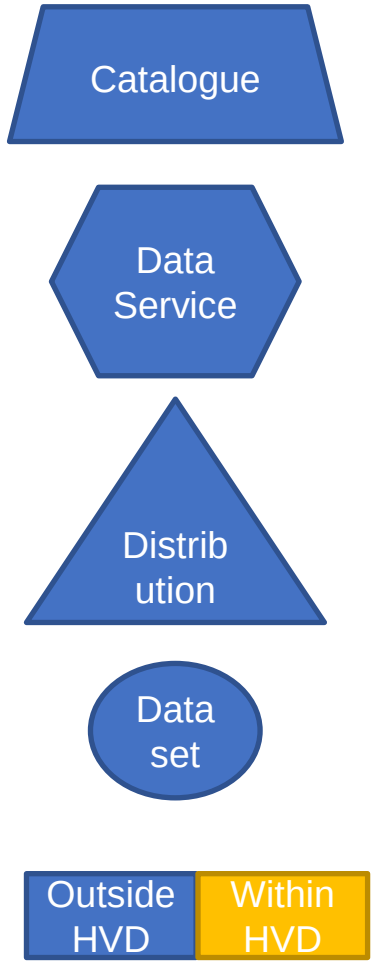
- **Bold statement**: data.europa.eu already contains all HVD
- **Portal perspective**: how to make these visible? What is their shared story?
- **Publisher perspective**: how can we augment the metadata so that it becomes HVD compliant with minimal impact on the metadata, but with maximal impact on the engagement

# Visual Representation
# (DCAT-AP creates a knowledge graph)

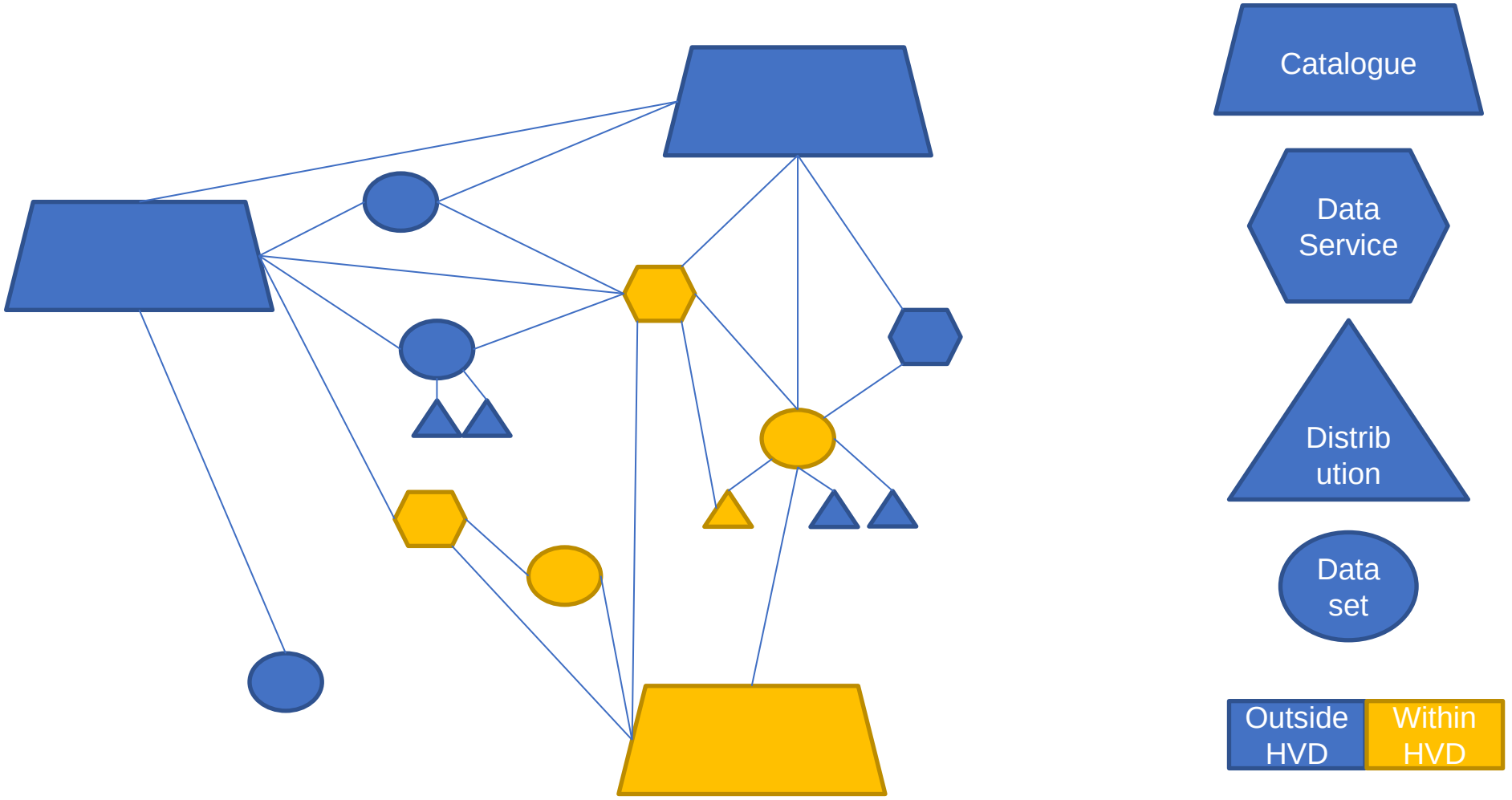


# Visual Representation (HVD indication)





# Visual Representation (HVD reporting)



# DCAT(-AP) is a knowledge graph

- DCAT(-AP) forms a loose network: a knowledge graph
- The notion (embedded in) or (in scope of) entity is not derivable from the knowledge graph. Each entity has its own life cycle
- From the DCAT-AP knowledge graph one cannot derive which entities are managed together (as an atomic unit)
- Portal content is not equal to the DCAT-AP knowledge graph
- Flexibility allows to capture many situations

# Common agreements

# Topics

C1. How to know which dataset is within the scope of the HVD directive?
C2. Reference to metadata descriptions
C3. Legal information

Break

C4. Bulk download
C5. API
C6. Point of Contact
C7. Adhere to specific information requirements
C8. How to know which MS is taking up the HVD responsibility

GitHub #260

# C1. In scope of a legislation

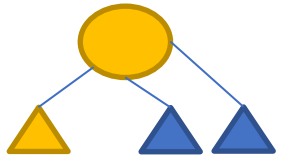
**Common use case:** legislative perspective on data

DCAT-AP does not have a clear way to state: *this entity addresses legal requirements derived from legislation.*

**Proposal:** a new property ***r5r:applicableLegislation***
with range eli:LegalResource
for the main classes: Dataset, Data Service, Distribution, Catalogue (and DatasetSeries).

```
<DGA_dataset> r5r:applicableLegislation <http://data.europa.eu/eli/reg/2022/868/oj>
<HVD_dataset> r5r:applicableLegislation <http://data.europa.eu/eli/reg_impl/2023/138/oj>
```

# C1. Part of HVD scope



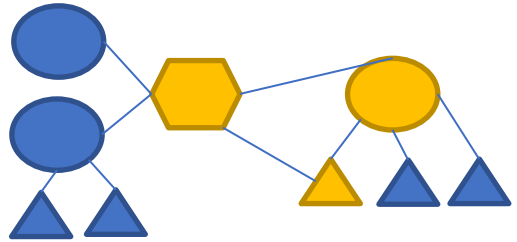
**Implementation guideline 1**

When a Dataset is within scope of HVD, it is not mandatory that all distributions are within scope of HVD.

**Motivation**

- Minimal impact on existing metadata in DCAT-AP
- Minimal impact on metadata documenting practices in other metadata frameworks (i.e. INSPIRE)
- Not imposed by the HVD implementing regulation
- It is a MS engagement

# C1. Part of HVD scope



**Implementation guideline 2**

When a Data Service offers access to multiple datasets and this Data Service fulfils the HVD requirements (e.g. *the HVD API for that dataset*) for a HVD then the HVD requirements apply only to that HVD.

**Impact**

- No requirement to change the data service setup because of the HVD implementing regulation (mixed use is allowed)
- No impact on the none-HVD that are served by this Data Service
- Nevertheless, via the HVD  which is served via this Data Service, operational and service level changes may be triggered or blocked.

**Motivation**

(same as for implementation guideline 1)

GitHub #251

# C1. Part of HVD scope

The HVD defines six thematic data categories:

*geospatial; earth observation and environment; meteorological; statistics; companies and company ownership; mobility*

**Proposal**

- Create new property **r5r:hvdCategory** defining **the HDV category to which this resource belongs**
- The codelist will be created and maintained by the Publications Office
- A resource may belong to more than one data category

GitHub #252

# C2. Identifiers for each HVD entity in the DCAT-AP metadata

Approved in webinar 1

Reporting requirement 5.3.a: with online reference to metadata

***Online***

- Dereferenceable: URI → URL (thus, no uuid)
- Guidelines on identifiers for DCAT-AP
- Best practices for URIs:
  - Organisation agnostic
  - System agnostic

GitHub #252

# C2. Identifiers for each HVD entity in the DCAT-AP metadata

Approved in webinar 1

**Implementation guideline 3**

The identifiers for a HVD can be different in different portals. What is the identifier to be provided?

Must rely on the whole community

- When metadata is maintained in a data category specific way (the master)

  → A mapping to DCAT-AP HVD should express how the source identifiers are represented in DCAT-AP HVD
- DCAT-AP guidelines for identifiers provide a recommendation for harvesters (and portals) to augment the list of alternative identifiers (adms:identifier) with the encountered identifiers

GitHub #253

# C3. Legal information in DCAT-AP

**Principle: *indicate legal information at the most precise level in the metadata description***

- Distribution and DataService, not at the Dataset or Dataset Series level
- Reasoning: the legal conditions for one representation can be different from another

**Usage Note**

- Use the codelist from the Publications Office, if no MS specific codelist is provided
- http://publications.europa.eu/resource/authority/licence/CC_BY_4_0 → Only machine-readable data

# C3. Legal information

**Basic guideline:**

The legal information shall be given by:

- Properties dct:license or dct:rights
- With a URI value (persistent link)
  → URI should be dereferenceable:
    - Machine-readable (provide RDF representation)
    - Human-readable text

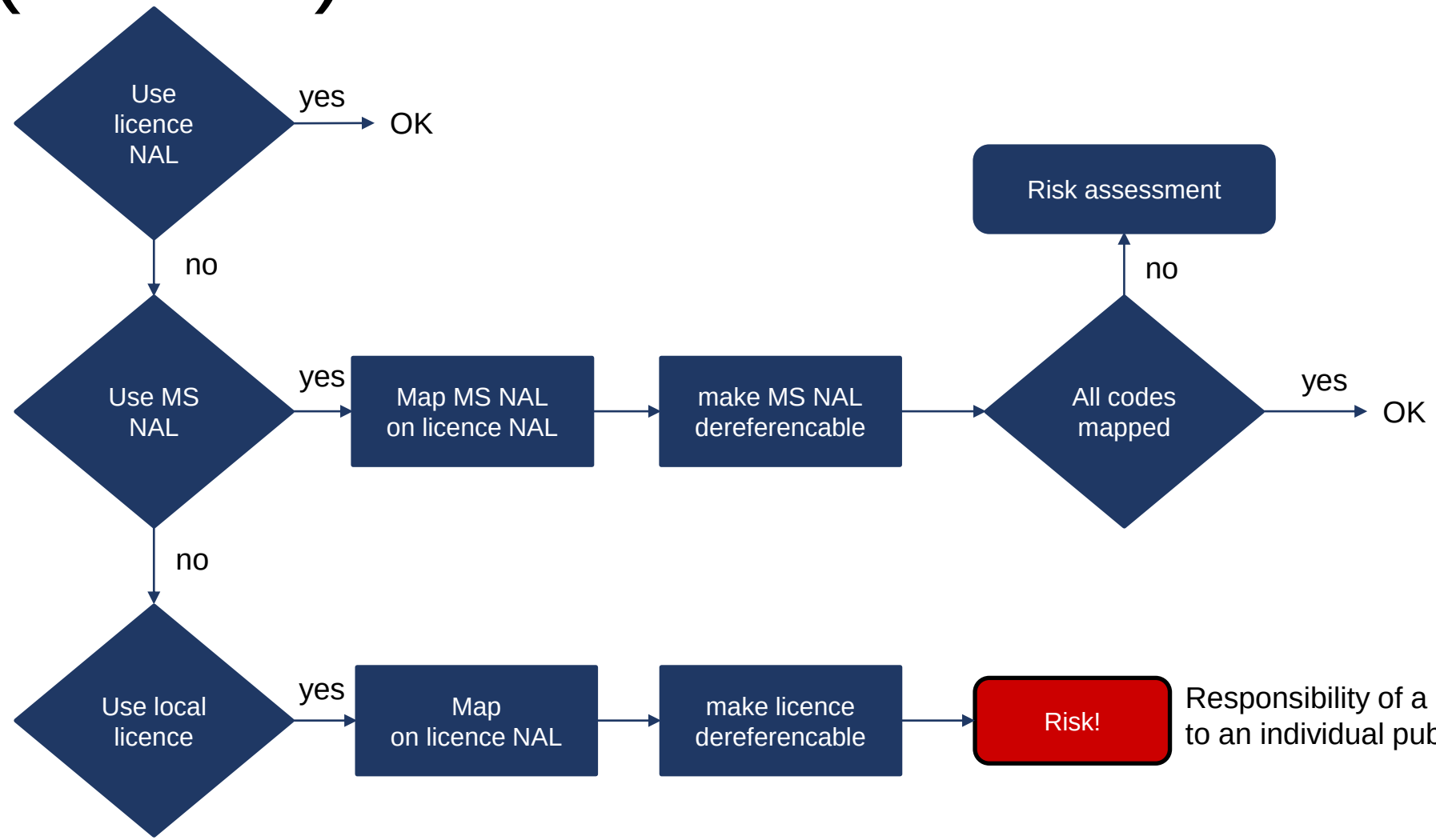
# C3. Legal information assessment objective

In the HVD reporting assessment, the provided legal information shall be assessed (using CC-BY 4.0 as reference).

**Proposal**: add mapping properties to *License Document & Rights Statement* to express a relationship with another licence / right.

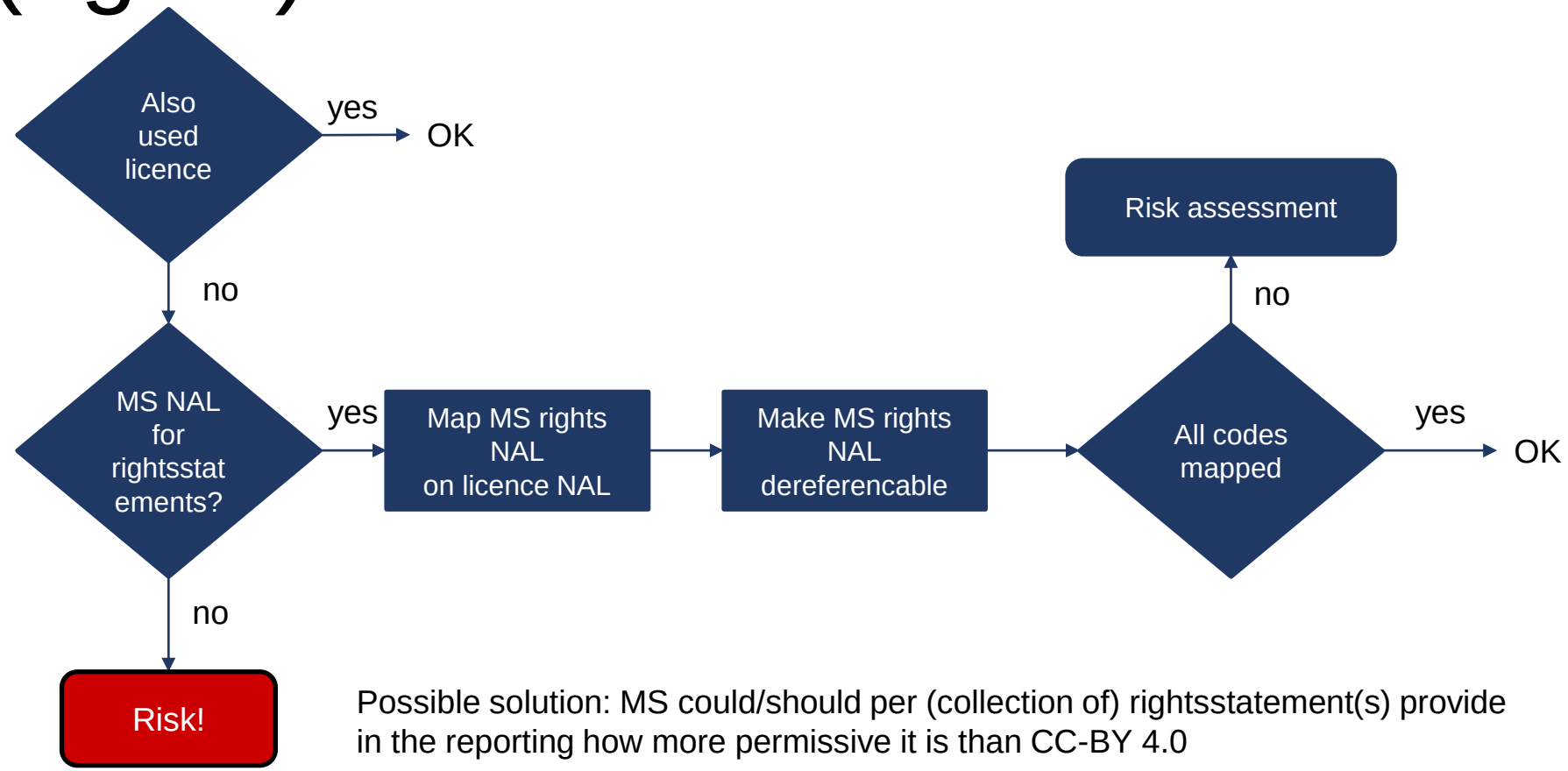
- Proposed properties: rdfs:seeAlso, owl:sameAs or skos matching properties
- restrict target list to the codelist of the Publications Office

# C3. Legal information workflow (licence)




Responsibility of a MS is shifted to an individual publisher

# C3. Legal information workflow (rights)




Possible solution: MS could/should per (collection of) rightsstatement(s) provide in the reporting how more permissive it is than CC-BY 4.0

# C3. legal information workflow

(In Annex) under the conditions of the Creative Commons BY 4.0 licence or any equivalent or less restrictive open licence

**Implementation guideline 5**

Although the HVD implementing regulation only uses the term licence, the formulation includes alternative expressions.

Thus the DCAT-AP approach of expressing rights instead of licence is acceptable.

# C3. legal information workflow

**Implementation guideline 5**

Other codelists for licence besides the PO NAL are allowed. However, they might fail all implementation requirements.

- E.g. can one use https://creativecommons.org/licenses/by/4.0/
- Issue is the requirement for dereferencable URIs
  → curl -H “accept:text/turle” https://creativecommons.org/licenses/by/4.0/

# Short break (5’)

GitHub #254

Approved in webinar 1

# C4. Bulk download

Bulk download is a functionality that allows to receive the whole dataset as a local copy.

- In DCAT-AP terminology: a Distribution.

**Proposal**

A HVD bulk download is denoted as a Distribution for a HVD.

# C4. Bulk download

Approved in webinar 1

**Assumption**

- The download URL is not always public, it might be only given on request.

**Use the current DCAT-AP practice**

- Property *access URL* is mandatory
- Property *download URL* is optional

GitHub #255

Approved in webinar 1

# C5. API representation

**Terminology**

- API = *a set of functions, procedures, definitions and protocols for machine-to-machine communication and the seamless exchange of data (source HVD)*
- Data Service = *a collection of operations that provides access to one or more datasets or data processing functions. (source: DCAT-AP)*

**Proposal**

A HVD API is a DataService

# C5. API representation

**DataService has properties:**

- Endpoint URL = *the root location or primary endpoint of the service (a Web-resolvable IRI)*.

  → i.e. the physical endpoint to be used
- Endpoint description = *a description of the services available via the end-points, including their operations, parameters etc.*

  → i.e. the documentation in the form of OpenAPI, OGC GetCapabilities,...

# C5. Reference

Reporting requirement 5.3.c: persistent link to APIs

DCAT-AP recommends persistency of the DataService URI, but does not impose any requirements on the Endpoint URL

**Proposal**

It is recommended to perform efforts to maintain persistency for both.

# C5. API

**The HVD mentions:**

- Terms of use
- Quality of service

# C5. API: Terms of use

**Proposal**

Documenting the *terms of use* is considered as the same activity as documenting the *legal information.*

# C5. API: Quality of Service

- DCAT-AP does **not** have a specific property for quality of service

- Quality of service covers a broad spectrum of topics

**Proposal**

**Add a usage note** on the generic documenting property **dcat:landingPage** that it should contain information or a reference to information about the quality of service.

GitHub #256

# C6. Point of contact

Public sector bodies shall designate **a point of contact** for questions and issues related **to the API** with a view to ensure the availability and maintenance of the API and ultimately the smooth and effective publication of the high-value datasets

## Proposal

A contact point is **mandatory** for *HVD DataServices* and **recommended** for *HVD* either in the form of *an (persistent) email-address* or a *webform address* (service desk).

GitHub #257

# C7. Adhere to specific data requirements

(Annex 3.2): the datasets shall be described in a complete and publicly available online documentation describing at least the data structure and semantics

## Proposed approaches

- **Option A**
  - Implicit
- **Option B**
  - Explicit with state of the art information

# C7. Adhere to specific data requirements

**Option A**

Implicit

- The associated label identifying it as a HVD is sufficient to know that it follows the directive
- No technical check can be performed if this is satisfied or not (self declaration)

**Option B**

Explicit with state of the art information

# C7. Adhere to specific data requirements

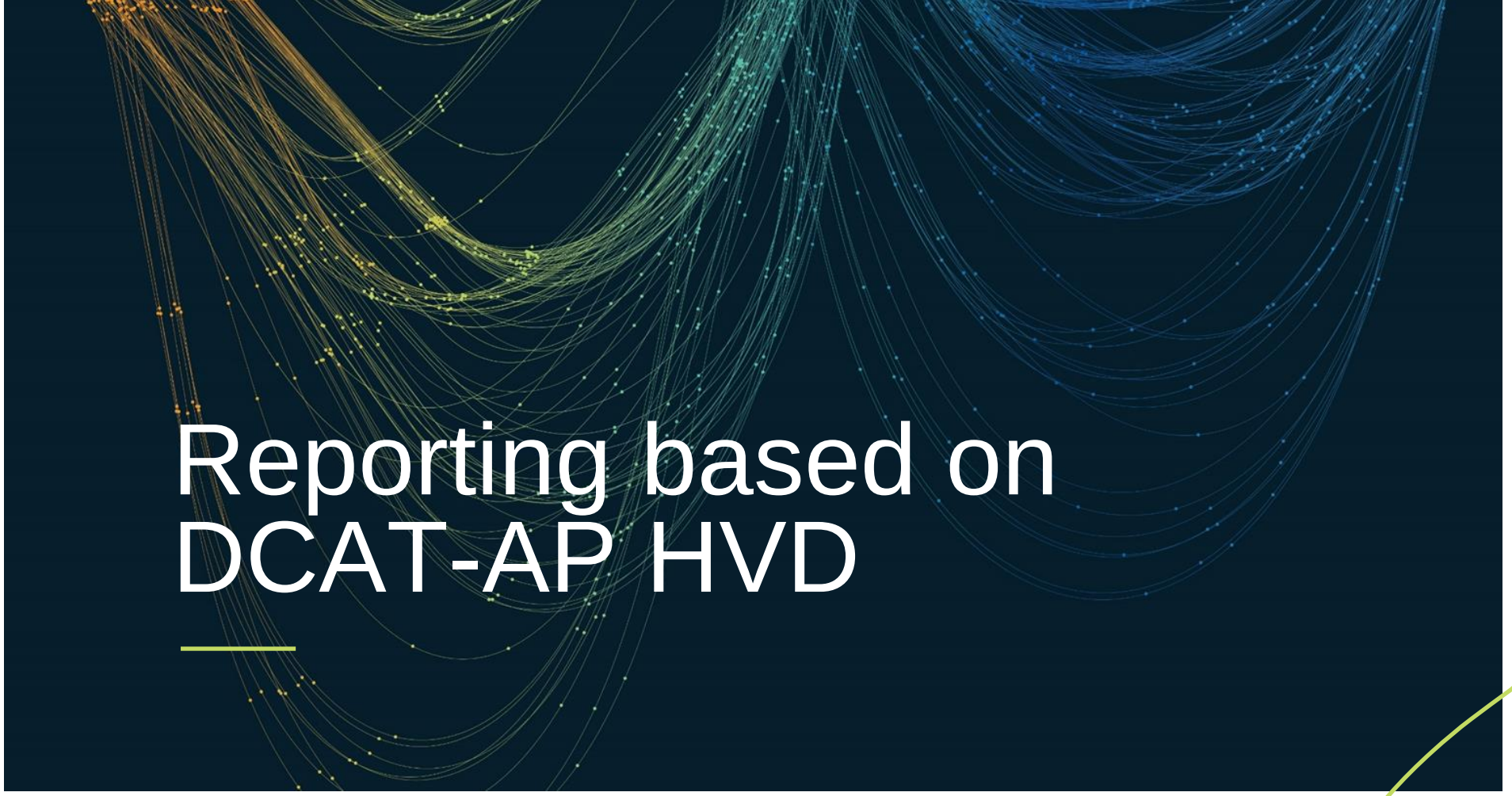
**Option A**

Implicit

**Option B**

Explicit with state of the art information

- (From DCAT-AP) use dct:conformsTo to express relationship with a **public** document (standard) that describes the internals of the resource.
- This ensures that the information about the content is made publicly accessible for reusers.
- Technical check can be performed by experts when using the provided information to assess if it matches the HVD requirements.

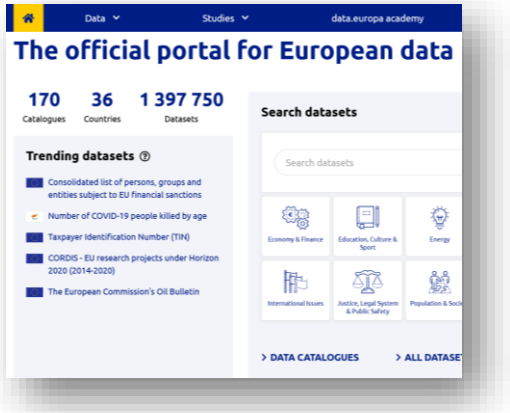


# Reporting based on DCAT-AP HVD

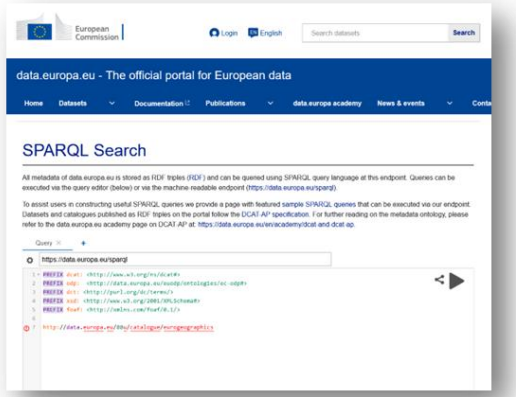
Share your reporting needs on GitHub

# Reporting

All metadata is harvested by data.europa.eu.




Data.europa.eu offers a SPARQL endpoint




**Objective**

Design queries that can highlight the HVD status for a national and EU-level.

# Reporting queries

- Find all datasets that are in scope of HVD
- Find all APIs that are in scope of HVD
- Find all bulk downloads that are in scope of HVD, including their category
- Find all datasets that are in scope of HVD and have a bulk download & API
- Find all licences assigned to entities in scope of HVD
- …

GitHub #258

# C8. Indicating MS

- DCAT-AP does not provide the notion of MS responsibility.
- MS data in data.europa.eu is the result of multiple catalogues.
- Some catalogues may provide data from multiple MS.
- DCAT-AP dct:spatial has a broader meaning than jurisdiction/legal responsibility, i.e. the “geographical coverage”.

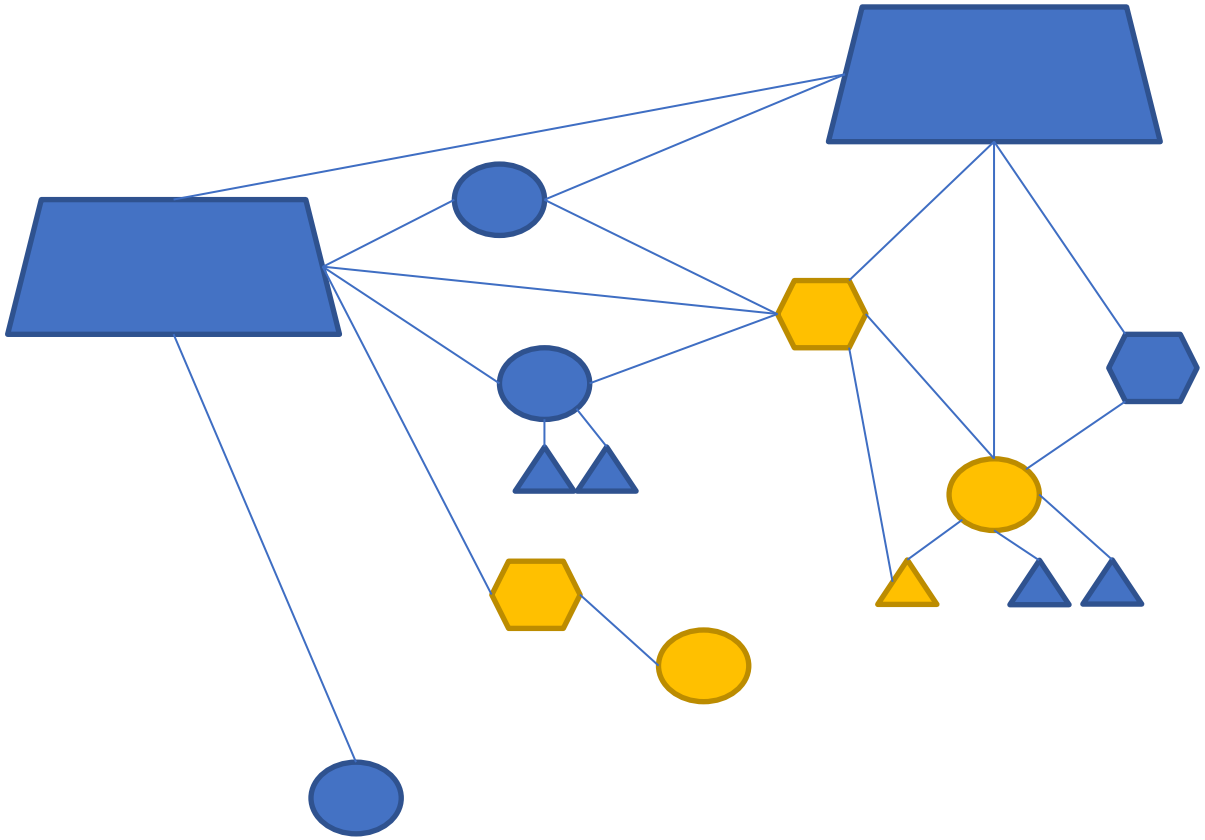
# C8. Indicating MS

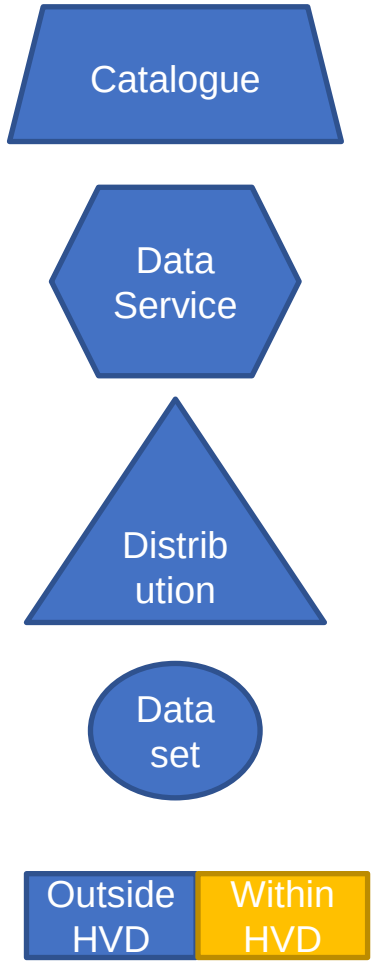
**Proposal**

MS provide a separate DCAT-AP HVD catalogue containing only the metadata that is relevant in the context of the HVD.

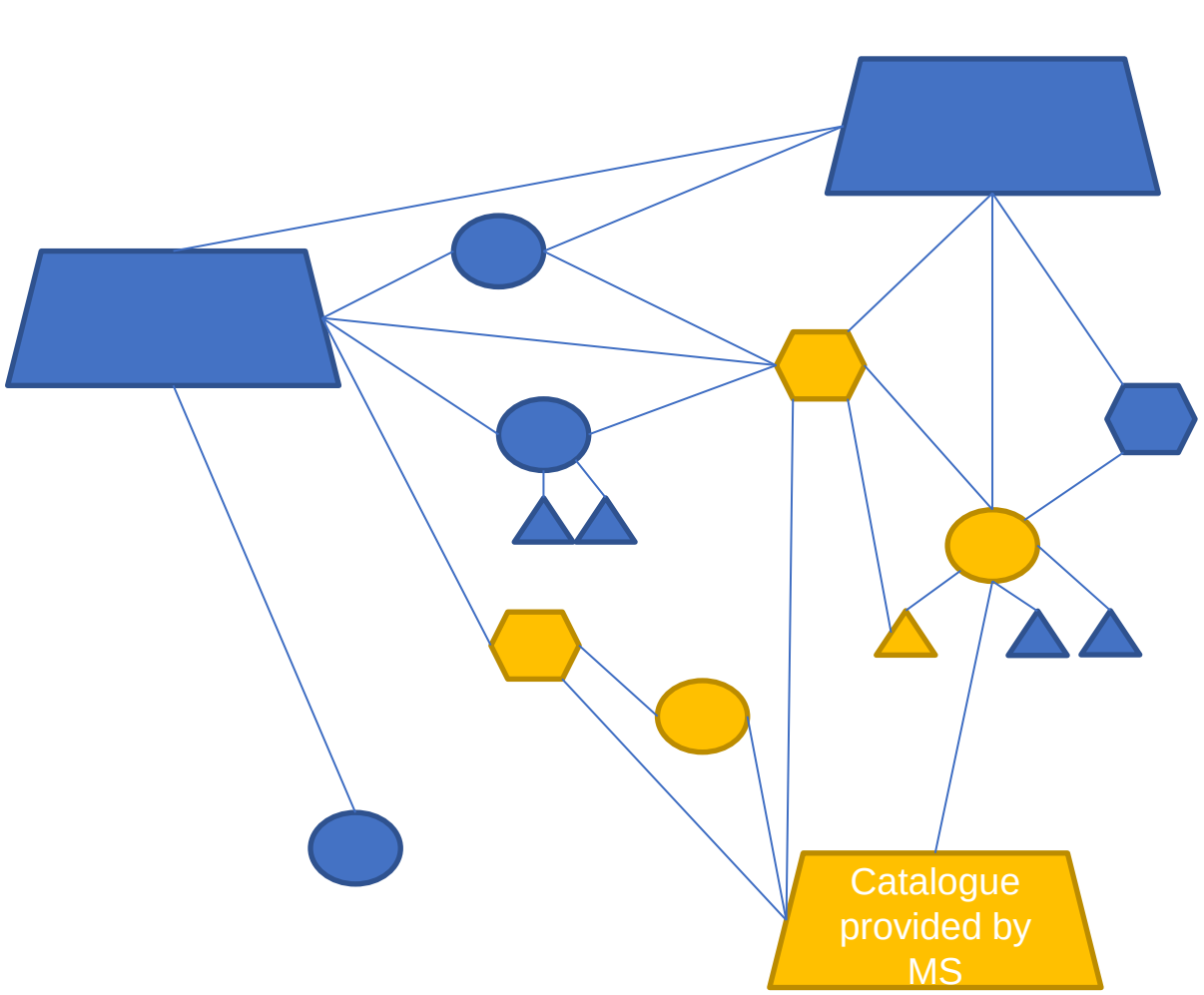
- Note: In this case, the HVD rules on persistent identifiers will never lead to duplicates in the aggregated data.europa.eu. The same data can be supplied via different ways.
- No new property introduced

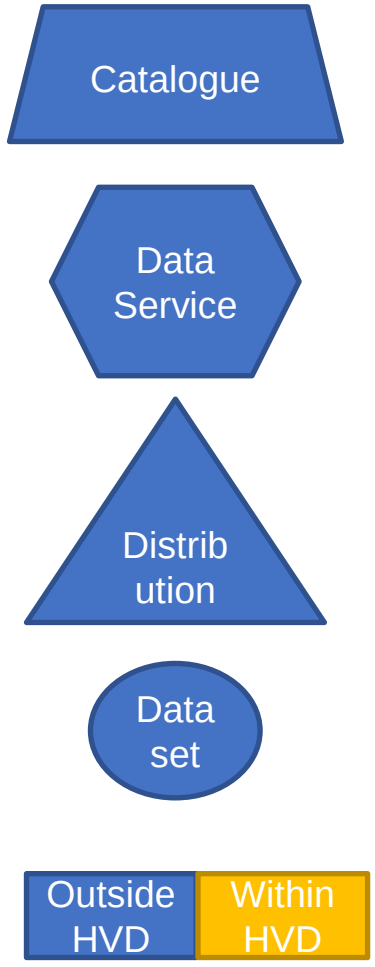
# Visual Representation





# Visual Representation



# C8. Indicating MS - example

```
geo.gov:cat a dcat:Catalog ;
   dcat:dataset geo.gov:d1.

geo.gov:d1 a dcat:Dataset;
   dct:title "Buildings in Gov";
   ...
   dcat:distribution geo.gov:d1-bulk.
```

```
data.gov:hvd a dcat:Catalog;
    dcat:dataset geo.gov:d1;
    ...
    dcat:dataset data.gov:dbus.
```

```
data.gov:cat a dcat:Catalog ;
   dcat:dataset data.gov:dbus1.

data.gov:dbus1 a dcat:Dataset;
   dct:title "Business register of Gov" ;
   ...
   dcat:distribution data.gov:dbus1-bulk.
```

# C8. Indicating MS

## Proposal

MS provide a separate DCAT-AP HVD catalogue containing only the metadata that is relevant in the context of the HVD.

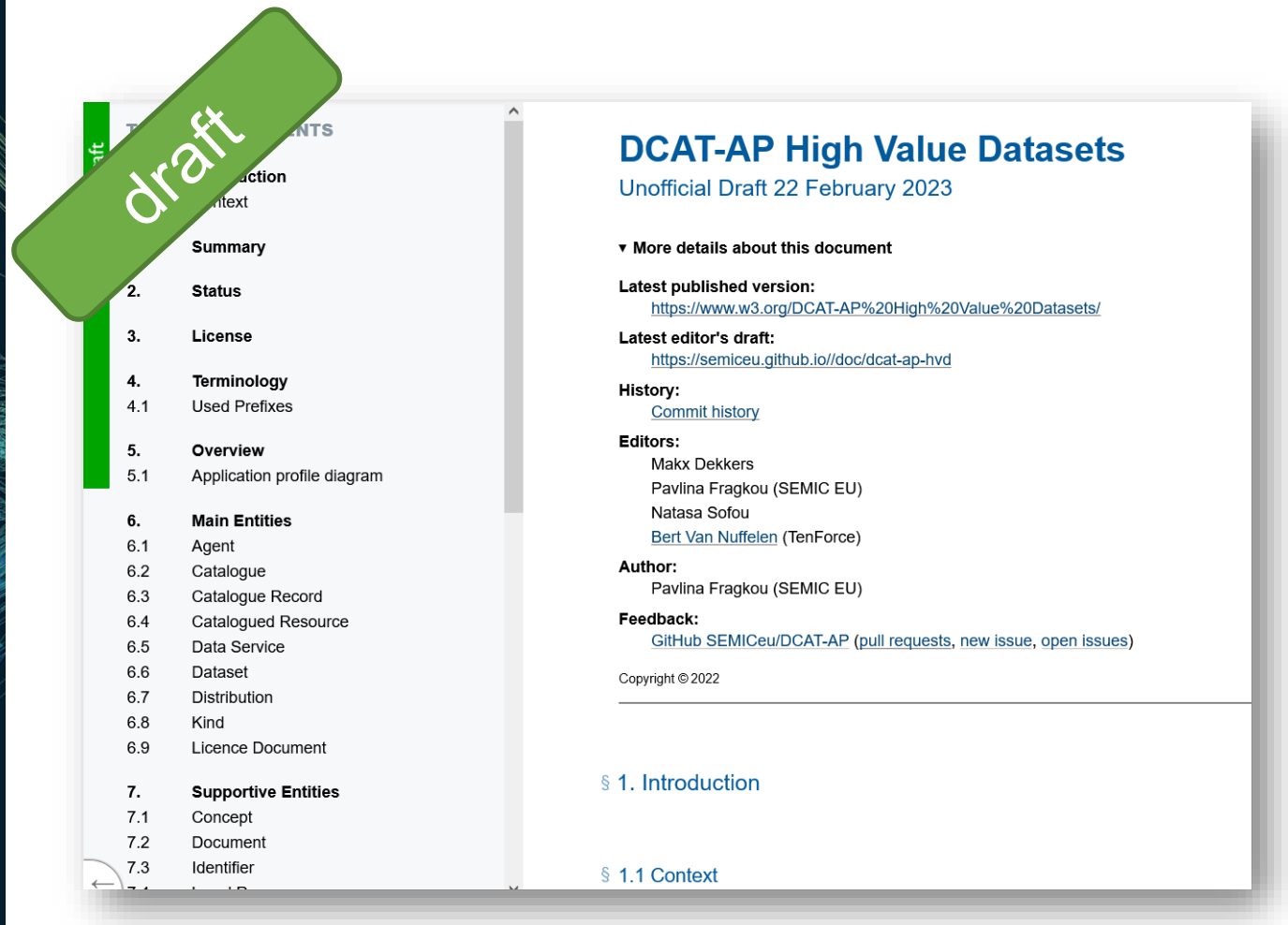
- Note that in case the HVD rules on persistent identifiers this will never lead to duplicates in the aggregated data.europa.eu. The same data can be supplied via different ways.
- No new property introduced

- Clear scope of reporting
- Use of persistent identifiers is mandatory
- MS can include information from other catalogues

- Portal system implementers may have to provide support for such separate catalogue.
- Visitors to a portal who would like to see the MS perspective need access to this catalogue. This must be provided by portal implementers.

# Wrap-up and next steps

# Walkthrough guidelines document




We welcome feedback!

Decisions and feedback of today will be integrated

https://semiceu.github.io/DCAT-AP/releases/2.2.0-hvd/

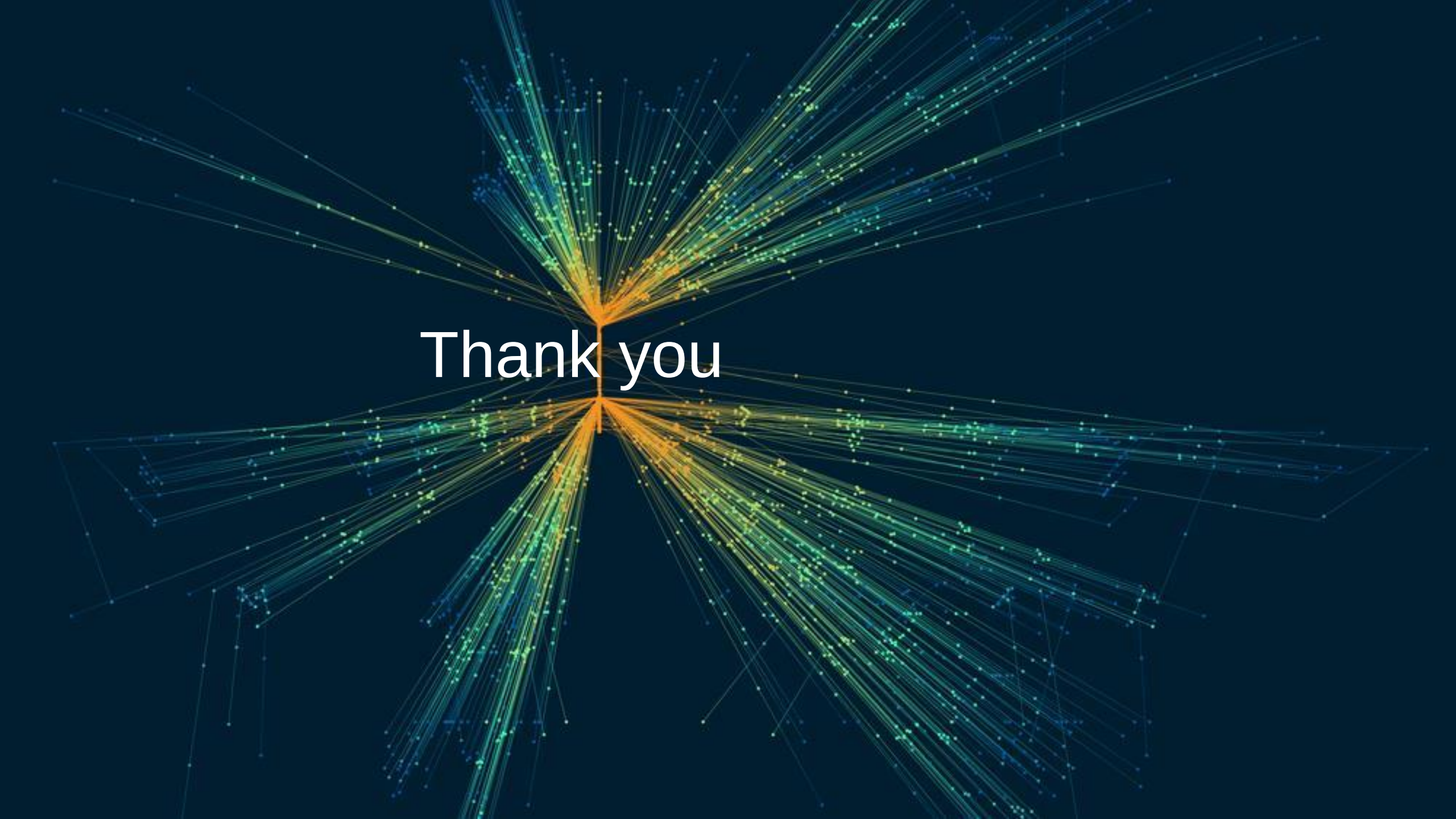
# DCAT-AP HVD guidelines

- Provides a global perspective on the metadata requirements (the publishing context) expressed by the HVD implementing regulation
- Useful as **support** for the **MS coordination** on the implementation across the different data categories
- A step towards the reporting that must happen

# Wrap up

- Discussed 8 proposals on how to apply DCAT-AP in order to satisfy the general metadata requirements expressed in the Implementing Regulation for High Value Datasets.
- Available for discussion on GitHub (open issues)
- Welcome any feedback on the preliminary draft https://semiceu.github.io/DCAT-AP/releases/2.2.0-hvd/
- All participants are mentioned in the **acknowledgements** section → Contact us if you do not want to be part of the acknowledgment section.

https://github.com/SEMICeu/DCAT-AP/issues


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

# Proposal summary

# Additional profile statements: HVD

| property | uri | range | card | definition | Usage note |
|---|---|---|---|---|---|
| **HVD category** | **r5r:hvdCategory** | **Concept** | **1..n** | **The HDV category to which this dataset belongs** | **The codelist <HVDCategory> must be used** |
| contact point | dcat:contactPoint | **HDV Kind** | **0..n (rec)** | contact information that can be used for sending comments about the Dataset | |
| conforms to | dct:conformsTo | Standard | **1..n** | refers to an implementing rule or other specification. | **The provided information should enable to the verification whether the detailed information requirements by the HVD is satisfied.** |
| Dataset distribution | dcat:distribution | **HVD BulkDownload Distribution** | **1..n** | links the Dataset to an available Distribution | **At least one of the associated distributions must satisfy the Bulk Download requirement of HVD.** |

# Additional profile statements: HVD Bulk download Distribution

| property | uri | range | card | definition | Usage note |
|---|---|---|---|---|---|
| **HVD category** | **r5r:hvdCategory** | **Concept** | **1..n** | **The HDV category to which this dataset belongs** | **The codelist <HVDCategory> must be used** |
| conforms to | dct:conformsTo | Standard | **1..n** | refers to an implementing rule or other specification. | **The provided information should enable to the verification whether the detailed information requirements by the HVD is satisfied.** |
| Access URL | Dcat:accessURL | Resource | 1..n | a URL that gives access to a Distribution of the Dataset. The resource at the access URL may contain information about how to get the Dataset. | |
| Download URL | Dcat:downloadURL | Resource | 0..n | a URL that is a direct link to a downloadable file in a given format. | |
| Licence | Dct:license | **HVD Licence Document** | **1..1** | the licence under which the Distribution is made available. | **The licence must be a dereferenceable URI, provided as machine and human readable description.** |

# Additional profile statements: HVD Licence Document

| property | uri | range | card | definition | Usage note |
|---|---|---|---|---|---|
| Same as | owl:sameAs | URI | 0..n | An alternative formulation for the licence with the same legal conditions | From a legal perspective both licences are the same. |
| For more information | rdfs:seeAlso | URI | 0..n | For a more information about the legal statements. | The provided information gives more insight in the legal conditions. |

# Additional profile statements: HVD Contact Point

| property | uri | range | card | definition | Usage note |
|---|---|---|---|---|---|
| Contact page | Foaf:page | URI | 0..1 | A webpage that either allows to make contact (i.e. a webform) or the information contains how to get into contact. | |
| email | Vcard:hasEmail | URI | 0..1 | A email address via which one can make contact. | |

# Additional profile statements: HVD Data Service

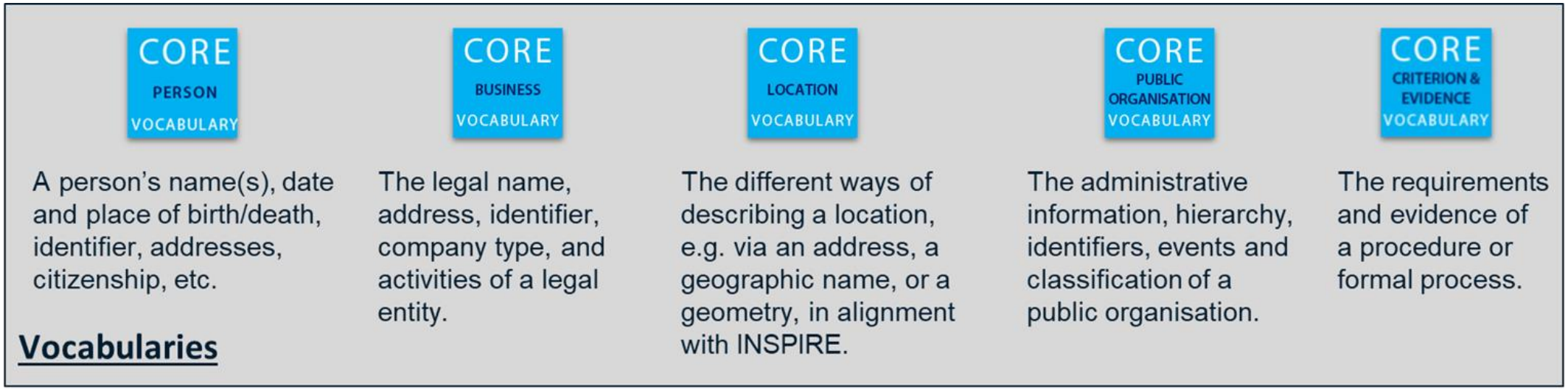
| property | uri | Range | Card | definition | Usage note |
|---|---|---|---|---|---|
| **HVD category** | **r5r:hvdCategory** | **Concept** | **1..n** | **The HDV category to which this dataset belongs** | **The codelist <HVDCategory> must be used** |
| conforms to | dct:conformsTo | Standard | **0..n (rec)** | refers to an implementing rule or other specification. | **The provided information should enable to the verification whether the detailed information requirements by the HVD is satisfied.** |
| Endpoint description | dcat:enpointDescription | Resource | **1..n** | a description of the services available via the end-points, including their operations, parameters etc.The property gives specific details of the actual endpoint instances, while dct:conformsTo is used to indicate the general standard or specification that the endpoints implement. | |
| Endpoint URL | Dcat:endpointURL | Resource | 1..n | The root location or primary endpoint of the service (an IRI). | |
| Licence | Dct:license | **HVD Licence Document** | **1..1** | the licence under which the Distribution is made available. | **The licence must be a dereferenceable URI, provided as machine and human readable description.** |

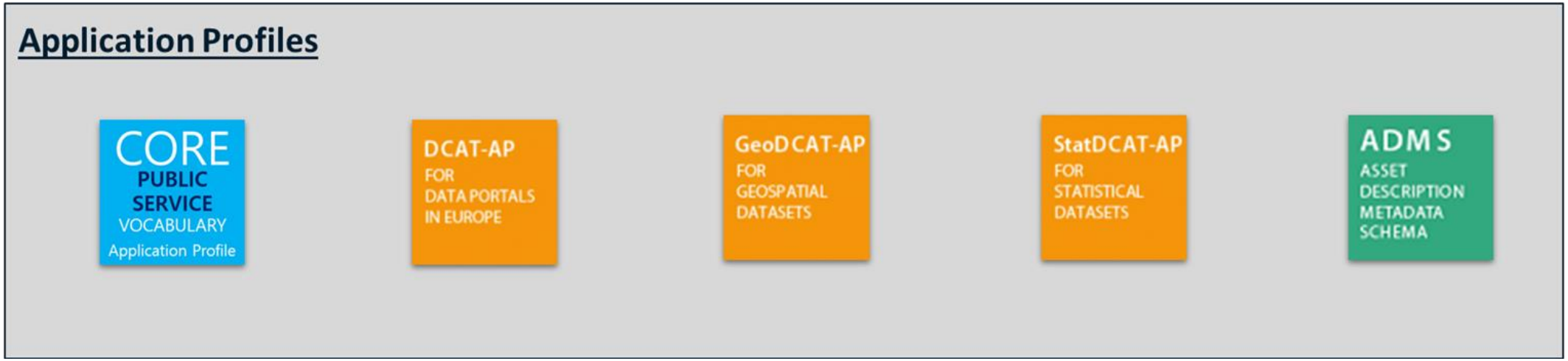
# Introduction to SEMIC

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

- Promoting, 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.

# Current SEMIC assets



# Objectives of DCAT-AP

- Supporting the discovery of/access to (open) data in a cross-border and cross-domain environment, by harvesting data from distributed portals.

- In the form of an application profile of W3C DCAT, by:
  - Expressing constraints and usages on DCAT properties and classes, and
  - Including additional properties and usages of controlled vocabularies

  In such a way that the metadata descriptions are maximally harmonised across Europe, and provide a reliable source for the European Data portal.

- Extensions exist to serve different communities better: BregDCAT-AP, GeoDCAT-AP, StatDCAT-AP.

# Expected outcome of today's activity

A document that expresses how to apply DCAT-AP in order to satisfy the metadata requirements expressed in the Implementing Regulation for High Value Datasets.

**Note**: not foreseen that this becomes an extension, such as GeoDCAT-AP, but more an expression of use.

# Governance

This activity is supported by

- DG CNECT for endorsement and alignment with policy implementation
- DIGIT/SEMIC for editorial and community facilitating services

This document will become a part for the DCAT-AP ecosystem

- SEMIC will take care that this document will be up to date with the latest DCAT-AP evolution.
- **Note:** *as the objective is to impose specific usage rules on the existing DCAT-AP there is a long term stability expected*.

# Community

Initiated as a joined activity of the

- PSI expert group (DG CNECT)
- DCAT-AP community (SEMIC)

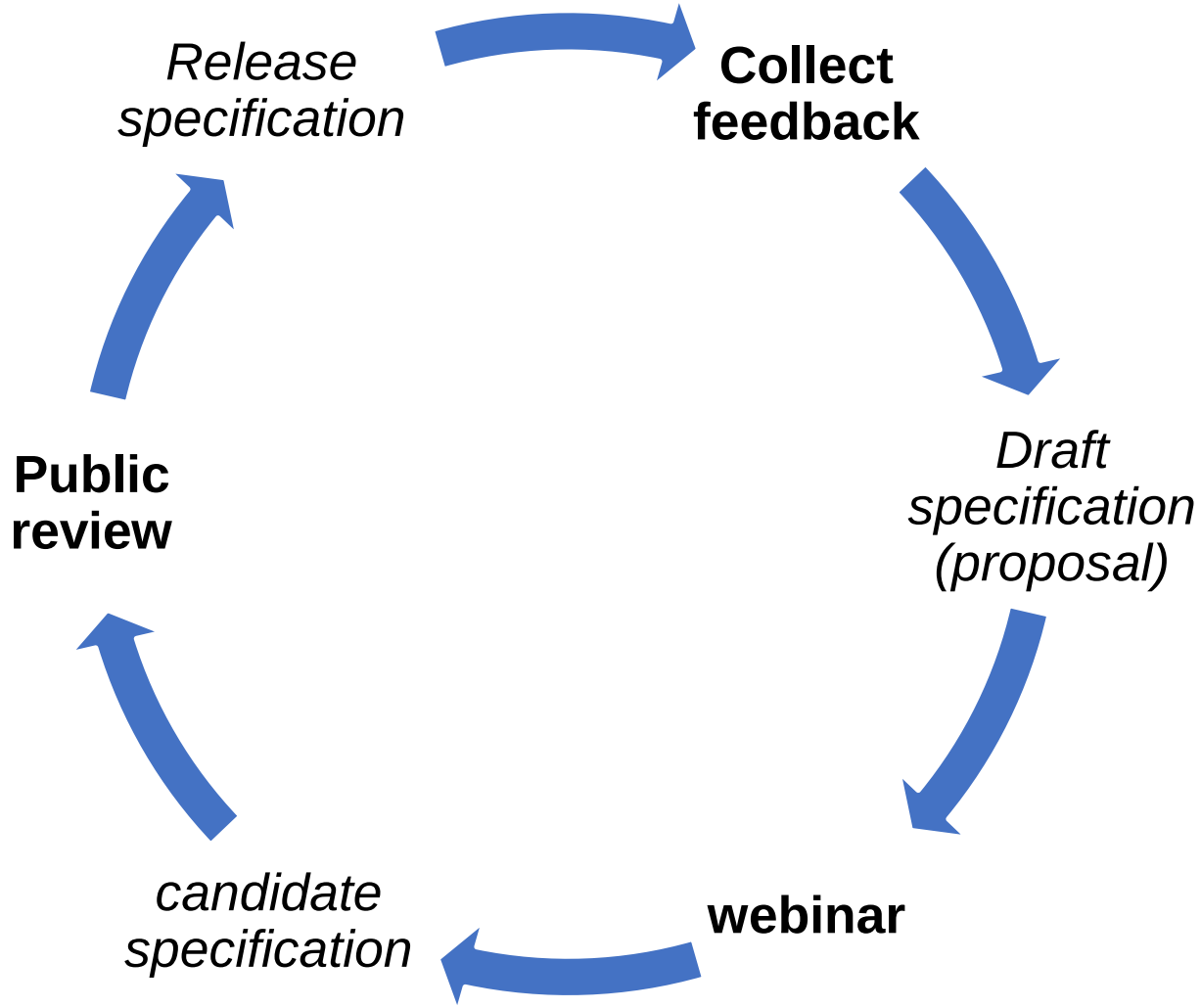
Feedback via

- GitHub
- Regular webinars

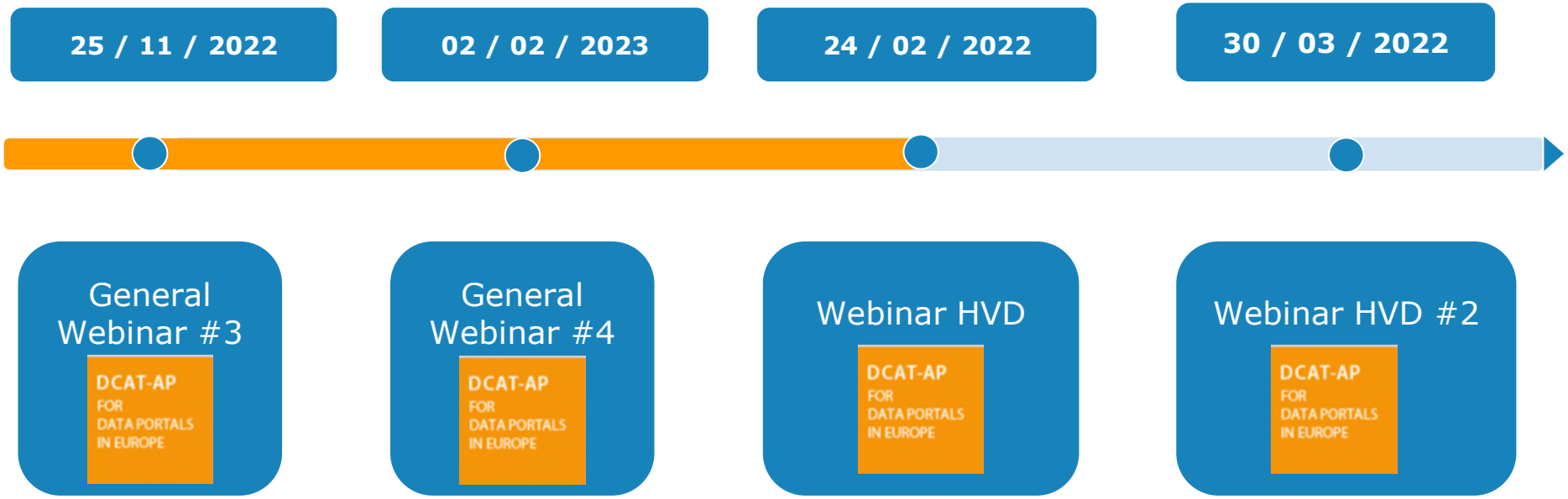
DCAT-AP data specification consists of

- Human readable text
- Support for implementers:
  - JSON-LD context
  - SHACL shapes for validation

# Consensus building



# Webinars


25 / 11 / 2022
02 / 02 / 2023
24 / 02 / 2022
30 / 03 / 2022
General Webinar #3
DCAT-AP FOR DATA PORTALS IN EUROPE
General Webinar #4
DCAT-AP FOR DATA PORTALS IN EUROPE
Webinar HVD
DCAT-AP FOR DATA PORTALS IN EUROPE
Webinar HVD #2
DCAT-AP FOR DATA PORTALS IN EUROPE

# Ongoing activities

Issue handling

DCAT 3.0

High Value Datasets

# Long term planning

- Further addressing issues raised by the community
- Supporting data spaces
- Supporting other legislative requirements

# Participation (today) in this activity

- Assess the proposals and share your perspective

- How?
  - Each proposal will be introduced
  - Followed by an open discussion
  - If consensus reached, proposal is adopted.
  - If no consensus, or discussion takes too long
    - Take it off-line to github
    - Develop for the next webinar a new proposal.

- More feedback, questions, suggestions, ...
  - Meeting minutes and recording will be shared
  - File them on GitHub (or reach out via any of the channels) for open discussion with the community,

# HVD metadata requirement assessment

# Legislative background

## Open Data Directive

*https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019L1024*

## Implementing Regulation for High Value Datasets

*https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023R0138*

# General requirements (article 3)

1. *Public sector bodies holding high-value datasets listed in the Annex shall ensure that the datasets described or referenced in the Annex are made available in machine-readable formats via **APIs** corresponding to the reasonable needs of re-users. Where indicated in the Annex, the datasets shall also be made available as **a bulk download**.*
2. *Those public sector bodies referred to in paragraph 1 shall set out and publish the **terms of use** of the API and **the quality of service** criteria on its performance, capacity and availability. The terms of use shall be available **in a human-readable and machine-readable format**. Both the terms of use and the quality of service criteria shall be compatible with the arrangements for the re-use of high-value datasets laid down in accordance with Article 4.*
3. *API terms of use shall be accompanied **by API documentation in a Union or internationally recognised open, human-readable and machine-readable format**.*

# General requirements (article 3)

4. *Public sector bodies referred to in paragraph 1 shall designate **a point of contact** for questions and issues related to the API with a view to ensure the availability and maintenance of the API and ultimately the smooth and effective publication of the high-value datasets.*
5. *Public sector bodies holding high-value datasets listed in the Annex shall ensure that the datasets are **denoted as high-value datasets** in their metadata description*

# Requirements Assessment Annex

**Publish datasets with**

- ‘open legal conditions’, i.e. CC-BY 4.0
- With a bulk download
- With an API
- According to the best practices and agreements imposed by the relevant legislation

# Requirements Assessment Annex

**Publish the identified datasets with**

- ‘Open legal conditions’, i.e. CC-BY 4.0
- With a bulk download
- With an API
- According to the best practices and agreements imposed by the relevant legislation
- **Identified by a circumscription of the content**

# Example **identification**

**The HVD for Buildings are**

- The building datasets that are in scope of the INSPIRE data theme building (Directive 2007/2/EC)
- With the granularities up to the scale of 1:5000
- And must provide the following content:
- Unique identifier;
  - Geometry (footprint of the building);
  - Number of floors;
  - Type of use.

# Requirements Assessment Annex

**Publish the identified datasets with**

- ‘Open legal conditions’, i.e. CC-BY 4.0
- With a bulk download
- With an API
- According to the best practices and agreements imposed by the relevant legislation
- Identified by a circumscription of the content
- **With documentation about the structure and semantics of the data**

  *The datasets shall be described in a complete and publicly available online documentation describing at least the data structure and semantics (e.g. Annex 3.2.c)*

# Reporting requirements (article 5.3)

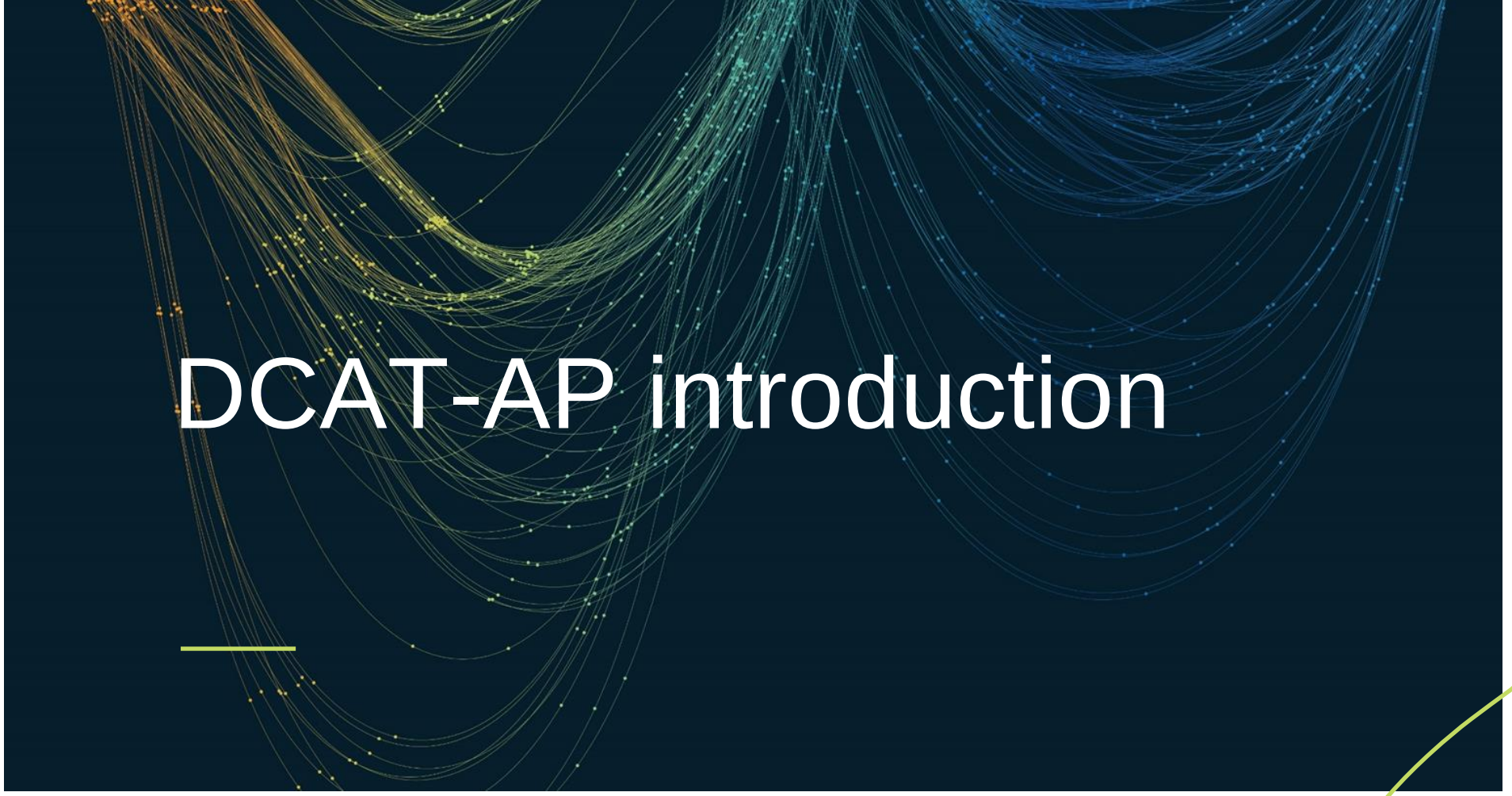
1. A **list of specific datasets** at Member State level (and, where relevant, subnational level) corresponding to the description of each high-value dataset in the Annex to this Regulation and with **online reference to metadata** that follow existing standards, such as a single register or open data catalogue;
2. A **Persistent link** **to the licensing conditions** applicable to the re-use of high-value datasets listed in the Annex to this Regulation, **per dataset** referred to in point a);
3. A **Persistent link** to **the APIs** ensuring access to the high-value datasets listed in the Annex to this Regulation, per dataset referred to in point a);

# Reporting requirements (article 5.3)

4. Where available, guidance documents issued by the Member State on publishing and reusing their high-value datasets;
5. Where available, the existence of data protection impact assessments carried out in accordance with Article 35 of Regulation (EU) 2016/679;
6. The number of public sector bodies exempted in accordance with Article 14(5) of Directive (EU) 2019/1024.

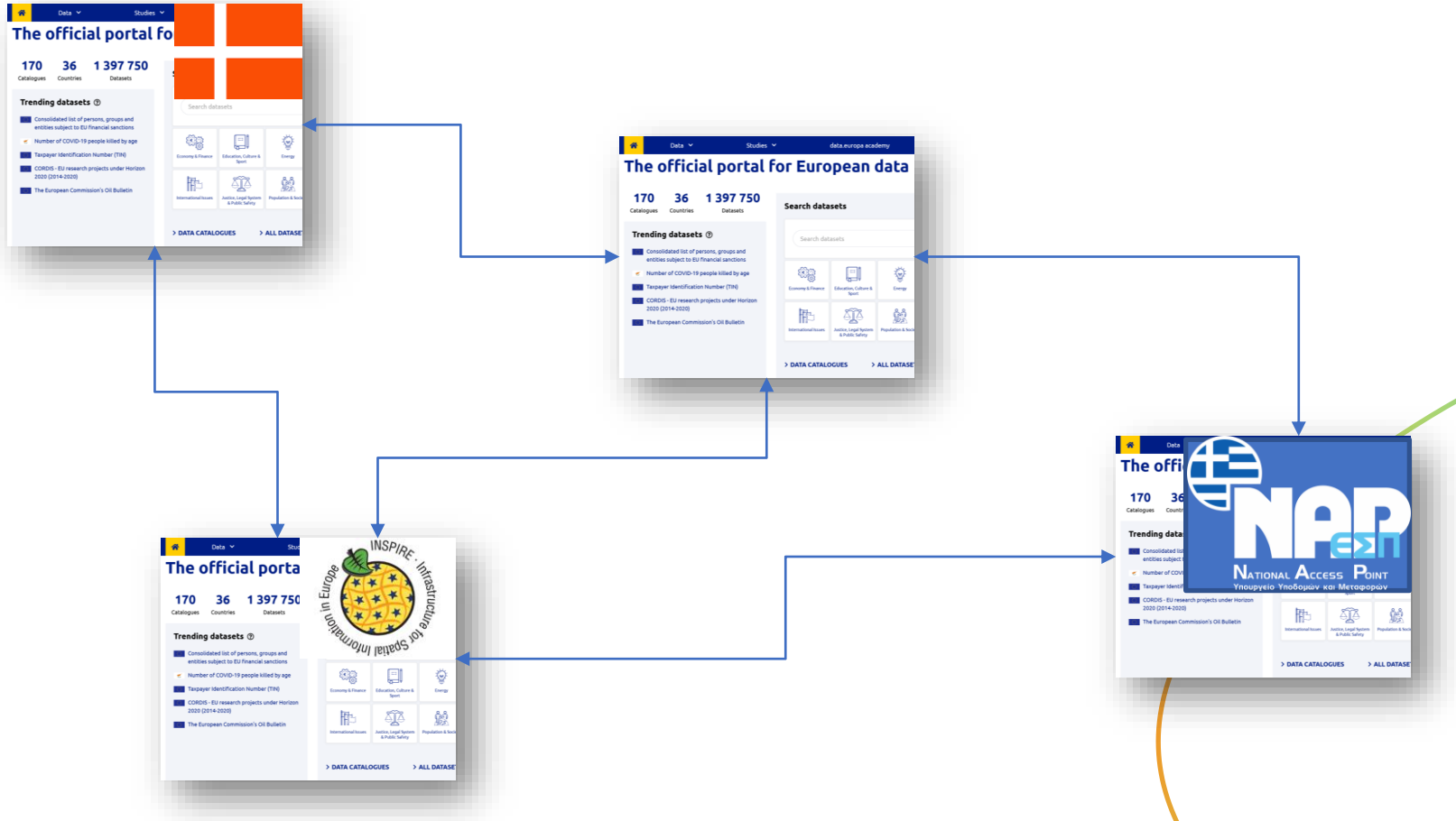
# Challenges

- How can MS adapt current metadata descriptions to satisfy the mentioned requirements?
- How can MS fulfil the reporting requirements?
- How can citizens/businesses see the HVD status (outer world perspective)?



# DCAT-AP introduction

# DCAT-AP Use Case



# DCAT-AP overview

https://github.com/SEMICeu/DCAT-AP/