

Meeting Minutes – Webinar 4 on the SDG Search Service Model 09/06/2021

Action 2016.29 Catalogue of Services Action

Project	Action 2016-29 Catalogue of services	Meeting Date/Time	09/06/2021 09:30-11:30 (GMT+1)
Meeting Type	Webinar	Meeting Location	Google meet
Meeting Coordinator	Miguel Alvarez Rodriguez Florian Barthélemy	Issue Date	18/06/2021

Meeting Agenda

1. Welcome
2. Discussing the model: tackling the GitHub issues
3. Pilot
4. Next steps

Webinar's participants

Attendee Name	Abbreviation	Organisation/Country
Kutt Kommel	KK	Estonia
Miguel Alvarez Rodriguez	MAR	European Commission
Valentini Paparrizou	VP	Greece
Davide Porrovecchio	DP	Italy
Vincenzo Travascio	VT	Italy
Lisdália Sanches	LS	Portugal
Florian Barthélemy	FB	PwC EU Services
Maxime Pittomvils	MP	PwC EU Services
Ana Rosa Guzmán	AG	Spain

Summary of the meeting

Topic	Summary
Welcome	<ul style="list-style-type: none"> Miguel Alvarez Rodriguez (MAR) started the webinar by welcoming all attendees to the fourth SDG Search Service model webinar organized by the Catalogue of Services Action of the ISA² Programme. MAR asked if there was an objection to record the webinar. There was no objection and the webinar has been recorded. MAR explained the agenda of the webinar.
Discussing the model: tackling the GitHub issues	<p>Updated model</p> <ul style="list-style-type: none"> Maxime Pittomvils presented the updated model based on the input received during the last webinar. See the figure below for the updated model. <pre> classDiagram class PublicOrganisation { + Identifier: Literal [1..*] + preferredLabel: Literal [1..*] } class Location { + geographicIdentifier: URI [1..*] + geographicName: Literal [1..*] } class PublicService { + Identifier: Literal [1..n] + name: Literal [1..n] + description: Literal [1..n] + policy code: Code [1..*] } class LinguisticSystem { + language: Code } class Webpage { + Identifier: URI [1..*] + type: Code [0..2] } class Output { + Identifier: Literal [1..*] + name: Literal [1..*] + description: Literal [1..*] + type: Code [0..*] } class Event { + type: Code [0..*] } PublicOrganisation "1" -- "1" PublicService : +hasCompetentAuthority PublicService "1" -- "1" Location : +spatial PublicService "1" -- "1" Webpage : +hasWebpage PublicService "0..*" -- "0..*" Output : +produces PublicService "0..*" -- "0..*" Event : +isGroupedBy LinguisticSystem "1" -- "1" Webpage : +hasLanguage </pre> <ul style="list-style-type: none"> Maxime Pittomvils gave an overview of the change log with regards to the previous webinar: <ul style="list-style-type: none"> Added policy code under Public Service (#38) Removed Contact Point (#41, #42) Agreed to use NUTS/LAU for Location (#40) Added geographic name to Location (#40) Changed cardinality hasLanguage of Webpage from 1..* to 1 (#45) Added Event as class and property isGroupedBy from Public Service to Event (#38) Cardinality of spatial of Public Service updated from 0..* to 1..* Added type to Webpage to cover different kind of pages.

{information/procedure}

- No additional Channels were added to the model ([#43](#))
- Updated type of Output to cover the outputs as mentioned in Annex II of the SDGR ([#37](#))

Open modelling discussions

- **Location ([#40](#))**

Context

Location is defined by NUTS/LAU, in alignment with the metadata for URLs for the repository of links. The name is in alignment with the codes from NUTS/LAU.

Proposition

We propose to add an additional property for capturing alternative labels of location to enrich the search.

Decision

An additional property 'name' will be foreseen within the model.

Discussion

/

- **Classification of Output ([#39](#))**

Context

The classification of outputs is according to the outputs defined in Annex II, which are not types but exact outputs depending on the procedure.

Proposition

We propose to create a classification which is more generic than the ones defined in Annex II in order to improve the search results:

- Analyse the existing types of output from Annex II
- Identify additional outputs from Annex I
- Compare with existing sources, such as the CPSV-AP
- Create a more generic list of outputs

Decision

We put the discussion on hold and will continue based on the results of the pilot. We keep the types of output currently in the model (outputs as defined in Annex II) as a placeholder. During the pilot, we will decide whether to use it or not.

Discussion

Ana Rosa Guzman: I agree that there should be more generic types for the different outputs defined in the SDGR. I would like to see the proposal of these more generic types.

Florian Barthélemy: That is the reason why the current classification in the model (the one based on annex II) is used as a starting point. These output types are already very specific and don't cover the outputs coming from annex I. We need to make sure we create a codelist of output types that cover both annex I and II. What we want to validate now is: Do we want to create more generic output types and which level of granularity do we want to achieve? Starting from the annexes of the SDGR, we will come up with a proposed list of output types.

Ana Rosa Guzman: The procedure number 16 of Annex II is quite extensive. For example, the output about managing licenses and permits for business activities is not only about confirmation of registration, but also about the cancellation, modification, etc. of the registration. In this view annex II is too generic.

Florian Barthélemy: Is it useful from a searchability point of view to have the types of output?

Valentini Paparrizou: We don't want to end up with a lot of optional fields, tackling the cases where only a few Member States fill them in. Will the user actually use such an extensive list of output types?

- **Policy code / types of Public Service / life and business events ([#37](#), [#38](#))**

Context

We are currently describing public services through the policy code and the types of events according to Annex II. Do we need to provide additional classification on top of the policy code to arrive at the right level of usability/granularity? Should each public service have a life or business event/output?

Proposition

We propose to analyse further the different life/business events captured in the annexes and other sources (CPSV-AP).

Decision

We will put the issue on hold and keep as classification the policy codes of Annex I and II for the public services within the model. We will discuss internally how to proceed and how to align between the classifications of the SDGR, the classifications defined in CPSV-AP (business/life events, types of public service) and those that are defined to navigate within the YourEurope portal.

Discussion

Valentini Paparrizou: We currently have the CPSV classification, the SDGR policy code and another classification from the YourEurope portal (the structure defining the navigation when entering the YourEurope portal: https://europa.eu/youreurope/citizens/index_en.htm). We are interested in a mapping between the policy codes of the SDGR and the categories in the YourEurope portal.

Florian Barthélemy: We should avoid creating multiple classifications which are targeting the same thing. The idea is to first use the classifications in the context of this pilot, but also align with the categories defined on the YourEurope portal.

Valentini Paparrizou: It is not clear where I can find the information of the annexes of the Regulation under the structure/navigation of the YourEurope portal, as there is no 1-1 mapping.

Miguel Alvarez Rodriguez: For this pilot, we can start of by focusing on the Regulation (Annex I and II). Regarding the classification within the YourEurope portal, we will discuss internally how to proceed and how to align between the classifications of the SDGR and the YourEurope portal.

- **Webpage ([#44](#))**

Context

The classification of webpages is currently defined whether it is information about a service or you can perform the procedure. Is there a need to have additional classification to enrich the information and improving the findability?

Proposition

We propose to keep the classification as it is defined for the repository of links.

Decision

It was agreed to keep the classification as it is defined today in the repository of links.

Discussion

/

- **Webpage ([#45](#))**

Context

Currently several webpages can be indicated for a specific public service. The question arises whether you should only indicate one webpage for each public service or multiple ones?

Proposition

The number of pages that you want to show can be chosen by the Member States, multiple services for one policy code and multiple pages for the same service.

Indicating more webpages for a service will provide multiple options, but the search engine could benefit from the additional information provided. (e.g. by also providing additional tags)

It depends on the amount of granularity that you want to achieve at European level.

Decision

We decided to keep this option open for the Member States to choose one or multiple webpages.

Discussion

Vincenzo Travascio: in relation to the pointing of the page, do

we also provide for the indication of a national hub to manage the mapping between the Annex II procedure and the specific national use case?

Florian Barthélemy: This is exactly the decision the Member States need to decide at national level.

- **Identifiers**

Context

Explicit identifiers have been included within the model for classes such as Public Organization, Public Service and Output. These allow for identification of the instances. (E.G in json or xml)

Within an RDF distribution, unique identification is facilitated by the value of the RDF node. In this case, the identifier should best be equivalent to the value of the RDF node.

For example:

```
<rdf:Description rdf:about="http://127.0.0.1:3362/TCC-001">
```

```
< dct:identifier>TCC-001</dct:identifier>
```

Proposition

We propose to have the identifiers for national identification. Do you agree?

Decision





It was agreed to keep the possibility for national identification.

Discussion

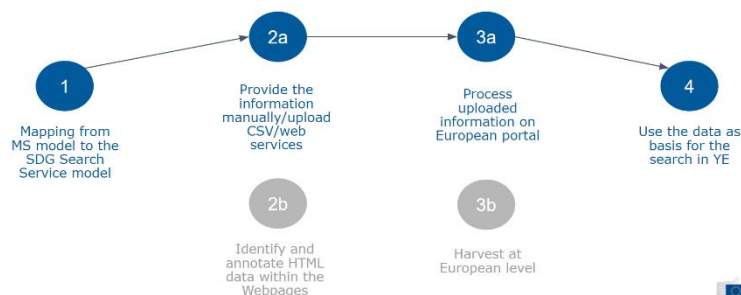
Vincenzo Travascio: I believe that identification at a national level can be useful for managing the variations in responsibility that may occur during operation, so I agree with the proposal.

Valentini Paparrizou: Should these identifiers be optional or mandatory?

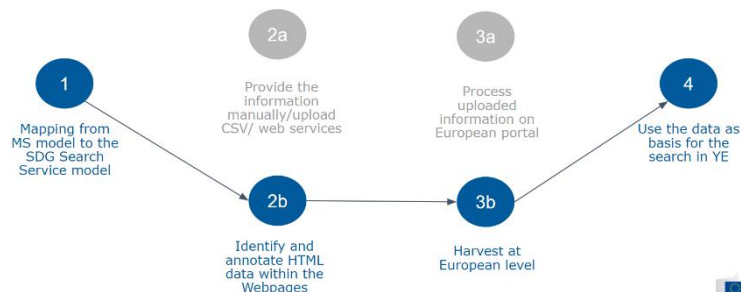
Florian Barthélemy: At the moment we have put them as mandatory.

	<p>Davide Porrovecchio: Webpages URL may vary, would be useful to have an identifier also for them?</p> <p>Florian Barthélemy: This is a valid point, we will update webpage to have both a unique identifier and the URL.</p>
Pilot	<p>PILOT EXPLANATION</p> <ul style="list-style-type: none"> Maxime Pittomvils explained the objectives of the pilot: <ol style="list-style-type: none"> Can you easily provide the additional data agreed? Are the search results improved at national and/or European level thanks to: <ol style="list-style-type: none"> Metadata tags? More public services described? Can the usability of the search be improved by: <ol style="list-style-type: none"> Showing additional data? Proposing additional filters? Can the provision of the data be (partially) automated? Maxime Pittomvils gave examples of structured data and filters that could be used within search engines. Maxime introduced the proposed approach for the pilot, which consists of 4 major steps: <div style="display: flex; justify-content: space-around; align-items: flex-end; margin-top: 20px;"> <div style="text-align: center;">  <p>Mapping from MS model to the SDG Search Service model</p> </div> <div style="text-align: center;">  <p>Prepare information of MS in collectable format</p> </div> <div style="text-align: center;">  <p>Collect the information at European level</p> </div> <div style="text-align: center;">  <p>Use the data as basis for the search in YE</p> </div> </div> Maxime Pittomvils introduced the two proposed approaches for conducting the pilot: <ul style="list-style-type: none"> Option 1: without annotation of webpages: This includes the option where people can upload their metadata to the links repository manually, via excel or through web services (API). Option 2: with annotation of webpages: Through this option, the Member States tag the required data within the webpages, and the crawler of the repository of links finds the tags on the websites indicated by the Member States. <p>Both options can be found in the figures below:</p>

Option 1: without annotation of webpages



Option 2: with annotation of webpages



INPUT CAPTURING

Further input and feedback on the pilot were captured through a collaborative tool, in the format of 3 exercises:

Exercise 1: Vote for the options that have priority within the pilot:

- Providing the additional data easily – **5 votes**
- Proposing additional filters to improve the usability of the search – **5 votes**
- Adding metadata tags to improve search results at national/European level – **3 votes**
- Increasing the number of public services described to improve search at national/European level – **1 vote**
- Increasing the automation of the provision of data – **1 vote**

- Showing additional data in the results (e.g. on Your Europe) to improve the usability of the search – **0 votes**

Exercise 2: How do you provide your data? Would you be interested in testing another method?

In the picture below is an overview of the way the Member States are providing their data to the repository of links: excel, crawler, manual and web services - API. (green dots)

Four additional countries indicated they would be interested in testing out the web services (API) option. (orange dots)

Countries	How do you provide data? (correct the green dots)				Would you be interested in testing another method? (Indicate in red)
	excel	crawler	manual	API	
Estonia	●			●	
Germany			●		
Netherlands		●	●		
Spain	●	●	●	●	
Sweden		●			
Italy			●	●	
Portugal			●		
Bulgaria					
Finland	●	●	●		
Greece		●	●	●	
Poland			●		
Slovenia					

Exercise 3: Is your data structured? Is your data according to a model (yes/no)? In which format is your data available?

- **Estonia:** Structured data in CSV format and in the future also other formats.
- **Spain:** There is a national data model that can be adapted. Legacy systems and some little competent authorities do not use data models. Each competent authority has a different situation (DBRel, for Catalog of Services, CMS, manual editing, ad-hoc solutions combination of some of the previous).

Formats used: DBRel, HTML, XML, Web Services.

- **Italy:** No structured data is used, it is included within the SDG implementation roadmap.
- **Finland:** No structured data, there is no common data model used by all competent authorities.
- **Greece:** Structured data, but only procedures data implements a data model. CPSV-AP has been used to develop the national register of procedures. Additional requirements placed from national actors led to the creation of an extension to CORE CPSV-AP (and thus additional properties). Formats used: xls, csv, json.

Below a figure with the captured input.

Exercise 3

Countries	Is your data structured? Is your data according to a model? (yes/no)	Comments	In which format is your data available?
Estonia	Yes		Yes, standard format, probably will be available in the future
Germany			
Netherlands			
Spain			DBRel, HTML, XML, Web services
Sweden			
Italy	No	According to the SDG implementation roadmap	
Portugal			
Bulgaria			
Finland	No	There is no common data model used by all competent authorities	
Greece	Yes, it is structured, but only procedures data implements a data model		xls, csv, json (development effort required)
Poland			
Slovenia			

Discussion points:

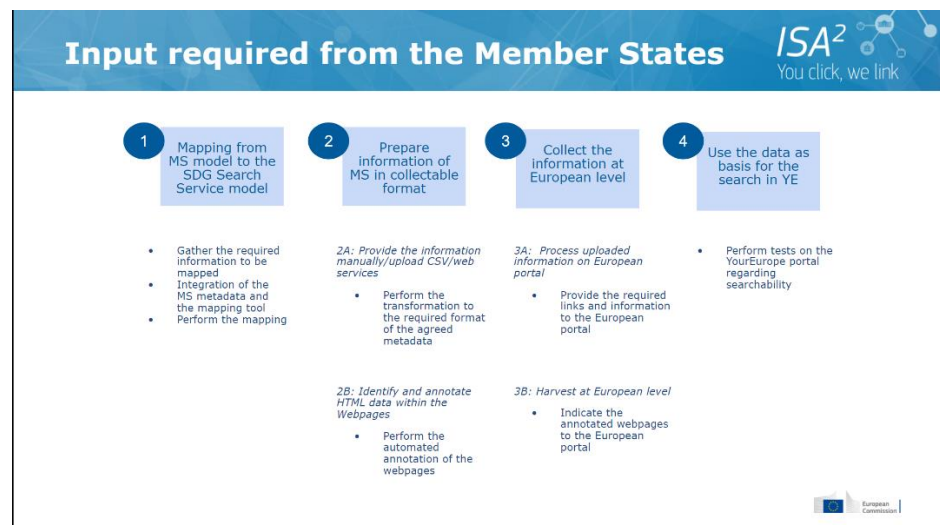
- Valentini Paparrizou: How are we going to measure the improvement in findability? Are we going to have test users? How do you measure your performance? This needs to be tested by people who are not involved within the SDGR work, because they are then not biased. (for example, because of the

knowledge of the specific terminology)

- Florian Barthélemy responded that this is indeed valid. We first need to check with DG GROW whether they log all the key words used by the end user. We need to check whether these key words match the metadata that we collect. Secondly, we need test persons without background knowledge of the SDG.
- Miguel Alvarez Rodriguez: Has Greece done testing along these lines?
- Valentini Paparrizou: No testing has been done so far. We so far have only received feedback. I believe we need something more specific for the testing.
- Florian Barthélemy added that Member States can help in checking at national level whether they have conducted already an analysis of the usability. We need to do the same at European level with the YourEurope team. The basis will be to look at the keywords used by the end user.

OVERVIEW REQUIRED INPUT PILOT

- Maxime Pittomvils gave an overview of input required from the Member States for the different steps of the pilot, see the figure below.



Next steps

- Florian Barthélemy informed the attendees on the next steps:
 - Share the scoping document for the pilot with the Member States with further details. Afterwards, the Member States can let the Catalogue of Services Action know whether they are interested in conducting this pilot.
 - Share the stable version of the SDG Search Service model via GitHub.
 - The meeting minutes and the recording will be shared with the attendees after the webinar.
 - Feedback can be provided through email or via GitHub:

<https://github.com/catalogue-of-services-isa/SDG-search-service-model>

- Miguel Alvarez Rodriguez closed the webinar by thanking all participants for collaborating.