

Evaluation of the Webinar Paving the way for new markets *Challenges and recommendations*

Synopsis

UN-GGIM: Europe | Work Group on Data Integration | Subgroup II

Version 1.0

2022-07-05



DOCUMENT HISTORY

Version	Date	Comment	Modified Pages	Responsibility
0.9	18-03-2022	Draft report	all	Sabine Afflerbach-Thom, Pier-Giorgio Zaccheddu (BKG), Nathalie Delattre (NGI-B)
1.0	29-06-2022	Final for approval	all	Nathalie Delattre (NGI-B), Sabine Afflerbach-Thom (BKG)
	01/07/2022	Final for approval		Comments from H Reuter integrated



CONTENT

DOCUMENT HISTORY	2
CONTENT	3
FIGURES.....	4
PURPOSE AND BACKGROUND OF THE DOCUMENT.....	5
INTRODUCING THE WEBINAR	6
MAIN STATEMENTS OF THE QUESTIONNAIRE.....	7
MAIN STATEMENTS OF THE GISCO BREAK-OUT SESSION	9
CONCLUSION.....	10
ANNEX	11
1. INTRODUCING THE WEBINAR	11
2. QUESTIONS CONCERNING THE WHY.....	13
3. QUESTIONS CONCERNING THE WHAT	17
4. QUESTIONS CONCERNING THE HOW	23



FIGURES

Figure 1: List of questions asked the participants within the webinar	6
Figure 2: Participants registered for the webinar	11
Figure 3: Distribution of the registered participants.....	11
Figure 4: Participants taking part in the poll	12
Figure 5: Distribution within the poll	12
Figure 6: Result of the Question on collaboration needs.....	13
Figure 7: Result of the Question on collaboration structure	14
Figure 8: Result of the Question on extending partnerships	15
Figure 9: Result of the Question on ontologies.....	16
Figure 10: Result of the Question on GKI	17
Figure 11: Result of the Question on PIDs.....	18
Figure 12: Result of the Question on Linked Data	19
Figure 13: Result of the Question on common definitions	20
Figure 14: Result of the Question on coordination of common definitions	21
Figure 15: Result of the Question on data integration rules.....	22
Figure 16: Result of the Question on knowledge about APIs.....	23
Figure 17: Result of the Question on how to improve the cooperation between geospatial and statistical experts	24
Figure 18: Result of the Question on how NMCAs and NSI can support European data spaces.....	25
Figure 19: Result of the Question on the moderization of NSDIs	26
Figure 20: Result of the Question on Knowledge graphs.....	27
Figure 21: Result of the Question on missing data integration methods within the document	28
Figure 22: Result of the Question on the knowledge and capacity for data integration methods	29



PURPOSE AND BACKGROUND OF THE DOCUMENT

The document provides a brief overview on the outcome and findings of the virtual webinar: ‘Paving the way for new markets – Challenges and recommendations’ held on 24 February 2022 and the discussion within the breakout sessions of the meeting of the Working Group on Integration of Statistical and Geospatial Information (GISCO) of Eurostat on 23 March 2022.

Data integration methods are already existing, applied and well documented. Moreover, the technology is rapidly evolving. Within the workplan 2019-2022 of UN-GGIM: Europe Working Group on Data Integration, the objective of the subgroup II on “Data Integration Methods” has therefore been to summarise the situation and provide a European perspective with recommendations for possible use, completed with an impact assessment at social, economic and organisational level.

As a result, the document [“Data Integration Methods – Analysis of future trends in geospatial data capture, creation, maintenance and management”](#) highlighting recommendations for increased application of good practices was published in early 2022 .

To promote the recommendations and to start a broader discussion within the NMCA’s and NSI’s community, a webinar (on 24-02-2022) was organised to guide through this document focusing on three aspects:

- (1) the importance of data integration
- (2) main tools, methods and requirements as well as
- (3) recommendations to enable data integration processes.

Interaction and knowledge exchange were important in this webinar and may guide further activities of the Working Group. A poll was prepared to ask the audience general, organisational and technical questions to data integration and its methods.

This webinar was open to NMCAs and NSIs, but also some people from international organisations took part.

Afterwards, the break-out session at the GISCO meeting (on 23-03-2022) has emphasized the improvement of collaboration and coordination between experts and partners; participants have been invited to reflect and give opinions based on a list of questions.



INTRODUCING THE WEBINAR

Around 100 persons have registered for the webinar, coming mainly from Europe. But some of them were also from USA and Asia.

In contrast to the registered 100 persons, only 36 took part in the poll answering the ice breaker questions. The questions to the Data integration methods were answered by 20-25 participants only and a majority was from NMCAs (see detailed figures in Annex).

The webinar was separated into three parts, dealing with the **Why**, **What** and **How** of “Data Integration Methods”.

In total, 17 questions have been prepared for all three parts.

Why	1	Do you think that strengthening collaboration on new data integration methods and processes is needed for your agency/company with other domains?
	2	Do you have a collaboration structure / coordination board that is helping you to evaluate and assess requirements on new data integration methods and processes?
	3	Do you consider extending your partnership with other domains (other than geospatial or statistical) in matter of data integration in order to gain new market?
	4	Have you/your institution already invested resources in defining common ontologies for a better linkage?
	5	Have you already invested resources for GKI or are you willing to do so in the future?
What	1	Do you already have PIDs integrated in your data and provide maintenance rules in agreement with your partners?
	2	Do you agree that Linked Data should be “seen as the key enabler for data integration”?
	3	Do you have common definitions of fundamental geographies elaborated and disseminated in the geospatial community?
	4	Do you think that the common definition for linked data should be coordinated on a European Level?
	5	Has your organization formulated data integration rules together with other organizations in order to facilitate the data integration processes?
	6	How can NMCA and NSI gain knowledge about APIs for smart geospatial data provision?
How	1	How can geospatial and statistical experts work better together in order to enable trusted and successful data integration processes?
	2	What can NMCA and NSI do to better support European data spaces?
	3	What do you think needs to be modernized most within NSDI?
	4	Do you have any experiences in modelling/describing knowledge graphs?
	5	Have you identified data integration methods suitable for your data integration processes that have not been tackled within the document?
	6	Does the staff within your organisation have the required knowledge and capacity for data integration methods?

FIGURE 1: LIST OF QUESTIONS ASKED THE PARTICIPANTS WITHIN THE WEBINAR

In the annex at the end of this document, the questions and answers are described in more detail.



MAIN STATEMENTS OF THE QUESTIONNAIRE

Why: a data integration methodology is a need

The majority agree on the usefulness to strengthen collaboration with other domains and investing on new data integration methods. However, most of them have not yet established any collaboration by mean of any structure or coordination board in the aim to evaluate and assess their requirements in that matter.

In matter of interest for extending the data integration with other domains (other than geospatial or statistical), more than half have already considered this by providing suitable data or services for data integration application. For the others, the question may come too early or is out of scope.

Data integration shall also be understood as an essential part of a geospatial knowledge infrastructure aiming at common data spaces. Common ontologies for better linkage of datasets across domains is a key request; but almost 60% of the voting persons have not already considered this aspect.

Evolving towards a Geospatial Knowledge Infrastructure, which establishes a network of spatial data infrastructures across domains, is the suitable framework allowing the efficient usage of our spatial data infrastructures to meet the objectives of the European Green deal; but a majority does not consider this possibility or does not have the willingness to invest resources in that matter because being currently out of scope.

What: data integration procedures and methods have already been implemented

The document [“Data Integration Methods – Analysis of future trends in geospatial data capture, creation, maintenance and management”](#) concludes that “linked data” methods should be seen as the key enabler for data integration. Half of the voting persons agree with the statement.

For successfully linking geospatial data and statistical data, the document recommends to define and implement valid Persistent Identifiers (PID) across domains. 1/3 of the voters have elaborated consistent rules with their partners for maintaining the PIDs, and 1/3 of them have partly implemented it.

Similarly, the document recommends to agree on common definition of fundamental geographies (definitions of terms and relations) disseminated in the geospatial community. Again, 1/3 of the voters have common registries and a majority is on the way to get some. In that perspective, the voters approve that common definitions of fundamental geographies should also be coordinated at European level.

Similarly, only 20% of the voters have already formulated data integration rules with other organisations.

- The document recommends to develop and implement standardized Open Application Programming Interfaces (Open APIs) for a smart geospatial data provision comprising service interface and API definitions, styles and processes including geospatial rights management. In that perspective, European organisations could strongly contribute to the acquisition of



knowledge for NMCAs and NSIs by determining some of their requirements in that matter, initiating projects or participating in standardisation bodies (ISO, OGC).

How: to progress

The participants agreed that the coordination between partners and coordination bodies will enable trustful and successful data integration processes.

Spatial is not special anymore; the spatial data infrastructures and the geospatial data have to adapt to an integration into the European Data Spaces (EDS). In order to better support the EDS, NMCAs and NSIs would like to get some leadership from the European institutions in order to determine requirements and to initiate projects on the definition of data spaces.

NSDIs and their implemented technical standards have to evolve and be adapted to new technologies and developments.

Organisations have to establish Knowledge graphs¹ to structure and document the way to transit and relate between the different approaches and data sources – for example geography, statistic, ontology. Unfortunately, most (70%) do not have any experiences in that respect yet.

Therefore, investing resources and capacity building are a must as a majority of the participants recognised a lack in that matter.

To conclude, a majority (3/4) recognises that the document has tackled the most valuable and used data integration methods, a list which – of course – is not exhaustive.

¹ The knowledge graph represents a collection of interlinked descriptions of entities – objects, events or concepts. Knowledge graphs put data in context via linking and semantic metadata and this way provide a framework for data integration, unification, analytics and sharing.
<https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph/>



MAIN STATEMENTS OF THE GISCO BREAK-OUT SESSION

The delegates at the GISCO break-out session on 23-03-2022 were invited to contribute to the discussion on the question how to contribute to the European Data Spaces (EDS). The discussion was framed along a number of guiding questions:

- ✓ What can we as NMCA and NSI do to better support European data spaces?
- ✓ How does your organisation face these new challenges?
- ✓ Is your staff directly involved in European working groups?
- ✓ Is there a national cross-domain interest community to identify and discuss challenges/requirements?
- ✓ Have you identified the need for working with an appropriate governance structure (coordination board)?
- ✓ Have you installed or prepared all needed roles in a common data space (e.g. identity management, broker)?

It became clear from the discussion that the EDS are still seen as an abstract/hypothetical concept at the European level, which is expected to deliver a) concrete results with b) to use already existing common infrastructure (e.g. open data, INSPIRE). As similar, thematic data spaces are also sometimes developed in the Member States, a business separation and also coordination needs to be ensured between National and European developments. In general, it was acknowledged that decoupling dataset/data-services towards a data theme to answer specific answers is a favourable way forward. Some members issued the request for further information sessions in the future.

On the collaboration areas, the participants reported the need to address methodological aspects for products based on both geographical and statistical data. Confidentiality for spatially disaggregated data is also an area to be addressed by NMCAs and NSIs together. The need for stronger cooperation on building registers was also mentioned, as well as INSPIRE implementation.

The need for a formalisation of the relationship between partners was also mentioned. This formalisation could be defined by law, with a dedicated legislation, or with a memorandum of understanding. Additionally, financial aspects and resource requirements in general should be addressed. The cooperation could also take place under the form of regular meetings, seminar, exchange forums, etc. involving both GIS and statistical communities.

It is important to clarify and consolidate the organisational/institutional structure around geographical levels. This starts from the regional level (when necessary, for countries like Germany and Spain for example) over coordination inside the country due to various organisations structures (e.g. NMCAs are split into several organisations, such as cadastral and mapping agencies) up to the European level. The need for a European organisation for geospatial data, comparable to Eurostat for statistics, was repeated several times from various participants.



CONCLUSION

In order to comply to the European political program and the European Green deal – which establishes a green transformation in the light of the UN Agenda 2030 for Sustainable Development Goals (SDG) –, NMCAs and NSIs have to break the silos of their SDIs in order to evolve towards a Geospatial Knowledge Infrastructure, which establishes a network of Spatial Data infrastructures across domains.

Investing in ‘Linked Data’ and common ontologies are the key factors setting up the first bridges between the geospatial infrastructures.

Clearly, this kind of investment is at the very beginning mainly for the NMCAs. Several reasons might occur: the quality of the cooperation between NMCAs and NSIs, the lack of knowledge, skills and expertises in matter of data integration methods and processes, and the strategy of the respective institutions considering the evolution towards a Geospatial Knowledge Infrastructure.

The way to move forward is paved of good willingness and a lot of initiatives in matter of data integration taken at national or sub-national levels have been recognised and promoted as best practices. However, this kind of national initiatives when kept isolated, will miss the goal of a political program at European level. European institutions must take a substantial and leading role by providing priorities and recommendations in matter of data integration, defining the European data spaces and initiating European projects on the provision of SDG indicators. Coordination needs to be ensured between National and European developments, for example on building registers.

The quality of the coordination and cooperation between NMCAs and NSIs is the main target. Only if different organisations from local to European level share responsibilities and work together under clarified, consolidated and formalised objectives, success will be inevitable.

In the technical domain, adaptation to new technologies and developments is key. In that perspective, the technical standards have to continuously be adapted to new technologies and developments. Investing resources and capacity building in data integration methods and processes is of main importance as well as exchanging knowledge, expertises and best practices under the form of regular meetings, seminar, exchange forums, etc. involving both mapping and statistical communities.

Altogether, we are stronger!



ANNEX

1. Introducing the webinar

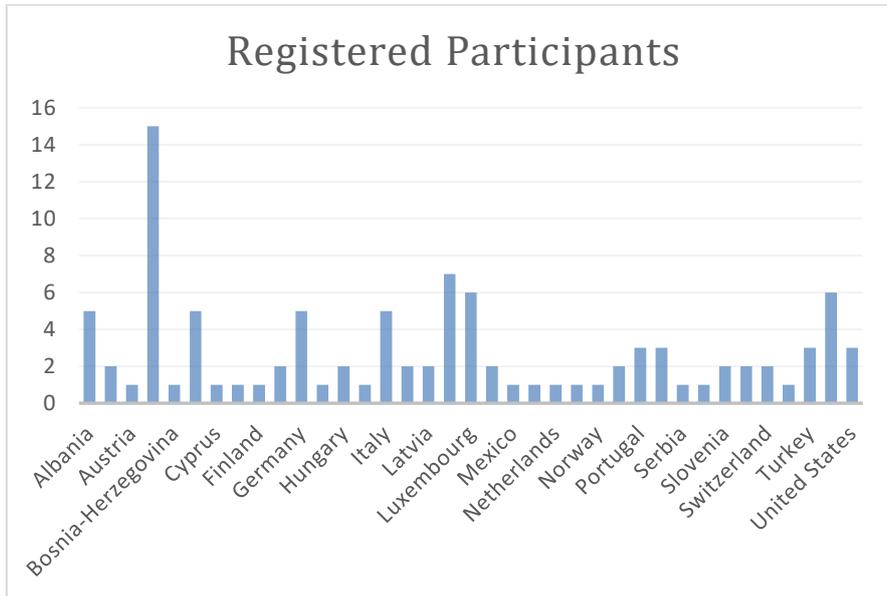


FIGURE 2: PARTICIPANTS REGISTERED FOR THE WEBINAR

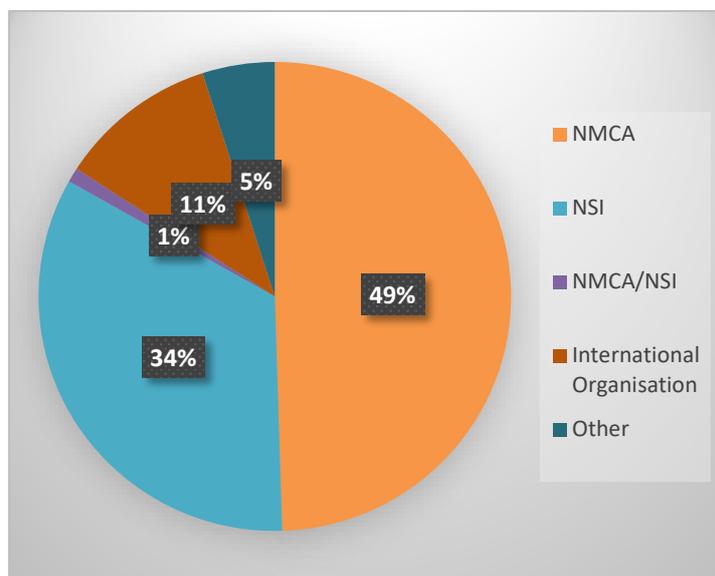


FIGURE 3: DISTRIBUTION OF THE REGISTERED PARTICIPANTS



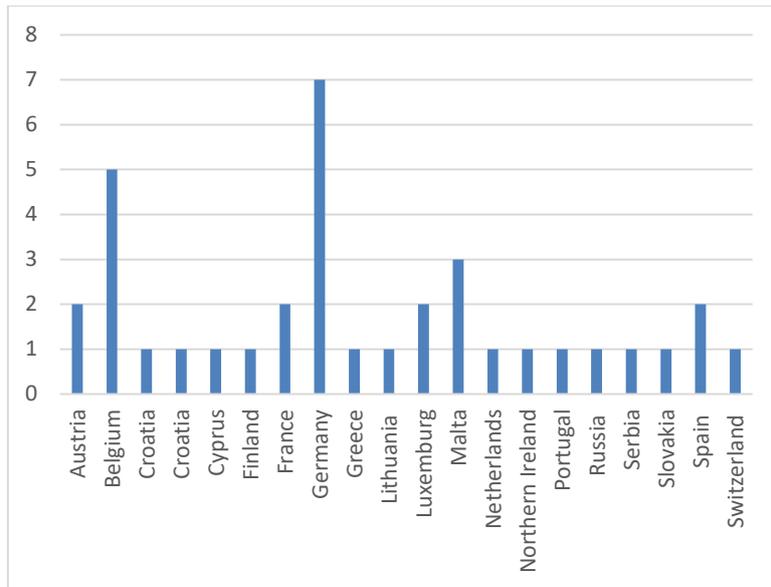


FIGURE 4: PARTICIPANTS TAKING PART IN THE POLL

As you can see from Figure 5 the majority of people taking part in the poll and trying to answer the questions are from NMCA's.

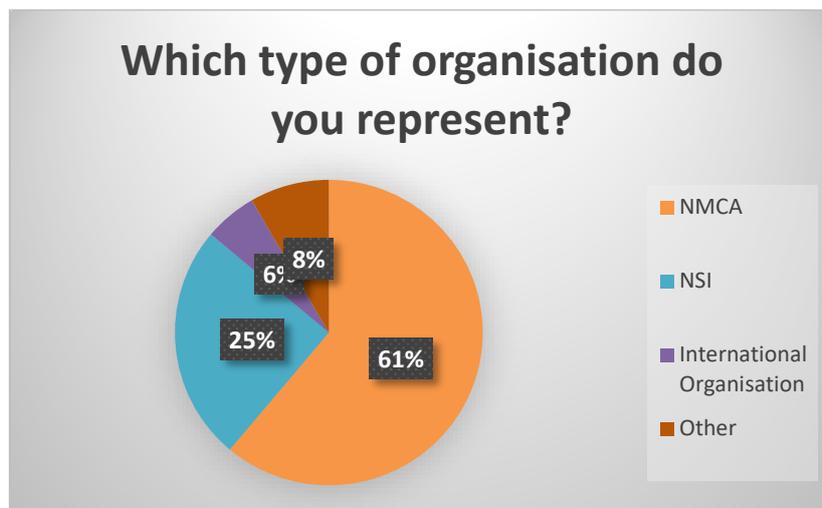


FIGURE 5: DISTRIBUTION WITHIN THE POLL



2. Questions concerning the WHY

Do you think that strengthening collaboration on new data integration methods and processes is needed for your agency/company with other domains?

Predetermined answers:

- Yes
- No

Result:

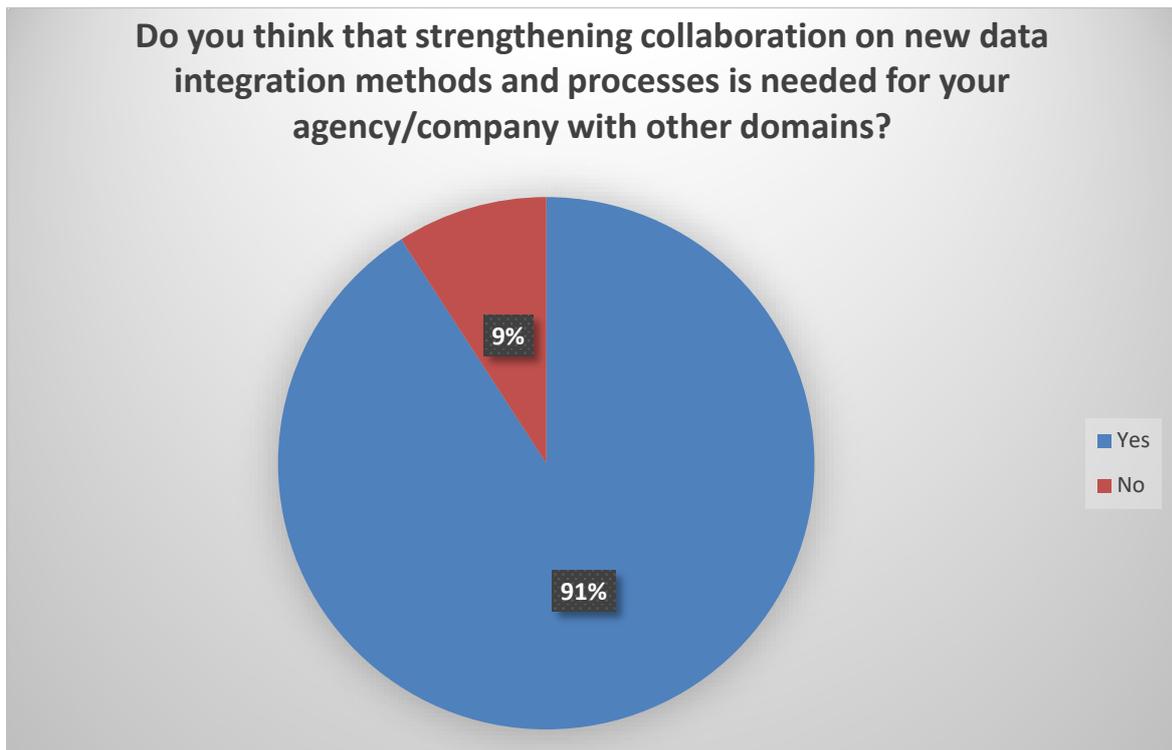


FIGURE 6: RESULT OF THE QUESTION ON COLLABORATION NEEDS

91% of the participants answered **Yes**.

Do you have a collaboration structure / coordination board that is helping you to evaluate and assess requirements on new data integration methods and processes?

Predetermined answers:

- Yes
- No

Result:

Do you have a collaboration structure / coordination board that is helping you to evaluate and assess requirements on new data integration methods and processes?

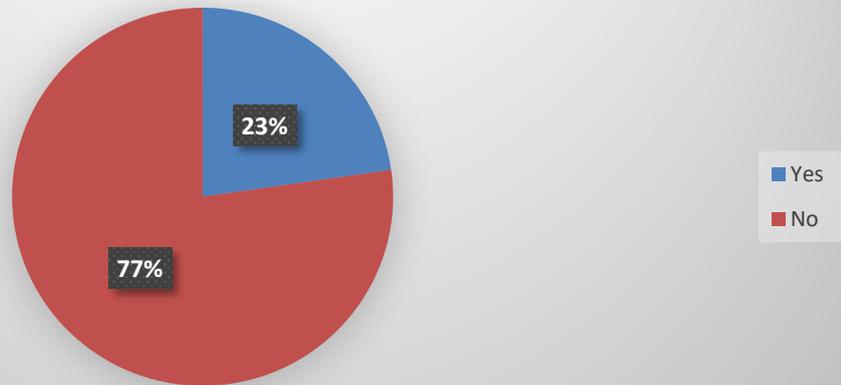


FIGURE 7: RESULT OF THE QUESTION ON COLLABORATION STRUCTURE

More than 2/3 of the participants answered this question with **No**. The follow-up question to be clarified if it is useful to have/install one collaboration structure, was unfortunately not asked within the webinar, but was discussed within the break-out session of the GISCO meeting on the 23 March 2022.

Do you consider extending your partnership with other domains (other than geospatial or statistical) in matter of data integration in order to gain new market?

Predetermined answers:

- Yes, providing suitable data
- Yes, providing services for data integration application
- It's not an issue yet
- No

Result:



Do you consider extending your partnership with other domains (other than geospatial or statistical) in matter of data integration in order to gain new market?

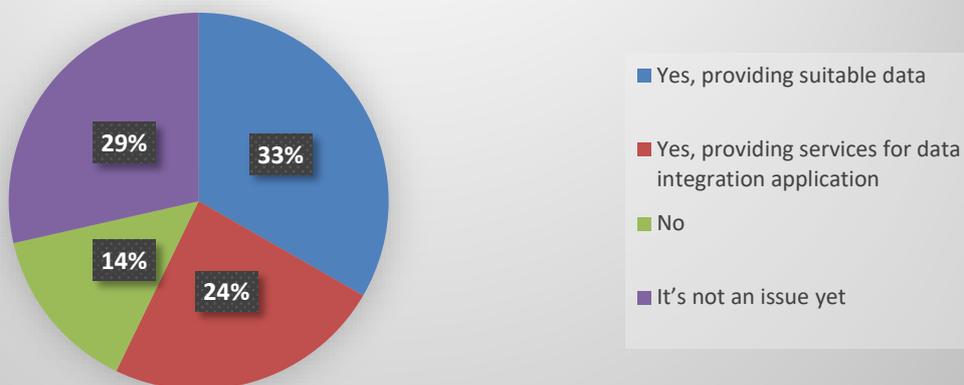


FIGURE 8: RESULT OF THE QUESTION ON EXTENDING PARTNERSHIPS

More than half of the people answered that they have already considered in extending the partnership with other domains and have a concrete idea how to do so. For the rest of the participants the question may come too early and the need for this topic is not there yet.

Have you/your institution already invested resources in defining common ontologies for a better linkage?

Predetermined answers:

- Yes
- No
- It's not an issue yet

Result:



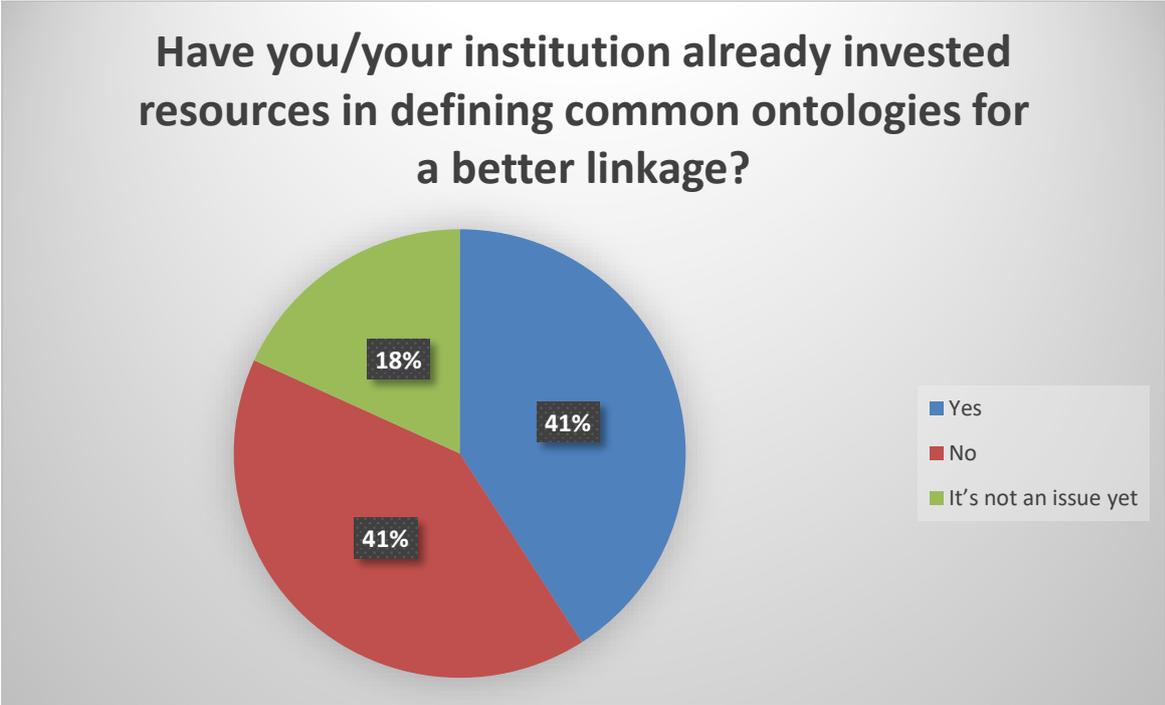


FIGURE 9: RESULT OF THE QUESTION ON ONTOLOGIES

Almost **60%** of the participants have answered that they haven't yet invested any resources in defining common ontologies. Within our elaborate investigation with data integration methods we as Working Group have recognized that this issue is of great importance and needs to be further developed. So, do we have to increase the awareness for NMCAs and NSI that it is necessary to work on this topic? How can this be done?

Have you already invested resources for GKI or are you willing to do so in the future?

Predetermined answers:

- Yes
- No
- It's not an issue yet

Result:

Have you already invested resources for GKI or are you willing to do so in the future?

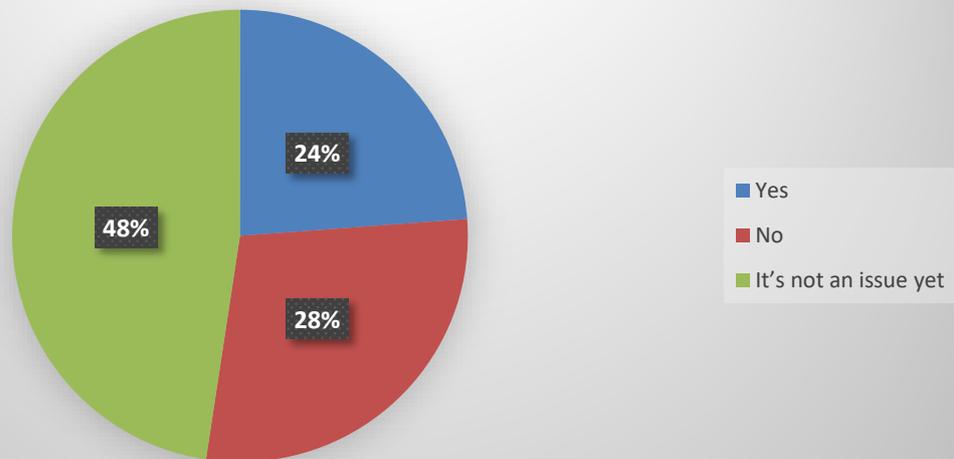


FIGURE 10: RESULT OF THE QUESTION ON GKI

Only **24 %** of the participants answered with **Yes**, the rest indicated **No** or **It's not an issue yet**.

It seems to be as if there are similarities concerning to question no. 4. So the question might be again: How to deal with this topic in future?

3. Questions concerning the What

Do you have already PIDs integrated in your data and provided maintenance rules in consistency with your partners?

Predetermined answers:

- Yes
- No
- partly
- It's not an issue yet

Result:



Do you already have PIDs integrated in your data and provide maintenance rules in agreement with your partners?

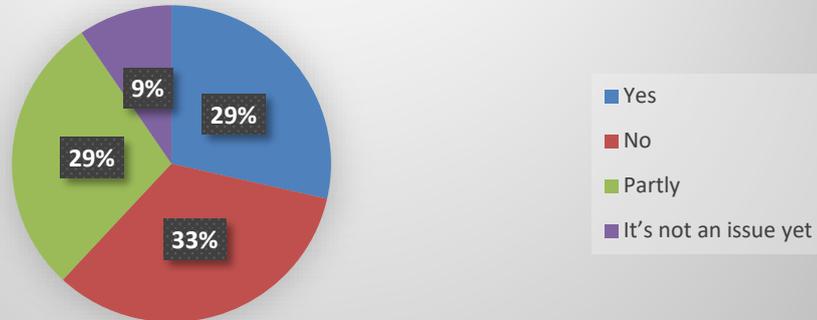


FIGURE 11: RESULT OF THE QUESTION ON PIDs

Only **29%** of the participants have answered **Yes**, but the rest indicated **No**, **Partly** or **It's not an issue yet**. Do we need more awareness in the importance of PIDs for our data? And if so, how can/should we do that?

Do you agree that Linked Data should be “seen as the key enabler for data integration”?

Predetermined answers:

- Yes
- No
- Partly
- It's not an issue yet

Result:



Do you agree that Linked Data should be “seen as the key enabler for data integration”?

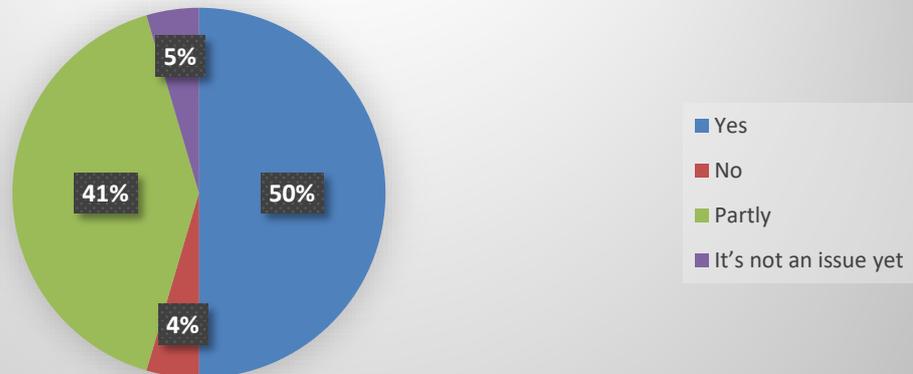


FIGURE 12: RESULT OF THE QUESTION ON LINKED DATA

50% of the participants voted **Yes**, **41%** voted **Partly**, the rest was for **No** or **It's not an issue yet**. What does this result mean for NMCAs and NSI for the future? Shall we elaborate this topic and try to invest more resources in this technology?

Do you have common definitions of fundamental geographies elaborated and disseminated in the geospatial community?

Predetermined answers:

- Yes
- No
- Partly
- It's not an issue yet

Result:



Do you have common definitions of fundamental geographies elaborated and disseminated in the geospatial community?



FIGURE 13: RESULT OF THE QUESTION ON COMMON DEFINITIONS

Only **36%** of the participants answered **Yes**, **52%** indicated **Partly**. It should be worth to invest more in the usefulness of common definitions to change **52% Partly** into **Yes**.

Do you think that the common definition for linked data should be coordinated on a European Level?

Predetermined answers:

- Yes
- No
- It's not an issue yet

Result:



Do you think that the common definition for linked data should be coordinated on a European Level?



FIGURE 14: RESULT OF THE QUESTION ON COORDINATION OF COMMON DEFINITIONS

88% of the participants answered **Yes**, only **12%** voted for **No** or **It's not an issue yet**. This is a clear statement and should be kept in mind.

Has your organization formulated data integration rules together with other organizations in order to facilitate the data integration processes?

Predetermined answers:

- Yes
- No
- Partly
- It's not an issue yet

Result:



Has your organization formulated data integration rules together with other organizations in order to facilitate the data integration processes?



FIGURE 15: RESULT OF THE QUESTION ON DATA INTEGRATION RULES

Only **20%** voted for **Yes**, **52%** indicated **Partly**. For the successful development of data integration processes, the importance of defining data integration rules between different organisations must be increased.

How can NMCA and NSI gain knowledge about APIs for smart geospatial data provision?

Predetermined answers, as multiple choice:

- Request European organisations to determine the requirements for NMCAs and NSIs
- Request European organizations to initiate projects on the definition of APIs
- Active participation in standardisation bodies (ISO, OGC) has to be done.

Result:



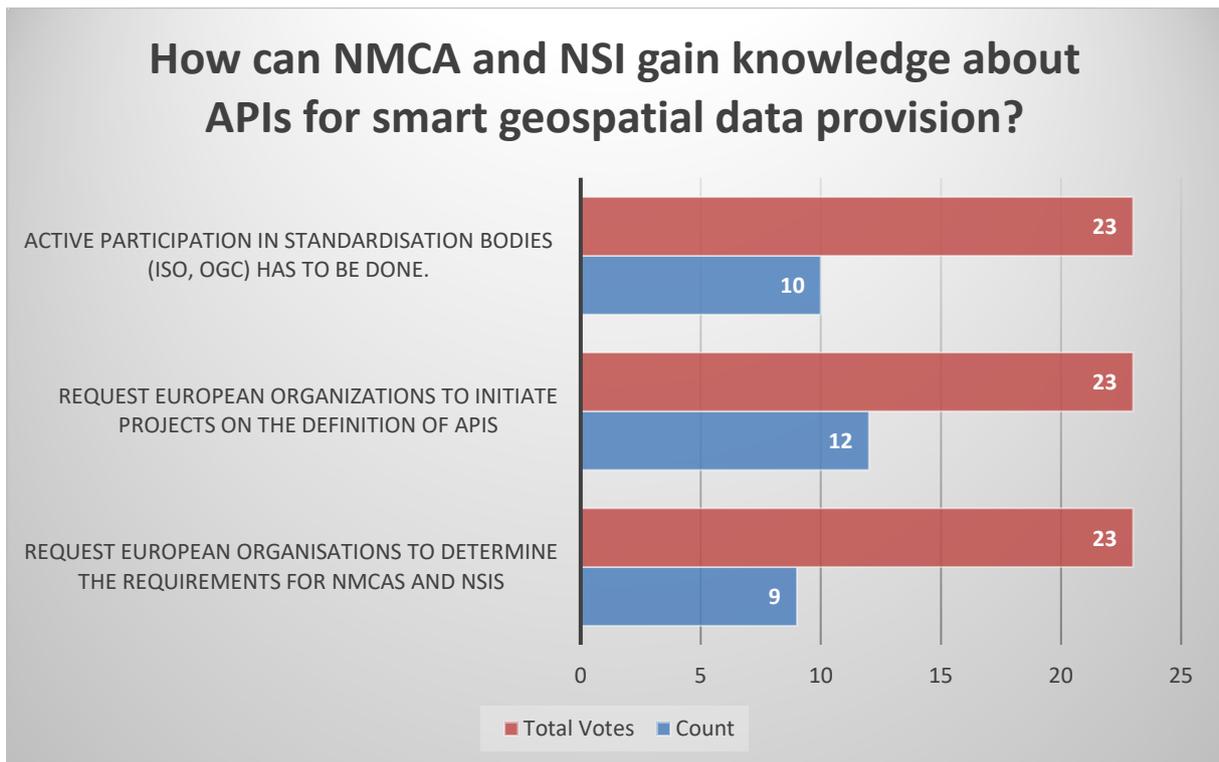


FIGURE 16: RESULT OF THE QUESTION ON KNOWLEDGE ABOUT APIS

The feedback to the 3 options of the question is quite similar, a clear trend to one option is not visible. The reason might be that NMCAs and NSI have not intensively elaborated this topic yet.

4. Questions concerning the HOW

How can geospatial and statistical experts work better together in order to enable trustful and successful data integration processes?

Predetermined answers, as multiple choice:

- Only one partner has the lead
- There is an integrated coordination between the partners
- Coordination bodies have to be created
- Other possibilities than the existing are needed.

Result:

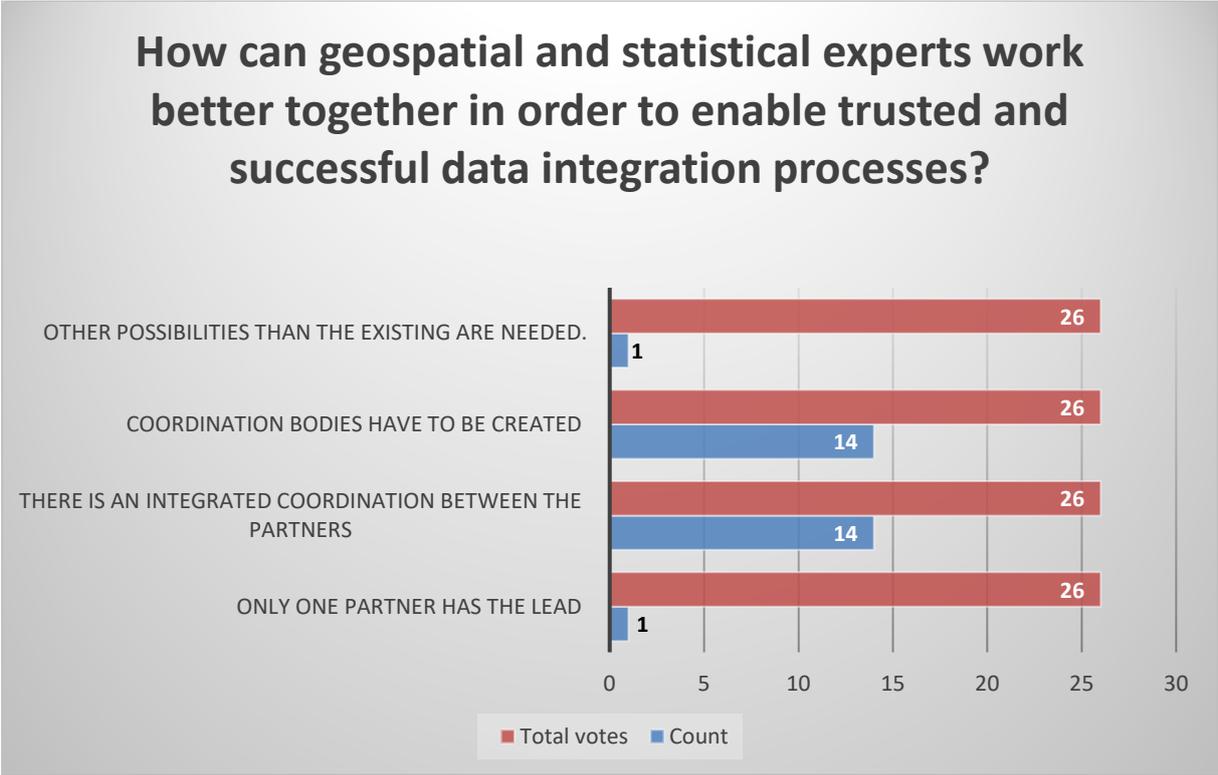


FIGURE 17: RESULT OF THE QUESTION ON HOW TO IMPROVE THE COOPERATION BETWEEN GEOSPATIAL AND STATISTICAL EXPERTS

2 options (“There is an integrated coordination between the partners” and “Coordination bodies have to be created”) are of great importance, the 2 other options (“Only one partner has the lead” and “Other possibilities than the existing are needed”) can be neglected. A trend is visible and needs to be kept in mind.

What can NMCA and NSI do to better support European data spaces?

Predetermined answers, as multiple choice:

- Request European organisations to determine the requirements for NMCAs and NSIs
- Request European organizations to initiate projects on the definition of data spaces

Result:

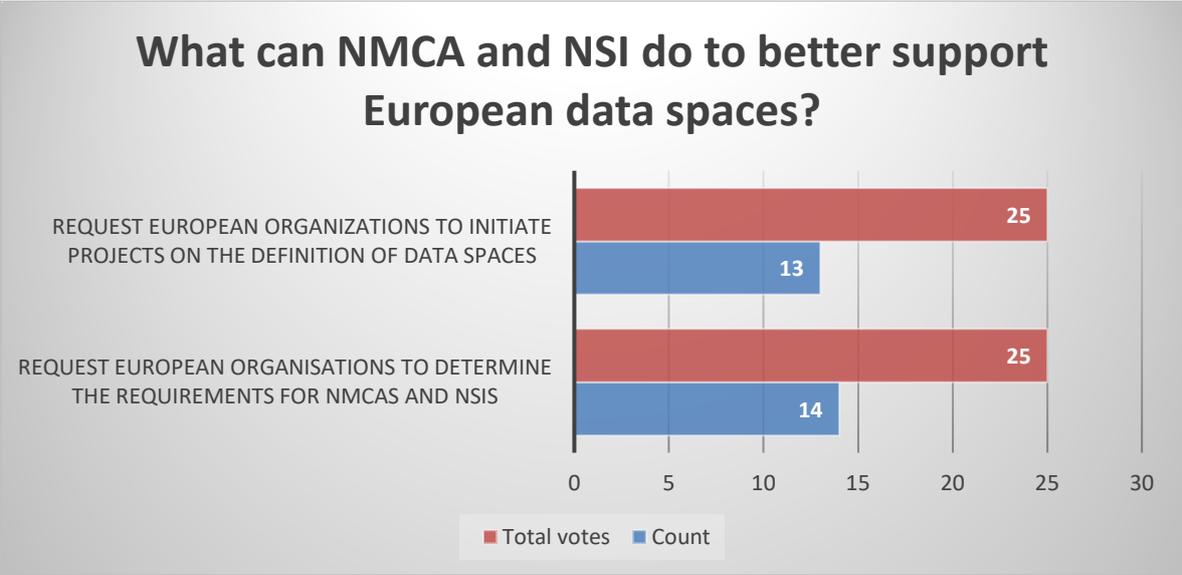


FIGURE 18: RESULT OF THE QUESTION ON HOW NMCAs AND NSI CAN SUPPORT EUROPEAN DATA SPACES

There is no significant difference between both answers.

The question, how NMCA and NSI can better contribute to European data spaces was more discussed within the break-out session of the GISCO meeting.

What do you think needs to be modernized most within NSDI?

Predetermined answers:

- Technical performance (of services)
- Technical standards adaptation to new developments

Result:

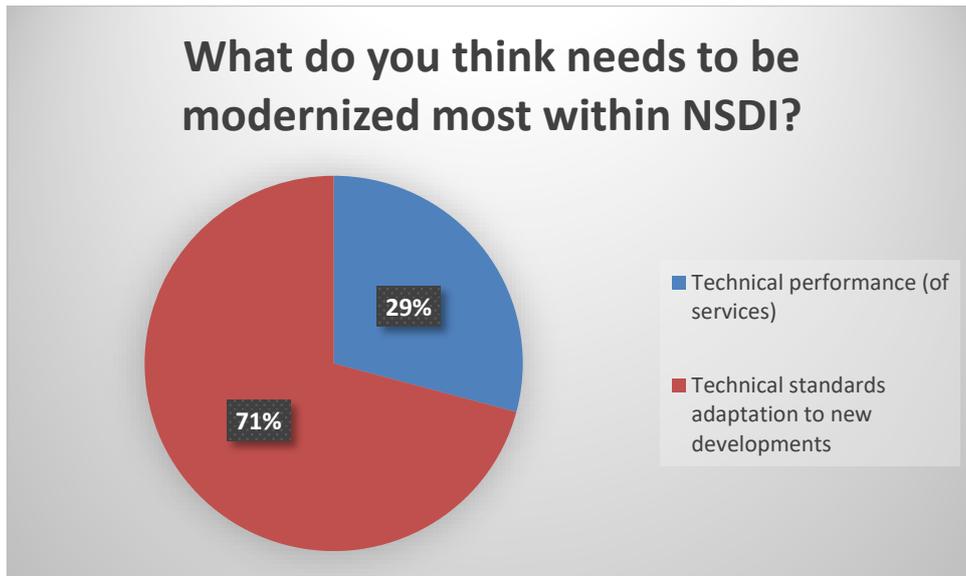


FIGURE 19. RESULT OF THE QUESTION ON THE MODERIZATION OF NSDI

71 % answered for **Technical standards adaptation to new developments**. That is a clear statement of the audience.

Do you have any experiences in modelling/describing knowledge graphs?

Predetermined answers:

- Yes
- No
- It's not an issue yet

Result:

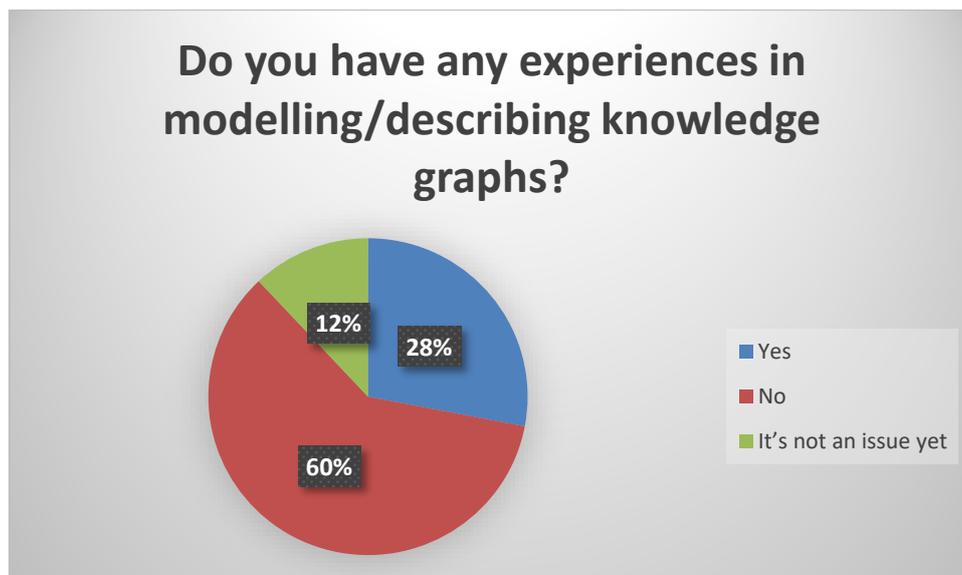


FIGURE 20: RESULT OF THE QUESTION ON KNOWLEDGE GRAPHS

Only **28%** of the participants answered **Yes**. Does this question come too early or are the NMCAs behind trends?

Have you identified data integration methods suitable for your data integration processes that has not been tackled within the document?

Predetermined answers:

- Yes
- No

Result:

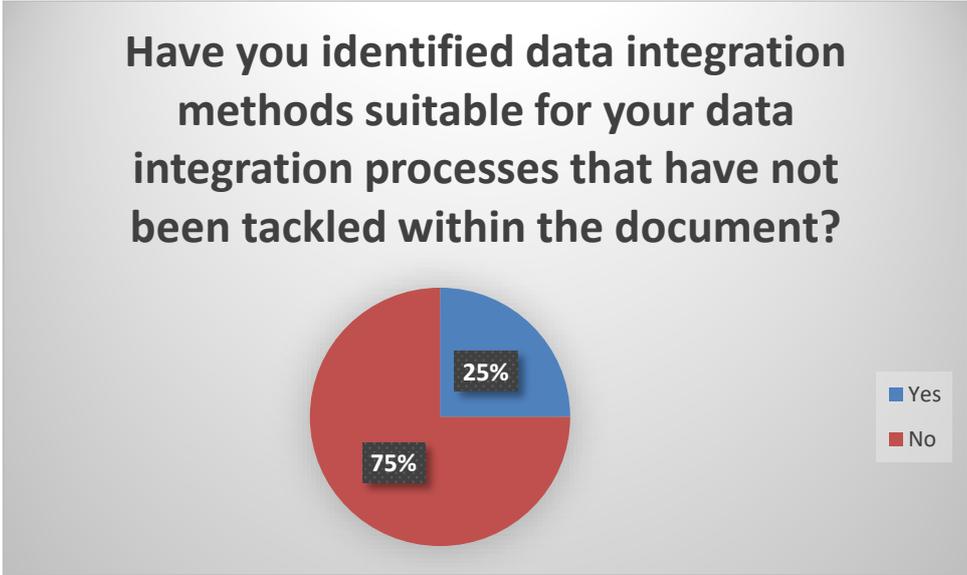


FIGURE 21: RESULT OF THE QUESTION ON MISSING DATA INTEGRATION METHODS WITHIN THE DOCUMENT

Only **25 %** answered **Yes**, the majority of **75%** voted **No**. It would have been interesting to learn from the 25 % what kind of data integrations have been left out or haven't been tackled within the document.

Does your staff have the required knowledge in and capacity for data integration methods?

Predetermined answers:

- Yes
- No
- Partly

Result:

Does the staff within your organisation have the required knowledge and capacity for data integration methods?

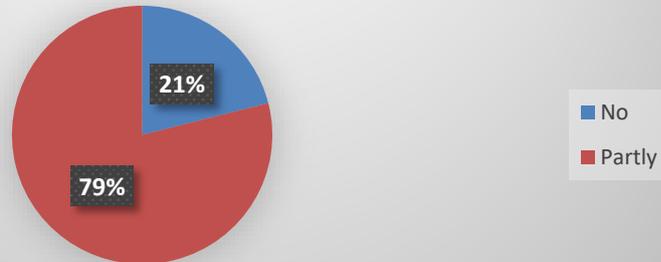


FIGURE 22: RESULT OF THE QUESTION ON THE KNOWLEDGE AND CAPACITY FOR DATA INTEGRATION METHODS

27% of the participants voted **Yes**, **73%** for **No** or **Partly**. It would be interesting to know if the majority is willing to gain knowledge in and capacity for data integration methods?

