19 April 2023

Style Guide for semantic data specifications

interoperable europe

Ground rules for the webinar

Audio Connection

Click "Connect audio" to hear the presenters but please mute your mic when you are not speaking.

Q&A Sessions

Use the function of raising your hand. We will enable your audio for you to address the speakers directly. You can also use the WebEx chat (optionally).
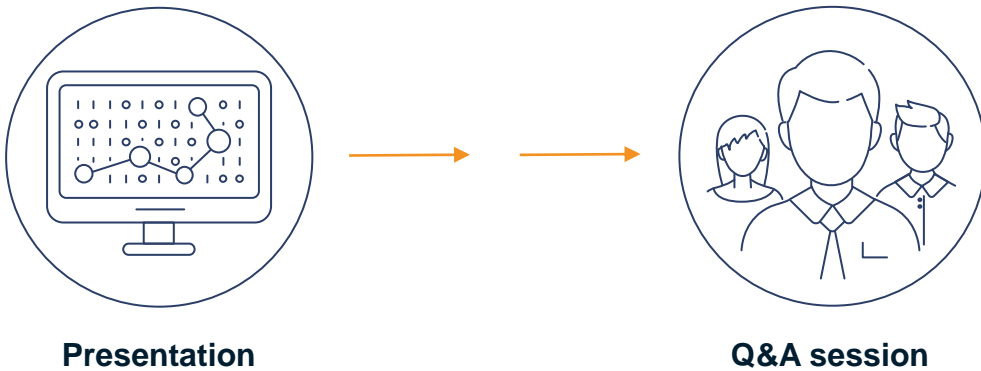
Technical difficulties?

If you have technical issues with WebEx, please submit your questions in writing to the HOST (privately) by using the WebEx chat.

European Commission
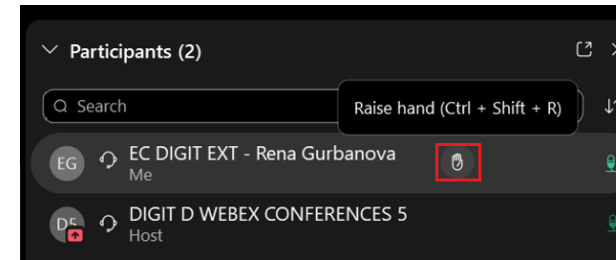
Meeting flow

Presentations of the meeting will be followed by Q&A sessions
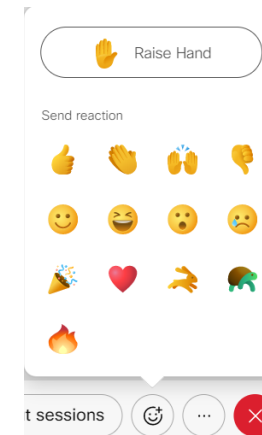
Presentation → **Q&A session**

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

How to raise your hand?

1. Click 'Participants' and then click 'Raise hand' next to your name.

2. You can also raise and lower your hand from the Reactions menu.

European Commission

Agenda

1   **Welcome**

2   **SEMIC overview & webinar introduction**

3   **Tour of the revised Style Guide**

4   **Changelog**

5   **Plenary discussion**

5   **What's next?**

European Commission

SEMIC and introduction to the webinar

Introduction to SEMIC

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

- Promotion, sharing 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

Vocabularies

CORE PERSON VOCABULARY

A person's name(s), date and place of birth/death, identifier, addresses, citizenship, etc.

CORE BUSINESS VOCABULARY

The legal name, address, identifier, company type, and activities of a legal entity.

CORE LOCATION VOCABULARY

The different ways of describing a location, e.g. via an address, a geographic name, or a geometry, in alignment with INSPIRE.

CORE PUBLIC ORGANISATION VOCABULARY

The administrative information, hierarchy, identifiers, events and classification of a public organisation.

CORE CRITERION & EVIDENCE VOCABULARY

The requirements and evidence of a procedure or formal process.

Application Profiles

CORE PUBLIC SERVICE VOCABULARY Application Profile

DCAT-AP FOR DATA PORTALS IN EUROPE

GeoDCAT-AP FOR GEOSPATIAL DATASETS

StatDCAT-AP FOR STATISTICAL DATASETS

ADMS ASSET DESCRIPTION METADATA SCHEMA

European Commission

Objective of this webinar

- Presenting the finalised Style Guide after the public review period

- Looking ahead on next steps

https://semiceu.github.io/style-guide/public-review/index.html

Road to this Style Guide

2012

2015

2020

2023

10 Rules For Persistent URIs

Core Vocabularies Handbook

SEMIC initiate activities to have more coherent and quality assets

Webinars on SEMIC Style Guide and toolkit

+ Growing experience through practice and piloting with(in) the SEMIC community

European Commission

European Commission

Motivation / Context

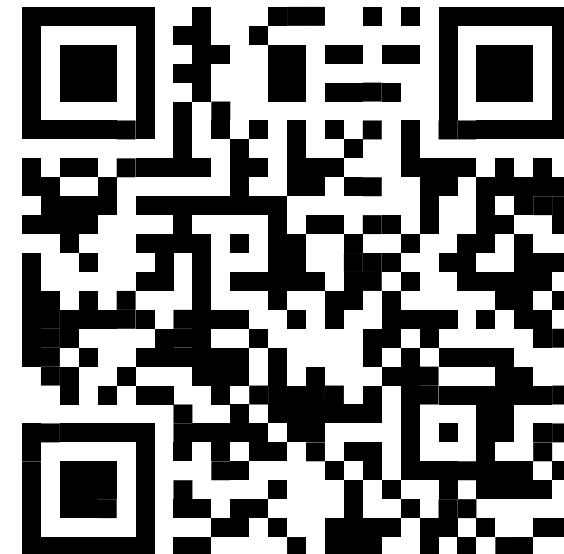**Share** the existing experience and current practices in a document to

- Improve the **quality** of the SEMIC assets
- Improve the **coherency** of the SEMIC assets
- **Streamline and support extension** of Core Vocabularies and Application Profiles by MS
- And, address the challenges of a growing need for data specifications through the **data spaces** initiative

Polling

Poll I

1. Have you read the Style Guide?

2. Have you shared the Style Guide?

3. Have you performed an assessment on your own semantic data specifications?

https://app.sli.do/event/2W6j85pFqMR7PJTFddmVQL

Results Q1

Results Q2

slido

Have you shared the Style Guide? 21

No

52%

Yes

48%

Results Q3

slido

Have you performed an assessment on your own semantic data specifications?   19

Not yet, but I'm planning to

53%

Yes
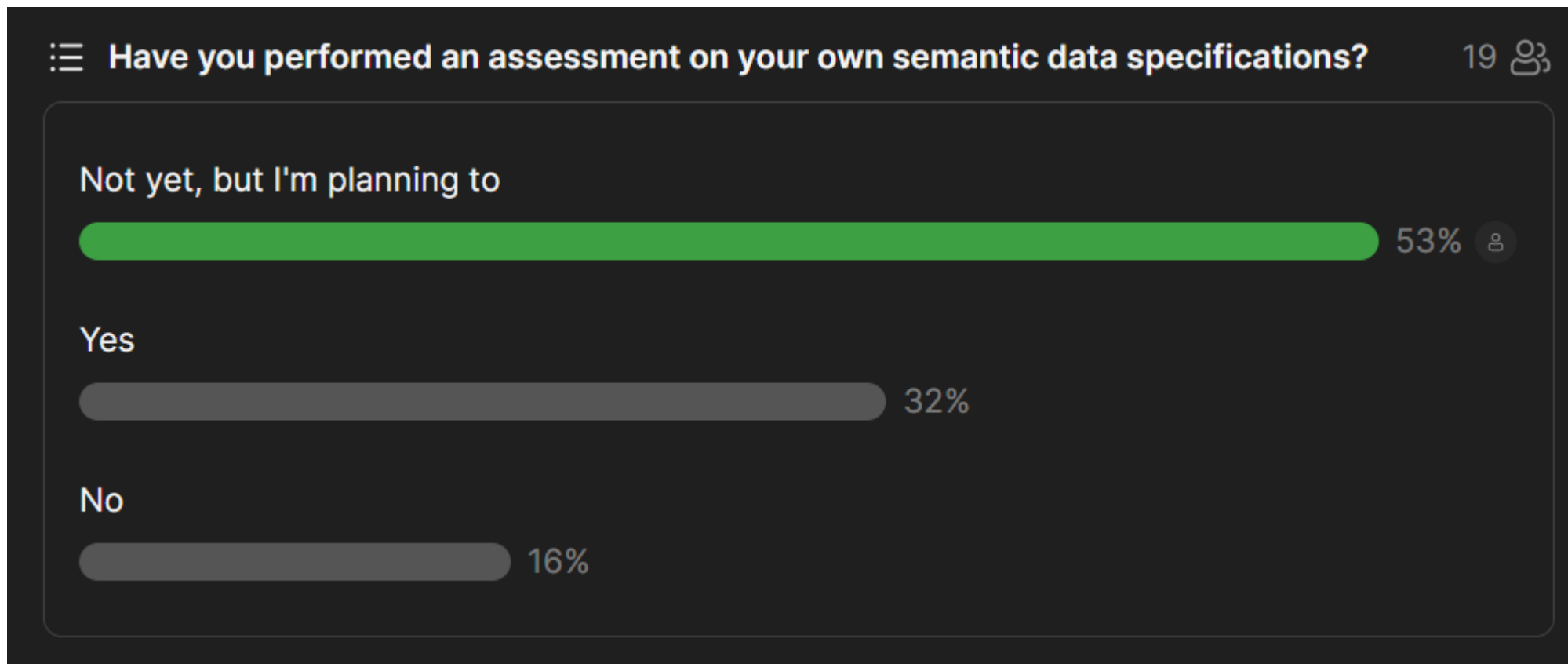
32%

No

16%
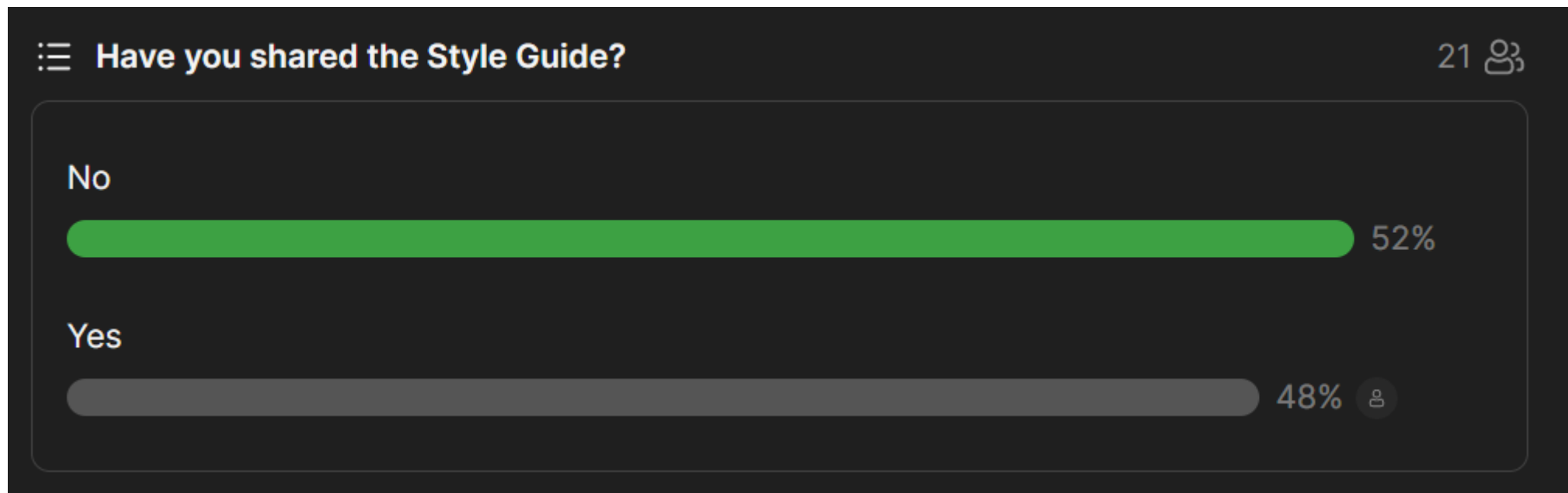
Motivation of the Style Guide

Audience Style Guide

Business/domain experts
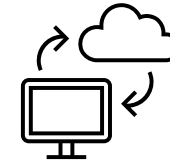
Machines

SEMIC Editors

Data specification
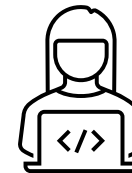
Developers

Audience (summary)

Mainly meant for the **editors** of the data specifications
(e.g. semantic engineers, data architects, knowledge modelling specialists).

The aim is to ensure the creation of **coherent** data specifications that can be read and used by *domain experts*, *developers* and *machines*.

Scoped to what is needed within SEMIC.

Semantic asset types

Semantic data specification

- Core Vocabulary
- Application Profile

Collection of terms
+
Structures/rules how to combine these terms
+
(Re)use expectations

Artefacts

- Persistent URIs
- RDF representation
- OWL representation
- SHACL representation
- HTML representation
- Pictures/Diagrams
- UML representation
- JSON-LD representation
- XML representation
- …

Format
+
Use with a purpose
+
Addressing one concern

Single Source of Truth for the Editor

Data specification

$=$

Conceptual model (UML)

Data shape (SHACL)

Formal Ontology (OWL/RDF)

Specification document (HTML)
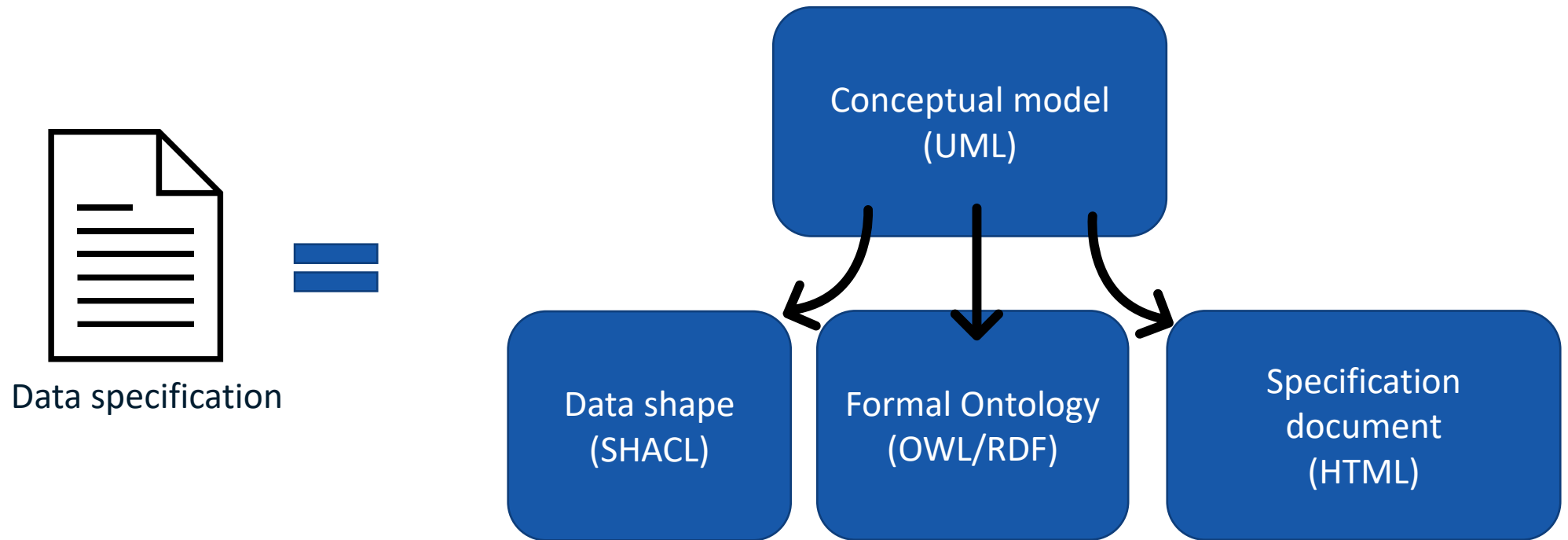
- We have the conceptual model as a SSoT to derive the artefacts
- UML becomes the representation language for the conceptual model

Reuse motivations

As-is

In the specification the term is used in the way as it is being defined in the source specification

With terminological changes

In the specification the term is used in a slightly other wording but with the semantics as defined in the source specification

- Typical case: domain specific words, preferred synonyms

With semantical adaptations

- The term strongly overlaps with the semantics in the source specification, but its usage is independent of that source term, or is narrower.

Organisation of the rules

- General conventions (7) (+2)
- Conceptual model conventions (18) (+3)
- Semantic conventions (5)
- Data shape conventions (5)
- Methodology conventions (3)
- Publication conventions (6)

Organisation/formulation of the rules

Conciseness of the document

- Avoid repetition: that is why more rules are in the conceptual model conventions

Balance between precise scope, strong guidance and leaving room for acceptable variations

- Coherent experience / way of working
- By promoting reuse, choices of others get imported:
  - E.g. opaque URIs versus human readable URIs
  - American English vs British English (dcat:Catalog vs Catalogue)

Provision of a minimal set of aspects to be considered

Tour of the Style Guide

Changes based on the public review

Feedback gathered

Channels
- Bilateral meetings
- Webinar & Mural
- GitHub issues
- External reports

Structured as
- General comments
- Architecture
- Reuse
- Specific feedback
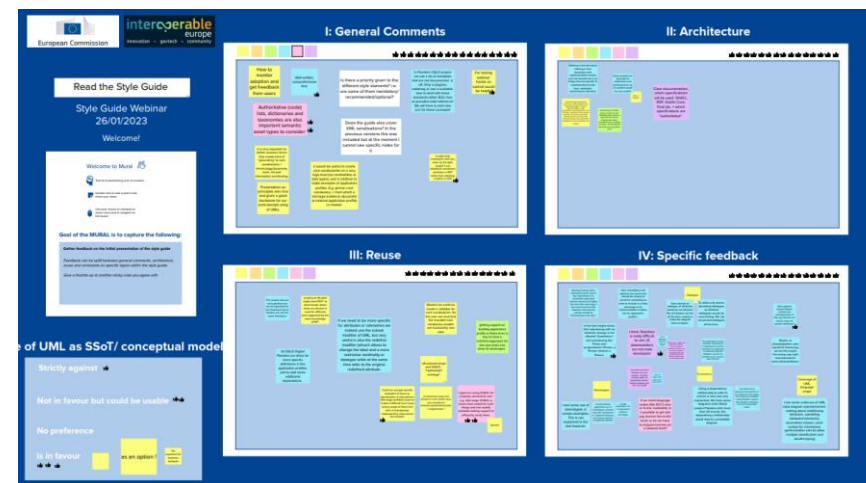
In & out of scope

Categorised the feedback:

- In scope
- Out of scope for the Style Guide
- Out of scope for now but possible to be done in the future

In scope

- Metadata & governance of the document
- Review of the architecture
- Review of the rules
- Review of the terminology
- Review of the reuse section

Out of scope

- [Handbook for using Core Vocabularies (2015)](#) will receive an update based on the lessons learned of the Style Guide

- Actions related to the tooling
  - Tutorial
  - Validator
  - Support on XML interoperability
  - Mapping artefacts
  - …

Possibly to be done in the future

Possible forms:

- bloglike articles,
- expansion to the current Style Guide,
  …

Topics

- Governance, monitoring and evolution of specifications
- Alternatives to the UML conceptual model
- Mapping & alignments
  (to be published by SEMIC team in coming weeks)
- Technical interoperability
- Anti-patterns and bad practices
  (i.e. why you should never treat XML as RDF)
- Do's and don'ts
- Dereferencing schemes / Recipes
- Authoritative standards
- Multilingualism support

European Commission

Major types of changes

- Added metadata & governance

- Extended and rephrased:
  - Terminological clarifications
  - Architectural clarifications (e.g. added contexts for different roles, deontic modals etc.)
  - Reuse compliance

- Editorial fixes
  - Missing titles, rewording of rules, styling and lay-out issues, broken links, …

Changes to Architecture

- Added explanation of the need for a conceptual model as different form plain diagrams (also in the Terminology)

- Added clear inventory of what is used from the UML

- Changed to only use OWL 2 instead of both OWL and OWL 2

Changes to Reuse

- Clarified reuse of URIs and definitions

- Clarified subtyping of Classes
  - emphasised more specific definitions in application profiles being allowed.

- Clarified 'domain' and 'range' constructs

Changes to Rules

- Rephrased the rules
- Changed the rules on enumerations
- Expanded and clarified on stereotype usage
- Clarifications on Datatypes & Classes
- Clarification on 'lightweight ontology'

Changelog

- https://github.com/SEMICeu/style-guide/blob/main/CHANGELOG.md

Break

Assessment of specifications

Self-assessment

Started doing the assessment on SEMIC semantic data specifications ourselves starting with the Core Location Vocabulary:

Total rules	Rules applicable	Total alignment	Percentage of alignment
44	42	39	92.86%

Metric	Definition
Aligned	The asset follows the guideline.
Unaligned	The asset does NOT follow the guideline
Partially aligned	The asset is only partially following the guideline.
To be revised	The asset will be aligned to this guideline.
Not applicable	This guideline cannot be applied for a valid reason

Self-assessment

Started doing the assessment on SEMIC semantic data specifications ourselves starting with the Core Location Vocabulary:

Total rules	Rules applicable	Total alignment	Percentage of alignment
44	42	39	92.86%

Outcome
• Lexicalisation (needs attention)
• Awareness of reuse type (needs improvement)
• Versioning (needs clear version governance)

Self-assessment

Started doing the assessment on SEMIC semantic data specifications ourselves starting with the Core Location Vocabulary:

Total rules	Rules applicable	Total alignment	Percentage of alignment
44	42	39	92.86%

Template & example available on the GitHub Repository

Plenary discussion

What's next?

Governance

- Should be a living document

- To be revised on a biannual basis

- Future revisions based on needs raised via GitHub.

  - Will address the new issues as much as possible but the intention is to release a first stable version.

Adoption

Blog posts

- **Phil Barker**

Adoption/Reuse

- Apply all rules on the SEMIC assets in the coming year
  - Adoption of the style guide in the assets
  - Adoption of the style guide in the tooling

- Already aligned tools
  - model2owl (Publications Office), OSLO toolchain (Flanders)

- Assisting others in their own assessments

Conclusion

- Style Guide is
  - the result of a journey of the past 15 years
  - built on experience, lessons learned, best practices, challenges and guidelines
  - addresses the need for coherency across all SEMIC data specifications
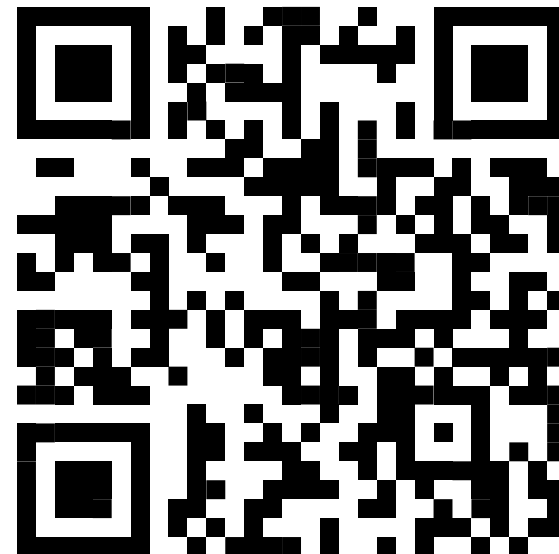
  many thanks to everyone that contributed to SEMIC

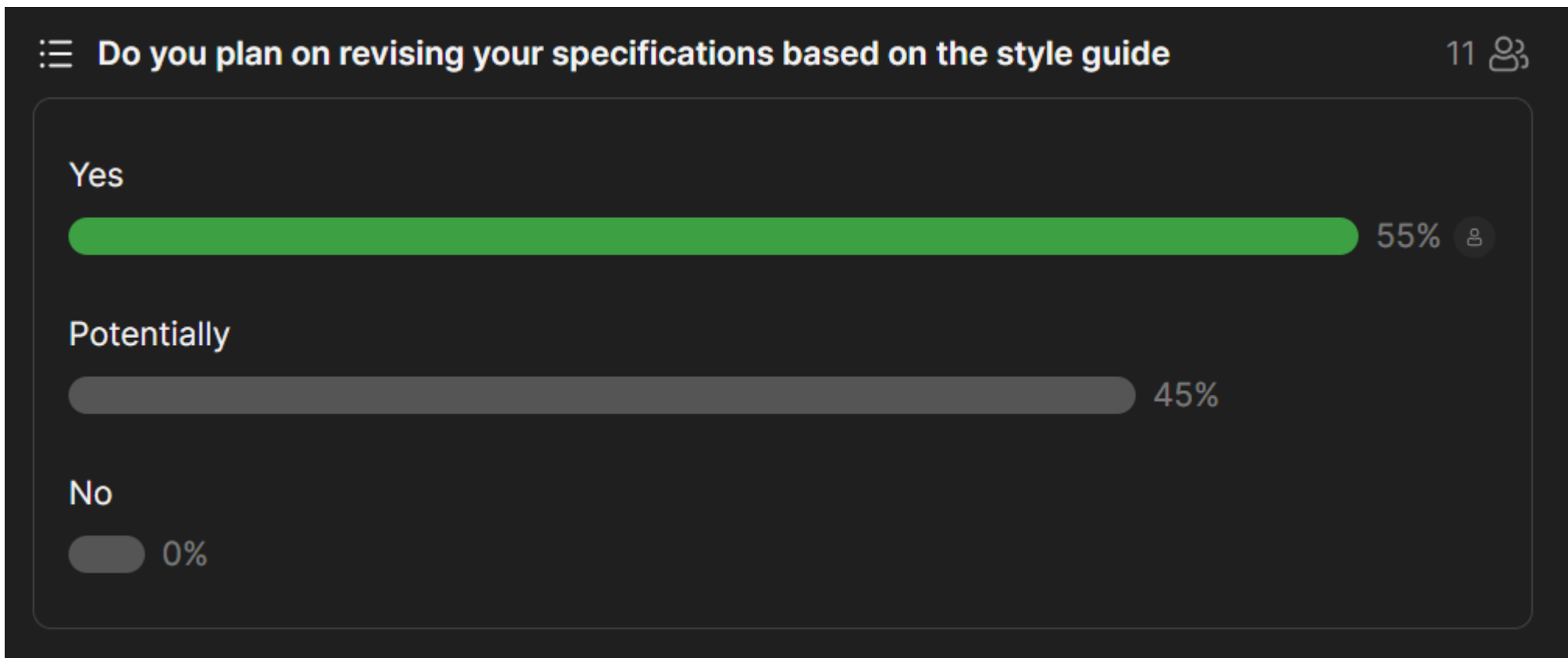- **Release v1.0.0 – 28 April 2023**

Polling

Poll II

1. Do you plan on revising your specifications based on the style guide?

2. Are you planning on adapting the style guide based on your own (national/regional/organisational) needs?

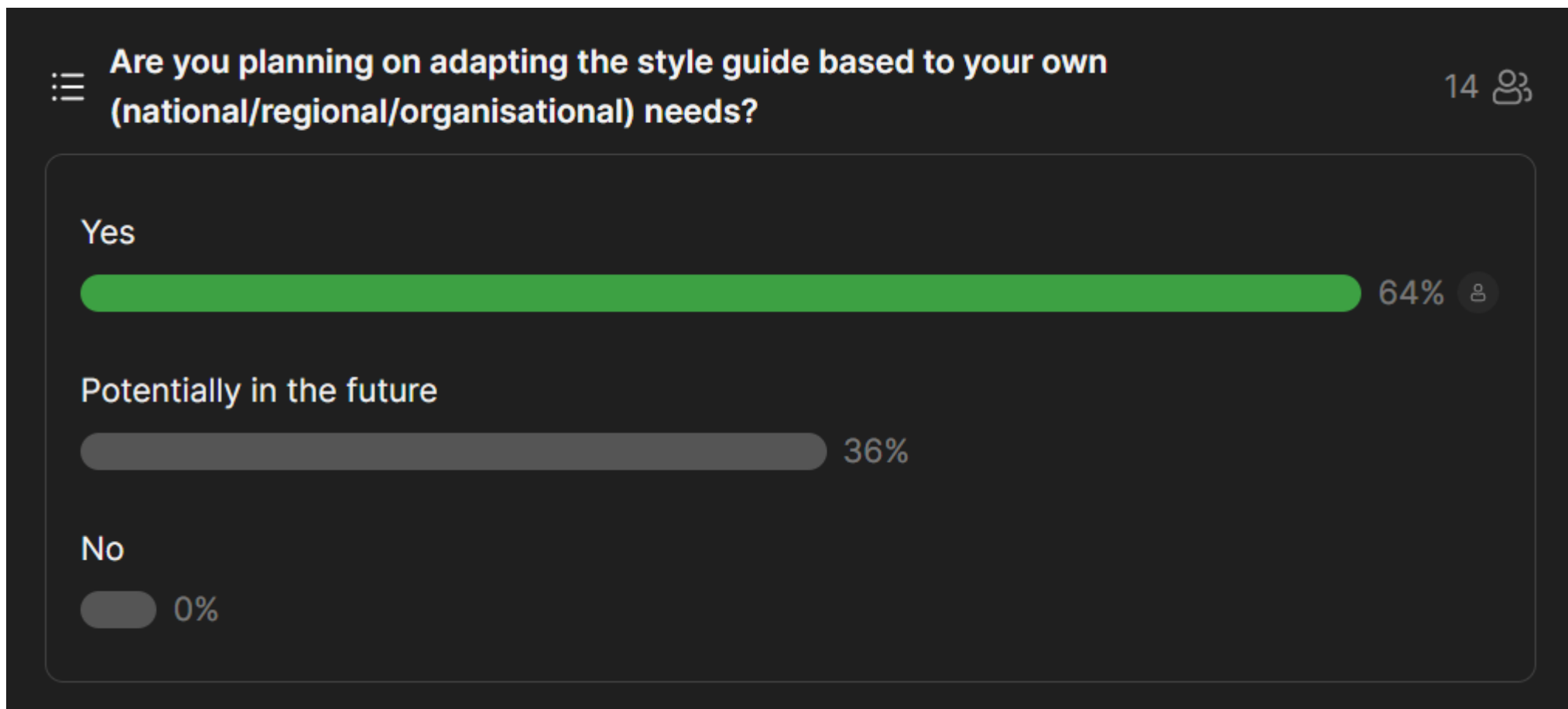3. Which topics would you like to read blog articles about?

https://app.sli.do/event/ivRenPaMCjw25x4893vGGL

Results Q1

slido

Do you plan on revising your specifications based on the style guide   11

Yes

55%

Potentially

45%

No

0%

Results Q2

slido

Are you planning on adapting the style guide based to your own (national/regional/organisational) needs?

14 👥

Yes

64% 👤

Potentially in the future

36%

No

0%

Results Q3

Which topics would you like to read blog articles about? — 14 participants

- Governance, monitoring and evolution of specifications — 57%
- Controlled vocabularies management — 57%
- Mapping & alignments — 50%
- Authoritative standards — 50%
- Bad practices/ do's & don'ts — 43%
- Technical interoperability — 36%
- Recipes — 29%
- Dereferencing schemes — 21%
- Domain/range — 14%
- Multilingualism support — 14%

Wrap up

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