

Metadata in de praktijk(code)

Dag van het Netwerk Statistiek Vlaanderen – 21 maart 2024

TOM DE WINTER



Kwaliteit @ Statistiek Vlaanderen

“ We zorgen ervoor dat de statistieken die we produceren van een zo hoog mogelijke kwaliteit zijn. (...) Er bestaan internationale kwaliteitsstandaarden over de productie en verspreiding van openbare statistiek. Momenteel worden die meestal impliciet gevolgd, maar soms ook niet. Onze ambitie is om kwaliteit centraal te stellen in wat we doen. ”

(Strategisch kader Statistiek Vlaanderen)

Kwaliteit @ Statistiek Vlaanderen

- “In België werden de principes uit de **praktijkcode** van de Europese statistiek door een samenwerkingsakkoord toepasselijk gemaakt op alle openbare statistieken die worden opgemaakt door de statistische autoriteiten van de federale staat, de gewesten en gemeenschappen.”
- “De **protocollen** van Statistiek Vlaanderen zijn een meer gedetailleerde uitwerking van de principes en werkwijzen uit de decretale en Europese regelgeving (...)”

Kwaliteit @ Statistiek Vlaanderen

“ De Praktijkcode heeft tot doel het **publiek vertrouwen** in de Europese statistiek te waarborgen en stelt daartoe vast hoe Europese statistieken in overeenstemming met de statistische beginselen (...) en de beste internationale statistische praktijken moeten worden ontwikkeld, geproduceerd en verspreid. ”

‘European Statistics Law’ – Verordening (EG) Nr. 223/2009 van het Europees Parlement en de Raad - Artikel 11: Praktijkcode voor Europese statistiek)

Kwaliteit @ European Statistical System

- Praktijkcode voor Europese Statistieken
 - 3 domeinen
 - 16 principes
 - 84 indicatoren
- Quality Assurance Framework
 - 440 aanbevelingen



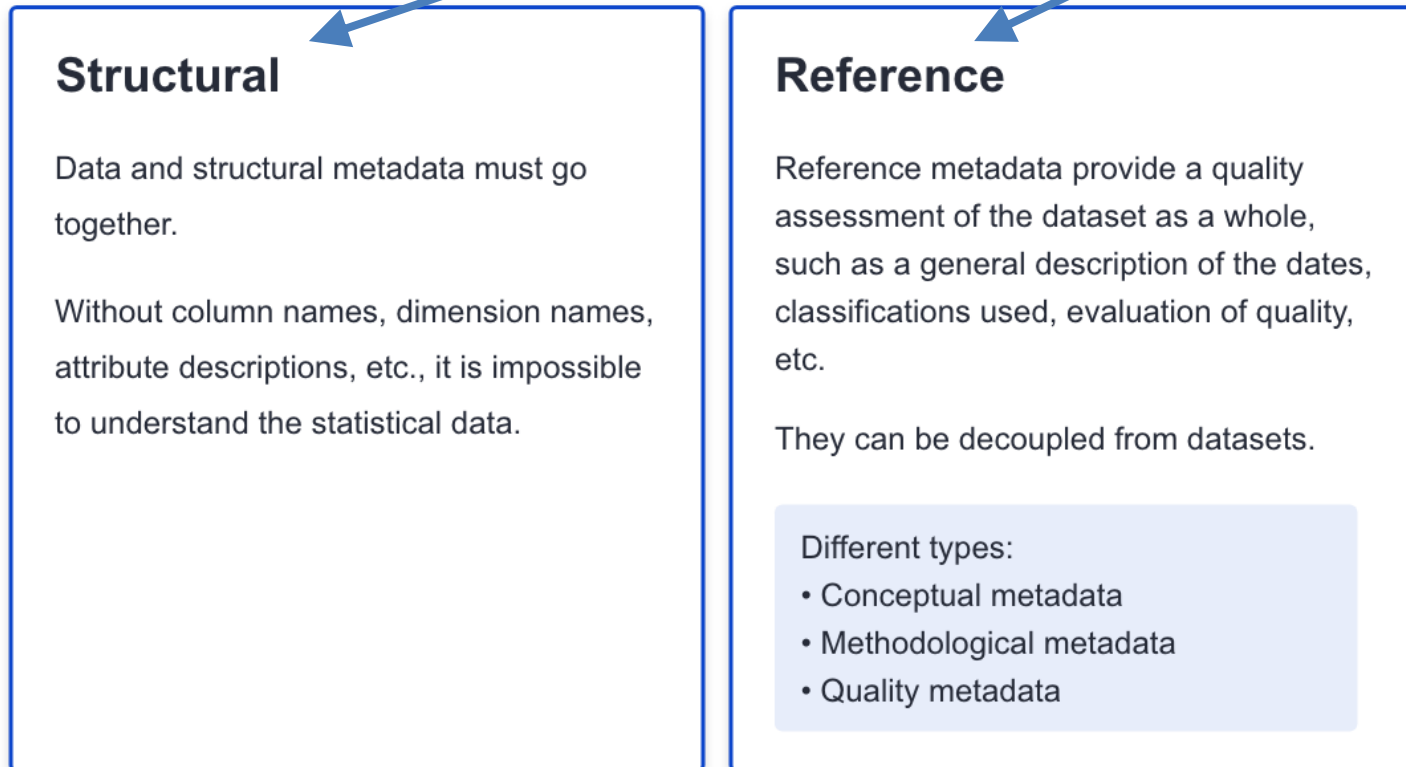
Kwaliteit @ European Statistical System

- **Belang van documentatie!**

- 6.4. Informatie over de gebruikte gegevensbronnen, methoden en procedures is openbaar.
- 8.4. Metadata over statistische processen worden gedurende de gehele statistische verwerking op een passende manier beheerd en verspreid.
- 15.1. De statistieken en de desbetreffende metagegevens worden gepresenteerd en gearchiveerd in een vorm die een juiste interpretatie en zinvolle vergelijkingen mogelijk maakt.
- 15.5. De statistische autoriteiten beheren en verspreiden metadata die aan output zijn gerelateerd, op basis van de Europese normen.
- 15.6. De gebruikers worden op de hoogte gehouden van de methoden die voor statistische processen zijn toegepast, waaronder het gebruik en de integratie van administratieve en andere gegevens.

ESS Metadata

”Gegevens die andere gegevens **definiëren** en **beschrijven**” (Eurostat Glossary)



<https://ec.europa.eu/eurostat/web/metadata/overview>

ESS Metadata: SIMS

- **Single Integrated Metadata Structure (SIMS) (v2)**
 - Samenvoeging van twee metadatastructuren
 - Euro-SDMX Metadata Structure (ESMS) > user report
 - ESS Standard for Quality Reports Structure (ESQRS) > producer report
 - Geharmoniseerde en efficiënte voorbereiding van producer- en user reports
 - Vergelijkingen van processen en outputs vergemakkelijken
 - Faciliteert uitwisseling van metadata

EURO-SDMX Metadata Structure	
5	Reference period
6	Institutional mandate
6.1	Legal acts and other agreements
6.2	Data sharing
7	Confidentiality
7.1	Confidentiality - policy
7.2	Confidentiality - data treatment
8	Release policy
8.1	Release calendar
8.2	Release calendar access
8.3	User access
9	Frequency of dissemination
10	Accessibility and clarity
10.1	News release
10.2	Publications
10.3	On-line database
10.4	Micro-data access
10.5	Other
10.6	Documentation on methodology
10.7	Quality documentation
11	Quality management
11.1	Quality assurance
11.2	Quality assessment
12	Relevance
12.1	User needs

Single Integrated Metadata Structure	
S.5	Reference period
S.6	Institutional mandate
S.6.1	Legal acts and other agreements
S.6.2	Data sharing
S.7	Confidentiality
S.7.1	Confidentiality - policy
S.7.2	Confidentiality - data treatment
S.8	Release policy
S.8.1	Release calendar
S.8.2	Release calendar access
S.8.3	User access
S.9	Frequency of dissemination
S.10	Accessibility and clarity
S.10.1	News release
S.10.2	Publications
S.10.3	On-line database
S.10.3.1	AC1. Data tables - consultations
S.10.4	Micro-data access
S.10.5	Other
S.10.5.1	AC 2. Metadata - consultations
S.10.6	Documentation on methodology
S.10.6.1	AC 3. Metadata completeness - rate
S.10.7	Quality documentation
S.11	Quality management
S.11.1	Quality assurance
S.11.2	Quality assessment
S.12	Relevance
S.12.1	User needs

ESS Standard for Quality Reports Structure	
11	Confidentiality
11.1	Confidentiality – policy
11.2	Confidentiality – data treatment
9	Accessibility and clarity
9.1	News release
9.2	Publication
9.3	On-line database
9.3.1	Data tables - consultations
9.4	Micro-data access
9.5	Other
9.7.2	Metadata - consultations
9.6	Documentation on methodology
9.7.1	Metadata completeness – rate
9.7	Quality documentation
4	Quality management
4.1	Quality assurance
4.2	Quality assessment
5	Relevance
5.1	User needs

ESS Metadata: SIMS

- **19 concepten > 80 subconcepten**

- S.01 Contact
- S.02 Metadata update
- S.03 Statistical presentation
- S.04 Unit of measure
- S.05 Reference period
- S.06 Institutional mandate
- S.07 Confidentiality
- S.08 Release policy
- S.09 Frequency of dissemination
- S.10 Accessibility and clarity
- S.11 Quality management
- S.12 Relevance
- S.13 Accuracy and reliability
- S.14 Timeliness and punctuality
- S.15 Coherence and comparability
- S.16 Cost and burden
- S.17 Data revision
- S.18 Statistical processing
- S.19 Comment



Nederlandstalige handleiding
'Implementatie SIMS'

ESS Metadata: SIMS

- Voorbeeld 1

SIMS	Concept name	Definition	Guidelines
S.09	Frequency of dissemination (U)	The time interval at which the statistics are disseminated over a given time period.	State the frequency with which the data are disseminated, e.g. monthly, quarterly, yearly. The frequency can also be expressed by using a code from the harmonised ESS code list so long as this is considered to be easily understandable by users.

This concept is included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).

ESS Metadata: SIMS

■ Voorbeeld 2

SIMS	Concept name	Definition	Guidelines
	Non-sampling error	Error in survey estimates which cannot be attributed to sampling fluctuations.	Summarise the most important aspects of coverage, measurement, non-response, processing and model assumption errors. Discuss the corresponding bias risks and actions undertaken to reduce them. <i>European level</i> Provide a summary of the above across countries.
S.13.3	A4. Unit non-response - rate (U)	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units, at a level of detail appropriate for a user report.	<i>For probability and census surveys:</i> <ul style="list-style-type: none">• Report A4: unit non-response rates. <i>For repetitive surveys</i> <ul style="list-style-type: none">• also briefly describe the trend for A4. (Note: for producer reports A4 is reported in S.13.3.3.1.)

ESS Metadata: SIMS



Income, consumption and wealth - experimental statistics (icw)

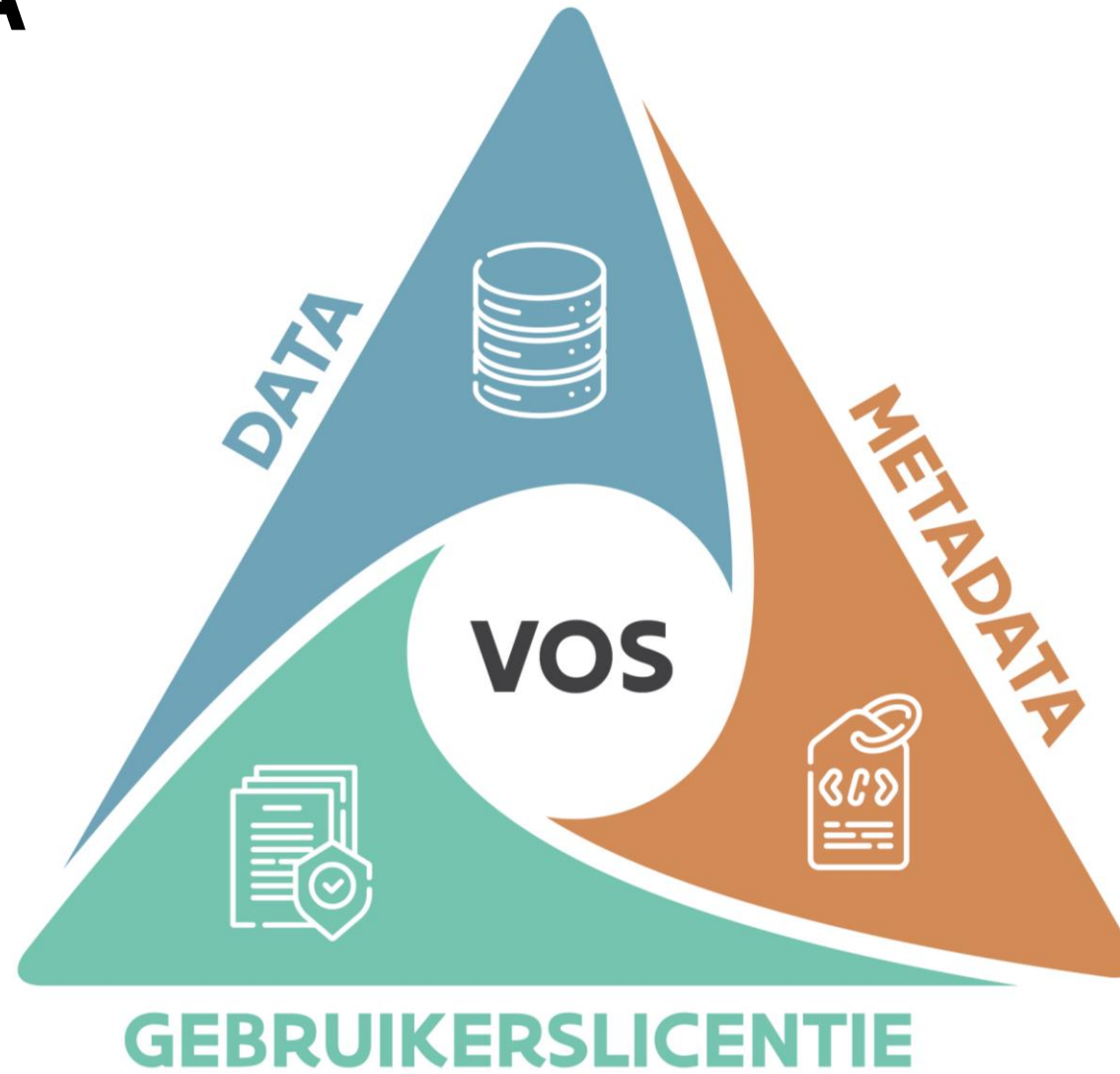
Reference Metadata in Euro SDMX Metadata Structure (ESMS)
Compiling agency: Eurostat, the statistical office of the European Union

Eurostat metadata

Reference metadata

- [1. Contact](#)
 - [2. Metadata update](#)
 - [3. Statistical presentation](#)
 - [4. Unit of measure](#)
 - [5. Reference Period](#)
 - [6. Institutional Mandate](#)
 - [7. Confidentiality](#)
 - [8. Release policy](#)
 - [9. Frequency of dissemination](#)
 - [10. Accessibility and clarity](#)
 - [11. Quality management](#)
 - [12. Relevance](#)
 - [13. Accuracy](#)
 - [14. Timeliness and punctuality](#)
 - [15. Coherence and comparability](#)
 - [16. Cost and Burden](#)
 - [17. Data revision](#)
 - [18. Statistical processing](#)
 - [19. Comment](#)
- [Related Metadata](#)
[Annexes](#) (including footnotes)

SIMS @ VSA



Waarom metadata volgens standaard?

- Omdat het volgt uit de praktijkcode en indicatoren
- Omdat we zo concrete richtlijnen / good practices hebben rond statistiekdocumentatie (Wat is goede documentatie? Wat is voldoende documentatie?)
- Maakt systematische kwaliteitsmonitoring mogelijk (Heb ik alle info die ik nodig heb om de kwaliteit te kunnen evalueren?)
- Faciliteert technische implementatie en data-uitwisseling

✉ tom.dewinter@vlaanderen.be

in be.linkedin.com/in/tom-de-winter >>>>>>>>>>>>>>>>

