Government data quality framework

Essi Kaukonen

Quality Knowledge Exchange Network 1.12.2022

Statistics Finland

Government data quality framework

- Context
- Data Quality Criteria and Metrics
- Implementation

- The amount of data is ever-increasing and so are the data opened and governed by government institutions.

- In the following, it is essential to recognise and describe the quality of the government data uniformly.

- Moreover, the more we wish to reuse the data for some other purposes than the original one, the more important it is to verify the quality of the data

Statistics Finland

The project on opening up and using public data (TiHA) 30.4.2020 – 31.12.2022

- The aims set out in the Government Programme of Prime Minister Marin's Government include adding depth to the management of information policy.

- TiHA project will put into effect the aims given in the Government Programme by promoting wider and more effective use of public data throughout society.

- Set up by the Ministry of Finance.

- TiHA project Work packages
  - WP1 Strategic aims
  - WP2 Opening up data, including common practices
  - **WP3 Data Quality**: led by Statistics Finland
  - WP4 Semantic and technical interoperability, including guidelines for APIs

Statistics Finland

NLS NATIONAL LAND SURVEY OF FINLAND

Valtiokonttori Statskontoret State Treasury

TULLI (Finnish Customs)

VERO SKATT (Tax Administration)
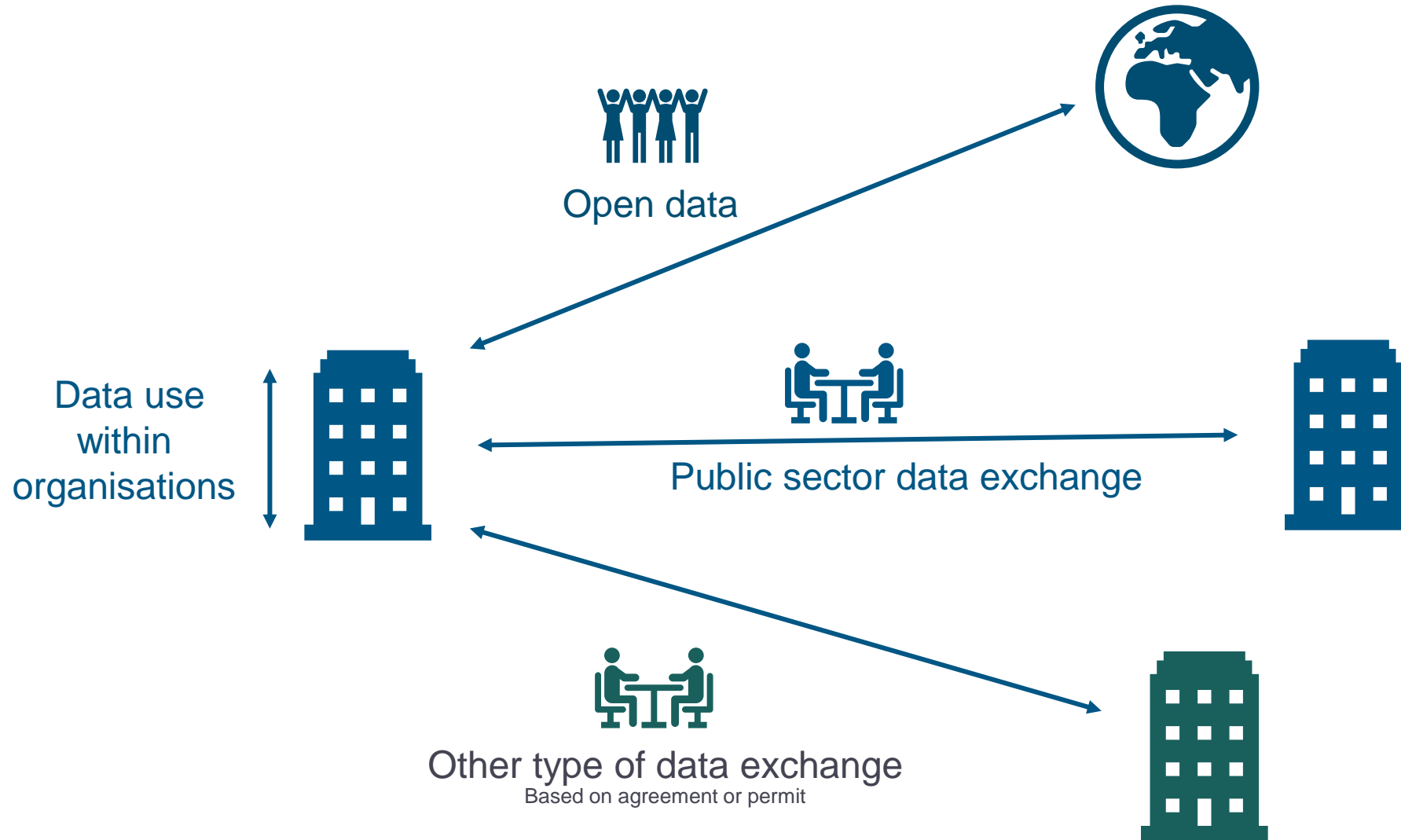
DIGITAL AND POPULATION DATA SERVICES AGENCY

Finnish Institute of Occupational Health

PRH FINNISH PATENT AND REGISTRATION OFFICE

Kela (The Social Insurance Institution of Finland)

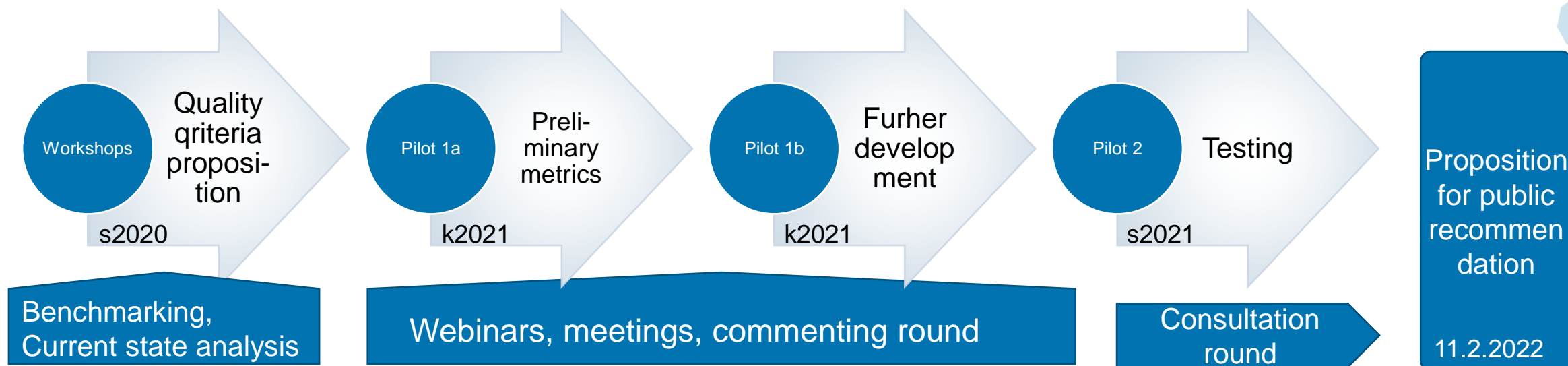FINNISH NATIONAL AGENCY FOR EDUCATION

Statistics Finland

Broad scope of the data quality framework

Open data

Data use within organisations

Public sector data exchange

Other type of data exchange
Based on agreement or permit

Statistics Finland

Developing the data quality criteria

Principles:
– Based on standards
– Applying
– Discussion and Collaboration
– Piloting
– Iteration

Workshops | Quality qriteria proposi-tion | s2020

Pilot 1a | Preli-minary metrics | k2021

Pilot 1b | Furher develop ment | k2021

Pilot 2 | Testing | s2021

Proposition for public recommen dation | 11.2.2022

Benchmarking, Current state analysis

Webinars, meetings, commenting round

Consultation round

Statistics Finland

The Data Quality Criteria and Metrics

- The data quality criteria and indicators are a tool for describing and evaluating data quality; answers to the question: What does data quality mean?

- The aim is to support extensive and versatile use of data throughout society.

- Quality criteria and indicators always reflect the current situation. User experiences, needs and changes in the operating environment effect the content of this tool.

Statistics Finland
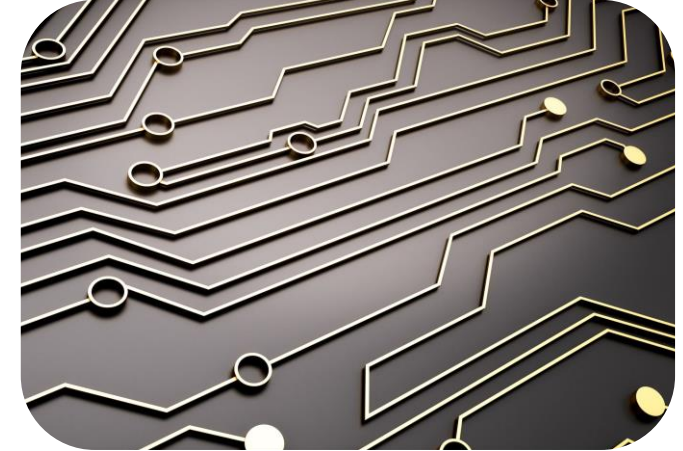
The Data Quality Criteria Groups

How well do the data describe reality?

- Correctness
- Accuracy
- Completeness
- Consistency
- Currentness

How well are the data described?

- Traceability
- Understandability
- Compliance

How can the data be used?

- Portability
- User rights
- Punctuality

Statistics Finland

How well do the data describe reality?

Correctness
methodically produced values
incorrect values
misclassification

Accuracy
standard deviation
outliers

Completeness
temporal target coverage
regional target coverage
target units
shortcomings in characteristics
missing units
additional units
incomplete units
incomplete characteristics

Consistency
Logic of data reviewed

Currentness
Baseline period
creation period
review period
change period

How well are the data described?

Traceability
data lifecycle
data source
change management

Understandability
dataset descriptions
definitions of concepts
data descriptions of characteristics
customer feedback on comprehensibility

Compliance
regulations and standards to be complied with

How can the data be used?

Portability
The dataset data model
permanent identifier of the target unit
customer feedback on portability

User rights
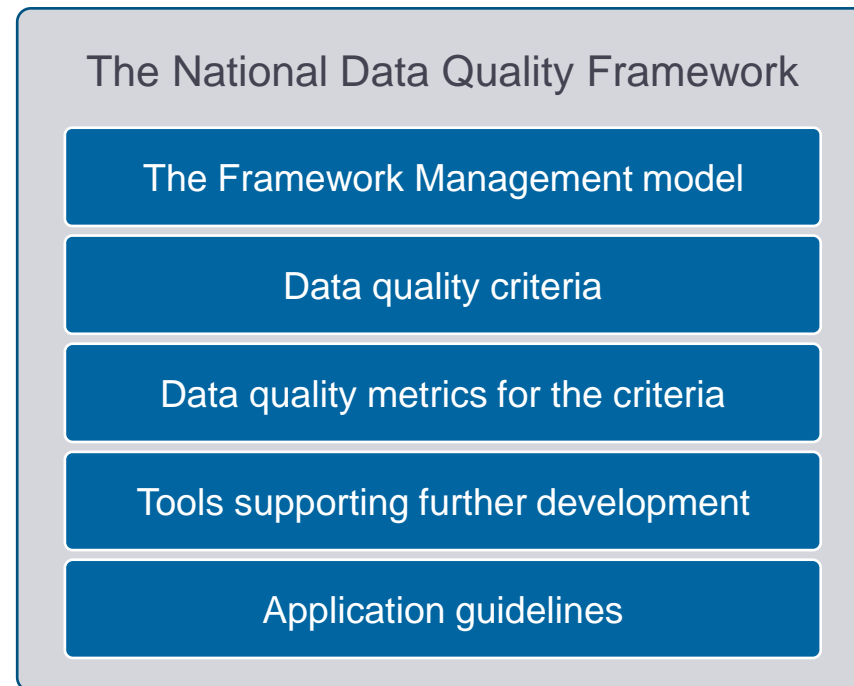access rights
restrictions on use

Punctuality
compliance with due dates
frequency of updates
values changed in the update

Preliminary translations into English
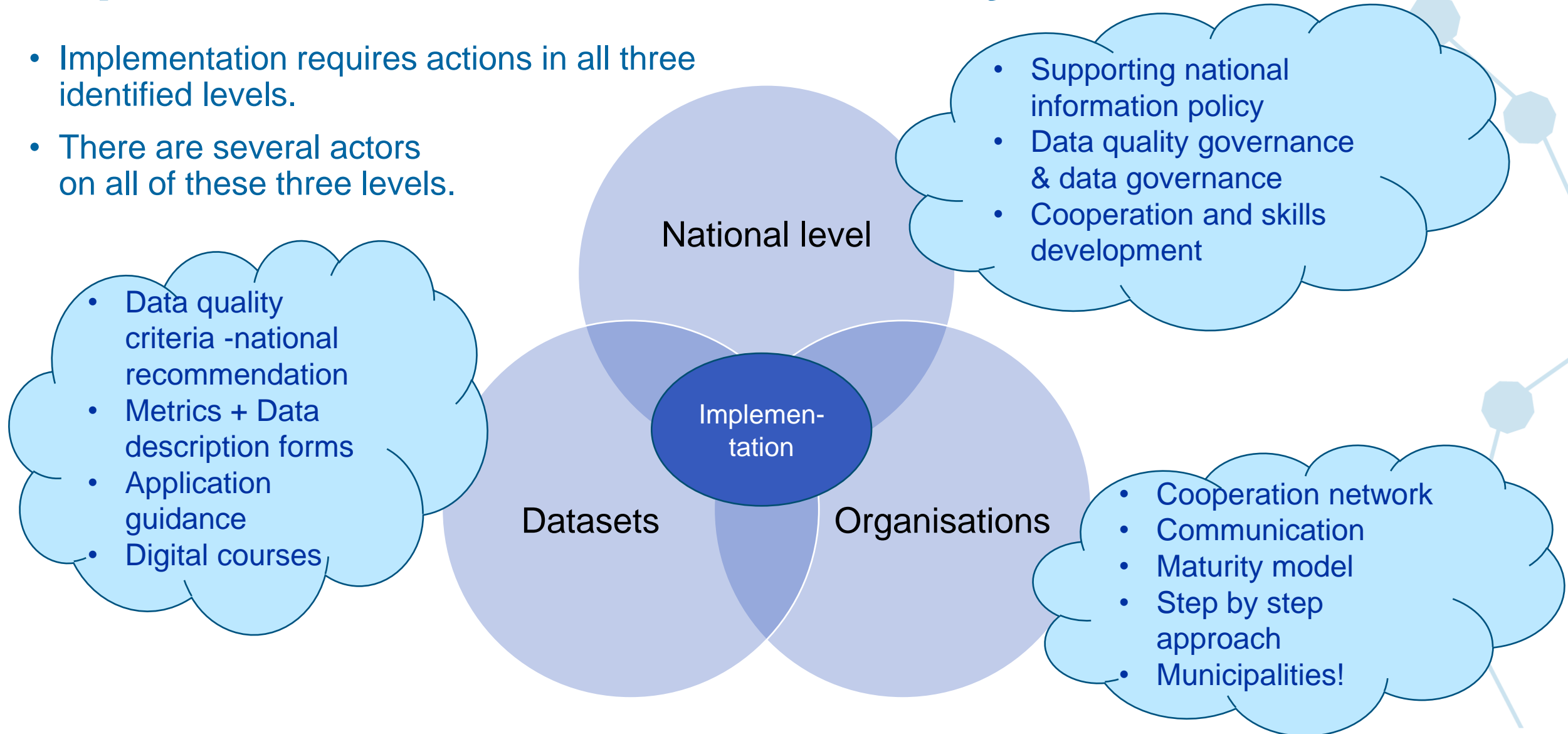
Statistics Finland

The National Data Quality Framework

- The defined data quality criteria and the metrics represent the present understanding about data quality.

- It is important to gather user experience and from time to time, review up-to-dateness of the criteria and the metrics.

- The data quality criteria and the metrics together with models and tools supporting their implementation and management form together the National Data Quality Framework.

The National Data Quality Framework

The Framework Management model
Data quality criteria
Data quality metrics for the criteria
Tools supporting further development
Application guidelines

Statistics Finland

Implementation of the Data Quality Framework

- Implementation requires actions in all three identified levels.

- There are several actors on all of these three levels.

National level

- Supporting national information policy
- Data quality governance & data governance
- Cooperation and skills development

- Data quality criteria -national recommendation
- Metrics + Data description forms
- Application guidance
- Digital courses

Implemen-tation

Datasets

Organisations

- Cooperation network
- Communication
- Maturity model
- Step by step approach
- Municipalities!

Statistics Finland

Further information

- [The project on opening up and using public data](#)
- [The data quality framework](#) (site will be updated soon)

Taking evidence-based decisions require high quality data

Statistics Finland