


# Survey results: Data quality in the application of agricultural system data

Jannes Uhlott (Julius Kühn-Institute, JKI), Anne Sennhenn (Institute for Agricultural Engineering and Bioeconomy, ATB), Markus Möller (Julius Kühn-Institute, JKI) on behalf of the FAIRagro consortium

## Methods

 online available from 14.11.2023 to 21.01.2024

 321 participants (in total)

 LimeSurvey

 211 participants (complete)

## Conclusion

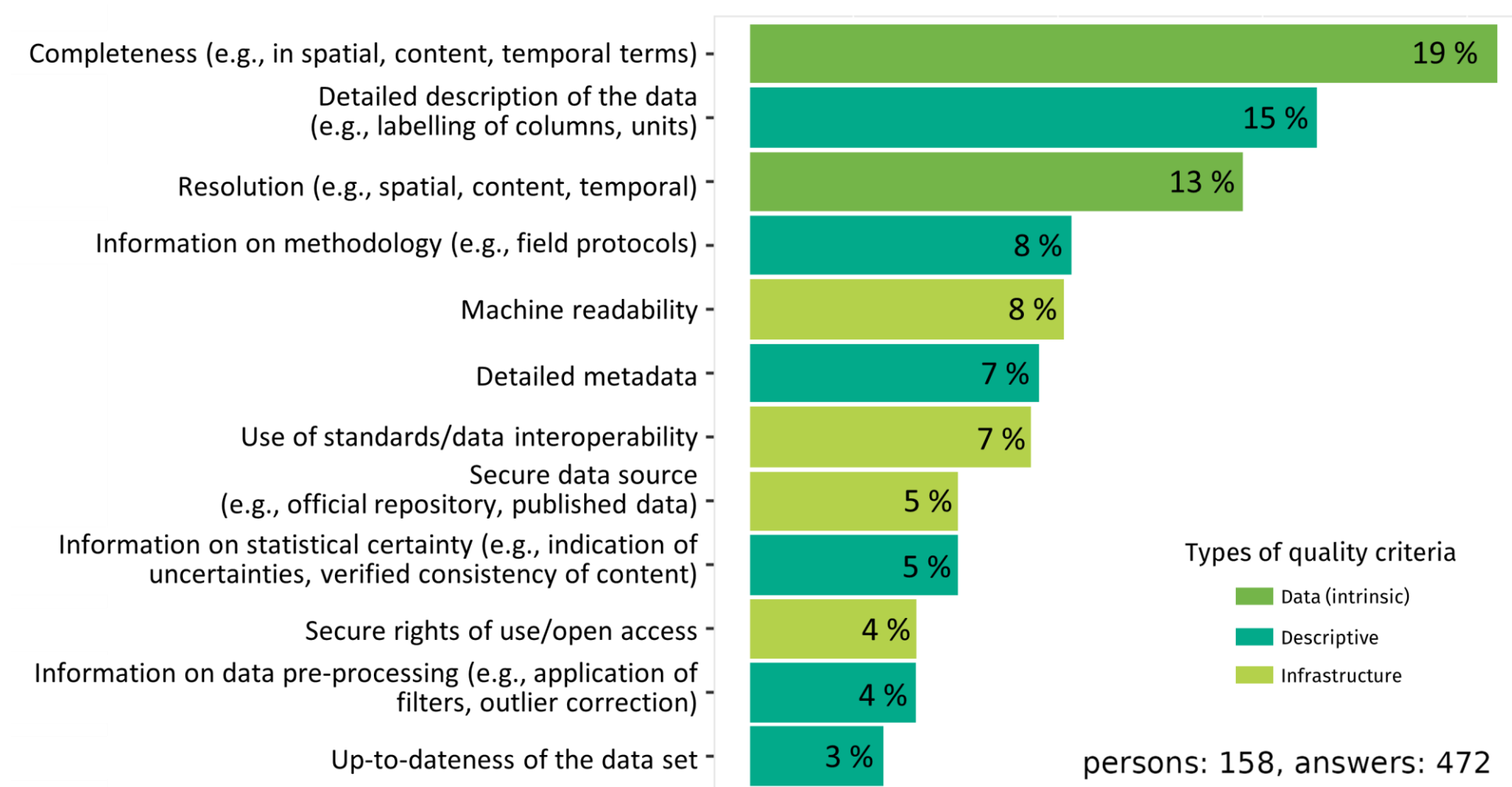
- Identification of key data quality criteria (completeness, detailed description of data, resolution) for future focus.
- Community awareness regarding the importance of data quality and documentation is established.
- Data quality and metadata remain areas of limited understanding.
- Limited time availability, but willingness to invest 1-2 hours per data set.

## Results

*'The idea of FAIR is good, but it means that some researchers have to invest a lot of time and effort in data documentation so that statistics-oriented researchers and modellers can use it - without any benefit.'*

### Data quality criteria

What are the three criteria for data quality that should be met at a minimum for you to be able to reuse data effectively?



- Three relevant criteria
- Intrinsic criteria have the highest importance
- Descriptive criteria are more important than metadata itself

## Outlook

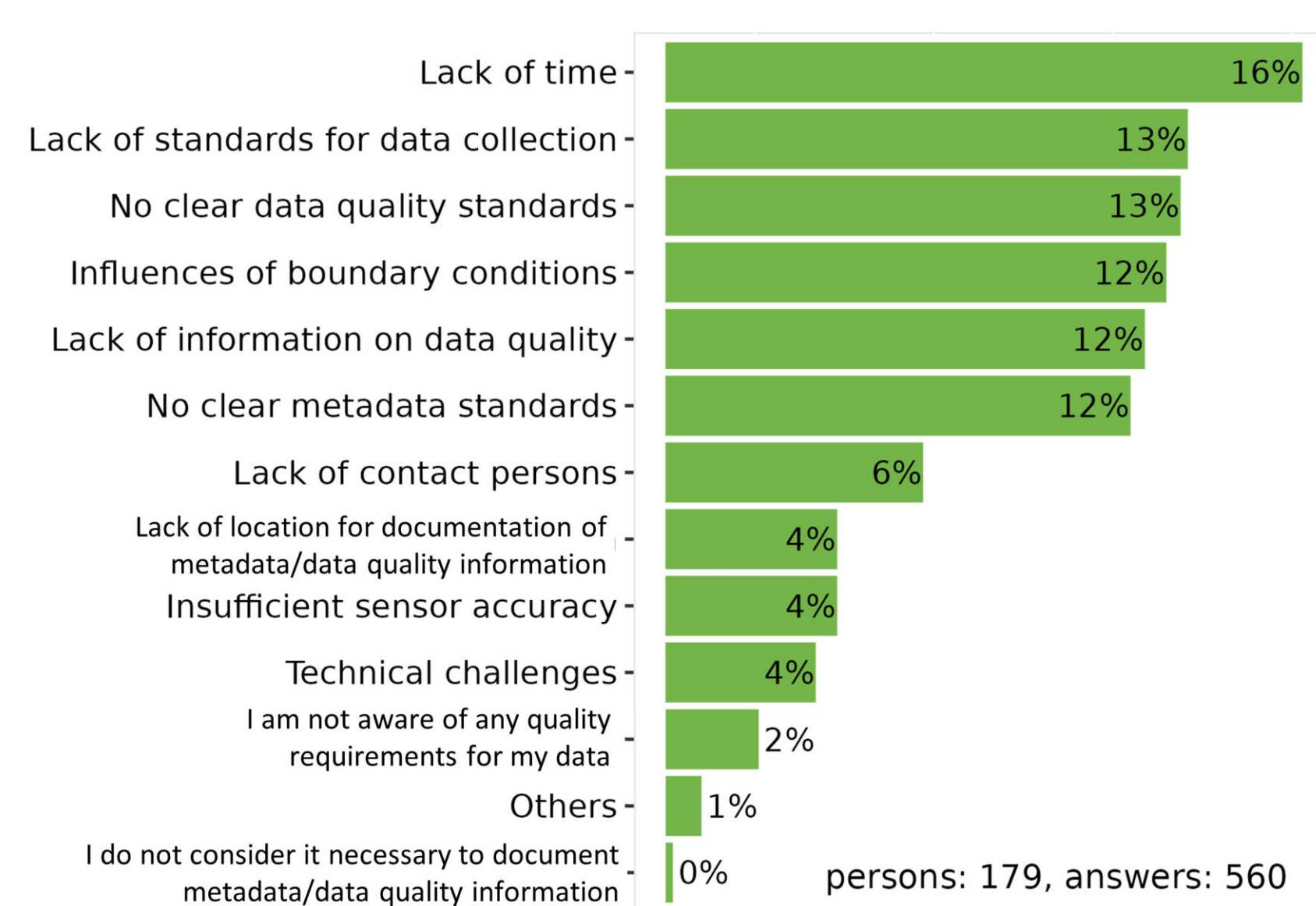
- Not all data types, categories and application areas can be handled with one generic solution  
→ Focus on different areas (e.g. geodata)
- Provide information about metadata and data quality as well as explain why they are a benefit for everyone  
→ Workshops, Trainings, Blog
- Create standards and quality labels

*'I am shocked how less importance is attached to the actual quality of the data and its validity in our institution, although there is talk of 'good scientific practice' etc. everywhere.'*

## Challenges

*'The poorer the data quality, the longer it takes to process and the more assumptions have to be made.'*

What are the biggest challenges for you in data collection regarding data quality?



- High answer rate (3.1 average answers per person) → Multiple important challenges
- Main challenges:
  - Lack of time
  - Missing standards (data collection, data quality, metadata)
  - Influences of boundary conditions (e.g. weather)

What are the biggest challenges for you in data reuse concerning data quality?

- *'Too many systems, each with their own requirements, no 'easy to use' output system, data constantly requires different specialized knowledge'*
- *'I have no idea about it.'*
- *'It takes a lot of time to assess/check whether the data quality is satisfying.'*
- *'The data I need for my work is not available in the appropriate resolution or the costs are too high.'*
- *'Missing/inaccurate data labelling, careless errors during data entry.'*
- *'Non-transparent data collection methods, poor collection protocols including incomplete data.'*
- *'Data protection.'*