



Contribution ID: 46

Type: **not specified**

## Quality aspects of machine learning in official statistics

*Wednesday, 15 May 2024 16:10 (20 minutes)*

Advanced statistical and machine learning models methods are becoming increasingly important in applied data science. At the same time, their trustworthiness is critical for the progress and adoption of data science applications in various fields, including official statistics.

„Bad quality reduces trust very, very fast.“ Taking up this dictum, official statistics in Germany have considered what a quality concept for the use of machine learning could look like. Six quality dimensions (including explainability and robustness) and two cross-sectional aspects (including fairness) were developed, as well as concrete guidelines on how compliance with these can be measured. All this with the aim of deriving a binding standard for official statistics.

The talk will highlight motivation, genesis and first practical implementations of this concept of quality.

Main references:

- Yung W, Tam S-M, Buelens B, Chipman H, Dumpert F, Ascari G, Rocci F, Burger J, Choi I (2022) A quality framework for statistical algorithms. *Stat J IAOS* 38(1):291–308. <https://doi.org/10.3233/SJI-210875>
- Saidani Y, Dumpert F, Borgs C et al (2023) Qualitätsdimensionen maschinellen Lernens in der amtlichen Statistik. *ASTa Wirtsch Sozialstat Arch*. <https://doi.org/10.1007/s11943-023-00329-7>

### Type of presentation

Contributed Talk

**Primary authors:** DUMPERT, Florian (Federal Statistical Office of Germany); SAIDANI, Younes (Federal Statistical Office of Germany)

**Presenter:** DUMPERT, Florian (Federal Statistical Office of Germany)

**Session Classification:** Contributed session

**Track Classification:** Spring Meeting