Open-Source Software implementing Statistical Methods for Official Statistics

Alexander Kowarik
Head of Statistical Methods and the Future Lab

Rome, 2 December 2025 Fourth Workshop on Methodologies for Official Statistics

www.statistik.at

Independent statistics for evidence-based decision making

STATISTICS AUSTRIA
The Information Manager

Sharing methodologies has a long history (my biased list)

- 1969 **X-11** Fortran code for seasonal adjustment and **X11-ARIMA 1980** by StatCan
- Since early 1990s **CALMAR**, **CLAN** … model assisted estimation / calibration
- 1998 **POULPE** … variance estimation
- Late 1990s … **tauArgus** and **muArgus** co-developed in European projects
- 2001 **CANCEIS** .. Editing software
- 2007 **sdcMicro** … Microdata protection
- 2008 **VIM** … R-package for Imputation
- 2009 **sdcTable** … Tabular data protection
- 2011 **editrules** … Fellegi-Holt editing (suceed by validate and errorlocate)
- …
- 2019 **surveysd** … Calibration and Bootstrap error estimation
- 2021 **GaussSuppression**: Tabular Data Suppression using Gaussian Elimination
- 2025 **banffprocessor** … Python BANFF reimplementation

A shared direction is emerging across institutions

- Clear methodological foundations
- Reuse of existing solutions (e.g., validate, VIM, errorlocate)
- Practical implementation experience in real statistical processes
- Community mindset → from isolated development to an ecosystem of tools

Open-source in official statistics is not only a technical shift, but a community evolution.

- Open-source as an enabler of transparency and reproducibility
- Strong alignment with UNECE Open Source Guiding Principles
- Increasing maturity: moving from development/experimentation of methods to structured governance

Principles

1. OSS by default
2. Work in the open
3. Improve and give back
4. Think generic statistical building blocks
5. Test, package and document
6. Choose permissive
7. Promote

Collaboration and reuse accelerate methodological progress

- Strengthen cross-institutional reusability and alignment
- Broader adoption of maturity indicators (testing, documentation, governance)
- Sustainable maintenance models and shared ownership
- Moving toward composable, generic building blocks rather than bespc

uRos2026: Paris, 16-18 November 2026
The Use of R in Official Statistics

The key challenge: Turning community into a sustainable ecosystem

- How can we scale governance and collaboration beyond single NSIs?
- How do we make reuse the default rather than the exception?
- What does a sustainable shared maintenance model look like?
- How do we ensure long-term alignment with the OS principles?

- Sharing implemented methodologies has a long running history in methodology, often based on single developer or a small group. How can this be change to a multi-institutional team?