# The Romanian Experience: Building Effective Collaborations Between Academia and Official Statistics

**Ana-Maria Ciuhu**

Romanian National Institute of Statistics (INS Romania)
Institute of National Economy, Romanian Academy
E-mail: dobre.anamaria@hotmail.com

## Abstract

The integration of academic research and official statistics is important for advancing both fields and improving the quality of data and analyses. This presentation explores the successful collaboration between the National Institute of Statistics (INS) Romania and academic or other statistical institutions, highlighting the mutual benefits derived from this partnership. By bridging expertise in sampling methodologies and statistical applications, these collaborations enhance research methodologies and improve data quality. A core aspect of this collaboration involves sampling methodologies, which require rigorous research and academic input. By working closely with academic researchers, INS Romania aims to improve its sampling procedures. This partnership not only improves the accuracy and reliability of statistical data but also enriches academic research with real-world applications and data. Working at the Romanian Academy provides valuable insight into using statistical data, allowing for a comprehensive understanding of the entire statistical process. This dual perspective helps to identify and address key challenges in the field, driving improvements in both data production and utilization. The uRos (Use of R in Official Statistics) conferences serve as proof to these collaborative efforts, fostering international cooperation between official statistics bodies and academia. These conferences provide a platform for sharing knowledge, methodologies, and experiences, showcasing how INS Romania collaborates with other national statistical institutes worldwide to enhance the global statistical community. This paper discusses key case studies and initiatives that exemplify effective collaboration between INS Romania and the academic sector, underscoring the importance of such partnerships in driving innovation and excellence in both official statistics and academic research.

**Keywords:** official statistics; academia; research; sampling methodology; capacity building; open-source; R

## 1. Introduction

National statistical institutes face rising expectations for timely, granular, policy-relevant indicators, while maintaining strict requirements for methodological soundness, confidentiality, and international comparability. Academic research evolves rapidly in domains such as sampling theory, small area estimation, statistical disclosure control, and reproducible computation. Connecting these ecosystems is therefore strategic: NSIs gain methodological innovation and independent validation. Meanwhile, academia gains access to real-world constraints and relevant applications.

This paper documents the Romanian experience, focusing on collaboration mechanisms between INS Romania and academia, as well as cooperation with other statistical institutions. The description is anchored in: methodological co-development; capacity building; and international community-building via the uRos conference series, which has become a practical and regular vehicle for open-source cooperation.

## 2. Particularities of academia-NSI collaboration

Collaboration creates value because it combines complementary capabilities at institutional level. For NSIs, partnerships provide early access to advanced methods, independent peer review, and a talent pipeline trained in official-statistics constraints. For academia, partnerships provide complex, high-impact problems and an environment where methods must be robust under operational and legal constraints.

Beyond direct technical benefits, collaboration supports trust and legitimacy. Methodological choices are easier to explain to users and stakeholders when they are clearly documented, tested, and reviewed by

external experts. This is especially important for advanced methods, where there is a need to communicate uncertainty and keep the full workflow auditable.

## 3. INS Romania partnership model

INS Romania's collaboration approach is best understood as an applied case. It emerged from a concrete methodological requirement, not necessarily as a top-down strategy. A decisive driver was migration statistics. INS needed to apply small area estimation (SAE) to estimate emigrant stocks. Direct estimates at detailed territorial levels were often unstable because of small samples. This created an immediate need for stronger modelling capacity and reproducible computational workflows. In response, INS began adopting R programming language in 2012. The choice was practical: R enabled implementation of advanced methods and supported transparent, reproducible analysis suitable for documentation and quality assurance. In 2013, INS scaled this effort through internal capacity building by organizing introductory R training for beginners and delivering internal training on SAE, with a focus on operational implementation (Dobre, Adam, 2014).

The same year, INS launched the uRos (Use of R in Official Statistics) international conference. uRos was designed as a structured platform for collaboration between official statistics, academia, and other statistical institutions. It transformed ad hoc exchanges into recurring cooperation, focused on methods, tools, and reproducible practice.

Over time, these developments matured into a multi-channel collaboration model. INS combined formal partnerships with universities and research institutes, targeted project-based work (often with international partners), and community platforms such as uRos. This structure provides multiple entry points, from joint research and methodological development to applied implementation and skills transfer, while remaining committeed in production needs.

Three enabling conditions have consistently supported successful collaboration. First, partners establish a shared problem definition that translates production challenges into research questions and concrete deliverables. Second, data governance is built in from the start, including confidentiality-by-design, access arrangements, disclosure control, and clear legal boundaries that protect respondents and preserve trust. Third, collaboration includes a research-to-production pathway, ensuring that outputs are not only methodologically sound but also reproducible, well documented, and maintainable in operational environments.

This model has supported methodological co-development across several priority areas. Census modernization is one example, where robust methodological decisions are required alongside operational planning, under confidentiality and comparability constraints. Sampling methodology is another, where academic input strengthens frame diagnostics, stratification and allocation, calibration, and systematic assessment of design alternatives under realistic cost and fieldwork constraints. SAE remains central, both as the initial driver and as an ongoing area of innovation, because it requires careful model specification, diagnostics, auxiliary-data integration, and credible uncertainty communication.

Finally, open-source tooling, especially the R ecosystem, has become a key interface between academia and official statistics in Romania. It enables rapid prototyping, transparent peer review through code, and reuse of validated solutions across teams and institutions. When combined with disciplined testing and documentation, these workflows reinforce transparency and reproducibility as core elements of quality management in official statistics.

## 4. uRos as an international network and collaboration platform

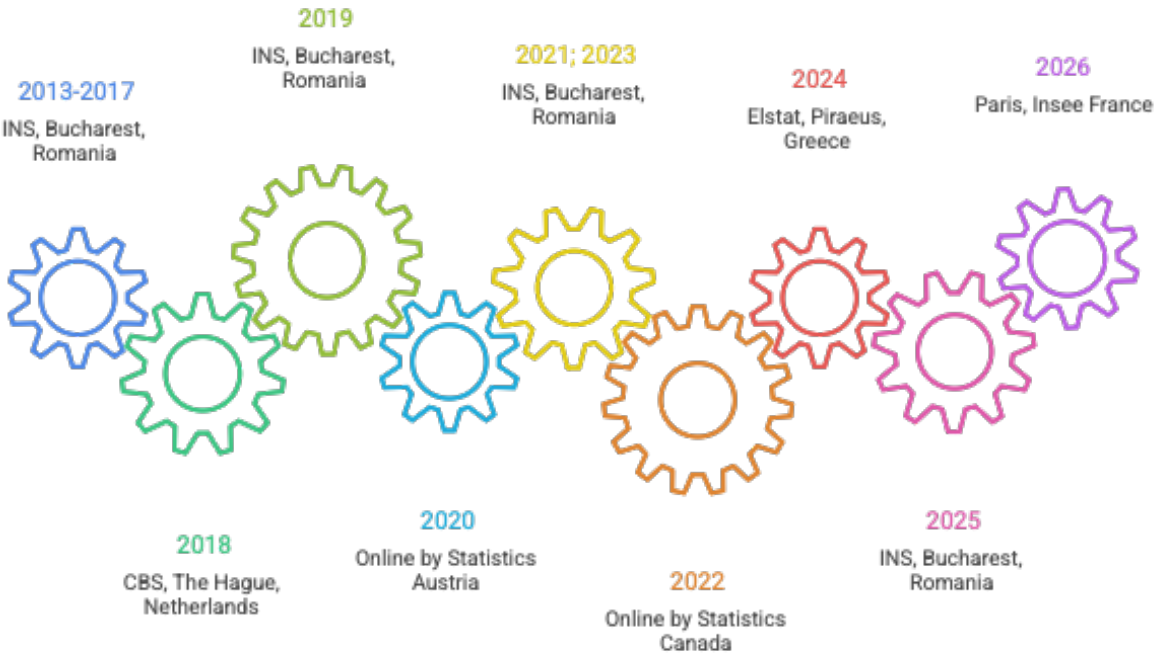
The uRos conference series is a central element of Romania's collaboration ecosystem, positioning open-source methods and reproducible computing as practical enablers for official statistics. Its format

combines tutorials, contributed sessions and workshops to move knowledge from presentation to implementable code and workflows.

The figure below presents the uRos conference series as a set of interconnected gears, suggesting a community that advances through continuity, coordination, and the transfer of knowledge from one edition to the next. Its most important feature is the rotation mechanism introduced in 2018, when uRos deliberately shifted from a predominantly Romania-based event to a clearly international hosting model. From that point onward, uRos follows an alternating pattern: one edition hosted in Romania, followed by an edition hosted in another country. This rotation was not merely a logistical choice. It was a strategic decision to institutionalize international cooperation, broaden participation, and ensure that expertise and good practices circulate across national statistical systems rather than remaining concentrated in a single host environment.

The timeline reflects this rotational scheme. After the initial editions hosted at Ecological University of Bucharest (EUB) in 2013 and INS Romania between 2014–2017, the first step of the rotation is 2018 at CBS in The Hague, signaling the expansion of the initiative beyond Romania. The following year, uRos returns to INS Bucharest, reinforcing Romania's role as a stable anchor in the series. The rotation then continues through 2020 in online mode hosted by Statistics Austria and 2022 hosted by Statistics Canada also in online mode, demonstrating that the international model remained intact even when the format had to adapt to pandemic-related constraints. Subsequent editions further illustrate the intended alternation and international reach, including 2024 hosted by Elstat in Piraeus (Greece), a return to INS Bucharest in 2025, and the next international edition hosted by Insee in Paris (France) in 2026.

**Figure 1. uRos editions**




*source: author's materials*

Overall, the schema communicates a key message: uRos is built around planned continuity through rotation. Romania provides organizational stability and institutional memory, while the international host editions expand the network, diversify perspectives, and accelerate the diffusion of methods, tools, and reproducible workflows across countries. In this way, the rotation principle functions as a governance device for collaboration, ensuring that uRos remains both rooted and genuinely international.

The institutional organizing framework is built as a partnership between official statistics producers and academic institutions, ensuring both operational relevance and methodological depth. uRos was initiated by the Romanian R Team, an enthusiastic user group with members from EUB and INS Romania. It was initially hosted by EUB, reflecting a strong academic origin focused on methodological exchange and skills development. The conference was then taken over by INS Romania, which scaled and institutionalized the initiative, expanding it beyond a national event into an international series with a clear collaborative mission. Today, uRos functions as a broad collaboration platform connecting academia with multiple national statistical institutes (NSIs). On the official-statistics side, the organizing network includes INS Romania as a stable anchor, alongside peer NSIs that have contributed to hosting and co-developing the series, such as Statistics Austria, CBS Statistics Netherlands, Elstat Greece, Insee France, and Statistics Poland (GUS). On the academic side, the partnership is reinforced by universities such as EUB, the University of Bucharest, and the Bucharest University of Economic Studies, which contribute as a sustained pipeline of students and early-career researchers, including EMOS-related engagement. Starting with the 2025 edition, uRos also became the first non-Eurostat-organised conference to receive Eurostat sponsorship, enabling EMOS students from other countries to be funded through the dedicated EMOS mobility programme to attend in person and present their work (Ciuhu, 2025). Taken together, this institutional setup provides a coherent governance model that combines continuity, international reach, and capacity building. Thereafter, uRos is positioned as a mature interface for joint methodological development and knowledge transfer between academia and the official statistics community.

Furthermore, we will summarize uRos participation over time, showing the number of participating countries, the number of attendees (useRs), and the delivery mode (in-person versus online).

**Figure 2. uRos participation**

| Year | Countries | useRs | Mode |
|---|---|---|---|
| 2013 | 1 | 12 | |
| 2014 | 6 | 24 | |
| 2015 | 15 | 50 | |
| 2016 | 15 | 29 | |
| 2017 | 20 | 43 | |
| 2018 | 32 | 100 | |
| 2019 | 30 | 105 | |
| 2020 | 16 | ≈ 300 | |
| 2021 | 28 | ≈ 200 | |
| 2022 | 24 | ≈ 250 | |
| 2023 | 13 | 54 | |
| 2024 | 22 | 92 | |
| **2025** | **26** | **90** | |

*source: author's materials*

The series starts in 2013 as a small, nationally concentrated event (1 country, 12 participants) and expands rapidly in international context, reaching 15 countries by 2015 and 20 countries by 2017. A clear step change appears in 2018–2019, when uRos consolidates as a large international conference, with 32 countries and 100 participants in 2018 and 30 countries and 105 participants in 2019. During the pandemic period, the shift to an online format is associated with substantially higher attendance (approximately 300 participants in 2020, around 200 in 2021, and around 250 in 2022), while still maintaining broad international representation. After the return to in-person editions, participation stabilizes at a strong level, with 22 countries and 92 participants in 2024 and 26 countries and 90 participants in 2025, indicating sustained international engagement even after the exceptional peak in online attendance.

## 5. International principles, governance, and lessons learned from other NSIs

In June 2025, the Conference of European Statisticians (CES) endorsed the Statistical Open Source Software (SOSS) Guiding Principles at its seventy-third plenary session in Geneva (UNECE, 2025). The principles are intended to guide both the production and the adoption of open-source software in official statistics. They position open-source development as a practical enabler of transparency, reproducibility, and peer review of statistical procedures. The Guiding Principles also translate these values into a concrete

operational agenda by setting out seven commitments: using open source by default, working in the open, improving and giving back to the community, designing generic building blocks, testing/packaging/documenting for reuse, choosing permissive licensing where possible, and actively promoting open-source developments. UNECE documentation further indicates strong international support for the updated principles through the consultation process, during which 37 countries provided feedback and endorsed the revised version.

Beyond international principles, experience from other statistical offices shows that scaling open-source adoption requires explicit governance. Recent contributions by Istat authors (D'Orazio et al, 2025) describe a structured pathway toward governance of open-source statistical tools. Related materials presented in major official statistics events emphasize the importance of standardized procedures across the full tool lifecycle, including assessment, development or acquisition, validation, documentation, dissemination, and long-term maintenance of the code. The central message is that open-source tools can deliver substantial value, but only when responsibilities and processes are clearly defined and embedded in an institutional framework. For INS Romania, this governance lesson is directly actionable. As collaborations with academia increasingly generate reusable code, scripts, and packages, INS needs governance mechanisms to ensure methodological validity and robust security and confidentiality safeguards. Clear documentation, defined maintenance ownership, and clarity on licensing and intellectual property are equally essential.

The governance and sustainability agenda is reinforced by foundational references on the role of R in official statistics. Templ and Todorov (2016) discuss R's usefulness in the day-to-day work of statistical offices and review commonly used packages for official statistics and survey methodology. Their paper illustrates practical capabilities such as producing social cohesion indicators and applying statistical disclosure control through dedicated tools. At the ecosystem level, the CRAN Task View "Official Statistics & Survey Statistics" (Templ et al, 2025) offers a curated and regularly updated map of R packages used and recommended in official statistics. It structures packages around key production tasks, aligned with the Generic Statistical Business Process Model, alongside access to official statistics data and related methodological areas. In practice, this supports standardization of internal tool stacks and faster onboarding for both NSI staff and external collaborators.

Finally, a growing frontier for collaboration between academia and official statistics is the responsible integration of machine learning (ML) into statistical production. Dumpert (2025) synthesizes recent advances and stresses that ML adoption must remain auditable and quality-assured, consistent with official statistics standards. The volume highlights several needs that fit NSI–academia cooperation particularly well. These include embedding ML in error and quality models and establishing robust protocols for performance estimation and uncertainty quantification. It also emphasizes the use of interpretable ML approaches to support transparency and accountability in production workflows.

## 6. Conclusions

The Romanian experience shows that long-term collaboration between academia and official statistics brings clear methodological and organisational benefits.

Several practical lessons stand out. Collaboration helps scale expertise. INS can combine skills with academic and international partners, learn faster, and avoid doing the same work twice. Governance is also critical, especially for access to administrative data. Cooperation needs clear rules on access, roles, and protection measures. Meeting international standards is easier when methods are peer reviewed and when INS is active in international communities, where approaches can be compared and clarified. Finally, open-source work can improve quality, but only if it is well managed. Simple procedures for review, testing, documentation, maintenance, and licensing help keep code reliable and support the move from prototypes to production.

## References

Ciuhu, A.-M. (2025). Promoting Open-Source Solutions in Official Statistics: the uRos Conference Series. Poster presented at 11th EMOS workshop, Bergamo, 20-21 November 2025, DOI: 10.13140/RG.2.2.11228.24961

Dobre, A.-M., Adam R. (2014). The Progress of R in Romanian Official Statistics, Romanian Statistical Review, 2/2014, pp. 45-54, https://www.revistadestatistica.ro/wp-content/uploads/2014/07/RRS_2_2014_a05.pdf

D'Orazio, M., Di Zio, M., Scanu, M., & Luzi, O. (2025). The journey towards a governance of open source statistical tools at Istat. Statistical Journal of the IAOS. https://doi.org/10.1177/18747655251390274

Dumpert, F. (Ed.) (2025). Foundations and Advances of Machine Learning in Official Statistics. Society, Environment and Statistics. Springer, Cham. https://doi.org/10.1007/978-3-032-10004-7

Templ, M., & Todorov, V. (2016). The Software Environment R for Official Statistics and Survey Methodology. Austrian Journal of Statistics, 45(1), 97-124. https://doi.org/10.17713/ajs.v45i1.100

Templ, M., Kowarik, A., & Schoch, T. (2025). CRAN Task View: Official Statistics & Survey Statistics (Version 2025-03-11). https://CRAN.R-project.org/view=OfficialStatistics

United Nations Economic Commission for Europe (UNECE). (2025). Statistical Open Source Software Guiding Principles (endorsed by CES at the Seventy-third plenary session, Geneva, 16-18 June 2025). https://unece.github.io/OSS/principles.html

United Nations Economic Commission for Europe (UNECE). (2025). Results of the consultation on the principles on open source software in official statistics (ECE/CES/2025/9/Add.1).

uRos conference series, https://r-project.ro/conferences.html