Funded by the European Union

COARA Boost

1958 COGNOSCERE EST MUTARE

Advancing Research Assessment in Serbia

Evidence and Policy Recommendations from the ARISE Pilot Project

Isidora Beraha
Jovan Zubović
Sonja Đuričin
Jelena Williams
Lazar Živković
Petar Vićentić

Institute of Economic Sciences

Belgrade, September 2026

Funded by the European Union

COARA Boost

1958 COGNOSCERE EST MUTARE

Table of Contents

Foreword 4

Introduction 5

1. Research Assessment in Serbia: Current Framework and Institutional Context 7

1.1. The Serbian Research Assessment Framework 8

1.2. Current Assessment Practices and Key Limitations 12

1.3. Research Assessment Practices at the Institute of Economic Sciences 21

2. Global and European Framework for Research Assessment Reform 29

2.1. From Metrics to Responsible Research Assessment: DORA and the Leiden Manifesto 29

2.2. CoARA and the European Research Assessment Reform 30

2.3. European Pathways to Responsible Assessment: Lessons from France and Romania 32

2.4. Bridging the Gap: Lessons from Europe and Opportunities for Reform in Serbia 38

3. From Researchers' Perspectives to the ARISE Model 42

3.1. Beyond Metrics: Broadening the Understanding of Research Quality and Contribution 42

3.2. Making the Full Range of Research Contributions Visible 43

3.3. Context Matters: Research Profiles, Career Stages and Institutional Diversity ..44

3.4. Research Assessment as a Developmental and Learning Process 45

3.5. Combining Narrative Evidence, Quantitative Indicators and External Peer Review 46

3.6. Translating Evidence into the ARISE Model Design 47

4. The ARISE Research Assessment Model 50

5. Pilot Implementation of the ARISE Research Assessment Model at IES 60

5.1. What the External Peer Review Revealed 60

5.2. From Individual and Team Strengths to Institutional Capacity 61

5.3. Researchers' Experience and Perceptions of the ARISE Pilot 62

6. Extending the Evidence Base Beyond IES 66

6.1. A Shared Demand for Broader Research Assessment 66

6.2. One Framework, Different Disciplinary Pathways 67

6.3. What the Evidence Means for the ARISE Model 69

6.4. From Institutional Pilot to Broader Adoption 70

7. From Evidence to Dialogue: The ARISE Forum and Pathways for Reform..... 72

7.1. The ARISE Forum: Dissemination and Stakeholder Dialogue.....72

7.2. Round Table: From the ARISE Experience to Pathways for Reform 73

8. Policy Implications and Recommendations: Towards Responsible Research Assessment in Serbia..... 75

Conclusion..... 83

References..... 85

Funded by
the European Union

Funded by the European Union

CoARA Boost

1958 COGNOSCERE EST MUTARE

Foreword

This policy document was developed within the project *Advancing Research Assessment in Serbia – A Pilot Project for the Institute of Economic Sciences (ARISE)*, financed through the Horizon Europe CoARA Boost programme, to bring together the evidence, experience and lessons generated through the project and translate them into a basis for policy discussion and continued dialogue on the future of research assessment in Serbia.

The document is the result of an iterative process of evidence building, model development, piloting and dialogue with the research community. The process began with an analysis of the Serbian research assessment system and selected European approaches, providing the institutional and comparative basis for identifying the need and scope for change. Evidence was then collected directly from the IES research community through an Institute-wide survey and focus groups, while the multidisciplinary ARISE Advisory Board provided continuous external reflection throughout the model-development process. These inputs informed the design of the ARISE qualitative research assessment model, which was subsequently piloted across all six IES research departments through self-assessment and external peer review. The experience of implementation was then examined through external evaluators' findings and feedback from IES researchers, allowing both the model and the assessment process to be critically considered. The evidence base was further extended beyond IES through a cross-disciplinary survey involving researchers from all broad scientific fields in Serbia. Finally, the ARISE Forum brought the project experience, international perspectives and the wider research community into dialogue on the conditions and possible pathways for broader reform. Each stage therefore built on and informed the previous one, progressively moving ARISE from analysis and consultation, through model development and testing, towards evidence-based policy implications and recommendations.

The policy recommendations presented in this document are the outcome of this cumulative process rather than conclusions drawn from a single pilot exercise. They bring together evidence from practical implementation with researchers' perspectives, multidisciplinary expert reflection and broader stakeholder dialogue. The purpose of the document is not to propose a finished national assessment model, but to provide an evidence-based foundation for the next stage of reform: identifying what can be learned from the ARISE experience, what requires further testing and adaptation, and which methodological, institutional and governance conditions will be necessary for moving towards a broader and sustainable system of responsible research assessment in Serbia.

Funding Acknowledgement

This policy document was developed within the project *Advancing Research Assessment in Serbia – A Pilot Project for the Institute of Economic Sciences (ARISE)*, supported through the Horizon Europe CoARA Boost Programme (Cascade Funding).

Introduction

Research assessment has been undergoing significant reform internationally, with particularly intensive developments across Europe. Initiatives such as the San Francisco Declaration on Research Assessment (DORA), the Leiden Manifesto and, more recently, the Coalition for Advancing Research Assessment (CoARA) have contributed to challenging excessive reliance on publication and journal-based metrics and to promoting broader recognition of research contributions. Within the European reform context, increasing emphasis is placed on qualitative assessment in which peer review plays a central role, supported by the responsible use of quantitative indicators. Greater attention is also given to recognising the diversity of research contributions and career paths, as well as their societal impact, and to the developmental role of assessment in supporting learning and improvement.

It was within this broader reform context that the project "Advancing Research Assessment in Serbia – A Pilot Project for the Institute of Economic Sciences (ARISE)" was implemented by the Institute of Economic Sciences (IES), Belgrade, from September 2025 to August 2026, within the Horizon Europe CoARA Boost Cascade Funding Programme. The project was conceived as an institutional pilot aimed at advancing research assessment practices in Serbia through the development and practical testing of a qualitative, peer-review-based approach to research assessment.

The main objective of ARISE was to develop and pilot a qualitative, peer-review-based research assessment model for research departments at IES and subsequently promote its broader adoption, initially among social science research institutes in Serbia and, ultimately, across the wider Serbian research community. The project sought to translate the principles of responsible research assessment into an operational framework that could be applied and tested in an actual institutional setting, while taking into account the characteristics of the Serbian research system and the specific context of social science research.

ARISE was implemented through a sequence of interconnected activities. The project began with an analysis of existing research assessment practices in Serbia and at the Institute of Economic Sciences, complemented by a comparative analysis of research assessment practices in selected European countries, followed by research into IES researchers' perceptions of existing assessment practices and participatory discussions on research values and appropriate assessment approaches. This evidence informed the development of the ARISE Research Assessment Model, which was subsequently piloted across all six IES research departments. The pilot involved narrative self-assessment by research departments, supporting evidence, external qualitative peer review by international evaluators, and interviews with IES management, heads of research departments and researchers at different career stages. The experience of the pilot was further examined through researcher feedback, while a cross-disciplinary survey involving researchers from four broad scientific fields explored the relevance of the ARISE approach beyond the immediate IES context. In July 2026, the ARISE Forum brought together representatives of the research community and relevant stakeholders, initiating a broader discussion on the lessons from the ARISE pilot, the potential for wider application of the model, and possible directions for the further development of research assessment in Serbia. Together with other stakeholder dialogue and dissemination activities, the Forum extended the discussion beyond the institutional pilot, engaging the wider research community in considering the implications of the ARISE experience and possible pathways for research assessment reform in Serbia.

The ARISE Research Assessment Model represents the project's central methodological output. It established research departments as the unit of assessment and introduced a multidimensional framework comprising six broad assessment dimensions, narrative self-assessment, qualitative peer review and clearly defined assessment procedures. Expert judgement is at the core of the assessment process, while quantitative and bibliometric indicators are used responsibly as contextual and supporting evidence rather than as substitutes for qualitative assessment. The model also adopts a developmental and learning-oriented approach: assessment is intended to support reflection, institutional learning and strategic development rather than to rank research departments or serve punitive purposes. These principles were operationalised and tested through a complete assessment cycle at IES.

This policy document brings together the evidence, findings and lessons generated throughout the ARISE project. Rather than proposing a ready-made model for national implementation, it consolidates the experience gained through the development and pilot implementation of the ARISE approach and the evidence collected within and beyond IES. On this basis, it identifies implications for the further development of qualitative, peer-review-based research assessment in Serbia and formulates policy recommendations for its gradual adaptation and broader application.

1. Research Assessment in Serbia: Current Framework and Institutional Context

KEY INSIGHTS

- **Research assessment in Serbia remains predominantly quantitative and indicator-driven**, with limited recognition of diverse research contributions and institutional context.
- **European reform is moving towards qualitative, peer-review-based assessment**, with metrics used as supporting evidence rather than substitutes for expert judgement.
- **France and Romania illustrate different pathways towards more qualitative and context-sensitive assessment**, combining self-assessment, peer review and attention to institutional development.
- **Reform in Serbia requires a broader change in assessment logic** — from counting outputs to understanding contributions, and from administrative control to development.

Research assessment plays an important role in shaping research funding, academic careers, institutional development and scientific priorities. Traditionally dominated by quantitative and bibliometric indicators, assessment practices are increasingly being reconsidered in light of their limitations and unintended effects. Contemporary approaches emphasise qualitative peer review, the responsible and contextualised use of quantitative indicators, and recognition of diverse research contributions and career paths. This shift reflects a broader understanding of research assessment as a strategic and developmental instrument that supports institutional learning and improvement rather than primarily measuring outputs and ranking performance.

In Serbia, research assessment remains predominantly quantitative and indicator-driven. At the institutional level, the formal assessment framework relies primarily on citation-based performance and the amount of funding secured by research institutes, while the assessment and career progression of individual researchers are strongly grounded in publication categories and bibliometric criteria. Although the regulatory framework is detailed and formally developed, it provides limited space for systematic consideration of research quality in context, institutional mission and strategy, the research environment, societal relevance and impact, or the diversity of contributions through which researchers and research organisations contribute to society.

This points to a key policy challenge for Serbia: how to move beyond a predominantly metric-driven approach towards a more comprehensive research assessment system while retaining the responsible use of quantitative indicators. Quantitative indicators can provide useful and comparable information on research performance, but they cannot by themselves capture the quality, relevance and diversity of research contributions. They may also create incentives to prioritise the accumulation of countable outputs over activities whose value is less easily captured through bibliometric indicators, including mentoring, collaboration, institutional capacity building, interdisciplinary research, policy engagement and other forms of societal contribution. The issue is therefore not whether quantitative indicators should be abandoned, but how they should be repositioned within a more comprehensive assessment framework in which they support, rather than substitute for, informed expert judgement.

The baseline analysis conducted within ARISE examined the Serbian research system, the regulatory framework governing research assessment, and assessment practices at the Institute of Economic Sciences. It provided the institutional and policy context for identifying the

limitations of existing approaches and the opportunities for introducing more qualitative, peer-review-based and context-sensitive assessment practices.

The analysis shows that the limitations of the Serbian system extend beyond the choice of individual indicators and concern the broader logic and function of research assessment. The system is normatively developed but substantively indicator-driven, administrative and context-blind. As such, it provides limited opportunities for contextualised assessment, constructive feedback, institutional learning and strategic development. These structural limitations point to the need for a transition towards a more qualitative, peer-review-based and context-sensitive approach, while retaining quantitative indicators as supporting evidence where they provide meaningful information on research performance.

This chapter examines the Serbian research assessment framework and existing assessment practices, including those at the Institute of Economic Sciences. It identifies key limitations and institutional challenges that informed the development of the ARISE Research Assessment Model.

1.1. The Serbian Research Assessment Framework

The scientific research system in the Republic of Serbia is shaped by the strategic framework established by the Strategy for Scientific and Technological Development of the Republic of Serbia 2021–2025 (The Power of Knowledge)[1] and regulated by the Law on Science and Research[2]. The Strategy establishes national priorities and long-term objectives for scientific and technological development and is aligned with Serbia's broader economic and societal strategies, including the Smart Specialisation Strategy (2020–2027). The Law provides the legal framework for organising and developing scientific research activities, covering the planning and implementation of the public interest in science, quality assurance mechanisms, the legal status and governance of research institutes, procedures for acquiring scientific titles, institutional and programme-based funding, and the maintenance of official research records.

System oversight, accreditation, and quality assurance are performed by the National Council for Science and Technological Development, the Accreditation Board for Research Organisations, the Commission for the Promotion to Scientific Titles, and the Scientific Field Councils. Scientific research activities are carried out under multiple ministries: the Ministry of Science, Technological Development and Innovation (MSTDI), which oversees 67 accredited research institutes; the Ministry of Education, which governs research at university faculties; as well as several institutes operating under the Ministry of Health, the Ministry of Defence, and the Serbian Academy of Sciences and Arts (SASA). The National Council is the highest expert and advisory body in this domain. It consists of a president and 13 members representing six scientific fields (natural and mathematical sciences, technical and technological sciences, medical sciences, biotechnical sciences, social sciences, and humanities) and is responsible for strategic planning,

[1] Government of the Republic of Serbia. (2021). *Strategy of Scientific and Technological Development of the Republic of Serbia for the period 2021–2025: "The Power of Knowledge"*. Belgrade: Ministry of Education, Science and Technological Development.
[2] Republic of Serbia. (2019). Zakon o nauci i istraživanjima [Law on Science and Research]. Službeni glasnik Republike Srbije, br. 49/2019.

policy monitoring and preparing the annual Report on the State of Science, the latest of which (for 2024) was adopted in September 2025[3].

Research Funding and Key Institutions

Research funding in Serbia is ensured through institutional financing and programme-based support provided primarily by the Science Fund and the Innovation Fund, both of which operate as independent legal entities cooperating with MSTDI.

According to the available data for 2024, the total budget allocated for institutional funding of research in Serbia amounted to approximately 22 billion dinars. Within this framework, funding for the DMT2 category reached 881.77 million dinars, representing only 4% of the institutional funding budget. Although the nominal amount allocated to DMT2 increased gradually between 2020 and 2024, its share within the institutional funding program consistently declined, remaining below 5% throughout the observed period. When compared to 2018 and 2019, years in which DMT2 accounted for 7.59% and 6.41% of institutional funding, respectively, the downward trend becomes even more pronounced. Adjusted for inflation, the real decrease in DMT2 funding would be even greater.

Since its establishment in 2019, the Science Fund has funded 553 projects through 13 programmes, involving 3,400 researchers from 145 scientific research organisations. In 2024, it implemented six programmes at different stages of execution: Ideas (third year), Identities (second year), the Green Science–Industry Collaboration Programme (second year), Prisma (first year), Promis 2023 (first year), and the Proof of Concept Programme (first year). The Joint Research Projects Programme for Cooperation with the Serbian Scientific Diaspora underwent a two-stage evaluation procedure, resulting in a published list of approved projects. Preparations were also undertaken for upcoming calls within the Diaspora Collaboration Programme, the Ideas 2024 Programme, and four new programmes of which two in artificial intelligence, the second phase of the Green Programme, and the China–Serbia Joint Research Programme. Throughout 2024, the Science Fund monitored the implementation of 320 active projects.

In parallel, for the Serbian Academy of Sciences and Arts (77 million dinars) and Matica Srpska (19 million dinars), the two most prominent national scientific institutions, a total of 96 million dinars was allocated in 2024.

The Innovation Fund is the only state organisation dedicated primarily to supporting innovation activities and managing financial instruments that foster innovation. In 2024, MSTDI allocated approximately 1.62 billion dinars to the Fund. A total of €37.7 million was distributed across several programmes, including Smart Start, Early Development, Katapult (Matching and Co-investment Grants), Innovation Co-financing, Science–Industry Collaboration, GovTech, Innovation Vouchers, Technology Transfer, and Serbia Ventures AI.

[3] National Council for Science and Technological Development. (2025). *Report on the state of science in 2024, with proposals and recommendations for the following year*. Republic of Serbia.

Public Investment and R&D Performance

Public investment in science has remained below 0.40% of GDP over the past three years (0.39% in 2022; 0.40% in 2023 and 2024). Total R&D expenditure amounted to 0.95% of GDP in 2023 and 0.94% in 2024. These levels are significantly below the EU and global averages (2.63%), and lower than in several regional peers such as Slovenia (2.13%) and Croatia (1.24%).

According to the Global Innovation Index (GII) 2024, Serbia ranks 52nd out of 133 countries, seventh among upper middle-income economies and classified as a moderate innovator. WIPO reports that Serbia performs above expectations relative to its economic development, indicating more efficient innovation output than investment levels would suggest. The European Innovation Scoreboard 2024 categorises Serbia as an Emerging Innovator, with performance at 62.8% of the EU average (30th of 39 countries).

Human Resources, R&D Expenditure and Research Output

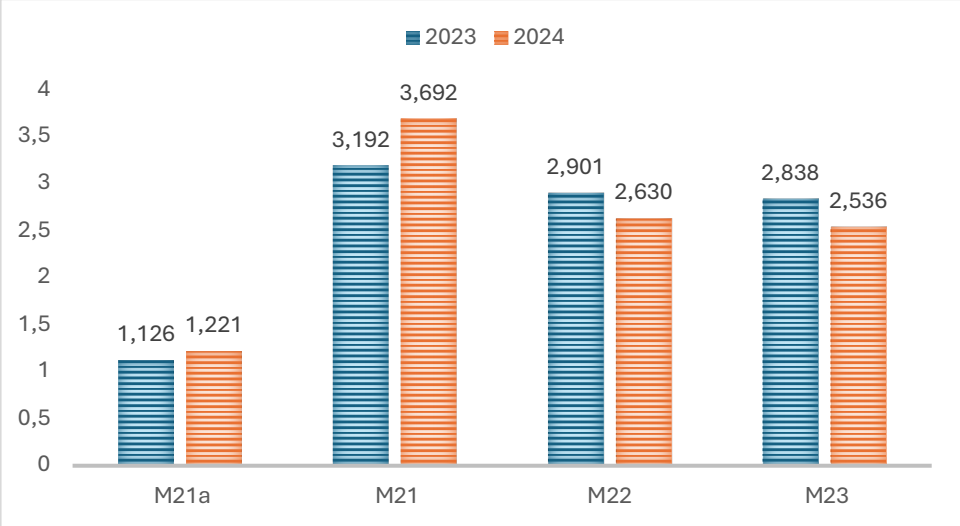
Data from the Statistical Office show that in 2024 Serbia had 30,810 R&D personnel, including 19,323 researchers (52.7% women), 3,130 professional associates, 5,668 technical staff and 2,689 other personnel. Total R&D expenditure exceeded 90 billion dinars (0.94% of GDP). Engineering and technological sciences employed the largest number of researchers, while most R&D expenditure was concentrated in the natural sciences. The largest gap between spending per capita and researcher numbers was observed in medical and health sciences.

The total number of scientific research outputs in 2024 was 30,852, representing a decrease of 4,510 compared to 2023. The majority of outputs were scientific papers. Category M20 (international journals) accounted for the largest share, followed by M30 (international conference proceedings), M60 (national conferences) and M50 (national journals). Within category M10, the largest number of outputs (774) were in subcategory M14 (chapters in internationally relevant monographs). Distribution of the number of papers published in 2023 and 2024 by category is presented below.

Category	2023	2024
M10 - Monographs, monographic studies, thematic edited volumes, and lexicographic and cartographic publications of international significance	1,423	1,265
M20 - Articles published in internationally significant scientific journals	11,735	11,215
M30 - Proceedings of international scientific conferences	10,768	10,209
M40 - Monographs of national significance	2,155	1,677
M50 - Articles published in nationally significant journals	3,681	2,321
M60 - Proceedings of national scientific conferences and critical editions of sources	4,032	2,683

Within category M20, researchers produced 3,692 outputs in M21 (top-tier international journals, ranked within the top 30% in their field), 1,223 in M21a (international journals of exceptional quality, top 10%), 2,630 in M22 (renowned international journals, ranked between the top 30% and 60%), 2,536 in M23 (international journals listed in JCR but ranked below the top 60%; in SSH fields this also includes WoS journals without an IF and SJR Q1 journals), and

887 in M24 (national journals of international significance; top 5% in the National Index, and in SSH also SJR Q2–Q3). Altogether, the 10,079 outputs in categories M21–M23 broadly correspond to Serbia's 2024 Scopus-indexed publications. In nationally relevant journals (M50), 2,321 outputs were published, with subcategory M51 dominating this group. Distribution of the number of papers published in 2023 and 2024 in top-tier journals by category is presented in the following graph.



A total of 636 doctoral dissertations (M70) were defended. Scientific articles remain the primary form of output, with 43.8% of all results published in peer-reviewed venues (M20 + M50). Outputs with potential economic relevance remain limited: patents and technical solutions (M80 and M90) accounted for only 1.2% of all results (239 technical solutions and 138 patents or similar outputs), lower than in 2023.

According to the ranking of Eastern European countries by the number of scientific articles published in 2024[4], the h-index of Serbian science was 384 (compared to 350 in 2023), which remains lower than the h-index of most countries in the region: Hungary (630), Slovenia (439), Romania (454) and Croatia (413).

Despite improvements in publication productivity, i.e., 1.7% growth relative to 2023 and 1,590 Scopus-indexed publications per million inhabitants (comparable to developed EU countries) citation rates remain below the European average. Serbia has the lowest GDP per capita and the lowest number of registered patents among the regional comparator countries, and the number of researchers per million inhabitants is roughly half that of Slovenia, though significantly higher than in Romania.

Overall, the limited level of investment in science and technological development constrains the advancement of Serbia's research system and its capacity to contribute to technological progress. Quantitative indicators continue to dominate assessment practices. However, the analysis above shows that this model has reached its limits. The persistent emphasis on increasing the number of publications no longer produces positive outcomes, as evidenced by the recent decline in scientific output. The reliance on quantitative metrics has

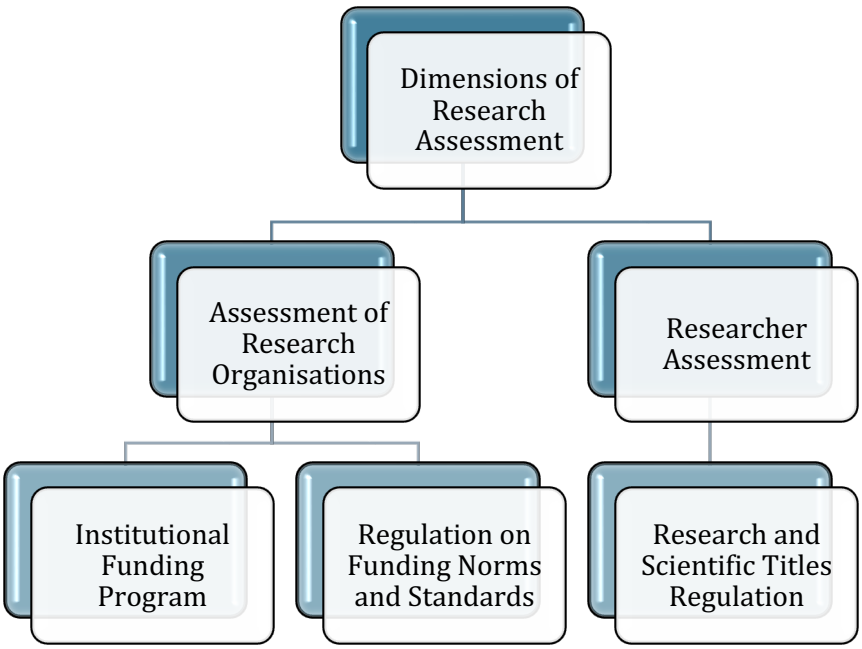
[4] https://www.scimagojr.com/countryrank.php

become exhausted as a driver of research quality and productivity, underscoring the need for a more balanced and qualitative approach.

This systemic overview provides the context for understanding current research assessment practices in Serbia, their institutional foundations, and the key challenges the ARISE project aims to address.

1.2. Current Assessment Practices and Key Limitations

Serbia's national framework for research assessment is based on the Law on Science and Research, which regulates the adoption of the Strategy for Scientific and Technological Development and defines institutional responsibilities and public policy instruments within the research and innovation system. The evaluation rules are implemented through subordinate acts adopted by the Government and the relevant ministry. The analysis is based on two interconnected dimensions of performance evaluation. The institutional evaluation of scientific research organizations is governed by the Program for Institutional Funding of Accredited Institutes and the Regulation on Norms and Standards for the Allocation of Funds to Accredited Scientific Research Organizations. The evaluation of individual researchers is regulated by the Rulebook on the Acquisition of Research and Scientific Titles.



The systematization of assessment at both the institutional and individual levels provide a clear overview of the criteria, procedures, and evaluation periodicity, and establishes a basis for examining the degree of systemic coherence among funding, assessment, and career progression.

Evaluation of Scientific Research Organizations

1. **The Program for Institutional Funding of Accredited Institutes** (**Program**) represents subordinate legislation adopted by the relevant ministry in accordance with the Law on Science and Research. The Program outlines the priorities and indicators for evaluating the

performance of accredited research institutes, providing a basis for monitoring and allocating public funds based on achieved results.

The Program for Institutional Funding guides the development of the research and innovation system by strengthening institutes and teams, enhancing their competitiveness and international visibility, and promoting more efficient use of resources, collaboration with industry, and a focus on societal priorities. Institutes are required to develop five-year scientific research programs and, through their activities, are expected to foster the development of new research fields and capacities, open science and transparency, gender equality and research integrity, infrastructure maintenance, international cooperation, and commercialization. National institutes are required to develop ten-year scientific research programs and, through their activities, are expected to promote the development of a fund of knowledge, shared use of resources, and support for evidence-based policymaking. Institutes conduct basic and applied research in natural sciences and mathematics, medical sciences, humanities, and social sciences, as well as technological development in engineering and technology, and agricultural and biotechnical sciences.

The Program defines two quantitative indicators for evaluating the performance of accredited research institutes: the total number of citations and the overall amount of funding received from all sources. These metrics are intended to be used to calculate a composite Institutional Performance Quality Index, which should serve as the basis for ranking research institutes at the national level.

- For scientific research institutes, citations account for 90% of the score, while total financial resources account for 10%.
- For research-and-development institutes, total financial resources account for 70%, while citations account for 30%.

Total citations represent the aggregate number of citations for all publications in which at least one researcher employed at the institute is listed as an author, with the institute's affiliation explicitly indicated. Total citations for the current year are calculated based on publications issued over the previous ten years, using one of the recognized citation databases (Scopus, Web of Science, or Google Scholar). The institute selects one database in the first evaluation year and must use the same source throughout the entire evaluation period. Total funding includes all financial resources the institute generates from all legally recognized sources. For institutes in the humanities, the quality assessment of research performance is based on the parameters reported in their annual reports during the first five years of this program's implementation. In parallel, the Ministry, through its expert bodies, evaluates additional indicators to develop appropriate analytical frameworks for the humanities. Citation impact is excluded from the assessment if this is more favorable for the institute. In evaluating research outputs in the social sciences and humanities, special consideration is given to the national, historical, cultural, and societal importance of the results, as well as their regional or international relevance and their contribution to preserving the country's cultural and scholarly heritage. In the social sciences, additional attention is paid to the extent to which research results contribute to addressing societal challenges and meeting national policy needs (through outputs categorized as M120 – Public Policy Analysis).

If the institute's Quality Index exceeds the average value of the previous four-year period by 20%, the institute becomes eligible for a 5% increase in material research expenditures, or, in exceptional cases, for a higher increase if permitted by the Programme's allocated funds and subject to the budget's liquidity constraints. Data for the evaluation process have been collected since 2024. Based on these data, a baseline Institutional Quality Index is established. For each subsequent year, the provisions of the Programme are applied, always using the most recent four-

year period as the reference. Furthermore, if the indicators (citations and total financial resources) show an increase compared to the previous year, the institute becomes eligible for an increase in funding for material research expenditures. Any increase to the institute's budget requires a positive opinion from the National Council for Science and Technological Development.

The Program clearly defines the criteria for evaluating the work of accredited research institutes and specifies two quantitative indicators for this purpose: the total number of citations and the overall amount of funding received from all sources. These measures are intended to form a composite Institutional Performance Quality Index, which should serve as the basis for ranking institutes at the national level. However, despite the clarity and precision of the framework set out in the Program, it has not been implemented in practice. The evaluation of research institutes does not follow these prescribed indicators, and a systematic assessment based on the defined criteria has not taken root in the national research system.

Research institutes prepare five-year activity reports containing key information on staff, research outputs, and revenues, including all elements required for calculating the index and assessing compliance with the institutional plan. These reports are submitted to the Ministry through a dedicated electronic form. The reports are first reviewed by the relevant Scientific Boards, followed by the opinion of the National Council for Science and Technological Development, and the final decision is issued by the Minister. Institutes also prepare an annual work plan based on their long-term activity program. The structure of the annual plan and the institute's activity report, as well as the submission deadlines, are governed by the Regulation on Norms and Standards for the Allocation of Funds to Accredited Research Performing Organizations. The evaluation of research performance for researchers holding research and scientific titles at accredited faculties is conducted on the basis of their annual research activity reports, submitted in accordance with the same Regulation on Norms and Standards for the Allocation of Funds to Accredited Research Performing Organizations.

In 2020, as part of the broader reform of Serbia's research system, the Ministry of Education, Science and Technological Development launched a nationwide initiative aimed at assessing the institutional capacities of public research institutes. For the first time, the Law on Science and Research introduced institutional funding as a core financing mechanism, requiring institutes to establish a clear understanding of their internal capacities, performance, funding sources, partnerships, and strategic positioning. As a prerequisite for accessing institutional funding, each institute was asked to conduct a structured self-assessment using a standardized questionnaire prepared by the Ministry. The questionnaire, provided in both Serbian and English, included methodological guidance and required institutes to evaluate their current status, identify key challenges, and articulate the conditions necessary for improvement. Institutes were instructed to appoint a Self-Assessment Committee of three to ten members to coordinate the process and submit the completed materials. As part of this process, institutes were also asked to declare whether they wished to undergo an independent external evaluation in the next phase. Within the SAIGE project for the transformation of scientific institutes implemented by the Ministry with the support of the World Bank and the European Commission the initial phase consisted precisely of such independent evaluations, conducted by the consulting firm Inomer, whose reports later served as the basis for drafting institutional transformation plans. This self-evaluation exercise was therefore designed not only to support national reform but also to help Serbian research institutes strengthen their strategic orientation and position themselves as credible partners within the European Research Area.

2. **Regulation on Norms and Standards for the Allocation of Funds to Accredited Research Performing Organisations (the Regulation)** is a piece of subordinate legislation adopted by the Government in accordance with the Law on Science and Research. The Regulation sets

out the norms and standards for allocating funds to accredited research performing organisations.

Norms and standards are general rules and guidelines that govern the allocation of funds for researchers' salaries, overhead costs, and material research expenditures. These funds are disbursed based on the annual work plan of the research-performing organization and the agreement between the organization and the ministry responsible for research activities. The annual work plan includes a description of the institution's activities with a SWOT analysis for the upcoming year, a research plan with expected results, a plan for staff and infrastructure development, a plan for attracting capital investments, mobility and collaboration activities (with industry, academia, and through projects), and dissemination and teaching activities. It also provides an overview of anticipated changes in the number and status of researchers. When submitting the annual activity report, the director or dean proposes one relevant Scientific Board for evaluation. The competent Scientific Board reviews the report and submits its opinion to the Ministry. The annual activity report of a research-performing organization includes basic information about the institution and its facilities, a detailed structure of researchers by title and type of engagement, research results, participation in national and international projects, and cooperation with academic and industry partners. It also contains the indicators required for calculating the Institutional Quality Index: the selected citation database (Scopus/ Web of Science/ Google Scholar), total citations over the past ten years, up to five most-cited outputs and researchers, and all revenue sources classified by prescribed categories. The reporting system functions on a binary "yes–no" evaluation of institutional reports. To date, no research institution has ever received a "no," indicating that the process does not meaningfully differentiate performance or identify shortcomings. As a result, the entire reporting mechanism operates largely as a formality, without generating actionable feedback or contributing to quality assurance.

a) **Funding of researchers' scientific and research activities**[5] in accredited organizations is based on coefficients assigned to scientific, research, and professional titles as defined by the Regulation. These coefficients are multiplied by the base value, which is determined annually by the minister responsible for science, following the prior opinion of the Ministry of Finance. The excellence bonus is determined by applying a corrective coefficient, the calculation of which is defined in the Guidelines for Ranking Researchers in Scientific Titles for Determining the Excellence Corrective Coefficient, which form an integral part of this Regulation. Funds for financing scientific and research activities are allocated to accredited research-performing organizations for researchers in scientific, research, and professional titles employed on a full-time basis, including prescribed allowances (transport, years of service, and a 30% management allowance for directors). The allocation, disbursement, and monitoring of the designated use of these funds are regulated by the annual agreement concluded between the Ministry and the organization.

b) **The Guidelines for Ranking Researchers in Scientific Titles for Determining the Excellence Corrective Coefficient**[6] (Guidelines for Ranking Researchers) define:

- Data sources
    - For researcher citations: citations of published articles indexed in the Web of Science.

[5] Official Gazette of the Republic of Serbia", No. 96/2023
[6] Official Gazette of the Republic of Serbia", Nos. 90/2019 of 20 December 2019; 96/2023 of 2 November 2023; 110/2023 of 8 December 2023; and 16/2024 of 1 March 2024.

- For international journals: the Web of Science is used (with DOAJ/Scindeks, and the categories defined by the Scientific Boards and by eNauka applied only in the social sciences and humanities). In the Web of Science, only Articles, Reviews, and Proceedings Papers are evaluated, except in the humanities, where all indexed outputs are considered. For journals with publishers based in Serbia, the annual List of Categorized Scientific Journals published by the competent ministry is used.
- For monographs: eNauka.

- Fields of science and ranking groups — A researcher is ranked within the group corresponding to the field in which they obtained their current scientific title:

Fields of Science	Researcher Ranking Groups
Natural sciences, Medical and health sciences	Physics, Chemistry, Biology, Mathematics, Computer Science and Mechanics, Medical Sciences, Earth and Space Sciences
Engineering and technology, Agricultural sciences	Biotechnology and Agriculture, Electronics, Telecommunications and Information Technologies, Energy, Mining and Energy Efficiency, Materials and Chemical Technologies, Mechanical Engineering and Industrial Software, Transport, Urban Planning and Civil Engineering, and Water, Soil and Air Management and Protection
Social sciences	Law, Economics and Political Sciences; Philosophy, Psychology, Pedagogy and Sociology
Humanities	History, Archaeology and Ethnology; Language and Literature

- Scope of researchers subject to ranking

The excellence corrective coefficient is applied to up to 10% of researchers employed exclusively in positions corresponding to scientific titles. The ranking uses unified lists created for each scientific title within each of the four scientific field groups (four field groups across three scientific titles). A researcher promoted to a higher scientific title retains the previously awarded excellence corrective coefficient until the next ranking cycle.

- Types of research outputs included in the scoring and aggregate calculation for ranking researchers within scientific field groups

Research outputs are classified into categories A and B for ranking purposes within each scientific field group. In the humanities, an additional category C is also used:

Scientific field group	Category A	Category B	Category C	Notes
Natural sciences, Medical and health sciences	M21a–M23	Citations (2013–2022) of articles indexed in the Web of Science	—	
Engineering and technology,	M21a–M23	Citations (2013–2022) of articles	—	

Agricultural sciences		indexed in the Web of Science		
Social sciences	M21a–M23	M24 + citations (2013–2022) Web of Science	—	
Humanities	M21a–M24	M11, M12, M13, M14, M41, M42	M43, M44, M51	For categories B and C verification by the Scientific Boards and entry into eNauka.

During data processing, 24 scoring lists are generated for categories A and B, covering 4 scientific field groups across 3 scientific titles, along with 3 C-category lists in the humanities.

- Scoring scale for research outputs
  - The K–Value of scientific results from Annex 2 of the Regulation on the Appointment to Research and Scientific Titles is applied.
  - Each citation, regardless of type, is assigned 1 point on the citation list.

- Normalization of contribution in co-authored outputs
  - The rules for normalizing the number of co-authored outputs from Annex 2 of the Regulation on the Appointment to Research and Scientific Titles are applied.
  - Normalization is applied to output types A, B, and C.

- Separate lists for collaborative research

A unified collaboration list is created, including all researchers from all scientific field groups who have outputs produced in collaborations with more than 100 authors. For ranking, only the total points earned from these outputs are considered, based on output types A and B, for the period 2013–2022, and applying the rules from these Guidelines relevant to the Natural Sciences and Medical and Health Sciences.

- Time period for scoring
  - Research outputs: All researchers holding scientific titles are ranked based on results from the past 5 years, except Principal Research Fellows, who are ranked based on results from the past 10 years. If a researcher has taken leave longer than one year during the relevant period, the time window is extended to the past 7 years (or 12 years for a Principal Research Fellow). For the extended period, the points earned are multiplied by ¾.
  - Citations: Citation counts are based on all Web of Science–indexed articles authored by the researcher and cited during 2013–2022. The exception for the title of Research Associate is explicitly defined in these Guidelines.

- Ranking of researchers in scientific titles within scientific field groups

Researchers are ranked within their scientific titles and scientific field groups according to the total number of points earned from the A and B lists, and, for the humanities, also from the C list. In scientific field groups where ranking is based solely on the A and B lists, the scientific excellence criterion is met by the top seven percent of researchers. In the humanities, this criterion is met by the top four percent, based on the combined results from the A, B, and C lists. All Project Leaders of ERC projects automatically qualify for scientific excellence. A returnee from

abroad – a researcher employed in a scientific title who obtained a PhD abroad or has at least five years of relevant professional experience abroad, (provided the researcher is included in the funding cycle between two rounds of ranking), also obtains the right to scientific excellence and retains this status until the next evaluation.

- The final list of researchers who have obtained the status of scientific excellence

The final list of researchers who have obtained the status of scientific excellence is created by combining the lists of excellent researchers generated within each scientific field group with the list derived from results achieved in international research collaborations.

The ranking procedure for determining the excellence corrective coefficient is conducted every two years, with the following deadlines: administrative data are retrieved from the Researcher Register by December 1; scientific data are retrieved from the eNauka repository by December 10; and the M-values (scientific values of the results) are submitted by December 13. The preliminary list, (containing twice the number of ranked researchers compared to the number who receive the bonus, that is, a total of 20%), is published by December 20, and the final list by December 25.

Publicly available sources of research outputs recorded in the eNauka information system are used to determine researcher excellence, and the list of ranked researchers who have met the excellence criterion is publicly available on the website of the ministry responsible for research activities.



Evaluation of Researchers

1. **The Regulation on the Appointment to Research and Scientific Titles (the Regulation)** defines the procedure, criteria, and methods for evaluating researchers' scientific and research outputs, including the quantitative and qualitative assessment of results for appointment, reappointment, and advancement within research and scientific titles.

The right to obtain a research or scientific title is granted to all individuals who meet the requirements set by the Law on Science and Research. The procedures are conducted by the scientific and teaching-scientific councils of accredited organizations, while decisions on appointments are made by the Scientific Boards and the Commission for the Appointment to Scientific Titles, according to the level of the title.

In the evaluation process, candidates submit detailed documentation that includes biographical information, a record of their scientific activities, an overview of their most

significant research outputs with descriptions of their contributions, quantitative indicators, a complete bibliography classified according to the M-categories, and a table demonstrating compliance with the minimum requirements. Committees appointed by the scientific or teaching-scientific councils, as well as the Scientific Boards, depending on the level of the title, assess the candidate's compliance with the prescribed criteria.

The evaluation of research outputs is based on the system of scientific field groups (Natural Sciences, Medical and Health Sciences, Engineering and Technology, Agricultural Sciences, Social

Sciences, and Humanities) and requires meeting the minimum quantitative and qualitative criteria specified in Annex 3 of the Regulation.

- **Quantitative evaluation –** M-values represent standardized categories of scientific outputs that enable quantitative assessment of a researcher's work based on the type and significance of the publications.

The minimum point thresholds are consistent across all scientific titles: 16 points for Research Associate, 50 points for Senior Research Associate, and 70 points for Principal Research Fellow. However, the required categories of research outputs vary according to the scientific field. In the Natural Sciences and the Medical and Health Sciences, emphasis is placed on publications in high-ranking international journals. In the Engineering and Technology and Agricultural Sciences, additional value is given to innovations, technical solutions, and patents. In the Social Sciences and the Humanities, greater weight is assigned to monographs, books, and articles published in national journals. This approach ensures a unified framework for advancement across scientific titles while providing discipline-specific criteria for the mandatory M-categories.

- **Qualitative evaluation –** Qualitative indicators of achievement are criteria used to assess the scientific excellence and professional contributions of the researcher, and their verification is conducted by the competent Scientific Board.

These indicators cover nine areas that together provide a comprehensive picture of the candidate's scientific reputation: (1) Influence, measured by citation counts and the Hirsch index according to the Scopus or Web of Science databases; in the Social Sciences and Humanities, Google Scholar may also be used, or, in the Humanities, documented citations of the candidate; (2) International collaboration, demonstrated by at least one of the following: a research stay at a foreign institution lasting more than three months, participation in international projects, or co-authored publications with researchers from foreign institutions in the relevant M-categories; (3) Leadership of projects and work packages, evaluated based on an established classification, ranging from European projects (ERC, Horizon Europe, Digital Europe, EIC/EIT) to national projects (Science Fund, Innovation Fund, bilateral projects, the Serbian Academy of Sciences and Arts, and Matica Srpska) and projects with industry valued at over 30,000 euros; (4) Editing of scientific publications, including roles such as editor-in-chief, managing editor, editor, or guest editor in M20-category journals indexed in WoS or Scopus, and, in the Humanities, also in M24 and M51 journals, as well as in thematic proceedings and monographic series; (5) Invited lectures, referring to lectures delivered at domestic or international institutions in science, higher education, or culture (excluding conferences); (6) Peer review of projects and scientific outputs, including the review of scientific projects and publications within the defined M-categories, with specific category definitions for the Humanities; (7) Training of scientific staff, including mentoring doctoral students and participating in teaching. Mentorship is recognized if the candidate has been formally appointed as a mentor by a university or research-performing organization and has co-authored publications with the doctoral candidate. Participation in teaching refers to delivering regular instruction at doctoral, master's, and specialist programs, and at the undergraduate level only if the candidate also holds an academic teaching title; (8) Awards and recognitions, including national and international awards granted by relevant scientific institutions through a formal procedure involving a qualified expert committee; (9) Contribution to the development of a scientific field, assessed based on research outputs demonstrating the candidate's role as a leading author in advancing a particular scientific area or discipline. The relevant publications must have been produced after completion of the doctoral dissertation, without the mentor as a co-author, and their authenticity and significance are verified by the competent Scientific Board.

For appointment to the title of Senior Research Associate, the candidate must meet at least three criteria from the combined A and B lists. Appointment to the title of Principal Research Fellow requires at least four criteria, with at least one from list A. Fulfillment of a single criterion from the A+ list is sufficient to meet all qualitative requirements.

- **List A** includes the core criteria of scientific excellence, evaluated through the candidate's career record. The requirements include project leadership, mentoring, achievements in categories M11 and M12 (for the Humanities), and the Hirsch index, which must be at least 9 for appointment as Senior Research Associate and 13 for appointment as Principal Research Fellow, except in Engineering and Technology, Agricultural Sciences, and Humanities, where the Hirsch index must be at least 7 for Senior Research Associate and 10 for Principal Research Fellow.
- **List B** includes supplementary criteria evaluated through two groups of indicators. The career record is used to assess the citation impact of the candidate's articles (with prescribed thresholds by field and title); international cooperation achieved through participation in projects or professional training abroad; leadership of sub-projects and work packages; editing of scientific publications; peer-review work on international and national projects; national and international awards and recognitions granted by relevant institutions; and the candidate's contribution to the development of a scientific discipline through a leading authorship role in scientific publications. During the assessment period, the following are also evaluated: international cooperation demonstrated through co-authored publications with foreign researchers; invited lectures (excluding conference presentations); and participation in teaching.
- **List A+** (exceptional scientific contribution) represents the highest level of qualitative evaluation and replaces all criteria from lists A and B for appointment to the titles of Senior Research Associate and Principal Research Fellow. The criteria in this list are: (1) leadership of international projects according to the established classification, with specific rules for the Humanities (career record); (2) a Hirsch index of at least 30 in the Web of Science or Scopus databases, excluding publications resulting from collaborations with more than 100 authors (career record); and (3) two publications ranked within the top 1% of papers in the relevant field according to the Web of Science (highly cited) in the year of publication, with the candidate required to be the first, last, or corresponding author (evaluation period).

In addition to Annex 3, "Minimum Quantitative and Qualitative Requirements for the Appointment to Individual Scientific Titles", the Regulation also includes:

- Annex 1: "Criteria for Determining the Category of Scientific Research Outputs and the Method of Referencing." This Annex defines the system for categorizing scientific outputs (M-categories) used to evaluate researchers during the appointment process for research and scientific titles. It covers all types of scientific outputs, including monographs, book chapters, journal articles, conference papers, doctoral dissertations, technical solutions, patents, various forms of intellectual property, architectural and artistic works, studies, exhibitions, and public policy documents. The Annex sets out: (1) the criteria for assigning scientific outputs to the appropriate M-category; (2) the conditions a publication or output must meet to be recognized; (3) the rules for referencing and verifying outputs; and (4) the method for normalizing outputs and allocating points among co-authors, including specific rules for publications with many co-authors and for technical solutions and patents. In this way, the Annex forms the foundation of the entire M-value system and provides the reference framework for consistent, standardized, and comparable evaluation of scientific outputs in the appointment process for research and scientific titles.

- Annex 2: "Types and Quantification of Individual Scientific Research Outputs." This annex defines the method for quantifying all individual scientific research outputs. It provides numerical values (K-values) representing the weight of each type of scientific output within the researcher evaluation system. Its purpose is to ensure uniform and consistent evaluation of outputs across all scientific fields, as well as their objective comparability in all procedures where scientific outputs are formally assessed. In other words, Annex 2 describes the fundamental mechanism for converting all scientific outputs into numerical values used in formal evaluation procedures.
- Annex 4: "Report Template of the Committee of the Research Performing Organisation (RPO)" prescribes the format for the committee's report on evaluating a candidate for a scientific title. The report must include basic information about the candidate, level of education, employment details, current and requested scientific titles, scientific field and subfield, a description of the candidate's previous research work, an overview of the most significant scientific outputs, an outline of research performance indicators (including citation data and other relevant indicators), quantification of outputs according to prescribed categories and values, and the committee's conclusion and recommendation. The template is prescribed to ensure a uniform, transparent, and verifiable method for preparing reports in the candidate evaluation procedure.

Although the evaluation system provides a unified framework for monitoring scientific outputs, certain elements related to methodological consistency, interpretation of qualitative criteria, and measurement of broader research impact remain insufficiently elaborated and require further operationalization.

1.3. Research Assessment Practices at the Institute of Economic Sciences

Following the analysis of the national research assessment framework, ARISE examined existing assessment practices at the Institute of Economic Sciences as the institutional setting in which the new assessment model would be developed and piloted. The analysis considered how research performance was assessed within the Institute, the extent to which existing practices incorporated qualitative elements, and whether mechanisms were in place for assessing research at the level of departments. Particular attention was given to the relationship between formal national requirements, which necessarily shape assessment practices at IES, and the Institute's own internal procedures and capacity to introduce broader approaches to research assessment.

The analysis showed that IES already contained the "seeds" of reform: elements of qualitative judgement were present in internal procedures, values-based discussions were emerging, and institutional leadership was engaged. At the same time, these elements were not yet embedded in a coherent, transparent and systematic framework for research assessment. What was missing was an integrated approach capable of linking research values, assessment criteria and methods with institutional objectives and supporting research development at the level of departments and the Institute as a whole.

Institutional Overview

The Institute of Economic Sciences (IES) is a public research institution operating under the jurisdiction of the Ministry of Science, Technological Development and Innovation of the

Republic of Serbia**.** Established with the aim of advancing excellence in economic research and contributing to evidence-based policymaking, IES combines fundamental and applied research in all major fields of economics and related social sciences. Guided by the principle of excellence and the motto *"Cognoscere est mutare"* (*"To know is to change"*), IES has earned a strong reputation as one of the leading research organizations in Southeast Europe.

IES conducts high-quality research and delivers measurable impact across multiple areas of society, primarily in macroeconomics, welfare economics, environmental economics, digital economics, economic history and theory, and innovation economics**.** Through these research fields, the Institute contributes to improving the competitiveness of the Serbian economy, supporting sustainable development, and strengthening the role of social sciences in evidence based public policy formulation.

The vision of IES is to become a leading scientific research institution in the field of economics in Southeast Europe, fully integrated into the European Research Area. Its mission has two dimensions:

- Internally**,** to be an attractive and stimulating environment for high-quality scientific research;
- Externally**,** to provide excellent and relevant research that supports economic development and informs public decision-making.

IES is guided by a set of core institutional values**,** which include:

- Adherence to the highest scientific standards and research ethics;
- Alignment with national priorities for scientific and technological development;
- A reward system based on measurable research outcomes;
- Partnership-oriented cooperation with the Ministry of Science and the Science Fund of the Republic of Serbia;
- Respect for all employees and maintenance of a positive working environment that fosters creativity;
- Academic freedom and encouragement of innovation.

The Institute's strategic goals for the period until 2025 are:

1) Advancing research excellence**;**
2) Expanding institutional capacity**;**
3) Strengthening societal influence and improving organizational performance.

These goals are operationalized through specific objectives, such as:

1) Increasing the number of publications in top-ranked international journals;
2) Enhancing the capacity to implement Science Fund projects;
3) Raising the competence level of researchers;
4) Expanding international cooperation and strengthening strategic partnerships;
5) Disseminating scientific results and increasing visibility through media and digital resources;
6) Promoting cooperation with the business sector;
7) Improving organizational efficiency and internal coordination.

A new Strategy for the Development of the Institute of Economic Sciences for 2026–2030 – Strategy '30 is currently in the adoption process. The Strategy sets out the overarching goal of building a modern, sustainable, and internationally recognized research institute that contributes to socio-economic development through high-quality economic research.

The general objective is operationalized through specific objectives, such as:

1) **Strengthening International Recognition -** position the Institute as a regional and international leader in economic research through alignment with global standards, strategic partnerships, and active participation in the European Research Area (ERA).
2) **Increasing Research Relevance and Impact -** ensure that research outputs address key societal challenges, including sustainable development and economic resilience, thereby enhancing their contribution to national and regional development.
3) **Enhancing Knowledge and Innovation Transfer -** build stronger links between research and practice by improving mechanisms for knowledge exchange, collaboration with industry and policymakers, and the development of solutions for complex economic issues.
4) **Ensuring Sustainability and Institutional Efficiency -** modernize infrastructure, diversify funding sources, and strengthen management practices to support effective decision-making and resource use.

The organizational structure of IES is designed to support both flexibility and accountability. Governance is ensured through a Steering Committee (seven members appointed by the Ministry of Science, Technological Development and Innovation, three from IES and four external experts, serving a four-year term). The Director is supported by a Collegium (comprising the Director, Deputy Director, Head of the Sector for Fundamental Research, and Head of the Sector for Applied Research) and an Extended Collegium**,** which includes the Collegium members and the six Heads of Research Departments. The Scientific Council**,** composed of 15 members elected for a two-year term, serves as the Institute's academic body.

IES currently employs 46 researchers**,** including 6 Principal Research Fellows, 16 Senior Research Associates, 18 Research Associates, 5 Research Assistants, and 1 Junior Research Assistant, supported by a part-time HR Manager and eight administrative staff members. Research is conducted within six departments:

- Macroeconomics
- Digital Economics
- Environmental Economics
- Welfare Economics
- Innovation Economics
- Economic History and Theory

This organizational structure ensures an effective balance between strategic oversight, scientific autonomy, and operational flexibility, allowing the IES to pursue its mission of contributing to national and European research excellence.

Current Research Assessment Practices

At the Institute of Economic Sciences, the assessment of researchers is conducted in accordance with the national regulatory framework**,** specifically the Rulebook on the Acquisition of Research and Scientific Titles**,** adopted by the Ministry of Science, Technological Development and Innovation. The IES does not independently determine criteria for the appointment of researchers but acts in line with national regulations.

The Scientific Council of IES initiates procedures for the appointment to research and scientific titles, adopts decisions on appointments to research titles, and forwards proposals for

appointments to scientific titles to the competent national bodies. The final decisions on appointments to scientific titles are made by the Scientific Field Committee (for the title of *Scientific Associate*) and the Commission for the Election to Titles (for the titles of *Senior Research Associate* and *Principal Research Fellow*).

Internally, research quality assurance is primarily the responsibility of the Sector for Fundamental Research (SFR), which systematically monitors and enhances the quality of research output. The key performance indicators used at the IES include the number of publications in indexed journals (SCI/SSCI), citation impact (Web of Science, Scopus, Google Scholar), research project involvement, and dissemination of research results.

The SFR conducts quantitative and qualitative assessments of research performance on an annual basis, supported by quarterly and semi-annual monitoring of individual researchers, departments, and the Institute as a whole.

- **Quantitative assessment** focuses on the number of published papers across result categories (M10–M60) and departments, as well as productivity indicators calculated by relating total publications to the number of researchers in a given year.
- **Qualitative assessment** is based on citation analysis in leading databases (WoS, Scopus, Google Scholar) and the quality of journals where the papers are published.

In addition, the Institute has adopted the Rulebook on Minimum Expected Results and the Evaluation of Research Performance[7], which defines minimum performance expectations for researchers and departments. These expectations are aligned with the strategic goals of IES for the 2021–2025 period: enhancing excellence, expanding capacities, strengthening impact, and improving organizational efficiency.

The Rulebook outlines minimum annual achievements (expressed through points) based on:
- published papers (adjusted by scientific title),
- participation in projects,
- contribution to the organization of scientific and professional events,
- participation in conferences,
- involvement in the preparation of public policy documents,
- activities that increase departmental visibility, and
- networking and collaboration efforts.

Although this document is not of a binding character, it serves as an internal guideline and benchmarking tool. It also forms the basis for the annual IES Research Excellence Award, granted in two categories: *Best Young Researcher* and *Best Researcher*. The award is financial and aims to recognize and incentivize researchers who demonstrate performance exceeding the expected results.



Internal Rulebook on the Evaluation of Research Performance

The Rulebook on Minimum Expected Results and the Evaluation of Research Performance at the Institute of Economic Sciences establishes internal standards and procedures

[7] Institute of Economic Sciences. (2023). *Rulebook on Minimum Expected Results and Evaluation of Research Performance*. Belgrade: Institute of Economic Sciences.

for assessing research achievements of both individual researchers and research departments. Although the Rulebook is not legally binding, it serves as a key internal mechanism for promoting excellence, transparency, and merit-based recognition within the Institute. Its provisions are fully aligned with the strategic objectives of IES defined in the *Programme of Work and Development 2021–2025: Strategy '25*[8], which emphasizes excellence, capacity expansion, impact, and organizational efficiency.

The Rulebook defines:

- Minimum expected annual results for individual researchers and research departments.
- Quantitative and qualitative criteria for evaluation.
- Procedures for scoring, normalization, and ranking researchers based on their annual performance.
- Eligibility for the annual *IES Research Excellence Award* (Best Researcher and Best Young Researcher).

The purpose of the Rulebook is to ensure objective and comparable evaluation of research output, stimulate higher performance, and recognize exceptional results.

A. Minimum Expected Annual Results for Individual Researchers

Each researcher is expected to achieve a minimum annual score based on the value of their research outputs (classified under the *M-categories* defined by the national *Rulebook on Scientific Titles*). The thresholds vary by academic title:

Research Title	Minimum Points
Junior Research Assistant (first year)	No minimum
Junior Research Assistant	4 points
Research Assistant	6 points
Research Associate	5 points (journals) or 6 points (obligatory category 2)
Senior Research Associate	5 points (journals) or 6 points (obligatory category 2)
Principal Research Fellow	6 points or leadership of a major project / program committee

B. Minimum Expected Results for Research Departments

Departments are evaluated collectively through six key elements:

1. **Scientific events** – At least one academic event (conference or round table).
2. **Projects** – Application or implementation of at least one competitive project involving ≥50% of department members.
3. **Conference participation** – Number of conference contributions (M30/M60 or M40) equal to the number of members.
4. **Preparation of M120 outputs** – Each newly promoted PhD must submit one policy-related or analytical paper (M120).
5. **Department visibility** – Preparation of a brand/visibility development plan.
6. **Networking** – Development of a list of relevant institutions and experts in the department's field.

[8] Institute of Economic Sciences. (2021). *Work and Development Program of the Institute of Economic Sciences for the period 2021–2025: Strategy '25*. Belgrade: Institute of Economic Sciences.

C. Evaluation Methodology

The evaluation system combines quantitative and qualitative indicators, each contributing up to 50 points, with a total of 100 normalized points per researcher.

1. **Quantitative Results (up to 50 points)**
   - Based on publications in journals and monographs (M11–M54 categories).
   - Points are distributed according to the author's position:
     - 1st co-author: 0.5 × K
     - 2nd co-author: 0.4 × K
     - 3rd co-author: 0.3 × K
     - 4th and later: 0.1 × K
   - The total score per researcher is normalized so that the top-performing researcher receives 50 points.

2. **Qualitative Contribution (up to 50 points)**
   - Points are awarded for:
     - Project leadership (2 points for competitive projects, 1 for others).
     - Participation in projects (0.3–0.7 depending on type).
     - Editorial and organizational activities (e.g., journal editing, conference organization, chairing committees).
     - Citations (based on the share of total IES citations in WoS).
     - Mentorship (PhD supervision, internal mentoring).
     - International collaboration (co-authorships, joint project proposals, membership in boards).
     - Organization of research activities (heads, chair of Scientific Council, KT officer).
     - Member of the Transformation team, member of the Scientific Council, member of the Project office.
     - Contribution to public policy papers (M121–M124 categories).
   - Each activity has a defined maximum annual value, ensuring balance across different types of contributions.

The total researcher score is the sum of normalized quantitative and qualitative points. Ranking is based on total points, and results are compiled annually by the *Department for Fundamental Research* using internal databases and reports.



Use of Evaluation Results

- The evaluation informs internal decision-making, recognition, and annual awards.
- The *IES Research Excellence Award* is granted annually in two categories:
  - Best Young Researcher
  - Best Researcher
- Awards are financial and serve as motivation for achieving results above the expected thresholds.

In summary, the *Rulebook on Minimum Expected Results and the Evaluation of Research Performance* represents an important internal mechanism for promoting a culture of quality, accountability, and transparency within the Institute of Economic Sciences. Although it does not

constitute a formal instrument for career advancement since appointments to scientific titles are governed by national regulations it provides a structured framework for monitoring progress, motivating researchers, and aligning individual achievements with institutional strategic goals. By combining quantitative indicators with qualitative dimensions such as project involvement, policy engagement, and collaboration, the Rulebook contributes to building an institutional environment oriented toward excellence and continuous improvement.

Alignment with European Principles

The Institute of Economic Sciences demonstrates a growing degree of alignment with European principles of responsible research assessment, particularly those promoted through the Agreement on Reforming Research Assessment and the Coalition for Advancing Research Assessment (CoARA). Although Serbia's national system remains predominantly quantitative, several institutional practices at IES already reflect emerging European standards. These include the combination of quantitative and qualitative indicators, systematic monitoring of research output, and the consideration of broader aspects of academic work such as project participation, dissemination activities, mentoring, and contribution to policy development.

At the European level, current reform trends emphasise reducing the reliance on journal-based metrics, recognising a broader diversity of research contributions, and strengthening qualitative, peer-reviewed assessment processes. IES has begun to incorporate elements of this approach through its internal evaluation mechanisms, most notably the Rulebook on Minimum Expected Results and the Evaluation of Research Performance, which balances quantitative measures with qualitative contributions related to collaboration, visibility, project leadership, and societal engagement.

Furthermore, the Institute's participation in the ARISE project reinforces its commitment to aligning internal assessment practices with European reform efforts. The project explicitly aims to establish a qualitative, context-sensitive and peer-review-based evaluation model for IES, consistent with the core principles advocated by CoARA, including responsible use of metrics, transparency, recognition of diverse outputs and career paths, and strengthening institutional research culture.

Through these activities, IES is gradually integrating European norms and expectations into its internal assessment system, positioning itself as an institution ready to adopt advanced evaluative practices and to contribute to Serbia's broader transition toward a modern, qualitative and coherent research assessment framework.

Identified Challenges

Although the Institute of Economic Sciences has made important steps toward modernising its internal research assessment system, several gaps remain that limit the ability to fully align with contemporary European standards. One of the notable strengths of the Institute's current practice is the introduction of qualitative indicators within its internal evaluation procedures. Unlike the national research assessment framework which is almost entirely quantitative and based on citation counts, publication categories and financial inflows the Institute has incorporated elements such as contributions to institutional development,

participation in international networks, mentoring activities, policy engagement and broader forms of academic service. This represents a meaningful shift toward a more balanced assessment model and demonstrates the Institute's willingness to align with evolving European practices.

However, despite this progress, the internal evaluation framework still relies predominantly on quantitative measures, particularly through the Rulebook on Minimum Expected Results and the criteria used for annual institutional awards. These instruments provide transparency and accountability, but they do not sufficiently capture the complexity of research performance in the social sciences, nor do they recognise the diversity of research trajectories and contributions. Assessment remains focused on numerical thresholds and publication counts, leaving limited space for qualitative judgement, contextual interpretation or recognition of outputs with high societal relevance, policy impact or innovative character.

A persistent challenge is the absence of structured peer-review processes that would bring external perspective and disciplinary sensitivity into institutional assessment. Departments and research groups are not evaluated through expert committees, nor is there a formal mechanism for assessing research strategy, governance quality, the research environment or long-term developmental capacities. Furthermore, although qualitative elements exist, they are not yet systematically defined or consistently applied across all departments, which limits their effect on institutional planning and strategic steering.

Digitalisation also remains partial. Research data and records are collected through a combination of manual and semi-structured processes, making it difficult to integrate qualitative evidence, track performance or provide a comprehensive evidence base for more advanced forms of evaluation. This gap is particularly important in the context of Serbia's transition toward modern, data-enabled research governance.

The Institute has demonstrated clear readiness for change by introducing qualitative indicators ahead of national reforms and by actively participating in the Horizon Europe CoARA Boost project. Yet, further development is needed to establish a coherent, qualitative, peer-reviewed and contextually grounded evaluation model. The ARISE project is designed precisely to address these challenges by developing a new qualitative assessment framework for research departments, aligning institutional practices with CoARA principles, and creating a model that can later be shared with other social science institutes and eventually integrated into national reform efforts.

2. Global and European Framework for Research Assessment Reform

Research assessment reform has become an increasingly prominent international and European policy agenda, driven by growing recognition of the limitations of assessment systems that rely heavily on journal-based metrics, citation indicators and other quantitative proxies for research quality. A series of international initiatives has contributed to a broader understanding of research quality and contribution, emphasising the responsible use of metrics, qualitative expert judgement, recognition of diverse research outputs and activities, and greater attention to research context, mission and societal relevance. In Europe, these principles have increasingly been translated into a coordinated reform agenda through the Agreement on Reforming Research Assessment and the Coalition for Advancing Research Assessment (CoARA). Together, these developments provide the broader conceptual and policy framework within which the ARISE project was developed and implemented.

2.1. From Metrics to Responsible Research Assessment: DORA and the Leiden Manifesto

Growing concern over the limitations of metric-driven research assessment has led to a broader international movement towards more responsible approaches to evaluating research quality and contribution. Two initiatives have been particularly influential in shaping this shift: the San Francisco Declaration on Research Assessment (DORA)[9] and the Leiden Manifesto for Research Metrics[10]. Although developed in different contexts, both challenge the inappropriate use of quantitative indicators as proxies for research quality and emphasise the need for assessment practices that recognise the substance, context and diversity of research contributions.

DORA was developed in 2012 in response to concerns about the widespread use of journal-based metrics, particularly the Journal Impact Factor, for purposes for which they were not originally designed. Its central recommendation is that journal-based metrics should not be used as surrogate measures of the quality of individual research articles, to assess an individual researcher's contributions, or in decisions concerning hiring, promotion and funding. Instead, DORA calls for research to be assessed on its own merits and encourages consideration of a broader range of research outputs and forms of impact. Its recommendations address funders, research institutions, publishers, organisations supplying metrics and individual researchers, thereby recognising that meaningful reform requires changes in assessment practices across the research system rather than action by individual researchers or institutions alone.[1]

An important implication of DORA is that the quality and significance of research cannot be inferred from the characteristics or reputation of the journal in which it is published. Assessment should therefore create space for consideration of the content and contribution of research itself and should recognise outputs and contributions beyond conventional research

[9] San Francisco Declaration on Research Assessment (DORA). (2012). *San Francisco Declaration on Research Assessment*. American Society for Cell Biology.
[10] Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520*, 429–431.

publications. In this sense, DORA does not simply challenge a particular metric; it challenges the broader practice of substituting journal-level indicators for direct assessment of research quality and contribution.[1]

The Leiden Manifesto, published in 2015, complements this perspective by providing ten principles for the responsible use of research metrics.[2] Its fundamental premise is that quantitative evaluation should support qualitative expert assessment rather than replace it. The Manifesto also emphasises that research performance should be assessed in relation to the missions of institutions, groups and researchers and that assessment should account for differences in research fields and practices. It calls for transparency in data collection and analytical processes, enables those being assessed to verify the data used in evaluation, and cautions against false precision and inappropriate comparisons across fields. It also stresses the importance of protecting excellence in locally relevant research and recognising that indicators themselves can influence research behaviour and may need to be regularly reviewed and updated.

DORA and the Leiden Manifesto are therefore complementary rather than competing approaches to research assessment reform. DORA primarily challenges the inappropriate use of journal-based metrics as proxies for research quality and calls for broader recognition of research outputs and contributions, while the Leiden Manifesto provides a set of principles for using quantitative evidence responsibly when metrics form part of assessment. Together, they helped establish a fundamental principle of responsible research assessment: quantitative evidence can inform assessment, but it should be interpreted in context and should not substitute for informed expert judgement. This represents a shift not from metrics to the absence of metrics, but from metric dominance to the responsible use of metrics within broader qualitative assessment.

These principles provide an important foundation for subsequent developments in European research assessment reform. The emphasis on direct consideration of research quality and contribution, contextual interpretation, qualitative expert judgement and the responsible use of quantitative evidence was further developed and institutionalised through the Agreement on Reforming Research Assessment and the establishment of the Coalition for Advancing Research Assessment (CoARA), discussed in the following section.

2.2. CoARA and the European Research Assessment Reform

At the European level, the principles of responsible research assessment have been consolidated into a more coordinated reform agenda through the Agreement on Reforming Research Assessment, published in 2022, and the subsequent establishment of the Coalition for Advancing Research Assessment (CoARA).[11] The Agreement sets out a shared direction for reforming the assessment of research, researchers and research organisations, while recognising that assessment serves different purposes and takes place at different levels, including individual researchers, research projects, research units and research-performing organisations. Rather than prescribing a single assessment model, it establishes common principles and commitments that organisations can implement in ways appropriate to their missions, disciplinary contexts and assessment purposes.

At the core of the Agreement is a broader understanding of what should be recognised and valued in research assessment. Its vision is that assessment should recognise the diversity of

[11] Coalition for Advancing Research Assessment (CoARA). (2022). Agreement on Reforming Research Assessment (Version 1).

research outputs, practices and activities that contribute to the quality and impact of research. This entails moving beyond narrow definitions of research performance centred predominantly on publications, journal prestige and citation-based indicators and creating greater space for the diversity of contributions through which researchers, research units and organisations contribute to science and society. The Agreement therefore places research assessment within a broader understanding of research quality, impact and contribution, rather than treating performance as reducible to a limited set of quantitative measures.[3]

A defining principle of the CoARA reform agenda concerns how assessment should be conducted. The Agreement states that research assessment should be based primarily on qualitative judgement, for which peer review is central, supported by the responsible use of quantitative indicators.[3] This principle does not imply abandoning metrics. Instead, it changes their role: quantitative indicators are expected to provide supporting and contextual evidence rather than substitute for expert judgement. In this respect, CoARA builds directly on the direction established by DORA and the Leiden Manifesto, while embedding these principles within a broader framework for institutional and systemic reform.

The Agreement translates this vision into ten commitments, including four core commitments that establish the foundation of reform. These call on signatories to recognise the diversity of contributions to, and careers in, research according to the needs and nature of research; base research assessment primarily on qualitative evaluation, with peer review playing a central role and quantitative indicators used responsibly; abandon inappropriate uses of journal- and publication-based metrics, particularly Journal Impact Factor and h-index; and avoid the use of research organisation rankings in research assessment. The remaining commitments concern the resources, institutional changes, criteria and processes, awareness raising, exchange of practices, communication of progress and evidence needed to translate these principles into sustained changes in assessment practice.[3]

Importantly, CoARA frames research assessment reform as a process of institutional change and learning, rather than simply the replacement of one set of indicators with another. The Coalition provides a collaborative framework within which organisations can experiment with and pilot new assessment approaches, exchange practices and experience, develop assessment criteria, methods and tools, and learn from implementation while retaining autonomy over how reforms are introduced in their particular contexts. This makes reform inherently iterative: organisations may start from different positions and adapt their approaches to different disciplinary, institutional and national circumstances while working towards a common direction of change.

This implementation-oriented dimension is particularly relevant to ARISE. The CoARA Boost initiative was established to strengthen the capacity of organisations to translate the commitments of the Agreement into changes in research assessment practice, including through the development and testing of new assessment approaches, models and procedures. Within this framework, ARISE represents an institutional pilot through which the broader principles of European research assessment reform were translated into a concrete assessment approach in the Serbian research context. Its emphasis on multidimensional research contribution, narrative self-assessment, external peer review, responsible use of metrics and attention to institutional and disciplinary context therefore reflects not only the general direction of responsible research assessment, but the specific reform logic articulated through the CoARA Agreement.

2.3. European Pathways to Responsible Assessment: Lessons from France and Romania

The comparative analysis undertaken within ARISE examined research assessment practices in France and Romania as two European cases representing different pathways towards broader and more qualitative approaches to research assessment. The purpose was not to identify a foreign model that could be directly transferred to Serbia, but to examine how different assessment systems address the limitations of predominantly metric-driven approaches and to identify elements that could inform the development of a context-sensitive model in Serbia.

France and Romania provide particularly relevant comparative perspectives because they represent different stages and approaches to research assessment reform. The French system offers an example of a well-established qualitative approach centred on the assessment of research units, narrative self-assessment and external peer review. Romania represents a more transitional approach in which quantitative indicators and scoring retain an important role, while assessment also incorporates broader dimensions of institutional performance. The comparison therefore provides insights into both a mature qualitative model and a more gradual pathway towards broadening research assessment within an existing regulatory framework.

The French Model of Research Assessment

France provides a particularly relevant example of an established research assessment system in which qualitative peer review plays a central role. Its approach combines institutional self-assessment, external expert evaluation and the contextual use of quantitative evidence, while considering research performance in relation to institutional missions, strategies and research environments. The following overview highlights the main features of the French model that are particularly relevant to the development of a more qualitative and context-sensitive approach to research assessment in Serbia.



Institutional Framework

France has a long-established system of external evaluation of higher education and research institutions, developed progressively since the 1980s. A major step in the consolidation of the system was the establishment of the Agency for the Evaluation of Research and Higher Education (AERES) in 2006, which brought previously separate evaluation functions covering higher education institutions, research units, research organisations and doctoral education within a single independent administrative authority. The present High Council for the Evaluation of Research and Higher Education (Hcéres) developed from this institutional framework and serves as the independent public authority responsible for the external evaluation of higher education and research structures in France.



Scope and Role of Hcéres

The French public research landscape comprises universities, grandes écoles and national research organisations such as the National Centre for Scientific Research (CNRS). Research is frequently conducted through joint research units involving universities and national research

organisations under shared governance. Within this system, Hcéres operates under a clearly defined mandate for evaluating research activities, established in the French Research Code and expanded by the Multi-Annual Research Programming Law (2020). Its responsibilities include[12]:

- Evaluating higher education institutions and their groupings, research units, research-performing organisations, scientific cooperation foundations, and the French National Research Agency (ANR);
- Assessing activities related to the dissemination of scientific, technical and industrial culture within institutions, research structures, units and programmes;
- Promoting research integrity and ensuring its integration into all evaluations conducted or validated by Hcéres.

A central feature of the French approach is the distinction between the assessment of research structures and the assessment of individual researchers. Hcéres does not evaluate individual researchers or their career progression. Its assessments focus on institutions and research units, their research activities, strategies and environments, while individual academic careers are assessed through separate national procedures. This distinction enables research assessment to address collective performance and the conditions in which research is conducted, rather than deriving institutional or unit-level quality from aggregated individual-level metrics.

Hcéres' approach also reflects the broader principles of responsible research assessment represented by the Leiden Manifesto, DORA and the Agreement on Reforming Research Assessment promoted by CoARA. In its guidance to expert panels, quantitative and bibliometric indicators are treated as complementary sources of evidence rather than as substitutes for qualitative assessment. Where journal-based metrics are used, evaluators are expected to recognise their limitations and interpret them cautiously.



A Qualitative and Context-Sensitive Approach to Research Unit Assessment

Research units constitute a central level of assessment within the Hcéres framework. They are evaluated on a five-year cycle through a process that combines institutional self-assessment, supporting evidence and external peer review. The approach is designed to assess research performance in relation to the specific scientific, institutional and disciplinary context of each unit, rather than against a uniform set of quantitative performance thresholds. Research performance is assessed through the analysis and interpretation of observable evidence in relation to the aims and context of the evaluation, with quantitative data serving as supporting rather than defining evidence.[13]

The assessment of research units is guided by five main principles: contextualised evaluation that takes account of disciplinary, institutional and regional specificities; qualitative assessment supported by the responsible use of quantitative indicators; peer review as the foundation of the assessment process; strong reliance on self-assessment; and consideration of the research unit's development and trajectory over the five-year assessment period.

[12] Hcéres. (2020). *Evaluation Framework for Higher Education and Research Institutions*. Paris: High Council for the Evaluation of Research and Higher Education.
[13] Hcéres. (2022). *Guidelines for the Evaluation of Research Outputs and Activities*. Paris: Hcéres.

Self-assessment is a central component of the process. Research units prepare a structured Self-Assessment Report covering their scientific objectives, organisation and resources; scientific results, visibility and attractiveness; and contribution of research to society. The report enables units to present and reflect on their activities and achievements within the context of their mission, resources and strategic objectives.

The Self-Assessment Report is complemented by supporting evidence, including a portfolio of selected research outputs and achievements. Rather than relying solely on an exhaustive count of publications or other outputs, research units select a limited number of achievements that they consider particularly significant in demonstrating their scientific quality and identity. These may include publications as well as other forms of scientific contribution, depending on the field and profile of the unit. An exhaustive list of scientific production is also provided as contextual evidence.

The assessment process also incorporates a SWOT analysis. Research units identify strengths and weaknesses in relation to the areas covered by the evaluation framework, while external evaluators consider opportunities and risks arising from the unit's broader context. This combination of self-assessment, selected evidence and external expert judgement enables the evaluation to consider not only what a research unit has produced, but also how its achievements relate to its objectives, resources, environment and future development.



Contextualized Expert Judgement

A defining feature of the Hcéres approach is that research units are assessed within their specific scientific, institutional and territorial contexts. Rather than applying uniform performance expectations across different types of research units, expert panels interpret achievements in relation to the unit's mission, objectives, resources and operating environment. Evaluative judgements must be supported by verifiable evidence, including the information and selected achievements presented by the research unit itself.

This contextual approach recognises that comparable expectations cannot necessarily be applied to research units with substantially different profiles and capacities. A large joint research unit comprising several teams and benefiting from substantial human and financial resources may, for example, be expected to make greater use of available funding and collaboration opportunities. A smaller unit operating with more limited resources may instead demonstrate strength through its capacity to sustain high-quality and original research despite those constraints. Research performance is therefore interpreted in relation to what a unit seeks to achieve and the conditions under which it operates.

Qualitative expert judgement makes it possible to recognise dimensions of research performance that cannot be adequately captured through standardised indicators alone. Depending on the unit's field, mission and context, relevant evidence may include scientific visibility, recruitment and researcher mobility, support for doctoral researchers, collaboration with external actors, contributions to public policy, dissemination of research, and multidisciplinary or interdisciplinary work. The assessment also considers the position of the research unit within the wider institutional research environment and the relationship between its priorities and the broader strategic orientation of the institution.

Balancing Qualitative and Quantitative Evidence

The emphasis on qualitative assessment does not imply the abandonment of quantitative evidence. Hcéres continues to use bibliometric and other quantitative data where they provide useful information about research activity, for example in mapping publication patterns and authorship contributions. Such data contribute to a broader evidence base, but are interpreted alongside qualitative and contextual information rather than used as the primary basis for evaluative judgement.

The French experience also illustrates some of the practical challenges involved in moving towards more qualitative research assessment. Despite the introduction of portfolios and greater emphasis on contextualised expert judgement, research units are still required to provide substantial amounts of quantitative information. This creates a tension between the ambition to move away from exhaustive, indicator-driven evaluation and the continuing demand for comprehensive data on research activity. The introduction of qualitative elements can therefore add to, rather than reduce, the assessment burden if existing reporting requirements remain unchanged.

Administrative burden is consequently an important consideration in the further development of the Hcéres approach. Some researchers perceive the qualitative portfolio as an additional reporting requirement rather than as a replacement for existing quantitative reporting. Hcéres has responded by exploring greater use of existing research information sources and by simplifying its evaluation framework, including greater flexibility in the format in which research units provide information on their scientific production.

The French experience therefore demonstrates that reforming research assessment involves more than introducing qualitative criteria or peer review into an existing system. The balance between qualitative and quantitative evidence, the volume of information required, and the administrative feasibility of the assessment process all need to be considered if qualitative assessment is to function as a useful instrument for research development rather than as an additional layer of reporting.

The Romanian Model of Research Assessment

Romania provides another relevant example of the gradual development of research assessment practices that combine quantitative evidence with qualitative and expert-based approaches. Its experience is particularly useful for considering how assessment frameworks can evolve within a national research system while taking into account institutional missions, research performance and broader organisational development. The following overview focuses on those features of the Romanian approach that offer useful points of comparison for the further development of research assessment in Serbia.

Institutional and Regulatory Framework

Romania's research and innovation system has undergone a series of reforms aimed at improving its performance, coherence and alignment with developments in the European Research Area. The National Research-Development and Innovation Plan 2022–2027 (PNCDI IV)

provides an integrated framework for research and innovation funding and defines strategic objectives for the national system. At the same time, Romania continues to face structural challenges related to fragmentation, limited funding and research governance, reinforcing the importance of effective evaluation and performance frameworks.

The evaluation of research, development and innovation organisations is established through a national regulatory framework rooted in Government Ordinance No. 57/2002 on scientific research and technological development and subsequent amendments. The framework applies to a broad range of research-performing organisations, including national research institutes, institutes of the Romanian Academy, public and private universities, public research centres and private-sector research entities. Evaluation is mandatory and periodic and is intended to assess institutional performance, strengthen research governance and support the development of the national research and innovation system.

The Romanian model is particularly relevant from a comparative perspective because it represents a different reform pathway from the French system. Rather than replacing quantitative assessment with a predominantly qualitative model, Romania combines quantitative indicators and scoring with qualitative dimensions of institutional performance, self-assessment and external expert review. It can therefore be understood as a transitional model in which broader and more context-sensitive dimensions of research performance are incorporated into a structured, performance-oriented assessment framework.

Evaluation Principles and Criteria

Romania applies a uniform criteria-based evaluation framework that is adapted to the profile of the research-performing organisation, distinguishing between organisations primarily engaged in basic research, applied research and technological development and innovation, as well as private higher education institutions. The framework combines quantitative and qualitative dimensions of institutional performance through six assessment criteria, with a maximum overall score of 200 points.[14]

The quantitative component accounts for a maximum of 100 points and covers development and innovation activities, as well as research results and their national and international impact. The remaining 100 points relate to broader dimensions of institutional performance, including management and institutional capacity, socio-economic impact, human resources and the quality of the institution's strategic development plan for the following five-year period.

Assessment criterion	Type of assessment	Maximum score
Development and innovation activities	Quantitative	34
Research results and impact at national and international level	Quantitative	66
Management and institutional capacity	Qualitative	46
Performance and socio-economic impact	Qualitative	26

[14] National Council for Science Research of Romania. (2023). *Performance Indicators and Evaluation Criteria for Research Organisations*. Bucharest: CNCS.

Human resources	Qualitative	18
Strategic development plan for the next five years	Qualitative	10
Total		200

This structure illustrates the hybrid character of the Romanian model. Conventional measures of research activity and output retain a substantial role, but they account for only one part of the overall assessment. Equal weight is formally assigned to dimensions that extend beyond research outputs, including institutional governance and capacity, societal and economic relevance, human resources and strategic development. The model therefore broadens the scope of institutional assessment without abandoning quantitative measurement and scoring.

The inclusion of a five-year strategic development plan is particularly relevant from the perspective of developmental research assessment. Institutions are assessed not only on past and current performance, but also on the quality and realism of their strategic objectives and the extent to which these objectives reflect developments in the relevant scientific fields and societal needs. In this respect, the Romanian framework introduces a forward-looking dimension into an otherwise strongly structured and score-based assessment system.

Institutional Architecture and Evaluation Process

The Romanian evaluation system combines central oversight with external expert assessment. The Ministry responsible for research oversees the national evaluation process and appoints the Evaluation Board, which includes international experts as well as representatives of higher education institutions, national research institutes, institutes of the Romanian Academy and other research organisations. The Evaluation Board approves the evaluation procedures and validates the final evaluation reports.

The assessment follows a multi-stage process combining institutional self-assessment, external expert review and formal validation.[15] Each research organisation first prepares a self-assessment report covering a five-year period, using a template adapted to its research profile. The report includes general institutional information, key indicators of research activity and performance, progress in implementing recommendations from previous evaluations, and the institution's development strategy for the following five years.

The self-assessment is subsequently reviewed by external experts, who assess the evidence provided by the institution and prepare a final evaluation report, including scores for the individual assessment criteria. The Evaluation Board then reviews the proposed report and either validates or rejects it. The methodology is adapted to different types of research activity, allowing the framework to be applied across organisations engaged in basic research, applied research, technological development and innovation.

The self-assessment is subsequently reviewed by external experts, who assess the evidence provided by the institution and prepare a final evaluation report, including scores for the individual assessment criteria. The Evaluation Board then reviews the proposed report and either validates or rejects it. The methodology is adapted to different types of research activity,

[15] Romanian Ministry of Research, Innovation and Digitization. (2023). *Methodology for the Institutional Evaluation of Research Organisations*. Bucharest: Government of Romania.

allowing the framework to be applied across organisations engaged in basic research, applied research, technological development and innovation.

Based on the overall score, research organisations are classified into three performance groups. Class 1 comprises institutions receiving more than 150 points, Class 2 those receiving between 60 and 150 points, and Class 3 institutions receiving fewer than 60 points. The classification is linked to institutional financing and access to national competitive funding, giving the assessment process consequences beyond the evaluation itself.

The Romanian model therefore combines elements that are characteristic of broader approaches to research assessment—self-assessment, external expert review, institutional strategy and consideration of dimensions beyond research outputs—with a strongly structured scoring and classification system. Unlike the French approach, where contextualised expert judgement has a more central role, the Romanian model retains explicit quantitative scoring and links assessment outcomes to institutional performance categories and funding. This combination makes Romania particularly relevant as an example of a gradual transition towards a broader assessment framework rather than a complete departure from metric-based evaluation.



Relevance of the Romanian Model for Research Assessment Reform

The relevance of the Romanian model lies primarily in its combination of continuity and gradual change. Quantitative indicators, explicit scoring and performance classification remain important components of institutional assessment, while the framework extends evaluation to dimensions that are less visible in predominantly output-based systems, including institutional capacity, socio-economic impact, human resources and strategic development. Self-assessment and external expert review further broaden the evidence considered in the evaluation process.

For Serbia, this experience is particularly relevant because it illustrates how the scope of research assessment can be broadened without requiring an immediate abandonment of an established indicator-based framework. The Romanian case shows that institutional strategy, governance, societal relevance and other dimensions of research performance can be incorporated progressively alongside quantitative measures. At the same time, its continued reliance on scoring and performance categories demonstrates the limitations of such a transitional approach when compared with systems in which contextualised qualitative expert judgement has a more central role.

The Romanian experience therefore provides a useful reference point for considering the sequencing and feasibility of research assessment reform in Serbia. It suggests that the transition towards broader and more context-sensitive assessment can be gradual, while also highlighting the importance of ensuring that the introduction of qualitative dimensions represents a substantive change in assessment practice rather than simply an expansion of the number of criteria included in a scoring system.

2.4. Bridging the Gap: Lessons from Europe and Opportunities for Reform in Serbia

The comparative analysis shows that the gap between current research assessment practices in Serbia and the European approaches examined within ARISE is not merely technical,

but conceptual and systemic. While the Serbian system remains predominantly indicator-driven and focused on measurable outputs, the French and Romanian cases illustrate different ways of broadening the scope and function of research assessment. France demonstrates the potential of a qualitative, peer-review-based and context-sensitive approach centred on research units, while Romania illustrates a more gradual pathway in which broader dimensions of institutional performance, self-assessment and external peer review are incorporated alongside quantitative indicators and scoring.

The comparison identifies several structural gaps in the Serbian system, but also points to reform opportunities that do not necessarily require an immediate replacement of the existing national framework. These concern not only the methods used for assessment, but also its level, evidence base and purpose: the introduction of assessment at the level of research units; greater use of narrative self-assessment and selected evidence of research contributions; a stronger role for external peer review and contextualised expert judgement; broader consideration of societal contribution, institutional capacity and strategy; and a more developmental use of assessment results.

The table below summarises the main differences identified through the comparative analysis and the corresponding opportunities for the further development of research assessment in Serbia.

Dimension	Serbia (Current)	France (Hcéres)	Romania	Gap for Serbia	Reform Opportunity
Core evaluation logic	Predominantly quantitative and indicator-driven	Qualitative, peer-review-based and context-sensitive	Hybrid quantitative – qualitative, score-based	Predominance of indicator-based assessment	Strengthen qualitative assessment and contextualized expert judgement
Role of peer review	Limited in institutional research assessment	Central to research unit assessment	External peer review integrated into a structured scoring process	Limited role of external expert judgement	Introduce structured external peer review
Unit of assessment	Institutional assessment exists, but researcher assessment remains strongly individualized	Research units	Research organizations	Limited assessment of collective research performance at research-unit level	Introduce research-unit-level assessment
Self-assessment and narrative evidence	Limited and not systematically embedded in the	Structured narrative self-assessment is central	Structured institutional self-assessment	Limited use of reflective and contextual evidence	Introduce structured narrative self-assessment

	national framework				
Use of quantitative indicators	Central and often determinative	Supporting and contextual evidence	Substantial role through indicators and scoring	Indicators tend to substitute for contextualized expert judgement	Reposition indicators as evidence supporting expert judgement
Societal contribution and impact	Limited consideration	Explicitly considered within research unit assessment	Explicit assessment dimension	Insufficient recognition of contributions beyond academic outputs	Broaden assessment to include societal contributions and impact
Institutional strategy and development	Limited role in assessment	Assessed in relation to unit objectives, trajectory and development	Five-year strategic development plan explicitly assessed	Weak link between assessment and strategic development	Connect assessment with institutional strategy and future development
Function of assessment	Predominantly administrative	Developmental, supporting reflection and improvement	Performance-oriented, with classification and funding consequences	Limited use of assessment for learning and development	Strengthen the developmental function of assessment

The comparative evidence suggests that the principal challenge for Serbia is not the absence of assessment mechanisms, but the limited scope and function of the existing framework. The experiences of France and Romania show that research assessment can incorporate a broader range of evidence, operate at the level of research units and organisations, and combine self-assessment and external peer review with quantitative information. They also demonstrate that different pathways towards broader research assessment are possible, ranging from a predominantly qualitative and context-sensitive model to a more gradual integration of qualitative dimensions within a structured, indicator-based system. These insights provided an important reference point for the development of the ARISE Research Assessment Model and its subsequent pilot implementation at IES.

3. From Researchers' Perspectives to the ARISE Model

KEY INSIGHTS

- **IES researchers support a broader understanding of research quality and contribution,** extending beyond publication counts and other quantitative indicators.
- **Research contributions are multidimensional,** encompassing projects, mentoring, societal and policy impact, collaboration and contributions to the research environment.
- **Researchers favour a combination of qualitative assessment and responsible use of metrics,** with external peer review and contextual evidence supporting informed expert judgement.
- **These findings directly informed the ARISE model,** translating researchers' perspectives into a multidimensional, peer-review-based and development-oriented assessment approach.

The development of the ARISE Research Assessment Model was informed not only by the analysis of existing research assessment practices in Serbia and selected European countries, but also by empirical evidence collected directly from the IES research community. This evidence was generated through an Institute-wide survey and two focus groups conducted as part of the ARISE project.

The survey, conducted in December 2025, involved 42 IES employees representing all six research departments and administrative staff. The questionnaire examined researchers' perceptions of research assessment at both national and institutional levels, including their understanding and experience of existing assessment practices, the extent to which current approaches recognise different forms of research contribution and career paths, and attitudes towards incorporating broader qualitative dimensions into research assessment. It also explored perceptions of assessment practices within IES and provided respondents with an opportunity to identify contributions and aspects of research work that they considered insufficiently recognised by existing assessment arrangements.

Two focus groups, held in January 2026 and involving 16 participants altogether, complemented the survey by enabling more in-depth discussion of the core research values at IES and possible approaches to research assessment. Participants represented different career stages, departmental affiliations and lengths of employment at IES, allowing a range of institutional perspectives and experiences to be reflected in the discussions. Both focus groups were conducted as moderated and structured group discussions, with recurring themes and shared perspectives identified and organised into thematic categories.

The combination of survey evidence and participatory discussion provided an empirical basis for identifying researchers' expectations, concerns and priorities and for translating them into the principles, dimensions and assessment approach subsequently incorporated into the ARISE model.

3.1. Beyond Metrics: Broadening the Understanding of Research Quality and Contribution

A central finding of the survey was that IES researchers recognise the role of quantitative indicators in research assessment but do not consider them sufficient for capturing research quality and contribution in their full complexity. The results point to substantial support for broadening the scope of assessment beyond conventional bibliometric and output-based

measures. In particular, 83.34% of respondents supported the inclusion of qualitative elements, such as research impact, relevance and strategic contribution, in national-level institutional assessment.

At the same time, the survey revealed an important distinction between the transparency of the existing assessment system and the breadth of what it recognises. Most respondents reported that they understood the national methodology governing academic career progression, and 88.1% considered the criteria for academic promotion clear and transparent. Yet 61.91% considered that the national assessment system does not sufficiently recognise different career paths and forms of research contribution. The issue identified by researchers therefore appears to concern not simply how clearly assessment rules are defined, but what is recognised as relevant evidence of research quality and contribution within those rules.

The focus-group discussions helped to explain this pattern. Participants did not argue for abandoning publications, citations, projects or other quantitative evidence. Instead, they repeatedly emphasised that such indicators provide only part of the picture. Publication counts and journal categories, for example, do not in themselves explain the significance of a research contribution, while citation counts do not necessarily capture its longer-term relevance. Similarly, the number of projects provides limited information about their contribution to knowledge development, research capacity, partnerships or the strategic development of a research unit.

3.2. Making the Full Range of Research Contributions Visible

The survey provided clear evidence that IES researchers perceive an important gap between what is formally recognised by the existing research assessment system and the broader range of contributions that they consider relevant to research quality and development. While respondents generally demonstrated a good understanding of the formal criteria governing research assessment and career progression, 61.91% considered that the national assessment system does not sufficiently recognise different career paths and forms of research contribution. This suggests that the perceived limitation of the current system lies not primarily in the clarity of its rules, but in the scope of what those rules are able to recognise and value.

This finding was reinforced by responses to the survey questions concerning broader approaches to research assessment. A large majority of respondents (83.34%) supported the inclusion of qualitative elements, such as research impact, relevance and strategic contribution, in institutional assessment at national level. The survey therefore indicates substantial support within IES for broadening assessment beyond conventional

output-based measures, while retaining quantitative evidence as part of a more comprehensive assessment framework.

The open-ended survey responses provide further insight into what such broadening would need to encompass. When respondents were given the opportunity to identify contributions that are insufficiently recognised by existing assessment arrangements, they pointed to mentoring and the development of younger researchers, leadership and team development, contributions to institutional development, preparation and coordination of research projects, international cooperation, collaboration with the business sector, and the development and transfer of new knowledge and methodologies. These responses reveal an understanding of research contribution that extends beyond final outputs to include activities that build research capacity, enable collective achievement and contribute to the longer-term development of the research environment.

The focus-group discussions provided an opportunity to explore these survey findings in greater depth. Participants similarly drew attention to contributions that often remain less visible in formal assessment, including internal knowledge transfer, seminars, peer review and editorial work, support for colleagues and early-career researchers, and longer-term investment in research and academic infrastructure. At the same time, the discussions introduced an important qualification: broadening assessment should not result simply in expanding the number of activities that are counted. Participants consistently emphasised the need to consider the quality, relevance and effects of such contributions rather than their frequency alone.

Mentoring illustrates this distinction particularly clearly. The survey identified mentoring and researcher development among the contributions that deserve greater recognition, while the focus groups showed why formal indicators alone may be insufficient to capture them. Participants distinguished formally recorded supervision from a wider range of developmental activities, including informal guidance, knowledge transfer, work with interns and support for early-career researchers. They also emphasised that assessment should take account of different career stages, since researchers at different stages of professional development contribute in different ways and operate under different opportunities and expectations.

The same pattern emerged in relation to teamwork and collaboration. Survey responses pointed to team development, international cooperation and other collective contributions as areas requiring greater recognition. Focus-group participants further stressed that visible research outputs often represent the final result of a much broader collaborative process involving project work, methodological exchange, networking and differentiated roles within research teams. Their concern was therefore not simply to attach additional numerical values to collaboration or authorship positions, but to create space for explaining the nature and significance of different contributions to collective research achievements.

Societal and policy contribution represented another important area in which respondents supported a broader understanding of research performance. The survey evidence favouring the inclusion of impact, relevance and strategic contribution was further elaborated in the focus groups, where participants discussed engagement in public policy, expert and working groups, cooperation with external stakeholders and public communication of research-based expertise. They emphasised that such engagement should be assessed through the significance and effects of the contribution rather than through the simple number of activities undertaken. Public and media engagement, in particular, was considered valuable when grounded in relevant expertise, research evidence and academic responsibility, rather than treated as visibility in itself.

The survey and focus-group evidence therefore pointed in the same direction: a meaningful broadening of research assessment requires recognition of a wider range of contributions, but should avoid converting those contributions into an expanded checklist of quantitative indicators. For the development of ARISE, this finding was particularly important. It supported the creation of assessment dimensions capable of capturing research quality alongside researcher development, collaboration, institutional capacity, societal engagement and other contributions, while allowing their significance to be explained and assessed in context.

3.3. Context Matters: Research Profiles, Career Stages and Institutional Diversity

The survey also pointed to the importance of context in research assessment. Researchers' responses indicated that a uniform assessment approach may not adequately reflect differences in research profiles, career trajectories and forms of contribution. This concern was

particularly evident in perceptions of the national system, where 61.91% of respondents considered that different career paths and forms of research contribution are not sufficiently recognised. The findings therefore raised a broader question for the development of the ARISE model: how to establish a common assessment framework while allowing meaningful differences between researchers and research units to be considered.

Career stage emerged as one important aspect of this issue. Research contributions and opportunities develop over time, and the expectations applied to early-career researchers cannot always be meaningfully interpreted in the same way as those applied to established researchers. The focus-group discussions reinforced this point, particularly in relation to researcher development, mentoring and networking. Participants noted that early-career researchers are often still developing their research profiles, establishing areas of specialisation and building professional networks, while more experienced researchers may contribute through mentoring, leadership, team development and the creation of opportunities for others.

Contextual differences were also considered relevant at the level of research departments. Participants in the focus groups emphasised that departments differ in their thematic profiles, research dynamics, project activities and organisational practices. A fully standardised set of indicators could therefore privilege certain forms of research activity while failing to reflect others adequately. This concern led to explicit discussion of whether assessment should be based on a single institution-wide framework or allow departments greater autonomy in defining evidence relevant to their specific profiles.

The discussions favoured a combination of the two approaches. A common institutional framework was considered important for ensuring coherence, transparency and a shared understanding of research quality across IES, while sufficient flexibility was needed to allow departments to explain their specific research profiles, priorities and contributions. Rather than applying different standards to different units, contextualisation was understood as enabling comparable dimensions of research performance to be interpreted in relation to the circumstances in which research is conducted.

This distinction became important for the design of ARISE. The evidence did not support either a fully standardised indicator-based model or an entirely decentralised approach in which each department would define its own assessment system. Instead, it pointed towards a shared institutional framework combined with contextualised evidence and expert judgement. This provided an empirical basis for designing a model with common assessment dimensions and procedures across IES while allowing departments to explain their achievements, challenges and development trajectories in relation to their own research profiles and operating contexts.

3.4. Research Assessment as a Developmental and Learning Process

The evidence collected at IES also raised a more fundamental question concerning the purpose of research assessment. Survey responses showed broad support for moving beyond a narrow assessment of measurable outputs and incorporating qualitative dimensions related to relevance, impact and strategic contribution. This implies a role for assessment that extends beyond verifying compliance with formal performance requirements and creates scope for a better understanding of how research performance can develop over time.

This issue was explored more explicitly in the focus groups. Participants emphasised that research assessment should have a developmental rather than primarily ranking or punitive function. At the level of research departments, its purpose should not be to identify “winners” and

"losers", but to provide meaningful feedback on strengths, areas for improvement and future development and to enable research units to reflect on their progress over time.

This developmental perspective also shaped discussions of research projects and other activities whose value may unfold over time. Participants noted, for example, that the significance of project activity cannot always be inferred from the number of grants obtained. Even an unsuccessful proposal may contribute to the development of a new research topic, formation of a team, creation of partnerships or preparation for future collaboration. Similarly, the value of internal research activities lies not in how often they occur, but in whether they contribute to knowledge development, stronger research capacities, improved coordination or institutional development.

The discussions therefore pointed towards assessment as a process that should not only establish what has been achieved, but also help explain how those achievements were produced, what capacities they created and what they imply for future development. This requires feedback that is sufficiently substantive to support reflection at both departmental and institutional levels and to identify areas in which existing strengths could be consolidated or further developed.

For the ARISE model, this represented an important shift in assessment logic. The purpose of assessment was defined not primarily as ranking research units or determining whether predefined numerical thresholds had been met, but as generating structured evidence and expert feedback that could support institutional learning and strategic development. This developmental orientation subsequently became one of the central principles guiding the design and pilot implementation of the model.

3.5. Combining Narrative Evidence, Quantitative Indicators and External Peer Review

Support for broadening research assessment was accompanied by an awareness of the challenges involved in implementing qualitative approaches in a fair, transparent and consistent manner. The survey showed substantial support for incorporating qualitative dimensions into research assessment, but the participatory discussions also demonstrated that moving beyond predominantly quantitative criteria raises important questions about how different forms of evidence should be presented, interpreted and assessed.

The focus groups provided a setting in which these questions could be explored in greater depth. Participants expressed support for an approach in which narrative evidence is used to explain the quality, relevance, impact and strategic significance of selected research contributions, while quantitative indicators provide supporting information and context. The discussion therefore did not suggest replacing quantitative assessment with unrestricted qualitative judgement. Rather, it pointed towards a combination in which different forms of evidence perform complementary functions.

Narrative assessment was considered particularly relevant because it enables research units to explain why selected achievements matter. In discussing research outputs, for example, participants emphasised the importance of explaining the problem addressed, the contribution made to the field, its longer-term significance and its relationship to wider research priorities. Similar reasoning was applied to projects, societal engagement and other contributions: the relevant question was not simply how many activities had taken place, but what they had contributed to research development, capacity building, partnerships, policy or society.

At the same time, participants raised concerns about the mechanisms required to ensure fairness, consistency and transparency when qualitative evidence is assessed. The existing regulatory framework and its predefined criteria remained an important reference point in the discussions, illustrating both researchers' familiarity with clearly specified assessment rules and a degree of caution towards approaches that rely more heavily on qualitative judgement. The discussions therefore highlighted the need for qualitative assessment to be sufficiently structured and transparent to build confidence among those being assessed.

External peer review emerged as an important component of such an approach. Independent experts were seen as able to examine the evidence provided by research units, ask questions, discuss the significance of selected contributions and make judgements that take disciplinary and institutional context into account. The focus groups also considered narrative CVs and research portfolios as possible mechanisms for presenting a broader range of contributions, while emphasising the need for clear procedures to ensure that qualitative assessment does not become arbitrary or opaque.

The evidence therefore pointed towards a structured combination of narrative self-assessment, supporting quantitative information and independent expert judgement. For the development of ARISE, this was a particularly consequential finding. It provided support not only for the principle of qualitative peer review, but also for the mechanisms needed to operationalise it: a common assessment framework, structured self-assessment, clearly defined dimensions, transparent assessment guidance and independent external reviewers. These elements were intended to preserve the contextual and explanatory advantages of qualitative assessment while addressing researchers' concerns regarding consistency, fairness and transparency.

3.6. Translating Evidence into the ARISE Model Design

The survey and focus groups were not conducted as stand-alone exercises, but as integral components of the development of the ARISE Research Assessment Model. Their findings complemented the analysis of the Serbian research assessment system and the comparative examination of European practices by providing evidence on how research assessment is perceived by those working within the institutional context in which the model would subsequently be piloted.

Several findings were particularly consequential for the design of the model. Researchers supported broadening assessment beyond conventional quantitative measures, but did not advocate abandoning quantitative evidence. They identified important research contributions that remain insufficiently visible in existing assessment arrangements, emphasised differences in research profiles and career stages, and highlighted the importance of interpreting achievements in context. They also saw value in assessment as a means of providing feedback and supporting development, rather than primarily as a mechanism for ranking. At the same time, participants raised legitimate concerns about fairness, transparency and consistency in qualitative assessment and therefore stressed the need for a sufficiently structured process and independent expert judgement.

These findings were translated into specific design choices within the ARISE model:

Evidence from the IES research community	Response in the ARISE model

Quantitative indicators are relevant but insufficient to capture research quality and contribution.	Qualitative peer review forms the core of assessment, with quantitative information used as supporting and contextual evidence.
Research quality cannot be understood solely through publication counts, journal categories, citations or other measurable outputs.	Departments provide narrative self-assessment and selected evidence explaining the quality, significance and context of their achievements.
Important contributions such as mentoring, researcher development, collaboration, institutional capacity building and societal engagement are insufficiently visible in existing assessment practices.	The assessment framework extends beyond research outputs to six dimensions encompassing research quality, project engagement, research capacities and careers, societal relevance, research culture and international cooperation.
Different research profiles and career stages require contextual interpretation.	A common institutional framework is applied across departments while allowing achievements and evidence to be interpreted in relation to the profile, objectives and context of each research unit.
Simply expanding the number of indicators would reproduce the limitations of the existing system.	Contributions are assessed through their quality, relevance and significance rather than through mechanical counting of activities.
Assessment should provide meaningful feedback and support improvement rather than primarily rank research units.	The model is developmental in purpose and focuses on identifying strengths, areas for improvement and opportunities for future development.
Qualitative assessment raises concerns regarding subjectivity, fairness, consistency and transparency.	Common assessment dimensions, structured self-assessment templates and peer-review rubrics provide a shared and transparent assessment framework.
Independent expertise is important for interpreting different forms of evidence and their significance.	Assessment is conducted by independent external peer reviewers who examine documentary evidence and engage directly with research units.

The resulting model therefore reflects the interaction of three sources of evidence: the analysis of the Serbian research assessment context, lessons derived from European reform principles and practices, and empirical and participatory evidence generated within IES. European experience provided important reference points, particularly concerning qualitative peer review, self-assessment and research-unit-level evaluation, but these elements were not simply transferred into the ARISE model. They were considered alongside the characteristics of the Serbian system and the views, experiences and concerns expressed by the IES research community.

Attention to disciplinary diversity was embedded in the ARISE development process from an early stage, including through the involvement of a multidisciplinary Advisory Board representing different scientific fields. The cross-disciplinary survey extended this approach by providing a broader empirical basis for examining whether the principles underlying ARISE resonate across different research communities in Serbia and where disciplinary adaptation may be required.

In this sense, the ARISE model was developed as a context-sensitive and evidence-based response to the assessment challenges identified through the project. The survey and focus groups helped determine not only *what* should be recognised within a broader assessment framework, but also *how* assessment should be organised to combine contextual judgement with sufficient structure, transparency and consistency. The model presented in the following section represents the operationalisation of these findings into a coherent research assessment framework that could be piloted in practice.

4. The ARISE Research Assessment Model

ARISE MODEL AT A GLANCE

Model element	ARISE approach
Unit of assessment	Research department / research unit
Primary assessment approach	Qualitative, peer-review-based and context-sensitive assessment
Purpose	Reflection, institutional learning and strategic development
Self-assessment	Structured narrative self-assessment supported by selected evidence
External assessment	Two independent external peer reviewers
Evidence base	Self-assessment report, selected supporting evidence, relevant quantitative indicators, site visit and interviews
Assessment framework	Six dimensions covering research quality, project engagement, research capacities and careers, societal engagement, research culture and international cooperation
Role of quantitative indicators	Supporting and contextual evidence; indicators do not determine assessment outcomes
Assessment method	Contextualized qualitative expert judgement supported by common peer-review rubrics
Expression of assessment	Reasoned narrative assessment and classes of merit across the six dimensions
Weighting	Two core dimensions: 25% each; two supporting dimensions: 15% each; two supporting dimensions: 10% each
Final output	Evaluation report identifying strengths, areas for improvement and recommendations for future development
Use of results	Departmental reflection, institutional dialogue, strategic planning and organizational learning
Not intended for	Ranking departments or automatically determining funding, sanctions or resource allocation

The ARISE Research Assessment Model was developed as the central methodological output of the project and provides a structured framework for qualitative, peer-review-based and context-sensitive assessment of research performance at the level of research units. Developed and piloted at the Institute of Economic Sciences, the model was designed to broaden the scope of research assessment beyond predominantly quantitative approaches.

The model shifts the focus of assessment from the accumulation of measurable outputs towards an informed understanding of research performance in context. It combines structured self-assessment by research departments with external peer review and the responsible use of

quantitative evidence. Research achievements are therefore not assessed in isolation, but interpreted in relation to the department's mission, objectives, resources, disciplinary profile and institutional environment.

A defining feature of the ARISE model is its developmental orientation. Assessment is intended to support reflection, institutional learning and strategic development rather than ranking research units or applying sanctions. The process enables departments to identify strengths, areas for improvement and future priorities, while providing the Institute with evidence that can inform research strategy and institutional development. Quantitative indicators may support this process where relevant, but they do not substitute for qualitative expert judgement.

The model was developed as a pilot framework that could be tested and refined through its implementation at IES, while also providing a basis for considering its future adaptation and broader application. Its design therefore combines methodological rigour with sufficient flexibility to accommodate differences between research units and, potentially, different institutional and disciplinary contexts.



Guiding Principles and Core Values

The ARISE Research Assessment Model is grounded in a set of guiding principles and core values that define both how assessment should be conducted and what dimensions of research performance should be recognised. The principles establish the methodological foundations of the assessment process, while the core values reflect the understanding of research quality and contribution that the model seeks to promote.

The model is based on six guiding principles:

- **Qualitative assessment** as the primary mode of evaluation, with expert judgement at the centre of the assessment process.
- **Peer review conducted by independent external experts** with relevant disciplinary and institutional knowledge.
- **Context-sensitive assessment,** in which research performance is interpreted in relation to the mission, objectives, resources and operating environment of the research unit.
- **Responsible use of quantitative indicators** as supporting and contextual evidence rather than as substitutes for expert judgement.
- **Self-assessment and reflection** as integral parts of the process, enabling research units to examine their achievements, challenges and future development.
- **Developmental orientation,** with assessment intended to support learning, improvement and strategic development rather than ranking or sanctioning research units.

These principles are complemented by six core values that provide the substantive basis for assessing research performance:

1. **Scientific quality and excellence** – the quality, originality, relevance and contribution of research and scholarly outputs.

2. **Strategic relevance and responsible project engagement** – the ability to participate in and develop research projects that are scientifically relevant and aligned with institutional and broader research priorities.
3. **Sustainability of research capacities and careers** – the development of researchers and research capacities, including mentoring, support for early-career researchers and conditions for sustainable research careers.
4. **Societal relevance and engagement** – contributions through which research interacts with and creates value for society, public policy, the economy and other relevant stakeholders.
5. **Collaboration, integrity and inclusive research culture** – research practices that support cooperation, responsible conduct of research, collegiality and an inclusive and supportive research environment.
6. **International openness and cooperation** – engagement in international research networks, partnerships and other forms of cross-border scientific cooperation.

The combination of these principles and values broadens the understanding of research performance beyond publication output alone. It enables assessment to consider scientific achievements alongside the conditions in which research is conducted, the development of research capacities, collaboration and research culture, engagement with society and the international positioning of research units.

Assessment Dimensions

The core values of the ARISE model are operationalised through six assessment dimensions that provide the structure for both departmental self-assessment and external peer review. The dimensions were designed to capture different and complementary aspects of research performance, extending assessment beyond research outputs to include capacity development, societal engagement, research culture and international cooperation.

The model distinguishes between two core dimensions, which reflect the principal research activities of the departments, and four supporting dimensions, which capture the broader environment, contributions and conditions that enable sustainable and socially relevant research performance.

Assessment dimension	Focus of assessment
Scientific quality and excellence of research outputs	Quality, originality, relevance and contribution of research outputs, considered in relation to the department's research profile and objectives.
Strategic relevance and responsible project engagement	Participation in and development of research projects, their scientific and strategic relevance, and the department's capacity to engage responsibly in project-based research.
Sustainability of research capacities and careers	Development of researchers and research capacities, including mentoring, support for early-career researchers, knowledge development and conditions for sustainable research careers.
Societal relevance and engagement	Contributions of research to society, public policy, the economy and other stakeholder

	communities, including different forms of engagement and knowledge exchange.
Collaboration, integrity and inclusive research culture	Internal and external collaboration, responsible research practices, research integrity, collegiality and the development of an inclusive and supportive research environment.
International openness and cooperation	International research collaboration, participation in networks and partnerships, researcher mobility and other forms of international scientific engagement.

The distinction between core and supporting dimensions does not imply that the latter are peripheral to research quality. Rather, it reflects the structure of the model and the relative emphasis assigned to different aspects of research performance. Scientific outputs and project engagement remain central components of research activity, while the supporting dimensions enable the assessment to recognise the wider conditions and contributions through which research quality, sustainability and relevance are developed.

Importantly, the dimensions are not intended to function as a checklist of isolated achievements. They provide a common structure within which departments can explain their performance and development in context, and within which peer reviewers can interpret the evidence presented through self-assessment, supporting documentation and discussion with the research unit.



Assessment Process

The ARISE Research Assessment Model follows a structured process that combines departmental self-assessment, external peer review and institutional reflection. The process is designed to ensure that assessment is evidence-based and systematic while preserving the central role of contextualised expert judgement.

The assessment begins with a structured self-assessment conducted by each research department. Departments reflect on their performance across the six assessment dimensions, identify their most significant achievements and challenges, and provide selected evidence to support their narrative. This stage is intended not only to provide information for external reviewers, but also to encourage collective reflection on the department's development, research environment and future priorities.

Independent external experts are subsequently selected to review the self-assessment reports and supporting evidence. Their assessment combines document review with a site visit and discussions with relevant members of the Institute. These interactions enable reviewers to clarify the evidence presented, better understand the context in which departments operate and explore issues that may not be fully captured in written documentation.

Based on the complete evidence base, the external reviewers assess performance across the six dimensions and prepare an evaluation report. Their conclusions are expected to identify strengths and areas for improvement and to provide recommendations relevant to the future development of the research department and the Institute.

The process concludes with feedback and institutional reflection. Assessment results are intended to inform dialogue between research departments and institutional leadership, contribute to strategic planning and support organisational learning. They are not designed to produce rankings between departments or to serve as an automatic basis for sanctions or resource allocation.

1. **Departmental Self-Assessment**
   Structured narrative self-assessment across the six assessment dimensions, supported by selected evidence



2. **Selection of External Peer Reviewers**
   Independent external experts with relevant disciplinary and institutional expertise



3. **Document Review**
   Review of departmental self-assessment reports and supporting evidence



4. **Site Visit and Interviews**
   Discussions with institutional leadership, researchers, early-career researchers and relevant support staff to contextualise and clarify the evidence



5. **Qualitative Peer Assessment**
   Contextualised expert judgement across the six assessment dimensions, supported by the responsible use of quantitative evidence



6. **Final Evaluation Report**
   Assessment of performance across the six dimensions, identification of strengths and areas for improvement, and recommendations for future development



7. **Feedback, Institutional Reflection and Strategic Development**
   Use of assessment results to support dialogue, organisational learning, strategic planning and future development.

Narrative Self-Assessment and Supporting Evidence

Narrative self-assessment is a central component of the ARISE model. Rather than requiring departments to compile an exhaustive inventory of outputs and activities, the self-assessment process asks them to reflect critically on their performance, identify the achievements and developments they consider most significant, explain their relevance in context, and support their assessment with selected evidence.

For the ARISE pilot, each of the six IES research departments prepared a self-assessment report covering the 2024–2025 assessment period. Reports were limited to ten pages, excluding supporting documentation, encouraging departments to prioritise the evidence most relevant to understanding their performance rather than reproduce comprehensive lists of activities and outputs.

The self-assessment follows a common structure covering the department's profile and strategic orientation and the six dimensions of the ARISE model. Departments are asked to address:

A. Department profile, mission and strategic orientation
B. Scientific quality and excellence of research outputs
C. Strategic relevance and responsible project engagement
D. Sustainability of research capacities and careers
E. Societal relevance and engagement
F. Collaboration, integrity and inclusive research culture International openness and cooperation
G. Overall reflection, key strengths, challenges and future development priorities

Within this structure, departments are encouraged to explain not only what they achieved, but why particular achievements matter in relation to their objectives, research profile and operating context. The narrative therefore provides space to explain developments that may not be visible through standardised indicators alone, including collective achievements, capacity building, mentoring, collaboration, societal engagement and changes in the research environment.

Supporting evidence is used selectively to substantiate the narrative. Depending on the dimension being assessed, this may include selected research outputs, information on projects and collaborations, evidence of mentoring and career development, policy or societal contributions, international activities and other documentation relevant to the claims made in the self-assessment. Quantitative information may also be provided where it contributes to understanding performance, but it is interpreted within the broader narrative and contextual evidence.

The self-assessment thus serves two related purposes. It provides external reviewers with a structured evidence base for their assessment, while also functioning as an internal reflective exercise through which departments can identify strengths, challenges and priorities for future development.

External Peer Review and the Responsible Use of Quantitative Evidence

External peer review constitutes the central evaluative mechanism of the ARISE model. Each research department is assessed by two independent external experts with relevant disciplinary expertise and experience appropriate to the profile of the unit being assessed. Their role is not to verify whether departments have met predetermined quantitative thresholds, but to interpret the evidence presented, assess the quality and significance of research performance in context, and provide an informed qualitative judgement across the six assessment dimensions.

The departmental self-assessment report and supporting documentation form the primary evidence base for peer review. Reviewers consider the achievements selected by the department, the explanations provided for their significance, and the broader institutional and disciplinary context in which the department operates. Documentary evidence is complemented by the site visit and interviews, allowing reviewers to clarify information, explore issues in greater depth and consider perspectives that may not be fully reflected in the written material.

Quantitative and bibliometric indicators may be used as complementary evidence where they contribute to a better understanding of research performance. They may, for example, provide contextual information on patterns of publication, project participation, collaboration or other aspects of research activity. However, no indicator or predefined numerical threshold determines the assessment outcome. Quantitative evidence is interpreted in relation to the department's profile, objectives and context and does not substitute for qualitative expert judgement.

This approach is intended to avoid both the mechanical use of metrics and the opposite risk of unstructured or insufficiently transparent qualitative judgement. Peer reviewers work within a common assessment framework, using the same dimensions, evidence requirements and qualitative rubrics, while retaining the expert discretion necessary to interpret different forms of research contribution in context. Assessment conclusions must therefore be evidence-based, explicitly reasoned and linked to the relevant assessment dimensions.

The combination of structured self-assessment, selected evidence, independent peer review and responsible use of quantitative information is intended to provide a more comprehensive understanding of research performance than can be achieved through indicators alone. It enables the assessment to recognise differences between research units while maintaining a common and transparent framework for evaluative judgement.

Assessment Methodology, Classes of Merit and Weighting

The ARISE model translates qualitative peer-review judgements into a structured assessment through classes of merit assigned to each of the six assessment dimensions. Importantly, the class of merit is determined only after reviewers have considered the complete evidence base and reached a reasoned qualitative judgement. The classes therefore provide a structured expression of expert assessment rather than a substitute for it.

Five classes of merit are used:

Class of merit	Coefficient	General interpretation
Exceptional	1.0	Performance of exceptional quality and significance
Excellent	0.8	Performance of very high quality, with substantial strengths
Standard	0.5	Solid performance meeting expected standards
Sufficient	0.2	Performance demonstrating some relevant achievements but also important limitations
Low relevance / Unacceptable	0.0	Performance showing major weaknesses or insufficient relevance

For each assessment dimension, reviewers assign a class of merit and provide a written justification grounded in the evidence examined. The use of performance classes is intended to improve consistency and transparency across assessments while preserving the primacy of qualitative expert judgement.

The model also assigns different weights to the six assessment dimensions, reflecting their relative role within the overall framework:

Assessment dimension	Weight
Scientific quality and excellence of research outputs	25%
Strategic relevance and responsible project engagement	25%
Sustainability of research capacities and careers	15%
Societal relevance and engagement	15%
Collaboration, integrity and inclusive research culture	10%
International openness and cooperation	10%
Total	**100%**

The two core dimensions, scientific quality and excellence of research outputs and strategic relevance and responsible project engagement, therefore account for half of the overall assessment, while the remaining half recognises dimensions related to research capacity, societal engagement, research culture and international cooperation.

The coefficients and weights provide a common structure for synthesising assessment results across dimensions. However, their role within ARISE is explicitly transitional and supportive. They are applied after qualitative peer assessment and are not intended to transform the model into an indicator-driven scoring system. Nor are the resulting assessments designed to generate rankings between departments or to trigger automatic decisions on funding, sanctions or resource allocation. Their purpose is to make qualitative judgements more transparent and comparable while retaining the explanatory and developmental value of the reviewers' written assessment.



Peer-Review Rubrics

To support consistency, transparency and comparability in qualitative assessment, the ARISE model provides peer-review rubrics for each of the six assessment dimensions. The rubrics serve as a common reference framework for external reviewers, helping them interpret the evidence presented by departments and translate their qualitative judgement into an appropriate class of merit.

The rubrics specify the aspects of performance that reviewers should consider within each dimension and describe how different levels of performance may be distinguished. They cover the quality and significance of research outputs; the relevance and quality of project engagement; the development of research capacities and careers; societal relevance and engagement; collaboration, integrity and research culture; and international openness and cooperation.

The rubrics are not intended to operate as checklists or as an additional system of numerical indicators. Reviewers are not expected to count the presence or absence of individual elements and derive an assessment mechanically. Instead, the descriptors provide reference points for holistic expert judgement, allowing different forms of evidence to be interpreted in relation to the department's profile, objectives and context.

Their role is therefore to introduce structure into qualitative peer review without eliminating expert discretion. By providing a shared evaluative framework, the rubrics support greater consistency between reviewers and departments, while the requirement for written justification ensures that the final class of merit remains grounded in evidence and explicit reasoning.



Site Visit, Final Evaluation Report and Use of Results

The site visit complements the documentary assessment by enabling external reviewers to examine the evidence presented in the self-assessment reports in greater depth and to develop a fuller understanding of the institutional context in which research is conducted. It provides an opportunity to clarify issues arising from the written documentation, explore differences in perspectives and gather additional contextual information relevant to the assessment.

During the visit, reviewers engage with different groups within IES, including institutional leadership, researchers, early-career researchers and relevant support staff. These discussions allow them to consider research performance not only through documented achievements, but also through the experiences and perspectives of those involved in the research environment. The site visit thereby contributes to the triangulation of evidence and strengthens the basis for contextualised expert judgement.

Following the documentary review and site visit, the external reviewers prepare a final evaluation report. The report presents the assessment across the six dimensions of the ARISE model, provides reasoned justification for the assigned classes of merit, identifies key strengths and areas for improvement, and formulates recommendations for future development. The written assessment is an essential part of the outcome: the performance classes provide a structured summary of the reviewers' judgements, while the narrative explains their meaning and the evidence on which they are based.

The results are intended primarily to support reflection, dialogue and strategic development. At departmental level, they provide feedback that can inform future priorities and strengthen research practices and capacities. At institutional level, the assessment provides evidence that can support strategic planning, identify shared strengths and challenges across departments and inform decisions concerning the development of the research environment.

The ARISE model does not use assessment results to rank departments or automatically determine funding, sanctions or resource allocation. Its primary function is developmental: to create a structured process through which research units and institutional leadership can reflect on performance, learn from external

The ARISE Research Assessment Model thus provides an integrated framework in which narrative self-assessment, selected evidence, external peer review, contextualised expert judgement and the responsible use of quantitative information form part of a single assessment process. Its six assessment dimensions broaden the scope of research assessment beyond scientific outputs to include research capacities and careers, societal engagement, research culture and international cooperation, while the common rubrics, classes of merit and weighting system provide structure and transparency to qualitative judgement. The model is designed not as an alternative system of metrics, but as a framework for understanding research performance in context and using assessment to support reflection, institutional learning and strategic development.

5. Pilot Implementation of the ARISE Research Assessment Model at IES

KEY INSIGHTS

- **The ARISE model was successfully operationalised across all six IES research departments,** demonstrating the feasibility of qualitative, research-unit-level assessment in practice.
- **Self-assessment and external peer review generated insights beyond conventional quantitative indicators,** capturing institutional strengths, development needs and diverse research contributions.
- **Post-pilot feedback showed broad researcher confidence in the assessment process and its results,** while also identifying areas for further methodological refinement.
- **The pilot confirmed the value of assessment as a developmental process,** providing evidence and recommendations that can support institutional learning and strategic improvement.

Following its development, the ARISE Research Assessment Model was piloted across all six research departments of the Institute of Economic Sciences between March and June 2026. The pilot provided an opportunity to test the model in a real institutional setting and to examine whether a qualitative, peer-review-based approach could generate meaningful insights into research performance that extend beyond those available through conventional quantitative assessment.

Each department conducted a structured narrative self-assessment using the common ARISE framework and provided selected evidence across the six assessment dimensions. The self-assessment reports were subsequently reviewed by independent external experts and complemented by direct interaction with researchers and institutional representatives during the peer-review process. The same assessment framework was applied across all departments, while allowing their different research profiles, activities and stages of development to be considered in context.

The purpose of the pilot was not to rank departments or compare them on the basis of their final assessments. Rather, it was to test whether the model could support a more comprehensive understanding of research performance, identify strengths and development needs that may not be visible through quantitative indicators alone, and generate evidence that could be used for reflection and strategic development at both departmental and institutional levels.

The external peer-review process produced detailed assessments for each department. For the purposes of this study, the findings are presented only at an aggregate level, focusing on recurring and cross-cutting observations across the Institute. This approach reflects the developmental purpose of the pilot while illustrating the type of institutional insight that can be generated through qualitative, research-unit-level assessment.

5.1. What the External Peer Review Revealed

The external peer review generated a considerably broader picture of research performance than could be derived from publication, citation or project indicators alone. Across

the six departments, reviewers were able to identify recurring institutional strengths, but also development needs related to the conditions in which research is conducted, the distribution and sustainability of research capacities, and the extent to which individual and team achievements are embedded in institutional structures.

Several strengths emerged consistently across the external evaluations. Reviewers highlighted the methodological competence of IES researchers, including the use of contemporary quantitative and qualitative methods and the capacity to work with complex data and analytical approaches. They also recognised the strong societal and policy relevance of research, particularly its engagement with major economic and societal challenges. A collaborative research culture, openness to teamwork and the involvement of early-career researchers in projects, publications and international activities were identified as additional strengths. The evaluations also pointed to the Institute's growing international integration through participation in European projects, COST Actions, international networks and mobility programmes.

At the same time, the peer-review process identified development needs that are less readily captured through conventional research metrics. These included the need to strengthen leadership roles in major international research projects, further improve international publication visibility, formalise mentoring arrangements, and reduce reliance on a relatively small number of researchers carrying substantial project and organisational responsibilities. Reviewers also identified a need for more systematic documentation of the societal and policy influence of research, so that contributions already being made outside academic publication channels can be more clearly demonstrated and communicated.

Importantly, these observations concern not only the volume or quality of research outputs, but the organisational conditions that shape the Institute's longer-term research performance. The external assessments drew attention to how expertise is distributed, how younger researchers are supported, how successful project experience is transferred across the organisation, how international relationships become institutional partnerships, and how research contributions beyond academic publications are recognised and documented. In this sense, the pilot demonstrated the capacity of research-unit-level peer review to generate evidence about research performance that is simultaneously scientific, organisational and developmental.

5.2. From Individual and Team Strengths to Institutional Capacity

A particularly important cross-cutting insight emerging from the external evaluations concerns the relationship between individual and team achievements and longer-term institutional capacity. Reviewers recognised substantial expertise, successful research teams, strong project experience and growing international engagement across IES. At the same time, their assessments indicated that the sustainability and further development of these strengths depend on the extent to which they are embedded in institutional structures and practices.

This was reflected in several recurring observations. Successful participation in international projects creates opportunities for the Institute to assume stronger leadership roles in future collaborations; mentoring practices could be strengthened through more systematic institutional arrangements; knowledge and networks developed through successful projects and mobility should be shared more widely across the organisation; and international contacts could increasingly be transformed into longer-term institutional partnerships. Similarly, reducing the

concentration of project and organisational responsibilities among a relatively small number of researchers would contribute to broader capacity building and greater organisational resilience.

The external review also highlighted the importance of connecting individual research activities and project-based achievements with longer-term research agendas. Greater coordination across departments and the consolidation of complementary expertise could strengthen the Institute's capacity to develop joint research initiatives, publications and internationally visible areas of specialisation. The findings therefore point to institutionalisation not as a replacement for individual initiative, but as a means of sustaining, connecting and amplifying existing strengths.

This cross-cutting finding illustrates an important developmental function of qualitative research assessment. Rather than establishing only whether predefined performance thresholds have been reached, the assessment can reveal how research capacity is organised, where it is concentrated, how effectively knowledge and experience are transferred, and which institutional conditions may support future development. The pilot therefore demonstrated how peer-review-based assessment can generate evidence that is directly relevant to institutional reflection and strategic planning.

5.3. Researchers' Experience and Perceptions of the ARISE Pilot

Following the completion of the pilot assessment, a post-pilot survey was conducted between 10 and 24 July 2026 to examine researchers' experiences with the development and implementation of the ARISE model. The survey was completed by 30 respondents representing all six IES research departments and different career stages and levels of research experience. It covered the process through which the model had been developed, the relevance and comprehensiveness of its assessment dimensions, experience with departmental self-assessment and external peer review, the usefulness of the final assessment reports, and researchers' overall confidence in the assessment findings and views on the model's potential applicability beyond IES.

The results indicate a generally high level of acceptance of both the ARISE model and the process through which it was developed. Across the survey items concerning model development, positive responses ranged from 73.3% to 86.7%. The open-ended responses help explain this assessment. Respondents particularly valued the participatory character of the process, the involvement of researchers in discussions about assessment, the opportunity to exchange different perspectives, and the shift towards a broader qualitative understanding of research performance. Several respondents also highlighted the value of moving the focus from individual researchers towards research units and of considering departmental performance from a perspective different from that provided by conventional quantitative assessment.

At the same time, the survey identified opportunities to strengthen the participatory process itself. Respondents suggested more continuous communication during model development, additional time for discussion and reflection, clearer and more detailed guidance, and further opportunities for researchers to contribute to the refinement of the framework. These observations do not challenge the participatory approach; rather, they suggest that participation should be treated as a continuing component of model development and implementation rather than as a one-off consultation exercise.

The assessment framework itself received particularly strong support. Across the five survey items addressing the design and relevance of the model, positive responses ranged from

76.7% to 96.7%. Most notably, 90% of respondents agreed that the assessment dimensions covered the most important aspects of the work of IES research departments. This finding was reinforced by a separate question asking whether any important area of departmental activity had remained insufficiently covered: only three respondents (10%) identified such a gap.

The responses nevertheless provide useful indications for further refinement. While 76.7% of respondents did not consider any existing assessment dimension to require improvement, seven respondents identified areas in which greater precision or adjustment could be useful. These included the organisation of departmental work, research projects, team formation, interdepartmental and multidisciplinary collaboration, publication practices and the relative weight given to scientific excellence. One response also drew attention to disciplinary differences in the relevance of societal, policy and economic contribution, indicating that such dimensions need to be interpreted in relation to the profile of the research field rather than applied mechanically across all areas of research. Rather than pointing to a need for fundamental redesign, these responses reinforce the importance of contextual sensitivity and continued refinement of the assessment framework.

The experience of departmental self-assessment provides further evidence of the model's developmental potential. The self-assessment component was evaluated positively across all three survey items addressing this stage of the process. Ninety per cent of respondents gave a positive assessment on one of the items, while 86.7% considered the guidance for preparing the self-assessment sufficiently clear and detailed. Importantly, 86.7% also agreed that the process encouraged discussion among department members about their department's strengths and other relevant aspects of its research performance. This suggests that self-assessment functioned not only as a mechanism for supplying information to external evaluators, but also as an opportunity for collective reflection within research units.

This reflective function is particularly relevant to the developmental logic of the ARISE model. Preparing a structured account of departmental performance required research units to consider their achievements across a broader set of dimensions, identify evidence supporting those achievements and reflect collectively on areas of strength and potential development. The survey results therefore indicate that narrative self-assessment can itself generate value before external evaluation takes place, by creating a structured opportunity for departments to examine their research profile and discuss aspects of performance that may receive limited attention in routine reporting based primarily on quantitative indicators.

Researchers also evaluated the external peer-review stage positively, although the results and open-ended responses reveal more clearly where implementation could be strengthened. Across the three survey items addressing external peer review, positive responses ranged from 76.7% to 80%. This indicates substantial support for the role of external expert judgement within the model, while also showing somewhat greater variation in researchers' experiences of this stage than in their assessment of some other components of the pilot.

The open-ended responses are particularly informative in explaining this variation. A recurring issue concerned the selection and matching of external reviewers. Respondents emphasised the importance of appointing experts whose disciplinary and subdisciplinary expertise closely corresponds to the research profile of the unit being assessed and who possess strong research credentials and international experience. The findings suggest that the credibility and usefulness of qualitative peer review depend not only on the assessment framework itself, but also on the reviewers' capacity to understand the scientific field, institutional context and specific characteristics of the research unit.

Respondents also identified opportunities to strengthen interaction between reviewers and departments. Suggestions included allowing more time for discussion and clarification during the evaluation process, creating a stronger feedback loop between reviewers and the research unit, and providing more specific and actionable recommendations. Some respondents also suggested considering a larger number of reviewers. These observations are particularly valuable as implementation lessons from the pilot: they indicate that the quality of peer review depends on both appropriate reviewer selection and sufficient opportunities for informed dialogue between evaluators and those being assessed.

The pilot experience therefore supports external peer review as a central component of the ARISE approach, while also identifying conditions that are important for its effective implementation. Careful reviewer selection and matching, adequate time for interaction, and feedback that is sufficiently specific to support subsequent action emerge as areas requiring particular attention in future assessment cycles.

The final assessment reports were among the most positively evaluated components of the pilot. All 30 respondents (100%) gave a positive response when asked whether the final report was clear and understandable, while 86.7% responded positively to the survey item concerning the value of the final assessment report. These results are particularly relevant to the developmental purpose of the ARISE model, since the usefulness of qualitative assessment depends not only on the quality of the evaluation process but also on whether its conclusions can be clearly communicated and subsequently used by the research units being assessed.

The open-ended responses provide further insight into how the reports could become even more useful. While their overall structure and clarity were viewed positively, respondents called for recommendations that are more specific, prioritised and actionable. Suggestions included clearer identification of the steps that departments could take following the assessment and, where appropriate, examples of good practice that could help translate recommendations into concrete improvements. The findings therefore suggest that the value of the final report lies not simply in documenting an assessment outcome, but in its capacity to provide research units with meaningful guidance for future development.

The survey also provides important evidence regarding researchers' confidence in the results produced through the ARISE model. Twenty-five of the 30 respondents (83.3%) agreed that they trusted the assessment findings obtained through the model, while four respondents were neutral and one mostly disagreed. Although the results do not indicate complete consensus, they demonstrate a substantial level of confidence in the findings generated through a newly developed and first-time piloted assessment approach. This is particularly significant given the model's reliance on qualitative evidence and expert judgement and the concerns regarding fairness and consistency that had been expressed during its development.

Researchers also expressed substantial support for the potential application of the ARISE approach beyond IES. Twenty-four respondents (80%) agreed that, with appropriate adaptations, the model could be applied in other research organisations in Serbia. Three respondents were neutral and three mostly disagreed. The qualification concerning adaptation is important. The survey does not support the mechanical replication of a single institutional model across different research organisations; rather, it indicates that the underlying approach is perceived as potentially transferable provided that differences in disciplinary profiles, institutional missions and organisational contexts are appropriately considered.

The open-ended responses reinforce this interpretation and identify several issues for future development. Suggestions included reconsidering the relative weight assigned to particular assessment dimensions, ensuring a sufficiently prominent role for scientific excellence, using a longer assessment period, such as four or five years, and maintaining continued dialogue with the research community as the model evolves. These responses are valuable precisely because they show that positive acceptance of the pilot was accompanied by critical reflection on how the model could be strengthened through subsequent assessment cycles.

Overall, the post-pilot survey provides evidence across several complementary aspects of the ARISE experience. Researchers evaluated positively the participatory process through which the model was developed, the scope and relevance of its assessment dimensions, the self-assessment process, external peer review and the final assessment reports. The survey also indicates substantial trust in the assessment findings and support for the potential wider applicability of the approach. At the same time, it identifies concrete areas for refinement, particularly in relation to reviewer selection and matching, interaction between reviewers and research units, the specificity and actionability of recommendations, contextual interpretation of assessment dimensions and continued participation of the research community in further model development. The pilot therefore provided not only evidence that the model can be operationalised in an institutional setting, but also an empirical basis for its further refinement and potential adaptation to other research organisations.

The two strands of evidence generated through the pilot provide complementary perspectives on the functioning and value of the ARISE model. The external peer-review findings demonstrate the model's capacity to generate contextualised insights into research quality, organisational capacities and development needs that extend beyond those visible through conventional quantitative indicators. The post-pilot survey, in turn, shows that researchers largely regarded the model and its main components as relevant, credible and useful, while also identifying concrete areas for further refinement. The pilot therefore served both as a test of the model in practice and as a learning process for its further development. It provided evidence that a qualitative, peer-review-based and context-sensitive approach can be operationalised at research-unit level and can generate information that supports reflection and institutional development, while also demonstrating that the model should continue to evolve through experience, feedback and adaptation to different research and institutional contexts.

6. Extending the Evidence Base Beyond IES

The development and pilot implementation of the ARISE model at IES provided evidence on the feasibility and value of a qualitative, peer-review-based and context-sensitive approach within a social science research institute. However, considering its potential relevance beyond IES requires attention to the diversity of research cultures, outputs and forms of contribution across scientific fields. Research quality may be expressed differently across disciplines, and the relative importance of particular outputs, activities and pathways to scientific and societal contribution may vary considerably.

To explore these differences and identify areas of convergence, the ARISE project extended its empirical evidence base beyond IES through a cross-disciplinary survey conducted in May 2026. The survey involved 162 researchers from four broad scientific fields in Serbia: Humanities (30 respondents), Social Sciences (64), Engineering, Technological and Biotechnical Sciences (30), and Natural Sciences and Mathematics (38). It explored researchers' perceptions of existing research assessment practices and their views on the types of contributions that should receive greater recognition in a broader assessment framework. It also examined preferences concerning the balance between quantitative indicators and qualitative assessment and provided respondents with opportunities to identify field-specific outputs and contributions that may not be adequately captured by conventional assessment approaches.

The purpose was not to test whether a single assessment model could be uniformly applied across disciplines, but to examine whether common principles and dimensions of broader research assessment could be identified while also determining where disciplinary specificities would require contextual interpretation and adaptation. The inclusion of researchers from different scientific fields therefore made it possible to examine both common patterns and disciplinary variation. Pasted text

The sample was not designed to be statistically representative of the entire Serbian research community, and the findings should therefore be interpreted as exploratory evidence rather than as nationally representative estimates. Nevertheless, the participation of researchers from all four broad scientific fields provides a valuable basis for examining cross-disciplinary patterns and for informing further testing and refinement of the ARISE approach beyond IES.

6.1. A Shared Demand for Broader Research Assessment

Despite differences in research cultures and disciplinary practices, the survey reveals substantial convergence across the four scientific fields regarding the need for a broader understanding of research quality and contribution. Researchers did not generally question the relevance of quantitative indicators, but the findings indicate that such indicators are not considered sufficient on their own to capture the diversity and significance of research performance. Across the fields, respondents expressed support for approaches that retain relevant quantitative evidence while creating greater space for qualitative assessment and contextual interpretation.

Scientific publications remained central to perceptions of research quality across all four fields. This is an important finding, as it shows that support for broader assessment should not be interpreted as a move away from scientific excellence or from the recognition of high-quality research outputs. Rather, respondents' preferences point towards a broader understanding of excellence in which publications remain a core component but are considered alongside other contributions and interpreted in relation to their quality, significance and disciplinary context.

At the same time, several forms of contribution emerged repeatedly across disciplinary boundaries. Interdisciplinary research was among the highly valued contributions across the four fields, while project leadership, mentorship, research leadership and management, and citations also featured prominently in researchers' responses. Although their relative importance varied between fields, their recurrence indicates that researchers understand research performance as encompassing not only the production of scientific outputs, but also contributions to the development of research agendas, people, teams and research capacity.

The findings also point to a shared interest in recognising the wider significance of research. The specific forms through which this significance is expressed differ across fields, ranging from policy and societal contribution to technological development, innovation and other forms of knowledge application. What is common is the view that research assessment should be capable of recognising contributions whose value cannot always be adequately represented through publication and citation metrics alone.

This cross-disciplinary convergence is particularly relevant to the broader applicability of the principles underlying ARISE. It suggests that the case for broadening research assessment is not confined to the social sciences or to the institutional context of IES. At the same time, convergence around broad principles does not imply that the same evidence, expectations or relative emphasis should apply uniformly across disciplines. The survey therefore provides support for a shared assessment logic while also pointing to the need for disciplinary context to shape how different forms of research contribution are evidenced and interpreted.

6.2. One Framework, Different Disciplinary Pathways

The areas of convergence identified across the four scientific fields coexist with meaningful differences in how research quality and contribution are understood and expressed. These differences reflect distinct research cultures, publication practices, types of outputs, relationships with external stakeholders and pathways through which research generates scientific, societal, policy or economic value. The survey therefore provides an important caution against interpreting support for broader research assessment as support for a uniform set of expectations across all disciplines.

Researchers in the Humanities expressed particularly strong reservations about assessment approaches dominated by quantitative indicators and demonstrated a clear preference for greater use of qualitative judgement. Their responses placed particular emphasis on the quality and significance of research outputs, interdisciplinarity, mentorship and contributions that are difficult to capture through conventional bibliometric measures. The findings indicate that assessment in the Humanities requires sufficient space for expert interpretation of the content, originality and significance of research rather than relying predominantly on numerical proxies for quality.

In the Social Sciences, respondents similarly supported a broader approach to assessment, while placing particular emphasis on the relationship between research and society. Alongside scientific publications and other academic contributions, policy impact and wider societal contribution emerged as particularly relevant. This reflects the multiple pathways through which social science research may generate value, including informing public policy, contributing expertise to public debate and institutions, and addressing societal and economic challenges.

Researchers in Engineering, Technological and Biotechnical Sciences showed comparatively greater acceptance of quantitative indicators, but their responses also demonstrated the limitations of assessment frameworks centred primarily on publications and citations. Patents, technical solutions, innovation, project-related achievements and cooperation with industry represent important forms of contribution in these fields. The findings therefore point not towards abandoning quantitative evidence, but towards broadening the types of outputs and contributions recognised and assessing them in relation to the specific research and innovation pathways characteristic of these disciplines.

A similar combination was evident in Natural Sciences and Mathematics. Respondents placed considerable value on objective and measurable evidence of research performance while also recognising the importance of qualitative interpretation and a wider range of research contributions. Research quality, interdisciplinarity, mentorship and other contributions extending beyond conventional output measures featured among the relevant priorities. The results suggest that maintaining robust quantitative evidence and introducing broader qualitative assessment should not be understood as competing alternatives, but as complementary elements of a more comprehensive assessment approach.

The differences across scientific fields become particularly visible when the contributions most frequently prioritised by respondents are compared. While several priorities recur across fields, their relative prominence and the presence of field-specific contributions illustrate different disciplinary understandings of research performance.

Scientific field	Highest-ranked contributions and assessment priorities
Humanities	Scientific publications; interdisciplinarity; mentorship; research leadership and management; project leadership
Social Sciences	Scientific publications; interdisciplinarity; policy impact and societal contribution; project leadership; research leadership and management
Engineering, Technological and Biotechnical Sciences	Scientific publications; patents and technical solutions; innovation; project leadership; cooperation with industry
Natural Sciences and Mathematics	Scientific publications; research quality; interdisciplinarity; mentorship; citations

The comparison confirms that scientific publications remain a central reference point across all four fields, while interdisciplinarity, project leadership, mentorship and research leadership or management recur across several disciplinary groups. At the same time, distinctive priorities are clearly visible. Policy and societal contribution is particularly prominent in the Social Sciences, while patents, technical solutions, innovation and cooperation with industry distinguish Engineering, Technological and Biotechnical Sciences. In Natural Sciences and Mathematics, research quality and citations retain particular prominence, while Humanities place strong emphasis on contributions whose significance requires qualitative and contextual interpretation.

These patterns do not suggest the need for separate assessment systems for each scientific field. Rather, they demonstrate why common assessment dimensions need to be sufficiently broad to accommodate different forms of evidence and different pathways to contribution. Scientific excellence remains fundamental across fields, but the evidence through which excellence and wider contribution are demonstrated cannot be assumed to be identical. The relevance of particular outputs and contributions depends on the nature of the field, the objectives of the research and the context in which it is undertaken.

The cross-disciplinary evidence therefore supports a distinction between common assessment architecture and discipline-sensitive implementation. A shared framework can provide coherence at the level of broad assessment principles and dimensions, while disciplinary context should inform the selection of evidence, its interpretation and the relative emphasis placed on different forms of contribution. This distinction is particularly important for considering how the ARISE approach might be adapted beyond the social sciences without reducing disciplinary diversity to a uniform set of indicators or, conversely, fragmenting research assessment into entirely separate disciplinary systems.

6.3. What the Evidence Means for the ARISE Model

The cross-disciplinary findings provide an important additional perspective on the architecture of the ARISE model. Although the survey was not designed as a formal validation of its six assessment dimensions, the contributions prioritised by researchers across the four scientific fields show considerable correspondence with the areas incorporated into the model. At the same time, the findings demonstrate that the relevance and expression of these dimensions may differ across disciplines, reinforcing the need to distinguish between a shared assessment architecture and the evidence through which performance is demonstrated within it.

The centrality of scientific publications across all four fields provides strong support for retaining Research Publications and Scientific Outputs as a core dimension of assessment. Importantly, however, the findings also suggest that this dimension should not be reduced to publication counts or bibliometric characteristics. Research quality and significance remain central concerns, while the types of outputs through which scientific contribution is demonstrated may differ across fields. This is particularly relevant in disciplines in which books and other forms of scholarly production play an important role, as well as in fields where patents, technical solutions and other research and innovation outputs complement conventional scientific publications.

The prominence of project leadership and related project contributions across several fields is consistent with the ARISE dimension Research Projects and Strategic Relevance. The survey indicates that project activity represents more than an additional measurable output: it can reflect research leadership, the capacity to develop research agendas, mobilise resources and partnerships, and connect research with scientific, technological or societal priorities. The specific significance of project activity may nevertheless vary across disciplinary and institutional contexts.

The recurring importance of mentorship provides particularly clear support for Mentoring, Development and Lifelong Learning. Its prominence across different scientific fields indicates that the development of researchers and research capacity is not specific to the institutional context of IES or to the social sciences. Similarly, the importance attached to research leadership, management, interdisciplinarity and collaborative forms of research corresponds closely with Collaboration and Research Environment, while also demonstrating that these contributions may be expressed through different organisational and disciplinary practices.

The findings relating to policy impact, societal contribution, patents, technical solutions, innovation and cooperation with industry are particularly relevant to Societal, Policy and Economic Engagement. They demonstrate why this dimension needs to remain broad and context-sensitive. In the Social Sciences, contribution may be expressed through policy advice, societal engagement and other forms of knowledge use, whereas in Engineering, Technological and Biotechnical Sciences it may be demonstrated through technological solutions, innovation

and interaction with industry. The underlying assessment dimension may therefore be shared, while the evidence used to demonstrate contribution should reflect different disciplinary pathways to impact.

The survey provides a somewhat different perspective on International Cooperation and Networking. Although internationalisation appeared in researchers' responses, it did not emerge among the highest-ranked priorities in the cross-field comparison. This should not be interpreted as evidence that international cooperation is unimportant, but it does suggest that its relative position within the assessment framework warrants further consideration. International collaboration may also operate as an enabling or cross-cutting element of research quality, project activity, interdisciplinarity and capacity development rather than always being experienced as a separate form of contribution. The finding therefore points to an area that should be examined further through subsequent piloting rather than assumed to have identical relevance or weight across all disciplinary and institutional contexts.

The cross-disciplinary evidence does not point to the need for a fundamentally different assessment architecture. Instead, it suggests that the six ARISE dimensions provide a sufficiently broad basis for recognising many of the contributions valued across scientific fields, while their operationalisation should remain adaptable. The types of evidence considered relevant, the interpretation of that evidence and the relative emphasis placed on particular dimensions may need to vary according to disciplinary profiles and institutional missions. This is consistent with the underlying logic of qualitative peer review, in which common assessment principles provide structure and transparency while expert judgement allows research contributions to be interpreted in context.

The findings therefore strengthen the case for further testing of the ARISE approach beyond IES, while cautioning against direct replication of the pilot model without adaptation. Broader application should preserve the core principles of multidimensional assessment, qualitative peer review, responsible use of quantitative evidence and contextual interpretation, while allowing the framework to evolve in response to disciplinary evidence and experience gained through subsequent pilots.

6.4. From Institutional Pilot to Broader Adoption

The cross-disciplinary survey extends the evidence generated through the IES pilot by showing that many of the principles underlying the ARISE approach resonate across different scientific fields in Serbia. Researchers across the four fields expressed interest in broader recognition of research quality and diverse forms of contribution, while also demonstrating that the ways in which research excellence and impact are expressed vary across disciplines. This combination of convergence and diversity has important implications for any further development of the model beyond IES.

The evidence supports a gradual approach to broader adoption rather than the immediate application of a fixed model across the research system. The IES pilot demonstrated that qualitative, peer-review-based assessment can be operationalised at research-unit level, while the cross-disciplinary survey indicates that its broader principles have relevance beyond the social sciences. However, the survey also confirms that disciplinary differences need to be reflected in the types of evidence considered, the interpretation of contributions and, potentially, the relative emphasis placed on particular assessment dimensions.

A logical next step would therefore be to test the approach in research organisations representing different scientific fields and institutional profiles. Such pilots would make it possible to examine how the common ARISE architecture functions in settings characterised by different research outputs, organisational structures, collaboration patterns and pathways to scientific, societal and economic contribution. They would also provide further evidence on which elements of the framework should remain common and where discipline- or institution-specific adaptation is necessary.

The survey findings also indicate that successful expansion would depend on more than the assessment framework itself. Researchers emphasised the importance of transparent procedures, clearly defined criteria, appropriate methodological guidance, organisational and technical support, and continued engagement of the research community. These conditions are particularly important for qualitative assessment, where confidence in the process depends on the clarity of expectations, the quality of expert judgement and the opportunity for research organisations and researchers to understand and participate in the assessment process.

The evidence generated through ARISE therefore supports an approach based on a common architecture with discipline-sensitive and institution-sensitive implementation. The six assessment dimensions can provide a shared starting point, while subsequent pilots can generate the evidence needed to refine their interpretation, supporting guidance and application across different research environments. This would allow broader implementation to develop through testing and learning rather than through the imposition of a uniform model.

The progression from the IES pilot to cross-disciplinary evidence consequently represents an important step towards considering the wider relevance of the ARISE approach. It does not establish a final national model, nor does it suggest that the framework developed at IES should be transferred unchanged to other research organisations. Rather, it provides an empirical basis for the next stage of reform: broader piloting, systematic learning across institutional and disciplinary contexts, and the gradual development of a research assessment framework capable of combining common principles with the diversity of Serbian research.

7. From Evidence to Dialogue: The ARISE Forum and Pathways for Reform

The evidence generated through the development and pilot implementation of the ARISE model provided a basis for extending the discussion beyond IES and engaging the wider research community in considering the future of research assessment in Serbia. An important step in this process was the Forum on Advancing Research Assessment in Serbia, organised within the ARISE project as both a dissemination event and a platform for stakeholder dialogue. The Forum provided an opportunity to present and discuss the experience gained through the project, place it in the broader context of European research assessment reform, and engage representatives of the research community in considering the conditions, challenges and possible pathways for the further development of responsible research assessment in Serbia. In this way, the Forum complemented the empirical evidence generated through the project with stakeholder perspectives on the potential transition from an institutional pilot towards broader and longer-term reform.

7.1. The ARISE Forum: Dissemination and Stakeholder Dialogue

The Forum on Advancing Research Assessment in Serbia was held on 9 July 2026 at Europe House in Belgrade. Its programme combined the dissemination of project results and international experience with a structured discussion on the future of research assessment in Serbia. The first part of the Forum presented the ARISE project and the experience gained through the development and pilot implementation of its qualitative research assessment model, the French experience with qualitative peer-review-based research assessment, and the current position of Serbia in the broader process of research assessment reform, including key challenges and future directions in the context of CoARA principles. Together, these contributions placed the ARISE experience within both the European reform landscape and the specific context of the Serbian research system.

The second part of the Forum was organised as a Round Table entitled *Towards Responsible Research Assessment in Serbia: Opportunities for Broader Application of the ARISE Model in the Serbian Research Community*. Moderated by members of the ARISE Advisory Board, the discussion shifted the focus from presenting project results and relevant experience to collectively considering their broader implications for research assessment in Serbia. The Advisory Board thus continued the role it had played during model development by helping to connect the institutional experience of the ARISE pilot with perspectives from different scientific fields and the wider Serbian research community.

This dual structure reflected the broader purpose of the Forum. Dissemination was not limited to communicating the results of the project, but served as a basis for informed discussion with the research community about what had been learned through ARISE and how this experience might contribute to future reform. The Forum therefore created a space in which evidence from the pilot, international experience and perspectives from the Serbian research system could be considered together. ARISE was presented not as a completed or universally applicable model, but as a tested approach and a basis for discussing how responsible and meaningful research assessment could be further developed in Serbia.

7.2. Round Table: From the ARISE Experience to Pathways for Reform

The Round Table discussion was structured to move from the experience generated through ARISE towards the broader question of how responsible research assessment could develop in Serbia. Participants were invited to consider the weaknesses of the existing assessment system and the need for change; the feasibility of developing a common assessment framework while respecting differences between scientific fields; the potential and challenges of qualitative assessment based on self-assessment and external peer review; the broader applicability of the ARISE dimensions and criteria; and the institutional support, methodological guidance, training and capacity building that wider implementation would require. The discussion concluded by considering the role of the research community in further developing the approach and the prerequisites for a common framework for responsible research assessment in Serbia.

The discussion confirmed considerable interest in continuing the reform process, while emphasising that broader implementation should proceed through further testing rather than immediate system-wide adoption. Further piloting of the ARISE approach in research units or groups in other research institutes was identified as a realistic next step, allowing the model to be tested in different institutional and disciplinary contexts. Participants also emphasised the research unit or group, rather than the individual researcher or the institution as a whole, as an appropriate level for qualitative assessment. At the same time, the discussion raised the possibility of developing a broader common set of assessment dimensions and indicators from which institutions could identify elements relevant to their missions, strategies and specific characteristics.

A second important message concerned the time horizon of reform. Participants stressed that changes in research assessment should be approached as a long-term process of institutional learning rather than as a one-off replacement of one methodology with another. The discussion proposed defining a longer-term vision of where the Serbian research assessment system should be in ten years and considering successive assessment cycles as a means of progressing towards that objective. Two five-year cycles were discussed as a possible framework through which the system could be progressively tested, refined and strengthened on the basis of accumulated experience.

The Round Table also moved beyond assessment methodology to the institutional architecture required to sustain reform. One proposal was to establish a working body bringing together relevant representatives of the research community, which could further develop the initiative and potentially serve as a precursor to a future national body for the assessment of research institutes. Participants emphasised that any future national assessment body should have sufficient autonomy and institutional independence from the Ministry responsible for science to safeguard professional independence, credibility and the trust of the research community. The experience of France was discussed in this context, particularly the importance of reform being supported and shaped by the research community itself.

Building such support was consequently identified as an important condition for further reform. Participants stressed the need to familiarise a wider community of researchers with the principles and experience of responsible research assessment and to develop a critical mass capable of supporting and contributing to change. The Association of Institutes of Serbia (ZIS) was identified as a potentially important platform for continuing dialogue among research institutes, including through presentation of the ARISE experience and further working

discussions. At the same time, the Forum highlighted the importance of establishing dialogue with the Ministry and other decision-makers while preserving the professional autonomy of the assessment process.

An additional prerequisite identified during the discussion was the strategic capacity of research organisations themselves. Meaningful qualitative assessment requires clearly articulated institutional missions, objectives and development priorities against which achievements and future potential can be considered. Participants noted that development strategies are not equally established across Serbian research institutes and highlighted experience gained through the SAIGE project, particularly in strategic planning, institutional transformation and expert support to research institutes. The discussion therefore suggested that wider introduction of qualitative assessment may need to be accompanied by support for institutions in developing or strengthening their strategies and should build on capacities and experience generated through previous reform initiatives.

The discussion also highlighted broader consequences of the way research is assessed. Participants raised concerns about the financial pressures associated with open-access publication charges and their relationship with incentives generated by publication-based assessment. Scientific communication and the visibility of research institutes were also identified as relevant considerations. By enabling institutions to demonstrate forms of contribution that are not visible through publication and bibliometric indicators alone, broader qualitative assessment may also help them communicate their scientific, societal and policy relevance more effectively to decision-makers and other stakeholders.

The Round Table discussions reinforced a central lesson emerging from the ARISE project: reforming research assessment cannot be reduced to replacing or expanding a set of indicators. It requires a longer-term vision, further experimentation and learning, institutional capacity and strategy, credible and professionally independent assessment structures, and broad engagement of the research community. The ARISE pilot provides a practical starting point for this process, but the pathway towards broader reform should remain gradual, evidence-based and open to adaptation as experience is accumulated across different institutional and disciplinary contexts.

8. Policy Implications and Recommendations: Towards Responsible Research Assessment in Serbia

The evidence generated through ARISE points to the need for a gradual but substantive shift in the way research assessment is understood and organised in Serbia. The recommendations presented in this section draw on the analysis of the existing Serbian research assessment system and selected European approaches, evidence collected from the IES research community during model development, experience gained through the development and pilot implementation of the ARISE model, external peer review and post-pilot feedback, the cross-disciplinary survey of researchers in Serbia, and stakeholder dialogue at the ARISE Forum. This evidence points not towards the replacement of one set of indicators with another, but towards a broader transformation of assessment logic, processes and governance. The recommendations therefore address both the assessment approach itself and the institutional conditions required for its credible, sustainable and progressively broader application.

Recommendation 1. Shift from metric dominance towards qualitative, peer-review-based and context-sensitive assessment

Serbia should progressively move from a predominantly indicator-driven approach towards a research assessment system in which qualitative peer review plays a central role. Quantitative indicators should remain an important source of evidence, but should support rather than substitute for informed assessment of research quality and relevance.

The evidence generated through ARISE consistently supports such a shift. Existing assessment practices in Serbia provide standardised and measurable information, but give limited recognition to the diversity of research contributions and to the institutional and disciplinary contexts in which research is produced. Evidence from IES researchers, experience from the ARISE pilot and the cross-disciplinary survey all indicate the need for assessment capable of recognising research quality and relevance alongside a broader range of scientific, organisational, societal, policy and economic contributions. At the same time, the cross-disciplinary evidence confirms that quantitative indicators remain relevant, and in some scientific fields particularly important. The reform objective should therefore not be to replace quantitative assessment with purely qualitative judgement, but to establish a more appropriate relationship between the two.

In such a system, indicators should function as evidence to be interpreted rather than as proxies that determine assessment outcomes by themselves. Publications, citations, projects and other measurable outputs can provide important information, but their meaning depends on disciplinary practices, research objectives and institutional context. Qualitative peer review makes it possible to examine the significance of these achievements, understand how different contributions relate to the mission and development of the research unit, and recognise forms of contribution that cannot be adequately represented through quantitative indicators alone.

This shift also requires attention to the incentives created by assessment systems. The ARISE Forum highlighted that publication-based incentives can influence researcher behaviour and generate wider financial consequences, including pressures associated with open-access publication costs. Responsible research assessment should therefore consider not only whether particular indicators are technically measurable, but also what behaviours and resource-allocation decisions they encourage.

The objective is consequently not less evidence, but better use of evidence: moving from counting outputs as an end in itself towards informed judgement about the quality, relevance and contribution of research. This principle should provide the foundation for the subsequent development of research assessment in Serbia.

Recommendation 2. Make research units the primary level of qualitative assessment

Qualitative research assessment should primarily focus on research units or groups rather than individual researchers. The experience of the ARISE pilot demonstrates that the research-unit level provides an appropriate basis for assessing not only research outputs, but also collective capacities, collaboration, research environment, strategic orientation and development needs that cannot be meaningfully understood through the aggregation of individual performance indicators alone.

The pilot assessment of the six IES research departments illustrated the added value of this perspective. External peer review made it possible to consider how individual and team achievements contribute to the broader capacity of a research unit, to identify organisational strengths and development needs, and to examine the conditions under which research quality is produced and sustained. The unit-level perspective therefore shifts attention from isolated outputs towards the collective research environment and the capacity of research groups to develop coherent research profiles, support researchers at different career stages, collaborate effectively and pursue longer-term scientific and institutional objectives.

This approach is also consistent with the conclusions of the ARISE Forum, where research units or groups were identified as an appropriate level for future qualitative assessment and further piloting in other Serbian research institutes. Assessing research at this level can provide institutions with evidence that is directly relevant to strategic development while preserving sufficient proximity to the actual research activities, communities and disciplinary contexts being evaluated.

Unit-level assessment should, however, be clearly distinguished from the assessment of individual researchers. Qualitative evaluation of research units should not replace procedures related to individual career progression, promotion or other forms of individual assessment, which serve different purposes and may require different criteria and evidence. Nor should unit-level assessment be reduced to the mechanical aggregation of individual scores. Its purpose is to understand the quality, contribution, capacity and development of the research unit as a collective research environment.

Establishing research units as the primary level of qualitative assessment would therefore create a complementary layer within the Serbian research assessment system: one capable of generating evidence for organisational learning and strategic development while allowing individual assessment processes to retain their distinct function.

Recommendation 3. Develop a common multidimensional assessment framework with discipline- and institution-sensitive implementation

Serbia should work towards a common multidimensional framework for qualitative research assessment that provides sufficient coherence across the research system while allowing its application to reflect differences between scientific fields, institutional missions and research profiles. A common framework should define broad principles and dimensions of

assessment, but should not prescribe an identical set of outputs, indicators, evidence requirements or expectations for all research organisations and disciplines.

The ARISE experience provides an evidence-based starting point for such an approach. The six dimensions developed and tested through the project (Research Publications and Scientific Outputs; Research Projects and Strategic Relevance; Collaboration and Research Environment; Mentoring, Development and Lifelong Learning; Societal, Policy and Economic Engagement; and International Cooperation and Networking) were designed to capture a broader range of research contributions than those visible through conventional publication-based assessment. Evidence from the IES pilot indicates that this multidimensional approach can generate a richer understanding of research-unit performance, while the cross-disciplinary survey provides additional support for many of the contributions represented within these dimensions.

At the same time, the cross-disciplinary evidence demonstrates why a common framework cannot be implemented as a one-size-fits-all model. Scientific publications remain central across fields, but other highly valued contributions vary in their prominence and form. Policy and societal contribution is particularly relevant in the Social Sciences, while patents, technical solutions, innovation and cooperation with industry are more prominent in Engineering, Technological and Biotechnical Sciences. Mentorship, interdisciplinarity, project leadership, research leadership and management, research quality and citations also assume different levels of importance across fields. A common framework must therefore be capable of recognising different pathways through which research quality, excellence and contribution are demonstrated.

This requires flexibility not only in the types of evidence accepted, but also in how evidence is interpreted. The same broad assessment dimension may be expressed through different activities and outputs depending on the field and institutional context. Societal, Policy and Economic Engagement, for example, may be demonstrated through policy advice and contribution to public debate in one setting and through technological solutions, innovation or cooperation with industry in another. Qualitative peer review is particularly important in this respect because it allows evidence to be interpreted in relation to disciplinary practices, research objectives and institutional missions rather than evaluated against a uniform numerical template.

The framework should also remain open to refinement as further evidence is generated. The cross-disciplinary survey, for example, showed that International Cooperation and Networking, although recognised as relevant, did not emerge among the highest-ranked assessment priorities across scientific fields. This does not provide a basis for removing the dimension, but it does raise a legitimate question about whether its role and relative emphasis should be identical across all disciplinary and institutional contexts. Further piloting should therefore be used to examine the relevance, boundaries and relative emphasis of individual dimensions rather than treating the initial ARISE architecture as fixed.

The ARISE Forum reinforced this direction by raising the possibility of developing a broader common set of assessment dimensions and indicators that could be adapted to the missions, strategies and specific characteristics of research organisations. Such flexibility should not, however, result in an entirely fragmented system in which every institution defines its own assessment rules. Common principles, broad dimensions, methodological standards and expectations concerning evidence and peer review are necessary to ensure credibility and coherence, while contextual adaptation should allow the framework to recognise legitimate differences in how research is conducted and contributes to science and society.

The objective should therefore be a common architecture without uniformity: a shared framework that provides coherence and transparency while preserving sufficient flexibility for discipline- and institution-sensitive interpretation. The six ARISE dimensions offer a tested starting point for this development, but their broader application should remain subject to further empirical testing, consultation and refinement.

Recommendation 4. Link research assessment to institutional mission, strategy and development

Qualitative research assessment should be designed not only to determine how research units have performed, but also to support their future development. Assessment should therefore be explicitly connected to the mission, strategic objectives and development priorities of research organisations, enabling research units to understand their strengths, identify areas requiring improvement and use evaluation findings as an input into subsequent strategic and organisational decisions.

The ARISE pilot demonstrated the value of this developmental function. Departmental self-assessment created an opportunity for research units to reflect collectively on their achievements, capacities and areas for development, while external peer review provided a broader and more contextualised perspective on their research profiles. Post-pilot feedback further showed that researchers valued this reflective dimension of the process and regarded the final assessment reports as useful, while also calling for recommendations that are increasingly specific, prioritised and actionable. These findings suggest that the value of qualitative assessment lies not only in reaching an evaluative judgement, but also in generating knowledge that research units can use.

For this developmental function to be meaningful, assessment needs a clear institutional reference point. Research performance cannot be interpreted fully independently of what an organisation or research unit is seeking to achieve. Institutional mission, strategic objectives and development plans provide the context within which achievements, constraints, priorities and future potential can be assessed. This does not mean that institutions should be evaluated solely against objectives they define for themselves, but that common standards of research quality should be interpreted alongside the institution's role, profile and strategic direction.

The ARISE Forum identified the existence of clearly articulated institutional development strategies as an important prerequisite for meaningful qualitative assessment. Participants also noted that strategic planning capacities are unevenly developed across Serbian research institutes. Experience generated through the SAIGE project, particularly in supporting strategic planning and institutional transformation, was identified as a relevant resource that should inform future reform. Broader introduction of qualitative assessment should therefore be accompanied, where necessary, by support to research organisations in developing or strengthening their strategies rather than assuming that all institutions already possess the same strategic capacity.

Assessment findings should, in turn, feed back into these strategies. Recommendations arising from peer review should be sufficiently specific to inform priorities and subsequent action, while future assessment cycles should consider how research units and institutions have responded to earlier findings. This would create a continuous relationship between assessment, reflection, strategic action and subsequent reassessment, rather than treating evaluation as an isolated periodic exercise.

Such an approach would also change the function of assessment within the research system. Instead of serving primarily as a mechanism for classification, compliance or control, evaluation can become an instrument of institutional learning and development while continuing to provide accountability and evidence of research quality. The longer-term objective should therefore be to establish an assessment culture in which accountability and development are mutually reinforcing rather than competing purposes.

Recommendation 5. Scale reform through successive pilots, assessment cycles and institutional learning

The broader introduction of qualitative research assessment in Serbia should proceed through successive pilots and learning cycles rather than through the immediate system-wide implementation of a fixed model. The next phase should extend testing beyond IES to research units in institutions representing different scientific fields and organisational profiles, using each pilot to generate evidence for further refinement of the assessment framework, methodology and implementation arrangements.

The ARISE experience provides a strong basis for such a gradual approach. The IES pilot demonstrated that qualitative, peer-review-based assessment can be operationalised at research-unit level and can generate findings that are meaningful for institutional reflection and development. The post-pilot survey showed substantial confidence in the assessment findings and support for the potential applicability of the approach in other Serbian research organisations, while also identifying areas requiring further improvement. The cross-disciplinary survey similarly indicates that many of the principles underlying ARISE have relevance across scientific fields, but confirms that disciplinary differences need to be reflected in how evidence is selected, interpreted and weighted.

The immediate priority should therefore be broader piloting rather than national roll-out. Future pilots should deliberately include research organisations and units with different disciplinary profiles, missions, sizes and organisational characteristics. This would allow the emerging framework to be tested under conditions different from those of IES and provide evidence on which elements can remain common, which require adaptation and what institutional capacities are necessary for implementation. Pilot results should be systematically documented and compared so that subsequent changes to the framework are based on accumulated evidence rather than isolated experience.

The ARISE Forum placed this process within a longer-term reform perspective. Participants emphasised the importance of defining where Serbia would like its research assessment system to be in ten years and then developing a pathway towards that objective through successive assessment cycles. Two five-year cycles were discussed as one possible framework through which the system could be progressively tested, refined and strengthened. This proposal should be understood as a possible strategic horizon rather than a predetermined timetable, but its underlying principle is important: meaningful reform requires sufficient time for institutions and the research community to develop experience with qualitative assessment and for the system to learn from successive rounds of implementation.

Each cycle should therefore be treated not only as an evaluation exercise but also as an opportunity for system learning. Evidence should be collected on the functioning of assessment dimensions, disciplinary adaptation, self-assessment, reviewer selection and matching, the quality and consistency of peer review, the usefulness of recommendations, institutional

responses to assessment findings, and researchers' confidence in the process. Lessons from one cycle should inform the design of the next.

A gradual approach also provides an opportunity to build legitimacy. Expanding the system only after experience has been generated, discussed and incorporated into subsequent design can help develop trust among researchers and institutions and reduce the risk that qualitative assessment is perceived as another administratively imposed requirement. Researchers and participating institutions should therefore remain involved in reviewing pilot experience and shaping subsequent adaptations.

The reform pathway should consequently follow an iterative logic: pilot, evaluate, learn, refine and expand. The objective should not be to replicate the initial ARISE model unchanged, but to use it as a tested starting point from which a more mature and broadly applicable Serbian approach to responsible research assessment can progressively develop.

Recommendation 6. Build the professional, methodological and organisational capacity for high-quality qualitative assessment

The introduction of qualitative research assessment requires more than a well-designed assessment framework. Its credibility and usefulness depend on the professional expertise of reviewers, the quality and transparency of procedures, clear methodological guidance, adequate organisational support and the capacity of research units to participate meaningfully in the process. Building these capacities should therefore accompany any further piloting or broader implementation of the approach.

The ARISE pilot provides several concrete lessons in this respect. Post-pilot feedback highlighted the importance of selecting external reviewers whose disciplinary and subdisciplinary expertise closely matches the profile of the research unit being assessed. Relevant research experience, strong professional credentials and the capacity to understand the institutional and disciplinary context are essential for credible expert judgement. Reviewer selection should therefore be treated as a core methodological component of qualitative assessment rather than primarily as an administrative step.

A broader assessment system would require transparent procedures for identifying, selecting and matching international reviewers, accompanied by clear provisions concerning their expertise, independence and potential conflicts of interest. Particular attention should be given to establishing a sufficiently broad and multidisciplinary pool of qualified international reviewers, ensuring appropriate expertise and disciplinary alignment with the research units being assessed. Reviewers should receive clear methodological guidance, including structured assessment rubrics, to ensure a consistent understanding and application of assessment dimensions and criteria, the responsible use of quantitative evidence, and the formulation of well-substantiated evaluative judgements and actionable recommendations.

Research units also need adequate support for self-assessment. The ARISE experience indicates that self-assessment can generate valuable internal reflection, but its quality depends on clear guidance, sufficient time and an understanding that the process is intended to produce an evidence-based and reflective account rather than an additional administrative reporting exercise. Templates and guidance should provide sufficient structure to support consistency while allowing research units to explain their specific context, achievements, constraints and development priorities.

The interaction between reviewers and research units should be recognised as an integral part of the assessment process. Post-pilot feedback pointed to the value of sufficient time for discussion and clarification and to the need for a stronger feedback loop between evaluators and those being assessed. Qualitative assessment should therefore provide structured opportunities for dialogue, enabling reviewers to clarify evidence and contextual factors and research units to understand the basis of evaluative judgements.

Particular attention should also be given to the quality of final assessment reports. The IES pilot showed a very high level of satisfaction with the clarity of these reports, while researchers simultaneously called for recommendations that are more specific, prioritised and actionable. Future methodological guidance should therefore support reviewers in moving from general observations towards recommendations that clearly identify areas for development and provide research units with a meaningful basis for subsequent action.

As implementation expands, mechanisms for quality assurance and methodological consistency will become increasingly important. Qualitative and context-sensitive assessment does not imply the absence of standards. Common procedures, clearly defined assessment criteria, structured assessment rubrics, reviewer guidance and mechanisms for reviewing the consistency and quality of assessment reports will be necessary to ensure that expert judgement remains credible while avoiding a return to rigid numerical standardisation.

Finally, capacity building should extend beyond reviewers and research units to the organisations responsible for coordinating assessment. Future pilots will require adequate administrative and technical support, clear allocation of responsibilities, appropriate documentation and mechanisms for collecting and analysing lessons from each assessment cycle. Investment in this infrastructure should be regarded as part of the reform itself rather than as a secondary implementation issue.

The objective should therefore be to build an assessment ecosystem, not simply an assessment methodology: one in which capable research units, appropriately selected and supported reviewers, transparent procedures, methodological guidance, quality assurance and organisational capacity work together to make qualitative peer review credible, useful and sustainable.

Recommendation 7. Establish a multidisciplinary working body to guide the next phase of reform and prepare the ground for an independent research assessment institution

The next phase of research assessment reform in Serbia requires a mechanism capable of maintaining continuity beyond individual projects and pilot initiatives, coordinating further experimentation and ensuring that the development of the assessment framework remains grounded in evidence and the experience of the research community. As a concrete next step, a multidisciplinary working body for research assessment reform should be established, bringing together relevant representatives of the Serbian research community and expertise from different scientific fields.

The proposal for such a body emerged directly from the stakeholder discussion at the ARISE Forum. Participants identified the establishment of a working body as an important next step and considered that it could serve as a precursor to a future national structure responsible for research assessment. The discussion also highlighted the importance of institutional autonomy and professional independence for the credibility and legitimacy of any future national assessment structure.

In the immediate term, the working body should provide continuity and coordination for the next phase of reform. Its role could include guiding and coordinating further pilots in different research institutes and scientific fields; reviewing evidence generated through successive pilots; contributing to the refinement of common assessment dimensions and methodological guidance; considering discipline- and institution-specific adaptations; supporting the development of reviewer selection and quality-assurance procedures; and identifying capacity-building needs across the research system. It should also provide a forum in which lessons from implementation can be systematically considered before decisions are taken about broader application.

The composition of the working body should reflect the diversity of the Serbian research system. It should be multidisciplinary and primarily grounded in the research community, bringing together expertise from different scientific fields, institutional profiles and relevant experience in research assessment. The experience of the ARISE Advisory Board demonstrates the value of multidisciplinary expert reflection during model development, while the next stage will require a broader structure capable of supporting experimentation, methodological development and dialogue at system level. The precise composition and mandate of such a body should be developed through consultation rather than prescribed in advance.

Building wider ownership of the reform should form an important part of this process. The ARISE Forum highlighted the need to develop a critical mass within the research community that understands and supports responsible research assessment and identified the Association of Institutes of Serbia (ZIS) as a potentially important platform for extending dialogue among research institutes. Continued engagement of researchers, research organisations and other relevant stakeholders would help ensure that subsequent stages of reform are informed by practical experience and retain the confidence of the community to which they apply.

In the longer term, however, a working body should be understood as a transitional rather than permanent governance solution. As experience is accumulated through further pilots and successive assessment cycles, Serbia should progressively prepare the institutional foundations for an independent national institution for research assessment, capable of ensuring continuity, professional autonomy and methodological integrity of the assessment system. The comparative European experience examined within ARISE demonstrates the importance of stable institutional arrangements for qualitative, peer-review-based assessment and provides relevant reference points for considering the future Serbian model.

The precise institutional form, mandate and governance of such a body should not be predetermined at this stage. These should emerge from further piloting, consultation and accumulated implementation experience. Nevertheless, identifying an independent research assessment institution as a longer-term objective would provide strategic direction to the reform process and help ensure that successive pilots and methodological developments contribute to the construction of a sustainable system rather than remaining a series of disconnected initiatives.

The proposed working body can therefore serve as the institutional bridge between experimentation and a future sustainable assessment system: maintaining continuity between pilots, accumulating institutional knowledge, coordinating methodological development, engaging the research community and preparing the evidence and capacities required for a future independent national research assessment institution. In this way, the experience generated through ARISE can provide not only a basis for further testing of a new assessment approach, but also a starting point for the gradual development of the institutional architecture needed to sustain responsible research assessment in Serbia.

Conclusion

Over the past several years, the need to reconsider how research quality and contribution are assessed has gained increasing attention within the Serbian research community, reflecting both the limitations of existing national practices and the broader European movement towards research assessment reform. Two forums organised in Serbia prior to ARISE contributed to opening and advancing this dialogue, bringing together members of the research community to discuss emerging international and European principles and approaches and their relevance for the Serbian research system. Building on this growing reform momentum, ARISE represented a logical next step: moving the discussion from recognising the need for change towards developing, testing and critically examining a concrete alternative approach to research assessment in the Serbian institutional context.

The particular significance of ARISE lies in its character as a pilot project. Rather than proposing a new assessment model solely on the basis of international practice or conceptual principles, ARISE translated the broader direction of European research assessment reform into an approach tailored to national context and tested in practice at the Institute of Economic Sciences. The model was informed by analysis of the Serbian assessment system and relevant European experience, evidence collected directly from researchers, multidisciplinary expert reflection through the Advisory Board, and an explicit commitment to qualitative, peer-review-based and context-sensitive assessment.

The pilot demonstrated that such an approach can be operationalised at research-unit level and that it can generate insights that cannot be captured through conventional quantitative indicators alone. Self-assessment and external peer review provided a basis for considering not only scientific outputs, but also research capacity, collaboration, mentoring, projects, societal and policy contribution, international engagement, institutional strengths and development needs. Importantly, the pilot also generated evidence about the assessment process itself — what researchers value, what makes external peer review credible and useful, and which elements require further refinement.

ARISE subsequently moved beyond the boundaries of the initial institutional pilot. Post-pilot feedback, the cross-disciplinary survey and dialogue with the wider research community provided additional evidence of the broader relevance of the approach while also demonstrating the importance of disciplinary and institutional adaptation. The ARISE Forum brought these different strands together, combining dissemination of project experience with stakeholder deliberation on the future of research assessment in Serbia. In this sense, the project did not simply produce and test an assessment model; it created an evidence base around which the next stage of reform can be organised.

This is perhaps the most important contribution of ARISE. The project does not provide a finished national model, nor does it suggest that the framework piloted at IES should simply be transferred unchanged across the Serbian research system. It provides something that was previously missing: a tested institutional example, empirical evidence from implementation, and a concrete starting point for further reform. It demonstrates that a broader, qualitative and developmental approach to research assessment is possible in the Serbian context, while also identifying the conditions that need to be addressed before broader implementation.

The next phase should build on this foundation through further pilots across different scientific fields and institutional settings, continued methodological refinement, capacity building and systematic learning from successive assessment cycles. Establishing a multidisciplinary

working body would provide continuity to this process and create a bridge between the experience generated through ARISE and the longer-term establishment of an independent national institution capable of sustaining credible and professionally autonomous research assessment.

ARISE should therefore be understood not as the end point of research assessment reform in Serbia, but as an important transition from debate to experimentation, from experimentation to evidence, and from evidence towards systemic reform. Its lasting value will depend on whether the knowledge, experience and momentum generated through the pilot are carried forward. By providing a practical and evidence-based foundation for that next stage, ARISE has created an opportunity for Serbia to develop a research assessment system that combines scientific excellence with recognition of diverse contributions, common principles with disciplinary sensitivity, and accountability with institutional learning and development.

References

Coalition for Advancing Research Assessment (CoARA). (2022). *Agreement on Reforming Research Assessment* (Version 1).

Government of the Republic of Serbia. (2021). *Strategy of Scientific and Technological Development of the Republic of Serbia for the period 2021–2025: "The Power of Knowledge"*. Belgrade: Ministry of Education, Science and Technological Development.

Hcéres. (2020). *Evaluation Framework for Higher Education and Research Institutions*. Paris: High Council for the Evaluation of Research and Higher Education.

Hcéres. (2022). *Guidelines for the Evaluation of Research Outputs and Activities*. Paris: Hcéres.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for Research Metrics. *Nature*, 520, 429–431.

Institute of Economic Sciences. (2021). *Work and Development Program of the Institute of Economic Sciences for the period 2021–2025: Strategy '25*. Belgrade: Institute of Economic Sciences.

Institute of Economic Sciences. (2023). *Rulebook on Minimum Expected Results and Evaluation of Research Performance*. Belgrade: Institute of Economic Sciences.

Republic of Serbia. (2019). *Zakon o nauci i istraživanjima* [Law on Science and Research], "Sluzbeni glasnik RS", No. 49/2019.

National Council for Science Research of Romania. (2023). *Performance Indicators and Evaluation Criteria for Research Organisations*. Bucharest: CNCS.

National Council for Science and Technological Development. (2025). *Report on the state of science in 2024, with proposals and recommendations for the following year*. Republic of Serbia.

Romanian Ministry of Research, Innovation and Digitization. (2023). *Methodology for the Institutional Evaluation of Research Organisations*. Bucharest: Government of Romania.

Rulebook on the Acquisition of Research and Scientific Titles. (2024–2025). *Pravilnik o sticanju istraživačkih i naučnih zvanja*, "Sluzbeni glasnik RS", No. 80/2024 and 70/2025. Government of the Republic of Serbia.

San Francisco Declaration on Research Assessment (DORA). (2012). *San Francisco Declaration on Research Assessment*. American Society for Cell Biology.

SCImago Journal & Country Rank: Country Rankings. https://www.scimagojr.com/countryrank.php