

May 2015

R&D Evaluation Methodology and Funding Principles

Summary Report



R&D Evaluation Methodology and Funding Principles

Summary Report

May 2015

Erik Arnold, Bea Mahieu – Technopolis Group



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Preface

This study has been undertaken under a contract to the Ministry of Education, Youth and Sports of the Czech Republic by a team from Technopolis, the Technology Centre ASCR, NIFU, and Infoscience Praha. Our panel of expert advisors served as an internal sparring partner for the team and comprised

- Prof. Diana Hicks, Georgia Tech University
- Mr Paul Hubbard, formerly HEFCE
- Prof. Keith Jefferey, EuroCRIS
- Dr Gunnar Sivertsen, NIFU

A team of scientific experts supported us in the design and testing the field-specific evaluation methodology, comprised of: Prof. Christofer Edling, Prof. Milena Horvat, Prof. Ron Perrott, Prof. Roland Pochet, Prof. Naomi Segal, and Prof. Ken Thomson.

During the work, we have received constant help from the IPN Metodika project team members and staff, and have received input and comments from hundreds of stakeholders in the Czech Republic.

We are grateful to all the people who contributed for their helpful and friendly support. We hope our report justifies the effort they put in and will be of use in the next phase for the implementation of a new evaluation methodology and institutional funding system in the Czech Republic.

Erik Arnold and Bea Mahieu

Brighton 27 May 2015, for the evaluation team:

Erik Arnold, Oliver Cassagneau-Francis, Kristine Farla, Barbara Good, Malin Jondell Assbring, Peter Kolarz, Bea Mahieu, Göran Melin, Anke Nooijen, Fritz Ohler, Martijn Poel, Xavier Potau, Caspar Roelofs, Tammy-Ann Sharp, Brigitte Tiefenthaler, Frank Zuijdam – Technopolis Group

Vladislav Čadil, Michal Pazour, Ondřej Pecha, Ondřej Pokorný, Tomáš Rättinger, Tomáš Vondrák – Technology Centre ASCR

Tomáš Chudlarský, Jan Dvořák - InfoScience Praha,

Liv Langfeldt, Kyrre Lekve, Gunnar Sivertsen – NIFU

Summary

This report synthesises the outcomes of the study developing a new R&D evaluation methodology and funding system for the Czech Republic. It was undertaken under contract to the Ministry of Education, Youth and Sports in 2014-15. The main and background reports upon which this summary report is based are published on the ministry's IPN Metodika project website.

This executive summary reflects the structure of the report. International practice and experience was an important source of information for the design of the new evaluation and funding system, so we first set the new evaluation and funding system in its international context. We then describe the overall and specific principles that guided us in the design of the evaluation and funding systems that we propose and give an overview of their characteristics. We go on to cover some key topics relevant for the legitimacy and implementation of the evaluation and funding systems, and conclude with recommendations for the next steps.

The Czech RD&I system in the international context

Evaluation broadly defined is part of the policy process. In practice, evaluations have different functions and objectives, depending on policy needs and strategies, and these are reflected in the choice of methods and indicators used to assess performance. A common denominator is, however, that evaluation should analyse social effects of intervention, not just focus on outputs – in contrast to the current approach in the 'Metodika' that has been used to allocate institutional funding for research in the Czech Republic.

Methods for assessing research are widely debated. The research community in general tends to prefer peer review to other methods. Many recent performance-based research-funding systems (PRFS) nonetheless focus on indicators rather than peer review, though they also tend to reallocate only a modest proportion of total institutional funding so the imperfections associated with an indicators-based approach are not so important. 'Informed peer review', where peers make use of the best available indicators together with other information in reaching judgements, offers 'the best of both worlds' as well as allowing comparison between indicator-based and judgement-based views of performance and therefore a degree of 'triangulation' across methods. A stark lesson from international practice is that the former Czech approach (the Metodika or 'coffee mill') was unique in using indicators to allocate a high proportion of institutional funding for research.

In research assessment, a great deal of effort goes into tackling differences among fields. In peer review based systems this generally means that mechanisms are put in place to ensure that peer review panels all use assessment scales that have the same meanings, irrespective of discipline. In metrics-based approaches it means that the designer of the assessment has to produce some sort of bibliometrics-based technique for comparing across fields without, in reality, having a deep understanding of what these inter-field differences are. As in other situations where it is hard to define very precise rules, there is virtue in using (peer) judgement to find solutions to these problems. Whichever approach is used, assessment systems – like wider evaluation systems or systems for deciding on the quality of research proposals in research councils and other funding agencies – struggle with interdisciplinary research, so the assessment system has to have a way to address this.

Research organisations (ROs) tend to be steered and funded through various combinations of unconditional block funding and performance-based funding, which

may be based upon a performance contract and/or a system that counts or assess results of research, ie a PRFS. The use of performance contracts is widespread – not only on their own, but also in combination with a PRFS. They play an important role in institutional funding systems as instruments for dialogue between ROs and their principals (the Ministries). PRFS are essentially policy instruments. They can be tuned to achieve a range of different strategic policy objectives, which drive the focus and scope of the evaluation, the type of evaluation (summative and/or formative), the choice of assessment criteria and their indicators as well as the choice of institutional funding criteria and mechanisms.

Internationally, different categories of research organisation are evaluated and funded using different systems. They are not put into meaningless competition; rather, policy decisions are made about how much activity the state wants to support in each category. Also, the significance of single components in institutional funding tends to differ among different types of ROs. The Czech Republic is therefore unique in its current use of a single performance-based system to address all the different types of research organisation. We note also that a broader variety of types of research organisation can be considered for institutional research funding in the Czech Republic than in the international context. Based on their missions we categorised them as Scientific Research Organisations (ScRO), Industry & Business services Research Organisations (IBRO), Public Services Research Organisations (PSRO), and National Resources (NatRes).

The principles according to which various countries provide ROs with institutional funding are based in the economics of knowledge. Central principles are: institutional funding is justified according to the extent to which ROs produce public goods, or whether the state appoints an organisation to produce such knowledge on its behalf. Some cases arise in the Czech Republic that in our view violate some of these principles. Of course, provided it follows prevailing law the Czech state is fully within its rights in doing whatever it wants. Nonetheless, our advice would be to make use of internationally normal principles in deciding which organisations should be considered for institutional research funding.

The economics of knowledge also have a direct consequence on the level of institutional funding: the closer knowledge gets to market application, the more companies are able to monopolise aspects of the knowledge and secure economic returns. So the state plays a big role in funding basic research (usually paying all the costs) but invests a far smaller proportion in those cases where it intervenes in more applied areas that are closer to market.

An Evaluation Methodology and Funding System will only function properly when it is connected to the needs of society, as articulated by the ministries, the RD&I Council and other parts of government. A first key aspect of R&D governance systems that need to function well if they are to produce good policy is policy coordination. A second key aspect is the need for ‘distributed strategic intelligence’ – in the sense of a wide availability of data and sources of information. Individual ministries need to be able to analyse, express and lobby for the satisfaction of their own sector needs.

In this context, the organisation and governance structure for research funding in the Czech republic has some characteristics that undermine its ability to operate well. The 2008 Reform of the RD&I system had a major negative effect on the capacity of the system to govern sector RD&I. It reduced the number of Ministries and other public administration bodies with competences for RD&I funding, creating a situation where a number of sector ministries lack (or were deprived of) budgets to fund institutional R&D costs – and therefore to maintain a body of research capabilities and evidence relevant to developing their policies. The resulting centralisation of budgets has not been balanced by an effective mechanism able to coordinate needs and research priorities across sectors.

The new Evaluation Methodology and Funding System

As our discussion of international experience shows, research assessments and PRFSs need to be seen as parts of the policy system. They can allow policymakers to understand and influence the shape of the RD&I system, acting as sources of strategic information to all actors involved as well as an opportunity to steer behaviour and induce structural changes – both at the RD&I governance and research-performing levels. The effects of the new evaluation and funding systems in terms of the needs for public support they create, will need to be monitored and assessed, ensuring an on-going alignment of the policy strategy and priorities with these needs.

A corollary of this is that the Evaluation Methodology and Funding Principles do not stand alone but need to be under the coordinated control of the state funders, who ‘own’ the research-performing organisations and provide their institutional research funding. In this way, policy needs can serve as drivers for change – but also as drivers for the elements of stability and predictability that are needed in any well-functioning RD&I system.

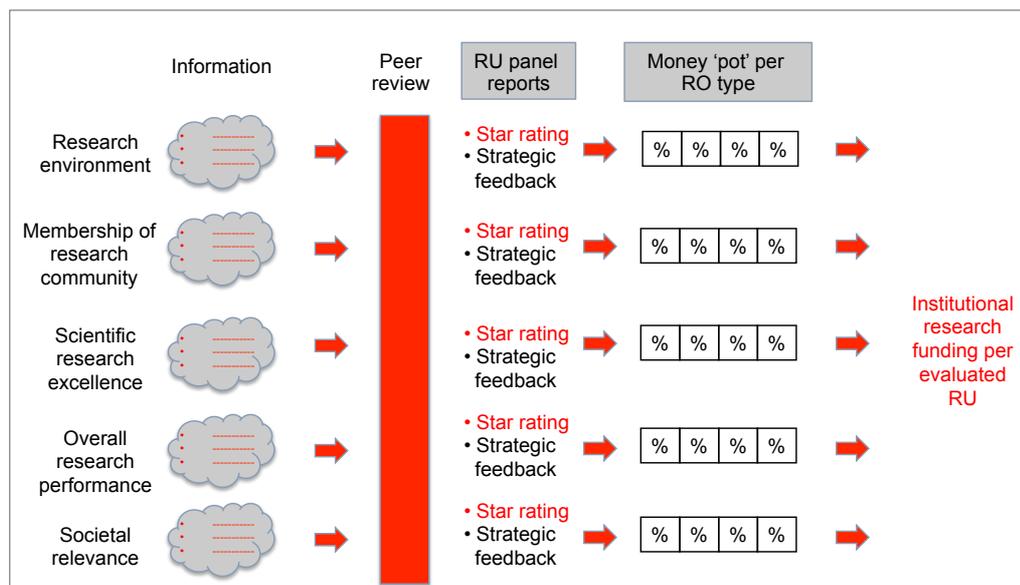
The history of the Metodika in the Czech Republic shows that a PRFS can itself induce instabilities in funding. The new Evaluation Methodology and Funding Principles are intended to reduce this instrument-induced instability. Historically, other key instabilities in Czech institutional funding have been policy-induced, underlining the importance of the link between the Evaluation Methodology and Funding Principles and the overall governance system.

We defined a set of ‘overall’ principles for the design of the Evaluation Methodology and Funding Principles, addressing some of the specific needs in the Czech Republic. Foremost is the need for fairness and transparency in the evaluation and funding system and the importance of taking the diversity of actors involved into account, in terms of the scientific fields in which they are active and the missions of their category of research organisation. The Evaluation Methodology and Funding Principles are designed to avoid setting different types of research organisation in competition with each other. The overall institutional funding system should reflect the needs for stability, while the evaluation system should provide formative as well as summative inputs to policymaking overall and to the specific development of individual research groups and organisations.

Most systems abroad make a clear distinction between evaluation (or ‘assessment’) and the use of assessment results in a funding formula and this distinction was central to the design of the new Evaluation Methodology and Funding System. This is reflected in the two-step structure of the Evaluation Methodology and Funding Principles. Scientific fields and the different Research Organisation types should not be set in competition to each other. We achieved this via the two-step approach: the evaluation system is centred on handling the scientific field specifics while taking the missions of the ROs into consideration; the institutional funding system centres on recognising the different missions of the RO types in society.

The expert panels will not assign an overall score to the evaluated units in their field. The evaluation results will consist of scores and related explanatory texts against each assessment criterion, together with panel conclusions and recommendations. The Funding Principles entail a mechanism to then translate the evaluation scores against each criterion into funding, in the context of the performance-based research funding system (PRFS) component of the institutional funding system (Exhibit 1).

Exhibit 1 Evaluation in the context of a funding mechanism



Implementing this approach, the core elements of the **R&D Evaluation Methodology** are as follows.

The evaluation takes place at the level of *Evaluated Unit (EvU)*, ie a Research Organisation or for the public HEIs, the organisational unit at the second level of the organisation's structure (faculties, institutes, centres etc). All Research Organisations and Evaluated Units can participate in the evaluation, on a voluntary basis, provided there is a minimum of critical mass. A minimum volume threshold for participation is set at 50 eligible research outputs over the evaluation period.

However, the fundamental building block of the assessment is the *Research Unit (RU)*. This sits at the intersection of the 'natural' dimension for peer-based evaluation of research - the scientific field - and the dimension determined by the need for information that can be used at the level of institutions. Each eligible Evaluated Unit may register one or more RUs for participation in the evaluation. A Research Unit is registered for a *field* of research and an EvU can register only one Research Unit per field. A Research Unit means the group or groups of staff in the Evaluated Unit that conduct their *primary* research in a specific field, and by extension, the structures and environment that support their research and its application or impact. An Evaluated Unit staff member can be part of one Research Unit only. The scientific fields and their categorisation in disciplinary areas are based on the OECD field classification.

The evaluation is a panel-based process of informed peer review. It is entrusted to a core structure of Main panels, Subject panels and referees. There will be 6 Main panels, organised at the level of disciplinary area, and 24 to 26 Subject panels, organised at the level of fields. Referees will assess submitted research outputs and will work at the level of sub-field.

Only international experts will be in charge of the performance assessments. A number of national experts will support these experts acting as advisors on the national context, but will not be directly involved in the assessments. We have set high demands on the international experts' selection criteria related to conflicts of interest.

Interdisciplinary research spanning different sub-fields will be handled directly by the relevant Subject panel. Where inter-disciplinary research in the RU covers different fields within *one* disciplinary area, an Evaluated Unit can recommend cross-referrals among the Subject Panels. An EvU that has a strong element of interdisciplinary

research in fields across disciplinary areas can apply for registration an Interdisciplinary Research Unit.

The evaluation process will be fully transparent. The working language is English. The evaluation will be partly remote and consists in remote reviews and remote RU assessments combined with panel meetings.

The evaluation is a fair and egalitarian system. It uses a single framework for assessment across all disciplines and types of research organisation, thereby ensuring full comparability of the evaluation results across all dimensions while allowing for a reasonable level of field- and RO type-specific variations. Unity in the method is key to achieving ‘field-neutral’ evaluation scores for each RU, applying common criteria across the different fields and types of Research Organisations (but enabling the panels duly to take into account variations to these criteria). It also constitutes a critical requirement of the Evaluation Methodology as a source of strategic information at the national level.

The evaluation is structured around five assessment criteria that jointly enable the fulfilment of the strategic objectives of the evaluation and funding system, ie to reward excellence in research while building capacity in research and for innovation. The main assessment criteria are: the research environment; membership of the national and global research community; scientific research excellence; overall research performance; and relevance for society. The expert panels draw on information submitted by the Evaluated Units, including a mix of quantitative and qualitative data as well as the outcomes of the Research Units’ self-assessment. The focus is comprehensive, covering all dimensions of the research activities and research outputs, outcomes and impact, as well as the conditions facilitating research performance.

The evaluation results will be predominantly qualitative. The Subject panel reports will show for each Research Unit the quality levels for each assessment criterion, with explanatory texts, and provide conclusions and recommendations for future development. Conclusive analytical reports will be developed at the level of EvU, field and disciplinary area, aggregating the information collected at the RU levels

A key value of the Evaluation Methodology lies in its formative function, ie in the development of strategic information. For this purpose, the Evaluation Methodology is not conceived as an arithmetic exercise. It stresses the importance of qualitative information to inform both the evaluation panels and the actors in the RD&I system responsible for the definition of policies and strategies. These include the national policymakers in the Government and the RD&I Council, the funding providers (the Ministries and the Academy of Sciences – further collectively called “principals”) and the institutional management of the Research Organisations. For the benefit of these actors, the Evaluation Methodology foresees the development of analytical reports at the Evaluated Unit, scientific field and disciplinary area levels, building upon the panel reports at the Research Unit level.

We envisaged the implementation of a full-scale evaluation every 5 to 6 years. A key principle for the design of the Evaluation Methodology (EM) was that the cost and burden of the evaluation was to be kept at the minimum possible to deliver a robust and defensible process. We estimate the costs of a full-scale evaluation of the EM at 12,929 k€. This implies a total cost per researcher of €702; the indirect costs that will need to be carried by the RD&I community represent 48% of the total. For both parameters, the EM constitutes a cost efficient evaluation system internationally.

A key principle for the **institutional funding system** is that it needs to provide continuity, stability, sustainability and resilience for institutional development. Institutional funding should provide a reliable basis for institutional development and therefore it must not be subject to short-term policy making. Ideally, public institutional research funding and its basic principles should be laid down in a long-

term RD&I strategy which is supported by all major political forces and thus holds beyond election terms.

A prerequisite for the proposed funding model is that there are separate *mission-based 'pots'* (budgets or budget lines) for the different types of Research Organisations (RO) in the Czech RD&I system. These pots are then sub-divided for the three funding components: a block grant; a performance agreement; and a PRFS. While the block grant ensures trust and continuity, the performance-based funding components, encompassing the performance agreement and the PRFS, reflect past performance as well as future developments and plans.

Block grants and funding for performance agreements are allocated to Research Organisations as legal entities and are determined on the basis of the budget allocated to Research Organisations in the preceding funding period. The allocation of the PRFS budget line is determined by the scores achieved by Research Units (RUs) for each assessment criterion determine the allocation of the PRFS budget line (see Exhibit 1, above). For this purpose, five *evaluation criterion-based "sub-pots"* are defined. This implies a weighting of the assessment criteria in line with the organisations' mission and in line with the PRFS's objectives and the ministries' strategies. This in turn implies that the weights have to be different for the different types of organisations because they have different missions and roles in the Czech RD&I system.

An important principle is also that the institutional funding system needs to take into account the different missions and specifics as well as the whole funding situation of Research Organisations. We conceptualise institutional funding in the context of a long-term working relationship between a ministry in charge (the principal) and a Research Organisation (the agent). Public institutional funding expresses responsibility and ownership by the state, with the state taking an interest and a stake in the Research Organisations it (co-)funds. The share of institutional funding in relation to the overall budget is different for different types of organisations. The rule of thumb is: the closer a research organisation is to the market, the less institutional funding it receives. In the end, it is a policy decision what share of institutional funding a research organisation should have in its overall funding mix.

Implementation of the evaluation and funding systems

The new R&D Evaluation Methodology (EM) and institutional funding system that we propose in this study constitute a significant change in the current approach to evaluation and institutional funding in the Czech republic. As with any reform, it will require time, patience and a constructive involvement of all actors in the RD&I system to be effective. Most important, it requires change management and measures that allow for the legitimisation of the EM and new funding system, providing the structure and means for capacity building and ownership among all actors involved.

In order for the evaluation and funding system to be 'legitimised', we see the need for a number of changes and actions. First of all, we considered that at least in the medium term, the Evaluation Methodology and Funding Principles would work more effectively in the social interest if both the governance structure and the capabilities of key ministries were strengthened. It is therefore our view that a number of changes in governance and ministry capabilities are needed to align the design and execution of Czech RD&I policies better with social needs. These range from a clear allocation of both budgetary and planning responsibility for sector research to ministries needing meaningful quantities of research, to the allocation of the responsibility for coordinating overall research funding and the Evaluation Methodology and Funding Principles to a single point in government, and the implementation of the civil service reforms - planned but not yet put in place.

The establishment of a *stable* structure for evaluation, allowing for the build-up and continuous use of skills and expertise, is an urgent point of action. The success of the evaluation and funding system will depend on the expertise, capabilities and capacities of the staff in this structure and there will be a need for training and learning and

building up an institutional memory from past experiences. In the first run of the evaluation, there will also be a need for external support and expert advice.

The research-performing community needs to build up evaluation capacity. There is little experience in evaluation, which risks advantaging those who did have the opportunity to build up capacity and have a better understanding of the type of information evaluation panels require. The ministries will need to build up expertise and knowledge, in order to take up their more active role in the governance of the ROs in their sector, but also for the translation of the evaluation results into the different policy agendas and vice versa.

Conclusive recommendations

Throughout this report we indicated various topics for which action is needed in order to guarantee a proper functioning of the new evaluation and funding system. There is a need for change in the RD&I governance system as a whole as well as for policy decision-making and action on some more operational aspects.

We recommend that national policy-makers

- Take a decision in principle to the adopt the new evaluation and funding system as soon as possible
- Strengthen the governance structure and the capabilities of key ministries, implementing a number of structural reforms in the governance system in order to align the design and execution of Czech RD&I policies better with social needs
- Set up a structure and processes for decision-making on the technical procedures as well as for a set of parameters that can be varied in the funding system
- Support the collection of strategic information for an improved understanding of the RD&I system and to support policymaking

We invite the RD&I Council

- To reach an agreement on a standard definition of an FTE researcher and the procedures for the calculations
- To review the definition of a 'research organisation' and the eligibility criteria for institutional research funding
- To take up the activities needed for a fair evaluation of research performance, including the development of a categorised journal register and the definition of applied research outputs
- To foster the implementation of updates and extensions to the RD&I Information System

We recommend the Research Organisations

- To set up internal structures and processes for evaluation, in line with the evaluation methodology

Table of Contents

1. Introduction	1
1.1 Background, goals and tasks of this study	1
1.2 Approach and methods	2
1.3 Guide to the report	5
2. The Czech RD&I system in the international context	6
2.1 R&D evaluation and institutional funding in international practice	6
2.1.1 R&D evaluation	7
2.1.2 Institutional funding for research	8
2.2 Research-performing organisations and their institutional funding	10
2.3 Research profile from a bibliometric perspective	14
2.4 The RD&I System in the Czech Republic	16
2.4.1 The RD&I state budget expenditures	16
2.4.2 The RD&I governance structure	17
3. The Evaluation Methodology and Funding Principles	21
3.1 Key principles and overall architecture	21
3.1.1 Key principles	21
3.1.2 The architecture	25
3.1.3 The evaluation structure	27
3.2 The Evaluation Methodology (EM)	29
3.2.1 Scope of the evaluation	29
3.2.2 The units of analysis and eligibility for participation in the evaluation	30
3.2.3 The evaluation panels	31
3.2.4 The panel evaluation process and its results	33
3.2.5 Overview of the assessment criteria	35
3.2.6 Overview of the indicators for the assessment criteria	37
3.2.7 Data on FTE researchers	40
3.2.8 Data and indicators related to research outputs	41
3.2.9 The potential role of the RD&I Information System (IS)	45
3.2.10 Implementing the National Evaluation of Research Organisations (NERO)	47
3.2.11 Cost estimate of the evaluation	49
3.3 The Institutional Funding System	52
3.3.1 Overview of the proposed Institutional Funding System	52

Summary report

3.3.2 The mission-based pots	54
3.3.3 The block grants	54
3.3.4 Performance agreements	55
3.3.5 The performance-based research funding system - PRFS	56
3.3.6 Handling the effects of the evaluation	57
<hr/>	
4. Implementation of the R&D evaluation and funding system	61
4.1 Legitimation of the evaluation and funding system	61
4.1.1 The R&D Governance system: need for expertise and ownership	61
4.1.2 The development of an evaluation management structure	62
4.1.3 Policy decisions related to the funding system	63
4.1.4 Decision-making by the RD&I Council	64
4.2 Capacity building in the RD&I system	65
4.2.1 Evaluation capacities in the Research Organisations	65
4.2.2 The capacities of the evaluation management structure	66
4.2.3 The capacities of the ministries	67
4.3 A pathway towards stability	67
4.4 Conclusive recommendations	68
<hr/>	
Appendix A The OECD field structure	69

Table of Exhibits

Exhibit 1 Evaluation in the context of a funding mechanism	v
Exhibit 2 The Final study report: main and background reports.....	5
Exhibit 3 RO type categories – shares in overall number and institutional funding....	12
Exhibit 4 Comparative analysis of the CR research profile, 2008-2013	15
Exhibit 5 RD&I State Budget Expenditures by the funding providers, 2014	17
Exhibit 6 The RD&I governance system in the CR	18
Exhibit 7 The Evaluation Methodology and Funding Principles in the policy system .	22
Exhibit 8 Two-step approach to ensure fairness in institutional research funding.....	25
Exhibit 9 Evaluation in the context of a funding mechanism	26
Exhibit 10 Evaluation in its formative function.....	27
Exhibit 11 The Research Unit as primary unit of evaluation	28
Exhibit 12 Organisation of the evaluation experts in main and subject panels	32
Exhibit 13 Research Units and the structure for interdisciplinary research	33
Exhibit 14 The evaluation process.....	34
Exhibit 15 Assessment criteria and sub-criteria	35
Exhibit 16 List of the main indicators and their relevance for the types of RO	39
Exhibit 17 Potential for extended use of the evaluation results.....	47
Exhibit 18 Costs of PRFS in other countries	50
Exhibit 19 Maximum total costs of the new evaluation system.....	50
Exhibit 20 Cost estimate for a full-scale evaluation	51
Exhibit 21 The core elements of the Institutional Funding System	53
Exhibit 22 Pros and cons of the options for the translation of scores into funding.....	57
Exhibit 23 Process for the management of Research Units with a low evaluation score	58
Exhibit 24 Process for the management of Evaluated Units that do not participate in the evaluation.....	59

List of Abbreviations

ASCR	Academy of Sciences of the CR
CR	Czech Republic
EM	Evaluation Methodology
EvU	Evaluated Unit
FEM	Field-specific evaluation methodology
GACR	Czech Science Foundation
HR	Human resources
IBRO	Industry and business research organisation
NatRes	National resource
NERO	National Evaluation of Research Organisations
PRFS	Performance-based research funding system
PSRO	Public services research organisation
R&D	Research and Development
RD&I	Research Development and Innovation
RD&I IS	Research, Development and Innovation Information system
RI	Research infrastructure
RO	Research organisation
RTO	Research and technology organisations
RU	Research Unit
ScRO	Scientific research organisation
SPE	Small pilot evaluation
SWOT	Strengths, weaknesses, opportunities, threats
TACR	Technology Agency
UK	United Kingdom
WP	Work package

1. Introduction

This report synthesises the work done in the study developing a new R&D evaluation methodology and funding principles for the Czech Republic (CR), which was undertaken in 2014-15.

The study is a component of, and provides service to, the broader and longer-term project “Effective system for the evaluation and funding of research, development and innovation” (the IPN Metodika project). The final and background reports on which this report is based are published on the IPN Metodika project website.

1.1 Background, goals and tasks of this study

The current system of R&D evaluation and institutional financing is one of the significant systemic obstacles to the development of the research system in the Czech Republic (CR) and thus a target of criticism from both the research community and the political level. All national strategic documents that touch upon research point out the need for major changes in the system of R&D evaluation and institutional financing. The Czech Republic National Reform Programme 2013 lists the creation and adoption of a new methodology for evaluating and financing RD&I as one of the 24 priority reform measures with a pro-growth potential for 2013. This measure was already discussed in the Strategy for International Competitiveness project proposal, which was approved in 2011 and which contained more accurate parameters that the new system should include. The fact that there has been a long-term consensus in the CR with regard to the need to reform the system of R&D evaluation and financing is also evidenced by the inclusion of this task in the National Research, Development and Innovation Policy of the CR for 2009-2015 approved in 2009 and its repeated emphasis in the Update of the National Research, Development and Innovation Policy of the CR in 2009-2015 with Outlook to 2020, which was approved by the government in 2013. The Policy tasks the RD&I Council, in cooperation with the Ministry of Education, Youth and Sports and other providers, with introducing “a new methodology for evaluating research organisations and their results, which will evaluate the results and the fulfilment of basic research goals and the results and the fulfilment of applied research goals using different methods for different groups of scientific fields and which will stimulate research organisations to achieve excellence, to produce applicable R&D results, and to realise R&D based on the needs of society and enterprises.”

The principal reason for the change in the system of R&D evaluation and institutional financing in the CR is the need to stabilise the research environment in the CR and motivate research organisations and researchers to do quality research that can stand international comparison and that have high potential for the application of results in practice. An equally important factor that intensifies the need to change the system is the dynamic change in the research environment in the CR which has come about in recent years in relation to the enormous investments into the building of new research centres and infrastructures using EU Structural Funds. The sustainability and continuous development of these research capacities, as well as the existing ones, will depend, *inter alia* on the parameters of the national system of institutional R&D funding.

A key step to the creation of the new system of R&D evaluation and financing was the formulation and approval of the Individual National Project, “Effective Evaluation and Financing System for Research, Development and Innovation” (“IPN Metodika”), which was started in February 2012. This Project aims to

- Propose and pilot-test a system for the evaluation of research organisations on the basis of field evaluation methodologies with elements of international, informed peer-review, which would replace the existing Evaluation Methodology with a

more efficient evaluation that would be better oriented towards international and field standards

- Propose a system for the evaluation of the performance of the system as a whole and mechanisms for the financing of the whole system, so that public support will contribute to improving the excellence of Czech R&D, motivate all actors, and support growth in the competitiveness of the CR
- Improve the quality of evaluation culture in R&D and establish a professional base for regular and systematic institutional evaluation in the future

The project responds directly to the current need of the Czech R&D governance system, which is the creation and adoption of a system of R&D evaluation and financing that will conform to international standards

In accordance with established foreign practices, with regard to the conclusions and recommendations of the International Audit, and in accordance with the Long-term Principles of Evaluating and Financing R&D as approved by the RD&I Council in 2011, the new methodology for R&D evaluation should

- Take into consideration field specifics
- Take into consideration the different missions of research organisations within the research system
- Cover outputs, impacts, and institutional projections of research development
- Use a peer-review evaluation process
- Actively involve the entities assessed in the evaluation
- Set up evaluation processes that will be resistant to clientelism and conflicts of interest
- Be set up so that total costs do not exceed 1% of public institutional support for R&D in a five-year period
- Fulfil both formative and summative functions

The new system of institutional financing for R&D will reflect proven foreign models, conclusions of the International Audit, and the Long-term Principles of Evaluating and Financing R&D. The system should

- Make use of R&D evaluation results
- Combine financing on the basis of retrospective and prospective views
- Take into account potential changes to the institutional base of research

Together the new systems of R&D evaluation and institutional funding should motivate research organisations and individual researchers to carry out research of high quality and create conditions for long-term strategic orientation of research organisations.

1.2 Approach and methods

The Terms of References (ToR) for this study specify

1. Development of detailed and structured drafts of field-specific methodologies for the evaluation of evaluated units (EvU) within research organisations with elements of international field-specific informed peer-review
2. The use of relevant bibliographic, citation and other databases of research findings and their analysis tools to ensure tasks set in 1), including the preparation of specialists to ensure tasks set in 7)

3. Development of a detailed draft of evaluation procedural rules
4. Implementation of a small pilot test of EvU evaluation within ROs or of entire ROs
5. Design of the system of institutional R&D funding (budget as prescribed and performance contract) and its links to the results of the evaluation: principles of mid-term funding and specification of annual adjustments enabling the formation and termination of research institutions, reflection of R&D priorities, orientation of targeted and other funding while ensuring a reasonable level of stability of institutional funding at research organisations
6. Analysis of targeted funding in relation to and in the interaction with institutional funding, compliance of the behaviour of beneficiaries with research priorities of providers and the priorities of the providers with the priorities at the national level, the determination of an appropriate mix of targeted and institutional funding
7. Recommendations for the establishment and functioning of the institution organising evaluation

There were four major intellectual building blocks, which we used in fulfilling the requirements of the Contract.

First, we envisaged a specific architecture for connecting the various evaluation activities envisioned together. That treated the evaluation of research groups within fields as the ‘core’ of the system and then specifies how the other requirements relate to it. (See below for a fuller discussion.)

Second, we relied on a systemic perspective concerning the evaluation and funding system. The key point was that different parts of a national funding and evaluation system are interdependent and cannot sensibly be treated in isolation.

Third, we emphasised the use of experience both from the Czech Republic and from a number of other countries. We concentrated on five comparator countries but also made use of lessons from elsewhere. Internationally, there is considerable experimentation and reform in progress in relation to evaluation and (performance-based) research funding systems; it is important to exploit and not ignore this growing body of evidence.

Fourth, this study was intended not only to design a revised system of evaluation and funding but also to assist in the initial stages of piloting and implementation. This was further reason to use international and national experience of implementation issues. Evaluation and funding are ‘hot topics’ in both policy and the research community and different interest groups are prone to try to ‘game’ the system. Practical experience of implementation and how certain systems have been implemented was therefore an important input into the work.

We structured our work around in nine Work Packages (WPs), reflecting the requirements of the Terms of Reference:

- WP1: EvU evaluation structure determination
- WP2: EvU field-specific evaluation methodology (FEM)
- WP3: EvU evaluation process rules
- WP4: Evaluation information support
- WP5: Definition of “institutional funding” according to good international practice
- WP6: Analysis of the institutional R&D base in the Czech Republic
- WP7: Draft models of formula-funding and performance-based-contract

Summary report

- WP8: Models for the impact assessment of the new institutional funding system
- WP9: Preparation and implementation of a small pilot evaluation (SPE)

The methods we used across the Work Packages were as follows.

- Document analyses, such as policy documents in the CR and documents in the comparator countries on the evaluation and funding systems and their results
- Literature review, for example on the features and effects of PRFS systems
- Interviews with people and at key organisations in all types of research organisations and various funding bodies
- Workshops and focus groups, for example on the topics for the narratives related to the societal relevance of research
- Statistical data analyses, for example on the R&D base, its research outputs and historical data on institutional funding
- Bibliometric analysis of publications and citations
- Modelling and simulation analysis of funding scenarios and options

The Work Packages were implemented in three major phases.

- Phase 1, setting the basis for the study, ie the evaluation methodology (WP1, WP2 and WP3) and resulting in the *First Interim Report*
- Phase 2, developing the institutional funding system (WP5, WP6 and WP7), described in the *Second Interim report*
- Phase 3, testing the evaluation methodology (WP9), the outcomes of which were set out in the *Third Interim Report*

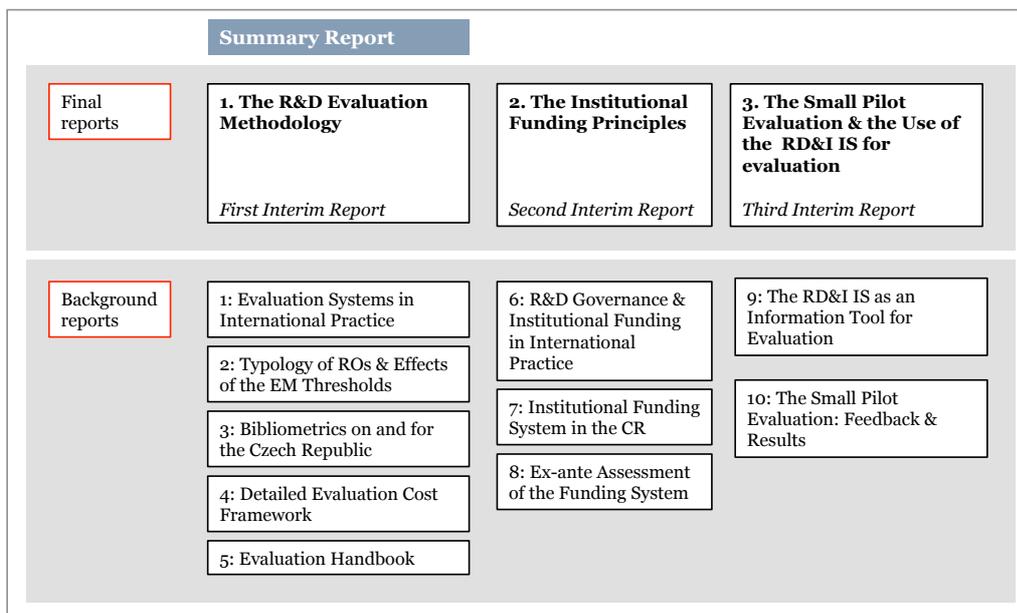
A number of activities were ‘horizontal’ and spanned over various WP, such as the work related to bibliometrics and the use of the RD&I Information System (IS) (WP4), the international comparisons, and the calculation of the potential costs of the evaluation system.

The design process of both the evaluation methodology and funding principles had a dynamic character. The WPs were strongly interrelated and in most cases took place more or less simultaneously. The study reports have gone through an intense process of consultation: the Concept versions of all reports were reviewed by the IPN project team members; the Draft versions of all reports were submitted for consultation to the broader research community. Two conferences offered the opportunity to discuss the evaluation methodology and the funding system with the Czech research community and last but not least, we ran a small pilot evaluation to test the evaluation methodology and its processes for implementation. This consultation and feedback process has led to various changes in the methodology and funding principles, updates to previously submitted reports and the conduct of additional analyses.

These continuous updates particularly affected the First Interim Report where the Evaluation Methodology is described. As a result, the denomination of the R&D Evaluation Methodology report as a “first” report is misleading; it constitutes both a first and final key result of this study.

Exhibit 2, below, shows the main and background reports that together constitute the final study report. Some of these background reports have been delivered in their final versions at the same time as the first interim report (background report 1) and the second interim report (background reports 2, 6, 7 and 8).

Exhibit 2 The Final study report: main and background reports



1.3 Guide to the report

To discuss the evaluation and funding systems, we need first to think about the international context of the Czech RD&I system. In Section 2.1 we discuss the performance-based research funding systems and its components in the context of international practice, while in Section 2.2 we cover the different types of research-performing organisations and their funding. Section 2.3 sketches the publication profile of the CR in international comparison. In Section 2.4 we report on our analysis of the RD&I system in the Czech Republic.

Chapter 3 describes the evaluation methodology and funding principles that we suggest. We set out the key principles and overall architecture of the evaluation and funding system in Section 3.1 and then give an overview of the evaluation methodology and the funding system in Sections 3.2 and 3.3, respectively.

Chapter 4 covers the needs for the legitimisation of the new evaluation and institutional funding systems and their integration into national research governance (Section 4.1) and for capacity-building in the RD&I system (Section 4.2) as well as the potential scenarios for the evaluation methodology management and eventual ‘extensions’ (Section 4.3). We conclude our report by providing recommendations for measures and actions that are to be implemented in the following period in order to ensure a proper functioning of the new R&D Evaluation Methodology and the Funding Principles (section 4.4).

Appendix A shows the OECD field categorisation that is used for the evaluation.

2. The Czech RD&I system in the international context

This Chapter is intended to provide the reader with some context based on international experience for the design of the new Evaluation Methodology and Funding System. Rather than burden the reader with details of practice in specific countries, we mostly focus on general principles here. The interested reader can learn more at the specific country level from the relevant background report to our Second Interim Report.

In this Chapter, therefore, we first discuss performance-based research funding systems in general and then analyse their two components: assessment on the one hand and institutional research funding on the other. Different types of research-performing organisations have different functions in the national research and innovation system so we go on to discuss a typology of research organisations (ROs) and describe the Czech organisations in these terms. Next we sketch some results from a bibliometric analysis of Czech research performance. We then go on to describe research-funding allocation in the Czech Republic and finally to make the link to the governance of the state's part of the research and innovation system. This is needed because an Evaluation Methodology and Funding System will only function properly when it is connected to the needs of society, as articulated by the ministries, the RD&I Council and other parts of government.

2.1 R&D evaluation and institutional funding in international practice

Most systems abroad make a clear distinction between evaluation (or 'assessment') and the use of assessment results in a funding formula and this distinction is central to the design of the new Evaluation Methodology and Funding System. The section on R&D evaluation is focused on drawing lessons from international practice in terms of concepts, methods and tools that can provide a basis for the future evaluation methodology. The section on the institutional funding systems, however, has a more pronounced comparative analytic focus. It also positions the funding system within the R&D governance system described in Section 2.4, below, drawing the relevant conclusions.

Internationally, countries are making growing use of performance based funding – whether through a performance-based research funding system (PRFS) or performance contracts – as part of a broader movement to increase performance, make universities more autonomous and introduce more strategic university research management. Governments increasingly use PRFS to

- Stimulate efficiency in research activity
- Allocate resources based on merit
- Reduce information asymmetry between supply and demand for new knowledge
- Inform research policies and institutional strategies
- Demonstrate that investment in research is effective and delivers public benefits¹

In the tradition of New Public Management², PRFS seek to increase accountability for the expenditure of taxpayers' money. They are seen as a means for selectively

¹ Abramo, G., D'Angelo, C. D., & di Costa, F. (2011). National research assessment exercises: a comparison of peer review and bibliometric rankings. *Scientometrics* (89), 929-941

² Ferlie, E., Ashburner, L., Fitzgerald, L., & Pettigrew, A. M. (1996). *The New Public Management in Action*. Oxford: Oxford University Press; Boston, J., Martin, J., Pallot, J., & Walsh, P. (1996). *Public Management: The New Zealand Model*. Oxford: Oxford University Press

distributing research funds, but most governments also seek to use them to drive particular behaviours.

2.1.1 R&D evaluation

Evaluation broadly defined is part of the policy process. It may be conducted *ex ante* as well as *ex post* and be at a range of levels including projects, programmes, funding or performing organisations, research fields or disciplines. Assessment in the context of a performance-based research funding system (PRFS) tends to be narrower, since it aims specifically at the *ex post* evaluation of research performance and is intended to be part of an intervention that itself changes performance. Rather than solely providing an absolute judgement, it tends also to be involved in deciding the relative merits of different research groups or organisation so that they can be rewarded with differing amounts of money. Research assessment can of course also be done solely in order to provide information – intending to influence behaviour through publicity rather than via direct financial rewards, as the Dutch and Australian systems do. Research assessments are conducted at national level, but can often also provide information at the institutional level.

In practice, evaluations have different functions and objectives and these are reflected in the choice of methods and indicators used to assess performance. A common denominator is, however, that an evaluation should analyse social effects of intervention, not just focus on outputs. PRFS have different objectives, which influence the way they work. For example, the UK system fosters ‘excellence’ by skewing rewards towards those with strong publications; the Norwegian model has a linear relationship between the production of quality publications and monetary rewards, and has strengthened research capacity in the weaker parts of the system rather than focusing money on ‘excellent’ researchers or institutions.

Internationally, different types of research organisations are evaluated using different systems. This is as true of PRFS as of broader systems of evaluation. Thus there are separate systems for universities, scientific institutes, research and technology organisations (RTOs) and other research-performing organisations. Different types of research organisation have different roles in society so they are not put into competition; rather, policy decisions are made about how much activity the state wants to support in each category. Competition may then be fostered within each category.

Assessment systems generally use certain minimum thresholds, rather than being ready to tackle very small entities. This is important not only for efficiency but also because meaningful judgements cannot be made about the performance of groups if there are few data available. Notably, bibliometric methods break down where there are small numbers. While it is possible to look at characteristics of individual publications such as the extent to which they are cited, a widely used rule of thumb is that if there are fewer than 50 publications then little can be said that is systematically comparative.

The research community in general tends to prefer peer review to other assessment or evaluation methods – not only because this makes it possible to temper indicator-based information with human judgement but also because this allows the community itself to bring its own values to bear. Peers are legitimate in the scientific community, can tackle context, understand inter-field differences and can additionally provide formative information. On the downside, their use tends to mean that those evaluated also have to do some work such as preparing self-assessments or other kinds of submissions, the independence of the peers needs to be checked, they may be more expensive than indicator-based methods and they can be hard to recruit. And, of course, peers occasionally misbehave – though their work is done very much in the public eye and in the presence of their colleagues. In our experience, laziness is a bigger risk than cheating.

Many recent PRFS focus on indicators rather than peer review, though they also tend to reallocate only a modest proportion of total institutional funding so the imperfections associated with an indicators-based approach are not so important. Indicators are inexpensive and their cost is falling as bibliometric databases and tools become cheaper to use. Those assessed may not have to do much work in order to be part of the assessment process. Indicators can produce finer-grained distinctions and therefore rankings than peers – but the question has to be asked whether this amounts to real or spurious precision. Problems with indicators include the fact that inter-field comparisons are difficult, the use of indicators is usually unsophisticated and sometime methodologically incorrect, it is hard to get indicators that explain the context of research (as opposed to research outputs) and they lack a formative dimension.

There is therefore a lot to say for ‘informed peer review’, where peers make use of the best available indicators together with other information in reaching judgements. This offers ‘the best of both worlds’ as well as allowing comparison between indicator-based and judgement-based views of performance and therefore a degree of ‘triangulation’ across methods.

In principle, both peer-review and indicators based assessment can be extended down to the level of the individual researcher but this is rarely done in PRFS. Indeed, it is only done in specialised systems that assess individual researchers as part of a promotion or award process (as in current Spanish and German systems). Equally, not all assessment systems include all researchers. In some, institutions have to choose those researchers whom they regard as ‘excellent’.

In research assessment, a great deal of effort goes into tackling differences among fields. In peer review based systems this generally means that different fields are not put into direct competition with each other. Mechanisms are put in place to ensure that peer review panels all use assessment scales that have the same meanings, irrespective of discipline. Putting different fields into competition within metrics-based approaches means that the designer of the assessment has to produce some sort of bibliometrics-based technique for comparing across fields without, in reality, having a deep understanding of what these inter-field differences are.

Whichever approach is used, assessment systems – like wider evaluation systems or systems for deciding on the quality of research proposals in research councils and other funding agencies – struggle with interdisciplinary research. Peer panels may have to cooperate to address it. Indicator-based approaches do not have a good frame of reference for interdisciplinarity precisely because the statistical pictures they use to describe ‘good’ research are defined within disciplines. As in other situations where it is hard to define very precise rules, there is virtue in using (peer) judgement to find solutions to these problems.

A final stark lesson from how things are done abroad is that the former Czech approach (the Metodika or ‘coffee mill’) was unique internationally in using indicators to allocate a high proportion of institutional funding for research.

2.1.2 Institutional funding for research

The decentralised nature of innovation systems and their governance means that the mix of policies and funding instruments in RD&I systems tends to be emergent, ie it generally results from the interplay of many independent policy actors’ decisions and cannot necessarily be predicted in advance. It is generally hard to coordinate at any level, despite the fact that there is a genuine need for coordination. Correspondingly, changes in institutional funding principles, such as the introduction of a PRFS, generally affect only a small part of the overall mix. They are also introduced incrementally in order to provide sufficient time for adjustment on both sides – principals (funding providers) as well as agents (research organisations).

In international practice, research organisations tend to be steered and funded through various combinations of unconditional block funding and performance-based funding, which may be based upon a performance contract and/or a system that counts or assesses results of research, ie a PRFS. Introduction of different components of institutional funding makes it possible to combine elements of continuity, stability and incentives for a desirable change in behaviour. It also enables differentiation of institutional funding schemes for different types of ROs. Also noteworthy is that a considerable proportion of many institutional funding systems remains unconditional. The performance contracts may not even be linked to funding.

Performance contracting is widespread – not only on its own but also in combination with a PRFS. Performance contracts play an important role in the institutional funding systems in the comparator countries as instruments for dialogue between ROs and their principals (the Ministries). They also support the principals in the implementation of policy: they provide the principal with opportunities to encourage ROs to achieve agreed goals that are in line with the principal's policy goals, within the limits of the ROs' autonomy. However, none of the countries that use performance contracts appears to have clear expectations about the consequences of sub-standard performance.

Experience in comparator countries indicates that performance contracting brings with it a significant need for monitoring. The research performers need to report at least yearly and there tend to be monitoring or 'dialogue' meetings at regular intervals so that both the principal and the RO are clear about the extent to which the requirements of the contract are being fulfilled. This is in stark contrast to the current situation in the Czech Republic where the Ministries have no monitoring function in relation to research performance and play a far less active role when it comes to defining long-term strategic directions for research conducted by the ROs compared to their peers in most of the other countries.

PRFS appear to reduce the need for monitoring, since the principal is in principle only interested in results, rather than the processes through which the RO produces the results. However, for those PRFS that – like the UK REF – are only operated at extended intervals, this does mean that the responsible ministry can only periodically understand the research performance of the research organisations.

The PRFS are generally seen as useful ways to add an element of competitive pressure to institutional funding – but only at the margin. Most PRFS make up a small component of the overall funding system for research and higher education. Consequently the PRFS component affects only a small part of the total funding for a given RO and the introduction of PRFS abroad has normally resulted in only marginal changes to the funding mix. Our impression is that the systems that only drive a small part of the institutional funding nonetheless are effective at changing behaviour and performance. They are also likely to be less contentious because design flaws only have small effects on the funding system as a whole.

Internationally, the use of PRFS is believed to improve both the amount and the quality of the output from research organisations. PRFS can be tuned to reach different policy objectives, for example to concentrate research in a small number of institutions or to boost overall research capacity across the system. International experience with PRFS is that they tend to increase both the quantity and the quality of research outputs. It is nonetheless easy to encourage perverse effects and promote 'gaming' and there is evidence that they favour traditional and mainstream research approaches over heterodox and interdisciplinary ones.

The Czech Republic is unique in attempting to use a single performance-based system to address all the different types of research organisation. Despite the fragmentation of the Czech research system and heterogeneity of ROs in relation to their mission, there is to be one single evaluation methodology and set of funding principles for universities, institutes of the Academy of Sciences and for other research organisations

– all are evaluated according to the same methodology and funded according to the same principles.

In international practice, the significance of single components in institutional funding tends to differ among different types of ROs. In the case of PRFS there is a strict differentiation between different types of ROs in order to ensure competition among ROs with the same or similar function in the research system – and avoid meaningless competition between research organisations with different missions. RTOs and PSROs do not receive institutional research funding through the same type of PRFS as other actors, reflecting their rather different research missions compared with the university and scientific institute sector. Bringing them together in a single PRFS must be done with some care.

2.2 Research-performing organisations and their institutional funding

In international practice, we can distinguish among five different types of research-performing organisation. At some points in this report we refer to the first two collectively as ‘scientific research organisations’.

- Universities – or, strictly, research universities (since some universities only teach and do not do research)
- Scientific research institutes – which, like research universities, conduct fundamental or applied research but either have no teaching responsibilities or only provide education at PhD level (alone or in combination with a university)
- Research and technology organisations (RTOs) – which conduct applied research and experimental development and provide technical services to support industrial innovation. These can be distinguished from technical consultancies in that they receive institutional funding from the state
- Government labs or public service research organisations – these do research and provide technical services on behalf of government. They produce knowledge the government needs in order to legislate or regulate or they produce ‘public goods’ such as standards, certification or weather forecasts that society needs but that private companies lack the incentives to make. Some of these institutes are run by private organisations on behalf of government
- National resources or infrastructure – such as libraries and museums, which enable others to do research and which may otherwise be needed for social, educational or cultural reasons. Normally, these need to do some research of their own in order to support their infrastructural function but research is not their primary function

The comparator countries have rather stable systems of research performing organisations. Changes are introduced by government policy – for example, as part of a decision to set up a new university or to privatise a Government lab – though in practice these are rare events. Entitlement to institutional funding generally follows automatically as a consequence of the policy decision. It is not possible ‘bottom up’ to create a new research organisation and apply to the state for institutional funding.

The Czech Republic’s institutional structure is less stable than many others, in the sense that there have been changes since the transition of the political system and there is still a process of readjustment in progress between the Academy and the universities. Like Latvia and Lithuania, the Czech Republic permits a bottom-up process where research organisations can apply to become recognised as research organisations. In current practice, this makes them also eligible for institutional funding. The experience in all three countries is that this fragments the research performing system.

At the time of writing (March 2015), there are in total 223 Research Organisations, including those recommended by the RD&I Council to be recognised as such in February 2015. Research organisations are defined in line with the 2006 “Community framework for state aid for research, development and innovation”:

“Research organisation’ means an entity, such as university or research institute, irrespective of its legal status (organised under public or private law) or way of financing, whose primary goal is to conduct fundamental research, industrial research or experimental development and to disseminate their results by way of teaching, publication or technology transfer; all profits are reinvested in these activities, the dissemination of their results or teaching; undertakings that can exert influence upon such an entity, in the quality of, for example, shareholders or members, shall enjoy no preferential access to the research capacities of such an entity or to the research results generated by it.”

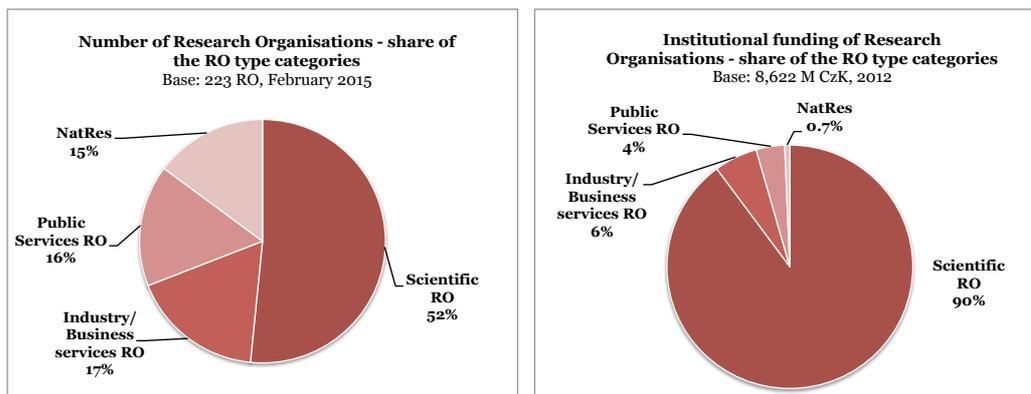
An in-depth analysis of the profile of these Research Organisations³ showed that *in the Czech Republic*, the current Research Organisation encompass a broader range of types of research institutions than the ones listed above. We categorised them as follows:

- Scientific Research Organisations – ScRO are institutions that have as primary function to conduct research to the benefit of the research community. This includes institutions that have as primary activity the conduct of research (the ASCR research institutes) and/or the teaching and training of future researchers, ie the public and private HEIs, and the research hospitals. It also includes institutions that conduct research in order to improve their services to the research community, ie the research infrastructures.
- Industry & Business services Research Organisations – IBRO are institutions that have as primary mission to develop and transfer technologies and knowledge to the benefit of the industry and business sector. This category includes RTOs and (an increasing number of) consultancies offering expert services or other professional services to industry and business entities.
- Public Services Research Organisations - PSRO are institutions that have as primary mission to develop and transfer knowledge and technologies to the benefit of the Public Sector. They are Government labs or consultancies offering services to the public sector.
- National Resources – NatRes. These Research Organisations provide cultural services: they collect and curate national or regional cultural public goods and provide access to the public and researchers. This category of ROs includes archives, museums, and galleries.

From a numeric perspective, the Scientific RO category accounts for approximately half of the ROs, while the three other RO categories take up a close-to-equal share of the RO base (between 15% and 17% - In terms of institutional funding actually allocated to the RO type categories, the Scientific ROs accounted for about 90% of the institutional funding for research organisations in 2012 (see Exhibit 3, below).

³ See the report *Typology of the Research Organisations and the Effects of the EM Thresholds* (Background report 2)

Exhibit 3 RO type categories – shares in overall number and institutional funding



Notes: Data on the number of ROs are based on the list of ROs in February 2015; data on institutional funding are based on 2012 data

Source: Návrh výdajů státního rozpočtu České republiky na výzkum, experimentální vývoj a inovace na rok 2012 s výhledem na léta 2013 a 2014 (tabulková část)

Since July 2014, a new EU Commission Regulation (GBER) is in force, extending the concept of ‘research organisation’ to ‘research and knowledge-dissemination organisation’. The Commission hereby (indirectly) recognises the important role of Knowledge Intensive Business Services (KIBS) for economic growth. KIBS include organisations offering ‘High Knowledge Intensive Services’, ie fundamental research, industrial research or experimental development, and ‘Knowledge Intensive Market Services’, ie technical and professional consultancy services.⁴ A technical amendment to the Czech Act is in the process of approval, which will transpose all changes in the Community Framework into the Czech legislation (including the definition of ROs).

The question that arises is whether this broader definition of ‘research and knowledge-dissemination organisation’ fits the purpose of identifying the organisations that can - and should - be eligible for institutional research funding.

The principles according to which various countries provide ROs with institutional funding are based in the economics of knowledge - we discuss them in the Final Second Interim Report. These principles are simple and apply in most countries. Central are

- Institutional funding is justified according to the extent to which ROs produce public goods, or
- Whether the state appoints an organisation to produce such knowledge on its behalf

As an overall principle, research organisations may receive institutional research funding irrespective of their legal form – but those that are organised as private companies must in practice operate as not-for-profit organisations (typically in the sense that they do not distribute dividends to their shareholders) and must recognisably be serving a public rather than a private purpose.

⁴ See the report *Typology of the Research Organisations and the Effects of the EM Thresholds* (Background report 2)

Some cases arise in the Czech Republic that in our view violate some of these principles – almost entirely in the case of certain industry and business services ROs (IBROs).⁵ Of course, provided it follows prevailing law the Czech state is fully within its rights in doing whatever it wants. Nonetheless, our advice would be to make use of internationally normal principles in deciding which organisations should be considered for institutional research funding.

- Research conducted in Scientific Research Organisations is generally expected to produce knowledge that is public goods. Inherently, the market cannot produce these so the state has to pay. That implies a need to provide institutional research funding. It is the business of the state to decide which organisations it regards as ScROs for this purpose. In most countries public universities are agencies of the Education Ministry or can otherwise be instructed by the government. Only in the Czech Republic are private universities eligible for state institutional research funding, applying the rulings that apply also for other private organisations
- IBROs also produce public goods – but usually only as a fraction of their overall activity. These organisations receive institutional funding so that they can develop knowledge and capabilities that are one step ahead of industry’s needs. Conceptually, they are provided with institutional funding to help pay for public goods production.

The subsidy or institutional funding distinguishes the IBRO from a commercial consultancy in two ways.

- First, it enables the production of public goods that would not be possible in a fully commercial organisation
- Second, it marks the IBRO as having been chosen by the state to play a particular role in the national innovation system

Based on these public goods, the IBRO then engages in support activities (RD&I, training, testing etc) for industry; these are normally priced so they at least cover their costs. Thus the IBRO has two types of activities – research and the exploitation of knowledge for the benefit of industry. A technical consultancy only has the latter and therefore receives no subsidy.

This also means that a company cannot spin its research department off into a subsidiary organisation in order to get institutional subsidy if it intends to monopolise the knowledge produced. In most countries such organisations would not in any case be recognised or appointed by the state as one that should produce research public goods on its behalf.

- Public Service Research Organisations (or ‘government labs’) are set up by the state in order to provide it with services and/or to generate public goods. There tends to be no other customer than the government for such services so the state normally owns such organisations. Whatever the form of ownership, the state pays through institutional funding (and often additionally through a contract for non-research services). However owned and irrespective of their legal form, PRSOs have some kind of performance contract or operating contract with their owning ministry. This reflects a decision by the state to have and to operate a particular facility. Private entities cannot simply set up as PRSOs and ask for institutional funding.

⁵ See the report *Typology of the Research Organisations and the Effects of the EM Thresholds* (Background report 2)

- National resources such as libraries and museums are not normally given institutional funding for research. Often the owner is part of the central or regional state but they may also be private. While it may be beneficial for them to conduct some research, this is not their dominant activity or their main purpose.

The economics of knowledge also have a direct consequence on the level of institutional funding: the closer knowledge gets to market application, the more companies are able to monopolise aspects of the knowledge and secure economic returns. So the state plays a big role in funding basic research (usually paying all the costs) but invests a far smaller proportion in those cases where it intervenes in more applied areas that are closer to market. The logic maps in a simple way across to different types of RO.

- Scientific ROs tend to produce rather fundamental knowledge where there is a high level or most failure, so the state pays all or most of the cost
- PSROs largely produce knowledge needed only by the state or that are public goods, so they are normally also heavily subsidised as are infrastructure, for which private funders are unlikely to be found
- In contrast, IBROs work close to market with industrial problems whose solution will enable one or more companies to make money. The need for subsidy is therefore smaller, so IBROs internationally tend to get 10-33% of their income as core funding whereas the other types of organisation considered here get a great deal more

2.3 Research profile from a bibliometric perspective

In the report *Bibliometrics on and for the Czech Republic* (Background report 3) we report on the outcomes of a bibliometric analysis in which the research performance of the Czech Republic is benchmarked against that of the USA, the UK, Germany, Poland and the Slovak Republic as well as the average of the European Union as a whole.

Research profiles are relatively stable over time as a result of long-term scientific traditions and priorities of a country. Exhibit 4, below, shows that the profile of the Czech Republic is most similar to the profiles of Poland and Slovakia, with relatively more emphasis on the Natural Sciences, Engineering and Technology, and Agricultural Sciences. The emphasis on Natural Sciences and Engineering is also shared with Germany. On the other hand, there is less emphasis on Medical and Health Sciences than can be seen in the EU in general and in the United States.

It is important to bear in mind that the Social Sciences and Humanities have limited coverage in the Web of Science or Scopus – the main databases of peer-reviewed research publications, conventionally used for bibliometric analysis⁶ There is also a concentration on Anglo-American journals in these databases. Consequently, these areas are probably more important in the Czech Republic than the exhibit below suggests.

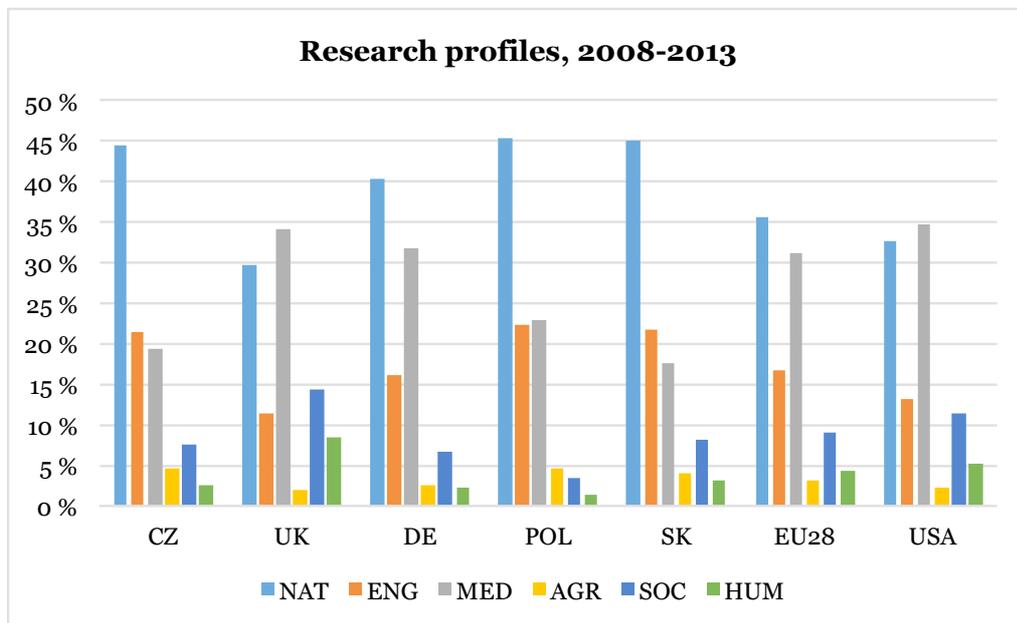
The trend in the last 5 years has been towards an increase in international co-publication and citation impacts in the Czech Republic, compared to the international level.

The average citation impact of publications from the Czech Republic recently has increased from a level below to a level above the average of the EU countries, making it

⁶ See the report *The RD&I Information System as an information tool for evaluation* (Background report 10)

comparable to the impact of publications from the USA and Germany. Our results also show that the citation impact of the Czech Republic is higher than that of Poland and Slovakia, although all three countries have been clearly improving.

Exhibit 4 Comparative analysis of the CR research profile, 2008-2013



Source: Web of Science publications, 2008-2013

Some field-specific differences appear.

- The increase in the citation impact of publications from the Czech Republic is seen in all major areas except Agricultural Sciences, where the situation is stable below world average. (The low scores here may be influenced by national journals in Agricultural science that have been included in the Web of Science.)
- The Medical and Health sciences stand out with an especially high average citation impact for the Czech Republic. However, the share of articles among the ten per cent most cited is on the average, indicating that there is regularly a few very highly cited articles from the Czech Republic in this area each year.
- The highest increase in impact is in the Social Sciences. However, in this relatively small area (within the Web of Science), the fluctuations may also be dependent on a few highly cited publications.
- There is a potential for increased impact, relative to the largest and most influential countries, in the two major areas that are heavily emphasized in the research profile of the Czech Republic: Natural Sciences and Engineering and Technology.

2.4 The RD&I System in the Czech Republic

2.4.1 The RD&I state budget expenditures

The Czech national RD&I state budget expenditure is subdivided into two major groups of instruments: institutional expenditures and competitive expenditures.⁷

- Institutional expenditures cover the funding lines for ‘the long-term conceptual development of Research Organisations’ (ie institutional funding that is distributed to the ROs, currently based on the evaluation of the results they have achieved) and a set of other budget lines, including costs related to the international co-operation of the CR in R&D; co-funding of the operational programmes; project management costs (appraisal, monitoring and evaluation); costs for major infrastructure; and expenditure related to the activities of the RD&I Council, the Czech Science Foundation (GACR), the Technology Agency (TACR) and the Academy of Sciences of the CR (ASCR)
- Competitive expenditures refer to the funding of grants, research programmes, specific university research, and major infrastructure

The Ministry of Education is the main actor providing public support to research. In 2014, it distributed the largest share of the state budget line for institutional funding for ROs (56%), followed by the ASCR (32%), and it is the main funder of competitive research programmes (28% of the competitive expenditure budget in 2014), followed by the Grant Agency (24%) and the Technology Agency (21%). The Ministry of Industry plays a minor role, allocating 8% of the competitive funding budget in 2014.

The process for the distribution of institutional funding among the funding providers (the Ministries and the Academy of Sciences) is similar to international practice: it is based on political negotiations and decisions rather than research organisations’ evaluation results or previous commitments for the funding of on-going RD&I activities.

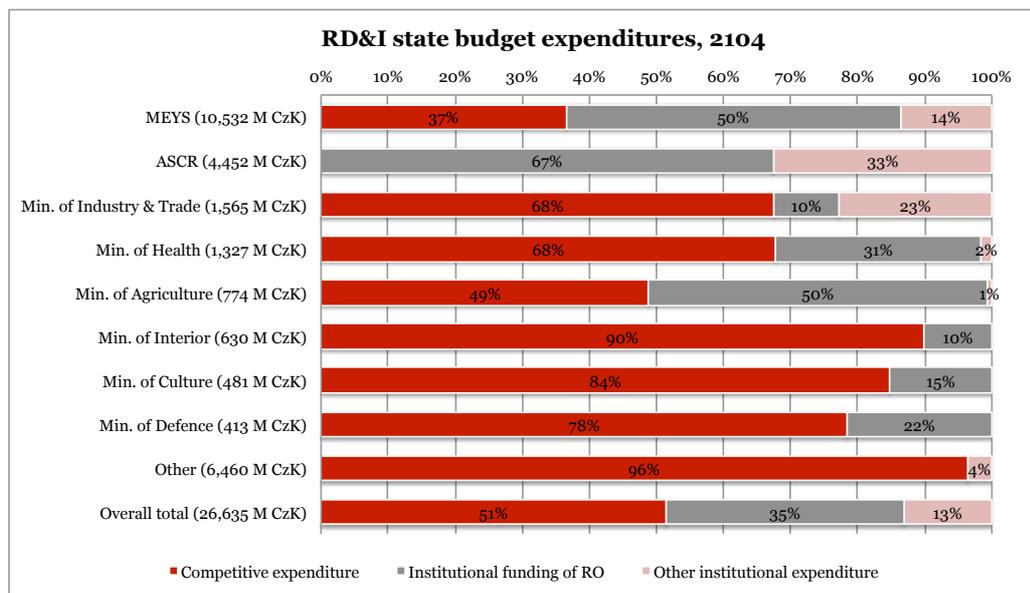
Ministries vary in their use of the RD&I funding mix (Exhibit 5, below): the Ministry of Education and Ministry of Agriculture invest 50% of their budget in the institutional funding of the research organisations in their field of competence, compared to the 10% of their RD&I budgets allocated for institutional funding by the Ministry of Industry. All of the ‘minor’ Ministries (the Ministry of Interior, Culture and Defence) allocate 80% or more of their RD&I budget through competitive funding programmes.

In the international landscape, the Czech Republic is unusual in being among the minority of countries with a high ratio of project-based to institutional funding of research. In 2014 ~50% of the RD&I State Budget was allocated in the form of competitive funding (Exhibit 5). When excluding the ‘other’ institutional expenditure lines from the total, the share of competitive funding versus institutional funding for ROs is 60% versus 40%, respectively. Generally, governments fund the majority of R&D through institutional rather than project-based funding. Most of the funding systems internationally involve a high proportion of institutional funding in the overall mix.⁸ International practice suggests therefore that it might not be wise to make a large proportion of Czech institutional funding contestable in the short term (as was the case with the old Metodika). Rather, an element of stability – in the sense of slow change rather than stasis – would be a good way to produce a balance between predictable institutional funds and more variable competitive funding.

⁷ See the Final report 2 – *The institutional funding system in the Czech Republic* (Background report 7)

⁸ See the Final report 2 -*The Institutional Funding Principles* (previously: Second Interim Report)

Exhibit 5 RD&I State Budget Expenditures by the funding providers, 2014



Notes: ‘Other’ stands for the funding allocated to and distributed by the Grant Agency, the Technology Agency and the Office of the Government

Source: State Budget Act, Technopolis analysis

2.4.2 The RD&I governance structure

As international practice indicates, a PRFS – and, indeed, the wider elements of research funding – is not a ‘neutral’ policy tool but is an instrument for encouraging the research system to behave in ways that are consistent with state policy. It is therefore important that the state apparatus is able to set policy and implement it through good governance of the research system.

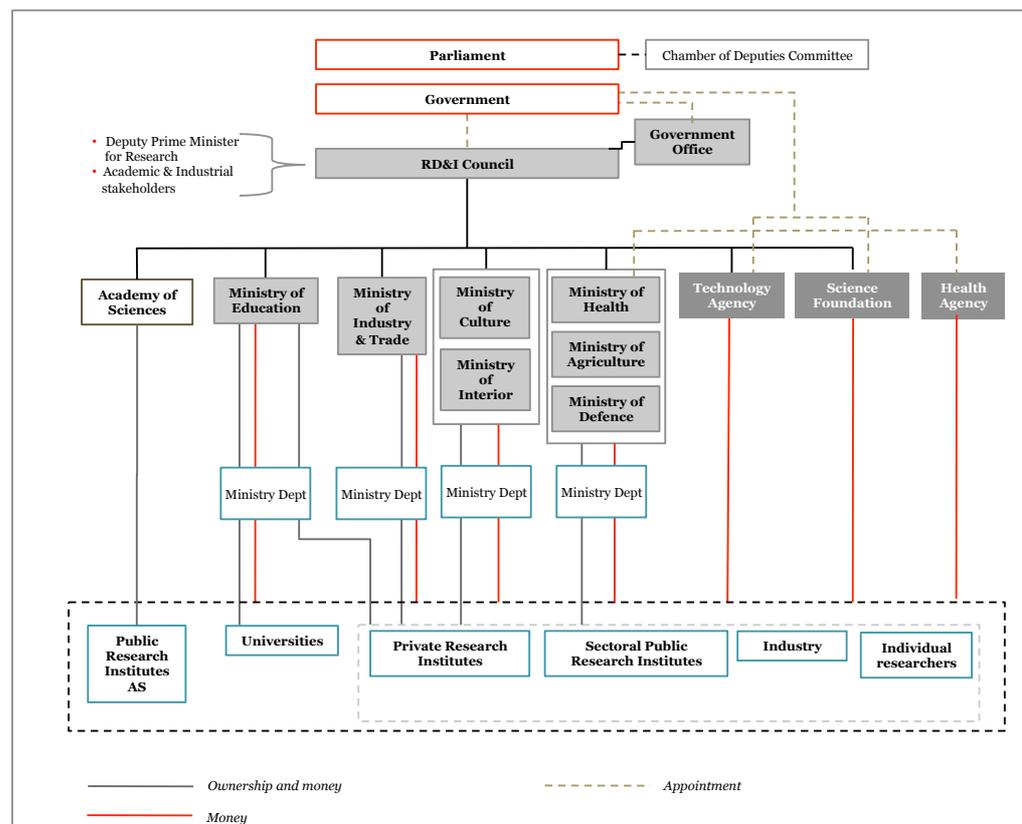
In the Czech Republic, the main national policy document is the Research, Development and Innovation Policy 2009 – 2015, implementing the 2008 Reform of the RD&I system and restructuring the RD&I governance system (Exhibit 6).

The Council for RD&I (hereafter: Council) is the highest body that advises the Government on the priorities, budget and the overall organisation of the RD&I System in the Czech Republic. It is the key ‘arena’ in which overall research and innovation policy can be coordinated. A member of the Government - currently the Deputy Prime Minister - acts as the Chairman of the Council, reinforcing its legitimacy.

A set of seven ministries, the Academy of Sciences and three agencies (the Science Foundation, the Technology Agency and the Health Agency), which are responsible for the implementation of the RD&I policy, constitute the second ‘intermediary’ level.

The 2008 Reform gave three ministries (the Ministry of Defence, the Ministry of Health, and the Ministry of Agriculture) responsibility for sector-specific RD&I. Four ministries (the Ministry of Education, Youth and Sport, the Ministry of Industry and Trade, the Ministry of Culture, and the Ministry of Interior) are in charge of “cross-sectoral RD&I”. The Ministry of Transport, the Ministry of Environment, the Ministry of Labour and Social Affairs, and the Ministry for Regional Development, instead, lost their status of public RD&I funding providers, which led to a dismantling of their RD&I departments. The responsibility and budget for RD&I of these ministries was transferred to the ‘cross-sectoral’ ministries and the TACR.

Exhibit 6 The RD&I governance system in the CR



Source: Technopolis, 2014

All the RD&I funding-providing ministries manage institutional funding for the research organisations in their area of competence. The Academy of Sciences has a similar status for its research institutes. Most of the ministries also develop and manage competitive R&D programmes ('targeted' research funding). An exception is the Ministry of Industry that officially does not have the responsibility for competitive funding programmes, even though it currently funds a competitive research programme (TIP), lasting until 2017.

The research community in the Czech Republic is composed of researchers employed in universities, Academy of Sciences research institutes, 'sectoral' public research institutions (including state research organisations), private research institutes, and industry. Key actors in public research are the public universities and the research institutes of the Academy of Sciences.

The system for institutional funding of research is rather decentralised with a high degree of autonomy for both funding providers and recipients (ie the research organisations). Laws guarantee the funding bodies a relatively high degree of autonomy in terms of institutional funding allocation and distribution. Factors such as budget constraints and lack of resources, however, have meant that the funding bodies have not always been able fully to utilise their autonomy.⁹ Only the Academy of Sciences has developed a specific evaluation methodology and strategies for the

⁹ See the report *The institutional funding system in the Czech Republic* (Background report 7)

internal distribution of institutional funding. Such an evaluation system is required if a funding organisation is to use other ways of distributing funding internally than the division implied by the Metodika. With the exception of the ASCR institutes, the outcomes of the Metodika evaluation therefore govern the allocation of institutional funding at the level of the specific research organisations rather than the level of funding bodies as it was originally intended.

The RD&I organisation and governance structure of the Czech Republic have some characteristics that put its functioning at risk.

A first key aspect of R&D governance systems that need to function well if they are to produce good policy is *policy coordination*. Policy coordination matters because it provides a way to ensure the overall coherence of research and innovation policy – an aim is to make sure that one part of the system does not rely on another part to deliver something, which it turns out is not delivered. It is increasingly important as research and innovation policies have to confront the ‘grand’ or societal challenges such as climate change, ageing, HIV/AIDS and so forth, which cut across the needs and abilities of individual parts of the system such as ministries to deal with them.

The organisation and governance structure is effectively the way in which the state connects R&D activities to social needs. Government is a major influence. But the individual ministries also have a strong say, with each representing a particular ‘sector’ of society. In principle, each ministry has an understanding not only of the overall needs of its sector but the kind of research needed to advance knowledge and develop policy. In many cases, some of this need will be expressed through a research institution ‘owned’ by the ministry.

In research policy budgeting as in policy more generally, therefore, the spending ministries make competing claims about their needs – and have an annual battle with the finance ministry about how many of these claims can be afforded. The relative power of government centrally and the individual ministries varies among systems but the outcome is rarely the result of a simple top-down decision. It emerges from the competition among claims – a competition that can in many systems benefit from being expressed in an ‘arena’ such as a policy council, where it is possible to coordinate and negotiate. Such an arena may also be helpful to focus the national effort in pursuit of a strategy, make sure all the needed parts of the system function and make it possible for the national system to change direction when circumstances change.

A second key aspect is the need for *‘distributed strategic intelligence’* – in the sense of a wide availability of data and sources of information. Especially for comparatively decentralised governance systems to operate well, the individual ministries need to be able to analyse, express and lobby for the satisfaction of their own sector needs. This means that people with relevant skills and capabilities and the independence not only to undertake the needed analysis but also to be able to present its results to the wider policy community must populate them.

Among the countries discussed in this study¹⁰, the Czech Republic might appear to be the best placed organisationally to coordinate research and innovation policy because it has an over-arching RD&I Council closely linked to government. However, a key weakness of the Council is the fact that it comprises almost entirely members of the research community, as opposed to incorporating wider stakeholder interests - inside and outside government. Further, it is an advisory body of the Government without real executive power. The RD&I governance structure in the Czech Republic appears therefore as decentralised, lacking an official central co-ordination body.

¹⁰ See the report *R&D governance and funding systems for research in international practice* (Background report 6)

In this context of a decentralised governance system, the 2008 Reform of the RD&I system had a major negative effect on the capacity of the system to govern sector RD&I. It reduced the number of Ministries and other public administration bodies with competences for RD&I funding, creating a situation where a number of sector ministries, such as the Ministry of Environment, lacked (or were deprived of) budgets to fund institutional R&D costs – and therefore to maintain a body of research capabilities and evidence relevant to developing their policies. A requirement for research funding and policy to be effectively connected to social needs via the state is for individual ministries to define and manage the funds they inject into the research-performing system through institutional funding as well as through ‘targeted’ funding. Effectively, this means that those ministries whose research budgets were earlier centralised need to be re-enfranchised by having their budgets restored to them. They may in some cases need to strengthen their internal policymaking capacity in relation to RD&I. This capability would be immeasurably strengthened if the Czech Republic would implement the overdue reforms aimed at strengthening the civil service.

3. The Evaluation Methodology and Funding Principles

This Chapter has three components. First, we discuss principles that underlie the design of the Evaluation Methodology and Funding Principles and its overall architecture. Second, we give a summary of the Evaluation Methodology and its main components (Section 3.2). Third, we give an overview of the institutional funding system (Section 3.3).

3.1 Key principles and overall architecture

Some of the key principles that we used in the design of the Evaluation Methodology and Funding Principles were given to us through the Terms of Reference for this study; others derive from our understanding of international experience and practice; yet others address specificities of the Czech RD&I system and its governance. We outline these principles in this section, addressing in turn: those that guide the development of the whole system; those that pertain to the assessment process; and those that underlie the design of the funding mechanism.

We then go on to provide an overview of the architecture of the new Evaluation Methodology and Funding Principles. The remaining sections of this chapter go into the specific design of the assessment and the funding components. Those readers needing more background or detail can refer to our main reports *The R&D Evaluation Methodology* and *The Institutional Funding Principles* (previously: the First and Second Interim Reports), which respectively refer to the assessment and the funding components of the new evaluation and institutional funding system.

3.1.1 Key principles

Performance-based research funding systems (PRFS) get a lot of attention internationally, from both policymakers and the research community. However, as our discussion of international experience shows, PRFSs are essentially policy instruments and they can be tuned to achieve a range of different policy objectives. These strategic policy decisions therefore drive the focus and scope of the evaluation, the type of evaluation (summative and/or formative), the choice of assessment criteria and their indicators as well as the choice of institutional funding criteria and mechanisms.

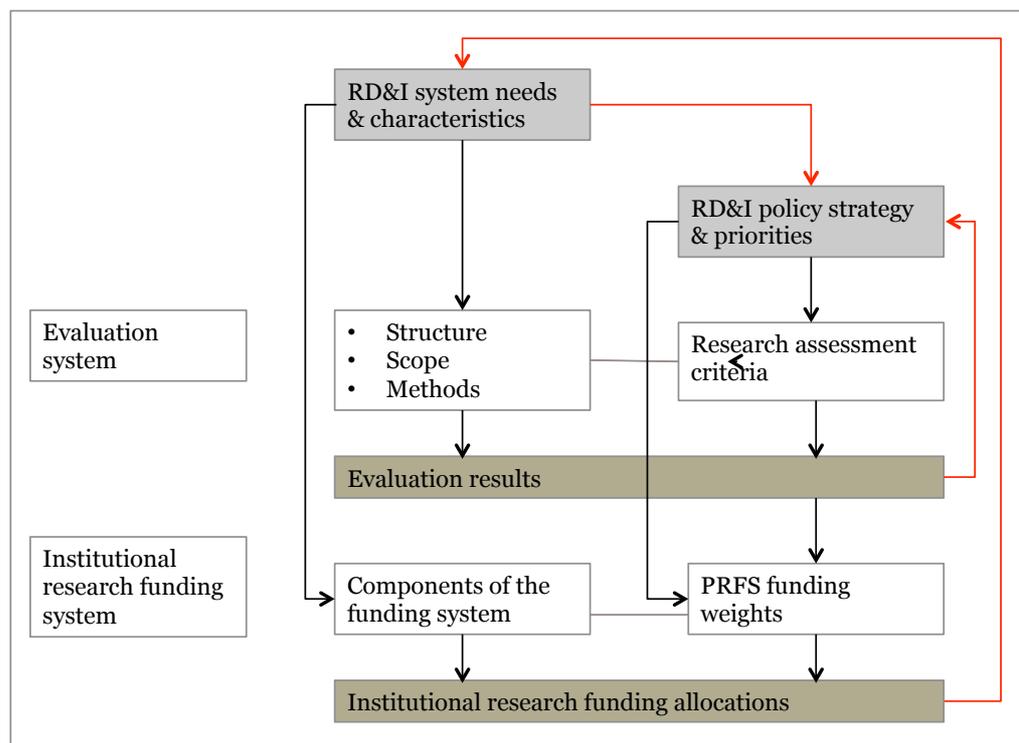
PRFSs need therefore to be seen as parts of the policy system (Exhibit 7). They can allow policymakers to understand and influence the shape of the RD&I system, acting as source of strategic information to all actors involved as well as an opportunity to steer behaviour and induce structural changes – both at the RD&I governance and research-performing levels.

A corollary of this is that the Evaluation Methodology and Funding Principles do not stand alone but need to be under the coordinated control of the state funders, who ‘own’ the research-performing organisations and provide their institutional research funding. In this way, policy needs can serve as drivers for change – but also as drivers for the elements of stability and predictability that are needed in any well-functioning RD&I system.

The history of the Metodika in the Czech Republic shows that a PRFS can itself induce *instabilities* in funding. The new Evaluation Methodology and Funding Principles are intended to reduce this instrument-induced instability. But historically the other key instability in Czech institutional funding has been policy-induced, underlining the importance of the link between the Evaluation Methodology and Funding Principles and the overall governance system.

The effects of the evaluation and funding systems in terms of the new needs for public support they create, will need to be monitored and assessed, ensuring continuing consistency between policy, strategy and priorities and these needs.

Exhibit 7 The Evaluation Methodology and Funding Principles in the policy system



Some of the specific needs addressed by the Evaluation Methodology and Funding Principles in the Czech Republic are

- De-fragmentation and modernisation of the structure of research-performing organisations. A key element is supporting the development of the universities, whose role has been under change since the transition
- Fairness and transparency – things that are important in all systems but that have a particular urgency in the Czech one because of a history that promotes distrust not only of government in general but also of some of the mechanisms the international research community uses in order to govern itself and maintain quality. For the Czech RD&I system to be part of the world community (as it must be, because you cannot do good science on your own), it must also embrace the norms of quality control and management that prevail in the rest of the world
- Both the evaluation and the institutional funding system need to take the diversities among the actors involved into account in order to ensure fairness in the assessment of research performance and public support given to these different Research Organisations. This requires adequate consideration of the specifics of the scientific fields in which the researchers are active as well as the different missions in society of the categories of Research Organisation
- Formative inputs to policymaking overall and to the specific development of individual research groups and organisations. The choice of assessment criteria and the way in which the results of using these criteria are reported have been conditioned by this need and are discussed in more detail below. In the context of the Czech Republic, an Evaluation Methodology and set of Funding Principles, Janus-like, looks in two directions at once. It looks backwards and rewards performance but it also looks forwards, providing advice and support for the future development of the system

- The Evaluation Methodology and Funding Principles are designed to avoid setting different types of research organisation in competition with each other. A decision to reallocate money between categories of research organisation is essentially a policy decision in which the state chooses among investments with different social purposes and should not be taken automatically through the operation of a PRFS

The overall institutional funding system should reflect the need for stability via so-called 'block' funding, performance-based change via the PRFS described in this report and development needs through a system of performance agreements concluded between research performing organisations and their principals.

A final overall principle pertains to the requirements in our mandate to design an integrated system that covers a range of different types of research organisation and a wide diversity of evaluation requirements. Other countries do not mix these things together; doing so inevitably has consequences for the extent to which the overall system is optimal in individual applications, eg in evaluations at an institutional level.

3.1.1.1 Key principles – Assessment

In this study we developed the methodology for a **National Evaluation of Research Organisations – NERO**. We defined the key principles for the NERO as follows:

- The evaluation is a national evaluation system focused on the assessment of performance in research and development. It will reflect the strategic policy objectives for the Czech RD&I system, take into account its needs and characteristics, and cover all types of research organisation
- The assessment of research performance will take place at the level of a field-defined Research Unit (RU) within an Evaluated Unit (EvU). Participation in the evaluation will require a minimum of critical mass
- The evaluation will be a panel-based process of informed peer review. The expert panels will draw on a mix of appropriate quantitative and qualitative information to make their professional judgements. The evaluation process will be conducted in English. It will consist of remote reviews and remote assessments, complemented by panel meetings
- The evaluation will be a fair and egalitarian system. It will use a single framework for assessment across all disciplines and types of research organisation while allowing for a reasonable level of field- and RO type-specific variations. It will be comprehensive, covering all dimensions of the research activities and research outputs, outcomes and impact, as well as the conditions facilitating good research performance
- The evaluation results will be predominantly qualitative. They will show the quality levels reached against each assessment criterion, with explanatory texts, and provide conclusions and recommendations for future development. Conclusive analytical reports will be developed at the level of EvU, field and disciplinary area, aggregating information collected at the RU levels
- The evaluation process will be fully transparent
- The cost and burden of the evaluation will be the minimum possible to deliver a robust and defensible process

Evaluation results need to be based on human judgement via expert panels. Only peers, experts in the fields, have the needed understanding of the field specifics and the role and positioning of the different ROs in these fields. They will be informed by metrics, but essential for their evaluation will be the qualitative information provided by the evaluated Research Units (RUs).

Unity in the assessment method is key to achieving ‘field-neutral’ evaluation scores for each RU, applying common criteria across the different fields and types of Research Organisations (but enabling the panels to take into account variations to these criteria). It also constitutes a critical requirement of the Evaluation Methodology as a source of strategic information at the national level.

3.1.1.2 Key principles – Funding

The most common understanding for institutional funding is that it needs to provide continuity, stability, sustainability and resilience for institutional development, and that a long term shortfall of institutional funding leads to a ‘hollowing out’ of research organisations.¹¹ Research is inherently a rather long-term activity, so there is a particular need for the institutional funding system to provide a degree of stability and freedom to make strategic investments (a) through the mechanism of performance contracts and (b) by damping excessive movements in institutional funding between funding periods. This latter element has to be built into the structure of the funding model: what proportion of institutional funding should be contestable in any one funding round?

Rationales commonly cited for granting institutional funding are that it allows research organisations to behave in a strategic, long-term oriented fashion, reducing the risk of converging on ‘hot areas’ with a consequent loss of diversity in the wider research system. It also provides space for researchers to develop ideas that may not be ready yet for exposure to external competition. In many countries, institutional funding also pays for equipment and support services of generic benefit that would not be easily attributable to individual grants.

We conceptualise institutional funding in the context of a long-term working relationship between a ministry in charge (the principal) and a Research Organisation (the agent). Public institutional funding expresses responsibility and ownership by the state, with the state taking an interest and a stake in the Research Organisations it (co-)funds.

This aspect seems to be particularly important in the Czech Republic where the public sector’s ‘ownership’ of Research Organisations can be weak. This also implies that context and history have to be taken into account and that institutional memory is important, on both the principal’s and the agent’s side.

Institutional funding in this context provides a basis for institutional development and therefore it must not be subject to short-term policy making. Ideally, public institutional research funding and its basic principles should be laid down in a long-term RD&I strategy which is supported by all major political forces and thus holds beyond election terms.

The significance of the different sources of funding (institutional, competitive) differs for different types of research organisation because they have different missions and functions in society and their research is embedded in different contexts, addressing different user communities (other researchers, industry, public sector etc.). Furthermore, for these and a number of additional reasons such as location or subject specialisation, different research organisation do not have equal access to competitive funding.

Hence, an important principle must be that the funding system needs to take into account the different missions and specifics as well as the *entire* funding situation of Research Organisations. Ideally, the ministries responsible base their funding

¹¹ Luke Georghiou, Effects of Public Funding Systems on Research in the United Kingdom. Presentation at ‘The Impacts of Impact Measurement’, Vienna, 25 September 2014

decisions on a thorough understanding of the Research Organisations’ ‘funding profiles’, ie a complete overview and analysis of the various funding sources per research institution, including competitive funding. This is necessary to understand how changes in institutional funding for Research Organisations affect their overall budget.

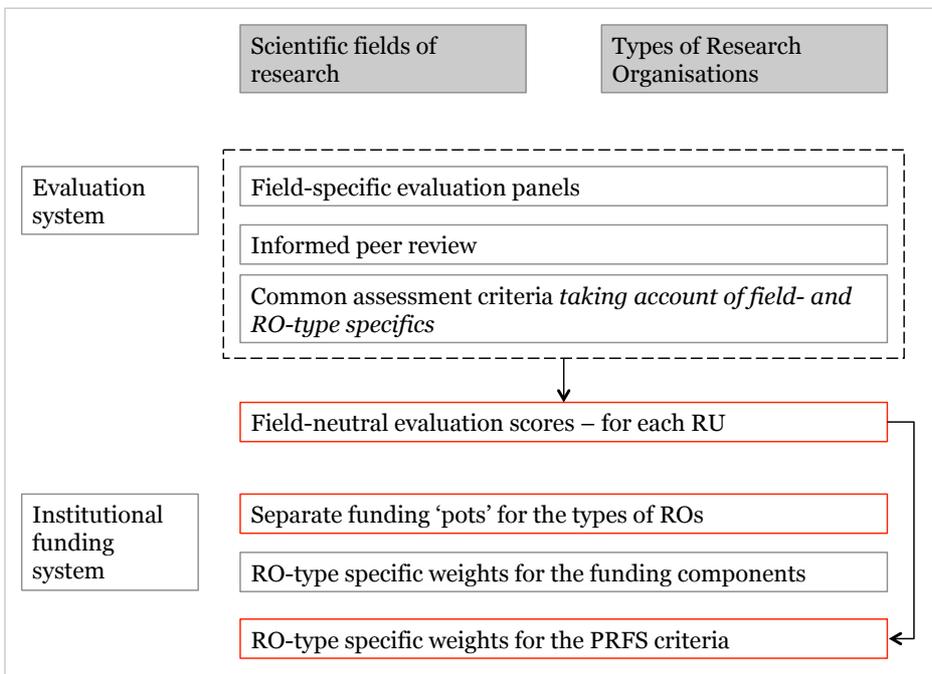
In international practice, the share of institutional funding in relation to the overall budget is different for different types of organisations. The rule of thumb is: the closer a research organisation is to the market, the less institutional funding it receives (see for example the Fraunhofer Society compared to the Max Planck Society in German). In the end, it is a policy decision what share of institutional funding a research organisation should get. In many countries, there are various arrangements at work, making the institutional funding landscape quite ‘messy’. However, in most countries this is considered the normality of a differentiated multi-level and multi-actor science and research system.

Institutional funding is awarded not only according to the assessment criteria discussed but also in proportion to the size of the respective units being evaluated. ‘Size’ in this context can be a tricky concept. In principle, it is the size of the ‘establishment’, the amount of permanent research capacity provided by the research unit. This is most conveniently conceptualised as the number of full-time equivalent (FTE) researchers, which the unit is committed to maintain. Inherently, therefore, it excludes students, temporary workers and so forth.

3.1.2 The architecture

Key to the overall design is the idea that assessment and funding are **separate** processes. This is reflected in the two-step structure of the Evaluation Methodology and Funding Principles (Exhibit 8).

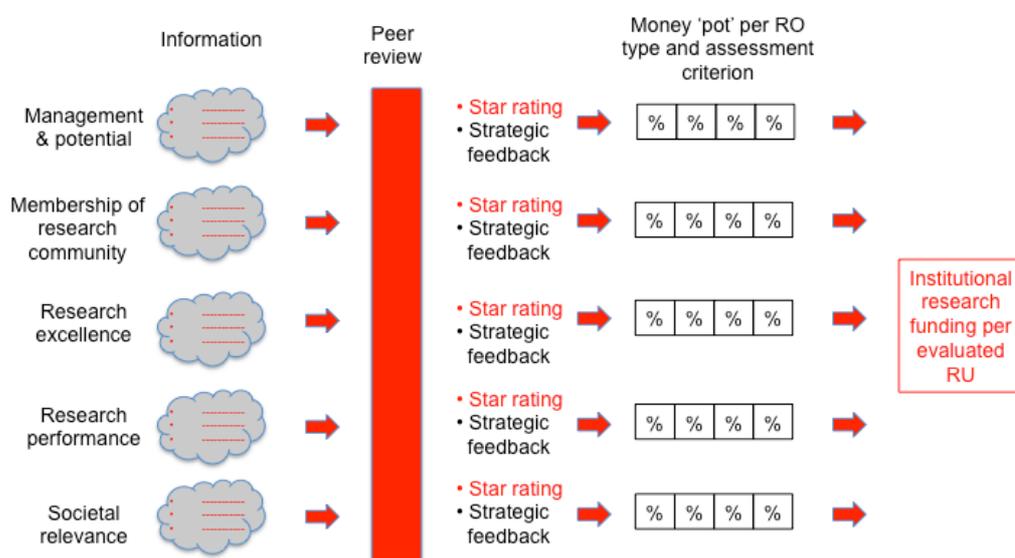
Exhibit 8 Two-step approach to ensure fairness in institutional research funding



Scientific fields and the different Research Organisation types should not be set in competition with each other. We achieved this by adopting a two-step approach: the evaluation system is centred on handling the scientific field specifics while taking the missions of the ROs duly into consideration; the institutional funding system centres on recognising the different missions of the RO types in society.

The expert panels will not assign an overall score to the evaluated RUs in their field. This avoids setting the Research Units that were registered by the different Research Organisations in competition with each other, which would risk a bias by the expert panels in favour of the Scientific Research Organisations. The evaluation results will consist for each RU of the scores and related explanatory texts against each criterion, together with the panel conclusions and recommendations. Exhibit 9, below, shows the mechanism for the translation of the evaluation scores against each criterion into funding, in the context of the performance-based research funding system (PRFS) component of the institutional funding system.

Exhibit 9 Evaluation in the context of a funding mechanism

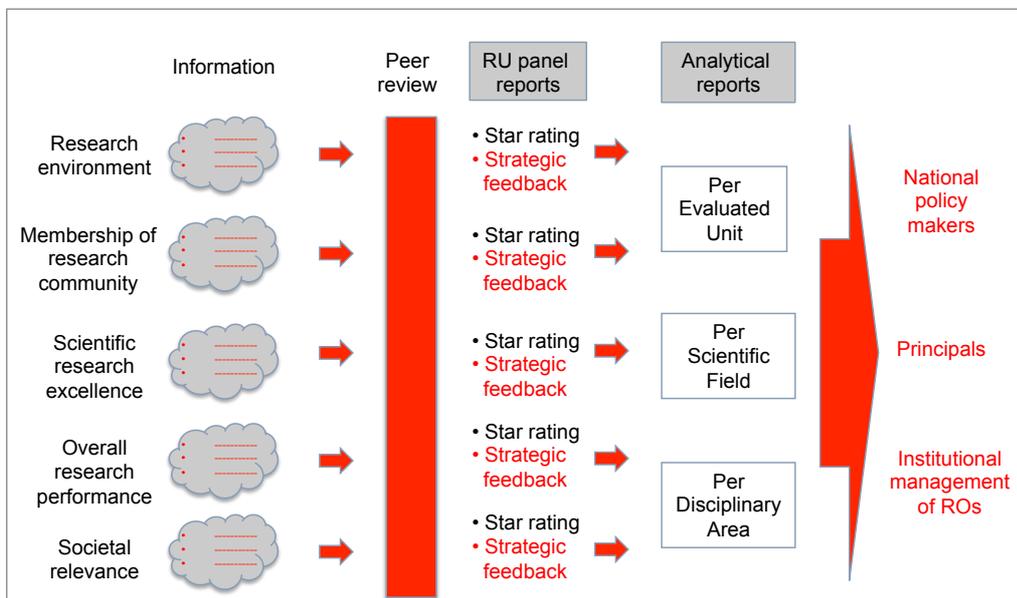


A key value of the Evaluation Methodology lies in its **formative** function, ie in the development of strategic information. For this purpose, the Evaluation Methodology is not conceived as an arithmetic exercise. It stresses the importance of qualitative information to inform both the evaluation panels and the actors in the RD&I system responsible for the definition of policies and strategies. These include the national policy-makers in the Government and the RD&I Council, the funding providers (the Ministries and the Academy of Sciences – further collectively called “principals”) and the institutional management of the Research Organisations. For the benefit of these actors, the Evaluation Methodology foresees the development of conclusive analytical reports at the Evaluated Unit, scientific field and disciplinary area levels, building upon the panel reports at the Research Unit level (Exhibit 10).

A major benefit that national policy-makers will get from this information is in the identification of the country’s fields of strengths and weaknesses in international competition and the eventual need for systemic policy interventions. For the principals a major value will be in a more in-depth understanding of the strengths and weaknesses of the ROs within their competence. This will support them in the development of a stronger principal-agent relationship with these ROs, which is a key requirement for the implementation of the institutional funding system. The evaluation results may, for instance, provide the basis for the discussion on the

performance agreements. Institutional management will also obtain, for example, an improved view on the positioning of the RO in the national and international landscape, informing its institutional research strategies.

Exhibit 10 Evaluation in its formative function



3.1.3 The evaluation structure

We have defined the evaluation structure as follows:

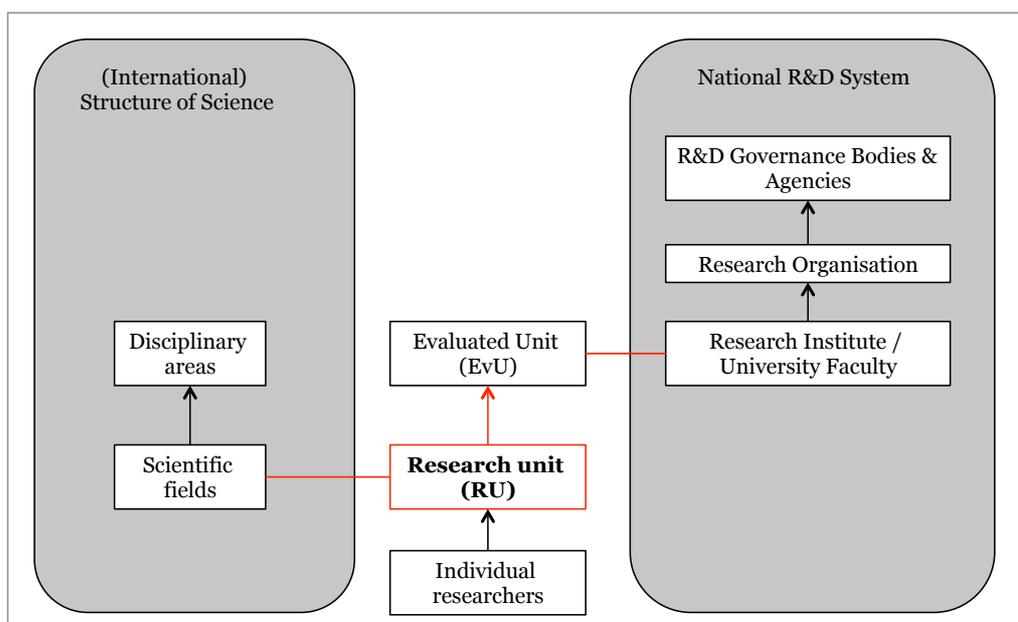
- An Evaluated Unit (EvU) is a research organisation, except for the public HEIs where the Evaluated Unit is a Faculty or Institute or any other organisational unit at that level such as Centres
- A Research Unit (RU) means the group or groups of staff in the Evaluated Unit that conduct their *primary* research in a specific field, and by extension the structures and environment that support their research and its application or impact. A Research Unit may comprise staff that work in multiple departments or organisational units in the Evaluated Unit

Each researcher can be assigned only to one Research Unit in an Evaluated Unit. An Evaluated Unit can register only one Research Unit for a specific field. The scientific fields are the ones defined in the OECD field classification, shown in Appendix A.

We have defined the primary unit of evaluation, ie the Research Unit, at the intersection of the ‘natural’ dimension for evaluation of research - the scientific field - and the dimension determined by the need for information at the level of institutions (Exhibit 11). This will enable the assessment of the RUs’ role, positioning, and competitive value in the national R&D and innovation system as well as in the international R&D landscape. As research is becoming more and more international, and competition – as well as collaboration – in research is at a global level, a view of the performance and position of the research actors in the Czech Republic relative to the international landscape is a critical piece of strategic information, for any actor in the R&D system.

The evaluation is not conducted at the level of individual researchers. International experience (in particular in PRFS) shows that any inclusion of even a single component of the evaluation methodology at the level of individual researcher risks having considerable negative unintended consequences. Any use of the evaluation results - or part of the evaluation results - at the individual researcher level is therefore inappropriate.

Exhibit 11 The Research Unit as primary unit of evaluation



The choice of the RU as the primary unit of analysis means that RU results can be combined in certain ways, in order to contribute to higher levels of analysis.

- First and most important, the combined results from analysis of its component RUs provide a basis for understanding the performance of EvUs. This will be the most frequent use of aggregated evaluation results
- EvU results can further be aggregated to the level of ROs, proving organisational management with a 'map' of performance and opportunities as an input to decision making and R&D agencies with an understanding of the capabilities of potential beneficiaries
- They can also be combined to provide a picture of national performance at the level of fields and scientific areas – another key input for science policy at the national level
- The proposed system is of more limited value in evaluating individual (targeted) programmes. They address individual researchers rather than RUs. The opportunities here are (1) to inform the programme evaluation about the quality and characteristics of the RUs within which their funds are spent and (2) to use 'tagged' data from the national database of R&D outputs as one input into the wider set of questions that programme evaluations should ask

3.2 The Evaluation Methodology (EM)

This chapter sets out the core components of the methodology that we developed for the National Evaluation of Research Organisations – NERO. It synthesises the content of the Main report *The R&D Evaluation Methodology* report (previously: First Interim Report) and its Background Reports.¹²

We start by describing the scope of the evaluation, ie the type of research-performing organisations that will be covered by NERO, and their categorisation (Section 3.2.1). In Section 3.2.2 we expand the description of the evaluation structure given in the previous chapter defining also the requirements for Research Organisations that are to take part in the evaluation. The key features of the evaluation panels, their structure and working processes, are described in Section 3.2.3 and Section 3.2.4, respectively.

Section 3.2.5 gives an overview of the assessment criteria, while Section 3.2.6 covers the information that will be collected for the assessment against these criteria. Two aspects that are of particular importance in this context are further described in more detail: the data on the FTE researchers (Section 3.2.7) and the data and indicators related to research outputs (Section 3.2.8).

Section 3.2.9 indicates the current and potential role of the RD&I Information System for collecting and processing this information for the evaluation panels. In Section 3.2.10 we cover some major topics related to the implementation of the National Evaluation of Research Organisations. We conclude with an estimate of the costs for the whole evaluation process in Section 3.2.11.

3.2.1 Scope of the evaluation

In this study we defined the methodology for the National Evaluation of Research Organisations (NERO). The term ‘Research Organisation’ (RO) stands for organisations that have been granted this status by the RD&I Council. These encompass a variety of types of organisations, which we structured in four categories based upon their mission in society: Scientific Research Organisations (ScRO), Industry & Business services Research Organisations (IBRO), Public Services Research Organisations (PSRO) and National Resources (NatRes) (see Section 2.2, above).

For the purpose of this evaluation, we consider the categorisation of the Academy of Sciences research institutes, universities, hospitals and medical research institutes, and research infrastructures in the category Scientific Research Organisations (ScRO) to be mandatory. The same is true for the categorisation of libraries, archives, museums and other similar Research Organisations *in the cultural sphere* in the category National Resources (NatRes). The other Research Organisations will be asked to indicate at the moment of registration for the evaluation whether they belong to the category of Industry & Business Research Organisations (IBRO) or Public Services Research Organisations (PSRO).

In line with the rulings for the recognition of Research Organisations, research infrastructures are explicitly covered in the EM only if they are independent legal entities, recognised as Research Organisations. Research infrastructures that are not

¹² Background reports specifically relevant for the design of the evaluation methodology are the reports *Evaluation systems in international practice* (Background report 1), *Typology of the Research Organisations and the effects of the EM thresholds* (Background report 2), *Bibliometrics on and for the Czech Republic* (Background report 3), *Detailed evaluation cost framework* (Background report 4), and the *Evaluation handbook* (Background report 5)

recognised as separate Research Organisations are included in the assessment if they are organisational units in the ASCR/universities.

In both cases, we categorise the research infrastructures (RI) as Scientific Research Organisations. This categorisation is based on the understanding that a specific evaluation methodology for research infrastructures is in place, developed by the IPN Team. This methodology covers the evaluation of *all* research infrastructures, independently of their legal status, and focuses on the specific features that define their efficiency, effectiveness and quality in performance.

In the case of the NERO, the results of which informs part of the institutional funding for *research*, the focus for assessment is the quality of the research activities conducted *within* the organisational structure of the RI. As with any other Research Organisation, the effects of the research on the user community will also be taken into account.

3.2.2 The units of analysis and eligibility for participation in the evaluation

The evaluation takes place at the level of Evaluated Unit (EvU), ie a Research Organisation or for the public HEIs, the organisational unit at the second level of the organisation's structure (faculties, institutes, centres etc). All Research Organisations and Evaluated Units can participate in the evaluation, on a voluntary basis, provided there is a minimum of critical mass. A minimum volume threshold for participation is set at 50 eligible research outputs over the evaluation period.

Each eligible Evaluated Unit may register one or more Research Units for participation in the evaluation. A Research Unit means the group or groups of staff in the Evaluated Unit that conduct their *primary* research in a specific field, and by extension, the structures and environment that support their research and its application or impact. A Research Unit may comprise staff that work in multiple departments or other organisational units in the Evaluated Unit. An Evaluated Unit staff member can be part of one Research Unit only.

A Research Unit is registered for a *field* of research and an EvU can register only one Research Unit per field. A minimum volume threshold for the registration of a Research Unit is set at 50 eligible research outputs¹³ over the evaluated period. For Interdisciplinary Research Units a minimum of 50 eligible research outputs across two or more fields applies.

The scientific fields and their categorisation in disciplinary areas are based on the OECD field classification (see Appendix A to this report). The use of the OECD fields enables comparability and compatibility of field-related data with international data and data systems. A major advantage of the OECD FOS fields is also that they are broader than the ones currently used in the Czech system. The OECD classification categorises approximately 190 sub-fields into 36 fields and 6 disciplinary areas (instead of the 123 fields and 11 groups of fields currently used in the Czech system). There is a system in place to re-categorise data on research outputs in the RD&I Information system following the OECD structure.¹⁴

¹³ In the field of the RU, based upon the indications provided at the moment of registration of the research output in the RD&I Information System

¹⁴ Appendix B to the report *The RD&I Information System as an information tool for evaluation* (Background report 10) shows the mapping of the fields currently used in the RD&I Information System (IS) to the OECD categorisation

A **minimum volume of research activity** is required for participation in the evaluation. Eligible research outputs¹⁵ that are taken into account for the minimum volume thresholds are

- Scholarly outputs - Papers in peer-reviewed journals (J), Conference proceedings (D), Monographs, books and book chapters (B)
- Non-traditional scholarly outputs - Results used by the funding provider, ie projected into legislation or norm, projected into non-legislative or strategic documents (H), Research report containing classified information (V), Certified methodologies, art conservation methodologies, specialized map works (N)
- Patents and other IP - Patents awarded (P), Plant/ breeders rights (Zodry & Zplem)

The following rules apply for the volume calculations:

- Books will count as 4 research outputs
- Co-publications of researchers employed in the same EvU will be de-duplicated (ie counted once)
- Co-publications of researchers active in different EvUs will be counted as one each, thus providing an incentive for collaboration among the Research Organisations

Taking into account the variety of Research Organisations in the Czech RD&I system and to ensure fairness in the evaluation system, we have made sure that the level of these thresholds and the type of research outputs included do not discriminate against applied research organisations.¹⁶

3.2.3 The evaluation panels

The use of evaluation panels, involving experts in the scientific disciplines, constitutes the primary tool by which the Evaluation Methodology implements a field-specific evaluation.

This section gives an overview of the evaluation panel structure, the profile of the experts involved and their working processes and rules. More details on the profiles of the experts and the rulings and processes for the panel evaluation can be found in the *R&D Evaluation Methodology* report (previously: First Interim report) as well as in the *Evaluation Handbook* (Background report 5).

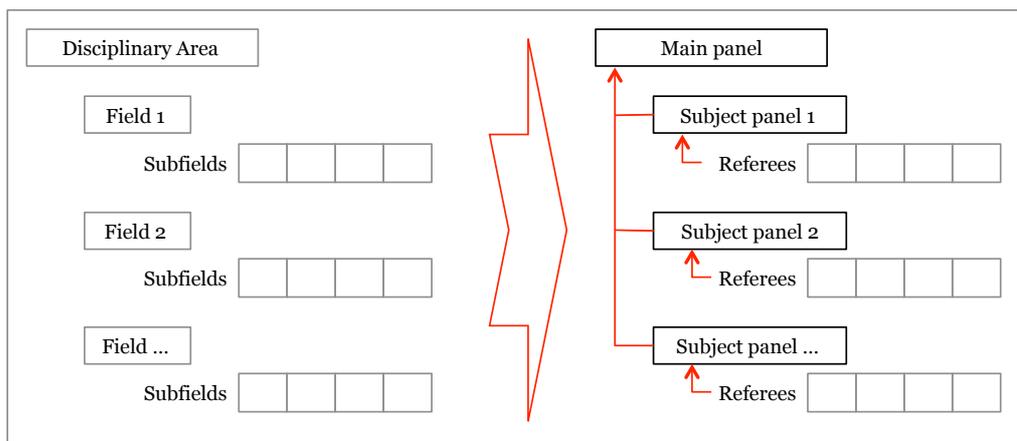
The evaluation is entrusted to a core structure of Main panels, Subject panels and referees (Exhibit 12, below).

- There will be 6 Main panels, organised at the level of *disciplinary area*, and 24 to 26 Subject panels, organised at the level of *fields*. Referees will assess submitted research outputs and will work at the level of *sub-field*
- The Main panels will have a Chair and 3 additional members
- Subject panels need to be small and high-level. Our recommendation is to keep the number to 5 max 6 members per panel

¹⁵ The definition of these outputs is the one provided for the Metodika 2013-15

¹⁶See the report *Typology of the Research Organisations and the Effects of the EM Thresholds* (Background report 2)

Exhibit 12 Organisation of the evaluation experts in main and subject panels



Subject panels are defined at the level of OECD field. However, this does not imply that 36 panels will be established. In the preparatory phase of the performance assessment, Subject panels will need to be defined, taking into consideration the volume of the research conducted in the CR in the specific fields, in terms of number of research units and research outputs produced over the evaluation period. The intent is to spread the assessment workload over the different panels as equally as possible as well as to reach the maximum level of efficiency during the evaluation process. There are various scenarios: two or more ‘small’ fields, ie for which a small number of Research Units are registered, can be covered by 1 subject panel; average sized fields are covered by one subject panel; particularly ‘large’ fields can be covered by two (or more) subject panels, ie the RUs in the field are distributed over multiple Subject Panels. In the latter case, specific measures will be taken in order to ensure consistency in approach among the different Subject Panels. We explain this in more detail in the *Evaluation Handbook* (Background report 5).

Only **international** experts will be in charge of the performance assessments. This takes account of the high risk felt in the Czech RD&I community that the assessments would be biased if they involved national experts because of nepotism and clientelism. Demanding requirements are set for the eligibility of the international experts in terms of conflicts of interest (see the main report *R&D Evaluation Methodology* for further details).

These international experts will be supported by a number of national experts, ie the Main panel members and Specialist advisors. The main role of these experts will be to act as *advisors* and provide knowledge on the national context; they will not play an active part in the assessments. Experience in the Small Pilot Evaluation confirmed the usefulness of this background information for the international experts in order to shape their judgments.

For the Subject panel members, breadth should be prioritised over depth. In other words, the experts involved in the Subject panels, and especially the panel Chairs, should have a broad overview of the field and sub-fields and major areas rather than expertise in one field only. In each Subject panel, some members should have expertise also in inter-disciplinary research or in application areas (for which different fields of disciplines are often combined or integrated).

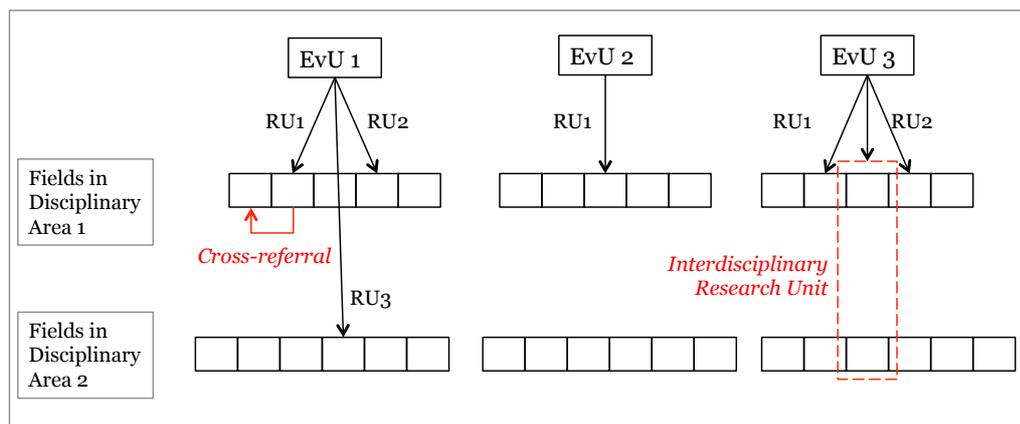
Referees will have fine-grained expertise. They will be in charge of reviewing the selected outputs submitted by the RUs only. They will work remotely, in order to keep costs down.

Interdisciplinary research can take place at different levels, ie between sub-fields in an OECD field, between 2 or more fields within a disciplinary area, or between 2 or more fields across disciplinary areas (Exhibit 13).

- Interdisciplinary research between sub-fields will be handled directly by the Subject panels. In case inter-disciplinary research in the RU covers different fields in a disciplinary area, an Evaluated Unit can recommend *cross-referrals* among the Subject Panels.
- An EvU that has a strong element of interdisciplinary research in fields across disciplinary areas can apply for registration an *Interdisciplinary Research Unit*. It will need to make a convincing case that interdisciplinary research is a significant feature of the research in the IRU and demonstrate that there is a strategic connection and strong collaboration between the researchers active in the different fields. If the application is accepted, an Ad-hoc panel will be set up, chaired by the most relevant Main panel Chair.

Readers needing more background or detail on cross-referrals or Interdisciplinary Research Units can refer to our main report *The R&D Evaluation Methodology* (previously: First Interim Report) and the *Evaluation Handbook* (Background report 5).

Exhibit 13 Research Units and the structure for interdisciplinary research



3.2.4 The panel evaluation process and its results

The evaluation will be implemented by means of remote reviews and remote RU assessments combined with evaluation panel meetings (Exhibit 14, below).

The evaluation process starts with a Subject panel meeting that will ensure a common understanding of the evaluation methodology and the assessment criteria and sub-criteria among the Subject panel members. During this meeting a calibration exercise will be performed during which the field- and RO type-specific variations to the Evaluation Methodology will be defined. We cover this further in Section 3.2.5, below.

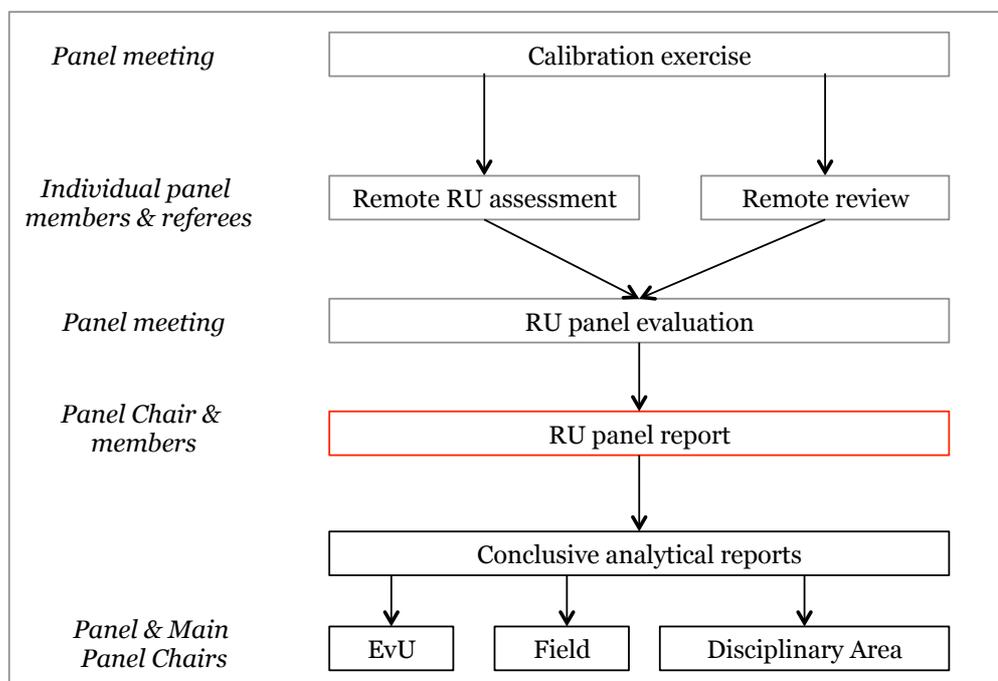
The next step consists of a remote assessment of the RU (by the Subject panel members) and a remote review of the selected research outputs submitted by the RU phase (by the Referees). Both the Subject panel members and the Referees will assign quality level scores and provide explanatory texts for their decisions about the quality levels assigned. Two referees will review each submitted publication; two Subject panel members will do the remote assessment for each RU.

The final RU panel evaluation is performed during a panel meeting. During this meeting, the individual RU assessments and their outcomes will be discussed and the panel members will reach a common, final view.

Summary report

The panel members provide judgments about the Research Units' performance for each of the 5 assessment criteria through a starred quality levels scoring system, using scores from 1 to 5. The values for the attribution of these quality levels, ie the criteria for the panels to take their decision (eg a quality level score of 2 rather than 3) are described for each assessment criterion. We list these in the report *The R&D Evaluation Methodology* (previously: the First Interim report).

Exhibit 14 The evaluation process



The **outcome** of the evaluation is an RU panel report, containing the quality level scores against each assessment criterion, explanatory texts for each score, conclusions and recommendations. The panels will not allocate an overall score for the RU; this makes it possible to use the 5 assessment dimensions to create incentives for 5 different but desirable behaviours in the funding system (see Section 3.1, above).

The evaluation will also result in a set of conclusive analytical reports at the EvU, field and disciplinary area levels. These reports will take the form of a panel-based aggregation of RU-level judgement to the higher levels.

- The Subject Panel Chair will be responsible for the analytical report at the field level
- The Subject Panel Chair for the major field of involvement of an EvU will be responsible for the analytical report at the EvU level, based upon a draft developed by the Evaluation Management Team and in cooperation with the other relevant Subject Panel Chairs
- The Main Panel Chairs will draft the analytical reports at the disciplinary area level, in co-operation with the Subject Panel Chairs

Further description of the panel evaluation process and its outcomes as well as the templates for the Panel reports and Analytical reports are provided in the *Evaluation Handbook* (Background report 5).

3.2.5 Overview of the assessment criteria

The EM uses a single framework for the assessment of all scientific disciplines and RO types, thereby ensuring full comparability of the evaluation results across all dimensions. Comparability is a fundamental condition for the use of evaluation results in a performance-based research funding system (PRFS). A starred quality level should have the same value for all field-specific RUs evaluated and the assessment should base itself on the same set of indicators and information. We explained this concept in Section 3.1, above.

The evaluation is structured around five assessment criteria that **jointly** enable the fulfilment of the strategic objectives of the evaluation and funding system, ie to reward excellence in research while building capacity, in research and for innovation. The main assessment criteria are: the research environment, membership of the national and global research community, scientific research excellence, overall research performance, and relevance for society (Exhibit 15).

The five assessment criteria are not independent. A (good) research environment, in particular (good) management, is the basis for (good) overall research performance while overall research performance is one basis for societal relevance and scientific research excellence. Societal relevance requires (good) overall research performance because societal relevance based on low-quality research is useless. Scientific research excellence may occasionally thrive without good overall research performance but this is rare and not normally sustainable. Finally, overall research performance and scientific research excellence are the basis for membership in the research community. The interconnectedness among the assessment criteria demonstrates the fundamental role of the research environment, in particular management, and overall research performance.

The formulation of these criteria, ie the description of the values for the 5-point scale starred quality levels, is inspired by international practice, but has been adapted to the realities of the Czech R&D system, while taking care to enable the exercise to describe a sufficient spread of quality within the range of national performance. An exception is the assessment of scientific research excellence.

Exhibit 15 Assessment criteria and sub-criteria

Assessment criteria	Sub-criteria
Research environment	The quality of the research management (including HR management)
	The adequacy of the research strategy
Membership of the global and national research community	International research presence and collaboration
	National research presence and collaboration
Scientific research excellence	
Overall research performance	Research output (including quantity and overall quality)
	Competitiveness in research
Relevance for society	

In principle, the assessment criteria and sub-criteria are relevant for all types of Research Organisations, no matter what type of research they conduct or the field in which they are active. However, scientific fields exhibit significant differences in the

way they conduct research¹⁷ and within those fields, the different types of Research Organisations differ in their roles and missions. Our reflections below highlight the relevance of the assessment criteria for all ROs and the differences in importance that we see from an overall perspective.

- The criterion *research environment* covers the availability of research infrastructure for the RU, but also research management. This is important as research management is typically a bottleneck in the Czech Republic¹⁸ (and in other countries) and is often not perceived to be important by the Research Organisations' management as a condition for good performance in research. The quality of the research strategy formulation for a RU is the second sub-criterion. Research strategy sets the basis for future development and its quality is of critical value for the panels. This is the case for all research institutions, including research universities and research hospitals where the formulation of a research strategy is additionally an important indicator for the balance between research and teaching
- *Membership of the national and global research community* is very important for Scientific Research Organisations because science is a globalised undertaking and reputation is a 'key currency' in the science community. However, one would expect also the other RO types, such as the Industry/Business Service ROs (IBRO) and the Public Services ROs (PSRO) to be embedded in international networks in order to ensure a state-of-the-art level of their research, even though they often work primarily with and for more local communities of partners and users
- The point of the *scientific research excellence* criterion is to capture the 'peak performance' of the RUs and to encourage them to perform at the highest possible levels of scientific quality. One could think of it almost as asking 'How close is this RU to being able to win ERC grants?' Mostly university and Academy RUs will be capable of operating at this level, but it is not excluded that an IBRO or PSRO can produce excellent papers in Nature. Most important in a PRFS, it should be encouraged to strive to achieve this
- *Overall research performance* is a multi-dimensional category and refers to productivity, quality and competitiveness in research. All of these factors are essential to the good functioning of Research Organisations, be they Scientific Research Organisations or more applied R&D actors that serve user communities.
- *Societal relevance* is different and in the context of evaluation important in particular for Research Organisations that conduct applied R&D and directly serve user communities such as industry sectors (Industry and Business Services Research Organisations), the public sector (Public Services Research Organisations) or other researchers (National Resources Research Organisations such as archives and libraries). For Scientific Research Organisations, societal relevance of the research depends on the nature of the research (fundamental versus applied) and the focus of the research activities. For some scientific fields and research institutions (such as for example the technical universities) societal relevance is related to the use of the research results for industrial R&D; for others it relies more on the teaching and training of (PhD) students, the establishment of

¹⁷ See the Final report 1 - *The R&D Evaluation Methodology* (previously: First Interim Report)

¹⁸ Erik Arnold et al., *The Quality of Research, Institutional Funding and Research Evaluation in the Czech Republic and abroad*. Thematic Report No. 3, International Audit of Research, Development & Innovation in the Czech Republic, 2011.

contacts with user communities in order to be informed about user needs, and in the use of media channels to inform and educate the citizens.

As explained in Section 3.1, above, these differences will be taken into account in the funding system where different weights are allocated to these criteria for the different types of RO. We explain this further in Section 3.3, below.

They are also taken into account in the evaluation process, by means of the **calibration exercise**. During the calibration exercise, the expert panels will

- Define key concepts for the assessment of the Scientific Research Excellence and Societal Relevance, setting them within the context of their field. These concepts are
 - Originality, significance and rigour – for the assessment of the submitted selected research outputs
 - Reach and significance – for the assessment of the knowledge transfer activities to non-scientific actors and their effects
- Define the importance of the sub-criteria (see Exhibit 15, above) for the final quality level score at the assessment criterion level, taking into account the specifics of the field and the types of Research Organisations.

For the sake of fairness in the evaluation, consistency in the interpretations by the different panels is important. This will be ensured through coordination among Subject panel Chairs in a disciplinary area and the participation of the Main Panel Chairs in the Subject Panels' calibration exercises. In addition, all Main panel members will receive and review the reports on Subject panels' calibration exercises. Coordination among the Main Panel Chairs will allow for consistency at the overall level.

The process for the calibration exercise is explained further in the *Evaluation Handbook* (Background report 5). We provide examples of the outcomes of such exercises conducted in the context of the Small Pilot Evaluation in the report *The Small Pilot Evaluation: Feedback and Results* (Background report 9).

3.2.6 Overview of the indicators for the assessment criteria

A core objective of the evaluation is the improvement of research and research management. To meet this objective, the evaluation system entails a *self-assessment* and *informed peer-review*.

In contrast to the current Metodika, metrics (ie quantitative data) will have a role only to inform the expert panels. The panels' judgment will be based on a mix of information – quantitative and especially qualitative. These two types of data complement each other in providing the panels with a comprehensive view of an RU's profile and performance. They also act as a control mechanism on the quality of the information provided: quantitative data can confirm or contradict qualitative statements and vice-versa. Last but not least, qualitative information in the form of self-assessments, performed by the researchers and their management that are effectively involved in the RU, is of great importance in particular in the absence of on-site visits.

As with the assessment criteria, the indicators used are **interconnected**. Several of the quantitative indicators will provide the panels with information that will be of use against multiple assessment criteria and are therefore collected in a consistent manner throughout the evaluation exercise. Quantitative data are collected for the following 'input' indicators

- Research and support staff and the researchers' profile, at Evaluated Unit and Research Unit(s) level

Summary report

- PhD students - enrolled, awarded, 'recruited' or trained by researchers in the Research Unit (only for Research Units teaching or training PhD students)
- Total institutional funding, at Evaluated Unit and (estimates) at Research Unit(s) level
- External competitive and contract research funding, at Research Unit(s) level

In order to reduce the burden on the Research Units, these data will be complemented with data extracted directly from the national RD&I Information System (IS) related to

- The number and types of research outputs by the researchers in the Research Unit during the evaluation period. These data will constitute the basis for the bibliometric analyses
- Competitive funding attained from national public sources for projects or grants where the researchers in the Research Unit acted as Principal Investigators

Qualitative information gives the panel experts a view of the reasons for certain trends in the data series, the overall institutional context, the profiles of the (national and international) research partners and the intensity and value of the collaboration, the profiles of the users of the research outputs and the channels for the transfer of knowledge to society at large. Qualitative information is requested for the following topics:

- Background information on the Evaluated Unit and the Research Unit(s) (organisational structure, scientific focus, history)
- Human resources management, for all researchers and for PhD students and postdoctoral fellows (the latter only for universities and university hospitals), at the Evaluated Unit and the Research Unit(s) level
- Research infrastructure, available and used by the Research Unit
- Research strategy and plans for the Research Unit
- National and international collaborations and esteem measures
- Knowledge and technology transfer activities to non-academic actors in society

Results of the self-assessment should be reported regarding

- The adequacy of the research infrastructure in the Evaluated Unit for research in the field
- Key value and relevance of the RU activities for research and development
- The RU competitive positioning in the national and international context
- Societal value of the RU activities and research
- The final SWOT analysis and conclusions

The indicators in the EM cover the dimensions that are most relevant for the collection of information against each assessment criterion. They are geared towards the assessment of performance in knowledge creation and knowledge transfer and cover in particular the mechanisms for knowledge transfer - for advancement in research, innovation in industry and the public sector, or the society at large. These range from publications and conference presentations to formal and informal collaborations and intellectual property (IP) outputs. These 'mechanisms for knowledge transfer' constitute the 'pathways to impact', that is they are the conditions for impact to occur. International collaborations and the knowledge exchange that takes place in these collaborations have a significant impact on the quality of research; publications are one of the means to create advancements in industrial R&D, and thus innovation; the

intensity of the relationship with future users of the research outcomes (be they industry, public administration or citizen associations) enable improved understanding of the users’ needs and alignment of research with those needs, which impacts the future use of the research results.

For the public HEIs, the assessment covers also the quality of their creation of future research capacity in the RD&I system, ie indicators related to the career development and education of PhD students.

Like the assessment criteria, the indicators are of relevance to **all types of Research Organisations**, but to differing degrees depending on their mission. In Exhibit 16, we list the main indicators and indicate our view on the extent to which they are of high importance (indicated by XX) or medium-to-low importance (indicated by X) for the different types of RO.

We make a distinction between organisations conducting basic or applied research in response to the concerns voiced by the organisations conducting applied research (such as the technical universities and the Industry & Business Services RO).

These importance attributions are indicative only. As mentioned above, the relevance of assessment criteria and specific indicators for the different types of Research Organisations will differ depending on the field of research in which they are active. In the context of the evaluation, it will be up to the expert panels to decide on this topic during their calibration exercise.

Exhibit 16 List of the main indicators and their relevance for the types of RO

	Scientific research organisations		Other applied research organisations
	Basic research	Applied research	
Research environment			
Research management			
Institutional funding	XX	XX	XX
Research capacity	XX	XX	XX
Infrastructure for research	XX	XX	XX
Human resource development	XX	XX	XX
Development of PhD students & postdocs (HEI only)	XX	XX	
Research strategy			
Quality, plans, use of resources, alignment with national policy	XX	XX	XX
Membership of the global & national research community			
International research presence and collaboration			
External funding (international competitiveness)	XX	XX	X
Reputation and esteem	XX	XX	X
International collaborations & partnerships	XX	XX	X
International co-publications	XX	XX	X
International mobility	XX	XX	
National research presence and collaboration			
External funding (national competitiveness)	XX	XX	XX
Reputation and esteem	XX	XX	XX
National collaborations & partnerships	XX	XX	XX
National co-publications	XX	XX	XX
Scientific research excellence			

Summary report

	Scientific research organisations		Other applied research organisations
	Basic research	Applied research	
Peak quality of the scholarly outputs	XX	XX	X
Overall research performance			
Research output			
Research productivity	XX	XX	XX
Publication profile	XX	XX	XX
Value for the advancement of R&D	XX	XX	XX
Competitiveness in research			
External funding	XX	XX	XX
Reputation (attractiveness PhD students)	XX	XX	X
National & international competitive positioning	XX	XX	XX
Relevance for society			
External funding (competitive & contract research)	X	XX	XX
Reputation and esteem	X	XX	XX
National collaborations & partnerships (user communities)	X	XX	XX
(Use of) research outputs	X	XX	XX
Use of media channels	X	XX	XX
Licence income	X	XX	XX
Spin-off companies	X	XX	XX
Value of the RU research activities for society	X	XX	XX
Overall self-assessment			
SWOT analysis	XX	XX	XX

Key: XX = high relevance; X = medium-to-low relevance, empty = no relevance

3.2.7 Data on FTE researchers

The number of full-time equivalent (FTE) researchers as an indicator of the size of a research 'entity' is an important element in both the evaluation and funding system. It gives the evaluation panels a general view of the critical mass in the RU as well as its positioning in the Evaluated Unit. It is also an indicator that is used for the calculation of research productivity (one of the elements informing the assessment of the sub-criterion 'Research Output'). In the Funding System, data on FTE researchers are needed for the translation of assessment scores into funding (see Section 3.3 below).

The Small Pilot Evaluation made it clear how big the challenge that current practice in the calculation of FTE researchers poses for attaining a fair assessment in a national context. The participating Evaluated Units had significant difficulties in collecting FTE researcher data and the evaluation panels spotted numerous mistakes in the data during their onsite visits. As a consequence, especially for the Higher Education institutes, the delivery of FTE researcher data constituted a greater burden for the Evaluated Units than it was expected to be – or should be.

The main challenges for the collection of accurate FTE researcher data are the multiple concurrent employment of researchers, which is a common phenomenon, and the uncertainties associated with the distinction between teaching and research time for university staff. There is also no common use or understanding of the term 'researcher'. In the Academy institutes, PhD students (with a contract) are not considered in the calculation of FTE researcher data, while they are in universities. In

the Academy institutes and some universities, employees who would normally be categorised as technicians are not counted for the calculation of FTE researchers; in other cases they are. As a consequence, accurate researcher FTE data are difficult to collect also from other sources such as the Czech Statistical Office, and the correctness of the data is hard to verify.

For the evaluation and funding system, this lack of a common definition of FTE researchers has the implication that gaming with FTE numbers is fairly easy. In order to minimise this risk, we have built a correction mechanism into the EM and Funding System combined. The basis is that data on FTE researchers *as they are submitted by EvUs* will be used for both the evaluation (in the context of the Research Productivity indicator) as well as the formula for calculating the PRFS component of institutional R&D funding. This implies that there are two opposite forces that discourage EvUs from over- or underestimating the number of FTE researchers: on the one hand, overestimation of FTE researchers will lead to a lower achievement for overall research performance (the more FTE researchers for the same research output, the worse the score for performance); on the other hand, underestimation of FTE leads to lower institutional R&D funding. For example, if the EvU registers only its best researchers for the evaluation, this will influence the number of FTE researchers that will be counted for the normalisation of the PRFS funding and so the EvU will receive less institutional funding.

This correction mechanism is far from optimal, though, so we urge the Czech RD&I community to tackle this issue and reach an agreement for a **standard definition** of a FTE researcher and the processes for the calculation.

3.2.8 Data and indicators related to research outputs

In this section we cover the data related to research outputs that will be collected to inform the assessment of the Research Units related to the criteria ‘Scientific research excellence’ and ‘Overall research performance’. We also briefly describe the bibliometric indicators that will be used. Bibliometric data are of relevance in particular to describe the Research Units’ ‘publication profile’, but are of use also for the profiling of international and national co-publications.

Readers who need more detailed information can refer to the Main report The R&D Evaluation Methodology as well as the reports *Evaluation Handbook* (Background report 5) and *Bibliometrics on and for the Czech Republic* (Background report 3).

Data and rules for the ‘Scientific Research Excellence’ assessment criterion

The objectives of the ‘Scientific Research Excellence’ criterion described above and the focus for the assessment as expressed in the description of the quality-starred levels, (ie the typically scientific criteria of originality, significance and rigour) implies that only *scholarly* outputs¹⁹ matter for this criterion. This relates both to the calculation of the volume on which to base the number of outputs selected and the selection of the outputs to submit. The same rules for the volume calculations related to books and co-publications as the ones indicated for the minimum thresholds for participation apply (see Section 3.2.2, above).

We have established upper and lower limits for the number of publications to be submitted for the assessment of scientific research excellence: each Research Unit will

¹⁹ Scholarly research outputs are defined as: Papers in peer-reviewed journals - J, Conference proceedings - D, Monographs, books and book chapters - B

submit for review a number of research outputs that accounts for minimum 1% and maximum 2% of the total number of scholarly outputs by the researchers in the Research Unit over the evaluation period - but however no less than 3 and no more than 20.

We defined a *range* of number of outputs to submit because the key concept is that the RU selects only its *most outstanding* research outputs. The quality of the selection process is of high importance here: the panel members will take into account the average of the scores assigned by the reviewers against all submitted outputs for the definition of the RU quality level against the criterion Scientific research excellence. In other words, the higher the quality of each submitted output, the higher the chance is to reach an overall high quality level.

For the calculation of the *number* of scholarly outputs that an RU can submit, the *total* volume of the RU scholarly outputs will count, no matter in which field. However, Research Units can *submit* only scholarly outputs that are within their field of research, so that they fall within the sphere of expertise of the Subject panel members.

RUs that are granted cross-referral can submit selected research outputs related to the fields covered by the 'main' Subject panel as well as the panel(s) of cross-referral. An Interdisciplinary Research Unit can submit selected research outputs related to all the fields it is active in.

The following rules apply for the research outputs submitted.

- EvU staff members should be among the *main* authors of the selected research outputs
- The selected research outputs are to be submitted *electronically*, including books – in their *entirety* (so, only the abstract or the contents page is not sufficient)
- In principle, the selected research outputs are to be in *English*. This makes it easier to contract competent peer reviewers; it also provides encouragement to the research community to internationalise. Our preference is to restrict the use of non-English outputs to the absolute minimum, ie a limited number of parts of the Humanities. It is up to the Evaluation Management to take a decision on this matter, as it will inevitably influence both the cost and the complexity of the evaluation process

Research output data for the 'Overall research performance' criterion

Exhibit 16, above, shows that the panels will base their assessment of the overall research performance upon two sub-criteria: research output and competitiveness in research. One of the three main indicators for the assessment of research output is research productivity.

The indicator 'Research Productivity' builds on two sets of data: the sum of the number of FTE researchers and contracted PhD students, and the volume of research output production. Both of these data are provided in time series. Some clarifications are needed.

- Data on research outputs are based on the information stored in the RD&I Information System and are not given in terms of total volume: they are given for each type of output separately. This allows the panel experts to take into account the types of research outputs that are most relevant in their field
- All eligible research outputs by the researchers included in the RU are counted, no matter the field
- The sum of the data on FTE researchers and contracted PhD students are considered in order to take into account that also publications by PhD students are

registered in the RD&I IS against the EvU and are therefore included in the bibliometric data

Eligible research outputs are the same ones as the ones used for the threshold calculations (see Section 3.2.2, above), ie

- Scholarly outputs: Papers in peer-reviewed journals (J), Conference proceedings (D), Monographs, books and book chapters (B)
- Non-traditional scholarly outputs: Results used by the funding provider, ie projected into legislation or norm, projected into non-legislative or strategic documents (H), Research report containing classified information (V), Certified methodologies, art conservation methodologies, specialized map works (N)
- Patents and other IP: Patents awarded (P), Plant/ breeders rights (Zodry & Zplem)

We also propose similar rulings for the volume calculations to the ones applied for the thresholds:²⁰

- Co-publications of researchers employed in the same EvU will be de-duplicated (ie counted once)
- Co-publications of researchers active in different EvUs will be counted as one each, thus keeping the incentive for collaborations among the Research Organisations

The rationale for the choice of eligible research outputs lies in the recent experience in the Czech Republic related to the reliability of data on research outputs. We adopted the approach only to include types of research outputs for which a verification process by means of external sources ensure the reliability of the data. The inclusion of the V-type non-traditional output was needed to cover security research in the evaluation. A methodological consideration was also that while bibliometric data can provide information about the quality of scientific research outputs (ie scholarly publications), the quality of outputs in the applied research sphere is shown through its value for future development (eg for industry R&D) and the uptake by or importance for the users.

Nevertheless, the research production of Research Organisations conducting applied research is broader than the types of research outputs currently covered. Hence, the inclusion of a broader range of research outputs for the assessment of Research Productivity would be **appropriate** for an improved fairness of the EM. We do not propose it in the EM because acceptance of such broader coverage by the whole RD&I community in the Czech Republic will depend on

- The understanding that quantitative data have only a relative importance in the Evaluation Methodology and that there is no direct linking between these data and the panels' judgment on scores. Hence, quantitative data do not define funding allocations
- The improvement of the definitions for these types of research outputs, in particular software

Once these conditions are fulfilled, also the broader set of applied research outputs can *and should* be eligible for the volume calculations in relation to the research productivity criterion.

²⁰ The rule for books is excluded here because this is no longer needed: books are listed separately

Bibliometric indicators and how to overcome their limits

The bibliometric indicators are grouped in three main categories: Publishing Profile, Citation Impact, and Collaboration. While the data related to the ‘publishing profile’ and ‘citation impact’ inform the panels for the assessment of the research output, data related to ‘collaboration’ inform the assessment against the ‘Membership of the RD&I community’ criterion.

For the bibliometric indicators covering publication-type research outputs, two complementary data sources are used:

- International databases, providing information on publications indexed for Web of Science and/or for Scopus
- National: the Czech RD&I Information System (IS)

The Czech RD&I Information System is the data source also for information on research outputs that are not covered in the international datasets, such as the ‘non-traditional’ scholarly research outputs and the IPR-related outputs.

Bibliometric indicators are calculated and presented per OECD field. Indicators can be aggregated (weighted for average citation rates) up to major area, research organisations, and country.

The bibliometric data will take account of the *full* research output production by researchers the RU, no matter in what field. In other words, research outputs in other scientific fields will be included in the analysis, showing for example the number of articles published by the RU per field. The panels are also given an overview of the characteristics of the field at country level, thereby providing a context for interpreting the indicators for a specific Research Unit in a specific field. This information will be of particular value for the (main) panel members drafting the overview and analytical reports at the national field and/or disciplinary area level. The same indicators are then used at the level of Research Units.

The process for the definition of these indicators and their specific focus is described in detail in the report *Bibliometrics for and on the CR* (Background report 3).

The social sciences and humanities (SSH) need special consideration in bibliometric analysis because the publication patterns of these areas – with relatively more use of publications in books and publications in other languages than English and in journals with a national rather than international coverage – are not as comprehensively covered in the international bibliographic and bibliometric data sources. Evaluation experts in the SSH panels need a broader view on the publication profile of the RUs than the one that can be provided by the international databases. For this purpose, the bibliometric data provided to the panels will include the full list of publications by the RUs, ranked per journal and/or author and complete with links to the abstracts stored in the RD&I Information System.

In this context, the RD&I Inform System (IS) stands out as a necessary supplementary data source in SSH, just as other current research information systems do in other countries. It allows for complementing the bibliometric data from the international databases with additional data on research outputs that are published locally or in the Czech language. The same is obviously valid for publications in other disciplines.

In the Main report *The Small Pilot Evaluation and the Use of the RD&I IS for evaluation* (previously: Third Interim Report) we suggest enhancing the value of the RD&I IS for evaluation through the development of a **standardised and dynamic register** of scholarly journals, series and book publishers, similar to the ones developed in countries like Belgium (the Flanders) and Norway. This would be especially useful for SSH, combining for example the current List of peer-reviewed non-impact journals published in the Czech Republic with ERIH PLUS. However, also

high-quality international journals prominent in other fields and included in specialised field-specific datasets or registers could be included.

This register would provide a basis for a more **self-determined evaluation** of the research outputs in the Czech Republic than the present categories provide - in particular in the field of SSH, but also for the benefit of other fields. Possibilities are

- To distinguish among journals that target scientific versus non-academic readers, ie journals aimed at *communication* for research versus *dissemination* and awareness raising on research
- To distinguish between quality levels in the register of journals, series, and the (book) publishers, similar to the ERIH levels of authorship, ie making a distinction between international, national and local 'authorship'

We refer to the Main report *The Small Pilot Evaluation and the Use of the RD&I IS for evaluation*' for further details on this topic.

3.2.9 The potential role of the RD&I Information System (IS)

In recent years, a clear trend is visible in Europe and world-wide towards the development of national research information systems, creating the opportunity for reaching a comprehensive and longer-term view on the performance of the research system in the country, its productivity and impacts. Based upon current information, in the EU member states only Germany, Austria and 2 of the 3 Baltic States have not considered (yet) developing a fully national research information system.²¹

Current international experience shows the value of these systems, in particular in providing access to information that may support strategy development, for the benefit of both research institutions and policy-makers – at a broader and more detailed level than was previously the case, and especially of higher quality. At the same time, they create efficiency gains in the reporting and collection of data. An additional benefit for the national funding agencies is the collection of information for the allocation of institutional core funding (PRF). Further benefits for research communities are services that allow re-use of the data, eg in grant applications, direct access to repositories, once-only input of data thanks to linkages with eCV databases, etc.

The Czech Republic is in an advantaged position from this perspective, having at its disposal an existing comprehensive and interconnected national research information system. However, improvements and extensions to the current system are needed²², in a first instance to keep pace with recent technological developments and remain state-of-the-art in the international context.

A key characteristic and value of the RD&I Information System (IS) is its capacity to interconnect data related to research organisations, individual researchers, research inputs and outputs, research programmes and their objectives, and national funding agencies. The potential value and role of the RD&I Information System for the purpose of evaluation is related to this interconnectedness.

²¹ Mahieu, B., Arnold, E., Kolarz, P. (2014) Measuring scientific performance for improved policy making - The feasibility and desirability of a European research information system. Study for the European Parliament - Science and Technology Options Assessment (ISBN 978-92-823-5531-2)

²² In the report *The RD&I Information System as information tool for evaluation* (Background report 9) we cover the potential extensions and improvements more in detail.

From an operational perspective, and similar to the experience in other countries, the RD&I IS has an important role to play in **informing** the collection of data for evaluation, leading to significant time and cost savings for all actors involved. The experience in the Small Pilot Evaluation proved its usefulness in this context.²³ It has the capacity to support the evaluated RU with data that is stored in the national information system, contributing to a significant reduction of the evaluation burden for the research organisations. It is an important source of bibliometric data, going beyond the research outputs that are registered in international databases, and is therefore a critical component for the analysis of research output to inform the evaluation panels. The RD&I IS also is an important tool for the implementation of eligibility and data quality checks during the evaluation, as well as for the validation of data submitted by the RU.

Most important, however, the RD&I Information System needs to be better exploited as a **source for strategic information**, supporting the use of the data it holds for other than purely administrative or accountability purposes – for the benefit of the research organisations, the funding bodies and national policy makers. Extensions improving the user-friendliness of the RD&I IS and its surveying capacities are key from this perspective.

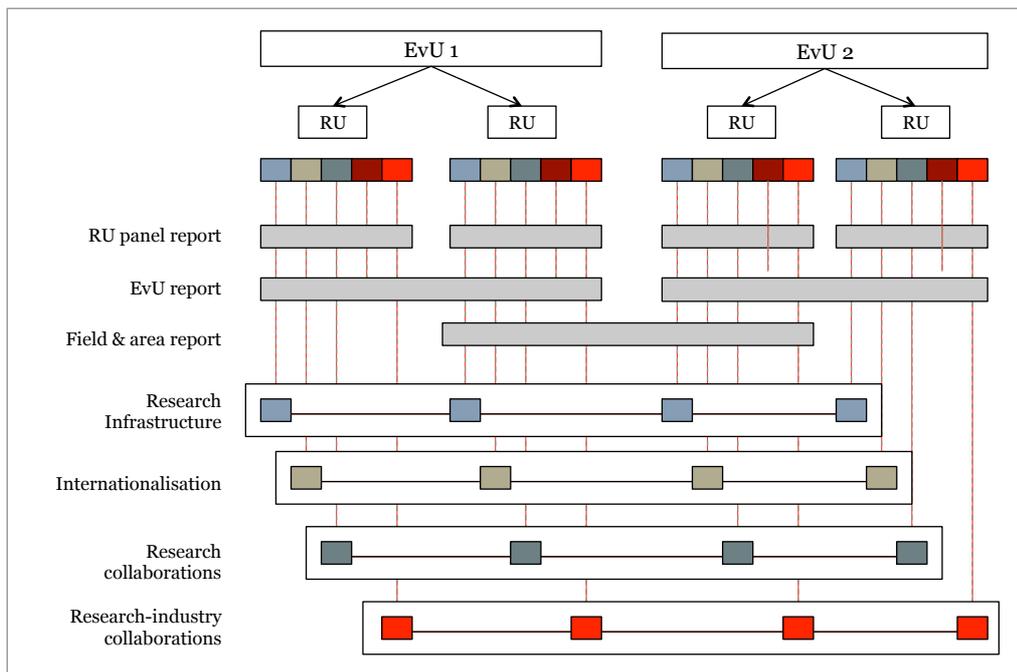
In this context, the use of the RD&I IS as the **platform for the implementation** of the National Evaluation of Research Organisations is of key importance. The collection of data provided by the evaluated Research Units and the storage of the evaluation results directly in the RD&I IS would not only allow for a longer-term analysis of these results and the developments, but also for the direct interconnection of the evaluation results with other data stored in the system, such as the competitive research programmes.

Exhibit 17, below, illustrates the concept of such use of evaluation results for strategic information purposes. It shows how storing the RUs' input of (qualitative and quantitative) data in the RD&I Information System would enable the exploitation of this information also at a 'thematic' level across all Research Units, thus providing a view at the national level. Information of potential interest includes information on internationalisation, collaborations in Centres or clusters, involvement in incubators, the use of research infrastructure etc.

Interconnection of the results from the NERO with the results from other evaluations, for example the evaluation of the Research Infrastructures, can improve policy makers' overview on the needs for and values of their policy interventions.

²³ See the Final report 3 - *The Small Pilot Evaluation and the use of the RD&I IS for evaluation* (previously: Third Interim Report) and the report *The Small Pilot Evaluation: feedback and results* (Background report 10)

Exhibit 17 Potential for extended use of the evaluation results



3.2.10 Implementing the National Evaluation of Research Organisations (NERO)

The implementation of the Evaluation Methodology requires a specific set of activities that go beyond the evaluation itself, such as, for example, the management of conflicts of interest. We cover these in detail in the Main report *The R&D Evaluation Methodology* and in the *Evaluation Handbook* (Background report 5).

Two topics merit further reflections in this summary report: the evaluation management structure and the consequences of the ‘multifunctional’ focus of the evaluation methodology.

The **management structure** for the implementation of the NERO, both in terms of the methodology definition and its implementation, is still an open question. In the current set-up of the Metodika, the RD&I Council has responsibility for the definition of the evaluation methodology, while the implementation is entrusted to the Office of the Government. However, there are various approaches in international practice, as outlined in the International Audit of RD&I in the Czech Republic²⁴ that are worth considering. Concretely, the management of the NERO can be entrusted to an independent and autonomous agency, an agency established by a ministry, or a unit within a ministry or the Government Office. There are pros and cons for all approaches, but in any case it requires a policy decision. The IPN project team, which is responsible for this study, has among its tasks to analyse the feasibility of establishing such a structure and its possible variants.

Approaching the topic in a ‘neutral’ manner, in the Main report *The R&D Evaluation Methodology* we suggest an evaluation management structure comprising an

²⁴ See, for example, Arnold, E., Mahieu, B., Horvath, A., (2011) *R&D Governance in the Czech Republic*, International Audit of RD&I in the Czech Republic, Final Report -2

Evaluation Management Board, acting as the overall supervising body, and an Evaluation Management ‘Team’. We consider that in order to reach the maximum level of legitimisation, the Evaluation Management Board should be composed of representatives of the R&D governing bodies (funding ministries and ministries with responsibilities for research institutes, the Academy and the agencies), chaired by a representative of the Government Office.

We also advise against a pronounced executive role of the RD&I Council, both in the development and implementation of the evaluation, and see its role more in terms of providing scientific advice and support. This consideration reflects the concept of evaluation following a ‘*waterfall principle*’ where actors at each level of the RD&I system evaluate actions at the level below them²⁵. The RD&I Council, being the representative body of the Research Organisations, can therefore advise on the evaluation methodology and any evaluation should be based on an as large as possible consensus in the RD&I community, but ultimate decision-making should be the task of the actors at the higher level in the RD&I system, ie the national policy makers.

The Evaluation Methodology that we propose is ‘**multifunctional**’, ie it is designed to provide input both at the national level and institutional level in terms of strategic information and indications for institutional funding allocations. It uses a single framework for assessment across all disciplines and research organisation typologies, with a common set of data required in all submissions, standard definitions and procedures, and assessment by expert panels against broad generic criteria. In Section 2.1.2, above, we mentioned that such ‘multifunctional’ evaluation and funding system is exceptional in the international context. It implies an all-round evaluation, simultaneously for all Research Organisations in the system. It also implies that there is a limit to the evaluation system to meet the differing needs for evaluation of the individual types of Research Organisation.

In the Main report *The R&D Evaluation Methodology* we recognise the value of *complementing* the Evaluation Methodology with ‘mission-specific’ evaluations or methods that would assess in more depth specific features of the different types of Research Organisations. In this context, the Evaluation Methodology constitutes the necessary *common spine* for these ‘mission-specific’ evaluations.

We also considered the possibility - and the effects - of spreading the evaluation effort and distributing the administrative burden over several years. Possibilities are to run the evaluation some fields at a time, or to have different evaluation runs for the different types of ROs.

Both of these options present the advantage of a more continuous workload for the evaluation office / agency and a less disruptive evaluation procedure for all parties involved. They would also create a better foothold of the evaluation in the RD&I system, and especially easier appropriation of the evaluation results by both the Research Organisations and the funding providers. However, both options present also significant disadvantages, creating additional complexity in the evaluation and especially the funding system. Information for the funding allocations would be collected over time; the overall cost of the evaluation would increase; the fairness of the evaluation would be set at risk by having to use different panels over time inhibiting the overarching coordinating role of the Main panels; and last but not least, a core value of the Evaluation Methodology would go lost, ie the capacity to give a comprehensive national view at a specific point in time informing the national RD&I policy. As a result, we are of the opinion that a full-scale evaluation covering

²⁵ See Arnold, E., Mahieu, B., Horvath, A.,(2011) *R&D Governance in the Czech Republic*, International Audit of RD&I in the Czech Republic, Final Report -2

simultaneously all Research Organisations and fields remains *the most efficient and effective approach*.

3.2.11 Cost estimate of the evaluation

The Terms of Reference for the study required us “to design the evaluation in a way that the total cost of its implementation (direct and indirect costs, costs incurred by research organisations (ROs), evaluation organizers and providers) does not exceed 1.0 per cent of public institutional funding for R&D for a five-year period (ie estimated according to SB and 2012 prices, ca CZK 677 million).”

We therefore designed the operational set-up for the evaluation bearing in mind the need for cost and time efficiency, for all actors involved.

We estimated the costs of a full-scale evaluation using an activity-based cost approach. We took into consideration all costs, that is both

- The ‘direct’ ones for the organisation of the evaluation, the handling of the process, the evaluation panels, and the costs related to relevant updates to the RD&I Information System, and
- The indirect ones, ie the costs related to the time that the organisations evaluated will need to invest in their self-assessments

Our cost estimate is based on our own, experience, the outcomes of the Small Pilot Evaluation, and international experience. Information on the full costs of national evaluations is hard to find: only a few countries look into the indirect costs of these exercises. Through literature review and interviews, we collected sufficiently useful data on the UK RAE 2009 and the Italian VQR (Exhibit 18).

While the evaluations in Italy and New Zealand are based on the UK system, which acts as the model for PRFS worldwide, the three evaluation systems show some significant differences that influence their overall costs²⁶:

- The UK RAE/REF is a mature system that has gone through many steps of improvement and enhancement (ie after each evaluation round) in response to the comments from the research community and policy needs. It has become more complex every time– and more costly, some say outrageously.²⁷ It is an expert panel-based system and covers universities only, who select the best of their researchers for the competition. It runs every 6 years covering a five-year period
- Italy has run three national peer review-based evaluations so far, the VTR in 2006 and the VQR 2004-2010 in 2013. Both are inspired by the UK system, but the VQR has a strong element of bibliometrics (for cost-saving reasons) while bibliometrics are only marginal in the UK system (even in the latest REF) due to opposition in the research community. All universities and some interuniversity research institutes are covered and all researchers participate, submitting a **maximum** number of publications (4 per researcher). The assessment is done predominantly through bibliometrics for the ‘hard’ sciences and peer review for the other. It is meant to run every 3 years

²⁶ We compare the features of PRFS internationally in the Final report 1 - *The R&D Evaluation Methodology* and cover them in detail in the report *Evaluation systems in international practice* (Background report 1)

²⁷ Martin, B., The Research Excellence Framework and the ‘impact agenda’: are we creating a Frankenstein monster?, *Research Evaluation*, 20(3), September 2011, pages 247–254

Summary report

- We included New Zealand in Exhibit 18, below, as it is the only evaluation system that we know off where the panel-based evaluation also includes onsite visits. It covers all universities and all researchers. As can be seen, this leads to huge costs. The evaluation runs every 6 years.

Exhibit 18 Costs of PRFS in other countries

	UK RAE 2008	IT VQR	NZ QE 2006
Direct costs (staff & panels) (k€)	€ 15,120	€ 10,570	
Indirect costs (k€)	€ 74,340	€ 54,029	
Total direct & indirect costs (k€)	€ 89,460	€ 64,600	€ 46,960
Indirect costs % of total costs	83%	84%	
<i>Nr FTE researchers</i>	68,563	61,822	8,671
Total costs per researcher	€ 1,305	€ 1,045	€ 5,416

Source: Technopolis calculation on multiple sources²⁸

As mentioned above, the target maximum budget for the EM was set at “1.0 percent of public institutional funding for R&D for a five-year period, estimated at CZK 677 million.” This estimate refers to the total of the national R&D institutional funding expenditure budget, of which Institutional funding for Research Organisations is one of the budget lines.

Based respectively on the total institutional funding budget for R&D and the ‘specific’ one, Exhibit 19, below, shows that the maximum cost per researcher would be either €1,336 or €876,.

In other words, the 1% limit calculated against the total institutional funding for R&D budget would imply that the evaluation system is more expensive than the UK and the Italian ones, while the one based on the ‘specific’ budget line indicates more cost efficiency than the UK RAE but only slightly so compared to the IT VQR.

Exhibit 19 Maximum total costs of the new evaluation system

	R&D Institutional funding expenditure – total	Institutional funding for conceptual development of ROs*
2012 budget (m CZK)	13,536 Kč	8,874 Kč
Total for 5 years based on 2012 budget (m CZK)	67,680 Kč	44,370 Kč
1% limit (m CZK)	676.8 Kč	443.7 Kč
1 % limit (k €)	€ 24,611	€ 16,135
<i>Nr FTE researchers</i>	18,422	18,422
Maximum total costs per researcher	€ 1,336	€ 876

incl. research intentions

²⁸ Main sources of information were: On the RAE: RAE 2008 Manager Report (2009); RAE (2008) Accountability Review, PA Consulting, 2009; On the VQR: Geuna, A., Piolatto, M., *The development of research assessment in the UK and Italy: costly and difficult, but probably worth (for a while)*, Department of Economics and Statistics “Cognetti de Martiis”, Università degli Studi di Torino, Italy: 2014

A key principle for our design of the Evaluation Methodology was that the cost and burden of the evaluation should be the minimum possible to deliver a robust and defensible evaluation process. We also took into consideration that as in any other country, the evaluation system is bound to become more complex and sophisticated in response to the needs and comments of the research community. This is especially the case when evaluation results are linked to institutional funding, as the experiences in, the UK with the RAE and the Czech Republic with the Metodika show.

Our approach was therefore to design an evaluation system that would set the **basis** for these future enhancements, keeping it as simple as possible and ensuring methodological robustness while responding adequately to the needs of the Czech RD&I community and policy makers.

We have estimated the total costs of a full-scale evaluation at **€12,929k** (Exhibit 20), ie approximately €3,000k under the targeted limit.

This implies a total cost per researcher of €702; the indirect costs that will need to be carried by the RD&I community represent 48% of the total. For both parameters, the EM therefore constitutes a cost efficient evaluation system by international standards. It foresees an evaluation running every six years, covering a five-year period.

Exhibit 20 Cost estimate for a full-scale evaluation

	Total – in k€
Total direct costs	€ 6,672
Evaluation panels	€ 3,905
Evaluation management	€ 2,767
Total indirect costs	€ 6,257
Estimated overall costs for a full-scale evaluation	€ 12,929
Indirect costs % of total costs	48%
Nr FTE researchers	18,422
Total costs per researcher	€ 702

The costs for the evaluation panels are based on using 26 Subject panels and a maximum of 20 days involvement for each expert, of which 5 days in the Czech Republic. The Evaluation management costs include 5 year salaries of the core staff in the evaluation management team given their on-going involvement in the evaluation cycle; 3 years salaries for the evaluation management board members (approximately one FTE/year in total); additional staff for 1 year (24 FTE/year); costs for external bibliometric analysis support and the costs for the RD&I Information System use and updates (cumulative essential and recommended updates). The indirect costs envisage the involvement of 10 researchers and 1 administrative staff per Research Unit.

Readers wishing to be informed more in detail can refer to the report *Detailed cost evaluation framework* (Background report 4).

3.3 The Institutional Funding System

In this chapter we describe the system and principles that we propose for the future institutional funding of Research Organisations in the Czech Republic. We refer to the Main report *The Institutional Funding Principles* (previously: the Second Interim Report) for a full and detailed description. Our proposal is based on the key concepts that are described in Section 3.3, above, and are to be set in the overall conceptual framework for the Evaluation Methodology and Funding Principles, described in Section 3.1.

This chapter is structured as follows: we first give an overview of the Institutional Funding System and its core components and principles (Section 3.3.1). We then go in more detail and describe the process and principles for the division of the national institutional funding budget into mission-based ‘pots’ in Section 3.3.2 and the subsequent sub-division over the three funding components, ie the block grant (Section 3.3.3, the performance agreement (Section 3.3.4) and the PRFS component (Section 3.3.5).

Finally, we describe how the Funding System will take account of the effects that the Evaluation Methodology can have on the future institutional funding of some Research Organisations (Section 3.3.6).

3.3.1 Overview of the proposed Institutional Funding System

Exhibit 21, below, shows the core elements of the Funding System that we propose.

A pre-requisite for the proposed funding model is that there are separate **mission-based ‘pots’** (budgets or budget lines) for the different types of Research Organisations (RO) in the Czech RD&I system (see Section 2.2, above). As a starting point the size of these mission-based pots will be based on the funding bodies’ current expenditure for institutional funding per type of Research Organisation.

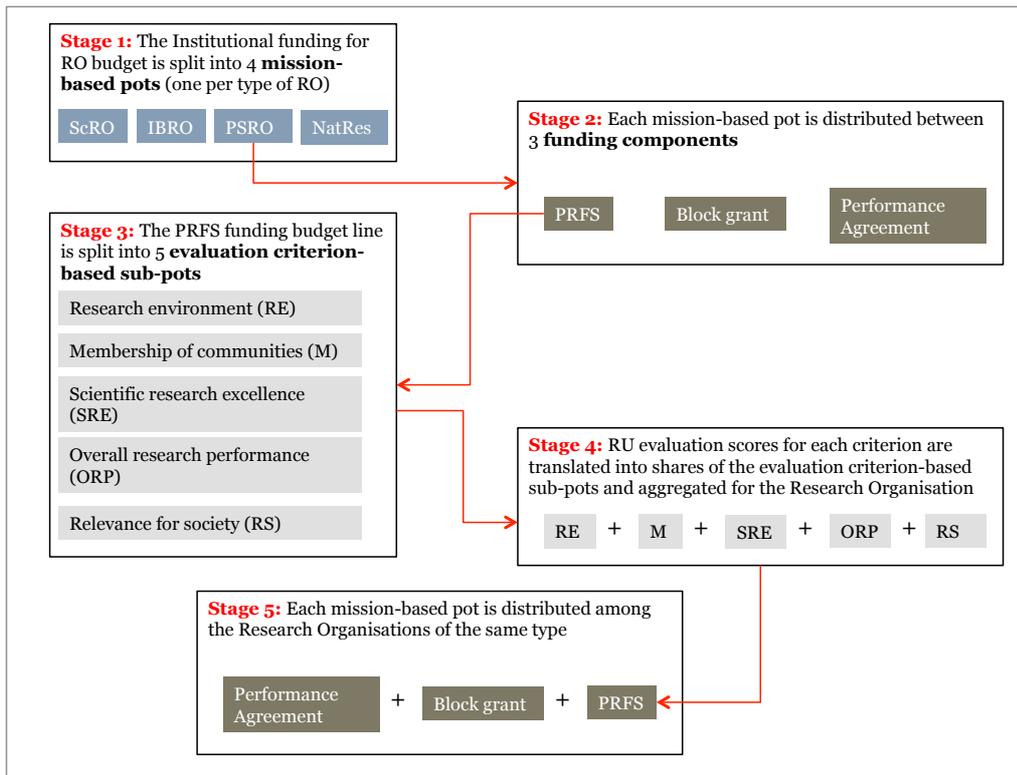
We propose a funding system that encompasses three **funding components**.

- A block grant (Block), ie a fixed sum or proportion of the public funding budget allocated to a Research Organisation (RO)
- A (negotiated) performance agreement (PA), mainly based on the negotiation between the government (a ministry) and the respective Research Organisation
- A performance-based research funding system (PRFS), which allocates funding on the basis of the evaluation results, and specifically the scores reached by the RU against each assessment criterion

While the block grant ensures trust and continuity, the performance-based funding components, which encompass the performance agreement and the PRFS, reflect past performance as well as future developments and plans.

Block grants and funding for performance agreements are allocated to Research Organisations as legal entities and are determined on the basis of the budget allocated to Research Organisations in the preceding funding period. The thinking for PRFS money is different. The money is determined on the basis of scores achieved by Research Units (RUs) and then aggregated to the level of the Research Organisation as the legal entity that receives the PRFS money. In order to translate the scores into funding, our proposal is to use the five assessment criteria and define **evaluation criterion-based “sub-pots”**.

Exhibit 21 The core elements of the Institutional Funding System



In Section 3.1.2, we explained that the overall architecture for the new evaluation and funding system is geared towards avoiding scientific fields and different Research Organisation types being set in competition to each other. For this purpose we adopted a two-step approach: while the evaluation system is centred on handling the scientific field specifics, the institutional funding system centres on recognising the different missions of the RO types in society.

The definition of **mission-based funding pots** at the very start of the budget allocation process allows for a transition from one single budget for institutional funding for RO of all kinds to *different budgets for different types of Research Organisations*. This is in line with international practice where policy makers decide on different budget pots for different types of research organisations. The distribution is generally contingent on: laws (responsibilities), history (eg size of budgets), politics, policies, (entrepreneurial) individuals (or the opposite).

The definition of mission-based funding pots also enables a differentiation of the funding allocation criteria for the Research Organisation types later in the process.

In a first instance, this relates to the **shares of the funding components** in the ‘mission-based’ pots (block grant, PRFS, performance agreement). These can be varied for different types of Research Organisations, reflecting their differences. We have not done so in our proposal. Instead, we propose for all types of Research Organisations a share of 80% for the block grant, 15% for the PRFS component, and 5% for the performance agreement.

In doing so, we considered that at the moment all types of Research Organisations are in need of stability. Together, the performance-based components will make up 20% of total institutional R&D funding. Given the high percentage of competitive funding (approximately 50%) in the Czech R&D system (see Section 2.1.2, above), we think that the competitive element in the institutional R&D funding system should not be

too pronounced, in order to guarantee stability and allow capacity building. Obviously, this situation may change in the future, reflecting the importance attributed to competition versus stability in funding for the different types of ROs.

The specific characteristics and missions of the Research Organisation types can and should be considered also when defining the **shares of the evaluation criterion-based sub-pots**. The different weights given to the assessment criteria should also be in line with the PRFS objectives and the ministries' strategies.

In other words, policy decisions are needed at various stages in this process. In particular this relates to

- The weights given to each of the assessment criteria for PRFS funding, determining the size of the evaluation criterion-based 'sub-pots'
- In future, the shares of the three components in each RO-type funding pot
- In future, the definition of pots for different types of Research Organisations and their size

3.3.2 The mission-based pots

The establishment of separate mission-based pots for the different types of Research Organisation is core to the proposed funding model. The thinking behind this prerequisite is that different types of Research Organisation fulfil different missions and functions in society. Hence, they should not be made to compete for the same budget pot as this may lead to some Research Organisations not being able to fulfil their roles any more. This is in line with common international practice and constitutes a break with the current funding system where all Research Organisations compete for institutional funding for RO irrespective of their type, a practice criticised in the Czech Audit.

We suggest taking the current expenditure on institutional funding for RO as a starting point (mean values or weighted mean values over several years). We have opted for this starting point for the sake of transparency and stability while being aware that the Metodika, through its own specific problems, has led to an allocation of funding among research providers and Research Organisations that some may consider deformed or unfair.

Our suggestion to define one funding pot for institutional research funding for each type of RO implies that there is a common funding pot for the ASCR institutes and the universities – which has been contested in the public debate. We cover this in the Main report *The Institutional Funding Principles*. Ultimately, the decision whether or not to have one pot for all Scientific Research Organisations or to have two pots, one for the ASCR and one for all other ScRO, is a policy decision that cannot be taken within the funding system itself. In the one-pot-solution, universities and the ASCR institutes would compete directly with each other, while in the two-pot-solution they would only compete among themselves.

3.3.3 The block grants

We suggest that, as in the current system, in the first funding period after the first evaluation based on the Evaluation Methodology, the block grant should make up 80% of the pot (or pots for the different Research Organisations) and 80% of what each Research Organisation received in the previous funding period.

The previous funding period should refer to an average of what Research Organisations received over a period of 3-5 years (a 'reference funding period'), in order to take into account trends. Moreover, these 3-5 years should cover the years when institutional R&D funding was based on the evaluation methodology valid for the years 2013-2015 (or later), encompassing an 80% block grant and a 20% performance based part. Using the years before the Evaluation Methodology 2013-

2015, when the Metodika allocated a higher share of institutional R&D funding, would not make sense because fluctuations were too large.

3.3.4 Performance agreements

The objective of these 'light touch' performance agreements is to promote institutional development and capacity building.

The performance agreement is an agreement between the ministry in charge (the principal) and the Research Organisation (the agent) on a small number of strategically relevant undertakings that the Research Organisation agrees to implement in the performance agreement period in order to increase or improve R&D capacity and to improve working conditions for research as well as for support staff. The main idea behind this approach is that a leverage effect is expected: if the strategic projects succeed, they will have a positive influence on the whole or at least essential parts of the Research Organisation.

Those ministries that lost their R&D budget as a consequence of the 2008 Reform (see Section 2.4.2, above) need also be involved in the negotiation of the performance agreement to ensure its relevance. Therefore, both the funding and the founding ministry will represent the principal and both will sign the performance agreement on the ministry side. In the long term, the possibility to transfer R&D funding responsibility back to the founding ministries should be envisaged (see Section 4.1, below)

The performance agreements are set up individually, taking into account the specifics of each Research Organisation (role and mission, research profiles, institutional setup). The performance agreements should be negotiated between the principal (ministry/ies) and Research Organisation for a period equal to half of the PRFS periodicity. For example, if the evaluation takes place every six years, the performance agreements should be re-negotiated after 3 years.

The Research Organisation reports on progress made in the strategic undertakings agreed upon in the performance agreement on a yearly basis. After three years, ie mid-term if an evaluation takes place every six years, the implementation of the performance agreement is evaluated, and based on the results of the evaluation a new performance agreement is negotiated. Substantial strategic projects may take more than three years; therefore it will also be possible to conclude agreements for 6 years with a half-term assessment of interim results after three years.

Both the Research Organisation (in preparing) and the ministry in charge (in assessing) shall be guided by the following criteria.

- Does the Research Organisation have a convincing overall set of goals and related strategies? Are these goals / strategies based on sound analysis of the current situation?
- Are the proposed strategic projects most relevant to support the stated goals / strategies (fit-for-purpose)? Are they feasible in terms of resources and timing?
- Is the profile of the key staff credible to successfully implement the proposed projects?

In the Main report *The Institutional Funding Principles* we provide more detailed guidelines and a basic template for the performance agreements.

3.3.5 The performance-based research funding system - PRFS

The PRFS funding system we are proposing is two-dimensional, with one dimension being the different evaluation criteria and the other being the types of Research Organisations. The basic idea is to allocate the PRFS pot among RUs based on the scores they obtained in the evaluation. ROs of the same type will therefore compete for the budget allocated in 5 evaluation criterion-based sub-pots.

As mentioned in Section 3.3.1, above, the size of these sub-pots involves a policy decision. We provide guiding principles on which to base this decision in the Main report *The Institutional Funding Principles* and summarise them in Section 3.1.1.2. The main guiding principle will be that the weights of the different evaluation criteria need to be in line with Research Organisations' missions and the strategies of the ministries concerned. These strategies will typically go beyond the mission, emphasising certain policy objectives.

We have suggested three scenarios in the Main report *The Institutional Funding Principles*: the default scenario aiming to increase overall research performance in Research Organisations; a radical scenario pushing scientific research excellence in Scientific Research Organisations and societal relevance in applied Research Organisations; and a medium scenario located in the middle between these scenarios aiming to increase overall research performance while more strongly emphasising scientific research excellence for Scientific Research Organisations and societal relevance for applied Research Organisations.

The ex-ante analysis showed that determining the weights for the various evaluation criteria in the funding mechanism will have important implications for the distribution of the institutional funds to ROs. In particular the so-called radical scenario, putting a dominant emphasis on one criterion, may cause significant differentiation in institutional funding.²⁹

As a fall-back option, in case no policy decision can be reached, all evaluation criteria can be given the same weight.

The weights should be used solely for the purpose of funding. They will be decided before the evaluation exercise but published only after submission of the self-assessment report by the Research Organisations in order to inhibit gaming. This is because if Research Organisations know that a certain category has a particularly strong weight for them, they will do (almost) anything to show themselves in the best light in that category.

There are two options for translating scores into funding:

- Option 1: Allocation of PRFS based on size (person-scores). The funds are allocated based on the number of person scores of each evaluated RU in each evaluation category. This is based on aggregation of person scores and the calculation of the average person score value for each RO type and evaluation criteria at the provider level. With regard to normalisation of RU size, we would use a measure of the number of researchers that work at the RU, ie the FTE researcher data.
- Option 2: Allocation of PRFS based on relative scores. This is based on the calculation of the average scores in each category of RO type per evaluation criteria weighted by the number of researchers (presumably FTE). The relative distance of actual scores of an RU (EvU) from the average in each category is the base for distributing the PRFS funds.

²⁹ See the report *Ex-ante assessment of the funding principles* (Background report 8)

These two options are described in detail in the Main report *The Institutional Funding Principles*, including the algorithms that can be used. Both options have pros and cons. We summarise these in Exhibit 22, below.

Exhibit 22 Pros and cons of the options for the translation of scores into funding

	Pros	Cons
Option 1 – person-scores	Has the capacity to address some weaknesses of the previous system	Exact and correct figures on scientific labour are needed (FTE researchers)
	Might allow for higher flexibility	Will require substantial fine-tuning - model simulations are necessary that deploy good data
	Growth of Research Organisations during a funding period (size of staff) is considered	
Option 2 – relative scores	Can well appreciate good performance and penalise poor performance	Might start from a bad base
	Easy to implement	Is likely to engender less change than the first option
	FTE are used only for the purpose of weighting scores achieved by individual RUs, ie calculating average scores for Research Organisations. So the space for gaming with FTEs is negligible	Requires an additional decision by policy-makers on the limits of the bands (eg between 50% and 75% from the national average) and how large the changes relative to the previous funding period in these bands are
		Growth of Research Organisations during a funding period (size of staff) is not considered

In both options, we present a linear allocation of funding. Of course, a non-linear allocation of funding can be used too, in order to boost excellence and concentrate funding on the top performers. In the case of non-linear money allocation, as in the British Research Excellence Framework (REF), the top band gets a larger share of money than the ones below, which leads to a concentration of funding on the best performing Research Organisations.

For the first funding period, we would advise using a linear allocation in order to allow for stability, continuity and in particular organisational capacity building. In a second funding period, a non-linear allocation for funding may be considered for the Scientific Research Organisations.

3.3.6 Handling the effects of the evaluation

Damping factors on changes due to the panel evaluation results

Changes in the scores of the PRFS will lead to changes in the PRFS funding. Large increases and large decreases are both difficult to accommodate, so Research Organisations may need to be given time to cope with the lower or higher level of funding.

Hence, we are proposing the guiding principle that if changes in PRFS funding exceed 5% of a Research Organisation’s total funding after the new evaluation round, the new level of PRFS funding should be implemented gradually over the years of the funding period. If figures on Research Organisations’ total income are not available, a gradual implementation of the new funding level may be introduced in the case that total institutional funding for RO is smaller than 92% or higher than 108% compared to the previous funding period.

The issue of RUs with low scores

RUs that receive low scores, in particular the quality level 1, in one or several evaluation criteria ought to get special attention, as the low scores are indicative of a problem. The immediate consequence is, of course, that the RU with low score will contribute little or nothing to the PRFS share of funding allocated to the Research Organisation it belongs to. This is already an incentive for management to get active.

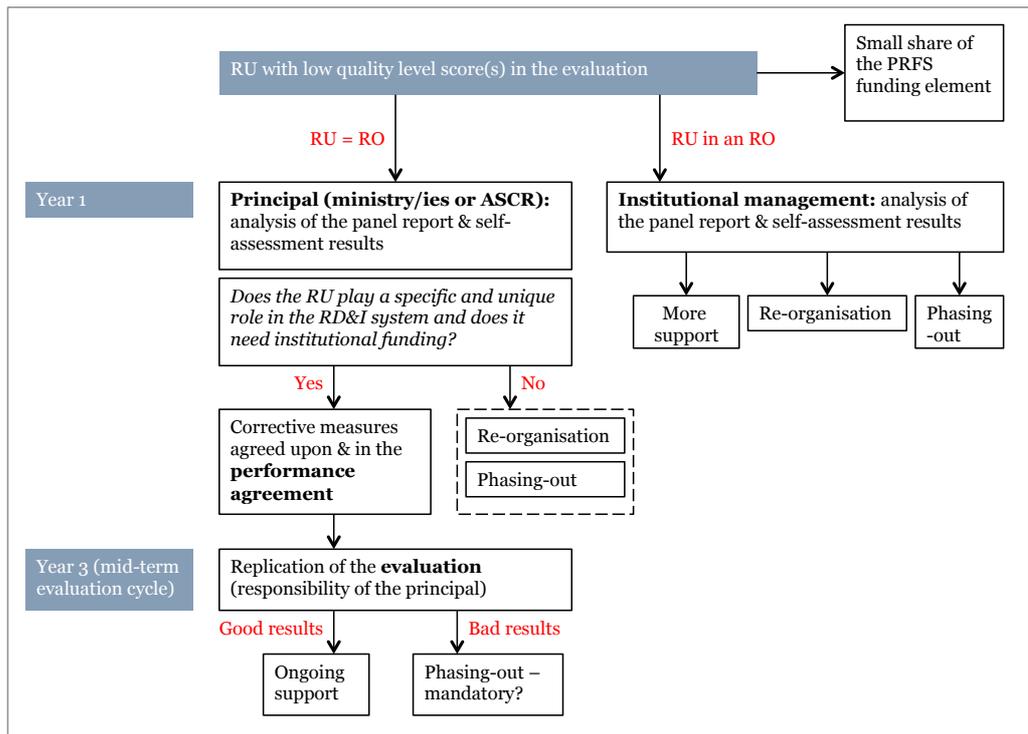
We propose a process as shown in Exhibit 23, below.

If RUs belong to a university then we would recommend the university management to take a closer look at the RU.

If an RU is equivalent to a Research Organisation, for example an RTO, and receives low scores in the evaluation, then we would invite the principal – the responsible ministry (and founding ministry if that is a different ministry) – to take a closer look at the Research Organisation. In both cases, the evaluation report would be the starting point. In the Main report *The Institutional Funding Principles* we give suggestions on what the actions by the university management or ministries could entail.

We also drafted a more severe scenario for phasing out institutional funding to weak Research Organisations. While we prefer the default scenario described above because it does not mandatorily prescribe phasing out of institutional R&D funding, the stricter scenario may also be applied given the problems in the Czech RD&I system (such as the fragmentation of the R&D base and the institutional funding, deformation of institutional funding allocations due to the problems of the Metodika).

Exhibit 23 Process for the management of Research Units with a low evaluation score



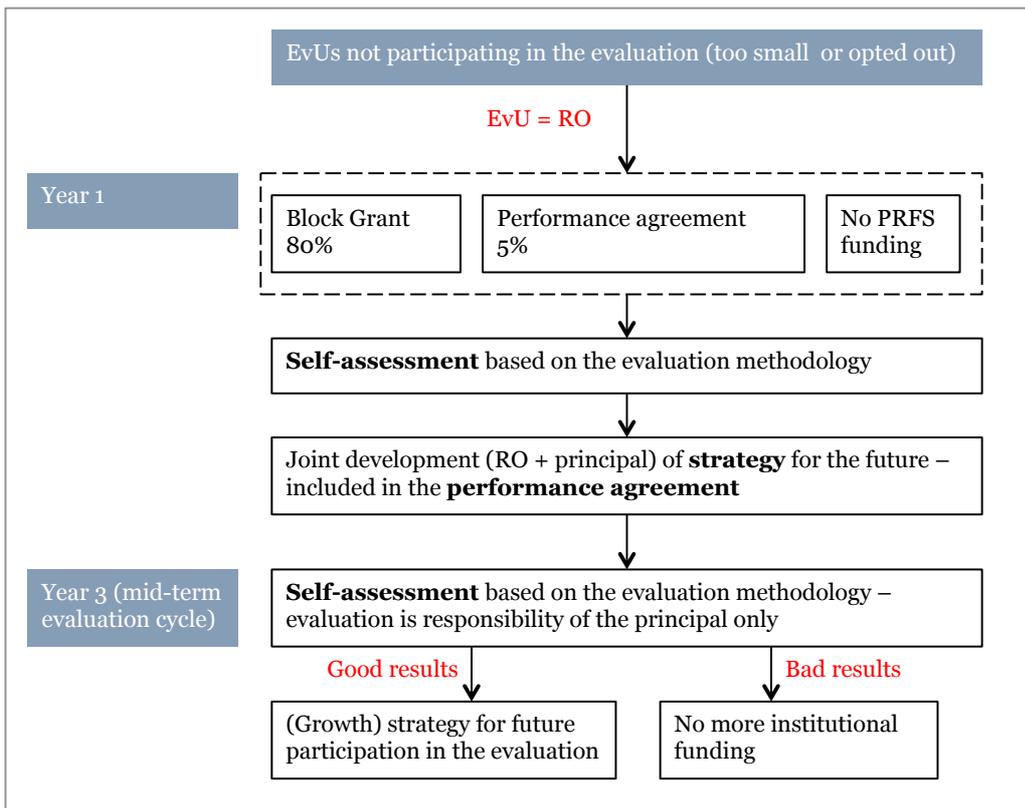
The issue of Evaluated Units (EvU) not participating in the evaluation

Small EvUs that do not pass the minimum volume threshold or EvUs that ‘opt out’, ie choose not to participate in the evaluation, will not have access to the PRFS pot of money. Without any regulation beyond those outlined above, these EvUs will (on average) lose a potential 15% of their institutional funding for RO. This will also negatively affect the calculation of their Block grant in future exercises.

In the Main report *The Institutional Funding Principles* we consider two cases: the case of an EvU that is part of an RO and the case where the EvU is an entire RO. The process for the management of the Evaluated Units that are legal entities on their own is shown in Exhibit 24.

In both cases, our suggestion is that the EvU should be given some time for change and a strategy for the future should be defined, with the principal or the RO management. This strategy should be part of the performance agreement. In both cases, the specific situation of ‘not-evaluated Evaluated Units’ has to be taken into account in order not to unintentionally reduce subject coverage or variety in the Czech R&D system.

Exhibit 24 Process for the management of Evaluated Units that do not participate in the evaluation



Entry of new research organisations into the Funding System

The use of a PRFS linked to an evaluation implies that a procedure and principles need to be defined for the entry in the institutional funding system of new research organisations. ‘New research organisation’ in this context means a research organisation that is a legal entity not eligible for public institutional R&D funding at the moment.

The first guiding principle we are proposing is that the founder of a new research organisation will be responsible for the funding of the new organisation (including institutional R&D funding), either by funding the organisation or by providing access to other funding sources or by a combination of both. This holds for any founder of new research organisations, be it a ministry, the Academy of Sciences in its role as a funding provider, a private company or a university. The advantage of this guiding principle is that it creates responsibility and ownership for the founder.

As a second guiding principle we propose that the decision about whether or not a new research organisation should receive public institutional R&D funding has to be taken outside the system established for allocating this institutional R&D funding (ie outside the PRFS). In fact, a PRFS system cannot accommodate the exit and entry of new research organisations as PRFS systems in other countries show. The second guiding principle means that a procedure and criteria have to be set up by which to decide about the eligibility of a research organisation for public institutional R&D funding (see also Section 2.2, above).

In the Czech Republic at present there is an existing two-step procedure defining the eligibility of research organisations to receive institutional R&D funding, with the first step assessing the formal features that are needed in order to be recognised as a Research Organisation and the second step assessing the scientific features of research organisations which should trigger their eligibility for institutional funding. This is quite an open procedure developed by the RD&I Council together with the Office for the Protection of Competition. However, the second step is not yet operational. Either this procedure (perhaps in an adapted version) will be followed under the new funding system to check the eligibility of research organisations or a new procedure will replace it. Eligibility would, of course, have to be verified for existing RO as well at the beginning of each funding period.

A third guiding principle is that policy makers need to set up guidelines as to how and when a new research organisation enters the system of institutional funding for RO (after having passed the eligibility test). For example, a new research organisation might participate in the new panel-based evaluation (based on the EM) once it has passed the threshold of 50 outputs in the preceding five years. Thus it would enter into the new system of institutional R&D funding we propose and it would have access to PRFS money (while the block grant will be covered by the founding ministry). The state budget (or the relevant chapters respectively) might need to be adjusted accordingly if the institutional R&D funding available to the research organisations that are already in the system is to be stable.

4. Implementation of the R&D evaluation and funding system

As with any reform, the adoption of the new evaluation and funding system will require time, patience and the constructive involvement of many actors in the RD&I system. Most important, it requires change management and measures that allow for the legitimisation of the new evaluation and funding system. We cover this in Section 4.1. There is also a need for the development of structures and means for capacity building and ownership among all actors in the RD&I system. We describe these in Section 4.2.

In a first instance, however, a **decision in principle** defining the evaluation methodology for future implementation (ie NERO) and the principles of the new institutional funding system is needed. Such a decision is key for creating a sense of stability and continuity in the RD&I community and most important, for setting up the needed structures and practices that will allow a robust implementation of the evaluation and funding system.

In Section 4.3 we set out our view on the process towards the achievement of an effective and stable news evaluation and funding system.

Conclusive recommendations for all actors in the RD&I system are given in Section 4.4.

4.1 Legitimisation of the evaluation and funding system

4.1.1 The R&D Governance system: need for expertise and ownership

The Czech RD&I governance system varies in important respects from the West European norm. In our view, this weakens the links between policy and its implementation on the one hand and the needs of society on the other. At least in the medium term, the evaluation and funding system would work more effectively in the social interest if both the governance structure and the capabilities of key ministries were strengthened. We developed this view in the so-called International Audit³⁰ but the analysis remains valid today.

The West European governance model has been strongly influenced by the tradition of New Public Management in recent years. Key principles include

- The government comprises ministers, each of whom generally has responsibility for a sector of society
- The job of the corresponding ministry is to support the minister in the development of policy – and in the past twenty years or so there has been increasing pressure for policy to be based not only on political ideology but also on empirical evidence about performance and what works in the way of policies and policy instruments
- There is a separation between the politically appointed level at the top of the ministry and the civil servants who support the minister. Rarely is the minister surrounded by more than a handful of political appointees and the bulk of the ministry's work is done by professional civil servants who are independent of the minister and who are appointed by the head of the civil service. The civil service develops and contracts the analytical capabilities needed to make policy and is

³⁰ Arnold, E., Good, B., Ohler, F., Tiefenthaler, B., Vermeulen, N., *The Quality of Research, Institutional Funding and Research Evaluation in the Czech Republic and abroad*. Thematic Report No. 3, International Audit of Research, Development & Innovation in the Czech Republic, 2011

therefore expert in relation to the sector of society for which the ministry is responsible

- Ministries do not themselves implement policies at the detailed level but employ agencies to do so, which they manage by objectives. That is, ministries set broad goals and specify performance parameters but do not involve themselves in decisions that involve allocating money to specific individuals or groups (such as project funding). This makes it relatively easy to understand how well the agencies perform their jobs while protecting the system from some kinds of corruption and improper political interference
- In general, a weakness of this kind of governance is that the fragmentation among ministries with different missions imposes a need for coordination if the government and the state are to address cross cutting issues such as overall research policy or the need to address 'societal challenges' such as climate change. Countries use a range of mechanisms including RDI councils, the appointment of 'lead' ministries or lower level coordination committees. It is not clear than any one of these mechanisms is inherently superior – effective arrangements appear to be dependent upon the national context

In this context, aligning the design and execution of Czech RD&I policies better with social needs requires in our view a number of changes in governance and ministry capabilities.

- Clear allocation of both budgetary and planning responsibility for sector research to ministries needing meaningful quantities of research
- These ministries are periodically to produce a research strategy and budget, updated annually
- Allocation of 'ownership' of research organisations to specific ministries, along with the responsibility to provide relevant institutional research funding and to maintain relevant performance contracts
- Allocation of responsibility for basic research funding to a single ministry or a partnership of ministries/the Academy
- Over time, allocation of responsibility for the RD&I funding organisations (including the Academy) to ministry principals
- Responsibility for coordinating overall research funding and the Evaluation Methodology and Funding Principles to a single point in government
- Implementation of the civil service reforms planned but not yet put in place

4.1.2 The development of an evaluation management structure

The establishment of a **stable** structure for evaluation, allowing for the build-up and continuous use of skills and expertise in evaluation design and management, is an important component of the change in the Czech evaluation system.

It is an open question whether such a structure should take the form of an independent agency, an agency established by a Ministry, or a unit within a ministry or the Government Office. No matter the choice, the success of the evaluation and funding system depends critically on the expertise, capabilities and capacities of the staff members of this structure. There will be a need for training and learning and building up an institutional memory from past experience.

In this context it is important that the evaluation should not be considered as an stand-alone, ad-hoc exercise, to be repeated every 5 to 6 years. Instead, it is to be designed and implemented in the context of an **evaluation cycle**, ie a continuous process of evaluation design, implementation, monitoring of its (intended and unintended) effects, advising policy-makers, and taking 'lessons learned' into

consideration, design an updated methodology, eventually responding to new policy priorities and needs. We describe this in more detail in the report *Detailed Evaluation Cost Framework* (Background report 4), indicating also how the combination of a core team and temporary staff members can guarantee the sustainability of the evaluation structure and its stability.

A stable structure with evaluation expertise can also be expected to be valuable for the implementation of other evaluations and the build-up of strategic intelligence that is needed in the RD&I Governance system. One can also envisage an extension of the tasks of the evaluation management structure to acting as an evaluation body for all RD&I governance entities.

4.1.3 Policy decisions related to the funding system

In the transition from the current to the new institutional funding system, policy decisions are needed in relation to some technical procedures as well as for a set of parameters that can be varied in the funding system. This implies the need for the development of a structure and processes for decision-making among the national policy-makers involved, ie all institutional funding bodies as well as the ministries that were the founders of Research Organisations.

Decisions that should be taken *before* the new funding system can be implemented include

- The weights given to each of the assessment criteria for PRFS funding, determining the size of the evaluation criterion-based ‘sub-pots’
- The approach for the translation of the evaluation scores into funding
- The definition of the funding pots for the types of Research Organisations

Future policy decisions can address

- The size of the pots for the different categories of Research Organisations
- The shares of the three components in each RO-type funding pot
- Changes to the previously defined size of the evaluation criterion-based ‘sub-pots’

It should be noted that ‘policy’ does not stand for ‘political’. While final decision-making is the task of the national policy makers, good practice in policy making implies that any decision on the matter should be taken upon **consultation** and consensus-building with the RD&I community.

The policy decisions should also be based upon **evidence and strategic information**. In this context, the role of the evaluation management structure and its activities in the context of the evaluation cycle are of key importance (see the Main report *The R&D Evaluation Methodology*). Evidence needs to be collected on the major failures in the RD&I system and their causes as well as on the intended and non-intended effects of the previous funding system.

Decision-making on the shares of the funding components for the different types of ROs, for example, should be based on evidence related to the needs and requirements of particular types of Research Organisations. If a certain type of Research Organisation displays a particular need for institutional development to improve its performance, the performance agreement component should be given a higher weight. Similarly, if a certain group of Research Organisations has a need to improve quality in its performance and is homogeneous enough to benefit from more direct competition, the PRFS part of the funding could be increased.

Consideration of the funding mix is another important guiding principle for decision making, on the shares of the funding components as well as on the size of the mission-based funding pots. If there already is a high share of competitive funding in a group of Research Organisations, then the performance-based components, in particular the

PRFS, should not be over-emphasised in order to ensure stability and opportunities for capacity building. Similarly, if the share of competitive funding is fairly low, then the performance-based components can be increased more. The share of competitive funding is typically higher in more applied Research Organisations compared to basic Research Organisations.

Of equal importance is the **ex-ante modelling** of the envisaged changes to the institutional funding system in order to ensure that the intended effects are reached while taking account of the eventually non-intended ones. The use of a non-linear funding allocation, for example, would have a strong effect on the RD&I policy system; its likely consequences should be understood *before* it is implemented.

An ex-ante modelling may be problematic in the transition from the Metodika to the new funding system, in particular in relation to the PRFS component because of the radical differences in the evaluation system. Ex-ante modelling of future changes to the funding system should be easier.

Central to any decision should be the national RD&I policy and its objectives and priorities, as well as the strategic objectives established at the ministry level. Especially the size of the mission-based funding pots should be based upon – and preferably expressed in – a long-term RD&I strategy or basic long-term principles for the RD&I policy accepted by all major political forces, ie a strategy or policy valid beyond single election periods.

Finally, we suggest the following principles should be considered when thinking about increasing or decreasing the shares of the funding components.

- Because stability is the most important function of institutional RO funding, the state budget available for that purpose should not be subject to radical cuts
- A prerequisite is that the PRFS is well established among stakeholders and needs to work well before its share can be increased. The same is true for the performance agreements
- Experience in other countries shows that in countries where there is an increase in institutional funding, this increase tends to be allocated through the performance-based part, which implies that the share of the block grant decreases – while not decreasing in absolute terms, thus ensuring stability. In countries where there is little increase in institutional funding, the shares of the block grant and the performance-based part typically remain fairly stable

4.1.4 Decision-making by the RD&I Council

As an arena for discussion and agreement among the research-performing actors, the RD&I Council has an important role to play in setting up the conditions for the evaluation to be run as well as possible.

In the preceding sections of this report we indicated the need and possible value of the following actions that are in the RD&I Council's field of competence.

- The review of the definition of a 'research organisation' and the eligibility criteria for institutional research funding, including the development of guidelines on the process and criteria for new RO to enter the funding system (see Section 2.2)
- The development of a register of scholarly journals, series and book publishers and their quality categorisation, allowing for a more self-determined evaluation of research outputs - in the Social Sciences and Humanities fields and beyond (see Section 3.2.8)
- The review of the definitions of applied research outputs (see Section 3.2.8)
- The updates and extensions needed for the RD&I Information System (see Section 3.2.9)

A major point of discussion and agreement that will need to be taken up before the evaluation and funding system can be implemented is the standard definition of a FTE researcher and the procedures for its calculation (see Section 3.2.7). Main challenges for the collection of accurate FTE research data are the multiple concurrent employment of researchers, which is a common phenomenon, and the distinction between teaching and research time for university staff. There is also no common use and understanding of the term ‘researcher’.

In a first instance, the procedures for the calculation of the time specifically dedicated to *research versus other activities* (such as teaching, but not only) should be defined. International practice shows two possible roads. In some countries, the community agreed upon a time allocation ‘on average’. Examples are Norway and Italy where researchers in universities are considered to devote half of their working time to research and the other half to teaching. In other countries, such as the UK, the expected time to be dedicated to research is agreed upon in each researcher’s employment contract and noted down in an institutional register (and eventually reported to a central national one).

Also a common definition of *what a researcher is* and *what categories of staff* are included as researchers should be agreed upon and used in all Research Organisations throughout the entire RD&I system. Whether or not to include technicians is a convention. Two questions need to be considered in this context: a) is the solution agreed upon fair to all parties concerned?; and b) what are the impacts of the solution agreed upon, for example in relation to gaming?

In relation to the inclusion of *PhD students*, as mentioned in Section 3.1.1, above, our view is that the size of a research ‘entity’ is conceptualised through the number of full-time equivalent (FTE) researchers that the entity is committed to maintain and that constitute its longer-term critical mass for research. Inherently, therefore, it excludes PhD students.

4.2 Capacity building in the RD&I system

The need for building up capacity and expertise in evaluation is significant in the Czech RD&I system. The International Audit of the RD&I System³¹ identified in 2011 the need for the Czech RD&I system to build up an evaluation culture and for policy-makers and research organisations to come to see evaluation as a tool for learning and improving research. In the course of this study, we have noticed that five years after the introduction of the Metodika 2009 PRFS, the direct link between evaluation and funding heavily dominates the evaluation culture and has affected profoundly the Czech R&D evaluation and policymaking system as well as HR management practices and processes in the research-performing institutions.

4.2.1 Evaluation capacities in the Research Organisations

There is little experience of evaluation in the research-performing community. Evaluations similar to the NERO have been implemented so far only in the Academy of Sciences and in the context of the Structural Funds. Especially the practice of self-assessment is not a common practice in the Czech RD&I system and the Small Pilot Evaluation has shown that several less-experienced ROs struggled to understand what type of information was required - and expected by the evaluation panels. Also the fact

³¹ Arnold, E. (2011), International Audit of the R&D&I System in the Czech Republic, Synthesis report, Technopolis Group

that the reports will have to be presented in English may be a hurdle for some ROs to overcome. As a result, the panels risk receiving self-assessment reports of very variable quality. In other words, lack of experience in the RD&I community is a risk factor for a fair evaluation.

While the lack of experience in writing self-assessments is an objective disadvantage for several Research Organisations, the key factor that will allow for a fair evaluation of the Research Unit's performance will be the extent to which the Research Organisations have set up their research management structures and processes in line with international good practice.

In other words, the question is whether the evaluation is seen as a one-time exercise or whether its key concepts are taken up as a guideline for internal practice. The actions that Research Organisations can undertake in order to prepare themselves for the evaluation therefore consist first of all in the take-up of the core concept that underlies NERO in their institutional practice.

The evaluation methodology builds upon the key concept that the research-performing organisations should have developed a robust strategy for their research activities, built upon a sound analysis of their strengths and weaknesses within the various fields of their research, in the national as well as international environment. There is a direction set for the research that takes into account the opportunities offered by the external environment related to the fields the institution is active in. It is expressed in goals and targets that can be (and are) monitored on progress internally, and it guides and is supported by the overall institutional environment, including HR management.

We therefore encourage the Research Organisations to set up an internal structure and processes for evaluation. A first step would be the organisation of an internal SWOT analysis, per field of research and involving the key researchers in the institution. A next step in the process is the development of a research strategy and its formulation. The section on the Guidelines for the Research Organisations in the *Evaluation Handbook* (Background report 5) lists the topics that are of particular relevance for the evaluation panels in this context.

4.2.2 The capacities of the evaluation management structure

The evaluation management structure should be provided with an opportunity to gain experience with the evaluation management before the 'official' implementation of NERO takes place. A learning process for the evaluation management structure is needed because one can expect that there will be organisational challenges, which may create flaws in the evaluation outcomes.

We'd therefore suggest the implementation of a third pilot evaluation. This pilot evaluation does not need to have a broad scope; it can be focused on one field only and involve only a small number of Research Units. This small pilot evaluation would have the only intention for the evaluation institution to gain experience in implementing NERO.

For the participating research organisations, it would equally constitute a learning opportunity. A need from this perspective may exist in particular among the Industry & Business Research Organisations and more applied research-oriented universities, which had little opportunity to be involved in the two pilots organised so far.

No doubt the evaluation management structure will need external support and expert advice also in the first run of the evaluation. In this context it will be able to take stock of the expertise in setting up large-scale evaluations that the Academy of Sciences has developed over the years, as well as the expertise developed in the RD&I system with setting up evaluations in the context of the Structural Funds. Finally, also the IPN Team has gathered experience in the context of the Second Pilot Evaluation. In other words, the evaluation management structure is not working in a complete void.

4.2.3 The capacities of the ministries

The ministries will need to build up expertise and knowledge in order to take up their more active role in the governance of the ROs in their sector as required in the Funding System, but also for the translation of the evaluation results into the different policy agendas and vice versa.

This implies that an increase in financial resources for the ministries' departments responsible for RD&I support should be envisaged, allowing for the employment of experts or the contracting of external experts. Also for the ministries, continuity and stability in personnel are critical for building up institutional memory.

In particular for closing gaps in knowledge and understanding, we suggest as preparatory work for the implementation of the funding system the collection and analysis geared to reaching a complete overview of the funding sources per Research Organisation, so that the decision-making ministry can see how changes in R&D institutional funding affect the organisation's budget as a whole. Also a thorough analysis of the competitive funding and its role for the Research Organisations would be useful, in order to understand the reasons why some competitive programmes seem to be oversubscribed, with success rates around 20%, while others appear to be undersubscribed leading to only limited competition.

4.3 A pathway towards stability

In strategic documents to the European Commission such as the Czech Republic 2014 National Reform Programme³², Czech policy makers say that they expect to reach full implementation of the new evaluation and funding system “from 2016”, after a pilot evaluation in 2015. In other documents for the European Commission, the date of 2017 is envisaged.

The new evaluation and institutional funding system constitutes a significant change in the current approach to evaluation and institutional funding in the Czech Republic, which is both its value and challenge. While there is no ‘technical’ objection to a swift implementation, a hasty implementation of the evaluation system risks creating flaws in the evaluation outcomes, which would undermine the value and acceptance of the new evaluation and funding system in the Czech RD&I system. Time should be given to all actors involved to implement the preparatory activities set out above. In conclusion, while the date of 2017 may still be early, a full-scale implementation of the National Evaluation of Research Organisations in 2018 should be feasible.

Experience from other countries shows that any reform of an evaluation and funding system takes time: it demands capacity building and the collection of expertise by all actors involved. As in most PRFS systems there will be also a continuous need for adjustment of the evaluation and funding system in order to adjust eventual negative effects on the RD&I system and reflect eventually changing policy priorities.

A preparation phase is followed by the first implementation of the new evaluation and funding system, and this leads to a transition period of new experiences and learning from the experiences made in the first period. This will lead to a refinement of the evaluation and funding system while keeping firm the key principles, thus ensuring

³² The National Reform Programme is a ‘conceptual’ policy document developed by the Office of the Government in the context of the European Semester. It sets out the plan of key measures defined by the Czech government to reach the national targets in the context of the Europe 2020 strategy, in line with its strategic priorities and in response to the European Commission Specific Country Recommendations.

continuity and stability in the overall process and concepts. This is followed by the second implementation of the new evaluation and funding system..

All together it may take two to three funding periods before the new system will be fully effective, meaning that patience and persistence are key prerequisites of a successful public governance of institutional funding for RO.

We have envisaged that the NERO will be implemented every six years. In a first phase, the need to ensure a balanced and fair allocation of the institutional funding may imply a more frequent implementation of the evaluation, eg every three years. Once the funding system has reached a satisfactory level of maturity, a longer interval can be envisaged.

4.4 Conclusive recommendations

As a conclusion, we recommend that national policy-makers

- Take a decision in principle to the adopt the new evaluation and funding system as soon as possible
- Strengthen the governance structure and the capabilities of key ministries, implementing a number of structural reforms in the governance system in order to align the design and execution of Czech RD&I policies better with social needs
- Set up a structure and processes for decision-making on the technical procedures as well as for a set of parameters that can be varied in the funding system
- Support the collection of strategic information for an improved understanding of the RD&I system and to support policymaking

We invite the RD&I Council

- To reach an agreement on a standard definition of an FTE researcher and the procedures for the calculations
- To review the definition of a 'research organisation' and the eligibility criteria for institutional research funding
- To take up the activities needed for a fair evaluation of research performance, including the development of a categorised journal register and the definition of applied research outputs
- To foster the implementation of updates and extensions to the RD&I Information System

We recommend the Research Organisations

- To set up internal structures and processes for evaluation, in line with the evaluation methodology

Appendix A The OECD field structure

Disciplinary Area	Field
1 Natural sciences	1.1 Mathematics
	1.2 Computer and information sciences
	1.3 Physical sciences and astronomy
	1.4 Chemical sciences
	1.5 Earth and related environmental sciences
	1.6 Biological sciences
	1.7 Other natural sciences
2 Engineering and technology	2.1 Civil engineering
	2.2 Electrical engineering, electronic engineering, information engineering
	2.3 Mechanical engineering
	2.4 Chemical engineering
	2.5 Materials engineering
	2.6 Medical engineering
	2.7 Environmental engineering
	2.8 Environmental biotechnology
	2.9 Industrial Biotechnology
	2.10 Nano-technology
	2.11 Other engineering and technologies
3 Medical and Health sciences	3.1 Basic medical research
	3.2 Clinical medicine
	3.3 Health sciences
	3.4 Medical biotechnology
	3.5 Other medical sciences
4 Agricultural sciences	4.1 Agriculture, forestry, and fisheries
	4.2 Animal and dairy science
	4.3 Veterinary science
	4.4 Agricultural biotechnology
	4.5 Other agricultural sciences

Summary report

Disciplinary Area	Field
5 Social sciences	5.1 Psychology
	5.2 Economics and business
	5.3 Educational sciences
	5.4 Sociology
	5.5 Law
	5.6 Political Science
	5.7 Social and economic geography
	5.8 Media and communication
	5.9 Other social sciences
6 Humanities	6.1 History and archaeology
	6.2 Languages and literature
	6.3 Philosophy, ethics and religion
	6.4 Art (arts, history of arts, performing arts, music)
	6.5 Other humanities

technopolis |group|

In collaboration with

NIFU

Nordic Institute for Studies in
Innovation, Research and Education

