

Feedback for the Technopolis consortium, composed by the expert team of IPN Metodika on the Draft of the 1st Interim Report

STRUCTURE OF THE FEEDBACK

1. Our feedback is provided in three forms:
 - a. A list of general observations by the expert team of the IPN Metodika, based on reading the submitted Draft versions of the three documents (Draft, further denoted as “D”, Evaluation Implementation Draft, further denoted as “EI”, Country Analyses, further denoted as “CA”).
 - b. Detailed comments provided by the guarantors of individual Key Activities and RVVI. In case of conflicting opinions, consider ad a) as superior.
 - c. Feedback from research organisations, providers, and few other institutions. Sections of the feedback that deserve greater attention are marked in red font if they seem to be very relevant or blue font if they represent widespread reservations regarding such types of evaluation.

GENERAL COMMENTS OF THE EXPERT TEAM OF IPN METODIKA

2. Comments stressing the link to Terms of Reference are marked by [ToR]. Page references to individual documents are marked by pp.#.
3. Compared to the previous Concept, the Draft improved notably in terms of content and structure. But even the Draft still needs substantial improvement, as we highlight below.
4. **The texts are still sloppy:**
 - a. Numerous mistakes still need to be fixed. We do not provide a detailed list because it would be too long. But many of them are mentioned in other components of our feedback and previous comments to the Concept.
 - b. Various definitions (such as eligible researcher, thresholds, outputs, money source types, etc.) are still too vague. Contradictions appear in different parts of texts.
5. **On the structure of documents:**
 - a. The executive summary is mostly a set of general statements. This is fine, but we recommend that another type of summary is prepared, featuring the main attributes

of the proposed evaluation methodology, further denoted as “EM”, a reader could get familiar with before reading the detailed proposal.

- b. The first half of the Draft discusses pros and cons and trade-offs of different approaches to evaluation. Would it not be useful to conclude each section by a bullet point summary, stating key implications you make for the proposed shape of the EM?

6. Some components highlighted in our previous comments are still missing. We consider it to be a serious deficiency:

- a. An explicit and fully described template (clear examples) of a standard report informing about bibliometric and total R&D performance of RU and EvU. What is proposed in the Draft is still very vague to comment on [ToR].
- b. Examples of filled-in forms such as self-evaluation report are still missing [ToR].
- c. A field or at least discipline specific methodologies highlighting the key specific features to be considered by panels. These should inform evaluated subjects ex-ante and among other things, it will foster evaluation culture in the country since it will help convincing the whole academic community of the necessity of adopting a new approach to evaluation (ToR). We do not insist that methodologies are very detailed and some space should be left for the judgement and experience of panellists.
- d. The EM proposed does not include on-site visits [ToR]. We have serious worries that only “distant” peer-review will not effectively capture some important dimensions of the evaluation. Proposed w/o on-site visits component, the EM has to be defended by clear and convincing arguments.
 - i. It should include reference to other countries where only 'remote' peer-reviews are used while evaluating all the dimensions the proposed EM does.
 - ii. It should clarify the potential value added by on-site visits and highlight cases of EvUs, where it would be the highest,
 - iii. It should estimate additional costs of it beyond the budgetary limit imposed by the ToR.
 - iv. We remind that we had proposed incorporating on-site visits as an option in case of selected EvUs. For example, an on-site visit might be required in case of EvU planning to contract significant forward-looking funding or it might be used in case of doubts about the content of a self-evaluation report.
- e. Related to the point above is the following: You state that a key principle of the EM also is: *It will use a single framework for assessment across all disciplines and research organisation typologies while allowing for a reasonable level of field- and RO typology-specific variations.*? [pp. 59Draft] But you should clarify why the EM you propose should use the same level of detail of exploration for all ROs and EvUs.

Given huge diversity between and within Czech ROs, budgetary limits and efficiency considerations of evaluation, as well as different levels of deepness of evaluation, it is a natural approach to consider. The scale and scope of evaluation in particular ROs could be decided either by the superior institution, such as the Ministry of Education in case of HEIs, or university management in case of its EvUs (faculties).

- f. It is not shown how the quantitative information (scores) assigned to RUs and EvUs on research performance and excellence at the level of scientific fields and disciplines will be synthesized to the level of the whole country. Will the synthesis highlight particular RU among other RUs in the same field and how it will compare to the rest of the world?
- g. Detailed guidelines for main panels, panels, reviewers, etc. are not provided yet. We expect that these will be provided as soon as possible to inform the preparation of the large pilot.
- h. EU GBER issues and eligibility of institutions for public institutional support will hopefully be treated in the 2nd IR.

7. Information support for the evaluation:

- a. The guidelines for selecting indicators [pp.70Draft] state that: *Takes into account data availability, in particular in the Czech RD&I IS, in order to promote cost efficiency.* But in our past comments on the Concept, we had stressed that the proposal should go beyond the current content of the Czech RD&I IS. It is because the EM is being designed in a longer time horizon and necessary improvement and extensions of the RD&I IS should be implemented.
- b. R&D outputs reporting proposed in the Draft rely implicitly on the existing system of how data finally appears in the RD&I IS. It also means that it relies on existing definitions of eligible R&D outputs and eligibility of RO to report outputs (defined by Metodika), on existing reporting procedures, on existing system of data inspections and verifications (quite poor). We understand that you cannot provide statistics based on properly defined types of outputs, properly reported and carefully verified R&D outputs at this moment. But the EM proposed should also contain clear proposals of what has to be changed to get a more reliable and more informative content of the Czech RD&I IS in the future. An example: is it ok to continue the current practice of using approved lists of recognised journals which are not indexed by the TR and Scopus. Should it still exist or should all articles in other than WoS/Scopus be inspected before they enter the RD&I IS?

8. There are some elements in the EM which we find very questionable:

- a. Only some types of R&D outputs can be proposed for evaluation of excellence. In particular, only published types whereas non-published types are not to be considered. But the EM should identify excellence in R&D wherever it is and

whichever form it has. ROs strongly oriented on applications justly complain that the EM would discriminate them.

- b. Related to ad a), it is not clear how panellist will identify the quality profile of R&D outputs (other than those very small 1% or 2% submitted in the excellence channel). It seems that they will see just total counts of outputs of particular type + bibliometric report. Given that bibliometry is meant as additional information only and serves well only for some fields, how will panellists learn about the quality of outputs? Will the RUs report complete references for all of its outputs? Will RUs sort these lists by their perceived importance and quality?
- c. The proposal relies too much and explicitly on the WoS. While we understand that examples and small pilot are using WoS data, the proposal should be forward looking and should not be linked to WoS only, but it should be open to alternative data sources. In other words, the EM should explain what data and how processed should it be, but should not be linked to a particular data provider. The company providing the data will be decided in the future when nation-wide evaluation is implemented.

9. The notion of RU still creates some problems:

- a. The term Research Unit gives the impression that “unit” means a team, department or some other coordinated research group. But this is not so. If you want to use the term RU, you have to make sure that the readers understand this term properly from the very beginning. You should explicitly state that RU does not necessarily represent a coordinated or collaborating research group. This should be also stressed in the list of defining keywords of the EM.
- b. We understand well that the primary role of RU as defined (corresponding to individual fields of research) is to secure that evaluation does not mix apples and oranges. But at the same time, the proposed structure of the EM does not sufficiently recognise that only some features subject to evaluation can be linked to a particular RU.
 - i. Some assessment criteria of *research performance* such as *ability to attract PhD students, capacity to gain competitive funding from international and national sources* will be difficult to attribute to particular RU within an EvU.
 - ii. Research infrastructure (e.g. library) is usually shared by more RUs within an EvU. Its use should be described by the EvU, not each of its RUs individually.
 - iii. Assessment criteria of *institutional management* and probably *development potential* [pp. 64Draft] make sense only at the level of EvUs, and not at the level of RUs. Note that a RU, as you define it, is NOT an organisational unit but it is a mere set of researchers doing research in the same field. The claim that RUs have their institutional structures provided on pp. 65Draft is not clear. It makes sense to evaluate this dimension at the level of RU only when

a RU within an EvU represents some kind of governance unit like a department or team.

- iv. Also, note that requiring inadequate information from RU (and simultaneously all other RU within an EvU) imposes an unnecessary additional burden on the evaluated.
- c. The Draft proposes that the panel chair will write a report on all RUs in the field the panel evaluates. It is probably fine if the number of RUs is not too big. But it has to be clarified who will synthesize evaluation reports of individual RUs within particular EvUs to a report for the EvU. It cannot be done by the panel chair because he will be informed only about one RU, not about other RUs within the given EvU.
- d. Reporting templates in the Evaluation Implementation impose absolute limits on the number of reported items by RU (10 max). But note that there will be a huge heterogeneity in the sizes of individual RUs. Identical limits imposed on all RUs, irrespective of size, does not seem to be suitable.

10. There are several important rules in the EM which deserve more attention and more careful specification in the text to eliminate ambiguities and conflicting statements:

- a. Typologies of ROs still contain mistakes and ambiguities.
- b. There are three very different classifications of research fields the EM will have to deal with (OECD Frascati defining the main panels and panels and RUs; WoS or Scopus categories which will guide bibliometrics; current RIV structure of fields). For example, how will the EM deal with differences between Frascati and WoS field structures when creating bibliometric reports for RUs?
- c. Rules guiding participation of ROs, EvUs, RUs in the evaluation are vague, unclear, and contradictory. Implications of (non-)participation and allocation of decision making powers in that respect are missing. If it is to be clarified in the 2nd IR on funding, it should be stated.
- d. The proposal of the implementation scheme has improved, but the text still contains numerous ambiguities and inconsistencies in terminology (e.g. Management board/committee, Governing body). This is also due to leaving out the scheme that was in the Concept on p. 89.
- e. Thresholds to enter the evaluation are uniform, not reflecting field or discipline specifics (e.g. publication frequency), some need better justification, and some are subject to easy gaming.
- f. Quantitative assessment of each of the five basic criteria with a single grade for the whole RU will hardly inform of the quality of larger and heterogeneous RUs. It will not inform of the qualitative profile (distribution) of R&D performance (outputs and impacts), of membership of the national and global research community, and societal relevance within a RU [ToR]. On the other hand, a single grade seems

suitable to evaluate institutional management of an EvU, while grading this at the level of RU does not make sense [pp.64Draft].

- g. PRFS is still frequently mentioned as evaluation, although it means primarily funding (pp.20Draft).
- h. It is still not completely clear to what extent and how R&D infrastructures will be evaluated. For example, whether and how will panels and main panels evaluate infrastructures such as libraries, which also do research to enhance their operation?
- i. Should there not be an option for huge RUs within single EvUs, allowing their split into more than one RU in the same field?
- j. Will a small number of panel members, as proposed, manage to process so many RUs, EvUs and outputs?
- k. Will the evaluation also look at whether national R&D priorities are being achieved, who will do so and how?