

Institutional Responsible Research Assessment Efforts and the Response from SSH Disciplines – the Case of Masaryk University

Michal Petr, Natálie Hílek, Petra Mořkovská, Monika Kuchlei Sieberová
 Research Office, Masaryk University, Žerotínovo náměstí 9, 60177 Brno (Czech Republic)

Introduction

Since 2022, the Coalition for Advancing Research Assessment has become the central initiative for responsible research assessment (RRA). But how to implement such general ideas in practice and do they fit the disciplines' needs?

In 2022, Masaryk University (MU) conducted the first systematic **Internal Research and Doctoral Studies Evaluation (IRDE)** with a strong focus on principles of responsibility. We firstly want to describe the MU assessment universe and secondly to discuss the response from SSH faculties to this system as a part of "evaluate the evaluation" phase.

We operate with the definition of RRA as an approach that „incentivises, reflects and rewards the plural characteristics of high-quality research, in support of diverse and inclusive research cultures.“

RESEARCH EVALUATION



Internal Research and Doctoral Studies Evaluation (IRDE)
 formative; expert feedback for the improvement

TRIAD OF EVALUATIVE ACTIVITIES AT MU

35% performance indicators
 65 % contract (stabilization and research strategies)

Annual meeting: research, teaching, investment. Regular bibliometric analysis (monitoring trends)



FUNDING



MONITORING

Figure 1. System of evaluation activities at the Masaryk University

Research Assessment at MU

MU assessment has a triad of components: formative **research evaluation, funding, and monitoring** (Figure 1). Even though these components are complementary, they **operate independently for separate purposes**. Foremost, bibliometrics does not affect research assessment (IRDE) and IRDE does not directly influence the funding.

Internal Research and Doctoral Studies Evaluation (IRDE) has a **formative** mission: once in 5 years it provides **departments and doctoral programmes** with expert feedback, for their use and benefit. The overall design and the criteria were developed with attention to the principles of RRA; inspired by SEP in the Netherlands and SCOPE framework.

Faculties newly prepare a **strategic plan** that may be informed by IRDE results and that is linked to the major part of the core budget for research – **contract** (Figure 2).

Main features of IRDE

- Faculties/institutes participated in the planning and development.
- High level of flexibility: the setup and composition of panels were entirely up to the choice of the faculty/institute (disciplinary panel or ISAB); self-evaluation reports of units are predominantly narrative and customizable.
- Bibliometrics is intentionally reduced and panoramic and only aims to deliver basic underlying publication statistics.
- Evaluating the evaluation phase and communication: consultation with the creators of the SCOPE protocol, a conference dedicated to RRA, survey and summary report evaluating the IRDE.

"Responsible assessment at MU is not only about the setup and design of the IRDE, but also about the setup of this entire triad system."

Survey – The response from SSH

Publication patterns, channels, and types are much more heterogeneous in SSH disciplines than in science, technology, and medicine (STM). Research assessment in SSH is an issue discussed in various communities (e.g. ENRESSH). Here, we want to **disentangle if the implementation of the above-defined principles of RRA can bridge the gap to SSH and to what extent it fits SSH disciplines**.

We conducted six semi-structured interviews with MU SSH faculties' vice-deans for research: Faculty of Social Sciences (FSS), Faculty of Arts (ARTS), Faculty of Education (PED), Faculty of Business and Administration (ECON), Faculty of Law (LAW), Faculty of Sports Studies (SPORT). Each vice-dean represents the full range of disciplines in their faculty. **We asked four default questions:**

What is the notion of quality in your disciplines and how can be represented in the assessment?

PED and ECON confidently formulate the quality concept. Most respondents describe external signs that demonstrate the wanted performance rather than explicitly say "what is quality". We see two basic trajectories. The first, and major trajectory accepts journal publishing and journal-level indicator as a general and convenient signal of quality and the effort to publish in community-recognized prestigious journals as a wanted publication pattern (FSS, ECON), although they concurrently admit differences in this attitude between sub-disciplines (ARTS) or teams/individuals (LAW, SPORT). Secondly, societal relevance, i.e. the ability to address research topics with some influence on society, or "meaningfulness", is perceived as central (PED).

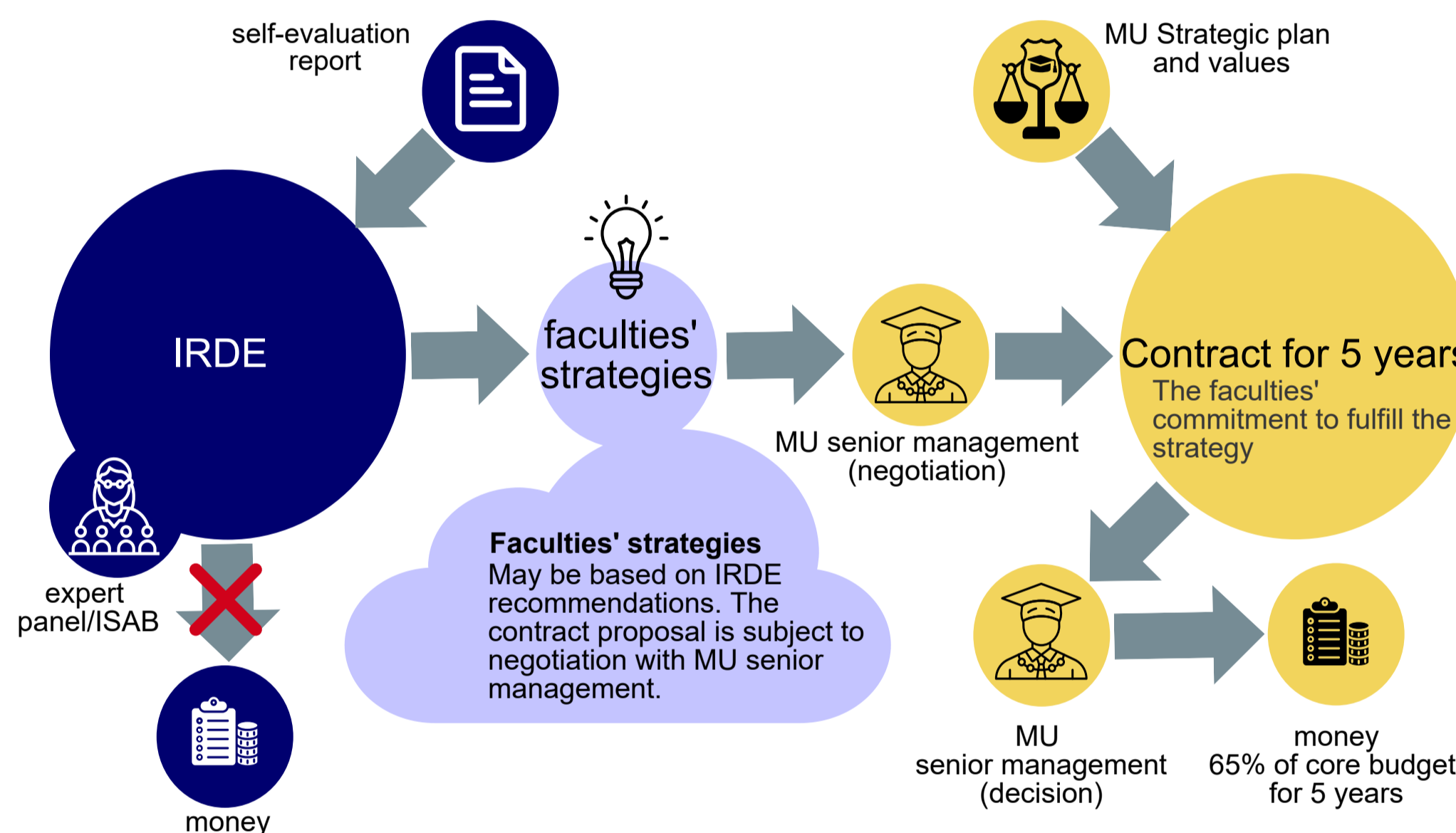


Figure 2. Impact of IRDE

To what extent does the concept of RRA correspond to your faculty's notion of how disciplines should be evaluated? What does responsibility in the assessment mean in SSH?

Most faculties agree with the concept of the RRA, PED even calls for intensifying this aspect. But there are some contradictions. On one hand faculties want to be evaluated like that, but according to ARTS, ECON, SPORT and LAW, the concept of diversity allows too broad interpretation leading to "adopting that definition to what they've been already doing" (LAW) or preventing from "breaking the encapsulation" (ARTS). According to SPORT, "...diversity and inclusivity need to be circumscribed because if it is boundless, it has no direction and no goal". FSS finds the RRA concept rather vague and prefers a bibliometric approach, also ECON views the quality through the prestige of sources.

To what extent was the design of IRDE sufficient for the notion of quality in SSH and for the responsibility in assessment in SSH?

According to all faculties the IRDE was sufficiently flexible and allowed to capture disciplinary diversity and quality in SSH. Three respondents (ARTS, LAW, SPORT) claim that aspects of diversity and inclusivity were so emphasized that the evaluation outcomes were blurred and difficult to work with at the faculty level. SPORT, however, is satisfied with the outputs thanks to constructive and tough feedback from the evaluators, whereas high level of inclusivity may elsewhere result in gaming with the panel members (ARTS). Thus, the choice of evaluators seems to be fundamental. Some faculties mentioned that they did not use IRDE's full potential, as they did not have previous experience with evaluation (PED).

As for the systemic impacts on academia, how did your faculty respond to IRDE?

Except for SPORT which proceeded to the fundamental restructuring around defined research areas with the position of "Scientific Leaders", other faculties didn't change traditional structures. However, IRDE had a direct impact on strategies or support mechanisms. Faculties created research topics that will be supported gradually (SPORT, ECON), changed their internal funding (ARTS, SPORT, ECON) or evaluation system (FSS). ARTS put differentiated requirements for departments and research support. The impact on the internal culture was primarily mentioned by SPORT "I think that in general people have accepted it and that IRDE has actually helped us to really kick off an increase in quality ..." and PED, although decided not to implement suggested structural reforms: "... that type of assessment is not the calibre to trigger things like that... it can only incentivise!"

Conclusion

- Evaluation reform, to be meaningful, must take place not only at the central level of the institution but at all levels and across all contexts (hiring, promotion, research evaluation etc.).
- The values of inclusivity and diversity are welcomed but must not relativize different evaluation missions for different interest groups. These values must be conceptualized in the evaluation system. RRA commitments are not to be used in a very straightforward way, but according to the desired mission and purpose of the evaluation.

- Despite recent trends and publication patterns in SSH, the journal-level metrics remain important in some SSH fields as a signal of quality and benchmarking tool for achieving desired performance in comparison with EU-15.
- The focus on formulating values for evaluation and formulating notions of quality in different disciplines is needed to allow respective interest groups to benefit from evaluation results, and, concurrently as a protection against possible inadequate influence of metrics.
- The new evaluation at MU was broadly accepted, and (mostly) honest and led to the transformation of evaluation culture.

Do you want to know more?



csse@muni.cz

Special thanks to

